HYBRID APPROACHES TOWARDS HIGH–SPEED, LOW–VOLTAGE FLASH MEMORY DESIGN

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Shantanu Rajaram Rajwade

January 2013

HYBRID APPROACHES TOWARDS HIGH–SPEED, LOW–VOLTAGE FLASH MEMORY DESIGN

Shantanu Rajaram Rajwade, Ph.D.

Cornell University 2013

Integration of discrete charge storage in nanocrystals (NC) or dielectric traps is shown to alleviate limitations on tunnel oxide scaling and operational voltage reduction in conventional Flash memories. With the advantage of reduced voltages, NC-based memories become viable for integration in conventional SRAM cells to provide nonvolatile (NV) functionality. 10-transistor NV-SRAM cell was proposed and validated with SPICE-level simulations to provide useful guidelines into system level design of a nonvolatile microcontroller.

However, discrete storage-based memory devices are known to increase variability in performance, restricting their progress towards full scale production. One method towards reducing variation investigates the effect of engineered nanopores in the tunnel oxide and charge-trap layer of Flash memory gate stack. Nanopores are shown to generate preferred tunneling paths for electron injection by field enhancement in the porous medium as well as render better retention characteristics to the memory at the expense of lower memory window.

In order to propel Flash memory into the low-voltage and fast-programmable device regime, integration of ferroelectric (FE) thin film with discrete charge storage into a single hybrid memory device was proposed. Storage of gate-injected electrons adds to the memory window generated by ferroelectric polarization and reduces the depolarization field in the ferroelectric during retention. The first generation of these

devices integrated ferroelectric PVDF polymer and HfO_2 trap layer to demonstrate larger memory window and longer retention compared to conventional FE-FETs.

The dynamics of ferroelectric switching in the hybrid gate stack naturally establishes a two-step program mechanism of faster polarization alignment followed by slower electron tunneling into the storage layer. This was verified by the fabrication of second generation of low-voltage hybrid devices with PZT thin film and Au NC. These devices demonstrated fast DRAM-like switching as well as slower Flash-like operation with distinct signatures from the two memory mechanisms in both program and retention dynamics. A statistical switching model describing ferroelectric switching was integrated in simulations for conventional Flash memory dynamics to corroborate the proposed dual-speed program mechanism in the hybrid device and provide realistic estimates of program and retention transients for the two distinct modes.

BIOGRAPHICAL SKETCH

Shantanu Rajwade is from the city of Mumbai, India. As soon as he showed early signs of scholastic aptitude, he was put on a continuing treadmill of academic excellence. His love for mathematics and physics motivated him to appear for the Joint Entrance Examination (JEE) and join the Dual Degree (B. Tech and M. Tech) program in Electrical Engineering at Indian Institute of Technology (IIT) Mumbai in July 2002.

Learning was abundant in the midst of brilliant peers and inspiring faculty during his five years in IIT, Mumbai. Shantanu developed keen interest in microelectronics through his junior level courses in semiconductor devices physics, device technology and VLSI design. He worked with Prof. V. Ramgopal Rao for his senior and master's level project on self-assembly of polymeric nanowires as interconnects in nanoelectronic applications.

After completing his dual degree, he joined the MS/Ph.D. program at the school of Electrical and Computer Engineering, Cornell University in Ithaca, NY from Aug 2007. After considerable brainstorming with Prof. Edwin Kan on various research directions, Shantanu found an interesting thesis-worthy topic in solid state memories. For the last four years of his PhD, he worked on the integration of conventional ferroelectric FETs and Flash memories into a single hybrid memory design.

His other interests include device modeling, quantum and mesoscopic transport, design of novel low-voltage transistors and memories, self-assembly techniques in polymer electronics and ferroelectric thin films.

Dedicated to my beloved parents, Rajaram and Manjiri Rajwade, brother Tejas and my wife, Shraddha

ACKNOWLEDGMENTS

The five-year haul of graduate school is an excellent opportunity into discovering oneself, one's likes and dislikes, as well as personal strengths and limitations. This journey would not have been enjoyable without the support and care from the all the people I connected with, near and away.

My wife, Shraddha, has been a strong pillar of support throughout my journey in graduate school. Her faith and affection kept me going through some difficult times. I am indebted towards my parents for their immense sacrifices and perseverance in giving the best they could during my upbringing. My brother Tejas has always been by my side, relieving those worries and providing me with ample smiles.

I was fortunate to find an excellent thesis advisor who constantly motivated me through his immense passion for research and innovation. Prof. Edwin C. Kan was always available for my needs and concerns, research and otherwise, giving me sufficient freedom in picking an interesting thesis topic. I am thankful to my special committee members, Prof. Sandip Tiwari and Prof. G. Edward Suh for their careful scrutiny and numerous useful suggestions for improvement in my research work.

Graduate school could not have been so much fun without my labmate Krishna and the infinitely many brainstorming sessions we had over lunch, coffee and even during experiments. I am grateful to former Kan group members, Alex Hou and Jonathan Shaw, and Joshua Phelps for training me in CNF as well as my junior Kshitij in assisting me with simulations through the latter half of my PhD. I will always

remember my interaction with all my group members including Udayan, Jaegoo, Keith, Nini, Fan, Xiaoyang, Sarah, Adrian, Phil, Yinglei, Yunfei and Fiona. Coffee breaks were real stress busters with Krishna, Tanay, Mustansir and Suresh. They truly served as a source of creative idea exchange outside of Kan lab. Learning through peers is always emphasized in Cornell ECE through numerous interactive platforms like Friday EDS talks and informal ECE-GO sessions. I am thankful to Rajeev, Sasikanth, Faisal, Kurshad, Xiao, Wacek, Eugene, Jahan, Dana, Joshua, Wei-min, Ishita, Shriram, Amit, Ajay, Siddharth, Bo, Omeed, Shriram, Adarsh and Albert.

I was fortunate to find forthcoming collaborations from KK Yu, Hitesh and Taro. I am grateful to Prof. R. B. van Dover and Prof. Ulrich Wiesner for their invaluable inputs towards our collaborative projects and Prof. Sunil Bhave for his friendly and helpful advice throughout my graduate schooling. Device fabrication and characterization would not have been possible without the assistance of CCMR, CNF and NBTC project staff.

I was blessed with the company of remarkable housemates during my five year stay at DFB. I will treasure the dinner table discussions and midnight chat-sessions with Praveen, Anirikh, Mihir, Umang and Rajesh. I will cherish my memories of those closely associated with our house including Debjeet, Manish, Kriti, Ashima, Poornima, Srinitya, Sushmit, Aadhar, Rohit, Salil, Akhil, Anuradha, Deepthi, Sahil, Nitin, Shreyasi, Megha, Chirag, Saranya, Sarvani, Rohil, Shubham, and Sachin. Lastly, I would like to thank everyone I interacted through Cornell India Association and my badminton and tennis sessions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Background

Arriving in the early 1980's, Electrically Erasable Programmable (E^2P) ROM Flash memories [1] have come a long way in the nonvolatile memory market that was primarily dominated by magnetic storage media. In their initial years, Flash memories were used with random access NOR architecture in computer boot drives and other read-only memory (ROM)-type applications. However, the last decade saw the second type of architecture for Flash, namely the NAND type, have an increasing impact on the state of electronic storage. Today, NAND Flash memories are ubiquitous in most mobile electronic devices like USB storage drives, gaming consoles and portable gadgets that include cell phones, music players and digital cameras. Regardless of the architecture, what make Flash memories so attractive are their extremely high bit-densities, compact package size, robustness against mechanical shocks and efficient memory access operations. Flash memory technology is compatible with complementary metal oxide semiconductor (CMOS) processing that has benefitted them in lowering the production cost significantly. Due to the regularity in their design, Flash memories have surpassed logic devices in their scaling and have truly become the semiconductor technology drivers. The ever increasing demand for portable storage is helping push the cost per gigabyte ($/GB) for these memories to be comparable to traditional magnetic storage. Chip manufacturing industries like Samsung, Toshiba and Intel are already shipping NAND Flash memory solid state

drives (SSDs) in the market, ushering in a new competition to hitherto exclusive magnetic hard-disk-drives (HDDs).

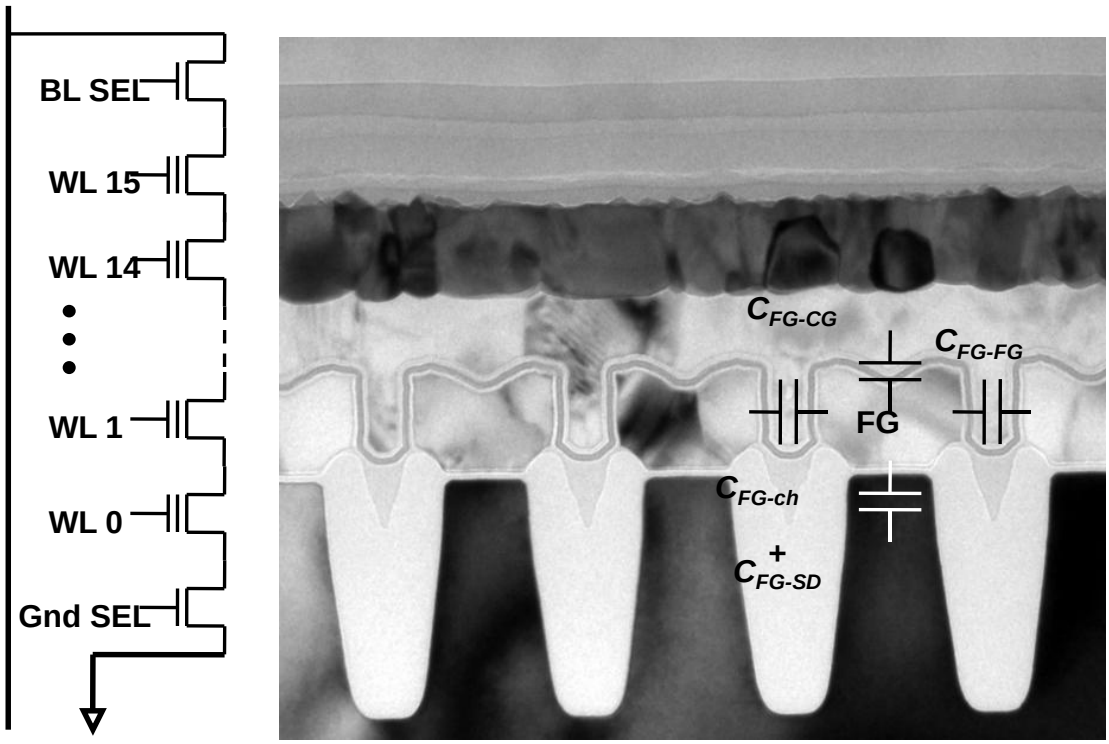


Figure 1-1 (a) Circuit schematic of NAND Flash memory (b) TEM cross-section along a word-line for 2 Gb Samsung Flash memory [11].

The relentless scaling of CMOS devices, including Flash memories, is fast approaching its physical limits and increasing device densities is getting increasingly difficult. Immense research has gone into investigating modifications to existing Flash design and material selection as well as proposing complete replacement to the current memory mechanism. Those involving modifications to the present design include substituting conventional polysilicon (poly-Si) floating gate with trapping dielectric layers [2]-[4] or engineered nanodots or nanocrystals (NCs) of semiconductors [5] or

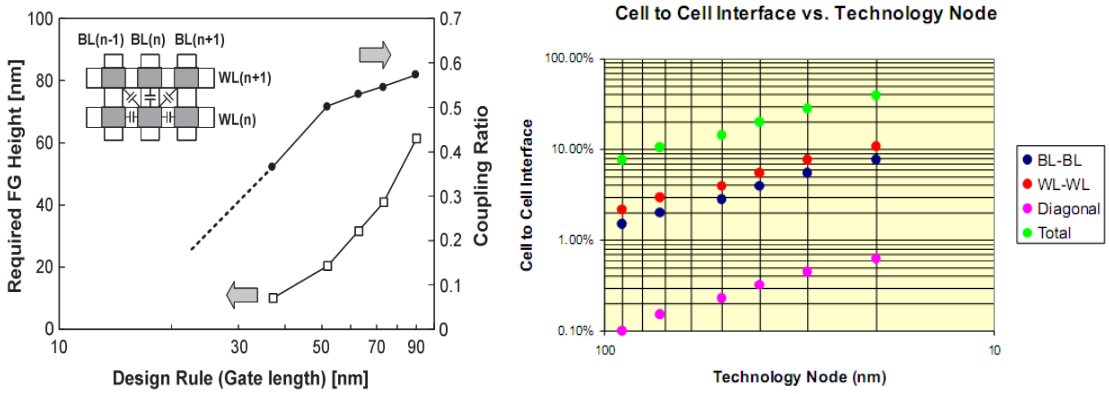
metals [6] as charge storage media. On the other hand, memories with new operational principles aiming for Flash replacement include Phase Change random access memory (PC-RAM) [7], Ferroelectric RAM (FeRAM) [8], Resistive (Re) RAM [9] and Magnetic Tunnel Junction (MTJ) RAM [10].

1.2 Fundamental limitations in Flash memory design

Fig. 1-1 (a) shows the schematic of NAND Flash memory array. Every bitline (BL) consists of 16 (or 32) memory cells connected in series (logic NAND-like configuration). Fig 1-1 (b) depicts the TEM cross section of the NAND array along the word-line direction fabricated in 90 nm technology [11]. The memory cell consists of a floating gate, electrically isolated from the control gate at the top and the sensing channel in the bottom. The floating gate (FG) is a conducting highly doped (p^+ or n^+) poly-Si film. The bottom isolating oxide, also known as the tunnel oxide, is a thermally grown silicon dioxide and less than 100 Å in thickness. The top blocking oxide, known as the control dielectric or the inter-layer dielectric (ILD) is usually about twice as thick and consists of layers of oxide-nitride-oxide (ONO) stack for a larger effective dielectric constant and better chemical isolation from the control gate.

Electrons from the sensing channel can be quantum mechanically tunneled in or out of the floating gate by applying a sizable positive or negative bias at the control gate, respectively. The cell experiences a low threshold voltage (V_T) in the erased condition when no electrons are stored on the floating gate. The V_T shifts to a higher value when the device is programmed by injection of inversion electrons from the channel into the floating gate.

The cell to cell spacing, known as the pitch of the cell, is defined by the technology node of the device and the current state of the art NAND Flash memories in production use a 32 nm technology [12]. Needless to say, the high aspect ratio of these devices poses a lot of challenges not only in the manufacturing of such densely packed arrays but also in their electrostatic coupling during memory operation. There are numerous difficulties that one runs into while designing these cells at such extremely small dimensions. We would point out a few most fundamental physical limitations that increasingly make tough challenges to scaling in the future technology nodes. The introduction of trap-based or NC-based floating gates comes as a natural progression in alleviating these limitations mentioned below.



(a) **(b)**

Figure 1-2 (a) Required FG height for acceptable cell interference and the resulting gate coupling ratio (GCR) [14] and (b) Floating gate coupling (per edge) from WL, BL and diagonal side of the cell relative to the cell V_{TH} plotted against technology node [13].

a. FG – FG coupling

The electrostatic potential of floating gates is susceptible to fluctuations due to capacitive coupling from floating gates in neighboring bitlines and wordlines [15]. This is shown by the capacitive network in Fig.1-2 (a). Program/erase operation on neighboring cells tends to disturb the V_T of the victim cells. This variation leads to increased errors in read operation by diminishing the read margin of the array. Further, it may become a significant concern in implementation of multi-level cells (MLC) that store more than one bit in each cell. Fig. 1-2 (b) shows the interference generated from each WL, BL and diagonal cell FG coupling normalized to the absolute threshold voltage [13]. The cell interference is seen to approach 50 % for 20 nm cell pitch.

b. Gate Coupling Ratio (GCR)

Gate Coupling Ratio (GCR) is defined as the amount of coupling seen from the control gate to the FG divided by the total capacitance measured at the FG. Mathematically, it may be defined as follows:

$$GCR = \frac{C_{FG-CG}}{C_T} \qquad (1)$$

$$C_T = C_{FG-CG} + C_{FG-SD} + C_{FG-FG} + C_{FG-ch}$$

In scaled Flash memories consistent with ONO control oxide, parasitic coupling to source/drain ($C_{\text{FG-SD}}$), channel ($C_{\text{FG-ch}}$) and neighboring FGs ($C_{\text{FG-FG}}$) do not scale to the same extent as coupling from the control gate ($C_{\text{FG-CG}}$). The net effect of scaling to a lower technology node is therefore a reduction in GCR. GCR can be improved by realizing a tall FG and forming a cap-like control gate that boosts the coupling to the

FG. The increase in FG thickness however brings greater FG-FG coupling and also complicates high aspect ratio cell design. Further, sidewall wrapping of FG by control oxide is getting extremely difficult as cell pitch scales much faster than ONO thickness.

c. Scaling limitation on bottom oxide

The large flux of electrons in and out of the FG during program and erase generates enormous stress in the tunnel oxide [16]-[17]. Over repeated cycling, such oxides tend to generate trap sites which aid in the leakage of the stored electrons during retention. This is called as stress induced leakage current (SILC). Increased electric field on the bottom oxides can also lead to hard breakdown that generates a conducting path between the FG and the channel. One leakage path is sufficient to drain the entire charge on the conductive poly-Si FG. Given all the above reasons, in achieving a 10-year retention with a low bit error rate, bottom oxides are not deemed as reliable below 80-90 Å thickness. This restricts the scaling of program and erase voltages even as the lateral dimensions continue to scale.

Electrostatics during program and erase in Flash memories is as important as that for logic devices. Thickness of the control oxide is chosen to have an acceptable GCR as well as block any electron current during program, erase and retention operations. Limitation on scaling of tunnel oxide stands out to be the primary bottleneck in scaling vertical dimensions in the gate stack for future generations of flash memories.

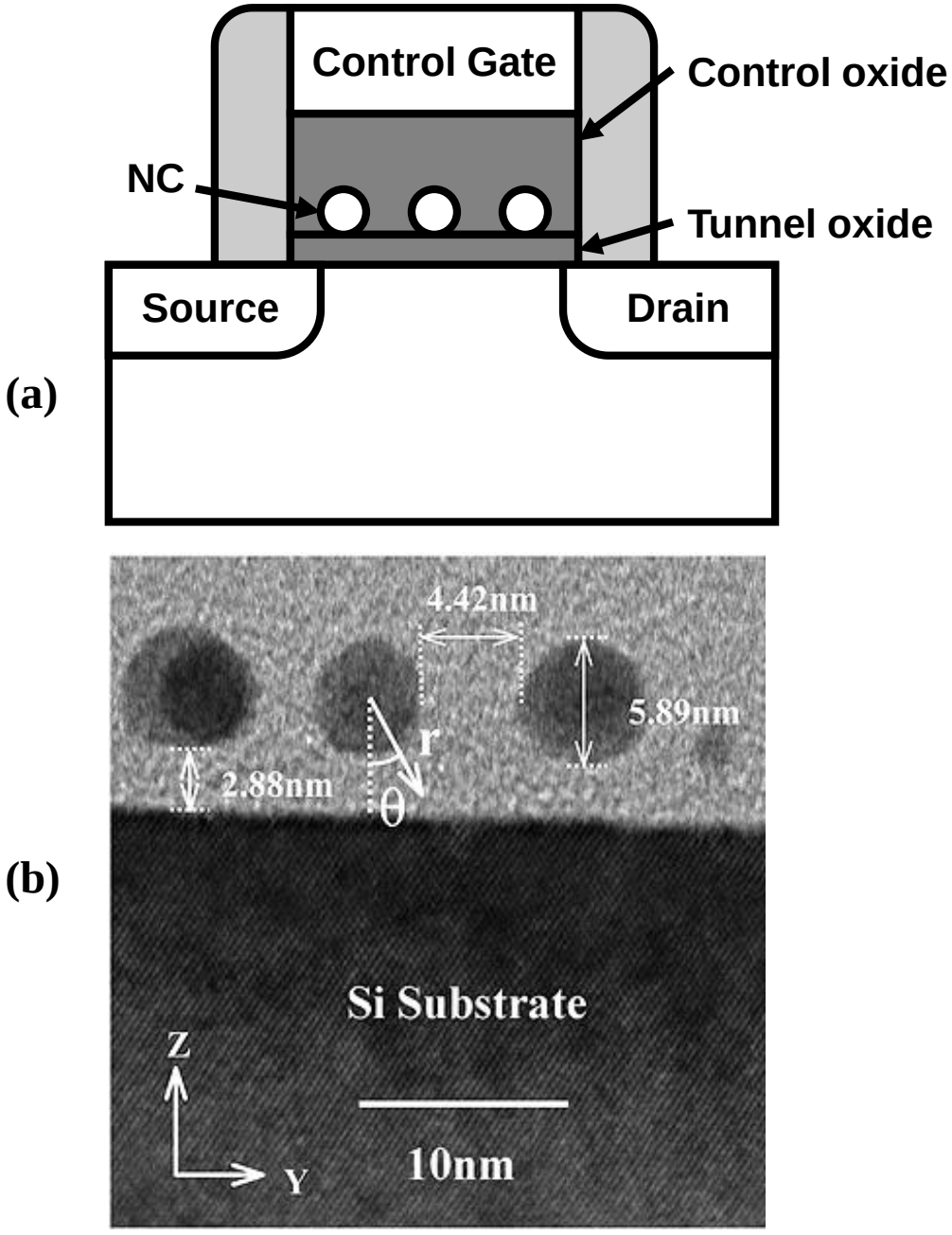


Figure 1-3 (a) Schematic of NC Flash memory (b) Cross-sectional scanning transmission electron microscopy (STEM) image of 6 nm diameter spherical Au NCs embedded in SiO_2 dielectric.

However, FGs can be made non-conducting by introducing dielectrics like silicon nitride [18] or aluminum oxide [19] that capture electrons at their interfaces or in the bulk volume. Alternatively, FGs can be comprised of NCs of semiconductors or metals. Charge captured on the trap-site or NC is localized and unaffected by leakage path generated at some other site. Thus, in the event of a localized oxide breakdown, only the charge surrounding the breakdown location is lost from the FG. In other words, FGs become immune to breakdown and SILC effects in the bottom oxide. This is the primary advantage in replacing conducting poly-Si FG with traps or NCs in future Flash technologies. Additionally, the poly-Si floating gate being highly conductive is prone to a higher FG-FG coupling coefficient. Insulating dielectrics with traps or NCs embedded inside the control oxide have a reduced capacitive coupling to FGs of the neighboring cells. Specifically, trap-sites or NCs only at the edges of the FG couple strongly to those in the neighboring cells; the rest of the trap-sites or NCs experience lower coupling from the adjacent FGs.

1.3 NC Flash Memory

Since their inception in the early 90's [20], NC-based memories have continued to stay in fashion on account of their promise in reaching full scale production at scaled nodes. Research on room temperature single electron effects and their sensing [21]-[22] was also the primary focus in developing these memories. In the span of fifteen years, there has been a considerable effort in unraveling the complete parameter space for design and fabrication of efficient and reliable NC-based Flash memories. Fig. 1-3 (a) presents the schematic of NC-based Flash memory and Fig. 1-4 (b) shows a high

resolution TEM image of 8 nm Ag NCs embedded in silicon dioxide dielectric. Immense research has been conducted in the past 10 years to establish semiconductor and metal NC Flash as a possible replacement to continuous floating gate Flash device [24]-[38]. The advantages of NC-based Flash memories may be summarized as follows:

a. Immunity against SILC enables scaling of tunnel oxide in NC-Flash below 70 Å. A thin tunnel oxide provides faster program and erase operations at reduced voltages. With a tunnel oxide thickness scaled below 30 Å, direct tunneling of electrons from the channel to the NCs becomes possible during program and erase. An electron undergoing direct tunneling does not enter the conduction band of the tunneling medium and hence offers lesser stress even after repeated cycling.

b. Charge localization on NCs lowers FG-FG coupling between the neighboring cells. This helps reduce the read disturb and improves the statistical distribution in V_T of densely packed arrays.

c. NCs are seen to have 3D geometries in a 2D assembly over the tunnel oxide [25]-[26]. They perturb the electric field around them, enhancing it in some places and diminishing it in other. The magnitude and location of enhancement depends on the dielectric constants of the NC and the dielectric surrounding it. This electrostatic enhancement is shown to offer higher programming efficiencies in NC Flash devices.

d. NC material selection is critical to retention characteristics of the device. A semiconductor (like Ge) or metal (like Pt) having conduction band offset with

Si provides deeper electron confinement inside the NC quantum well. This aids longer retention of stored charge [31].

e. Quantum confinement effect added to energy separation by Coulomb blockade may introduce a possibility step charging in NCs, an effect that would greatly benefit MLC implementation in NAND arrays [39].

1.4 The Energy, Voltage and Speed Dilemma

As the new disruptive memory technologies try to grab their piece of the nonvolatile memory market pie, let us put all the existing technologies in production today as well as in research into perspective. As the memory market proceeds to accommodate the demands of the mobile computing market in the next 10 years, there are a few performance benchmarks that need to be satisfied before their translation into full-scale manufacturing. They can be mainly placed into three categories:

a. Speed

Image and video processing applications that contribute to a majority of the memory storage capacity require extremely high bandwidths with minimal read/write latencies. The system-level speed of the memory technology not only depends on the individual program-erase time for the single device but also the array architecture and read-out circuitry. For example, a NAND-like configuration is inherently slower in read functionality over NOR architecture.

b. Power

The large asymmetry in program versus retention time arrives with a large energy overhead during P/E operations. This energy can be roughly estimated by the Joule heat ($V*I*t_{PROG}$) produced by the device. For two terminal devices like ReRAM, and PCRAM, that require large bias current (> 10 μA) during P/E [40], the energy overhead can often be demanding. For three terminal devices like Flash, majority of the power is expended in maintaining high voltages at the charge pumps. In mobile applications, the energy source is usually limited to a small battery and large P/E energy overhead is strictly undesirable.

c. Voltage

The aforementioned asymmetry in program and retention may alternatively be achieved by applying a large voltage. This is accomplished in conventional Flash devices that require > 20 V on the word lines during P/E operations. Although there is no bias current consumed at every cell level, on chip voltage amplifiers account for the majority of the power consumption. For a typical charge pump, the energy efficiency reduces with increase in the voltage amplification as well as with larger capacitance at the pumping node.

For every memory technology, the three criteria presented above are always competing against one another during optimization of performance. To understand the trade-offs involved, we need to classify the existing technologies based on the memory mechanisms and the sensing methodologies.

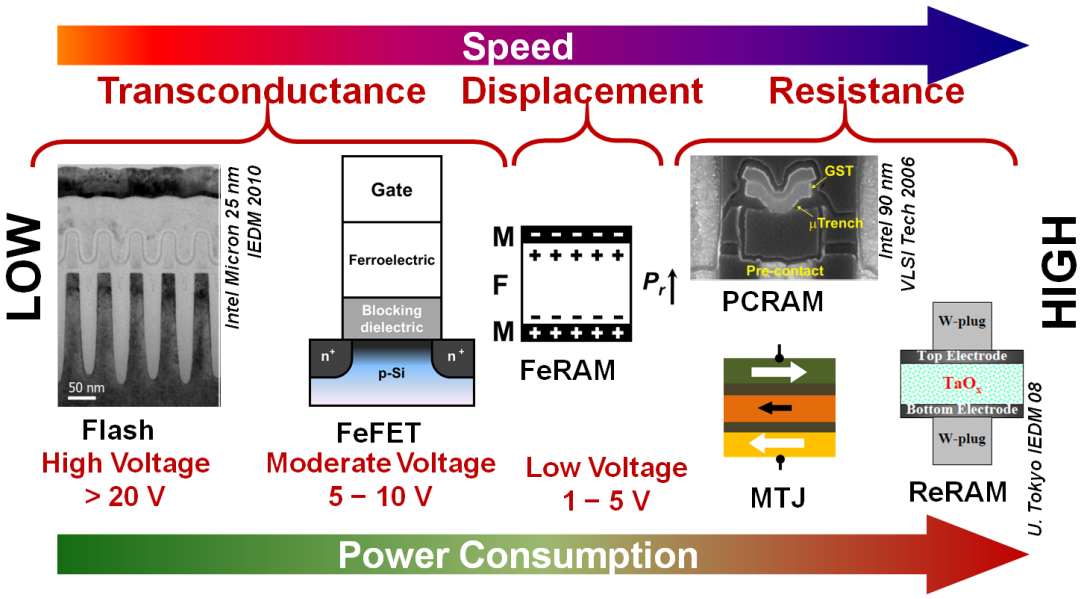


Figure 1-4 Overview of all existing memory technologies in production and research and their categorization based on device architecture and sensing methodology.

1.5 Competing Memory Technologies

According to the P/E mechanism and the read-out characteristics, the device architecture of all existing memory technologies can be divided into three types, as shown in Figure 1-4:

a. Transconductance-based sensing

These memory technologies have an underlying FET that senses the change in conductance of the channel. Evidently, this change in conductance is established by either a charge storage mechanism like Flash or a charge separation mechanism like polarization alignment in a ferroelectric (FE) film integrated in the FE-FET gate stack [41]-[43]. The transconductance-based device architecture allows non-destructive read-out with minimal interference of P/E to the sensing methodology.

Conventional Flash memories require over 25 V of P/E operating voltages. With the advent of NC Flash devices, these voltages have been demonstrated to scale below 10 V. However, charge injection through tunneling is an inherently slow process and therefore program speeds are often restricted above 10-100 µs. FE-FETs, on the other hand, have been shown to operate at moderate voltages (5-15 V). Polarization switching in inorganic oxides like PZT and SBT may be achieved below 100 ns, making FE-FETs viable for storage class memory applications [44]-[46]. However, the switching field asymmetry during short program and short retention time scales is inherently low in ferroelectric films. The depolarization field generated in the ferroelectric during retention condition is often comparable to the coercive field and opposes the direction of the remanent polarization. This depolarization effect tends to promote back-switching in the aligned atomic dipoles and reduce the memory window of the FE-FET [47]-[48].

b. Displacement-based sensing

Ferroelectric RAM devices [49] measure the displacement current across the ferroelectric capacitors to distinguish between the two polarization states similar to DRAM operation. These memories are suitable for low voltage and low density applications and have found success in their integration into card readers and RFID tags.

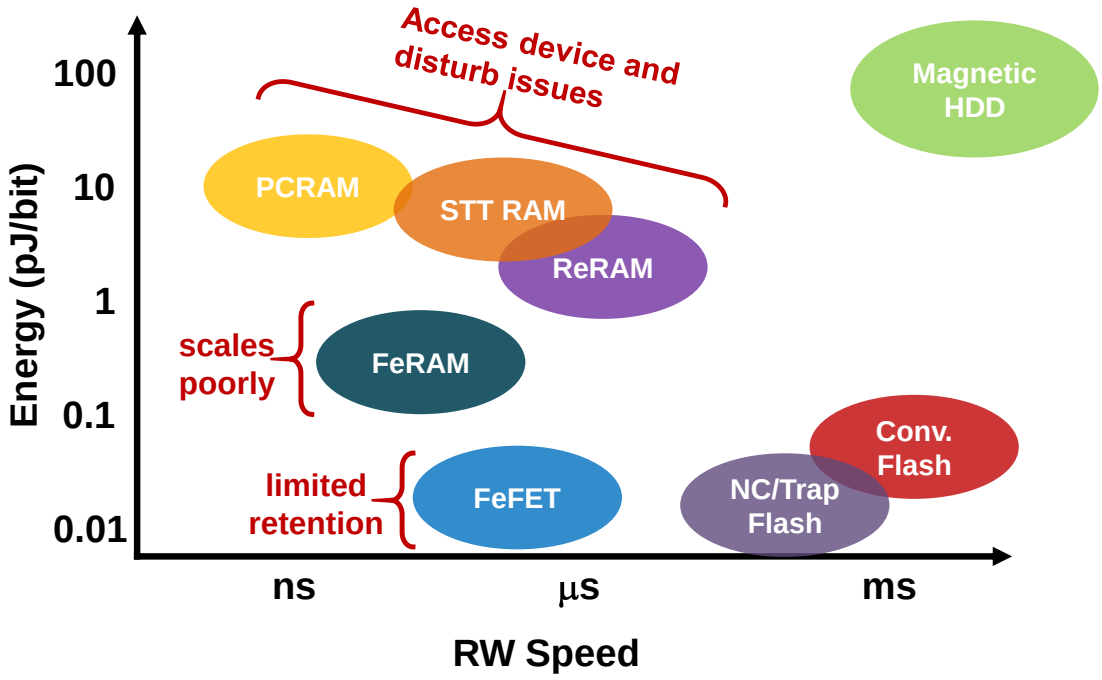


Figure 1-5 Comparison of energy required per bit during P/E operation and read-write (RW) speeds for all existing nonvolatile technologies

c. Resistance change-based sensing

These are the new emerging two-terminal memory technologies which rely on resistance switching between two stable states. The mechanism of switching differs from phase change in PCRAM, change in spin polarized tunneling current by reversal of magnetization in MTJ devices and electromigration or nano-filament growth in ReRAM devices. The simple crossbar architecture, fast switching speeds (~0.1-1 µs) and low voltage requirement of such devices makes them attractive for high-density integration [50]. However, each one of these devices relies on the bias current flowing through the device during P/E operations. As a result, switching energy for these

devices is about three orders of magnitude higher than that in Flash. Figure 1-5 depicts energy and P/E speeds of all existing technologies and their primary impediments (if any) towards their commercialization.

1.6 Scope of this work and Chapter Organization

Flash memory offers the unique advantage of large memory window, long retention and easy CMOS integration which makes it competitive to other emerging technologies. However, we need to address the inherent limitation of slow program and erase speeds in these devices. The scope of this thesis focuses on various techniques to improve the program to retention ratio in charge-trap and NC-based Flash memories.

Chapter 2 investigates the possibility of nonvolatile (NV) SRAM circuit [51] based on NC Flash memory for applications in circuits that experience frequent power outages and therefore require instant nonvolatility to preserve the current state in the register files and program counter of the microprocessor.

Chapter 3 studies the effect of ordered nanopores in the Flash memory tunnel oxide and storage layer in order to improve the program efficiencies and retention times [52]. The chapter details the electrostatic effects generated by nanoporosity, their impact on P/E and retention characteristics as well as a unique method of integration of ordered nanopores by block copolymer self-assembly in the Flash memory gate stack.

The remainder of the thesis then focuses on combining the two transconductance-based memory devices, namely Flash and FE-FETs to propose a new hybrid memory device. Chapter 4 explains the device design of the novel ferroelectric and charge hybrid nonvolatile memory and the rationale behind the complementary advantages offered by the two memory mechanisms [53]. 1-D electrostatic simulations are also presented to quantify the benefits derived from the hybrid mechanism. Chapter 5 focuses on the first generation of hybrid devices fabricated with PVDF copolymer as the ferroelectric and HfO_2 thin film as the charge trap layer [54]. The measurement results clearly established the superiority of the proposed device over conventional FE-FETs with larger memory window and longer retention times.

During program operation, the hybrid memory is shown to undergo unique dynamics in evolution of the electric field in the ferroelectric and the tunnel oxide. Electric field at the start of the program operation is shown to be enhanced in the ferroelectric and diminished in the tunnel oxide. This condition is however reversed, once sufficient polarization aligns to the applied program bias, naturally establishing a two-step program process: 1) fast DRAM-like mode arising from ferroelectric switching with limited retention and 2) slower Flash-like mode resulting from gate-injection of electrons into floating nodes that adds to the memory window in the DRAM mode and improves retention. Fabrication and measurement results confirming this dual mode hybrid memory device under low-voltage operation (second generation) by PZT and Au NC integration is presented in Chapter 6 [55]. Chapter 7 explains a generalized methodology to model the dynamics of program and retention characteristics in the hybrid devices. The model confirms the dual-speed

program process and the predicted trends show reasonable agreement with measurements performed on PZT and Au NC hybrid devices [56].

Chapter 8 highlights the main contributions of this work as well as provides suggestions for future research directions in these topics.

REFERENCES

[1] F. Masuoka, M. Asano, M. Iwahashi, T. Komuro and S. Tanaka, "A new flash E^2PROM technology using triple polysilicon technology", IEDM Tech. Dig., 1984, pp. 464-467.

[2] M. H. White, D. A. Adams and J. Bu, "On the go with SONOS", IEEE Circuits Devices Mag. pp. 22-31 (2000)

[3] C. T. Swift, G. L. Chindalore, K. Harber, T. S. Harp, A. Hoefler, C. M. Hong, P. A. Ingersoll, C. B. Li, E. J. Prinz and J. A. Yater, IEDM Tech. Dig. p. 927-930 (2002)

[4] Y. N. Tan, W. K. Chim, W. K. Choi, M. S. Joo, T. H. Ng and B. J. Cho, IEDM Tech. Dig. pp.889-892 (2004)

[5] J. D. Blauwe, "Nanocrystal nonvolatile memory devices", IEEE Tran. Nano. **1**, 72 (2002)

[6] Z. Liu, C. Lee, V. Narayanan, G. Pei and E. C. Kan, "Metal nanocrystal memories – Part I: Device design and fabrication", IEEE Tran. Elec. Dev. **49**, 1606 (2002)

[7] S. Lai, "Nonvolatile memory technology: the quest for ever lower cost", IEDM Tech. Dig., 2008, pp. 1-6.

[8] S. S. Eaton, D. B. Butler, M. Parris, D. Wilson and H. McNeillie, "A Ferroelectric nonvolatile memory", Solid-State Cir. Conf. Tech. Dig. 130 (1988)

[9] Y. Nishi and J. R. Jameson, “Recent progress in resistance change memory”, Proc. Dev. Res. Conf., pp. 271-274 (2008)

[10] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H, Nagao and H. Kano, “A novel nonvolatile memory with spin-torque transfer magnetization switching: spin-ram”, IEDM Tech. Dig. 459 (2005)

[11] D. James, “Nanoscale Flash in the mid-decade”, IEEE Adv. Semi. Manufac. Conf. 371 (2007)

[12] B. T. Park et al, “32 nm 3-bit 32 Gb Flash memory with DPT (double patterning technology) process for mass production”, Symp. VLSI Tech. 125 (2010)

[13] K. Prall, “Scaling nonvolatile memory below 30 nm”, Proc. NonVolatile Sem. Mem. Workshop, pp. 5-10 (2007)

[14] Y. Shin, “Non-volatile memory technologies for beyond 2010”, Symp. On VLSI Cir. Tech. Dig. pp. 156-159 (2005)

[15] J.-D. Lee, S.-H. Hur and J.-D. Choi, “Effects of floating-gate interference on NAND Flash memory cell operation”, IEEE Elec. Dev. Lett. **23**, 264 (2002)

[16] P. Olivo, T. Nguyen and B. Ricco, “High-field-induced degradation in ultra-thin SiO2 films”, IEEE Tran. Elec. Dev. **35**, 2259 (1988)

[17] R. Moazzami and C. Hu, “Stress-induced current in thin silicon dioxide films”, IEDM Tech. Dig. 139 (1992)

[18] C. H. Lee, K. I. Choi, M. K. Cho, Y. H. Song, K. C. Park and K. Kim, "A novel SONOS structure of SiO2/SiN/Al2O3 with TaN metal gate for multigigabit Flash memories", IEDM Tech. Dig. 613 (2003)

[19] M. Specht, H. Reisinger, F. Hofmann, T. Schulz, E. Landgraf, R. J. Luyken, W. Rosner, M. Grieb and L. Risch, "Charge trapping memory structures with Al2O3 trapping dielectric for high-temperature applications", Solid State Elec. **49**, 716 (2005)

[20] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé and K. Chan, "A silicon nanocrystals based memory", Appl. Phys. Lett. **68**, 1377 (1996).

[21] K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai and K. Seki, "Room-temperature single electron memory", IEEE Trans. Elec. Dev. **41**, 1628 (1994).

[22] S. Tiwari, F. Rana, K. Chan, L. Shi and H. Hanafi, "Single charge and confinement effects in nano-crystal memories", Appl. Phys. Lett. **69**, 1232 (1996)

[23] C. Lee, U. Ganguly, V. Narayanan, T.-H. Hou, J. Kim and E. C. Kan, "Asymmetric electric field enhancement in nanocrystal memories", IEEE Elec. Dev. Lett. **26**, 879 (2005).

[24] U. Ganguly, V. Narayanan, C. Lee, T.-H. Hou and E. C. Kan, "Three-dimensional analytical modeling of nanocrystal memory electrostatics", J. Appl. Phys. **99**, 114516 (2006).

[25] T.-H. Hou, C. Lee, V. Narayanan, U. Ganguly and E. C. Kan, "Design optimization of metal nanocrystal memory — Part I: Nanocrystal array engineering", IEEE Trans. Elec. Dev. **53**, 3095 (2006).

[26] T.-H. Hou, C. Lee, V. Narayanan, U. Ganguly and E. C. Kan, "Design optimization of metal nanocrystal memory — Part II: Gate-stack engineering", IEEE Trans. Elec. Dev. **53**, 3103 (2006).

[27] J. H. Chen, Y. Q. Wang, W. J. Yoo, Y.-C. Yeo, G. Samudra, D. SH. Chan, A. Y. Du and D.-L. Kwong, "Nonvolatile Flash memory device using Ge nanocrystals embedded in HfAlO high-k tunneling and control oxides: Device fabrication and electrical performance", IEEE Tran. Elec. Dev. **51**, 1840 (2004).

[28] J. J. Lee and D.-L. Kwong, "Metal nanocrystal memory with high-k tunneling barrier for improved data retention", IEEE Tran. Elec. Dev. **52**, 507 (2005).

[29] P. K. Singh, G. Bisht, R. Hoffman, K. Singh, N. Krishna and S. Mahapatra, "Metal nanocrystal memory with Pt single- and dual-layer NCwith low-leakage Al2O3 control dielectric", IEEE Elec. Dev. Lett. **29**, 1389 (2008).

[30] C. M. Compagnoni, D. Ielmini, A. S. Spinelli and A. L. Lacaita, "Modeling of tunneling P/E for nanocrystal memories", IEEE Tran. Elec. Dev. **52**, 569 (2005).

[31] T.-H. Hou, J. lee. J. Shaw and E.C. Kan, "Flash memory scaling: From material selection to performance improvement", Proc. Mat. Res. Soc. Symp. **1071**, F02-01 (2008).

[32] C. Compagnoni, D. Ielmini, A. S. Spinelli, A. L. Lacaita, C. Previtali and C. Gerardi, "Study of data retention for nanocrystal Flash memories", Proc. Int. Rel. Phys. Symp. pp. 506-512 (2003).

[33] C. Compagnoni, D. Ielmini, A. S. Spinelli and A. L. Lacaita, "Optimization of threshold voltage window under tunneling program/erase in nanocrystal memories", IEEE Tran. Elec. Dev. **52**, 2473 (2005).

[34] Y.-C. King, T.-J. King and C. Hu, "Charge-trap memory device fabricated by oxidation of $Si_{1-x}Ge_x$", IEEE Trans. Elec. Dev. **48**, 696 (2001).

[35] T. Baron, A. Fernandes, J. F. Damlencourt, B. De Salvo, F. Martin, F. Mazen and S. Haukka, "Growth of Si nanocrystals on alumina and integration in memory devices", Appl. Phys. Lett. **82**, 4151 (2003).

[36] K. C. Scheer, R. A. Rao, R. Muralidhar, S. Bagchi, J. Conner, L. Lozano, C. Perez, M. Sadd and B. E. White, "Thermal oxidation of silicon nanocrystals in O2 and NO ambient", J. Appl. Phys. **93**, 5637 (2003).

[37] M. W. Stoker, T. P. Merchant, R. Rao, R. Muralidhar, S. Straub and B. E. White, "A model of silicon nanocrystal nucleation and growth on SiO2 by CVD", Proc. Mat. Res. Soc. Symp. **830**, D5.7.1 (2005).

[38] C. Lee, J. Meteer, V. Narayanan and E. C. Kan, "Self-assembly of metal nanocrystals on ultrathin oxide for nonvolatile memory applications", J. Elec. Mat. **34**, 1 (2005).

[39] A. Thean and J. P. Leburton, "Stark effect in single electron charging in silicon nanocrystal quantum dots", J. Appl. Phys. **89**, 2808 (2001).

[40] K. Tsunoda, K. Kinoshita, N. Noshiro, Y. Yamazaki, T. Iizuka, Y. Ito, A. Takahashi, A. Okano, Y. Sato, T. Fukano, M. Aoki and Y. Sugiyama, "Low power and high speed switching of Ti-doped NiO ReRAM under the unipolar

voltage source of less than 3 V," Int. Elec. Dev. Meeting Tech. Dig., pp. 767-770 (2007).

[41] S. Mathews, R. Ramesh, T. Venkatesan and J. Benedetto, "Ferroelectric field effect transistor based on epitaxial perovskite heterostructures," Science, **266**, pp. 238–240, (1997).

[42] T. P. Juan, C. Chang and J. Y. Lee, "A new metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (Dy_2O_3)–semiconductor (MFIS) FET for nonvolatile memory applications," IEEE Elec. Dev. Lett., **27**, pp. 217-220, (2006).

[43] Y. Watanabe, M. Tanamura and Y. Matsumoto, "memory retention and switching speed of ferroelectric field effect in (Pb,La)(Ti,Zr)O3/La2CuO4:Sr hetereostructure," Jpn. J. Appl. Phys., **35**, pp. 1564–1568, (1996).

[44] S. Sakai, M. Takahashi, K. Takeuchi, Q. –H. Li, T. Horiuchi, S. Wang, K. –Y. Yun, M. Takamiya and T. Sakurai, "Highly scalable Fe (ferroelectric)-NAND cell with MFIS (metal-ferroelectric-insulator-semiconductor) structure for for sub-10 nm terabit capacity NAND Flash memories" International Mem. Workshop pp.103 (2008).

[45] M. H. Tang, Z. H. Sun, Y. C. Zhou, Y. Sugiyama and H. Ishiwara, "Capacitance-voltage and retention characteristics of $Pt/SrBi_2Ta_2O_9/HfO_2/Si$ structures with various buffer layer thicknesses," Appl. Phys. Lett., **94**, p. 212907, (2009).

[46] M. Tang, X. Xu, Z. Ye, Y. Sugiyama and H. Ishiwara, "Impact of HfTaO buffer layer on data retention characteristics of ferroelectric-gate FET for

nonvolatile memory applications," IEEE Trans. Electron Devices, **58**, pp. 370−375, (2011).

[47] R. R. Mehta, B. D. Silverman and J. T. Jacobs, "Depolarization fields in thin ferroelectric films," J. Appl. Phys., **44**, pp. 3379−3385, (1973).

[48] T. P. Ma and J. –P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," IEEE Electron Device Lett., vol. **23**, pp.386−388, (2002).

[49] R. Moazzami, C. Hu and W. H. Shepherd, "Electrical characteristics of ferroelectric PZT thin films for DRAM applications," IEEE Tran, Elec. Dev., **39**, pp. 2044–2049, (1992).

[50] R. F. Freitas and W. W. Wilcke, "Storage–class memory: The next storage system technology," IBM J. Res. Dev. 52, no. 4/5, pp. 439–447 (2008).

[51] S. Rajwade, W. Yu, Q. Xu, T.-H. Hou, G. E. Suh and E. C. Kan, "Low power nonvolatile SRAM circuit with integrated low voltage nanocrystal PMOS Flash", Proc. of the 24th IEEE SOC Conference (SOCC), Las Vegas, Sept 2010.

[52] S. Rajwade, H. Arora, J. Shaw, U. Wiesner and E. C. Kan, "Effects of engineered nanopores on charge-trap nonvolatile memory," submitted to Appl. Phys. Lett.

[53] S. Rajwade, K. Auluck, J. Phelps, K. Lyon and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part I: Device concept and modeling," IEEE Tran. Elec. Dev., **59**, pp. 441–449, (2012).

[54] S. Rajwade, K. Auluck, J. Phelps, K. Lyon and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part II: Experimental validation and analysis," IEEE Tran. Elec. Dev., **59**, pp. 450–458, (2012).

[55] S. Rajwade, T. A. Naoi, K. Auluck, K. Jayant, R. B. van Dover and E. C. Kan, "Ferroelectric-assisted dual-switching-speed DRAM-Flash hybrid memory," submitted to IEEE Tran. Elec. Dev.

[56] S. Rajwade, K. Auluck, T. A. Naoi, K. Jayant and E. C. Kan, "Modeling of program and retention dynamics in ferroelectric and charge hybrid memory," submitted to J. Appl. Phys.

CHAPTER 2 NONVOLATILE SRAM CIRCUIT USING LOW VOLTAGE NANOCRYSTAL PMOS FLASH

2.1 Abstract

This chapter presents a new nonvolatile SRAM design that incorporates low voltage nanocrystal PMOS Flash transistors. The design enables global store, restore and erase operation with negligible penalty on regular SRAM operation. Store/erase operations also do not consume power except in possible charge pump circuits. Circuit simulations based on experimental I-V characteristics demonstrate that 10 μs store/erase operation at ± 6 V is sufficient for correct restoration of the stored bit even under reasonable process variation.

2.2 Introduction

Low power systems with unreliable power supplies can greatly benefit from nonvolatile (NV) SRAMs. Instance check-pointing can eliminate static power consumption and enable continuous operation across power supply failures. Several designs that incorporate Resistive (Re) RAM [1]-[3], Phase Change (PC) RAM [4], Magnetic (M) RAM [5] and Ferroelectric memory [6] with SRAMs have been previously proposed. Although ReRAM and PCRAM provide low voltage operation, they rely on accurate high current pulses to switch their nonvolatile states. This puts additional burden over the on-chip power supply as well as increases design complexity of peripheral circuitry. In order to be truly viable, NV-SRAMs must not only be CMOS compatible but also have minimal performance and power overheads.

In the last decade, NAND Flash memory has proved to be the single most driving force in enabling high density nonvolatile storage in mobile applications. The high voltage program and erase operation for these Flash devices however necessitates off-chip data storage. This chapter proposes a novel NV-SRAM design which integrates low voltage nanocrystal (NC) Flash in every cell [7]. The circuit enables regular SRAM operation in stable power supply as well as performs store, restore and erase operations through globally generated interrupts to enable seamless computation over subsequent power suspensions. The design shows great potential for embedded SoC applications that inherit unstable power supply and therefore require low power and low voltage operation.

The chapter is organized as follows: Section 2.3 discusses NC Flash briefly; Section 2.4 describes NV-SRAM design and operation, Section 2.5 evaluates performance characteristics, Section 2.6 illustrates dependence of store/erase time on process variation to ensure correct restoration and lastly, Section 2.7 validates low power advantage of using NC Flash.

2.3 Nanocrystal Flash

To ensure ten year retention of the stored charge, scaling of tunnel oxide in conventional polysilicon floating gate Flash is restricted to 7-9 nm owing to stress induced leakage currents. This constraint can however be eliminated by use of discrete charge storage floating gates like nitride traps [8] or nanocrystals (NCs) [9]. Reducing tunnel oxide thickness to 2-3 nm can bring down program/erase voltages of such devices below 8 V [7]. Further, metal NCs provide significant electric field

enhancement in the tunnel oxide during program/erase assisting faster store/erase operation [10]. On chip high voltage generation of such magnitudes is achievable through charge pump circuits powered by supply voltage. Since the high voltages generated by charge pumps falls across a high impedence gate of Flash device, they can be designed efficiently for minimal power dissipation.

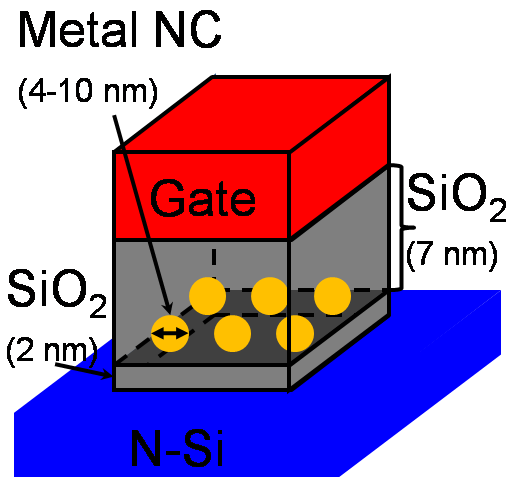


Figure 2-1 Schematic of NC PMOS Flash device incorporating metal NCs as nano-floating gates

Fig. 2-1 shows the schematic of PMOS NC Flash device. The PMOS NC Flash is programmed by applying a high positive voltage (~ 6 V) which puts the device in deep accumulation and enables tunneling of electrons from the n-substrate to the floating NC gate. Trapped electrons in the NCs favor easier depletion in the channel with applied negative bias (- 1 V with respect to bulk), shifting the threshold voltage in the positive direction (V_{TH} of PMOS is negative). The device is erased by applying a high negative voltage on the gate which pushes the electrons trapped in the NCs back to the substrate and restores V_{TH}.

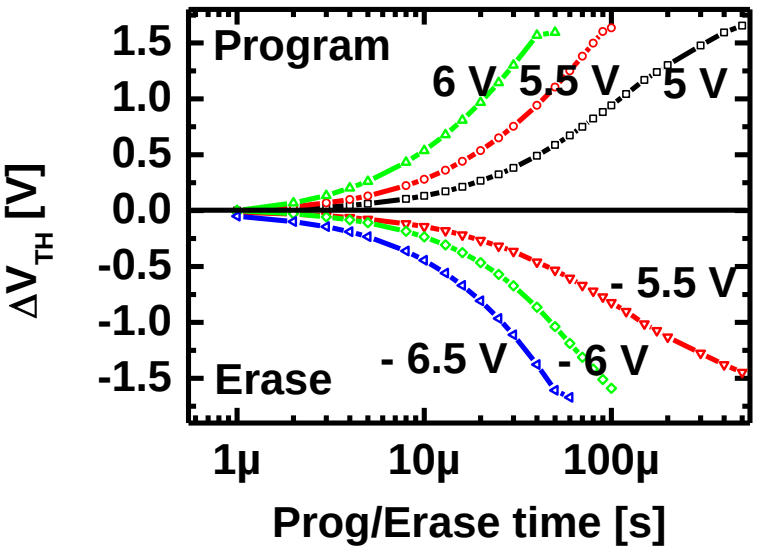


Figure 2-2 Simulated ΔV_{TH} against program/erase time for NC PMOS Flash incorporated in the NV-SRAM circuit

NC devices are shown achieve fast program/erase at low operating voltages as well as achieve a ten year retention mark. Memory window (ΔV_{TH}) against program/erase time is modeled based on experimental measurements performed on NC Flash devices on NMOS substrates and following the procedure outlined in [11]. Fig. 2-2 illustrates the simulation of memory window for three different program/erase voltages against time. The device is seen to achieve ΔV_{TH} of 0.5 V with a program (erase) voltage of 6 V (- 6.5 V) in less than 10 μs.

2.4 NV-SRAM Design and Operation

Fig. 2-3 presents the NV-SRAM design with NC PMOS Flash transistors. Nodes Q and Qb are loaded with the Flash transistors controlled by the PE (program-erase) signal. They are accessed by the NMOS transistors driven by EN signal. Circuit simulations are performed in SPICE with 70 nm BSIM3v3 model [12]-[13] and low power (high V_{TH}) design at V_{DD} = 1 V operation. The following subsections explain the working of the cell in detail.

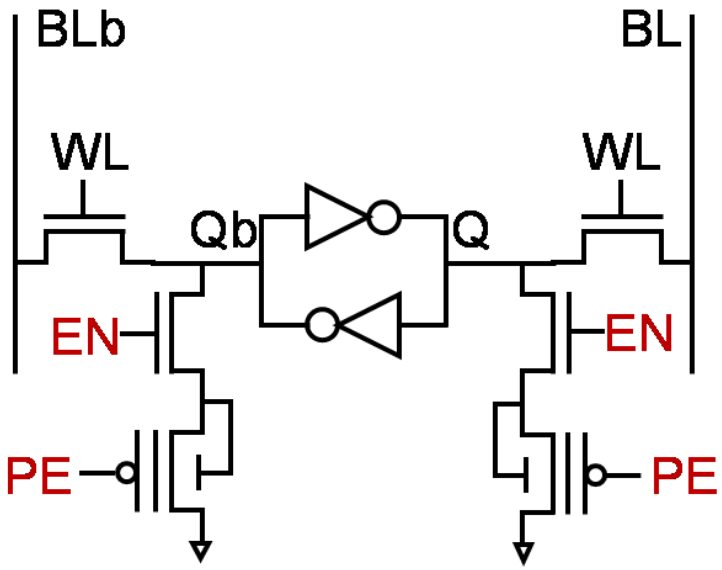


Figure 2-3 NV-SRAM cell integrating NC PMOS Flash devices controlled by the PE signal. They are accessed through NMOS pass transistors enabled by EN. Appropriate global switching of PE and EN signals achieves store/erase operation for each cell

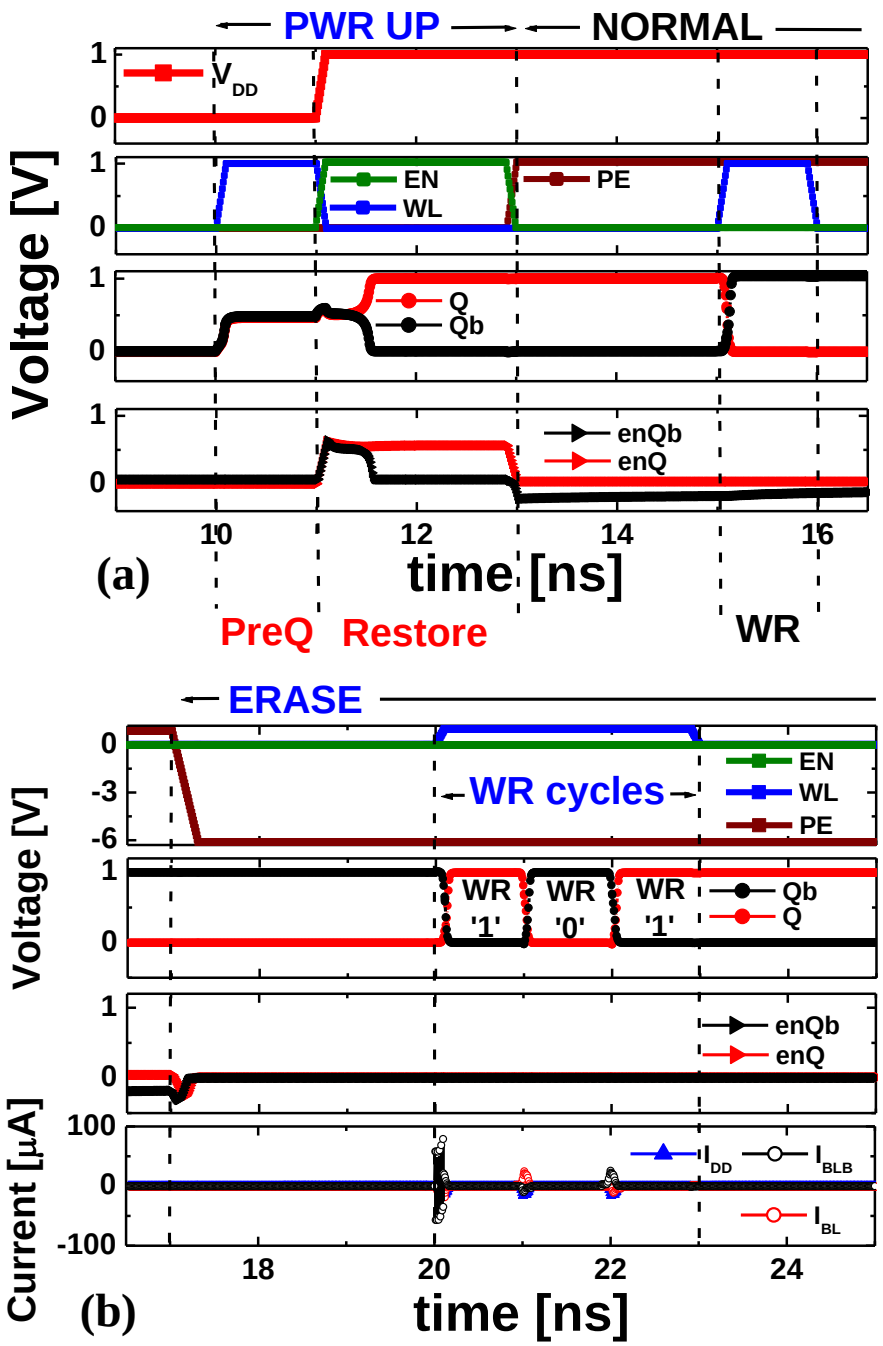


Figure 2-4 (a) Timing diagram for POWER UP operation. The cell stores bit '1' (Q = 'HIGH') before powering down (b) Timing diagram for ERASE operation

2.4.1. REGULAR Operation

The PMOS Flash transistors as well as the NMOS access transistors are turned off in REGULAR operation (EN = 0 V; PE = 1 V). The cell therefore resembles a volatile SRAM barring the extra capacitive loading at nodes Q and Qb.

2.4.2. STORE Operation

The memory controller initializes the STORE operation on sensing power supply failure. During this operation the WL is disabled and EN is enabled. The body bias of the PMOS Flash is now controlled by the respective Q or Qb node. A suitable program voltage (5 – 7 V) is applied on PE leading to tunneling of accumulated electrons into the floating NC gates. The tunneling flux of electrons is exponentially dependent on the voltage drop between the body and the floating NC gate of the PMOS Flash transistors. Therefore, the PMOS Flash with 'LOW' body bias is 'programmed' by exponentially higher electron injection than the other Flash device. V_{TH} of the programmed device shifts significantly in the positive direction. Both the PMOS Flash transistors remain in accumulation (switched off) and the cell does not consume any bias current through the Q and Qb branch, in exception to any subthreshold leakage. At the end of the STORE operation (~ 10 μs), the Flash transistors acquire a significantly different V_{TH} and the cell is ready to be powered down.

2.4.3. POWER UP Operation

The POWER UP operation achieves the restoration of the stored state in every cell immediately after revival of the power supply. Fig. 2-4(a) shows the detailed

sequence of control signals performing this global restoration. The POWER UP sequence begins with the precharge (PreQ) cycle, during which nodes Q and Qb are precharged to $V_{DD}/2$ by enabling all the WLs. This is followed by switching the EN signal as well the power supply V_{DD} in the restore cycle. The EN signal puts the access NMOS transistor in above threshold regime. The resistance of nodes Q and Qb to ground is therefore determined by the subthreshold leakage through their respective PMOS Flash transistors. The 'programmed' Flash offers lower subthreshold resistance to ground node. The internal feedback generated between nodes Q and Qb by the cross coupled inverter pair amplifies this asymmetry to restore 'LOW' state in the 'programmed' PMOS Flash branch. The cell resumes normal (REGLAR) operation from the subsequent cycle.

The time elapsed in amplifying the small resistive difference into rail to rail output depends on the V_{TH} difference between the 'programmed' and the 'non-programmed' PMOS Flash as well as the small signal gain of the inverters. For the low power (high V_{TH}) design in our cells, this time is observed to be 200 - 400 ps, but may be reduced significantly by using high performance transistor design in the inverter pairs.

2.4.4. ERASE Operation

The ERASE operation restores the V_{TH} of all PMOS Flash transistors. The asymmetry in 'programmed' and 'non-programmed' PMOS Flash does not participate in the REGULAR operation of the SRAM cell. Therefore, the memory controller has the liberty to perform ERASE during periods of low power utilization. Figure 2-4(b)

presents the timing diagram of the cell during ERASE operation. A suitable high negative voltage is applied at PE (-6 to -8 V) causing electrons trapped in floating NCs to tunnel back to the substrate. The PMOS Flash device is in deep inversion (low resistance) during this operation. The NMOS access transistor is tuned off (EN = 0 V) to ensure no bias current flows in either of the branches. Hence, no additional power is expended on erasing the stored state of the cell except in the high voltage charge pump circuits. The ERASE operation returns the V_{TH} of all PMOS Flash devices to the 'non-programmed' state.

Table 2-1 Scheme of Operation for proposed NV-SRAM cell

Operation	V_{DD} (V)	WL (V)	PE (V)	EN (V)	Notes
REGULAR	1	1/0	1	0	6 % performance penalty to read/write
STORE	1	0	5 to 6	1	Program time ~10 – 100 μs; Negligible power overhead
POWER UP	1 (delay)	1	0 to -1	1	V_{DD} enabled 1 cycle after PreQ
ERASE	1	1/0	-6 to -7	0	Erase time ~10 – 100 μs; Runs parallel to REGULAR operation with minimal power overhead

The NMOS access transistors also achieve a dual purpose. They decouple the ERASE operation from the normal functioning of the SRAM cell; in other words allow REGULAR operation to run in parallel to ERASE. Fig. 2-4(b) shows the write (WR) cycles performed in parallel with ERASE. The bulk and source potential of the PMOS Flash devices (nodes enQ and enQb) is held at ground potential due to the highly conductive inversion layer in the device.

It should be noted that PMOS Flash cannot be replaced with conventional NMOS Flash device. PMOS (NMOS) Flash maintains a high resistance state during program (erase) and a low resistance state during erase (program). This inherent difference makes PMOS inevitable to disable any static current during STORE and ERASE operation. Table 2-1 summarizes all the operations of the proposed NV-SRAM.

2.5 Performance Evaluation

The NV-SRAM design achieves nonvolatile backup and restore by means of globally controlled STORE, POWER UP and ERASE operations at negligible power penalty. The cell footprint is larger by 4 minimum sized transistors. Circuit simulations were performed at 1 GHz clock frequency, although this is not restricted by the capacitive loading at Q and Qb nodes. Frequency of operation is set by the low power (high V_{TH}) transistor switching time. The capacitive loading however is seen to offer a 6 % penalty to REGULAR operation.

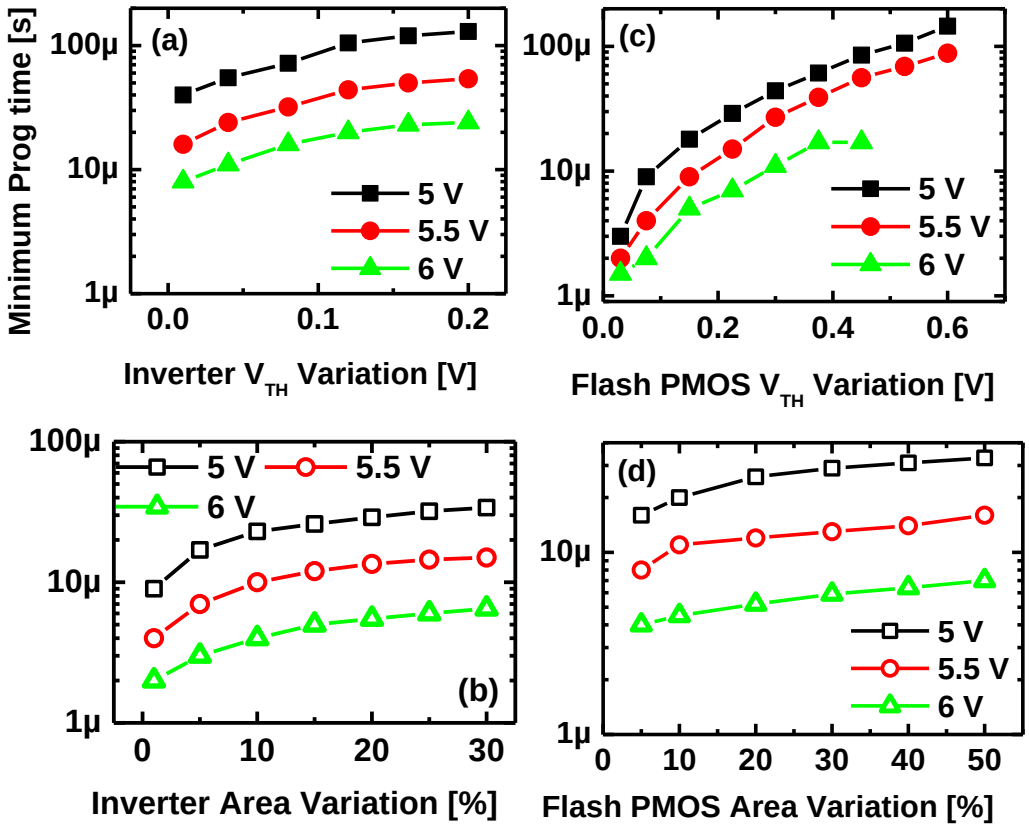


Figure 2-5 Dependence of minimum program time (STORE operation) on worst case (strong-p, weak-n) (a) V_{TH} mismatch and (b) area mismatch in inverter pair. Dependence of minimum time for STORE operation on (c) V_{TH} and (d) area mismatch in PMOS Flash transistors.

2.6 Process Variation and Circuit Performance

In the event of power suspension, the memory controller must utilize the remaining stored energy efficiently to maximize the backup volume of the current system state. In other words, time expended in high voltage STORE operation must be minimized. Minimum time for STORE operation, that ensures correct restoration of

the stored bit in the subsequent POWER UP operation, is determined by the degree of matching in transistor pairs within a cell.

Transistors in a single cell are subject to area and V_{TH} mismatch as a result of process variation. Therefore, difference in resistance to ground node seen from Q and Qb nodes during POWER UP operation must be sufficient to overpower any opposing latch-up condition occurring in the cross-coupled inverter pair due to device mismatch. Fig. 2-5(a) shows the dependence of minimum time for program to compensate for opposing latch-up in the worst case V_{TH} mismatch (strong-p and weak-n or vice versa) in inverter transistors. Fig. 2-5(b) illustrates the same constraint under worst case area mismatch in inverter transistors.

Fig. 2-5(c) and 2-5(d) demonstrate the minimum time required in STORE operation under V_{TH} and area mismatch in PMOS Flash transistor for correct restoration of the stored bit. Since program time has an inverse exponential relation with program voltage, V_{TH} variation is seen to be critical to minimum program time and especially serious to lower program voltages.

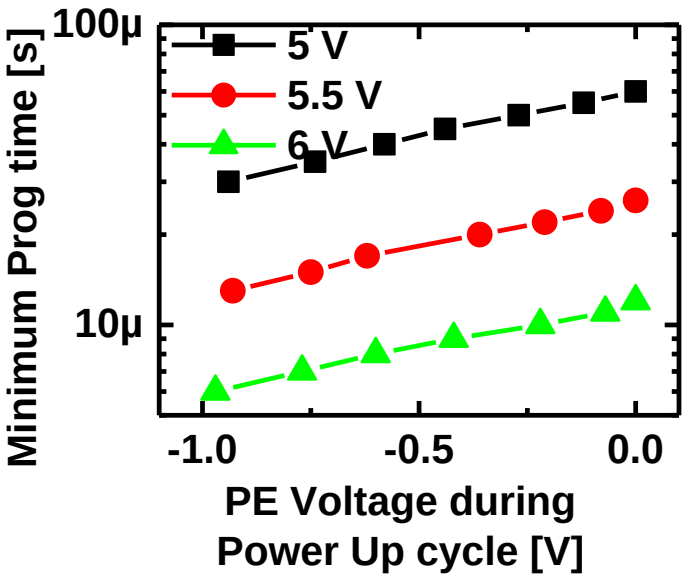


Figure 2-6 Simulation showing reduction in minimum time of STORE operation with decreasing PE bias

Flash devices are prone to V_{TH} variation due to effective oxide thickness fluctuation. During POWER UP, PMOS Flash operates in subthreshold and is the most resistive device in the ground path. This resistance decreases exponentially with the applied negative bias at PE and may compensate for device mismatch effectively. In other words, the asymmetry in the conductance to ground node can be boosted by applying a small negative bias on PE (but still keeping the device in subthreshold) which helps overpower opposing force resulting from inverter mismatch. This helps bring down the lower bound on STORE operation time. Fig. 2-6 illustrates the reduction in program time with applied negative bias at PE during POWER UP for a 0.1 V worst case mismatch in inverter pair.

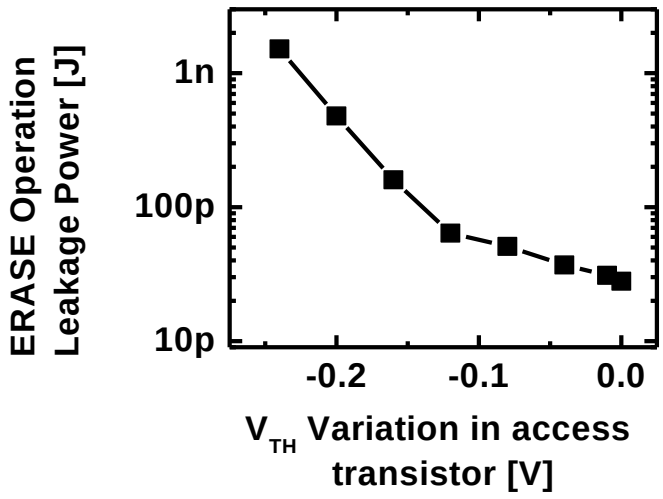


Figure 2-7 Leakage power during ERASE operation due to V_{TH} lowering in the NMOS access transistor

Leakage characteristics of NMOS access transistor become critical during ERASE operation. As seen from Fig. 2-7, low V_{TH} of the access transistor results in exponentially higher power dissipation stemming from the leakage through the device. NMOS access transistors should be designed to provide low leakage at zero bias (V_{TH} > 0.4 V).

Table 2-2 Comparative Study of Estimated Power Dissipation during STORE operation for various proposed NV-SRAMs

NV-SRAM Design	Estimated STORE operation Power (μW)
ReRAM	24.6
PCRAM	378
MRAM	32.2
FeRAM	0.124
This work	0.075

2.7 Power Evaluation

The STORE and ERASE operations in the proposed NV-SRAM design offer negligible power overhead at the cell level. Most other designs that incorporate ReRAM, PCRAM or MRAM consume high bias currents during these operations [14]-[15]. Delivering precise high bias currents introduces complexity of current mirrors, temperature compensation techniques and variable I-R drops in the word lines. On the other hand, generation of on-chip high voltages is well established by efficient charge pump designs [16]. This unique advantage in our design presents no additional complexity to the peripheral circuitry including decoder design.

Power estimates for the proposed design were performed with charge pump circuits included in the design. Table 2-2 presents a comparative summary of

estimated energy dissipation at 70 nm node for STORE operation against proposed NV-SRAM design which includes the high PE generating charge pumps.

2.8 Conclusion

This work has proposed a new low power low voltage NV-SRAM design with global STORE, RESTORE and ERASE operations with minimal power overhead. The minimum time required for critical STORE operation was evaluated under process variations. It offers 10 μs store/erase at ± 6 V operation with no additional cell level power dissipation. The design shows promise in low power embedded SoC architectures with minimal performance penalty.

REFERENCES

[1] W. Wang, A. Gibby, Z. Wang, T. W. Chen, S. Fujita, P. Griffin, Y. Nishi and S. Wong, "Nonvolatile SRAM cell," in IEDM Tech. Dig., Dec. 2006, pp. 1-4.

[2] X. Xue, G. Jin, J. Zhang, L. Xu, Y. Ding, Y. Xie, C. Zhao, B. Chen and Y. Lin, "Nonvolatile SRAM cell based on CuxO," in ICSICT Dig. Tech. Papers, Oct. 2008, pp. 869-871.

[3] S. Yamamoto, Y. Shuto and S. Sugahara, "Nonvolatile SRAM (NV-SRAM) using Functional MOSFET merged with resistive switching devices," in Proc. CICC, Sept. 2009, pp. 531-534.

[4] M. Takata, K. Nakayama, T. Izumi, T. Shinmura, J. Akita and A. Kitagawa, "Nonvolatile SRAM based on phase change," in Proc. Nonvolatile Semicond. Mem. Workshop, Feb. 2006, pp. 95-96.

[5] N. Sakimura, T. Sugibayashi, R. Nebashi and N. Kasai, "Nonvolatile magnetic flip-flop for standby-power-free SoCs," IEEE J. Solid-State Circuits, vol. 44, no. 8, pp. 2244-2250, Aug. 2009.

[6] T. Miwa, J. Yamada, H. Koike, H. Toyoshima, K. Amanuma, S. Kobayashi, T. Tatsumi, Y. Maejima, H. Hada and T. Kunio "NV-SRAM: A Nonvolatile SRAM with backup ferroelectric capacitors," IEEE J. Solid-State Circuits, vol. 36, no. 3, pp. 522-527, Mar. 2001.

[7] J. Lee and D. Kwong "Metal nanocrystal memory with high-κ tunneling barrier for improved data retention," IEEE Trans. Electron Devices, vol. 52, no. 4, pp. 507-511, Apr. 2005.

[8] M. H. White, D. A. Adams and J. Bu, “On the go with SONOS,” IEEE Circuits Devices Mag., pp. 22-31, Jul. 2000.

[9] S. Tiwari, F. Rana, K. Chan, H. Hanafi, W. Chan and D. Buchanan, “Volatile and nonvolatile memories in silicon with nano-crystal storage,” in IEDM Tech. Dig., Dec. 1995, pp. 521-524.

[10] C. Lee, U. Ganguly, V. Narayanan, T. Hou, J. Kim and E. C. Kan, “Asymmetric electric field enhancement in nanocrystal memories,” *IEEE Electron Device Lett*., vol. 26, no. 12, pp. 879-881, Dec. 2005.

[11] T. Hou, C. Lee, V. Narayanan, U. Ganguly and E. C. Kan, “Design optimization of metal nanocrystal Memory–Part I: nanocrystal array engineering” *IEEE Trans. Electron Devices*, vol. 53, no. 12, pp. 3095-3102, Dec. 2006.

[12] Device Group, UC Berkeley, “Berkeley Predictive Technology Model.”

[13] L. Larcher, P. Pavan, S. Pietri, L. Albani and A. Marmiroli, “A new compact DC model of floating gate memory cells without capacitive coupling coefficients,” *IEEE Trans.Electron Devices*, vol. 49, no. 2, pp. 301-307, Feb. 2002.

[14] I. G. Baek, M. S. Lee, S. Seo, M. J. Lee, D. H. Seo, D. Suh, J. C. Park, S. O. Park, H. S. Kim, I. K. Yoo, U. Chung, and J. T. Moon, "Highly scalable non-volatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses," *IEDM Tech. Dig.*, Dec. 2004, pp. 587 - 590.

[15] S. Lai, “Current status of phase change memory and its future,” in *IEDM Tech. Dig*., Dec. 2003, pp. 255-258.

[16] K. Choi, J. Park, J. Kim, T. Jung and K. Suh, “Floating-well charge pump circuits for sub-2.0 V single power supply Flash memories,” in *Symp. VLSI Circuits Dig. Tech. Papers*, June 1997, pp. 61-62.

CHAPTER 3 EFFECTS OF ENGINEERED NANOPORES IN CHARGE-TRAP NONVOLAILTE MEMORY

3.1 Abstract

This chapter investigates the effects of controlled nanoporosity in the tunnel oxide and the storage layer of charge–trap–based nonvolatile memory. Electric field and the resulting tunneling current is enhanced in regions between the nanopores (NPs), thereby constraining the location for charge capture. Aluminosilicate films with 8–10 nm diameter NP arrays formed using block copolymer self–assembly were integrated in the Flash memory gate stack. Samples with NPs demonstrated higher injection efficiency, lower memory window and longer retention over control devices without NPs. Measurement of conductance through the NP–film confirmed the strong enhancement observed in program efficiencies.

3.2 Introduction

The conventional poly–Si floating gate Flash memory scaling is soon predicted to reach its fundamental scaling limitations. Inter–cell coupling due to floating gate interference can cause severe read and program failures [1]. The high program/erase (P/E) voltages applied between adjacent word lines severely affect the oxide reliability as cell–pitch continues to scale below 20 nm [2]. Discrete charge storage in nanocrystals or thin dielectric traps–layers alleviates this limitation to an extent, by enabling aggressive scaling of the tunnel oxide [3]-[5]. As lateral dimensions keep shrinking, variability in device performance becomes the key bottleneck towards

commercialization of such new technologies. Specifically for trap–based memories like SONOS [5], TANOS [6] and other dielectrics with high trap densities, the discrete nature of the electron trapping and detrapping event results in inherent statistical variation of observed memory window [7] (ΔV_{TH}). One of the primary reasons for such variability is the uncertainty in the charge–trap location during each P/E cycle. Further, trap sites also have a statistical spread in their capture cross–sections and trap–energy levels with significant dependence on process parameters. Lastly, the location of the trap–site also influences the ease of capture and the resulting ΔV_{TH} of the programmed cell. For example, electron capture is favored at the edge traps along the active area due to the enhanced fringing fields. Also, electron trapped at the source–injection barrier is likely to generate higher ΔV_{TH} than other sites [8]. Such effects become increasingly dominant with shrinking device geometry.

This chapter studies the effects of spatial asymmetry in the electric field during P/E operations on charge–trap memory by engineered nanopores (NPs). Ordered porosity in the trap layer was introduced by block copolymer self–assembly [9] over the tunnel oxide in the gate stack. In recent years, researchers have demonstrated block copolymer self–assembly as a viable option for low–cost bottoms–up electronics [10]. Particularly, this approach was successfully used in obtaining ordered arrays of metal nanocrystals [11]-[12]. This work modifies the same method to produce 8–10 nm diameter NP arrays embedded in aluminosilicate matrix over the tunnel oxide. NP charge–trap devices were benchmarked against control devices without NPs by *CV* measurements to evaluate the effect of porosity on ΔV_{TH} and retention [13]. Pulsed program measurements were used to quantify the enhancement in electrostatics and

the resulting influence on tunneling efficiencies. Conductance measured directly through the spun cast aluminosilicate film provided useful insights on the influence of NPs in charge transport mechanisms during program and retention states.

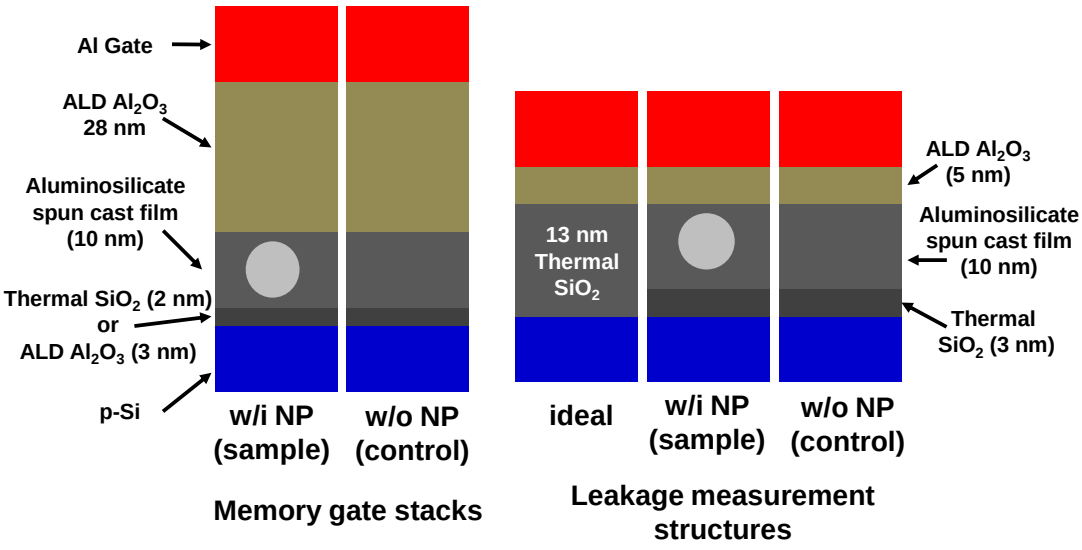


Figure 3-1 Overview of all the devices fabricated with the self assembly process. Sample devices with NPs (w/i NP) as well as control devices without NPs (w/o NPs) were designed for all memory and leakage structures to calibrate the improvement in performance

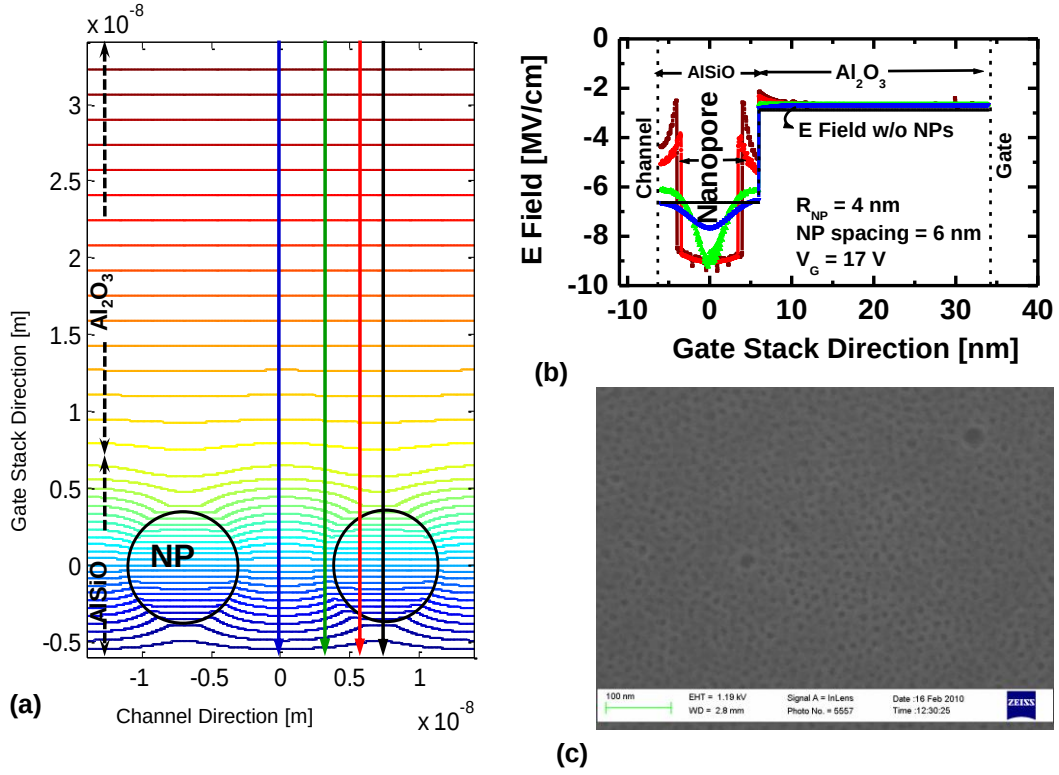


Figure 3-2 (a) Electric potential contours for the aluminosilicate (AlSiO) sample NP memory structure with 2 nm SiO_2 tunnel dielectric (shown in Fig. 3-1) during program condition. (b) Vertical electric field along cut lines (shown in Fig. 3-2 (a)) in the NP gate stack. Electron capture is preferred in high-field regions between the NPs. (c) A representative SEM image of a calcined aluminosilicate film revealing 8–10 nm NPs.

3.3 Electrostatics Effects

Fig. 3-1 portrays the schematic of the fabricated memory gate stacks and leakage measurement structures. Memory gate stacks comprised of 2 nm SiO_2 (simulated in Fig. 3-2 (a) and (b)) or 3 nm Al_2O_3 as tunnel dielectric. Sample devices with nanopores (w/i NP) consisted of 8–10 nm diameter NPs in aluminosilicate matrix spun–cast over the tunnel dielectric by block copolymer self–assembly. Control

devices consisted of homogeneous aluminosilicate matrix without nanopores (w/o NP). Al_2O_3 was chosen as the control dielectric with 28 nm thickness. For leakage measurement structures, self–assembly was performed on 3 nm thermal SiO_2 followed by deposition of 5 nm ALD Al_2O_3 capping layer. Measurements from an ideal tunneling structure that replaced the aluminosilicate film with thermal SiO_2 were used to benchmark the leakage through the spun–cast layer.

Fig. 3-2 (a) depicts the simulation of potential contours during program condition for the memory gate stack with NPs assembled over 2–nm SiO_2 tunnel oxide. The simulation assumes 8–nm diameter NPs arranged in hexagonal symmetry over the tunnel oxide with 14 nm center to center spacing. The NPs perturb the electric field in the surrounding medium, diminishing the strength in the regions facing the gate or substrate and enhancing it in regions between the pores. Fig. 3-2 (b) shows the vertical electric field on cut lines at various positions, as displayed in Fig. 3-2 (a). The electric field is significantly enhanced on cut lines passing through the NPs due to the contrast in their dielectric constant from the surrounding medium. However, the vacuum barrier prevents any significant electron tunneling through such paths. For regions between the NPs, the improvement in electric field for SiO_2 can be about 25 % compared to homogenous dielectrics (without NPs), as confirmed from Fig. 3-2 (b). Carrier injection efficiency is exponentially dependent on the electric field in the barrier [14]-[15]. This inhomogeneity therefore strongly favors electron tunneling and capture between the pores, and thereby helps pinning their capture location.

The effect of enhancement is proportional to the difference in the dielectric constants of the NP and the surrounding medium and is expected to be larger for high-*k* nanoporous dielectrics like silicon nitride or hafnium oxide. It should be noted that the carrier injection efficiency improves in regions between the NPs in spite of the reduction in total EOT for the NP–gate–stack. The position, diameter and the spacing between the NPs are crucial to the magnitude of improvement in electron tunneling efficiency. For SiO_2 tunnel barrier (ΔE_C ~ 3 eV) and program field of 6 MV/cm, the tunneling distance through the barrier is ~ 5 nm. Any field enhancement generated in the gate stack beyond this distance from the channel will have no effect on the injection efficiency. Therefore, reducing the distance of the NP matrix from the channel would boost program speeds. Secondly, the maximum improvement in electric field is achieved in the plane containing the centers of the NPs. Hence, smaller diameter of NPs (4–6 nm) would lead to larger enhancement over homogenous injection. Further, increasing NP density (reducing NP spacing at constant NP diameter) also increases the peak field at the expense of smaller available area for electron injection. These factors need careful scrutiny in optimization of memory performance.

3.4 Block Copolymer Self Assembly of Nanopores

Self-assembly process was used to generate thin films, with controlled and uniformly distributed NPs across the substrate area, for device fabrication. Monolayer thin films (thickness ~ 10–12 nm) were prepared using self-assembly of pluronic F127 block copolymer surfactants (polyethylene oxide $(PEO)_{101}$-*b*-polypropylene oxide

$(PPO)_{56}$-*b*-polyethylene oxide $(PEO)_{101}$) to structure direct sol-gel derived aluminosilicate nanoparticles (8:2, wt:wt, Si:Al) [9], [16]-[17]. A dilute solution (0.7 wt%) of F127 polymer and pre-hydrolyzed aluminosilicate nanoparticles (2-5 nm) in 1:6 weight ratio in tetrahydrofuran (THF) was prepared and then spun cast on silicon wafers (at 2000 rpm for 60 seconds) followed by heat treatment (130 ˚C under vacuum for 60 minutes) to remove residual solvents and cross-link the aluminosilicate nanoparticles. The volume fraction of polymer and aluminosilicate was controlled to target spherical phase with PPO block forming hexagonally packed spheres in a matrix of PEO and aluminosilicate nanoparticles. The samples were then coated with a capping layer of Al_2O_3 (5 nm) deposited using ALD process. Following this, the films were calcined under air (heat treated to burn off the organics) at 500 ˚C for 30 minutes to open up pores within the aluminosilicate matrix. Fig. 3-2 (c) shows the SEM image of the calcined film (without the capping layer) revealing ~8-10 nm diameter NPs in aluminosilicate matrix. The calculated density of NPs is estimated to be 2×10^{11} cm^{-2} which accounts for ~ 20 % porosity. Furthermore, 23 nm of ALD Al_2O_3 was deposited at 300 ºC to serve as the control dielectric. Devices were patterned with Al as the metal gate.

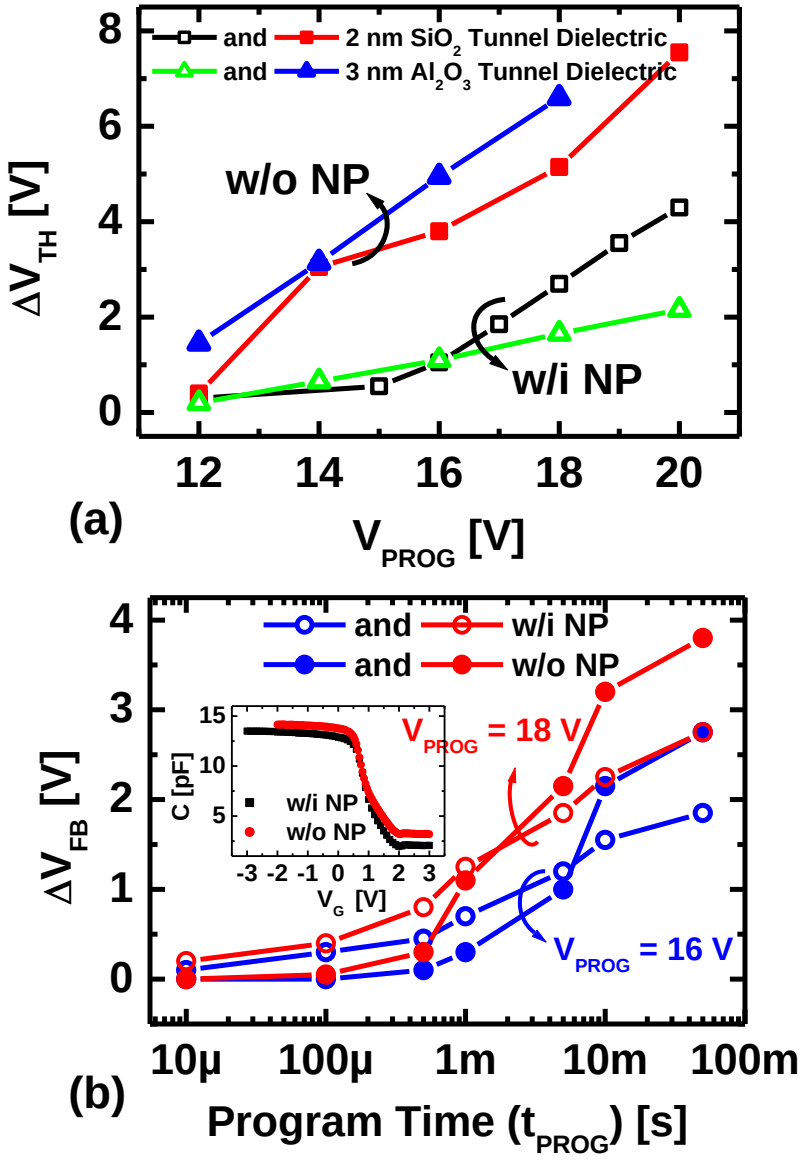


Figure 3-3 (a) ΔV_{TH} against V_{PROG} for memory structures in Fig. 3-1 (b) ΔV_{FB} plotted for sample (w/i NP) and control (w/o NP) device against program pulse time (t_{PROG}) for V_{PROG} = 16 V and 18 V (both devices shows the same initial V_{FB})

Figure 3-3 (a) ΔV_{TH} against V_{PROG} for memory structures in Fig. 3-1. All samples showed channel injection of electrons. NP samples consistently showed lower ΔV_{TH} against control samples due to smaller trapping volume in the charge storage layer. (b) ΔV_{FB} plotted for sample (w/i NP) and control (w/o NP) device against program pulse time (t_{PROG}) for V_{PROG} = 16 V and 18 V (both devices shows the same initial V_{FB}). NP–devices show higher injection efficiency than control devices (w/o NPs) for smaller ΔV_{FB}, indicating improved electrostatics for electron capture. However, larger trapping volume in control samples reverses the trend for higher ΔV_{FB}.

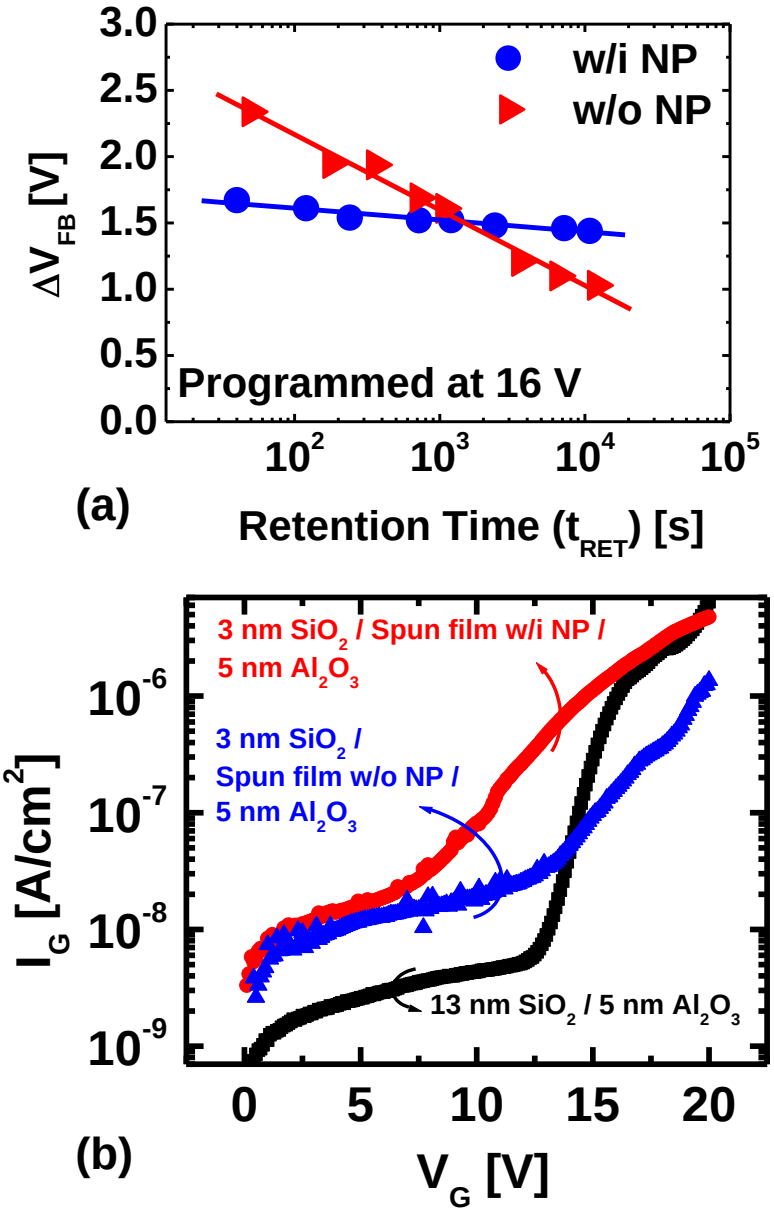


Figure 3-4 (a) Retention time measurements on memory structures with 2 nm SiO_2 tunnel dielectric. (b) Leakage current through spun cast films (both with and without NPs) compared to the ideal thermally grown SiO_2 with the same electrical thickness

3.5 Measurement Results

The low temperature processing of the aluminosilicate film naturally lead to numerous defect sites and charge capture locations in its bulk and therefore acted as an electron–trap–layer. Fig. 3-3 (a) depicts the ΔV_{TH} in all samples for various program voltages (V_{PROG}). Control devices (without NPs) exhibited larger ΔV_{TH} than samples with NPs for both types of tunnel dielectrics. ΔV_{TH} observed in devices with NPs at high V_{PROG} (> 15 V) is about half compared to the control devices. A larger trapping volume in control devices results in a higher density of stored electrons.

Pulsed program measurements were performed on both devices to estimate efficiency of electron injection during program operation. Fig. 3-3 (b) illustrates the flat band shift (ΔV_{FB}) against program time (t_{PROG}) for sample and control devices with SiO_2 tunnel dielectric. Both devices showed the same initial flat band voltage (V_{FB}). Further, as expected, samples with NPs displayed higher effective oxide thickness (EOT) (as confirmed by the lower maximum capacitance value in the *CV* measurements from the figure inset) due to the effect of low–*k* NPs in the trapping layer. For both V_{PROG} = 16 V and 18 V, at the start of the program operation (low ΔV_{TH}), electron injection in NP samples was seen to be more efficient than the corresponding control samples. This confirms the improvement in tunneling efficiency for inhomogeneous NP trapping medium, which was presented in Fig. 3-2 (b).

Due to a smaller trapping volume, as ΔV_{FB} continues to increase, a higher majority of trap sites in the NP–device are filled with injected electrons. These trapped electrons impede subsequent capture of incoming electrons and limit the rise in ΔV_{FB}.

However, control devices show continued increase in ΔV_{FB}, on account of higher trapping volume. In other words, at longer t_{PROG}, the control samples demonstrate higher ΔV_{FB} due to larger trapped charge density. For example, ΔV_{FB} in the NP–device is higher than the control device until 1 ms for V_{PROG} = 18 V, and 5 ms for V_{PROG} = 16 V. As ΔV_{FB} is dependent on the total trapped electrons, this crossover t_{PROG} is seen to diminish for higher V_{PROG}. Further, the observed ΔV_{FB} at crossover is consistent at ~ 1 V for both program voltages.

Fig. 3-4 (a) shows retention measurements for the same structures. The devices were programmed at 16 V for 100 ms. NP–devices show longer retention times over control devices, Further, the rate of ΔV_{TH} loss is about 5 times slower in NP–devices. Due to the inhomogeneous medium in NP–devices, each captured electron encounters a different effective barrier towards the channel. For example, electrons captured in the denser trapping medium need to overcome the SiO_2 barrier while those trapped above the NPs observe a significantly higher vacuum barrier and a correspondingly larger effective mass in the direction of the shortest escape path. The tunneling probability in retention condition is exponentially dependent on the barrier height as well as the effective mass of the particle in the barrier. Consequently, those electrons that have NPs in their tunneling path demonstrate a higher time constant of emission to the substrate. On the contrary, all electrons trapped in the homogeneous trap layer of the control device (without NPs) only perceive the SiO_2 tunnel barrier. Such inhomogeneous increase in the effective barrier height leads to longer retention time in NP–devices. Retention characteristics may be further enhanced by high temperature annealing to improve the quality of the spun cast films.

Transport through the aluminosilicate films was studied by making leakage measuring structures without the control barrier as detailed in Fig. 3-1. An ideal tunneling structure with thermally grown SiO_2 and having the same EOT was also prepared. At low fields (V_G < 10 V), there is no distinction between the conductance of both spun–cast films (with NPs and without NPs), as majority of the leakage results from trap–assisted tunneling However, at moderate fields (V_G > 10 V), films with NPs showed about an order of magnitude higher current densities compared to those without NPs, as displayed in Fig 3-4 (b). This direct measurement validates the enhancement in FN tunneling efficiency resulting from the field enhancement in the NP gate stack. However, both aluminosilicate films showed higher conductivity compared to thermal SiO_2 at moderate fields (V_G < 15 V). Fluence of electrons and the abundance of trap locations in both spun cast films may lead to significant dielectric charging. Such trapped charge in turn influences the Frenkel-Poole leakage through the film. The sudden transition (at V_G = 12 V) evident for thermal SiO_2 conductance as it transitions into the FN tunneling regime is smeared out due the trap–assisted leakage in these aluminosilicate films.

3.6 Conclusion

This chapter presents a study on the porosity in the gate stack and its influence on the charge-trap-based nonvolatile memory performance. Ordered nanoporosity was shown to modify the electrostatics crucial to program and retention characteristics. A simple and effective method of block copolymer self–assembly was integrated in the CMOS process flow to demonstrate the advantages of nanoporous dielectrics in

achieving higher programming efficiencies, smaller memory windows and longer retention times in charge–trap Flash devices. These findings may also be useful in understanding the outcomes of unintentional process–induced porosity in Flash memory gate stacks.

REFERENCES

[1] K. Prall, in Proc. NonVolatile Sem. Mem. Workshop, 2007, pp. 5-10.

[2] P. Olivo, T. Nguyen and B. Ricco, IEEE Tran. Elec. Dev. **35**, 2259 (1988).

[3] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé and K. Chan, Appl. Phys. Lett. **68**, 1377 (1996).

[4] Z. Liu, C. Lee, V. Narayanan, G. Pei and E. C. Kan, IEEE Tran. Elec. Dev. **49**, 1606 (2002).

[5] M. H. White, D. A. Adams and J. Bu, IEEE Circuits Devices Mag. **16**, 22 (2000).

[6] J. Choi, C. Kang, C. Lee, Y. Shin, B. Choi, J. Kim, S. Jeon, J. Sel, J. Park, K. Choi, T. Yoo, J. Sim and K. Kim, in Int. Elec. Dev. Meeting Tech. Dig., 2006, pp. 1–4.

[7] S. M. Amoroso, A. Maconi, A. Mauri, C. M. Campagnoni, A. S. Spinelli and A. L. Lacaita, IEEE Tran. Elec. Dev., **58**, 1864 (2011).

[8] Z. Liu, C. Lee, V. Narayanan, G. Pei and E. C. Kan, IEEE Elec. Dev. Lett., **24**, 345 (2003).

[9] P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka and G. D. Stucky, Nature, **396**, 152 (1998).

[10] B. A. Parviz, D. Ryan and G. M. Whitesides, IEEE Tran. Adv. Pkg., **26**, 233 (2003).

[11] K. W. Guarini, C. T. Black, Y. Zhang, I. V. Babich, E. M. Sikorski and L. M. Gignac, in Int. Elec. Dev. Meeting Tech. Dig., 2003, 22.2.1.

[12] A. J. Hong, C.-C. Liu, Y. Wang, J. Kim, F. Xiu, S. Ji, J. Zou, P. F. Nealey and K. L. Wang, Nano Lett. **10**, 224 (2010).

[13] S. R. Rajwade, H. Arora, J. Shaw, U. Weisner and E. C. Kan, in Proc. Dev. Res. Conf., 2009, pp. 235–236.

[14] M. Lenzlinger and E. H. Snow, J. Appl. Phys. **40**, 278 (1969).

[15] H. Bachhofer, H. Reisinger, E. Bertagnolli and H. von Philipsborn, J. Appl. Phys. **89**, 2791 (2001)

[16] M. Templin, A. Franck, A. D. Chesne, H. Leist, Y. Zhang, R. Ulrich, V. Schädler and U. Wiesner, Science, **278**, 1795 (1997).

[17] H. Arora, P. Du, K. W. Tan, J. K. Hyun, J. Grazul, H. L. Xin, D. A. Muller, M. O. Thompson and U. Wiesner, Science, **330**, 214 (2010).

CHAPTER 4 A FERROELECTRIC AND CHARGE HYBRID NONVOLATILE MEMORY—PART I: DEVICE CONCEPT AND MODELING

4.1 Abstract

This chapter a new one-transistor (1–T) hybrid nonvolatile memory based on the combination of two distinctive mechanisms, namely remanent polarization in ferroelectrics and charge injection into floating nodes. The gate stack design and the memory operation of the hybrid device are aimed to offer mutually complementing benefits between the two mechanisms, thereby presenting superior performance over conventional ferroelectric (FE) FET and gate injection-based Flash memory. During program operation, a high negative bias at the gate orients the ferroelectric polarization to the applied field. In addition, electrons at the gate electrode also tunnel into the floating nodes located between the ferroelectric thin film and the thin top tunnel dielectric and increase the total memory window. High electric displacement in the ferroelectric enables field enhancement in the tunnel dielectric for faster program and erase operations. During retention, the injected electrons reduce the depolarization field in the ferroelectric and the remanent polarization reduces the electric field in the tunnel oxide, which helps longer retention of the programmed state by the two additive memory mechanisms. This chapter evaluates the benefits of the hybrid gate stack through one dimensional (1−D) simulations incorporating the polarization-field (*P–E)* hysteresis in the ferroelectric layer. The simulations provide a guideline for

optimal gate stack design of the proposed hybrid memory. The following chapter then discusses the fabrication and experimental validation.

4.2 Introduction

While the conventional polysilicon Flash technology is fast approaching the sub-20 nm device nodes, overcoming future scaling challenges becomes increasingly difficult [1]. Stress from repeated program-erase (P/E) cycling makes the tunnel oxide susceptible to stress-induced trap generation [2], which causes undesirable charge leakage from the floating gate during retention. This is commonly known as stress-induced leakage currents (SILC) [3], which primarily restricts the scaling of tunnel oxide below 7–8 nm. The introduction of charge storage in discrete nodes like semiconductor [4]–[6] or metal nanocrystals (NCs) [7, 8], traps in dielectric like nitrides [9, 10] or even their heterogeneous integration [11] helps alleviate this drawback to a great extent. Nevertheless, charge injection in and out of the floating nodes requires extremely high fields (6–10 MV/cm for SiO_2) that results in high-voltage P/E operation and limits endurance cycling. In addition, these processes require longer P/E times ranging typically from 100 μs to milliseconds even at such high fields.

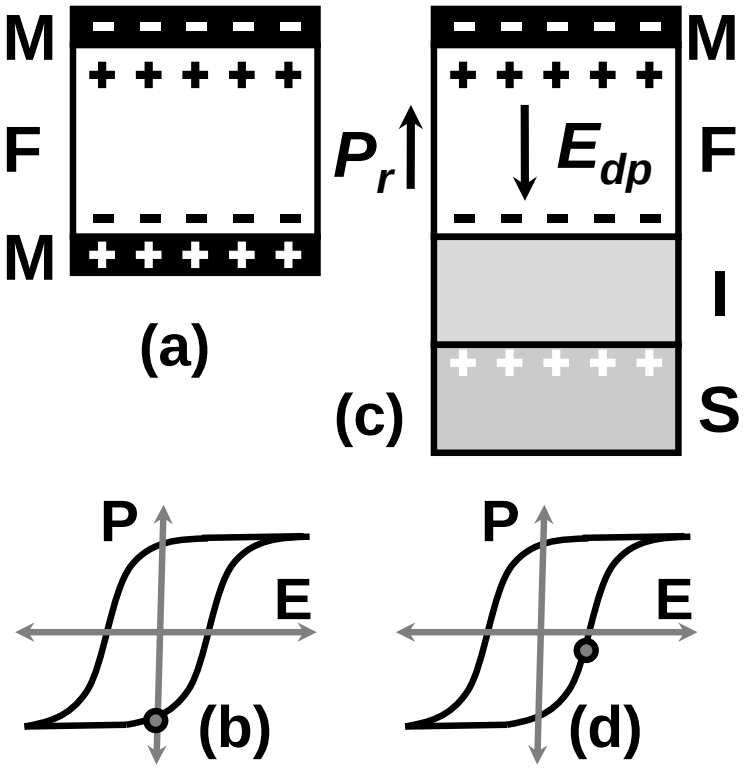


Figure 4-1 (a) Metal-ferroelectric-metal (MFM) capacitor with remanent polarization (P_r) in the ferroelectric film (b) Location of the total polarization on the *P–E* hysteresis in the ferroelectric during retention condition (c) Schematic of metal-ferroelectric-insulator-semiconductor (MFIS) design with P_r in the ferroelectric film. (d) The adjoining insulator prevents complete compensation of P_r in MFIS structure

In recent years, ferroelectric (FE) FETs have received much attention for extremely scaled memory applications [12]. FE–FETs have an advantage of non-destructive readout over ferroelectric random access memories (FRAM). Further, due to low coercive fields (E_c), FE–FETs demonstrate low–voltage operation even for reasonable memory windows [13, 14]. FE–FETs made from high-quality ferroelectric films like lead zirconium titanate (PZT) or $SrBi_2Ta_2O_9$ (SBT) exhibit fast switching (~

100 ns) at fields comparable to E_c [15, 16]. However, they are known to show poor retention characteristics on account of compensating charge tunneling at the ferroelectric-insulator interface and the depolarization field setup in the ferroelectric film during retention [17]. In addition, the large thickness of the ferroelectric film (> 500 nm) required to achieve acceptable memory window presents fabrication and scaling challenges due to very high device aspect ratios.

Fig. 4-1 explains the origin of this depolarization field. As seen in part (a), for a ferroelectric capacitor, the remanent polarization (P_r) is completely screened by the ideal metal electrodes at both ends. Consequently, the total polarization in the ferroelectric stabilizes very close to the zero electric field condition at retention state (Fig. 4-1(b)). But the presence of an insulator in the FE–FET gate stack (part (c)) prevents complete screening of this surface charge, which results in a field that opposes P_r and aids the dipole randomization process in the ferroelectric (part (d)). This depolarization field in the ferroelectric (E_{dp}) can be modeled by [17],

$$E_{dp} = P_r\left[k_{FE}\left(\frac{C_{ins}}{C_{FE}}+1\right)\right]^{-1} \qquad (1)$$

where k_{FE} is the dielectric constant of the ferroelectric, and C_{ins} and C_{FE} are the capacitance of the insulating film and the ferroelectric, respectively. Equation (1) implies that E_{dp} increases with the thickness of the adjoining insulator and may get comparable to E_c. It should be noted that incomplete compensation of polarization does not alter the shape of the polarization-electric field (*P–E)* hysteresis but changes the position of the total polarization state during retention condition, as shown in part

Fig. 4-1 (d) [18]. Such constant opposing field inside the ferroelectric tends to speed up the randomization in P_r and limit the retention time in FE–FETs [19].

In principle, C_{ins} represents the series combination of the insulator capacitance and the depletion or accumulation capacitance of the semiconducting substrate. For moderate substrate doping concentrations, semiconductor capacitance can be comparable to the capacitance of the insulating thin film. Depolarization field can be truly minimized by eliminating the insulator in the FE–FET gate stack as well as switching to oxide-based semiconducting substrates that offer extremely thin depletion thicknesses [20, 21]. However, these materials would make the sensing FET design and fabrication very difficult.

This chapter proposes a new hybrid memory transistor incorporating a charge storage layer adjoining the ferroelectric film that reduces the depolarization field by intentional charge injection from the gate electrode during the program condition [22]. The proposed gate stack configuration is similar to the metal-insulator-ferroelectric-insulator-semiconductor (MIFIS) structure presented previously [23, 24]. However, reduction of the leakage current through the gate stack and the associated fatigue was the primary motivation behind the MIFIS studies. Our work focuses on the influence of nonvolatile charge towards complementing the advantages offered by FE–FET memories. One dimensional (1–D) electrostatic simulations including the *P–E* hysteresis in the ferroelectric were performed to quantify the advantages of hybrid memory gate stack. These simulations also evaluate the effectiveness of hybrid design for varying ferroelectric material and geometry parameters.

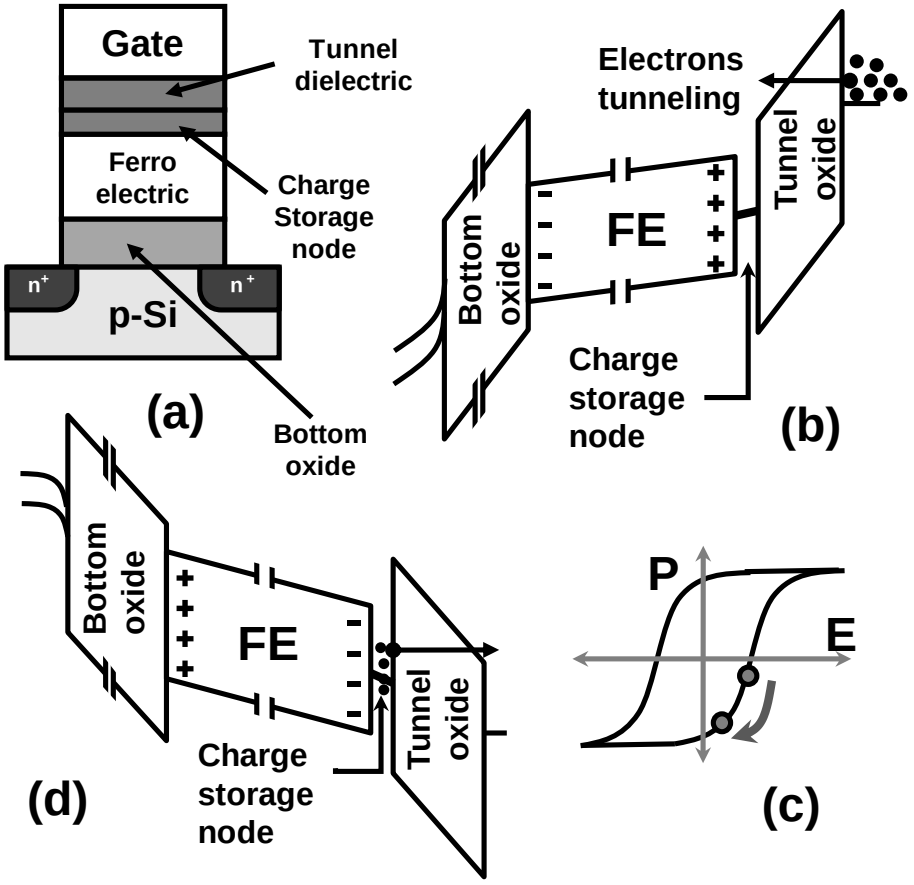


Figure 4-2 (a) Schematic of hybrid memory gate stack. (b) Band diagram in program condition. (c) Injected charge moves the location of stable polarization on the *P–E* hysteresis during retention. (d) Band diagram in erase condition

This chapter is organized as follows. Section 4.3 explains the device concept and working principles in detail. Section 4.4 outlines the electrostatic modeling as well as the choice of material and geometry parameters for the ferroelectric in improving the hybrid memory design. We then evaluate the expected electrostatic benefits for the fabricated hybrid devices. Section 4.5 presents the discussion on modeling and device

limitations. The next chapter will describe the fabrication and experimental validation of the enhanced memory characteristics.

4.3 Device Concept and Operation

Fig. 4-2 (a) shows the schematic of the proposed hybrid memory gate stack. The thick bottom dielectric serves the purpose similar to a control dielectric in the conventional Flash memory cell. It primarily blocks tunneling current from the channel to the ferroelectric interface and provides maximal gate coupling. The charge storage nodes (traps/NCs) are located directly above the ferroelectric layer. A thin top oxide (~5 nm) separates the control gate and serves as the tunneling dielectric for gate-injection (GI) of electrons during P/E operation [25]. The hybrid memory cell is programmed contrary to the conventional channel-injection Flash memory, namely by applying a high negative bias to the gate that aligns the ferroelectric polarization with positive surface charge facing the control gate (Fig. 4-2 (b)). The alignment of dipoles to the applied voltage creates a significant electric field in the top tunnel oxide that leads to electron injection from the gate into the charge storage layer. The gate-injected electrons add to the memory window originating from ferroelectric polarization. During retention, the stored electrons partially screen the remanent positive surface charge and reduce the depolarization field in the ferroelectric. In other words, the stable point for total polarization on the *P–E* hysteresis moves closer to $-P_r$, as depicted in Fig. 4-2(c). The erase operation performed by applying a high positive voltage at the gate reverses the polarization as well as removes stored electrons from the discrete storage nodes (Fig. 4-2 (d)).

4.4 Hybrid Memory Gate Stack Design

4.4.1. Electrostatic Modeling

Hybrid memory design needs to consider the interplay of the two memory mechanisms. In order to present a clear and simple assessment of the benefits and limitations of hybrid design over conventional FE–FET and GI Flash, we performed 1–D electrostatic simulations. Depending on the P/E voltages and the gate stack geometry, the electric fields experienced by the ferroelectric may not be able to drive the polarization to complete saturation. Taking this into consideration, electric displacement in the ferroelectric is modeled as proposed in [26]. These 1–D quasistatic expressions in (2) – (5) are based on those derived in [27] but include the additional effect from unsaturated *P*–*E* hysteresis. $P^{+}(E_{FE}, E_m)$ and $P^{-}(E_{FE}, E_m)$ are the forward and reverse unsaturated hysteresis branches that represent the sum of dipole and electronic polarization in the ferroelectric under quasistatic conditions. E_{FE} is the electric field in the ferroelectric, P_s is the saturation polarization, E_m is the maximum field experienced by the ferroelectric during P/E operations, ε_{FE} is the linear dielectric susceptibility of the ferroelectric, D is the electric displacement and ε_0 is the permittivity of free space.

$$P^{+}(E_{FE}, E_m) = P_s \tanh\left(\frac{E_{FE} - E_c}{2\delta}\right) + \varepsilon_{FE}\varepsilon_0 E_{FE} + \frac{1}{2}\left\{P_s \tanh\left(\frac{E_m + E_c}{2\delta}\right) - P_s \tanh\left(\frac{E_m - E_c}{2\delta}\right)\right\} \quad (2)$$

$$P^{-}(E_{FE},E_m)=P_s\tanh\left(\frac{E_{FE}+E_c}{2\delta}\right)+\varepsilon_{FE}\varepsilon_0 E_{FE}$$
$$-\frac{1}{2}\left\{P_s\tanh\left(\frac{E_m+E_c}{2\delta}\right)-P_s\tanh\left(\frac{E_m-E_c}{2\delta}\right)\right\} \quad (3)$$

$$\delta=E_c\left\{\ln\left(\frac{1+\frac{P_r}{P_s}}{1-\frac{P_r}{P_s}}\right)\right\}^{-1} \quad (4)$$

$$D=\varepsilon_0 E_{FE}+P^{\pm}(E_{FE},E_m) \quad (5)$$

These expressions are solved self-consistently with the Poisson's equation, $\nabla\cdot D=\rho$, including the band bending inside the channel for known sheet charge density ρ and applied gate voltage V_G. For simplicity, we assume the ferroelectric switches much faster than time scales relevant to electron tunneling. It should be noted that these assumptions may not be accurate at very low program voltages when switching field in the ferroelectric is less than E_c. Electron charge is considered to be stored at the interface of the trap layer and tunnel oxide. Charge centroid movement during all memory operations is ignored.

4.4.2. Gate Stack Material and Geometry Considerations

The choices regarding the material and thickness of the bottom and tunnel dielectric are critical to hybrid memory design. The bottom dielectric should provide maximal coupling of the ferroelectric layer to the channel but prevent any charge injection from the substrate during all memory operations. This criterion is not

different from conventional FE–FET design. However, the choice of tunnel oxide thickness is governed by two conflicting considerations. Unlike channel-inject Flash memory design, which requires minimal acceptable thickness of the tunnel oxide, GI-based hybrid design partially benefits from thicker tunnel oxide. Under 1–D approximation, memory window (ΔV_{TH}) in the GI Flash device is proportional to the product of stored charge density (Q) and tunnel dielectric thickness (t_{tox}) and inversely proportional to the dielectric constant (ε_{tox}) (i.e. $\Delta V_{TH} \sim Q t_{tox} / \varepsilon_{tox}$). Therefore, thicker tunnel oxide provides larger ΔV_{TH}. Secondly, the presence of high spontaneous polarization of the ferroelectric forces the adjoining tunnel oxide into the Fowler-Nordheim (FN) tunneling regime during P/E operations. This makes the P/E efficiency relatively insensitive to t_{tox}. On the other hand, thick tunnel oxide increases the operating voltage in any design.

For the purpose of consistency with the experimental measurements presented in the next chapter, simulations were performed with the bottom and tunnel dielectrics comprising of SiO_2 with thicknesses of 10 nm and 5 nm, respectively. Charge storage nodes are considered to be traps at the interface of the 2.5 nm HfO_2 layer and tunnel oxide. The importance of composite high-κ materials in forming excellent tunnel and control dielectrics for Flash memories [8, 25, 28] as well as their efficacy in reducing operating voltages in FE–FETs [13, 14] is well established. Optimization of material and geometry parameters for bottom and tunnel dielectrics in the hybrid device is however not presented in this article. Instead, this study focuses on exploring the parameter space for the ferroelectric material and the injected charge interplay in

providing superior performance of the hybrid device over conventional FE–FET and GI Flash.

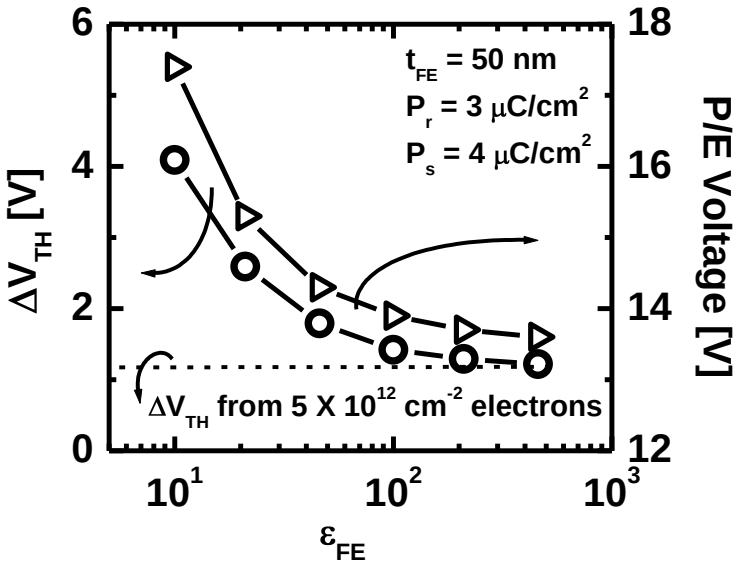


Figure 4-3 Total memory window (ΔV_{TH}) and P/E voltage against ε_{FE} in hybrid memory device. The thickness of the ferroelectric is fixed at t_{FE} = 50 nm. E_c is assumed to be inversely proportional to ε_{FE}. The electron sheet density of the storage layer is 5×10^{12} cm^{-2} that contributes 1.2 V to the total ΔV_{TH}

Fig. 4-3 shows the dependence of the total ΔV_{TH} and the P/E voltages in the hybrid device on ε_{FE}. The thickness of the ferroelectric film (t_{FE}) is fixed at 50 nm. ΔV_{TH} contribution from *P–E* hysteresis strongly depends on E_c and ε_{FE}. E_c is dependent on the choice of the ferroelectric material as well as t_{FE}. For example, oxide–based ferroelectric films have their ε_{FE} in the range of 100–200 and E_c in the range of 40–80 kV/cm [26, 29]. As ε_{FE} decreases, larger field is required to generate the electric displacement in the ferroelectric that initiates polarization reversal. Hence,

low ε_{FE} (~15) materials like polyvinylidene fluoride-trifluoroethylene [P(VDF–TrFE)] [30] tend to have high E_c. Further, in all ferroelectric materials E_c is sensitive to t_{FE} and generally scales proportional to $t_{FE}^{-2/3}$ [31].

We first attempt to depict the influence of ε_{FE} alone on ΔV_{TH} while still capturing the qualitative relationship between ε_{FE} and E_c. This is achieved by assuming E_c to be inversely proportional to ε_{FE}, with E_c = 500 kV/cm for ε_{FE} = 12 (typical for P(VDF–TrFE)). P_r is chosen to be 3 μC/cm^2 and the ratio P_r / P_s is fixed to 0.75 unless stated otherwise. Operating voltages are chosen to generate $E_{FE} = \pm 1.2E_c$ during quasistatic P/E conditions. The stored electron density of 5×10^{12} cm^{-2} contributes 1.2 V to the total ΔV_{TH}.

As seen from Fig. 4-3, both ΔV_{TH} and P/E voltages scale inversely with ε_{FE}. Contribution of ferroelectric polarization to the total ΔV_{TH} depends on the width of the *P–E* hysteresis (~ $2E_c$) as well as the saturation slope (proportional to ε_{FE}). For fixed P_r, this approximately scales as the product of E_c and t_{FE}. As ε_{FE} varies inversely with E_c in Fig. 4-3, this contribution to the total ΔV_{TH} becomes negligible above ε_{FE} = 100. Most complex oxides belong to this regime and therefore t_{FE} is required to be well over 200 nm to obtain reasonable memory window in conventional FE–FET. However, as discussed previously, ferroelectric materials like P(VDF–TrFE) or $Sr_2Nb_2O_7$ [32] depict sufficiently low dielectric constants which may reduce the device aspect ratio. For example, P(VDF–TrFE) has ε_{FE} in the range of 10–20 [33, 34]

and may provide total ΔV_{TH} over 4 V at ± 16 V P/E operations in the current hybrid design geometry.

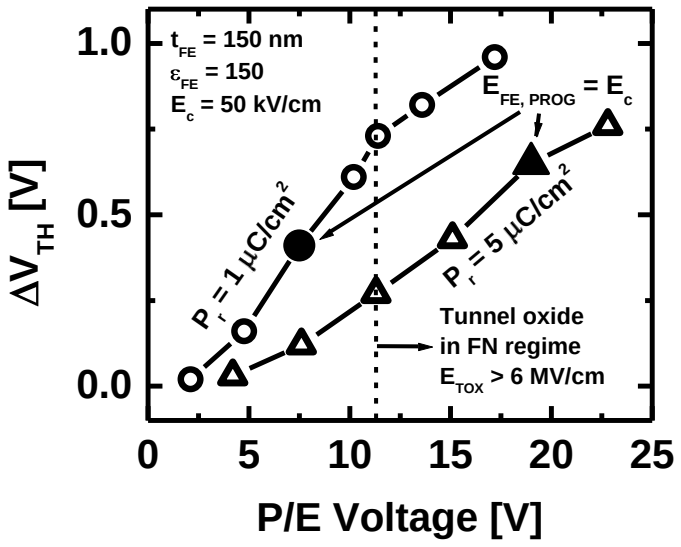


Figure 4-4 ΔV_{TH} due to *P–E* hysteresis plotted against the estimated P/E voltage for hybrid design. The ferroelectric material and geometry parameters considered for this analysis are t_{FE} = 150 nm, ε_{FE} = 150, and E_c = 50 kV/cm for two different values of P_r. No electrons are assumed to be stored in the charge storage layer

The height of the *P–E* hysteresis (~ $2P_r$) determines the P/E operating voltages. Fig. 4-4 examines the impact of P_r on the hybrid memory performance. For the remainder of this subsection, we fix t_{FE} = 150 nm, ε_{FE} = 150, and E_c = 50 kV/cm which are typical values to high–κ ferroelectrics [26], unless stated otherwise. The figure depicts the ΔV_{TH} contribution from *P–E* hysteresis as a function of P/E voltages for two different values of P_r = 1 μC/cm^2 and P_r = 5 μC/cm^2. Since the material and geometry parameters in the hybrid gate stack are fixed for this analysis, we can

estimate the gate voltage at which the tunnel oxide enters the FN regime and may therefore sustain appreciable gate injection thereafter. For the current 5 nm SiO_2 tunnel oxide, assuming 3 eV electron barrier from the gate, this condition is attained at V_G ~ 12 V. For ferroelectric with smaller P_r, the electric field surpasses E_c at a much lower P/E voltage (~7.5 V) compared to the electron injection voltage. On the other hand, E_{FE} barely attains $0.6E_c$ for high P_r (= 5 μC/cm^2) device at V_G ~ 12 V.

The choice of P_r is important to the design of hybrid devices. P/E voltages are chosen to cause reversal in ferroelectric polarization as well as generate high fields in the tunnel dielectric for gate injection. High values of P_r may result in severe charge injection from the gate (as well as the channel) even before V_G becomes sufficiently large to generate E_c in the ferroelectric material. This is undesirable as programming the ferroelectric below E_c may result in slow and incomplete domain reversal. The injected electrons would retard the reversal process by further decreasing E_{FE}. Lower P_r ensures that E_{FE} exceeds E_c at the chosen P/E voltages. In other words, ferroelectric domain reversal should saturate before appreciable electron injection into the floating nodes. However, extremely low P_r may also result in limited memory window and poor device performance.

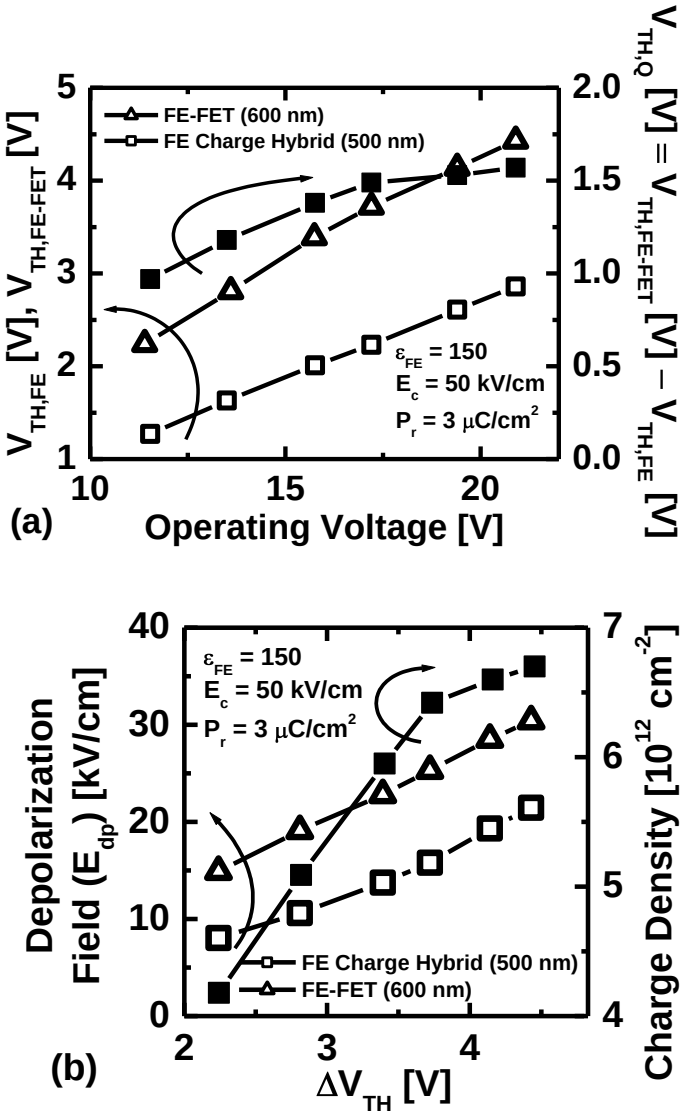


Figure 4-5 (a) Comparison of ΔV_{TH} from *P–E* hysteresis alone in FE-FET ($\Delta V_{TH,FE-FET}$) with t_{FE} = 600 nm and hybrid device ($\Delta V_{TH,FE}$) with t_{FE} = 500 nm respectively for operating voltage in the range of 10–25 V. The figure also shows the ΔV_{TH} deficit in hybrid device ($\Delta V_{TH,Q} = \Delta V_{TH,FE-FET} - \Delta V_{TH,FE}$) which is fulfilled by injected electrons during program operation. (b) Comparison of E_{dp} during retention for FE–FET and hybrid device operated at equal P/E voltages and correspondingly equal total ΔV_{TH}. The figure also shows the electron sheet density (Q) stored in the hybrid device required to match the ΔV_{TH} deficit arising purely from *P–E* hysteresis.

Suitable P_r for a particular hybrid gate stack therefore depends on the material and geometry of the adjoining layers. For the geometry under consideration, $P_r \sim 3$ $\mu C/cm^2$ would sustain $E_{FE} \sim E_c$ when the tunnel dielectric transitions to the FN tunneling regime. It should be noted that the above simulations are performed for quasistatic conditions and the actual electric fields in the ferroelectric immediately after applying P/E voltages may momentarily become much higher than the quasistatic value. However, the qualitative inference that low P_r ferroelectric material is beneficial for hybrid memory operation remains valid.

Memory window arising from *P–E* hysteresis alone for any given programming voltage is always smaller in the hybrid design compared to an FE–FET (with identical gate stack excluding the charge storage layer and tunnel oxide). However, correct design of the hybrid gate stack in principle may offer to compensate this deficit in ΔV_{TH} by addition of electrons to the storage nodes during program operation. Fig. 4-5 presents one such optimized hybrid design that yields significantly lower depolarization field over a conventional FE–FET for identical P/E voltages and ΔV_{TH}. To begin with, we consider a conventional FE–FET with t_{FE} = 600 nm. We choose to design the hybrid gate stack with t_{FE} = 500 nm and 5 nm tunnel SiO_2. Both devices have the same 10 nm SiO_2 bottom oxide. P/E voltage is varied between 10 V to 25 V in order to achieve ΔV_{TH} in the 2 – 5 V range for FE–FET device. At any operating voltage, due to the higher effective oxide thickness (EOT) and smaller t_{FE}, ΔV_{TH} for the hybrid device arising purely from *P–E* hysteresis ($\Delta V_{TH,FE}$) is smaller than the same for the FE–FET device ($\Delta V_{TH,FE-FET}$). However, electrons injected in the

floating nodes after program operation also contribute to the total ΔV_{TH} (charge contribution to memory window $\Delta V_{TH,Q} \sim Qt_{tox} / \varepsilon_{tox}$). This ΔV_{TH} deficit from *P–E* hysteresis is assumed to be compensated by the injected electrons (i.e. $\Delta V_{TH,Q} = \Delta V_{TH,FE-FET} - \Delta V_{TH,FE}$). Thus, the total ΔV_{TH} in the hybrid device is made to match with that in the FE–FET at every operating voltage. This is depicted in Fig. 4-5 (a), which plots $\Delta V_{TH,FE}$, $\Delta V_{TH,FE-FET}$ and the difference in the two ($\Delta V_{TH,Q}$) as a function of operating voltage for both devices. Both $\Delta V_{TH,FE}$ and $\Delta V_{TH,FE-FET}$ are seen to scale linearly in the 10–25 V operation range and their difference saturates at 1.6 V beyond 15 V.

The dependence of E_{dp} on total ΔV_{TH} (= $\Delta V_{TH,FE-FET}$ = $\Delta V_{TH,FE}$ + $\Delta V_{TH,Q}$) in both devices is displayed in Fig. 4-5 (b). As observed, hybrid device shows over 40 % reduction in E_{dp} for low ΔV_{TH} and about 30 % decrease at high ΔV_{TH}. The required electron sheet density to compensate for ΔV_{TH} deficit is also plotted against total ΔV_{TH}. The injected electrons observe a much lower field in the tunnel dielectric during retention due to the compensating positive surface charge in the ferroelectric.

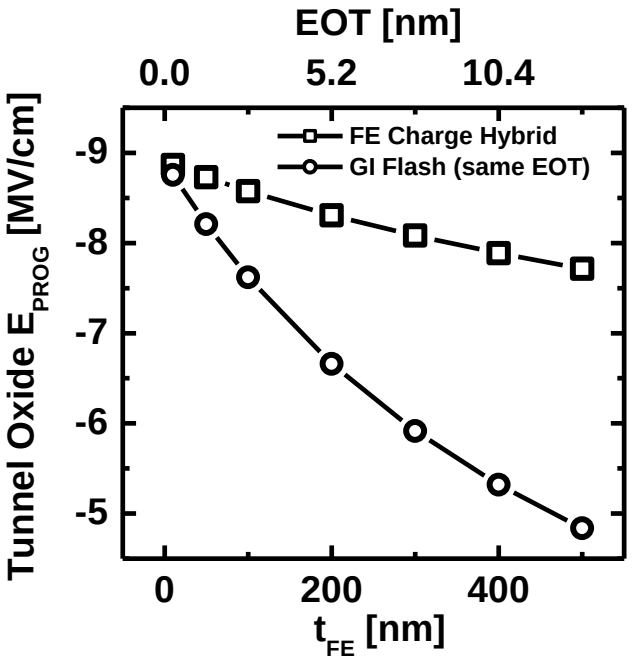


Figure 4-6 Comparison of the program field in tunnel oxide at –14 V for hybrid device and GI Flash with the same EOT as a function of t_{FE}

It should be noted that the hybrid device with the same t_{FE} (= 600 nm) as FE–FET would yield even higher reduction in E_{dp} at the expense of higher P/E voltages or larger contribution from $\Delta V_{TH,Q}$ (higher Q). The maximum electron density that can be stored at the floating node depends on the permissible electric field in the tunnel oxide that guarantees 10-year retention time. $\Delta V_{TH,FE}$ contribution to the total ΔV_{TH} can be reduced either by decreasing t_{FE} or by increasing the tunnel oxide thickness (thereby increasing $\Delta V_{TH,Q}$ for fixed Q). However, tunnel oxide thickness increase has a severe penalty on P/E voltage. Further, E_{dp} rises with reducing t_{FE}, that may have a negative impact on the device retention characteristics. In other words, biasing the device largely towards the GI Flash architecture reverses the proposed advantages of the

hybrid design. Conversely, the device can be biased towards the FE–FET by reducing tunnel oxide thickness or stored electron charge density. Tunnel oxide scaling is limited by the direct tunneling of electrons to the gate (~ 4–5 nm) during retention and charge density reduction significantly affects E_{dp}. All the above factors need careful consideration in optimizing hybrid design that can outperform a conventional FE–FET.

Fig. 4-6 compares the tunnel oxide electric field (E_{PROG}) in the hybrid device with the same ferroelectric film parameters against GI Flash for the program condition at –14 V. t_{FE} is varied from 0 nm to 500 nm. The ferroelectric layer is replaced by SiO_2 with the same EOT in the GI Flash device. For low t_{FE} in hybrid design (and corresponding EOT in GI Flash), E_{PROG} is equivalent for both devices. As t_{FE} increases, E_{PROG} diminishes rapidly in GI Flash due to the potential drop in the increased EOT. Electric displacement in the ferroelectric however arises mainly from the dipole polarization. As a result, the potential drop across the ferroelectric is less than that across a dielectric with the same EOT. This difference increases with t_{FE}, leading to significant field improvement in the tunnel oxide for hybrid devices. Field enhancement during P/E operations would lead to higher ΔV_{TH} and faster speeds for the hybrid device. It should be noted that the GI Flash device may be designed with much thinner bottom oxide thickness to improve its performance. However, this discussion merely brings out the influence of nonlinear *P–E* relationship in gate stack engineering of FE–FET and hybrid devices.

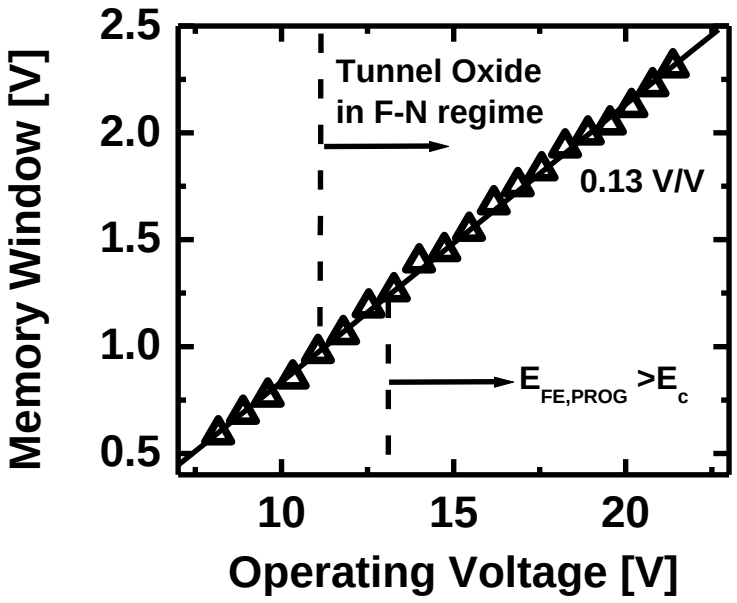


Figure 4-7 Dependence of memory window contribution from *P–E* hysteresis ($\Delta V_{TH,FE}$) on operating voltage in P(VDF–TrFE) hybrid memory. The electric field in the ferroelectric film crosses E_c for P/E voltages above 13 V. Tunnel oxide enters the FN tunneling regime above 11 V

The above analyses present all the important features in hybrid device design. Benefits arising from the *P–E* hysteresis improve as the ferroelectric layer EOT becomes comparable to the total EOT of the hybrid gate stack. Moderate value of P_r guarantees E_{FE} exceeds E_c before appreciable electron-injection from the gate. The introduction of tunnel oxide in the hybrid device necessitates higher P/E voltages. However, optimal gate stack design and contribution from charge storage can compensate for this penalty over conventional FE–FET as well as offer significant reduction in the depolarization field.

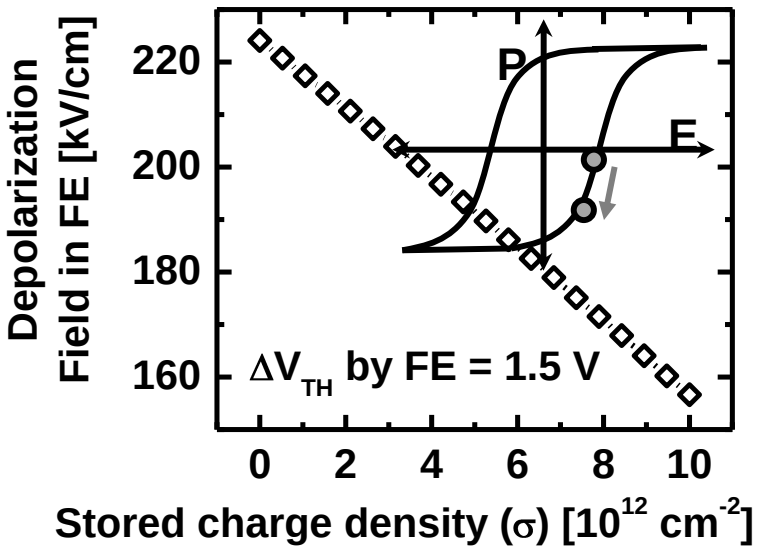


Figure 4-8 Reduction in the depolarization field in the P(VDF−TrFE) film as a function of injected electron sheet charge density for the hybrid memory. The simulated contribution of *P–E* hysteresis to the total ΔV_{TH} is 1.5 V. The inset schematic shows the movement of stable polarization location on the *P–E* hysteresis during retention.

Table 4-1 Gate Stack Composition of the Hybrid Memory and Control Devices

	FE Charge hybrid	FE–FET	GI Flash
Thermal SiO_2 bottom oxide	9.8 nm	17 nm	21 nm
P(VDF−TrFE)	35 nm	35 nm	
Evaporated SiO_2	1 nm		1 nm
Thermal ALD 110 ºC HfO_2 trap layer	2.5 nm		2.5 nm
Plasma ALD 110 ºC SiO_2 top tunnel oxide	5.4 nm		5.4 nm

4.4.3. Simulation Results for Fabricated Devices

Hybrid devices were fabricated with 35 nm P(VDF−TrFE) copolymer as the ferroelectric and 2.5 nm HfO_2 as the trapping layer, along with comparable FE–FET and GI Flash memory cells. Table 4-1 outlines the gate stack design and geometry parameters for all the memory devices. Details of the fabrication process and the experimental results are discussed in the following chapter. This section quantifies the expected benefits of hybrid design compared to conventional memory structures. P(VDF−TrFE) material parameters were extracted from experimental calibration of ferroelectric capacitors, also discussed in detail in next chapter. The values used in simulation are P_r = 3 μC/cm^2, P_s = 4 μC/cm^2, ε_{FE} = 12 and E_c = 500 kV/cm. It should

be noted that the hybrid memory without trapped charge is electrostatically equivalent to conventional FE–FET (with the same ferroelectric film thickness and equivalent EOT).

The contribution of charge to the total memory window ($\Delta V_{TH,Q}$) varies linearly from 0 V to 2.3 V when electron density in the storage layer varies from 0 to cm^{-2} in the hybrid device. Fig. 4-7 shows the memory window arising from *P–E* hysteresis alone ($\Delta V_{TH,FE}$) as a function of the P/E operating voltage for the device. The quasistatic P/E electric field experienced by the P(VDF–TrFE) film varies from 300–800 kV/cm. $\Delta V_{TH,FE}$ shows linear dependence of 0.13 V increase for every 1 V increase in the program voltage. The low dielectric constant of P(VDF−TrFE) enables reasonable memory window even at 35 nm thickness. The quasistatic program field in the ferroelectric film rises above E_c for P/E voltages above 13 V. The tunnel oxide enters the FN tunneling regime above 11 V. As discussed previously, hybrid design can truly benefit from the two complementary memory mechanisms when the P/E voltages and ferroelectric thin film P_r are designed to initiate significant charge tunneling as well as complete reversal of dipole polarization. For the remainder of the analysis, we assume the operating voltage of ±15 V for the hybrid device which generates $\Delta V_{TH,FE}$ = 1.5 V.

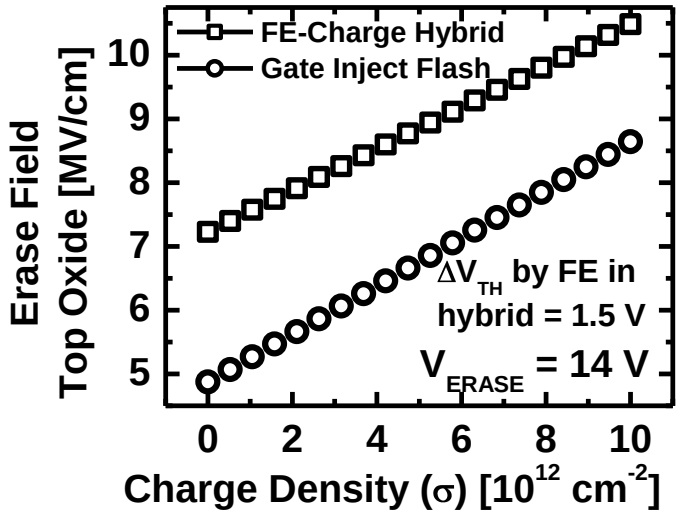


Figure 4-9 Enhancement in the erase field in the tunnel oxide for the hybrid device at V_{ERASE} = 14 V. The contribution of *P*–*E* hysteresis to the total ΔV_{TH} is set to 1.5 V.

Fig. 4-8 illustrates the reduction in E_{dp} for P(VDF−TrFE) against the injected electron sheet charge density. We can estimate significant reduction in the E_{dp} with 5×10^{12} cm^{-2} electron density. This trapped charge also adds 1.15 V to the memory window of the hybrid device. Fig. 4-9 depicts the electric field enhancement in the top tunnel oxide during erase operation at 14 V for hybrid memory over GI Flash. The hybrid memory shows a higher field offset by over 2 MV/cm in the tunnel oxide for all stored electron charge density. This enhancement is the consequence of switched dipole polarization in the ferroelectric film that adds to the electric displacement in the gate stack. Similarly, the hybrid memory shows better program fields for the same reason.

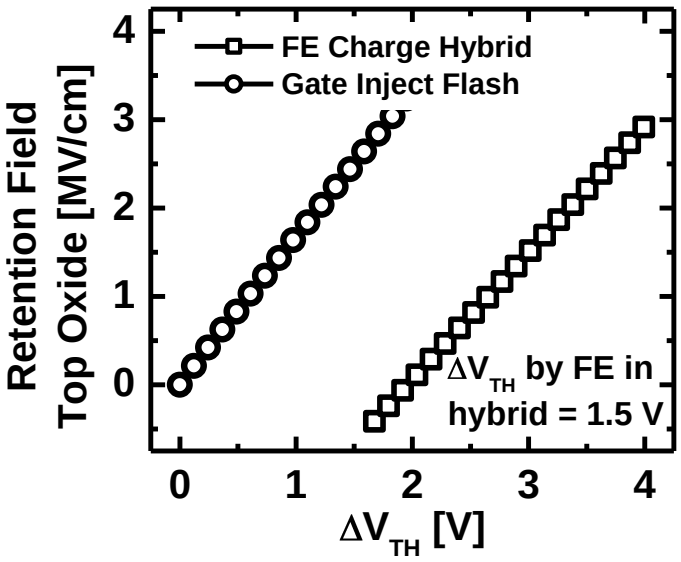


Figure 4-10 Electric field in the tunnel oxide during retention for hybrid memory and GI Flash as a function of total ΔV_{TH}

Any amount of stored charge is more stable at retention in the hybrid memory than in GI Flash due to the compensating positive surface charge in the ferroelectric. It is interesting to note that for a small amount of stored electron sheet charge, the tunnel oxide field during retention would actually repel the trapped electrons from escaping to the gate. The electron density near the zero-field condition (~ 0.5×10^{12} cm^{-2}) is however very low to cause any appreciable reduction in E_{dp} (as verified from Fig. 4-8) or any practical addition to the total ΔV_{TH}. In an ideal case, simultaneous zero electric field in the tunnel oxide as well as zero depolarization field in the ferroelectric is highly desired. However, the flat band condition in the tunnel oxide occurs much in advance compared to $E_{dp} = 0$ in any hybrid design due to the finite dielectric constant of the semiconducting substrate. Fig. 4-10 compares the retention field in the tunnel

oxide for hybrid and GI Flash as a function of total ΔV_{TH}. The hybrid device can sustain much higher charge density due to lower retention field in the tunnel oxide. With the current gate stack geometry, GI Flash may not provide $\Delta V_{TH} > 1$ V at realistic retention times.

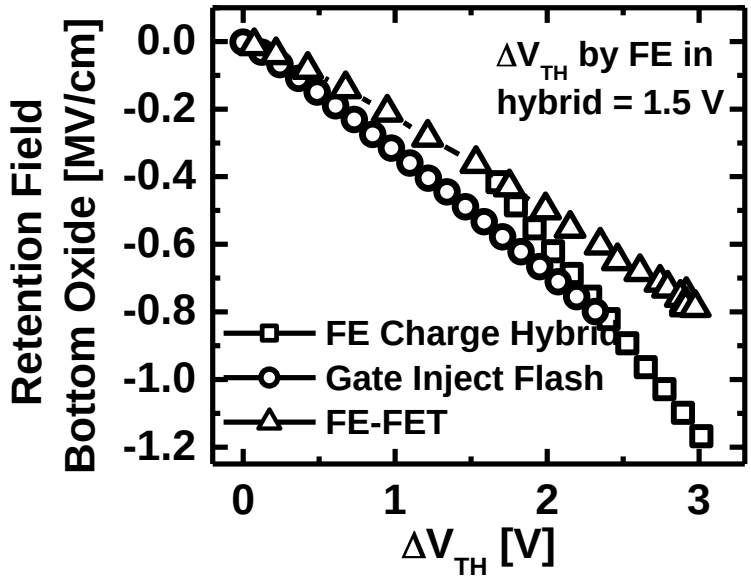


Figure 4-11 Electric field in the bottom oxide during retention for hybrid memory, FE–FET and GI Flash devices

The downside of reduced field in the tunnel oxide as well as the smaller depolarization field in the ferroelectric is the larger field in the bottom oxide at retention. Fig. 4-11 shows the comparison of field in the bottom oxide during retention in all three devices. The rate of increase in the retention field is higher in GI Flash than in FE–FET. The hybrid memory straddles in between the two devices until 5×10^{12} cm^{-2} of injected electron charge. By optimizing the contribution of injected electrons

and *P–E* hysteresis, hole tunneling through the bottom oxide during retention can be effectively inhibited.

4.5 Modeling and Design Limitations and Discussion

The 1–D electrostatic simulations presented above provide a preliminary analysis of the impact of gate stack material and geometry designs on P/E and retention characteristics of the proposed hybrid memory. Improvement in retention time over conventional FE–FET is surely one of the important figures–of–merit for the hybrid design. Direct tunneling of stored electrons through the tunnel oxide as well as the randomization of the aligned atomic dipoles has a bearing on the retention characteristics of the device. ΔV_{TH} loss due to charge leakage from the storage node depends on the charge density, tunnel oxide thickness and band structure of the tunneling medium and has been extensively studied for Flash memory devices. However, the effect of reduction in the depolarization field on the process of randomization in the aligned polarization requires detailed modeling through free energy calculations and the dipole-dipole interactions inside the polycrystalline ferroelectric thin film. Therefore, quantitative estimation of improvement in retention time resulting from this decrease in the depolarization field inside the ferroelectric has not been investigated in this work.

Further, these quasistatic simulations ignore the evolution of E_{FE} immediately after the onset of P/E operation. Accurate estimation of P/E speeds requires careful investigation of the kinetics of dipole rotation in time–varying electric fields. At the onset of program operation, the remanent polarization opposes the electric

displacement in the ferroelectric. As a result, E_{FE} increases significantly above the quasistatic value to maintain the requisite electric displacement in the gate stack. Enhancement in E_{FE} boosts the process of alignment of atomic dipoles to the applied field. The aligned dipoles in turn add to the electric displacement and reduce E_{FE}. This dynamic evolution of electrostatics in the gate stack coupled to switching models in ferroelectrics can help accurate estimation of P/E times. Additionally, the total P/E times would be strongly constrained by the relatively slow electron tunneling processes (~ 10 μs–1 ms) and the hybrid device may not be able to provide fast switching speeds inherent to FE–FETs.

As the device geometry scales down, 3–D fringing effects play a major role in P/E and read characteristics in scaled NAND–like arrays. Due to the large physical thickness of ferroelectric films and large ε_{FE}, these devices are prone to higher electrostatic interference from neighboring cells. Retention times in FE–FETs have been shown to have device area dependence [36] which points to careful scrutiny towards the use of wide area parameters in extremely scaled ferroelectric thin films. Accurate estimation of program and read disturb may further be obtained through 3–D electrostatic simulations that incorporate experimentally calibrated dielectric tensors for *P–E* hysteresis in ferroelectric thin films. All the above factors require detailed attention through comprehensive modeling to provide reasonable figures–of–merit for the hybrid device.

At last, reliability in hybrid devices is constrained by the combined effect of fatigue characteristics in the ferroelectric films as well as the integrity of the tunnel

oxide after repeated P/E cycling. Recent advances in thin film processing have enabled excellent fatigue endurance up to 10^{12} switching cycles in ferroelectric memories [37]. However, FN tunneling of electrons in and out of the floating nodes engages high energy transport across the thin tunnel oxide that may ultimately limit the reliability of such devices to 10^5 cycles. Such factors need thorough performance and reliability evaluation before their integration in scaled embedded nonvolatile memory applications.

4.6 Conclusion

This chapter presents a novel hybrid memory by integration of ferroelectric materials in charge–based Flash memory. The working principle is aimed to obtain complementary advantages from these distinct memory mechanisms. 1–D electrostatic simulations incorporating the *P–E* hysteresis were performed to evaluate the impact of hybrid gate stack geometry and material parameters on memory performance. The hybrid design is shown to offer maximal benefit at moderate values of remanent polarization and high ferroelectric film thicknesses. These devices exhibit substantial reduction in the depolarization field at much lower device aspect ratios. Contribution of the injected electrons and the *P–E* hysteresis to the total ΔV_{TH} may be tuned by altering the gate stack geometry to bring forth best performance enhancement over conventional FE–FET and GI Flash memories. Simulation results for hybrid devices and the corresponding control structures fabricated in the following chapter indicate significant field enhancement during P/E operations and field reduction during retention in the tunnel oxide compared to the GI Flash memory cell.

REFERENCES

[1] K. Prall, "Scaling non-volatile memory below 30 nm," in *Proc. NonVolatile Sem. Mem. Workshop*, 2007, pp. 5−10.

[2] P. Olivo, T. Nguyen and B. Ricco, "High-field-induced degradation in ultra-thin SiO2 films," *IEEE Trans. Electron Devices*, vol. 35, no. 12, pp. 2259−2267, Dec. 1988.

[3] R. Moazzami and C. Hu, "Stress-induced current in thin silicon dioxide films," *in IEDM Tech. Dig.*, 1992, pp. 139−142.

[4] K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai and K. Seki, "Room-temperature single-electron memory," *IEEE Trans. Electron Devices*, vol. 41, no. 9, pp. 1628−1638, Sept. 1994.

[5] S. Tiwari, F. Rana, K. Chan, H. Hanafi, W. Chan and D. Buchanan, "Volatile and non-volatile memories in silicon with nano-crystal storage," in *IEDM Tech. Dig.*, 1995, pp. 521−524.

[6] J. H. Chen, Y. Q. Wang, W. J. Yoo, Y.-C. Yeo, G. Samudra, D. SH. Chan, A. Y. Du and D.-L. Kwong, "Nonvolatile Flash memory device using Ge nanocrystals embedded in HfAlO high-κ tunneling and control oxides: Device fabrication and electrical performance," *IEEE Trans. Electron Devices*, vol. 51, no. 11, pp. 1840−1848, Nov. 2004.

[7] Z. Liu, C. Lee, V. Narayanan, G. Pei and E. C. Kan, "Metal nanocrystal memories, Part I: Device design and fabrication," *IEEE Trans. Electron Devices*, vol. 49, no. 9, pp. 1606−1613, Sept. 2002.

[8] J. J. Lee and D.-L. Kwong, "Metal nanocrystal memory with high-κ tunneling barrier for improved data retention," *IEEE Trans. Electron Devices*, vol. 52, no. 4, pp. 507−511, Apr. 2005.

[9] M. H. White, D. A. Adams and J. Bu, "On the go with SONOS," *IEEE Circuits Devices Mag.* pp. 22−31, July 2000.

[10] C. T. Swift, G. L. Chindalore, K. Harber, T. S. Harp, A. Hoefler, C. M. Hong, P. A. Ingersoll, C. B. Li, E. J. Prinz and J. A. Yater, "An embedded 90 nm SONOS nonvolatile memory utilizing hot electron programming and uniform tunnel erase," in *IEDM Tech. Dig*., 2002, pp. 927−930.

[11] C. Lee, T.-H. Hou and E. C. Kan, "Nonvolatile memory with a metal nanocrystal/nitride heterogeneous floating-gate," *IEEE Trans. Electron Devices*, vol. 52, no. 12, pp. 2697−2702, Dec. 2005.

[12] S. Wang, M. Takahashi, Q. –H. Li, K. Takeuchi and S. Sakai, "Operational method of a ferroelectric (Fe)-NAND flash memory array," *Semicond. Sci. Technol*. Vol. 24, p.105029, Sep. 2009.

[13] M. H. Tang, Z. H. Sun, Y. C. Zhou, Y. Sugiyama and H. Ishiwara, "Capacitance-voltage and retention characteristics of $Pt/SrBi_2Ta_2O_9/HfO_2/Si$ structures with various buffer layer thicknesses*," Appl. Phys. Lett.*, vol. 94, p. 212907, May 2009.

[14] M. Tang, X. Xu, Z. Ye, Y. Sugiyama and H. Ishiwara, "Impact of HfTaO buffer layer on data retention characteristics of ferroelectric-gate FET for nonvolatile memory applications," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 370−375, Feb. 2011.

[15] A. Chin, M. Y. Yang, C. L. Sun and S. Y. Chen, "Stack gate PZT/Al_2O_3 one transistor ferroelectric memory," *IEEE. Electron Device Lett.*, vol. 22, no. 7, pp. 336−338, July 2001.

[16] S. Sakai, M. Takahashi, K. Takeuchi, Q. –H. Li, T. Horiuchi, S. Wang, K. –Y. Yun, M. Takamiya and T. Sakurai, "Highly scalable Fe (ferroelectric)-NAND cell with MFIS (metal-ferroelectric-insulator-semiconductor) structure for sub-10 nm terabit capacity NAND Flash memories," in *International Memory Workshop*, 2008, pp.103−105.

[17] T. P. Ma and J. –P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," *IEEE Electron Device Lett.*, vol. 23, no. 7, pp.386−388, Jul. 2002.

[18] C. T. Black, C. Farrell and T. J. Licata, "Suppression of ferroelectric polarization by an adjustable depolarization field," *Appl. Phys. Lett.*, vol. 71, no. 14, pp. 2041-2043, Oct. 1997.

[19] R. R. Mehta, B. D. Silverman and J. T. Jacobs, "Depolarization fields in thin ferroelectric films," *J. Appl. Phys.*, vol. 44, no. 8, pp. 3379−3385, Aug. 1973.

[20] S. Mathews, R. Ramesh, T. Venkatesan and J. Benedetto, "Ferroelectric field effect transistor based on epitaxial perovskite heterostructures," *Science*, vol. 276, pp. 238–240, Apr. 1997.

[21] W. Choi, S. Kim, Y. W. Jin, S. Y. Lee and T. D. Sands, "Capacitive–voltage modeling of metal–ferroelectric–semiconductor capacitors based on epitaxial oxide heterostructures," *Appl. Phys. Lett.*, vol. 98, p. 102901, Mar. 2011.

[22] S. R. Rajwade, K. Auluck, J. T. Shaw, K. G. Lyon and E. C. Kan, "A hybrid

ferroelectric and charge nonvolatile memory," 69th *Device Research Conf. Digest*, 2011, pp. 169−170.

[23] E. Tokumitsu, Y. Takano, H. Shibata and H. Saiki, "Fabrication and characterization of M–I–FIS ferroelectric–gate structures using HfAlOx buffer layer," *Microe. Engg.*, vol. 84, pp. 2018–2021, 2007.

[24] M. Takahashi, H. Sugiyama, T. Nakaiso, K. Kodama, M. Noda and M. Okuyama, "Analysis and improvement of retention time of memorized state of metal–ferroelectric–insulator–semiconductor structure for ferroelectric gate FET memory," *Jpn. J. Appl. Phys.*, vol. 40, pp. 2923–2927, Jan. 2001.

[25] H.-T. Lue, E.-K. Lai, S.-Y. Wang, T. Yang, K.-C. Chen, K.-Y. Hsieh, R. Liu and C.-Y. Lu, "A novel gate-injection program/erase p-channel NAND-type Flash memory with high (10M cycle) endurance," in *Symp. VLSI Tech. Dig.*, 2007, pp. 140−141.

[26] H. –T. Lue, C. –J. Wu and T. –Y. Tseng, "Device modeling of ferroelectric memory field-effect transistor (FeMFET)," *IEEE Trans. Electron Devices*, vol. 49, no. 10, pp. 1790−1798, Oct. 2002.

[27] S. L. Miller, R. D. Nasby, J. R. Schwank, M. S. Rodgers and P. V. Dresendorfer, "Device modeling of ferroelectric capacitors," *J. Appl. Phys.*, vol. 68, no. 12, pp. 6463−6471, Dec. 1990.

[28] T.-H. Hou, C. Lee, V. Narayanan, U. Ganguly and E. C. Kan, "Design optimization of metal nanocrystal memory—Part II: gate-stack engineering," *IEEE Trans. Electron Devices*, vol. 53, no. 12, pp. 3103−3109, Dec. 2006.

[29] K. Ishikawa and H. Funakubo, "Electrical properties of (001)– and (116)–

oriented epitaxial SrBi2Ta2O9 thin films prepared by metalorganic chemical vapor deposition," *Appl. Phys. Lett.*, vol. 75, no. 13, pp. 1970–1972, Aug. 1999.

[30] S. Fujisaki, H. Ishiwara and Y. Fujisaki, "Low-voltage operation of ferroelectric poly(vinylidene fluoride-trifluoroethylene) copolymer capacitors and metal-ferroelectric-insulator-semiconductor diodes," *Appl. Phys. Lett.*, vol. 90, p. 162902, Apr. 2007.

[31] M. Dawber, P. Chandra, P. B. Littlewood and J. F. Scott, "Depolarization corrections to the coercive field in thin–film ferroelectrics," *J. Phys. Condens. Matter.*, vol. 15, pp. L393–L398, June 2003.

[32] Y. Fujimori, N. Izumi, T. Makamura, A. Kamisawa and Y. Shigematsu, "Development of low dielectric constant materials for the ferroelectric memory field effect transistor," *Jpn. J. Appl. Phys.*, vol. 36, no. 9B, pp. 5935−5938, Sept. 1997.

[33] T. Furukawa and G. E. Johnson, "Dielectric relaxations in a copolymer of vinylidene fluoride and trifluoroethylene," *J. Appl. Phys.*, vol. 52, no. 2, pp. 940–943, Feb. 1981.

[34] T. Yamada and T. Kitayama, "Ferroelectric properties of vinylidene fluoride-trifluoroethylene copolymers," *J. Appl. Phys.*, vol. 52, no. 11, pp. 6859–6863, Nov. 1981.

[35] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part II: experimental validation and analysis," submitted to *IEEE Trans. Electron Devices*.

[36] G. A. Salvatore, D. Bouvet, I. Stolitchnov, N. Setter and A. M. Ionescu, “Low voltage ferroelectric FET with sub-100 nm copolymer P(VDF−TrFE) gate dielectric for nonvolatile 1T memory,” in *Proc. European Solid-State Device Research Conf.,* 2008, pp. 162−165.

[37] S. Sakai and R. Ilangovan, “Metal-ferroelectric-insulator-semiconductor memory FET with long retention and high endurance,” *IEEE Electron Device Lett.*, vol. 25, no. 6, pp. 369–371, June 2004.

CHAPTER 5 A FERROELECTRIC AND CHARGE HYBRID NONVOLATILE MRMORY—PART II: EXPERIMENTAL VALIDATION AND ANALYSIS

5.1 Abstract

The previous chapter introduced the concept and operation of the novel hybrid memory, integrating ferroelectric polarization and nonvolatile charge injection. In this chapter, we demonstrate the experimental validation of this hybrid design. One transistor (1–T) memory cells were fabricated with polyvinylidene fluoride-trifluoroethylene [P(VDF−TrFE)] as the ferroelectric and HfO_2 as the charge trap layer. Hybrid devices showed larger memory window and longer retention time compared to conventional ferroelectric (FE) FETs with the same effective oxide thickness. Pulsed measurements were performed on metal-ferroelectric-metal (MFM) capacitors to estimate switching delay in the P(VDF−TrFE) thin film. Field enhancement in the tunnel oxide resulted in pronounced electron injection from the gate compared to gate-injection (GI) Flash memory cells. Hybrid devices also exhibited higher program efficiencies against FE–FET due to the contribution from these injected electrons. The presence of tunnel oxide in hybrid devices showed over 20 × reduction in gate leakage, which resulted in 100 × improvement in cycling endurance against FE–FETs.

5.2 Introduction

The rapid growth of mobile computing market has propelled the research in conventional and emerging nonvolatile memory devices that operate at high speeds and consume minimal power [1]–[3]. While the ubiquitous NAND Flash devices currently offer limited hope for low-voltage and high-speed operation, ferroelectric (FE) FETs [4] are considered as serious contenders for realizing low-power massively parallel memory architectures. However, these devices are known to suffer from poor data retention, limited program/erase (P/E) cycling endurance and huge device aspect ratios which limits their integration at scaled nodes [5]–[8]. In the last chapter, we have presented a novel hybrid nonvolatile memory design that draws advantages of low-power operation in FE–FETs as well as large memory window and long retention in conventional Flash devices. Electrons injected from the gate during program operation partially compensate the positive surface charge in the ferroelectric. The injected charge increases the memory window (ΔV_{TH}) as well as reduces the depolarization field (E_{dp}) during retention. Modeling results have shown that hybrid memory can demonstrate significant improvement over FE–FET at moderate values of remanent polarization (P_r). Based on the material and geometry selection guidelines established in the simulations, this chapter discusses the fabrication and experimental validation of the hybrid design.

In the recent years, polyvinylidene fluoride-trifluoroethylene [P(VDF−TrFE)] copolymer has generated considerable interest as a possible ferroelectric material in low-voltage FE–FETs and ferroelectric random access memories (FRAMs) [10]–[12].

Table 5-1 Gate Stack Composition of Hybrid Memory and Control Devices

	FE charge hybrid	FE–FET	GI Flash
Thermal SiO_2 bottom oxide	9.8 nm	17 nm	21 nm
P(VDF−TrFE)	35 nm	35 nm	
Evaporated SiO_2	1 nm		1 nm
Thermal ALD 110 ºC HfO_2 trap layer	2.5 nm		2.5 nm
Plasma ALD 110 ºC SiO_2 top tunnel oxide	5.4 nm		5.4 nm

By choosing the appropriate molar concentrations of VDF and TrFE, this polymer can be cast into the β phase [13] that demonstrates significant ferroelectric and piezoelectric properties. Low processing temperature and chemical stability enables its easy integration over various substrates. Owing to the low dielectric constant, P(VDF−TrFE) FE–FETs exhibit reasonable memory windows even at sub-100 nm film thicknesses [14, 15]. This chapter demonstrates the experimental validation of the proposed hybrid design by integration of sub-50 nm P(VDF−TrFE) film in the hybrid memory gate stack. Charge storage nodes were realized by a thin HfO_2 trap-layer deposited by atomic layer deposition (ALD) technique. It should be noted that in principle semiconductor or metal nanocrystals can also accomplish the function of

effective electron capture and retention in these devices [16, 17]. Our choice of the storage layer was purely based on their present ease of integration with low temperature processing. Measurement results highlight the importance of the complementary interplay between the two distinct mechanisms governing conventional FE–FET and gate-injection (GI) Flash [18]. The chapter also discusses the possibility of varying the contribution from polarization-field (*P*–*E*) hysteresis and stored charge to the total ΔV_{TH} by appropriate choice of P/E voltages to emphasize diverse benefits of hybrid operation.

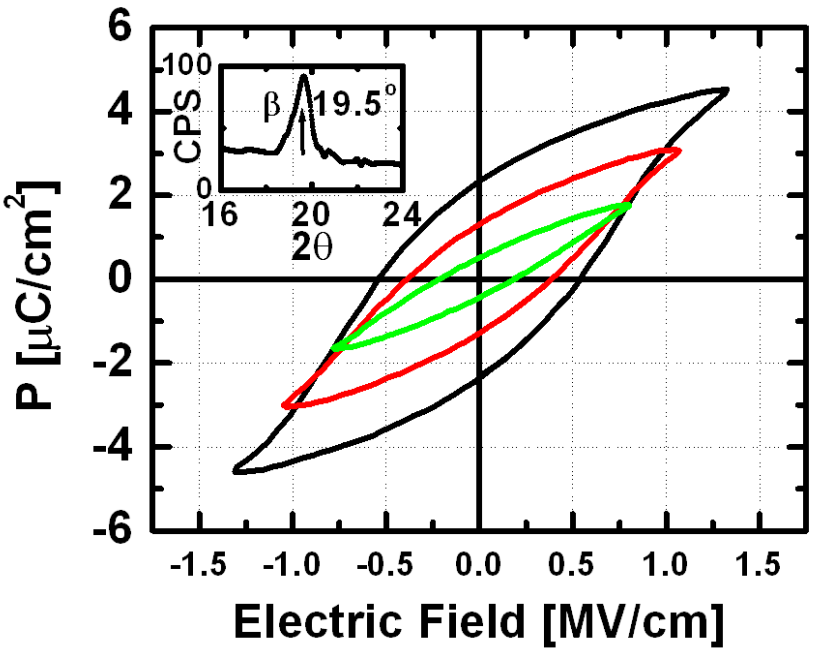


Figure 5-1 *P*–*E* hysteresis for 35 nm P(VDF–TrFE) metal-ferroelectric-metal (MFM) capacitors measured at 1 kHz and 3–5 V amplitude. The devices showed E_c of about 500 kV/cm. Inset shows θ–2θ XRD analysis that confirms the existence of ferroelectric β phase at 19.5 º.

5.3 Device Fabrication

FE–charge hybrid one transistor (1–T) memory devices and capacitors were fabricated with a modified gate-last process to accommodate lower processing temperatures after the deposition of organic ferroelectric film. P-type silicon (100) substrate with resistivity of 1-5 Ω-cm was first patterned to define active device region by the conventional shallow trench isolation (STI) procedure. The bottom dielectric was composed of 9.8 nm thermally grown SiO_2. This thickness ensures reasonable coupling of the remanent polarization to the channel and also blocks any direct tunneling current during retention. Source and drain regions were formed by phosphorus implantation and annealing followed by Ni/Al contact through the lift-off process. 0.5 % solution of P(VDF−TrFE) in 70:30 proportion was prepared in methyl-ethyl ketone (MEK) as described in [14, 15]. The filtered solution was spin–coated on the substrates at 2000 rpm. The film thickness measured by profilometer was 35 nm. In order to cast the film in the ferroelectric β phase, the substrates were annealed at 140 ºC for 10 min. A 1 nm SiO_2 layer was evaporated over P(VDF−TrFE) to serve as an incubation layer . The incubation layer is meant to facilitate the adsorption of reactant species for the following ALD process. ALD performed at 110 ºC was used to deposit the 2.5 nm HfO_2 trapping layer and the 5.4 nm SiO_2 top tunnel oxide. Lastly, gate metal deposition of Cr/Au/Al and wet etch followed by opening up source-drain contact regions concluded the fabrication. Control samples having the same effective oxide thickness (EOT) from the gate included conventional FE–FETs with the same P(VDF−TrFE) thickness but without the charge trapping layer, as well as the gate-injection (GI) Flash structures. Table 5-1 summarizes the gate stack composition for

all devices. Al / P(VDF−TrFE) / Pt–Cr metal-ferroelectric-metal (MFM) capacitors with 35 nm thick P(VDF−TrFE) were also fabricated to measure ferroelectric properties of the copolymer.

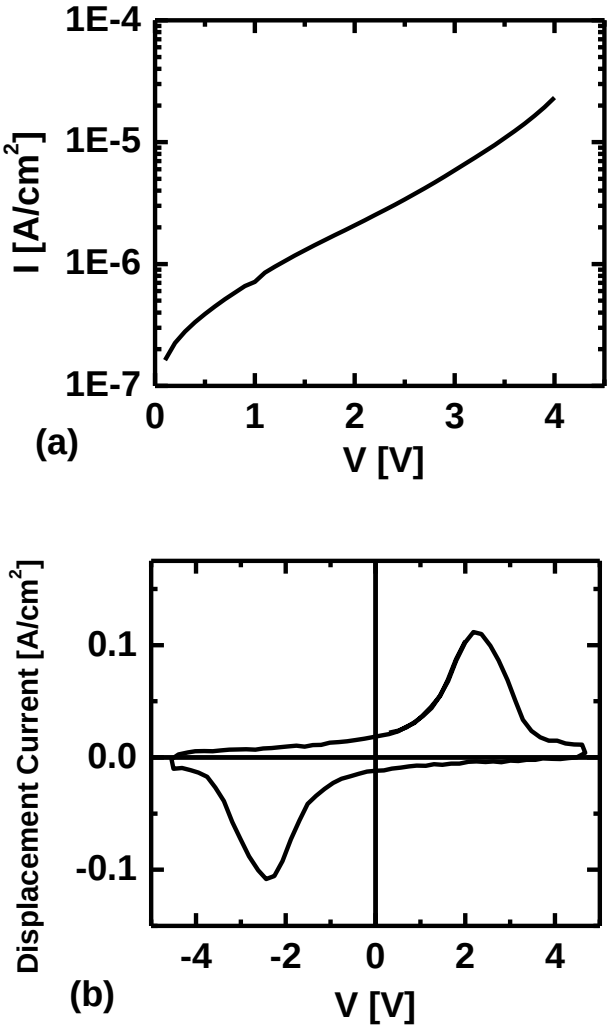


Figure 5-2 (a) Leakage current and (b) displacement current (measured at 1 kHz) through the 35 nm P(VDF–TrFE) MFM capacitor.

5.4 Results and Discussion

5.4.1. Ferroelectric Characterization

Measurement of polarization hysteresis for the 35 nm P(VDF−TrFE) MFM capacitors was performed by the Sawyer-Tower circuit arrangement [19]. Fig. 5-1 shows the *P–E* hysteresis extracted at 1 kHz. The applied peak amplitude was varied from 3 to 5 V. E_c was seen to be 500 kV/cm for poling voltage of 4 V. The remanent polarization (P_r), and the saturation polarization (P_s) were observed to be about 2.1 μC/cm^2 and 4.3 μC/cm^2 at 5 V. The low-frequency dielectric constant measured sufficiently above E_c was 13. Fig. 5-1 inset shows the XRD analysis of the P(VDF−TrFE) thin film and confirms the β phase formation [13].

Spin-cast P(VDF–TrFE) thick films (>200 nm) are known to demonstrate values of P_r in the range of 6 to 9 μC/cm^2 [20, 21]. However, as the thickness shrinks below 100 nm, ferroelectric response in such films is observed to reduce significantly [21, 22]. One of the reasons for such diminished performance at low thicknesses is the drastic loss in crystallinity. There exists a threshold thickness below which the crystallization process in severely affected. This critical thickness is in the range of 70–80 nm. In 35 nm spin cast films, total crytallinity can be even lower than 30 % [23]. The ferroelectric β phase forms plate-like crystallites (called lamellae) which are arranged perpendicular to the growth substrate. This lamellar size is observed to reduce from 100 nm in thick films to about 10–30 nm in thin films (<100 nm) which hinders complete crystallization upon annealing [20].Further, there is also a possible

formation of an interface dielectric layer which effectively occurs in series with the ferroelectric thin film and decreases the observed remanent polarization [20, 24].

Fig. 5-2 (a) presents the leakage current through the P(VDF–TrFE) film. The observed current density is above 10^{-5} A/cm^{-2} at 4 V. Fig. 5-2 (b) shows the displacement current measured by applying a 1 kHz saw tooth waveform on the MFM capacitor. The peak switching current arising from the polarization reversal is seen to be 0.1 A/cm^{-2}. Leakage through the spin-cast film is the result of defect-assisted transport through the semi-crystalline structure but is still 4 orders of magnitude lower than the switching current.

It is important to study the polarization switching response to pulsed electric fields in order to estimate program times for conventional FE–FETs or hybrid devices. The MFM capacitor is connected to a nanosecond pulse generator at one end and a resistance of suitable value to the other [25]. A long reset pulse is applied to the sample which saturates the remanent polarization in the direction of the applied field. Following this, double read pulses of opposite polarity and each of the same magnitude V and time t are applied to the film and the polarization reversal current is measured by monitoring the transient voltage across the resistor on the oscilloscope [26]. The integration of the total current over time normalized to the device area can be used to estimate the total switched polarization.

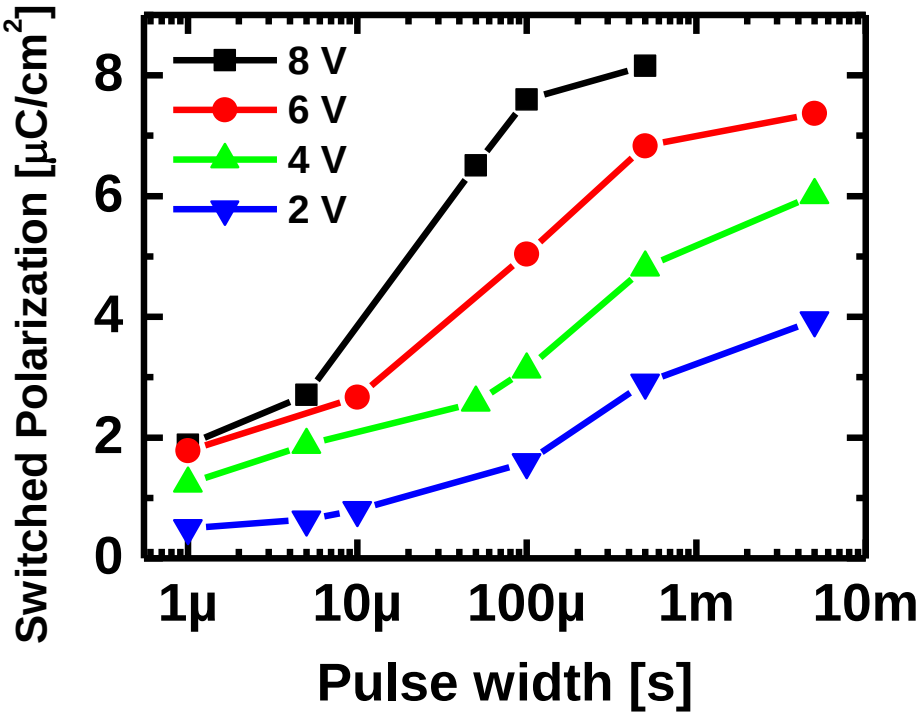


Figure 5-3 Total switched polarization measured for the 35 nm film for different pulse magnitudes and varying pulse widths

Measurements were performed with pulse durations ranging from 1 μs to 50 ms and amplitudes varied between 2 to 8 V. Fig. 5-3 shows the switching times for the P(VDF−TrFE) film for various pulse amplitudes. At 8 V (E_{FE} ~ 2.3 MV/cm) complete switching of polarization is seen for t = 1 ms. As the pulse magnitude decreases, switching time increases exponentially. The film only switches to half the maximum value for V = 2 V (E_{FE} ~ 0.57 MV/cm) and t = 50 ms. Switching time in thick P(VDF–TrFE) films (> 1 μm) exhibit a weaker dependence on the film thickness for a constant applied field. However, it has been observed that films below the critical thickness (60–80 nm) demonstrate thickness–dependent switching time for the same applied

field. In other words, as the film thickness reduces, switching time is shown to increase rapidly [20, 27] and may be at the scale of several seconds for applied fields close to E_c. The polarization reversal proceeds through two distinct mechanisms. The first step is called nucleation, where several new domains consisting of one or more dipole aligned to the applied field are formed at various sites in the film. The second step involves the growth of these nuclei by domain wall propagation through the adjacent dipoles [28, 29]. The domain growth in P(VDF–TrFE) becomes considerably slower with reduction in film thickness. The incomplete crystallization of the thin film as well as the smaller lamellar size impedes the domain wall growth. Furthermore, large defects in such films modulate the field in their vicinity which increases the local effective activation energy. As discussed earlier, the interface dielectric layer also exerts a strong restrictive force through the substrate which inhibits the reversal of molecular dipoles [20]. These effects overall manifested in extremely slow switching times in the 35 nm P(VDF–TrFE) films.

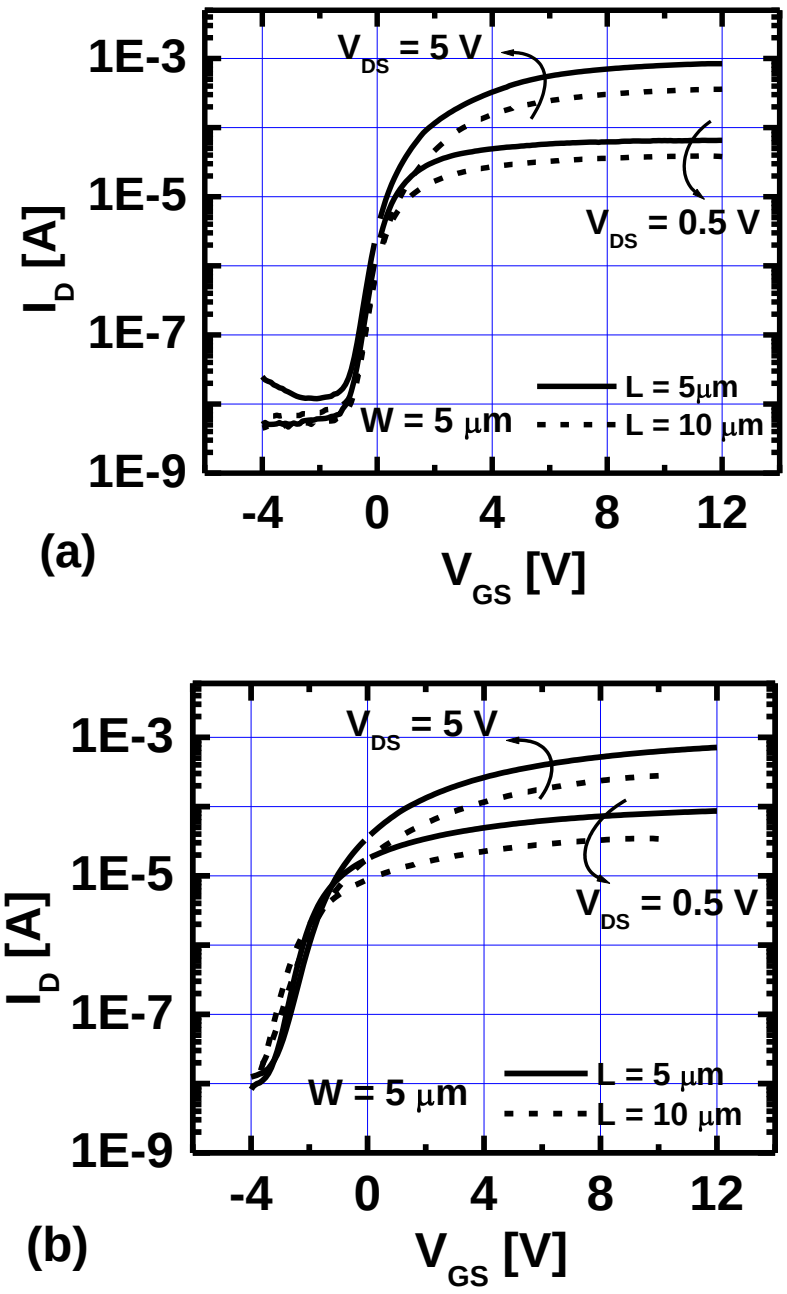


Figure 5-4 Transfer characteristics for (a) FE–FET and (b) hybrid devices at drain biases of 0.5 V and 5 V. The widths, *W* of all the devices were fixed at 5 μm

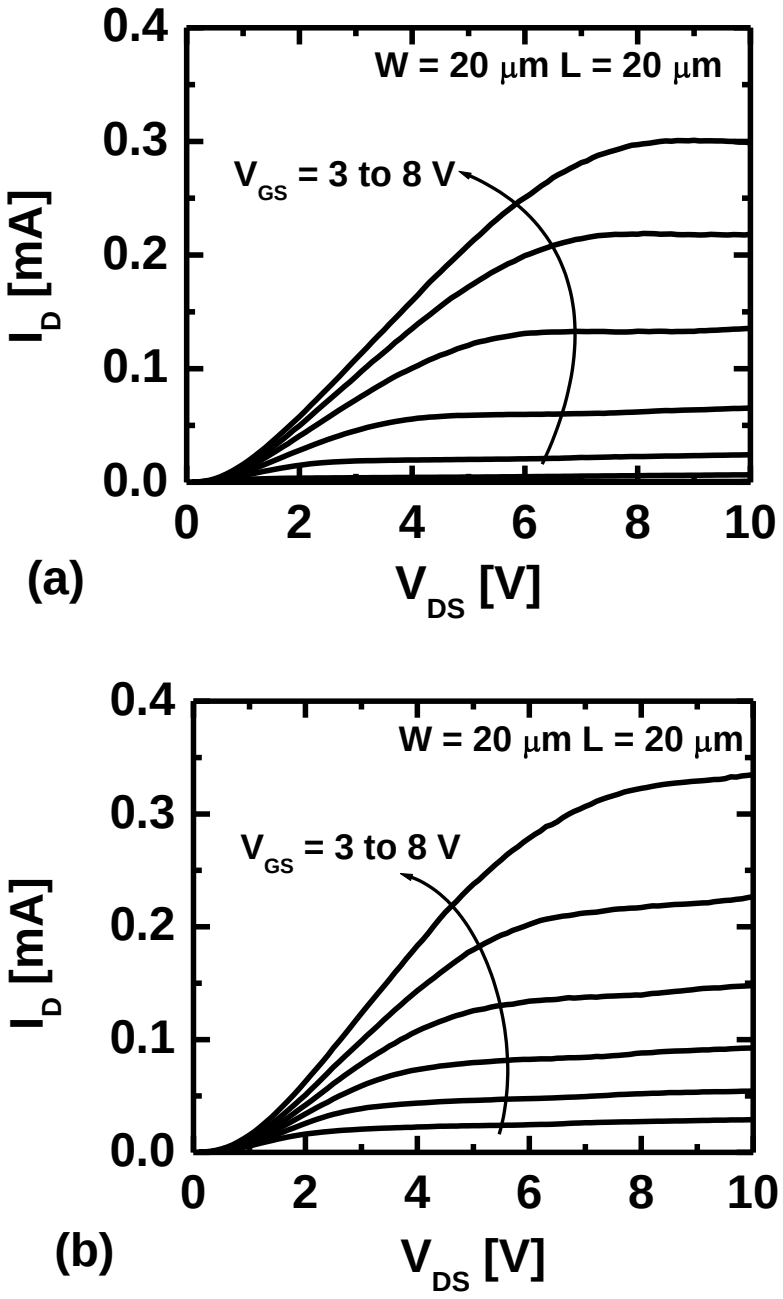


Figure 5-5 Output characteristics of (a) FE–FET and (b) hybrid memory for V_{GS} = 3 to 8 V for devices with $W = L$ = 20 μm.

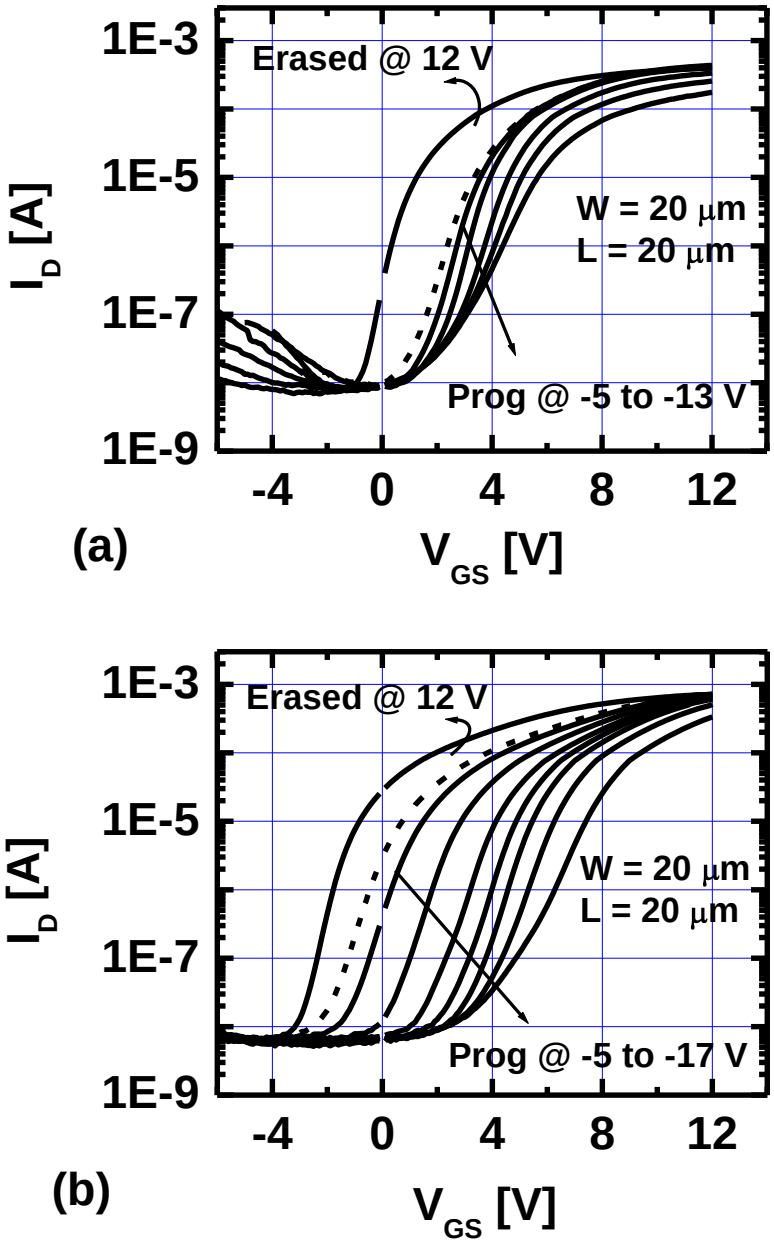


Figure 5-6 I_D–V_{GS} program characteristics for (a) FE–FET and (b) hybrid memory. Devices were programmed from −5 V to −17 V

5.4.2. Hybrid Memory Characterization

Fig. 5-4 presents the transfer characteristics (I_D–V_{GS}) in the erased state for the P(VDF−TrFE) FE–FET and the hybrid cell measured for two different lengths. Erase pulse at 12 V was applied for 5 seconds before sweeping V_{GS} to the device off state. Hybrid devices show current drives comparable to FE–FETs at both low and high drain biases. The ON current at V_{DS} = 5 V was observed to be 1 mA for devices with $W = L$ = 5 μm. Both devices showed an equivalent I_{ON}/I_{OFF} ratio of 10^5. Polarization in ferroelectrics is sensitive to the interfaces on either side of the thin film. The top interfaces for the two devices are different (evaporated SiO_2 for the hybrid device and Cr for FE–FET) and thus likely to generate distinct charge–trapping and screening effects. This may possibly result in 1 V difference in the observed threshold voltage for the two devices. Fig. 5-5 illustrates similar trends in FE–FET (Fig. 5-5 (a)) and hybrid (Fig. 5-5 (b)) device output characteristics measured at V_{GS} ranging from 3 to 8 V. The output current measured at $V_{DS} = V_{GS}$ = 8 V is about 0.3 mA for both devices with $W = L$ = 20 μm. These measurements confirm the same EOT for both FE–FET and hybrid memory gate stack designs.

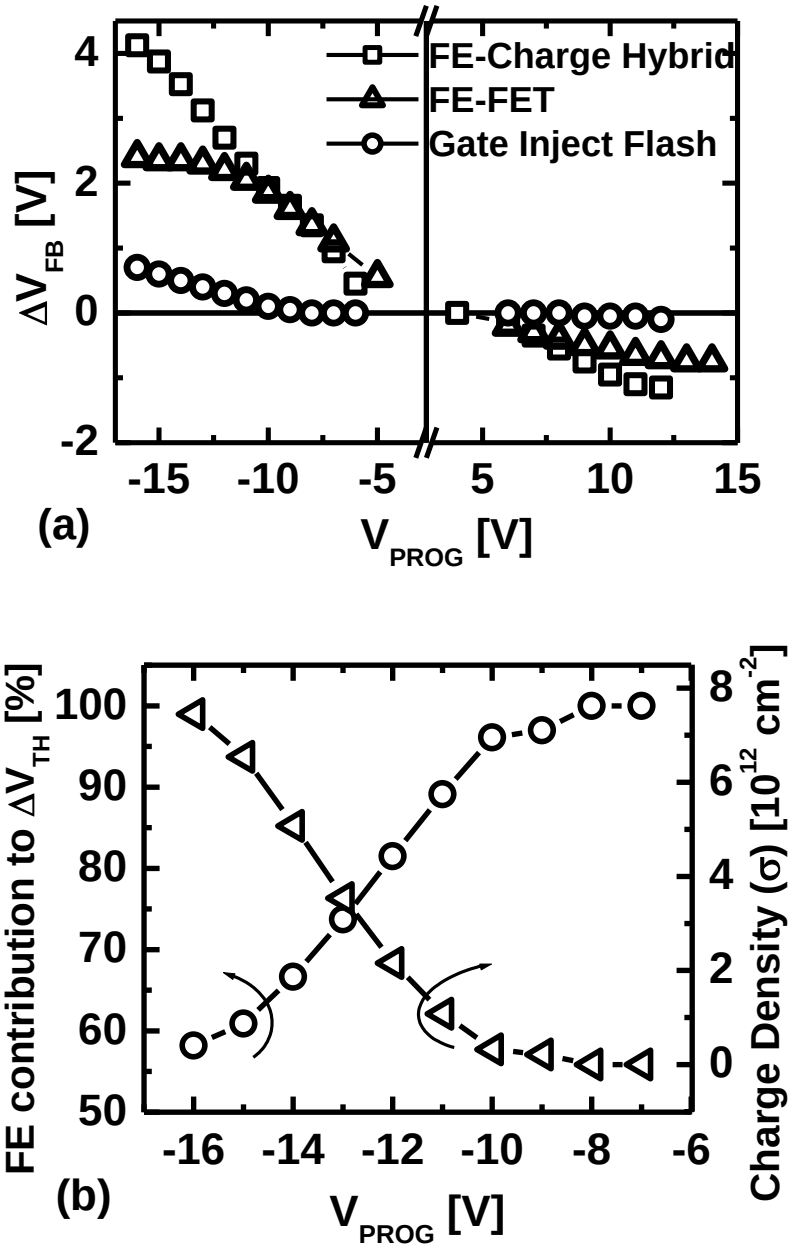


Figure 5-7 (a) ΔV_{FB} against V_{PROG} for FE–FET, hybrid memory and GI Flash devices (b) Estimated contribution of *P–E* hysteresis to the total ΔV_{FB} of hybrid device in part (a) and the electron sheet charge density as a function of V_{PROG}.

Figs. 5-6 (a) and (b) examine the memory window for the FE–FET and hybrid devices, respectively. Devices are initially erased at 12 V for 5 seconds followed by program operation at V_{PROG} ranging from -5 to -17 V (dashed lines show I_D–V_{GS} characteristics prior to erase operation). At lower program voltages (> –10 V), FE–FET devices show memory window (ΔV_{TH}) comparable to hybrid devices. However, for higher V_{PROG}, ΔV_{TH} is seen to saturate in FE–FET at ~ 2.5 V as seen from Fig. 5-6 (a). This value is about 1 V higher than the simulated ΔV_{TH} at the same operating voltage (discussed in the previous chapter). The difference in the observed and simulated values may arise from the contribution of unintended electron trapping in the ferroelectric during program operation, which has not been included in the simulation. Further, the current drive also starts degrading beyond V_{PROG} = –13 V as observed from the skewing of the I_D-V_{GS} characteristics.

Hybrid devices demonstrate continued increase in ΔV_{TH} (Fig. 5-6 (b)) and minimal skewing in transfer curves even at V_{PROG} = –15 V. Fig. 5-7 (a) summarizes the flat band shift (ΔV_{FB}) against V_{PROG} for FE–FET, hybrid as well as GI Flash devices. GI Flash devices show ΔV_{FB} of 1 V at V_{PROG} = –16 V against hybrid memory which demonstrates ΔV_{FB} over 4 V. Inadequate memory window in GI Flash is a result of insufficient program field in the tunnel oxide. Unlike dielectric materials, where electric displacement is proportional to the field, majority of the displacement in ferroelectric materials arises from the spontaneous polarization of aligned dipoles. As a result, the ferroelectric absorbs a smaller voltage drop as compared to a dielectric with the same EOT, enhancing the field in the adjoining dielectric layers. This

observation agrees well with the simulation results which predict about 2 MV/cm smaller electric field in the tunnel oxide of GI Flash at ± 14 V program/erase (P/E) operation.

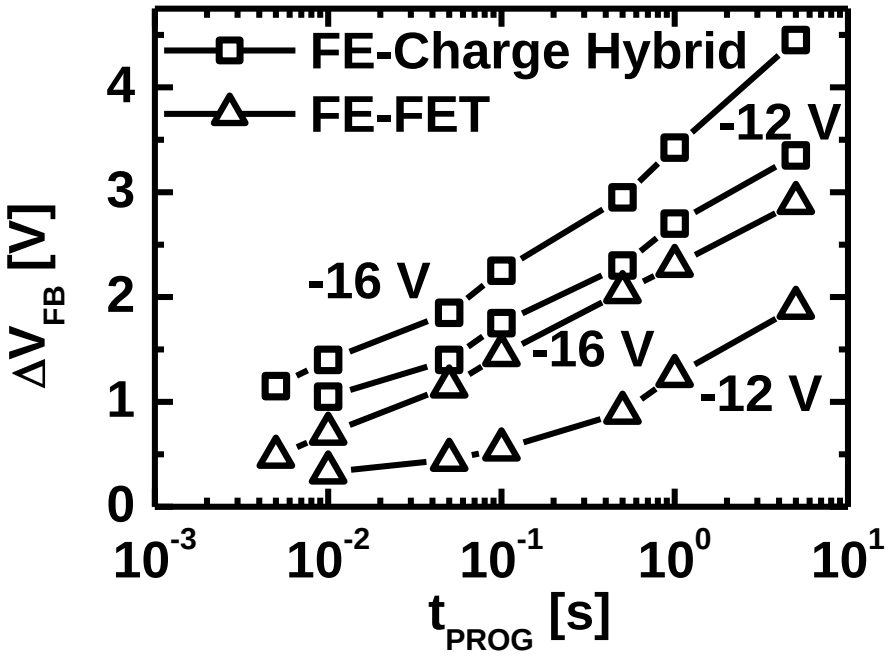


Figure 5-8 ΔV_{FB} against t_{PROG} for FE–FET and hybrid memory for V_{PROG} = −12 V and −16 V.

One of the concerns in the hybrid design as well as the FE–FET is the undesired tunneling of electrons from the channel during erase condition and a subsequent increase in the V_{FB} of the device. However, as seen from Fig. 5-7 (a), V_{FB} continues to decrease for erase voltages as high as 14 V, confirming negligible electron injection or trapping at the P(VDF−TrFE) and bottom oxide interface during erase. Similarly, hybrid devices with the present gate stack geometry do not exhibit

unintended channel hole injection for program operation. Fig. 5-7 (b) estimates the contribution of *P–E* hysteresis to the total ΔV_{FB} observed in Fig. 5-7 (a) for the hybrid device as well as the electron sheet charge density stored in HfO_2 trap layer as a function of V_{PROG}.

Fig. 5-8 depicts the ΔV_{FB} as a function of the program time (t_{PROG}) for FE–FET and hybrid memory devices for two different program voltages. Hybrid devices consistently showed higher programming efficiency over FE–FET for all V_{PROG}. ΔV_{FB} observed for 50 ms pulse width at V_{PROG} = –16 V is approximately 1.8 V for hybrid memory and 1 V for conventional FE–FET device. However, program saturation was no observed until 5 seconds in many devices, especially at lower V_{PROG} (> –15 V). The estimated quasi-static electric field in P(VDF−TrFE) during program condition is about 600 kV/cm at –16 V. As discussed earlier and seen from Fig. 5-3, complete alignment of ferroelectric domains under 1 ms requires electric field over 2 MV/cm. For pulse magnitude of 2 V (570 kV/cm), switched polarization only adds to about half the saturation value in 50 ms. This explains the long program times for these devices with the present illustrative choices of geometry and materials. Further, the time scales for ferroelectric switching and charge injection being similar, it is not possible to distinguish between the V_{FB} shift resulting from the two competing mechanisms. Integration of complex oxides such as PZT or $SrBi_2Ta_2O_9$ may however enable ferroelectric switching below 1 μs [8, 30] and thereby single out its contribution to the total ΔV_{FB} at small t_{PROG}.

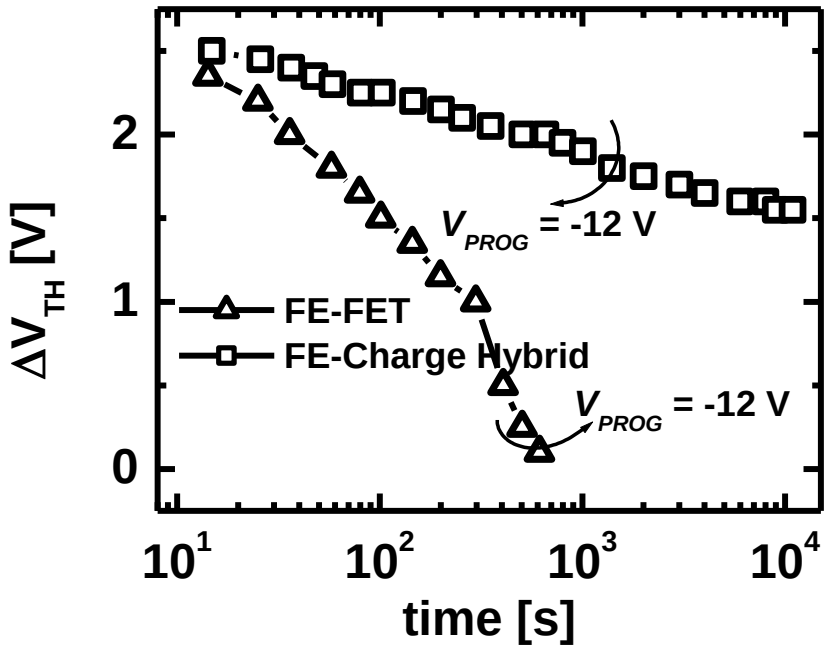


Figure 5-9 Retention characteristics for FE–FET and hybrid memory device after program operation at –12 V for 5 seconds

Retention measurements were performed by applying a suitable read gate voltage (V_{READ}) and measuring the drain current of the programmed memory cell. V_{READ} was chosen to provide maximal difference between the drain current for the programmed and erased state in both memory devices. Fig. 5-9 shows the retention measurements for the two devices. Both the devices were programmed at V_{PROG} = –12 V for 5 seconds. V_{READ} was suitably chosen to be 1 V. V_{TH} was extracted from the drain current monitored at V_{READ} for every time point. FE–FET device displayed poor retention characteristics as it lost more than 90 % of the ΔV_{TH} in less than 10^3 seconds. Hybrid device, on the other hand, demonstrated much better retention of the programmed state with less than 50 % loss after 10^4 seconds.

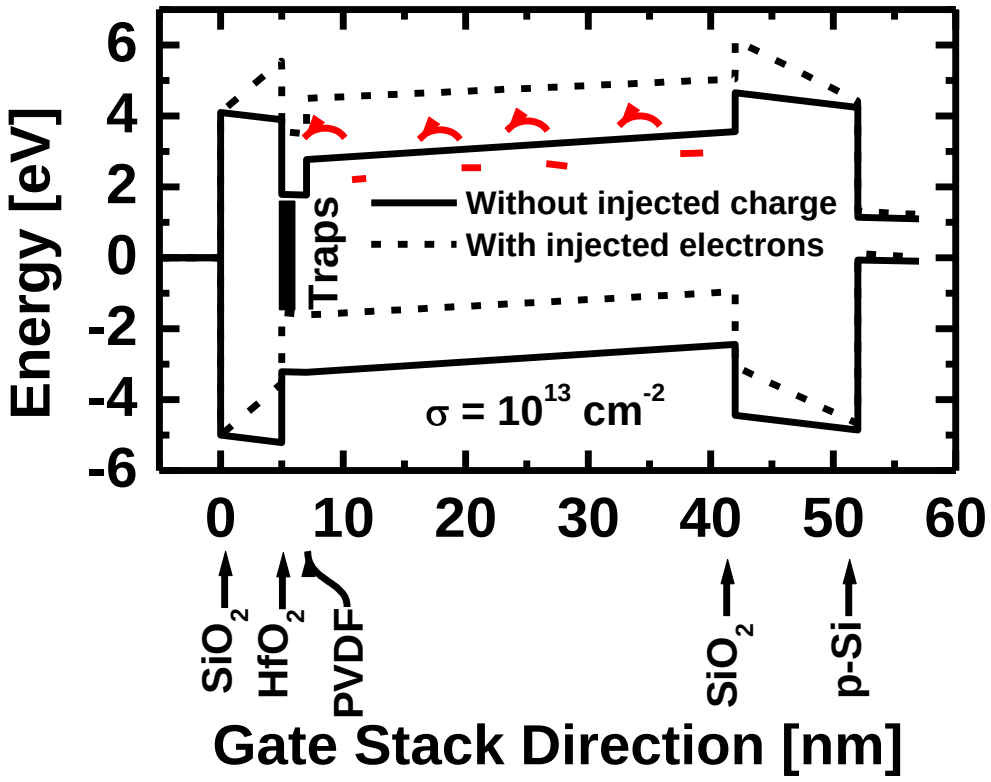


Figure 5-10 Band diagram of the hybrid gate stack during retention condition with and without injected electrons

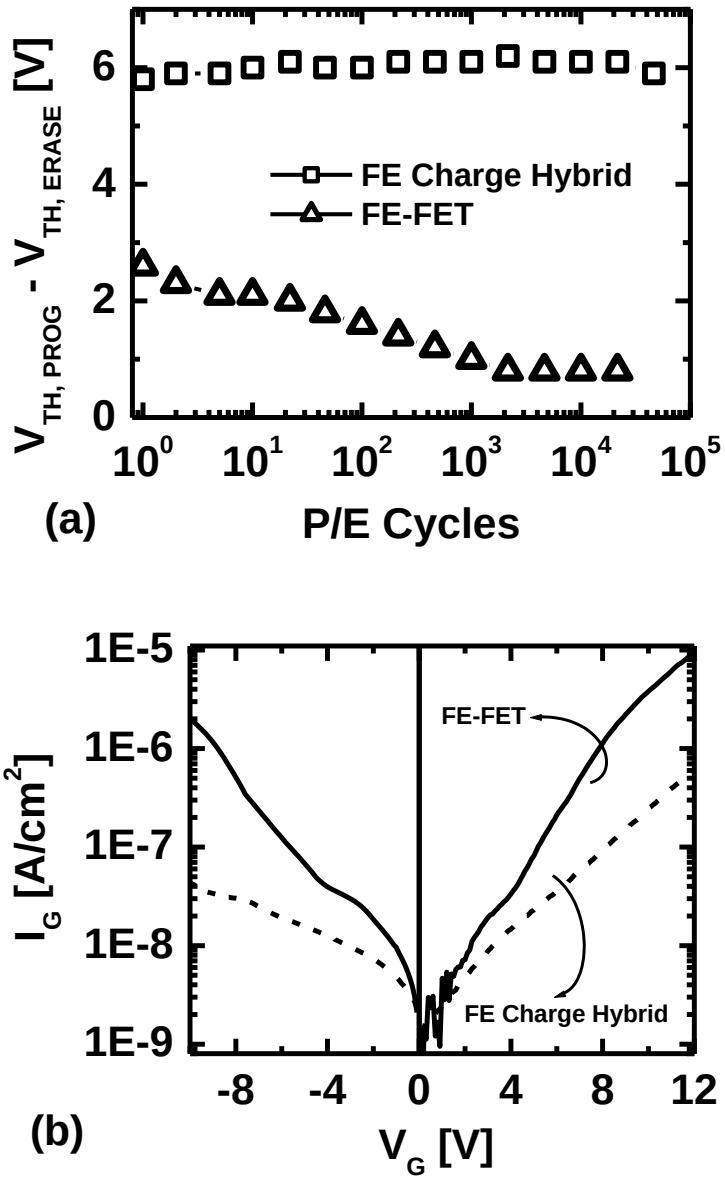


Figure 5-11 (a) Endurance of memory window against P/E cycling for FE–FET and hybrid device. The devices were programmed and erased at ±12 V with the pulse duration of 20 ms each (b) Gate current measurements for the devices show that FE–FET sustains 40 times larger current at –10 V and 20 times larger current at 12 V than the hybrid memory

Not all electrons injected during program condition may get captured in the thin trapping layer. Due to the polycrystalline nature of the adjoining P(VDF−TrFE), some of them may get captured in the bulk of the ferroelectric film. However, upon removal of program voltage, those trapped in the bulk experience the depolarization field developed in the ferroelectric. In case of FE–FETs, these electrons gradually thermalize or leak out as they are guided by the depolarization field to escape towards the gate. This leads to a substantial loss in memory window observed in FE–FET cells. The trap layer located between the ferroelectric and the tunnel oxide for hybrid memory acts as an ideal potential well for such emitted electrons. The depolarization field assists the gradual movement of these electrons towards the trap sites in the potential well. Those already captured in these trap sites increase the compensating negative charge at the ferroelectric surface and thus reduce the depolarization field. This helps longer retention of the programmed state in hybrid memory.

Fig. 5-10 compares the effect of electron storage on the band diagram in the hybrid device during retention condition. The contribution of *P–E* hysteresis to the total ΔV_{TH} is 1.5 V. The stored electron density of 10^{13} cm^{-2} is shown to reduce the depolarization field in P(VDF−TrFE) with no injected charge by over 25 % (225 kV/cm, without compensating charge and 160 kV/cm, with stored electrons). As discussed in the previous chapter, retention field in the tunnel oxide also changes polarity upon charge injection (–0.5 MV/cm, without compensating charge and 3 MV/cm with stored electrons). The arrows indicate the direction of gradual movement of injected electrons in the P(VDF–TrFE) film towards the trap layer.

It should be noted that the read voltage opposes the polarity of program voltage and results in finite disturbance of the programmed state during read operation. However, read disturb to the programmed state may be minimized by reducing the time of read operation. It can also be completely eliminated by immediate program operation or switching the present design to PMOS memory transistors.

The retention times observed even in hybrid devices with the current proof-of-concept choice of materials need major improvement for practicable implementation in nonvolatile memory systems. Absolute retention time is determined by the quality of the ferroelectric thin film as well as the tunnel oxide. For the current choice of ferroelectric material, we are restricted to the use of low temperature (< 140 ºC) processing for trap layer and tunnel dielectric deposition. However, low temperature thin film deposition techniques including ALD yield relatively porous films that may have significant trap densities. During retention condition, these traps may assist escape of captured electrons towards the gate. Significant improvement in the hybrid device may be attained by proceeding towards integration of oxide-based ferroelectrics that demonstrate enhanced performance at lower defect densities. The increased thermal budget (for e.g. ~ 600 ºC for PZT) would also permit the use of high temperature deposition processes like chemical vapor deposition (CVD) that are known to produce denser films at reduced trap densities. Thus, a combined effect of a superior ferroelectric with improved trap layer and tunnel dielectric can realize significantly longer retention time.

Reliability is a serious concern in ferroelectric-based devices due to their ability to permanently trap charges at defect sites and demonstrate shrinking hysteresis over repeated P/E operations [7, 19]. Fig. 5-11 (a) presents the endurance of FE–FET and hybrid memory against P/E cycling. The devices were programmed and erased at ±12 V with pulse width of 20 ms each. Hybrid devices show negligible loss of memory window until 10^5 P/E cycles. FE–FET devices however only retain 50 % of their memory window after 1000 cycles. Higher endurance can be attributed to lower fluence of charge in the hybrid memory during P/E cycling. As previously observed from switching measurements (Fig. 5-2(a)), P(VDF−TrFE) film can assist charge transport through grain boundaries or defect sites that may result in higher charging and discharging currents in FE–FET devices during cycling. This is confirmed from Fig. 5-11 (b), which plots gate current density in two devices. FE–FET supports 40 × higher current at –10 V and 20 × higher current at 12 V compared to the hybrid device. The presence of top tunnel oxide in hybrid memory blocks majority of this defect-induced leakage current and therefore reduces the fluence of charge during cycling.

All the above results bring out the major advantages of hybrid design over conventional FE–FET with the same EOT. The choice of ferroelectric material used in these devices was purely based on the ease of integration in the memory gate stack. This enhancement in principle may be achieved through any other ferroelectric material with appropriate film thickness (scales with dielectric constant). For example, oxide-based ferroelectrics may realize comparable electrostatic benefits with large film thicknesses (> 300 nm). Superior performance in programming efficiency and

retention characteristics can be additionally obtained by the use of such materials to demonstrate faster switching, lower leakage and higher cycling endurance. However, their integration at sub–20 nm nodes and beyond may ultimately be constrained by the high device aspect ratio (> 20) and the cell–to–cell interference generated thereof. Regardless of this limitation, hybrid devices can offer significant advantages in low–power embedded memory applications at relaxed device dimensions.

Table 5-2 Comparison of Operational parameters in Memory Devices for ΔV_{TH} = 3 V

Parameter	FE–FET	GI Flash	Hybrid 40:60	Hybrid 60:40
Operating Voltage [V]	15	> 20	12	17
Retention field top oxide [MV/cm]		4.78	2.41	1.39
Retention field bottom oxide [MV/cm]	–1.06	–1.06	–1.29	–1.15
Program field top oxide [MV/cm]		–5.3 (@ –14 V)	–6.7 (@ –12 V)	–9.35 (@ –17 V)
Depolarization field [kV/cm]	237		96	200
Trapped electron density [cm^{-2}]		1.2×10^{13}	8×10^{12}	5.5×10^{12}

5.5 Towards Flexible Hybrid Memory Operation

The alignment of atomic or molecular dipoles to the applied field in the ferroelectric creates huge electric displacement in the adjoining dielectric layers during program operation that helps gate injection of electrons. The magnitude of spontaneous polarization attained during P/E conditions depends on the applied gate bias. In the event of fast switching of the ferroelectric, a huge electric displacement in the tunnel oxide initiates the process of electron injection from the gate. The amount of electrons injected into the storage depends primarily on t_{PROG} besides material parameters like capture cross section and trap emission coefficients. Therefore, contribution from *P–E* hysteresis and injected charge to the total ΔV_{TH} may be tuned by an appropriate choice of P/E voltages and t_{PROG} for any given gate stack geometry and the choice of materials. This offers the possibility of maximizing one of the many criteria in the performance matrix of nonvolatile memories.

As an illustration, we consider the same hybrid gate stack and GI Flash designs mentioned in Table 5-1. However, they are compared against an FE–FET with identical bottom oxide but thicker P(VDF−TrFE) film (53 nm) adding up to the same EOT as the other memory gate stacks. It should be noted that the physical thickness of the hybrid gate stack is smaller than that for FE–FET and this difference would get even more pronounced with high-κ ferroelectric materials. Table 5-2 summarizes the operation parameters for all the memory designs with ΔV_{TH} = 3 V specification. FE–FET is seen to require operating voltage of ±15 V. In GI Flash architecture, this is clearly an impracticable design as this would need over 1.2×10^{13} cm^{-2} electron sheet

charge density and over ±20 V P/E operation. It may never achieve realistic retention times with such large ΔV_{TH} = 3 V due to high fields (> 4.7 MV/cm) in the tunnel oxide.

We investigate two regimes of operation for the hybrid design. For simplicity, we assume the ferroelectric switches much faster than the time scales for charge injection. Hybrid 40:60 is a possible low-voltage design that has a 1.2 V (40 %) contribution in total ΔV_{TH} arising from *P–E* hysteresis and the remaining 60 % from injected charge. The high-voltage design, hybrid 60:40, has just the opposite contributions to the total ΔV_{TH}. As seen from Table 5-2, hybrid 40:60 offers P/E operation at ±12 V and above 50 % reduction in E_{dp} over FE–FET. The electron sheet density is expected to be 8×10^{12} cm^{-2} which would result in longer t_{PROG} and larger fields in the tunnel oxide during retention. Hybrid 60:40 design relies more on the ferroelectric contribution and therefore operates on a larger *P–E* hysteresis. The operating voltage is estimated to be ±17 V. Owing to lower density of the stored electrons, the reduction in depolarization field is less than 20 %. The field enhancement in the tunnel oxide is significant on account of larger contribution from polarization, making it suitable for relatively fast P/E operation. Both designs show significant improvement over GI Flash in P/E and retention electrostatics as seen from the table.

The above illustrative example for flexible operation may provide additional benefit with fast switching ferroelectric thin films. As described earlier, due to slow switching in P(VDF–TrFE), ΔV_{TH} progression with t_{PROG} is indistinguishable between

the two memory mechanisms. However, the possibility of sub-microsecond switching in ferroelectrics can offer their distinct signature in t_{PROG}. Although the overall P/E time is still limited by slow electron tunneling processes, faster switching ferroelectric in principle can demonstrate dual P/E speeds that may open new opportunities for nonvolatile memory integration in embedded system applications.

Estimation of switching time in hybrid gate stack using quasistatic electric field approximation severely underestimates P/E speeds in ferroelectric-based memories. For example, at the onset of the program operation, the remanent polarization opposes the direction of the applied field inside the ferroelectric. In order to generate the same displacement in the entire gate stack, electric field in the ferroelectric is momentarily enhanced above the quasistatic value. The magnitude of enhancement depends upon the thickness, dielectric constant and remanent polarization of the film. Dipole switching is sensitive to the applied field and this enhancement accelerates the process of polarization reversal. Voltage drop in the remainder of the gate stack diminishes to accommodate the higher electric field inside the ferroelectric. Consequently, field sensitive electron tunneling process is suppressed. As time progresses, the aligned dipoles add towards the electric displacement to gradually reduce the effect of enhancement. Upon complete reversal of remanent polarization, electric field in the gate stack settles to the quasistatic value. In other words, ferroelectric switching is electrostatically favored over electron tunneling at the beginning of the P/E operation that may enable the dual speed process. Simulations incorporating simultaneous modeling of kinetics in ferroelectric switching

and charge tunneling can capture such transient effects to provide realistic estimates in P/E speeds in hybrid memory.

5.6 Conclusion

This chapter presents the fabrication and measurement results for the proposed ferroelectric and charge hybrid nonvolatile memory based on the simulation insights given in the previous chapter. Hybrid devices were fabricated with integration of P(VDF−TrFE) copolymer as the ferroelectric and HfO_2 thin film as the charge trap layer. Hybrid devices displayed better memory performance including larger memory window and longer retention times against conventional FE–FET devices. Spontaneous polarization induced field enhancement in the tunnel oxide enabled higher program efficiency over GI Flash. The inclusion of tunnel oxide also reduced the fluence of charge during P/E cycling that facilitated over 100 × improvement in endurance compared to FE–FET design. Hybrid design offers the flexibility of tuning the contribution of *P–E* hysteresis and charge to the total ΔV_{TH} and therefore target specific performance parameters with the same gate stack design.

REFERENCES

[1] S. Lai, "Nonvolatile memory technology: the quest for ever lower cost," *in IEDM Tech. Dig.*, 2008, pp. 1−6.

[2] Y. Nishi and J. R. Jameson, "Recent progress in resistance change memory," in *Proc. Device Research Conf.*, 2008, pp. 271−274.

[3] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, C. Fukumoto, H, Nagao and H. Kano, "A novel nonvolatile memory with spin-torque transfer magnetization switching: spin-ram," in *IEDM Tech. Dig.*, 2005, pp. 459–462.

[4] S. Wang, M. Takahashi, Q. –H. Li, K. Takeuchi and S. Sakai, "Operational method of a ferroelectric (Fe)-NAND flash memory array," *Semicond. Sci. Technol*. Vol. 24, p.105029, Sep. 2009.

[5] R. R. Mehta, B. D. Silverman and J. T. Jacobs, "Depolarization fields in thin ferroelectric films," *J. Appl. Phys.*, vol. 44, no. 8, pp. 3379−3385, Aug. 1973.

[6] P. Ma and J. –P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," *IEEE Electron Device Lett.*, vol. 23, no. 7, pp.386−388, Jul. 2002.

[7] R. Moazzami, C. Hu and W. H. Shepard, "Endurance properties of ferroelectric PZT films," in *IEDM Tech. Dig.*, 1990, pp. 417–420.

[8] S. Sakai, M. Takahashi, K. Takeuchi, Q. –H. Li, T. Horiuchi, S. Wang, K. –Y. Yun, M. Takamiya and T. Sakurai, "Highly scalable Fe (ferroelectric)-NAND cell with MFIS (metal-ferroelectric-insulator-semiconductor) structure for sub-

10 nm terabit capacity NAND Flash memories," in *International Memory Workshop*, 2008, pp.103−105.

[9] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part I: device concept and modeling," submitted to *IEEE Trans. Electron Devices*.

[10] R. C. G. Naber, C. Tanase, P. W. M. Blom, G. H. Gelinck, A. W. Marsman, F. J. Touwslager, S. Setayesh and D. M. Dee Leeuw, "High-performance solution-processed polymer ferroelectric field-effect transistors," *Nature Materials*, vol. 4, no. 3, pp. 243−248, Mar. 2005.

[11] D. Mao, M. A. Quevedo-Lopez, H. Stiegler, B. E. Gnade and H. N. Alshareef, "Optimization of poly(vinylidene fluoride-trifluoroethylene) films as non-volatile memory for flexible electronics," *Organic Electronics*, vol. 11, pp. 925–932, Feb. 2010.

[12] S. J. Kang, Y. J. Park, J. Sung, P. S. Jo, C. Park, K. J. Kim and B. O. Cho, "Spin cast ferroelectric beta poly(vinylidene fluoride) thin films via rapid thermal annealing," *Appl. Phys. Lett.*, vol. 92, p. 012921, Jan. 2008.

[13] J. Choi, C. N. Borca, P. A. Dowben, A. Bune, M. Poulsen, S. Pebley, S. Adenwalla, S. Ducharme, L. Robertson, V. M. Fridkin, S. P. Palto, N. N. Petukhova and S. G. Yudin, "Phase transition in the surface structure in copolymer films of vinylidene fluoride (70 %) with trifluoroethylene (30 %)," *Phys. Rev. B*, vol. 51, no. 8, pp. 5760−5770, Feb. 2000.

[14] S. Fujisaki, H. Ishiwara and Y. Fujisaki, "Low-voltage operation of ferroelectric poly(vinylidene fluoride-trifluoroethylene) copolymer capacitors

and metal-ferroelectric-insulator-semiconductor diodes," *Appl. Phys. Lett*., vol. 90, p. 162902, Apr. 2007.

[15] G. A. Salvatore, D. Bouvet, I. Stolitchnov, N. Setter and A. M. Ionescu, "Low voltage ferroelectric FET with sub-100 nm copolymer P(VDF−TrFE) gate dielectric for nonvolatile 1T memory," in *Proc. European Solid-State Device Research Conf.,* 2008, pp. 162−165.

[16] S. Tiwari, F. Rana, K. Chan, H. Hanafi, W. Chan and D. Buchanan, "Volatile and non-volatile memories in silicon with nano-crystal storage," in *IEDM Tech. Dig.*, 1995, pp. 521−524.

[17] Z. Liu, C. Lee, V. Narayanan, G. Pei and E. C. Kan, "Metal nanocrystal memories, Part I: Device design and fabrication," *IEEE Trans. Electron Devices*, vol. 49, no. 9, pp. 1606−1613, Sep. 2002.

[18] H.-T. Lue, E.-K. Lai, S.-Y. Wang, T. Yang, K.-C. Chen, K.-Y. Hsieh, R. Liu and C.-Y. Lu, "A novel gate-injection program/erase p-channel NAND-type Flash memory with high (10M cycle) endurance," in *Symp. VLSI Tech. Dig.*, 2007, pp. 140−141.

[19] S. L. Miller, R. D. Nasby, J. R. Schwank, M. S. Rodgers and P. V. Dresendorfer, "Device modeling of ferroelectric capacitors," *J. Appl. Phys.*, vol. 68, no. 12, pp. 6463−6471, Dec. 1990.

[20] Y. Tajitsu, "Effects of thickness on ferroelectricity in vinylidene fluoride and trifluoroethylene copolymers," *Jpn. J. Appl. Phys.*, vol. 34, no. 9B, pp. 5418–5423, Sept. 1995.

[21] F. Xia, B. Razavi, H. Xu, Z.–Y. Cheng and Q. M. Zhang, "Dependence of

threshold thickness of crystallization and film morphology on film processing conditions in poly(vinylidene fluoride–trifluoroethylene) copolymer thin films," *J. Appl. Phys.*, vol. 92, no. 6, pp. 3111−3115, Sept. 2002.

[22] H. Xu, X. Liu, X. Fang, H. Xie, G. Li, X. Meng, J. Sun and J. Chu, "Domain stabilization effect of interlayer on ferroelectric poly(vinylidene fluoride–trifluoroethylene) copolymer ultrathin film," *J. Appl. Phys.*, vol. 105, p. 034107, Feb. 2009.

[23] F. Xia, H. Xu, F. Fang, B. Razavi, Z.–Y. Cheng, Y. Lu, B. Xu, Q. M. Zhang, "Thickness dependence of ferroelectric polarization switching in poly(vinylidene fluoride–trifluoroethylene) spin cast films," *Appl. Phys. Lett.*, vol. 78, no. 8, pp. 1122−1124, Feb. 2001.

[24] Q. M. Zhang, H. Xu, F. Fang, Z.–Y. Cheng, F. Xia and H. You, "Critical thickness of crystallization and discontinuous change in ferroelectric behavior with thickness in ferroelectric polymer thin films," *J. Appl. Phys.*, vol. 89, no. 5, pp. 2613−2616, Mar 2001.

[25] P. K. Larsen, G. L. M. Kampschöer, M. J. E. Ulenaers, G. A. C. M. Spierings and R. Cuppens, "Nanosecond switching of thin ferroelectric films," *Appl. Phys. Lett.*, vol. 59, no. 5, pp. 611−613, May 1991.

[26] E. Tokumitsu, N. Tanisake and H. Ishiwara, "Partial switching kinetics of ferroelectric $PbZr_xTi_{1-x}O_3$ thin films prepared by sol–gel techniques," *Jpn. J. Appl. Phys.*, vol. 33, no. 9B, pp. 5201–5206, Sept. 1994.

[27] K. Kimura and H. Ohigashi, "Polarization behavior in vinylidene fluoride–trifluoroethylene copolymer thin films," *Jpn. J. Appl. Phys.*, vol. 25, no. 3, pp.

383–387, Dec 1985.

[28] A. K. Tagantsev, I. Stolitchnov, N. Setter, J. S. Cross and M. Tsukada, "Non-Kolmogorov-Avrami switching kinetics in ferroelectric thin films," *Phys. Rev. B*, vol. 66, no. 21, p. 214109, Dec. 2002.

[29] D. Mao, I. Mejia, H. Stiegler, B. E. Gnade and M. A. Quevedo-Lopez, "Polarization behavior of poly(vinylidene fluoride-trifluoroethylene) copolymer ferroelectric thin film capacitors for nonvolatile memory applications in flexible electronics," *J. Appl. Phys.*, vol. 108, no. 9, p. 094102, Nov. 2010.

[30] A. Chin, M. Y. Yang, C. L. Sun and S. Y. Chen, "Stack gate PZT/Al_2O_3 one transistor ferroelectric memory,*" IEEE. Electron Device Lett.*, vol. 22, no. 7, pp. 336−338, July 2001.

CHAPTER 6 FERROELECTRIC–ASSISTED DUAL–SPEED DRAM–FLASH HYBRID MEMORY

6.1 Abstract

This chapter presents a novel one–transistor low–voltage DRAM–Flash hybrid memory. The proposed device integrates ferroelectric thin film and nonvolatile charge injection, and demonstrates two modes of operations: 1) a fast (10–100 ns) DRAM–like mode with ~ 10^3 seconds retention, associated with ferroelectric switching and 2) a slower (1 ms) Flash–like mode with long retention time, from charge tunneling into floating nodes. The time evolution of the electric field in the ferroelectric and the tunnel oxide is shown to naturally establish a two–step mechanism during the program operation. The complementary characteristics of ferroelectric switching and gate charge–injection enable low–voltage program/erase (± 8 V), large memory window (0.8 V) and long retention time (> 10 years). Devices were fabricated with the lead zirconium titanate (PZT) thin film as the ferroelectric layer and Au nanocrystals for gate–injected electron storage. Pulsed programming measurements were also performed to distinguish the memory window obtained from the two mechanisms in DRAM and Flash operations.

6.2 Introduction

The rapid growth in mobile computing coupled with the demand for improved energy–efficient data centers has lead to a continued interest in the discovery of new nonvolatile memory devices with higher speeds and lower power consumption. This category of ‘storage–class memory’ is expected to eventually bridge the gap between

conventional nonvolatile devices like NAND Flash and fast–access memory like DRAM with short latencies [1]. Two–terminal devices that rely on resistance switching, such as the phase change (PC) RAM, resistive (Re) RAM and magnetic RAM, have demonstrated considerable promise in realizing fast program and long retention with sub–5 V operating voltages [2–6]. However, they still suffer from access device scalability limitation and peak switching current compliance issues that impede their application to low–power mass storage. For now, Flash memory still remains the major work horse for off–chip data storage due to its high density and low power architecture.

Ferroelectric (FE) FET–based memories present a viable approach to storage–class functionality [7–9]. They combine the high speed and low voltage operation inherent to ferroelectric switching with the non–destructive readout intrinsic to FET. However, the strong remanent polarization (P_r) cannot be compensated thoroughly by the underlying semiconductor during retention. This sets up a depolarization field (E_{dp}) that promotes randomization of aligned dipoles. The high field induced in the blocking layer can also induce tunneling from the substrate. These effects are known to limit the retention in FE–FETs [10, 11] to days. Compensating charge of opposite polarity placed near the ferroelectric surface polarization charge can effectively reduce E_{dp} and improve retention. This was demonstrated in a novel ferroelectric and charge hybrid device that incorporates gate injection of electrons during the program operation [12, 13]. This chapter extends the idea of these complementary memory mechanisms to demonstrate two modes of program characteristics that naturally result from the distinct time scales in ferroelectric switching and charge injection: 1) fast

DRAM–like switching that arises from ferroelectric polarization and 2) slower Flash–like electron tunneling to the floating nodes that increases memory window (ΔV_{TH}) and improves retention by reducing E_{dp} in ferroelectric and the electric field in the tunnel oxide, simultaneously.

The prospects of DRAM and Flash integration on a single transistor have been previously explored [14]. This unified RAM (URAM) combines the functionality of capacitorless DRAM and a charge–trap Flash in a one–transistor memory gate stack. The proposed ferroelectric and charge hybrid design differs from URAM in two main ways. The two memory mechanisms in URAM, namely storage of holes in the SOI body and storage of channel–injected electrons in dielectric traps, generate ΔV_{TH} with opposite polarities. Secondly, hole generation in URAM occurs by avalanche breakdown at the drain junction that necessitates high current and independent control of the drain node. The two memory mechanisms in the proposed hybrid device are additive in ΔV_{TH}, and do not require high drain fields, which enables both NAND and NOR configurations.

This chapter describes the concept and operation of the one–transistor DRAM–Flash hybrid memory as well as discusses the measurements obtained with the proposed design. The chapter is organized as follows. Section 6-3 explains the dual–speed program mechanism inherent to the hybrid gate stacks. Devices were fabricated with the lead zirconium titanate (PZT) thin film as the ferroelectric and Au nanocrystals (NCs) as charge storage nodes. Section 6-4 outlines the fabrication process of the device. Characterization of PZT switching and electrical measurement

results are discussed in Section 6-5. Section 6-6 describes the future prospects and limitations.

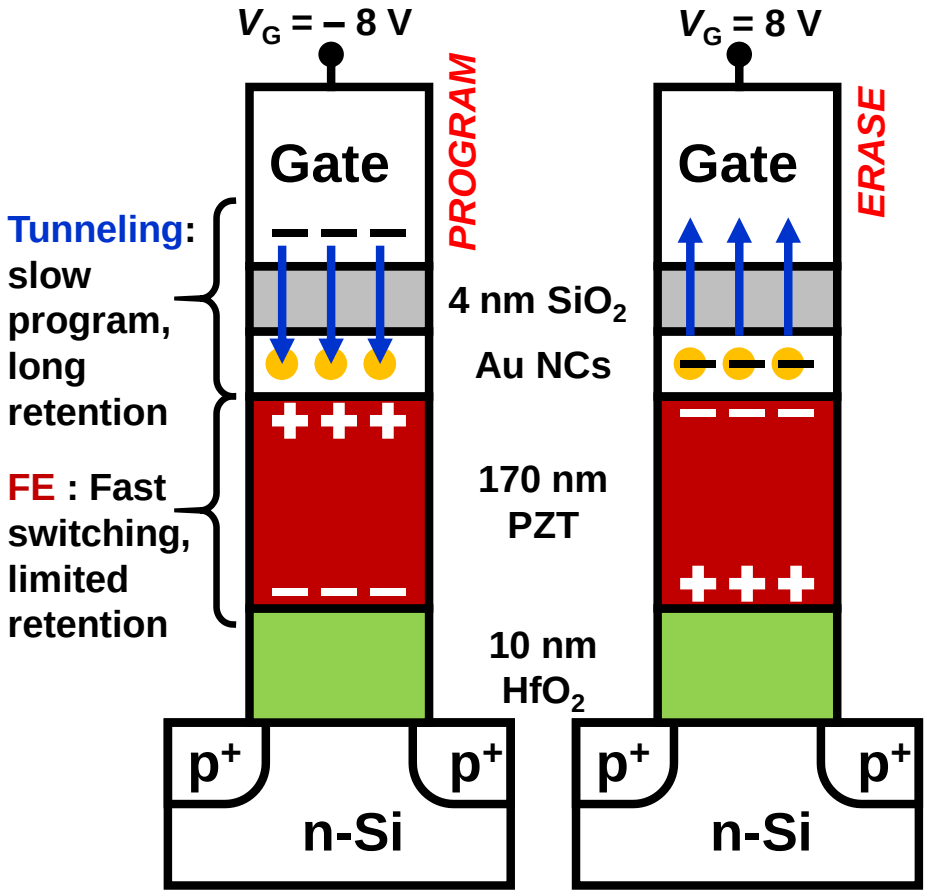


Figure 6-1 Schematic of ferroelectric–assisted DRAM–Flash hybrid memory. Program operation in the DRAM mode involves the short program pulse that aligns the ferroelectric domains in the PZT film. On prolonged program, the device enters the Flash mode with gate injection into the charge trap layer (Au nanocrystals here). Trapped charge adds to the memory window and reduces ferroelectric depolarization field during retention condition

6.3 The Dual-speed Operating Principle

A ferroelectric and charge hybrid nonvolatile memory has been previously demonstrated [13] with slow-switching polyvinylidene fluoride-trifluoroethylene [P(VDF−TrFE)] copolymer. Fig. 6-1 shows the revised device schematic for the proposed dual–speed DRAM–Flash version. The program operation is performed with a negative gate bias that favors the alignment of ferroelectric polarization to the applied field. DRAM–like switching arises from fast–responding PZT dipole polarization, while Flash–like operation is obtained by gate–injection of electrons into the Au NCs. The mode of operation (DRAM or Flash) for a specific program voltage (V_{PROG}) depends solely on the program time (t_{PROG}). The kinetics of ferroelectric domain switching and of electron tunneling is extremely sensitive to the electric field in the respective materials [15–18]. Therefore, in order to comprehend the dynamics of total ΔV_{TH} evolution with t_{PROG} as well as the contribution from each mechanism, we need to analyze the electric fields in the two coupling layers.

To simplify our analysis, we examine 1–D electrostatics in the gate stack direction. Furthermore, we assume no net charge is stored on the floating nodes at the start of the program operation, where the electric displacement is continuous across all gate stack layers. This is shown in Eq. (1). D_{FE}, D_{TOX} and D_{BOX} are the displacements in the ferroelectric, tunnel oxide and bottom dielectric, ε_{FE}, ε_{TOX} and ε_{BOX} are the dielectric constants, E_{FE}, E_{TOX} and E_{BOX} are the corresponding electric fields and P represents the dipole polarization.

$$D_{FE} = \varepsilon_{FE} E_{FE} + P = D_{TOX} = \varepsilon_{TOX} E_{TOX} = \varepsilon_{BOX} E_{BOX} \tag{1}$$

For the fixed V_{PROG} at any t_{PROG},

$$V_{PROG} = V_{Si} + E_{BOX} t_{BOX} + E_{FE} t_{FE} + E_{TOX} t_{TOX} \tag{2}$$

where V_{Si} is the voltage drop in silicon, and t_{BOX}, t_{TOX} and t_{FE} are the thicknesses of the bottom dielectric, tunnel oxide and the ferroelectric. E_{FE} can then be evaluated by solving Eqs. (1) and (2) to obtain,

$$E_{FE} = \left(\frac{V_{PROG} - V_{Si}}{t_{FE}} - \frac{P}{t_{FE} C_{||}} \right) \Bigg/ \left(1 + \frac{C_{FE}}{C_{||}} \right)$$

$$\frac{1}{C_{||}} = \frac{t_{BOX}}{\varepsilon_{BOX}} + \frac{t_{TOX}}{\varepsilon_{TOX}} \; ; \; C_{FE} = \frac{\varepsilon_{FE}}{t_{FE}} \tag{3}$$

where $C_{||}$ is capacitance of all other dielectrics in the gate stack. For simplification, we assume that P is initially randomized (P = 0 at $t_{\mathrm{PROG}} = 0^{-}$). At the beginning of program ($t_{\mathrm{PROG}} = 0^{+}$), E_{FE} attains the highest value. As t_{PROG} progresses, P gradually aligns to the applied electric field ($|P| > 0$) and adds to the displacement in the ferroelectric. Therefore, magnitude of E_{FE} diminishes with increasing P, as seen from Eq. (3). This generates field enhancement in the tunnel dielectric that initiates electron tunneling from the gate electrode. In other words, *ferroelectric polarization switching is favored* at the beginning of the program operation and electron tunneling becomes appreciable only after sufficient dipole alignment has been developed. It should be noted that this effect would be enhanced if initial polarization is aligned opposite to

the applied field from erase (P and E_{FE} have opposite polarity at $t_{PROG} = 0^+$). Secondly, the degree of reduction in E_{FE} (consequently the enhancement in E_{TOX}) depends on t_{FE} and $C_{||}$. This effect would be more prominent for larger contribution from non–ferroelectric layers to the total EOT (lower t_{FE} and smaller $C_{||}$).

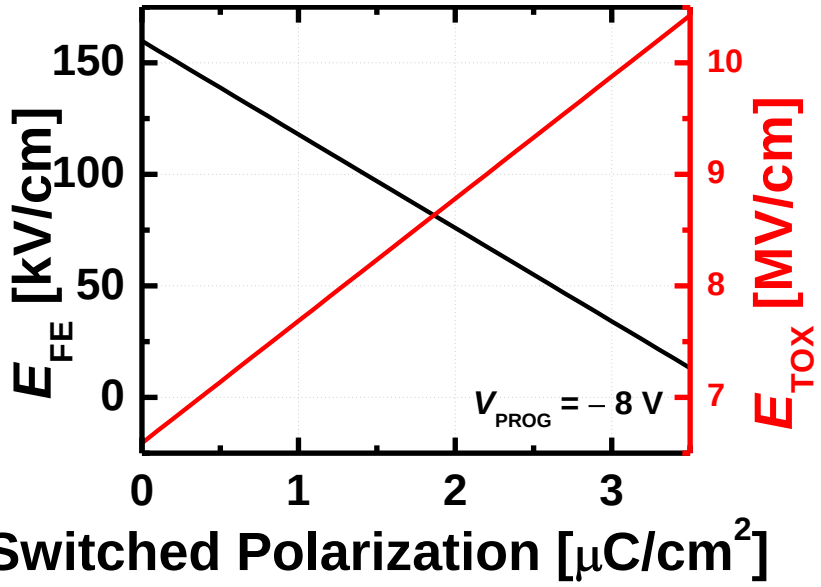


Figure 6-2 Magnitude of the electric field in the ferroelectric (E_{FE}) and tunnel oxide (E_{TOX}) as a function of switched polarization (P). Because P evolves from a randomized state, magnitude E_{FE} is highest at the beginning of program operation. As t_{PROG} progresses, P increases and the voltage drop in ferroelectric is transferred to the adjoining dielectric layers. E_{TOX} is boosted over 2 MV/cm after P rises above 2 μC/cm^2, initiating significant electron injection from the gate.

Let us consider the specific case of the proposed memory gate stack. Fig. 6-2 shows the dependence of the magnitude of E_{FE} and E_{TOX} on P. The electronic dielectric constants considered here are $\varepsilon_{FE} = 165$ and $\varepsilon_{TOX} = 3.9$, as observed from our

measurements in Section 6-5. It is known that switching times measured oxide–based ferroelectric thin film capacitors are exponentially dependent on applied E_{FE}. Particularly for PZT, which depicts high spontaneous polarization (P_s > 15 μC/cm^2), over 90 % of dipoles align to the applied field in 1 μs for E_{FE} = 160 kV/cm [15]. This switching time is drastically reduced below 100 ns for E_{FE} = 220 kV/cm. For V_{PROG} = – 8 V, the magnitude of E_{FE} is over 150 kV/cm. With increase in t_{PROG}, E_{FE} is reduced proportionally to the total P. Unlike ferroelectric capacitors, complete polarization switching cannot be achieved in FE–FET or hybrid memory structures. Maximum switched polarization is usually limited by the maximum displacement that can be sustained by the blocking dielectric, (~3 μC/cm^2 for SiO_2). Switching time required for such incomplete polarization (~ 0.2P_s) can be over an order of magnitude lower than that required for complete dipole alignment. Further, such incomplete polarization is sufficient to generate reasonable ΔV_{TH} in ferroelectric transistors. This is confirmed by the fast switching observed in PZT and other FE–FETs even under low voltage operation [8, 19, 20].

To summarize, preferential switching in the ferroelectric followed by field enhancement in the tunnel dielectric can bring forth a two–step program/retention characteristics. Due to the exponential dependence of switching rate on E_{FE}, the dipole alignment eventually saturates. For example, as seen from Fig. 6-2, E_{FE} equals 75 kV/cm when P reaches 2 μC/cm^2, at which point the rate of polarization switching is nearly diminished. This short t_{PROG} operation is designated as the DRAM mode. Both E_{TOX} and t_{PROG} are insufficient to cause any electron tunneling in this mode and the ΔV_{TH} in the hybrid device is entirely due to dipole polarization. Retention in this state

is usually poor (at best days depending on the gate stack parameters) owing to dipole randomization induced by E_{dp}. An ideal t_{PROG} (10 ns – 1 µs) for the DRAM mode would be constrained by the minimum ΔV_{TH} requirements as well as the quality of the ferroelectric film. Further, t_{PROG} may be reduced by optimizing the EOT of the ferroelectric or increasing V_{PROG}.

On continued programming beyond the DRAM mode, the hybrid device naturally enters the Flash mode. Because P is high, the value of E_{TOX} in the Flash mode is enhanced by over 2 MV/cm compared to the DRAM mode. Therefore, as t_{PROG} progresses (~ 100 µs – 1 ms), electron tunneling starts to dominate and adds to the total ΔV_{TH} of the hybrid device. The trapped electrons partially compensate the ferroelectric polarization to reduce E_{dp} during retention. The Flash mode is thus characterized by higher ΔV_{TH} with significant contribution from electron charge. The reduction in E_{dp} also facilitates longer retention times.

Polarization switching and electron storage in the floating nodes together enlarge the field in the bottom oxide. The maximum ΔV_{TH} obtained in the Flash mode would be determined by the maximum field supported by the bottom dielectric before substantial hole injection from the channel occurs. Erase operation is performed by applying a positive gate bias and follows similar dynamics as the program operation.

6.4 Device Fabrication

One–transistor hybrid devices were fabricated with the modified gate–last process. The PMOS transistor was chosen as the sensing channel in order to align read

voltage to V_{PROG} and thereby overcome read disturb. N–type silicon substrates with resistivity of 5–20 Ω–cm were patterned to define the active region by shallow trench isolation (STI). Boron implantation was performed in the source/drain (S/D) region followed by dopant activation. 10 nm HfO_2 was deposited at 300 ºC by plasma enhanced atomic layer deposition (PE–ALD). The bottom oxide needs to disable hole tunneling from the channel and also block interdiffusion of PZT in silicon during anneal. PZT was deposited on HfO_2 using a RF magnetron sputtering system in the on-axis configuration from a lead zirconate titanate [Pb $(Zr_{0.52}Ti_{0.48})O_3$] ceramic target with 20 % excess PbO [21]. The thickness of the ferroelectric film was 170 nm. The film went through rapid thermal annealing at 750 ºC for 60 seconds. NCs were self–assembled over PZT by evaporation of 12 Å of Au and subsequent annealing at 650 ºC for 60 seconds [22]. 4 nm PE–ALD SiO_2 deposited at 300 ºC formed the tunnel oxide. The Ti/Al gate electrode was patterned by wet etch and Ni/Ti/Au S/D contacts were formed by lift–off. We also fabricated FE–FETs with the same gate stack configuration as the hybrid device excluding Au NCs, and gate-inject (GI) Flash transistors with the PZT layer replaced by 4 nm SiO_2. The measurements from these control devices were used not only to benchmark the hybrid device under the same process flow but also isolate the contribution of individual memory components. Pt/PZT/Pt metal–ferroelectric–metal (MFM) capacitors were also made to study quasi–static and switching properties of the deposited ferroelectric film.

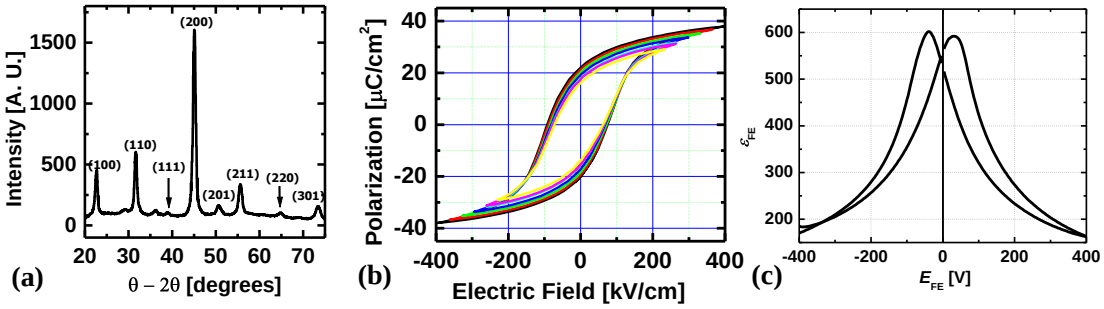


Figure 6-3 (a) X–ray diffraction measurement showing formation of ferroelectric PZT on HfO_2 after 725 ºC anneal for 60s. (b) Polarization–electric field hysteresis at 10 kHz in Pt/PZT/Pt capacitors measured for peak electric fields ranging from 200 kV/cm to 400 kV/cm. (c) Small–signal *CV* measurement showing 4× higher capacitance at coercive voltage compared to high fields measured at 10 kHz.

6.5 Results and Discussion

6.5.1. Ferroelectric Thin Film Characterization

Fig. 6-3 (a) shows the X–ray diffraction analysis of 170 nm PZT sputtered on 10 nm HfO_2 after annealing. The diffraction pattern shows reflections indicative of PZT formation with strong peaks at (100), (110), and (200). Intermediate phase formation was minimal and may be due to the lead–rich nature of the target and the annealing conditions. The maximum anneal temperature was limited to 750 ºC to avoid microcracking and interdiffusion. Fig. 6-3 (b) depicts the polarization–electric field (*P–E*) hysteresis in MFM capacitors measured by the Sawyer–Tower circuit [23]. The electric field was applied at 10 kHz with peak fields varied from 200 kV/cm to 400 kV/cm. The spontaneous polarization (P_s) and remanent polarization (P_r) were observed to be 28 μC/cm^2 and 16 μC/cm^2 at 200 kV/cm. The coercive field (E_c) is 75 kV/cm as seen from Fig. 6-3 (b). The dielectric constant as a function of the applied

field was obtained by small–signal capacitance measurements at 10 kHz. As shown in Fig. 6-3 (c), high capacitance was observed near $\pm E_c$ from the contribution of switching dipole to the total dielectric response. The electronic dielectric constant (ε_{FE}) measured after complete polarization switching was 165.

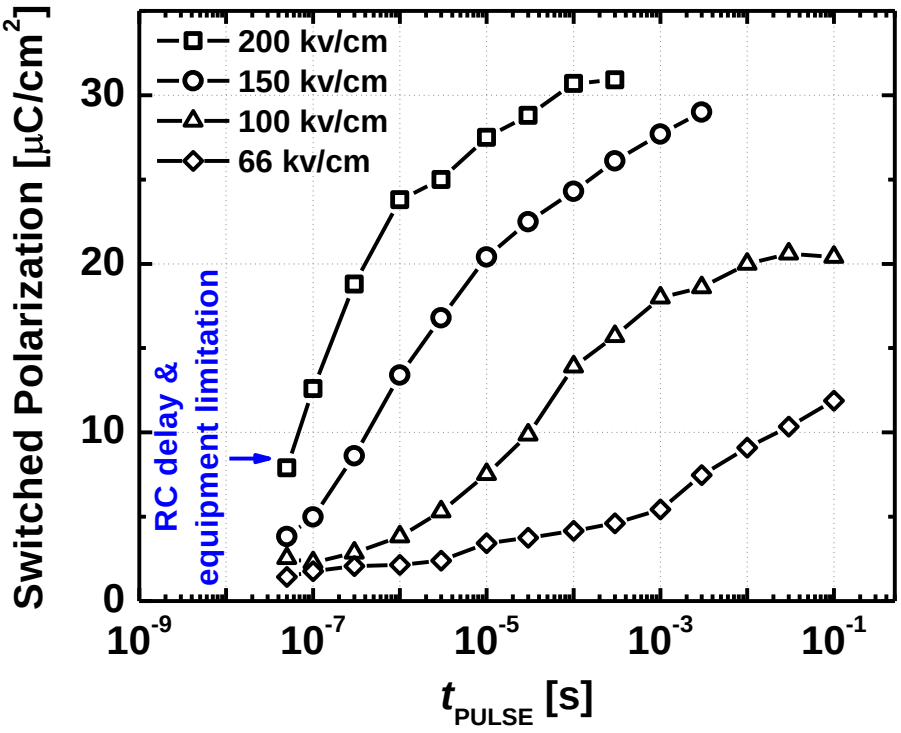


Figure 6-4 Pulsed measurements evaluating polarization switching in Pt/PZT/Pt capacitors for various E_{FE}. Switching time observed to align 80 % of the domains is 100 μs at 150 kV/cm and 2 μs at 200 kV/cm.

The switching response was determined by an altered positive–up–negative–down (PUND) procedure [24]. A positive 6 V, 1 μs pulse was applied to attain complete alignment of polarization. The following negative pulse of magnitude '– *V*' and time 't_{PULSE}' achieved partial reversal of domains. A subsequent 6 V, 1 μs pulse

was used to measure the polarization of reversed domains and a final identical pulse corrected for the electronic component of polarization charge as well as back–switching dipoles. Fig. 6-4 shows the switched polarization for t_{PULSE} ranging from 50 ns to 100 ms for various applied electric fields ($E_{FE} = V/t_{FE}$). As discussed earlier, the switching time depends exponentially on the applied voltage. At constant E_{FE} = 200 kV/cm, 80 % switching of the polarization is achieved below 1 μs. This time increases two orders of magnitude to 100 μs at E_{FE} = 150 kV/cm.

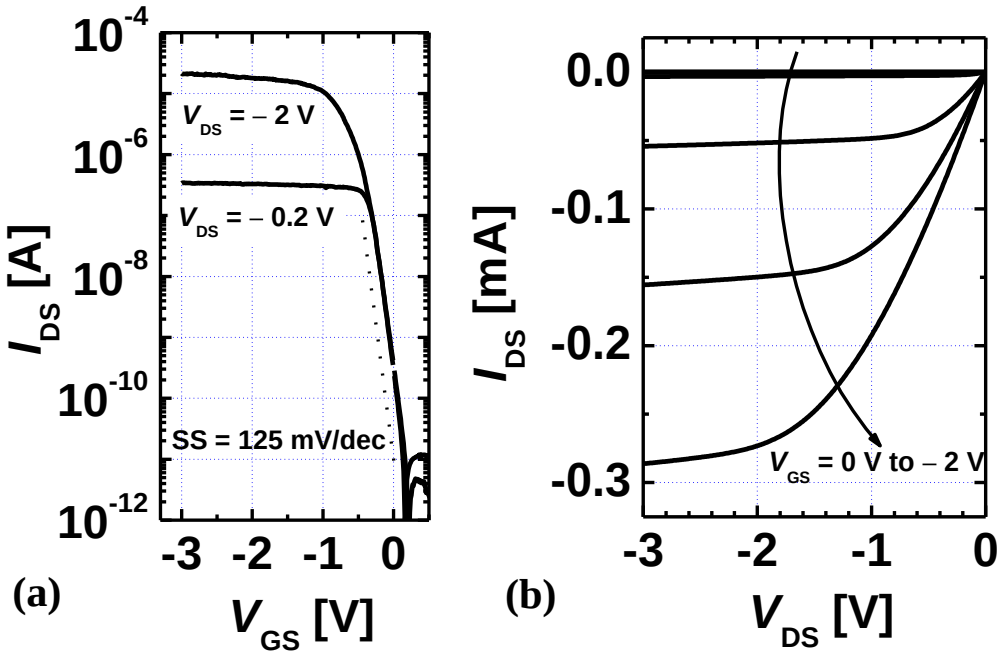


Figure 6-5 (a) Transfer characteristics for the hybrid device at V_{DS} = –0.2 V and –2 V. Subthreshold slope observed was 125 mV/decade. (b) Output characteristics for V_{GS} ranging from 0 V to –2 V.

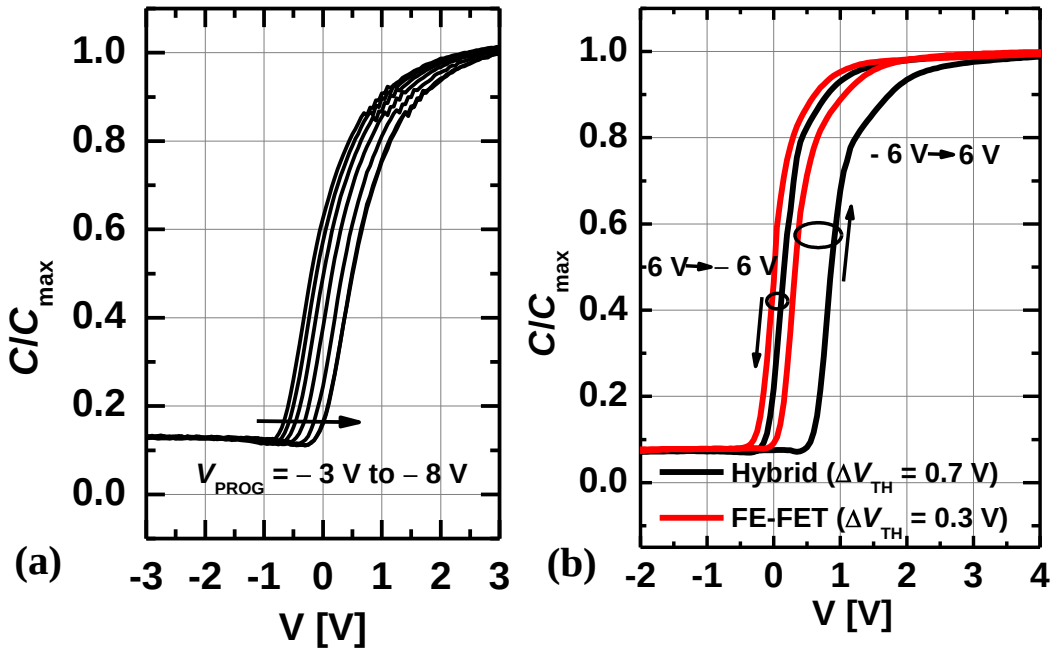


Figure 6-6 (a) *CV* characteristics showing the program window in the hybrid device for V_{PROG} ranging from – 3 V to – 8 V. (b) Comparison between *CV* hysteresis at ± 6 V reveals 0.4 V larger ΔV_{TH} for the hybrid device over FE–FET.

It should be noted that the switching experiment measures dipole switching from a completely reversed polarization state. In comparison, the hybrid device operation is limited to ± 3 μC/cm^2 from the randomized state. Therefore, the region of interest lies about 16 ± 3 μC/cm^2 on the switching data. Depending on the applied V_{PROG}, E_{FE} at the beginning of program operation can be ~ 150 – 200 kV/cm (based on Fig. 6-2 for V_{PROG} = – 8 V). At such fields, initial dipole alignment (~ 1 μC/cm^2) can be reduced below 100 ns to accomplish the DRAM mode operation.

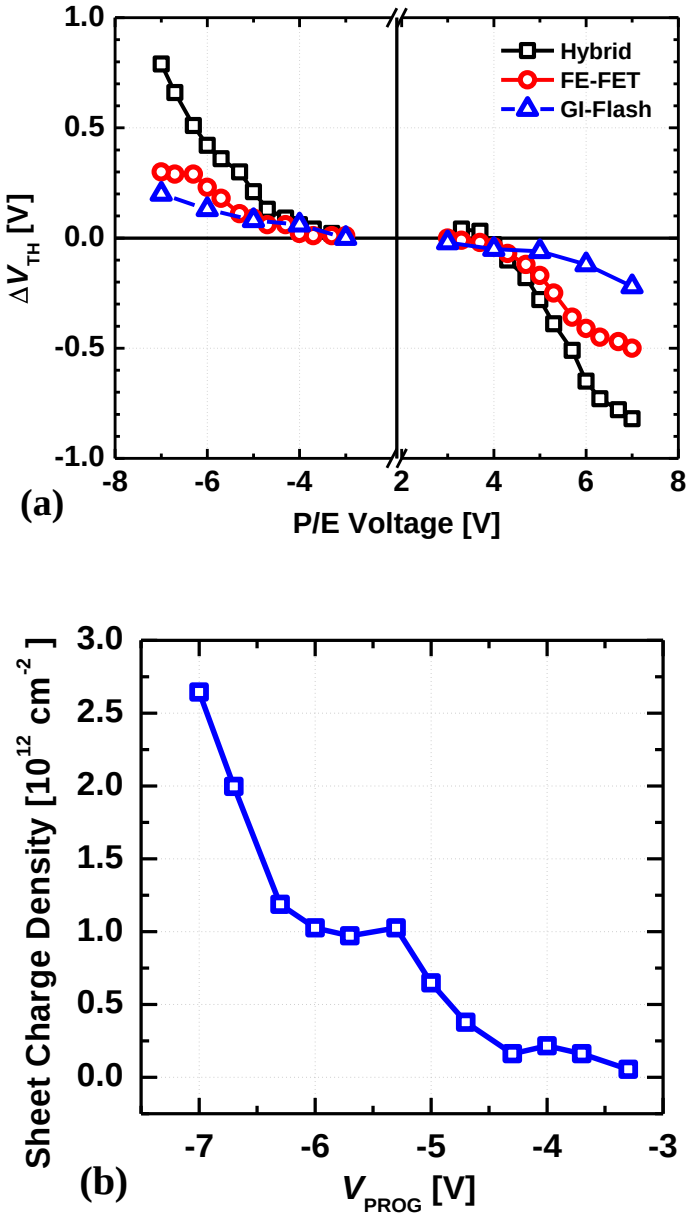


Figure 6-7 (a) Summary of program and erase ΔV_{TH} measurements for hybrid, FE–FET and GI–Flash device. (b) Estimated sheet charge density of electron stored in the Au NC layer as a function of V_{PROG} in the hybrid device.

6.5.2. DRAM–Flash Hybrid Device Measurements

Fig. 6-5 (a) represents the transfer characteristics of the hybrid device with W = 30 µm and L = 5 µm at V_{DS} = 0.2 V and 2 V. The devices achieved I_{ON}/I_{OFF} ratios above 10^6 with subthreshold slopes ranging from 125 – 200 mV/decade. Devices did not show any ΔV_{TH} hysteresis within ± 2 V operation. Fig. 6-5 (b) shows the output characteristics of the same device for varying V_{GS}. In practical memory operation, the drain bias was limited to – 2 V to avoid hole injection over the HfO_2 bottom dielectric.

Fig. 6-6 (a) depicts the quasi–static *CV* characteristics for the hybrid device after program operation. Devices demonstrate steady V_{TH} shift for V_{PROG} = – 3 V to – 8 V. Due to the contribution of injected electrons to the total ΔV_{TH}, the hybrid device shows a higher memory window compared to that of a FE–FET. In Fig. 6-6 (b), *CV* hysteresis is compared between the two devices at ± 6 V. Fig. 6-7 (a) summarizes the ΔV_{TH} analysis extracted from constant current measurements for varying program and erase voltages (V_{PROG} and V_{ERASE}) applied for 100 ms. The FE–FET shows comparable ΔV_{TH} to hybrid device for V_{PROG} > – 5 V. At such low voltages, E_{TOX} is too low to induce any electron tunneling in the hybrid device. However, a clear distinction can be observed between the two devices for V_{PROG} < – 5 V. The hybrid device shows continued rise in ΔV_{TH} over FE–FET due to injected–electron contribution. For V_{PROG} = – 8 V, ΔV_{TH} in the hybrid memory is 0.8 V in comparison to 0.35 V in the FE–FET. ΔV_{TH} in GI Flash transistors is small, and demonstrates weak dependence on V_{PROG}. Due to the large capacitance of the tunnel oxide from the gate (4 nm SiO_2), the injected electrons have modest influence on ΔV_{TH}. The fabricated FE–FET devices are electrostatically equivalent to the hybrid devices without injected charge. Therefore,

electron stored charge density and its contribution to total ΔV_{TH} (Fig. 6-7 (b)) can be estimated in the hybrid device as a function of V_{PROG}. The charge density is predicted to be below 10^{11} cm^{-2} at V_{PROG} = – 4 V but rapidly rises above 2.5 × 10^{12} cm^{-2} for V_{PROG} = – 7 V, at which point its contribution to total ΔV_{TH} is over 50 %. The erase characteristics also follow a similar trend across all devices.

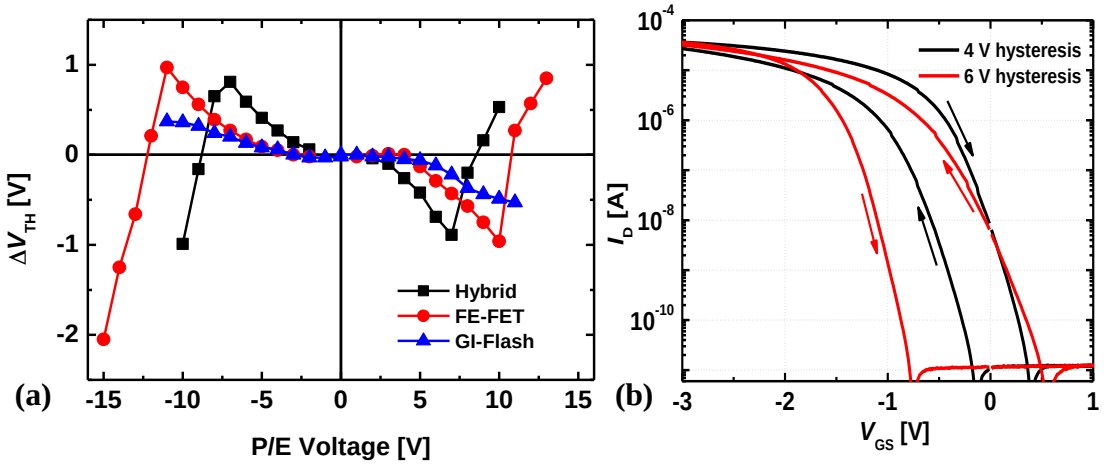


Figure 6-8 (a) High–voltage program and erase ΔV_{TH} measurements showing reversal of memory window above ± 8 V for the hybrid device and ±11 V for FE–FET. (b) Quasi–static *IV* hysteresis comparison before and after ΔV_{TH} reversal showing consequent skewing in *IV* characteristics.

For larger magnitude of V_{PROG}, an anomalous reversal in memory window is observed for both ferroelectric–based devices. Since the two devices are designed to be electrostatically equivalent, the reversal V_{PROG} only depends on the maximum positive ΔV_{TH} that the memory gate stacks can support. In other words, the reversal occurs at smaller $|V_{PROG}|$ for the hybrid device than that of FE–FET. Fig. 6-8 (a) shows

the reversal V_{PROG} of hybrid and FE–FET devices to be – 8 V and – 11 V, respectively. This phenomenon has been previously observed for other high–κ dielectrics integrated in PZT FE–FETs [25, 26].

Conventional Flash devices usually show saturation in ΔV_{TH} with increasing V_{PROG}. As more and more charge is injected in the floating gate, its potential rises and steadily enhances the field in the blocking dielectric. With continued increase in V_{PROG}, most of the injected charge is lost to the tunneling current out of the floating gate. The situation is different in case of the FE–FET and the hybrid device with the ΔV_{TH} reversal. During program operation, the field in the bottom dielectric gradually builds up as the result of polarization switching and electron gate–injection (in the hybrid device). This eventually initiates tunneling of holes from the substrate. Holes are injected over the bottom oxide into the dead layer formed at the ferroelectric interface as a result of high temperature annealing. This unintentional storage layer has low capacitance from the control gate. Therefore, ΔV_{TH} generated by these trapped holes is of an opposite polarity and can be larger compared to the ΔV_{TH} together caused by polarization and gate–injected electrons. This leads to an abrupt reversal in the threshold voltage. A similar trend is also observed in the erase characteristics and can be attributed to electron injection from the channel.

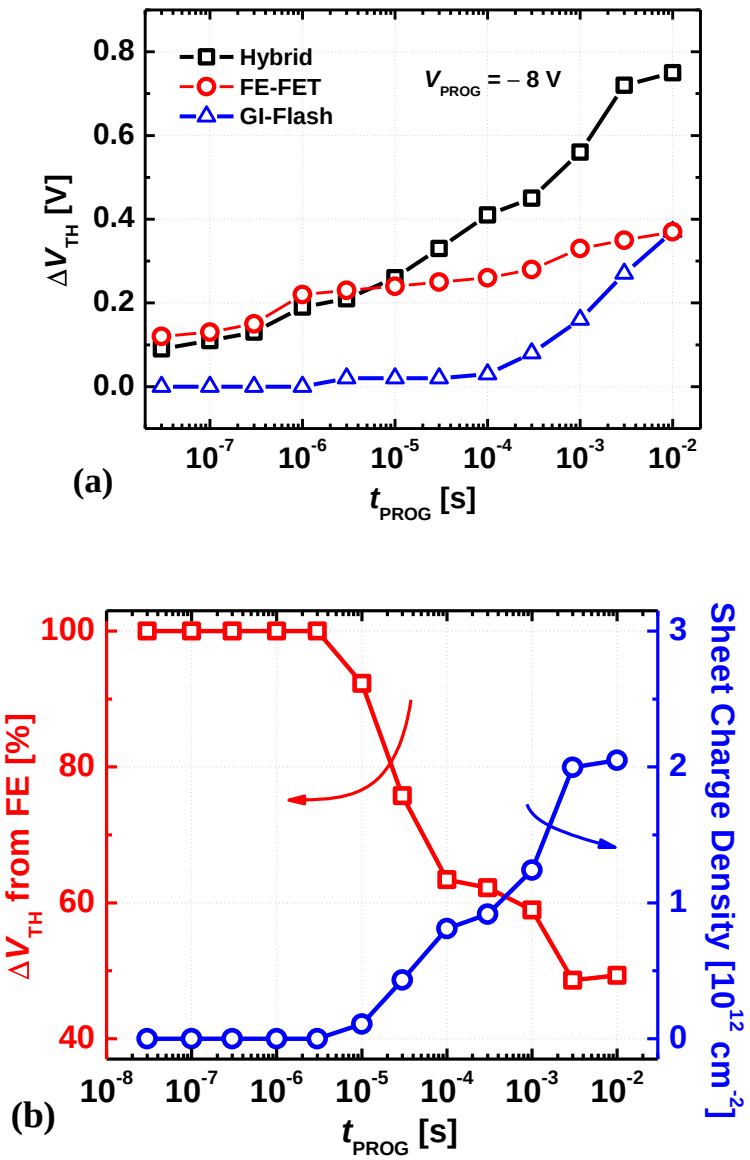


Figure 6-9 (a) Pulsed program measurements for hybrid, FE–FET and GI–Flash devices at V_{PROG} = – 8 V. Sub–microsecond ΔV_{TH} > 0.1 V is observed in both ferroelectric–based devices. (b) Estimated contribution of ferroelectric polarization to total ΔV_{TH} and sheet charge density as a function of t_{PROG}.

The reversal V_{PROG} and maximum positive ΔV_{TH} obtained depends on t_{PROG}, high–κ and PZT thickness as well as the material properties such as bottom dielectric band offsets with silicon, dielectric constant and interfacial quality. For example, Al_2O_3 shows negligible positive ΔV_{TH} before reversal [27]. On the other hand, La_2O_3 and Dy_2O_3 show high field tolerance to hole injection [28, 19]. However, this ΔV_{TH} reversal cannot be used for erase, as subsequent program will follow a different hysteresis loop with short retention time. Fig. 6-8 (b) shows the quasistatic *IV* hysteresis before (± 4 V) and after (± 6 V) ΔV_{TH} reversal in the hybrid device. The consequent skewing of *IV* characteristics upon reversal is a result of the trapped holes in HfO_2 responding to the applied bias.

Fig. 6-9 depicts the program pulse measurements on all devices at $V_{PROG} = -8$ V. The applied t_{PROG} was varied from 30 ns to 10 ms. As seen from Fig. 6-9 (a), both hybrid and FE–FET devices show DRAM–like operation ($\Delta V_{TH} > 0.1$ V) for t_{PROG} below 100 ns. This is attributed to an enhanced E_{FE} (~ 170 kV/cm) that causes rapid polarization dipole alignment at the start of the program operation (see Fig. 6-2). The estimated polarization in PZT for $\Delta V_{TH} = 0.15$ V is 0.5 μC/cm^2, as calculated from $\Delta V_{TH} = P/C_{FE}$. When these initial domains align, they diminish E_{FE} (< 125 kV/cm) to retard subsequent polarization switching. According to Fig. 6-4, switching time for $E_{FE} = 100$ kV/cm at 16 ± 1 μC/cm^2 is seen to be 5 – 10 μs and even higher at lower E_{FE}. This explains the gradual saturation in ΔV_{TH} observed for both devices beyond $t_{PROG} = 1$ μs. Optimal DRAM-mode t_{PROG} can be chosen to meet the minimal ΔV_{TH} requirements. For a fixed t_{PROG}, ΔV_{TH} can be improved by increasing V_{PROG} and t_{FE}.

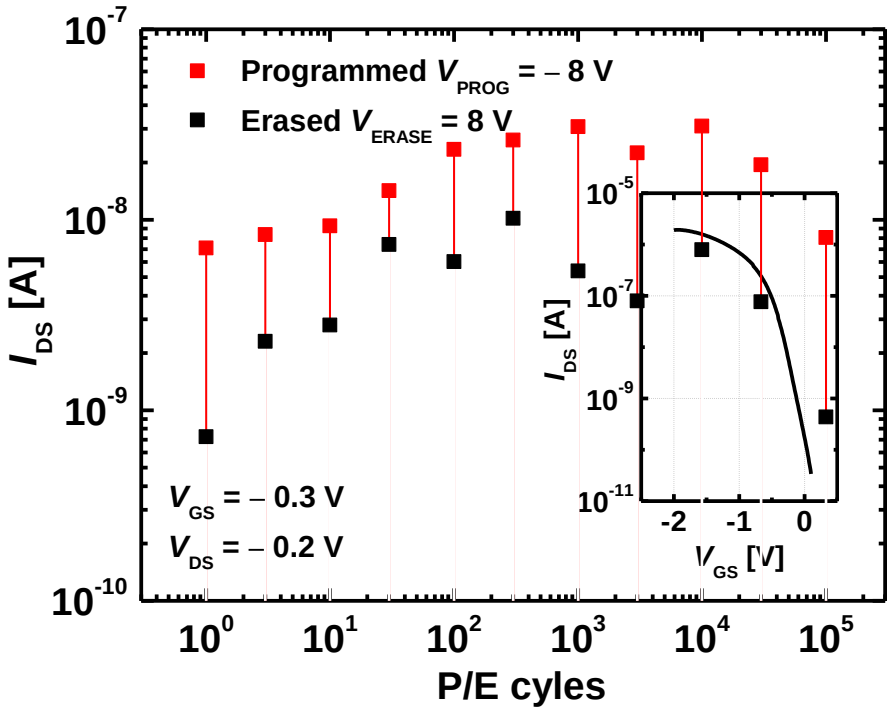


Figure 6-10 Program/erase cycling for the hybrid device with L = W = 5 µm in the DRAM mode at stress voltage of ± 8 V.

In the DRAM mode, ΔV_{TH} arises exclusively from polarization switching, and both ferroelectric devices show similar ΔV_{TH} for short t_{PROG}. There is however a clear distinction between the two devices beyond t_{PROG} = 100 µs due to the contribution of injected electrons in the hybrid device. Electron tunneling is also evident in GI Flash device with $\Delta V_{TH} > 0$ for $t_{PROG} > 100$ µs. The minimum t_{PROG} in Flash mode depends on the total ΔV_{TH} requirements after 10–year retention time. It can be optimized by increasing t_{TOX} and V_{PROG}. Fig. 6-9 (b) estimates the contribution of polarization to the total ΔV_{TH} and the injected sheet charge density in the hybrid device at every t_{PROG}. The ferroelectric contribution is seen to reduce from 100 % at 1 µs (DRAM mode) to

50 % at 10 ms (Flash mode). The estimated electron density in the Flash mode is ~ 2×10^{12} cm^{-2}.

Fig. 6-10 depicts the hybrid memory performance in terms of program/erase (P/E) cycling for the DRAM mode. 1 µs P/E pulses at ± 8 V were applied and the consequent modulation in channel conductance was measured. Devices displayed considerable variation in conductance especially after the erase operation. Phase formation and fatigue mechanisms of PZT on metal electrodes have been well studied [29, 30]. However, the growth and switching characteristics of PZT deposited on amorphous oxides are less understood. Variability in ΔV_{TH} after fast P/E cycling may result from movement of trapped charge in the bulk of the ferroelectric film. The cycling stability can be partially improved by optimizing the deposition and annealing parameters of the PZT. A lead–rich target was used to compensate the PbO loss during post-annealing. However, it has been shown that lead–rich films can sometimes cause the appearance of lead–rich intermediary phases such as AB_3O_7-type lead titanate and lead zirconate [31]. The weak reflections in XRD analysis at around 30^o and 36^o indicate the presence of intermediary phases, which might promote unwanted charge trapping and detrapping in the bulk during P/E operations. Further, domain switching is known to be dependent on the grain size as well as the nature of the interfacial dead layer [32, 33]. These material dependent effects need thorough scrutiny to reduce variability in device performance.

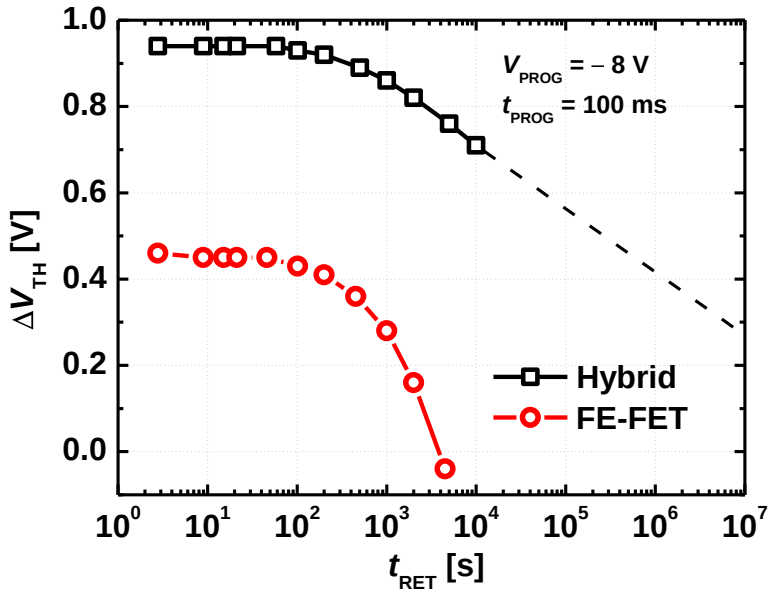


Figure 6-11 Retention measurement comparison for FE–FET (DRAM mode) and the hybrid device (Flash mode) for V_{PROG} = – 8 V at t_{PROG} = 100 ms. The hybrid device shows higher ΔV_{TH} and longer retention due to reduction in E_{dp} by injected electrons over FE–FET.

Fig. 6-11 represents the retention measurements on the hybrid and FE–FET devices. Devices were programmed at – 8 V for 100 ms to force the hybrid device in Flash operation. Due to lack of charge storage layer, the FE–FET remains in the DRAM mode even after such large t_{PROG}. Electron storage in Au NCs partially compensates the polarization charge in the ferroelectric to reduce E_{dp}. Therefore, the Flash mode shows much longer retention with lower rate of ΔV_{TH} loss compared to DRAM mode. There is a clear trade–off in increasing the contribution of electron storage to the total ΔV_{TH}. While storing more electrons on Au NCs reduces the extent of ferroelectric depolarization, it also increases the retention field in the tunnel oxide.

This aspect needs detailed investigation in order to optimize the contribution from the two mechanisms. The projected 10–year retention for the Flash mode is still insufficient but this inadequacy may be attributed to the charge loss by the Poole–Frenkel leakage in the ferroelectric layer and trap–assisted leakage across the 4 nm tunnel oxide. Electron retention can be improved by increasing t_{TOX} at the expense of higher V_{PROG}.

6.6 Device Prospects and Scaling Limitations

With the ability to program at two independent speeds, the hybrid memory can be a viable contender for the storage–class memory and for specific applications such as instance check–pointing and context switching in multithreaded operations. Due to the moderate operating voltages and low peak current requirements, these hybrid device systems can be designed using NAND or NOR architectures to enable massively parallel P/E operations. However, sub–100 nm scaling of the hybrid device is currently limited by the thickness of the ferroelectric film. Due to high dielectric constants, ΔV_{TH} scales rapidly with t_{FE}, and the memory window practically vanishes for thicknesses below 100 nm. Secondly, off–axis coupling between adjacent memory cells can also be significantly higher than floating gate structures. Such coupling causes read and program disturb issues that can ultimately limit scaling. Experimental measurements of off–axis coupling may provide useful insight in calibrating dielectric tensors for *P–E* hysteresis. Detailed 3–D electrostatic simulations of P/E and retention conditions will be necessary to identify disturb mechanisms and their impact on scaling. However, researchers have recently reported the possibility of ferroelectric

HfO_2 thin films by intentionally incorporating minute quantities of Si, Zr or several other elemental substitutions [34–36]. Further, they are shown to demonstrate reasonable P_r (10 – 15 μC/cm^2) with low dielectric constants (ε_{FE} ~ 20 – 40). Integration of such materials could in principle extend the hybrid device scaling below 20 nm at practical aspect ratios.

6.7 Conclusion

This chapter has introduced a novel dual–switching–speed hybrid memory design that benefits from two independent memory mechanisms, namely, ferroelectric polarization hysteresis and gate charge–injection. By integration of fast–switching ferroelectrics, the program operation is shown to naturally proceed in two stages, enhanced initial ferroelectric domain switching followed by voltage transfer to the tunneling dielectric. This unique mechanism was exploited to design the DRAM–Flash hybrid device that exhibits sub–100 ns polarization switching in the DRAM mode and 1 ms gate injection of electrons in the Flash mode. Devices fabricated with PZT as the ferroelectric and Au NCs as the storage nodes with high–κ integration demonstrated sub–10 V P/E operation. The DRAM mode operation achieved ΔV_{TH} > 0.1 V for V_{PROG} = – 8 V and t_{PROG} = 100 ns, while the Flash mode operation displayed a larger memory window and much longer retention. Program and retention characteristics of the two modes were thoroughly analyzed to provide a clear distinction between the two memory mechanisms.

REFERENCES

[1] R. F. Freitas and W. W. Wilcke, "Storage–class memory: The next storage system technology," *IBM J. Res. Dev*. vol. 52, no. 4/5, pp. 439–447 July/Sept. 2008.

[2] S. Tehrani, B. Engel, J. M. Slaughter, E. Chen, M. DeHerrera, M. Durlam, P. Naji, R. Whig, J. Janesky and J. Calder, "Recent developments in magnetic tunnel junction MRAM," *IEEE Tran. Mag.*,vol. 36, no. 5, pp. 2752–2757, Sep. 2000.

[3] D. Kau, S. Tang, I. V. Karpov, R. Dodge, B. Klehn, J. A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, S. Lee, T. Langtry, K. Chang, C. Papagianni, J. Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro and G. Spadini, "A stackable cross point phase change memory," in *IEDM Tech. Dig.*, 2009, pp. 27.1.1–27.1.4.

[4] F. Pellizzer, A. Benvenuti, B. Gleixner, Y. Kim, B. Johnson, M. Magistretti, T. Marangon, A. Pirovano, R. Bez and G. Atwood, "A 90–nm phase change memory technology for stand–alone non–volatile memory applications," *VLSI Tech. Dig.*, 2006, pp. 122–123.

[5] B. Govoreanu, G. S. Kar, Y. Chen, V. Paraschiv, S. Kubicek, A. Fantini, P. I. Radu, L. Goux, S. Clima, R. Degraeve, N. Jassart, O. Richard, T. Vandeweyer, K. Seo, P. Hendrickx, G. Pourtois, H. Bender, L. Altimime, D. J. Wouters, J. A. Kittl and M. Jurczak, "10×10 nm^2 Hf/HfO_x crossbar resistive RAM with excellent performance, reliability and low–energy operation," in *IEDM Tech.*

Dig., 2011, pp. 31.6.1–31.6.4.

[6] H. Akinaga and H. Shima, "Resistive random access memory (ReRAM) based on metal oxides," *Proc. IEEE*, vol. 98, no. 12, pp. 2237–2251, Dec. 2010.

[7] M. H. Tang, Z. H. Sun, Y. C. Zhou, Y. Sugiyama and H. Ishiwara, "Capacitance-voltage and retention characteristics of $Pt/SrBi_2Ta_2O_9/HfO_2/Si$ structures with various buffer layer thicknesses," *Appl. Phys. Lett.*, vol. 94, p. 212907, May 2009.

[8] M. Tang, X. Xu, Z. Ye, Y. Sugiyama and H. Ishiwara, "Impact of HfTaO buffer layer on data retention characteristics of ferroelectric-gate FET for nonvolatile memory applications," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 370−375, Feb. 2011.

[9] S. Sakai, M. Takahashi, K. Takeuchi, Q. –H. Li, T. Horiuchi, S. Wang, K. –Y. Yun, M. Takamiya and T. Sakurai, "Highly scalable Fe (ferroelectric)-NAND cell with MFIS (metal-ferroelectric-insulator-semiconductor) structure for sub-10 nm terabit capacity NAND Flash memories," in *International Memory Workshop*, 2008, pp.103−105.

[10] R. R. Mehta, B. D. Silverman and J. T. Jacobs, "Depolarization fields in thin ferroelectric films," *J. Appl. Phys.*, vol. 44, no. 8, pp. 3379−3385, Aug. 1973.

[11] T. P. Ma and J. –P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," *IEEE Electron Device Lett.*, vol. 23, no. 7, pp.386−388, Jul. 2002.

[12] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part I: device concept

and modeling," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 441–449, Feb. 2012.

[13] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part II: experimental validation and analysis," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 450–458, Feb. 2012.

[14] J. Han, S. Ryu, C. Kim, S. Kim, M. Im, S. J. Choi, J. S. Kim, K. H. Kim, G. S. Lee, J. S. Oh, M. H. Song, Y. C. Park, J. W. Kim and Y. –K. Choi, "A unified–RAM (URAM) cell for multi–functioning capacitorless DRAM and NVM," in *IEDM Tech. Dig.*, 2007, pp. 929–932.

[15] A. K. Tagantsev, I. Stolitchnov, N. Setter, J. S. Cross and M. Tsukada, "Non-Kolmogorov-Avrami switching kinetics in ferroelectric thin films," *Phys. Rev. B*, vol. 66, no. 21, p. 214109, Dec. 2002.

[16] J. Y. Jo, H. S. Han, J. –G. Yoon, T. K. Song, S. –H. Kim and T. W. Noh, "Domain switching kinetics in disordered ferroelectric thin films," *Phys. Rev. Lett.*, vol. 99, 267602, Dec. 2007.

[17] K. Auluck, S. R. Rajwade and E. C. Kan, "Switching dynamics in ferroelectric–charge hybrid nonvolatile memory," in *Proc. Dev. Res. Conf.*, 2012, pp. 133–134.

[18] Z. A. Weinberg, "On tunneling in metal–oxide–silicon structures," *J. Appl. Phys.*, vol. 53, no. 7, pp. 5052–5056, Jul. 1982.

[19] T. P. Juan, C. Chang and J. Y. Lee, "A new metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (Dy_2O_3)–semiconductor (MFIS) FET for

nonvolatile memory applications," *IEEE Elec. Dev. Lett.*, vol. 27, no. 4, pp. 217-220, Apr. 2006.

[20] Y. Kaneko, Y. Nishitani, H. Tanaka, M. Ueda, Y. Kato, E. Tokumitsu and E. Fujji, "Correlated motion dynamics of electron channels and domain walls in a ferroelectric–gate thin–film transistor consisting of ZnO/Pb(Zr,Ti)O_3 stacked structure," *J. Appl. Phys.*, vol. 110, no. 8, 084106, Oct. 2011.

[21] G. Velu, D. Remiens and B. Thierry, "Ferroelectric properties of PZT thin films prepared by sputtering with stochiometric single oxide target: comparison between conventional and rapid thermal annealing," *J. Euro. Cer. Soc.*, vol. 17, pp. 1749–1755, Feb.1997.

[22] C. Lee, J. Meteer, V. Narayanan and E. C. Kan, "Self–assembly of metal nanocrystals on ultrathin oxide for nonvolatile memory applications," *J. Elec. Mat.*, vol. 34, no. 1, pp. 1–11, Jan. 2005.

[23] S. L. Miller, R. D. Nasby, J. R. Schwank, M. S. Rodgers and P. V. Dresendorfer, "Device modeling of ferroelectric capacitors," *J. Appl. Phys.*, vol. 68, no. 12, pp. 6463−6471, Dec. 1990.

[24] A. Sheikholeslami and P. G. Gulak, "Transient modeling of ferroelectric capacitors for nonvolatile memories," *IEEE Tran. Ultrason. Ferroelectr. Freq. Control*, vol. 43, no. 3, pp. 450–456, May 1996.

[25] P. Juan, Y. Hu, F. Chiu and J. Y. Lee, "The charge trapping effect of metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (HfO_2)–silicon capacitors," *J. Appl. Phys.*, vol. 98, 044103, Aug. 2005.

[26] P. Juan, Y. Hu, F. Chiu and J. Y. Lee, "The electrical properties of metal–

ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator–silicon (MFIS) capacitors with different insulating materials," *J. Microelec. Engg.*, vol. 80, pp. 309–312, 2005.

[27] A. Chin, M. Y. Yang, C. L. Sun and S. Y. Chen, "Stack gate PZT/Al2O3 one transistor ferroelectric memory," *IEEE Elec. Dev. Lett.*, vol. 22, no. 7, pp. 336–338, July 2001.

[28] T. P. Juan, C. –L. Lin, W. –C. Shih, C. –C. Yang, J. Y. Lee, D. –C. Shye and J. –H. Lu, "Fabrication and characterization of metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (La_2O_3)–semiconductor capacitors for nonvolatile memory applications," *J. Appl. Phys.*, vol. 105, 061625, Mar. 2009.

[29] A. Gruverman, O. Auciello and H. Tokumoto, "Nanoscale investigation of fatigue effects in Pb(Zr,Ti)O3 films," *Appl. Phys. Lett.*, vol. 69, no. 21, pp. 3191–3193, Nov. 1996.

[30] M. Dawber and J. F. Scott, "A model for fatigue in ferroelectric perovskite thin films," *Appl. Phys. Lett.*, vol. 76, no. 8, pp. 1060–1062, Feb. 2000.

[31] C. V. R. Vasant Kumar, R. Pascual and M. Sayer, "Crystallization of sputtered lead zirconium titanate films by rapid thermal processing," *J. Appl. Phys,*. vol. 71, no. 2, pp. 864–874, Feb. 1992.

[32] M. Alexe, C. Harnagea, D. Hesse and U. Gosele, "Polarization imprint and size effects in mesoscopic ferroelectric structures," *Appl. Phys. Lett.*, vol. 79, no. 2, pp. 242–244, July 2001.

[33] H. M. Duiker, P. D. Beale, "Grain size effects in ferroelectric switching," *Phys. Rev. B*, vol. 41, no. 1, pp. 490–495, Jan. 1990.

[34] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Appl. Phys. Lett.*, vol. 99, 102903, Sept. 2011.

[35] J. Müller, U. Schröder, T. S. Böscke, I. Müller, U. Böttger, L. Wilde, J. Sandqvist, M. Lemberger, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectricity in yttrium–doped hafnium oxide," *J. Appl. Phys.*, vol. 110, 114113, Dec. 2011.

[36] J. Müller, T. S. Böscke, D. Bräuhaus, U. Schröder, U. Böttger, J. Sandqvist, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectric $Zr_{0.5}Hf_{0.5}O_2$ thin films for nonvolatile memory applications," *Appl. Phys. Lett.*, vol. 99, 112901, Sept. 2011.

CHAPTER 7 MODELING OF PROGRAM AND RETENTION DYNAMICS IN DUAL-SPEED FERROELECTRIC ANS CHARGE HYBRID MEMORY

7.1 Abstract

This chapter presents a physical model for program and retention transients in ferroelectric and charge hybrid nonvolatile memory. A part–by–part statistical model governing domain switching in polycrystalline ferroelectrics was incorporated with the tunneling current simulations to predict the memory window (ΔV_{TH}) evolution during program and retention operation in the hybrid memory. By using a specific case of the hybrid memory integrating thin film lead zirconium titanate (PZT) and Au nanocrystals, this model was calibrated to actual polarization switching measurements in PZT capacitors and program transients in the hybrid device. The resulting simulations validated the two–step program mechanism experimentally observed in such memories; namely, rapid initial domain switching in PZT on account of high fields in the ferroelectric followed by field enhancement in the tunneling dielectric which initiates electron injection into storage nodes. Further, these simulations were shown to accurately account for individual ΔV_{TH} from the two additive memory mechanisms at all program times. The depolarization effect was shown to be a dominant cause of ΔV_{TH} loss at short and moderate retention time scales (< 100 s). The model was also used to provide realistic estimates in achieving dual–speed DRAM–Flash nonvolatile memory operation.

7.2 Introduction

Ferroelectric (FE)–FETs have generated considerable interest in recently as possible alternatives to Flash in low–power embedded memory applications [1]-[4]. Inorganic polycrystalline thin-film ferroelectrics such as lead zirconium titanate (PZT) and $SrBi_2Ta_2O_9$ demonstrate high–speed (< 100 ns) polarization switching at low applied electric fields (100–200 kV/cm) that makes them a viable option in low–voltage nonvolatile memories [5]-[8]. Recently, researchers have also shown the possibility of ferroelectric HfO_2 with a few percent atomic substitutions by elements like Si, Zr and Yt [9]-[11]. FE–FETs made from such HfO_2 thin films have demonstrated sub–100 ns switching with less than 6 V program/erase (P/E) operation [12]. These P/E speeds, although insufficient for DRAM specifications, are attractive in 'storage class' memory applications that demand nonvolatility with ultra–scaled density and high throughput [13].

However, FE–FETs are known to exhibit limited retention on account of the randomization in the aligned polarization. Due to incomplete compensation of the ferroelectric surface charge by the free carriers in the sensing channel, the depolarization field (E_{dp}) is set up inside the ferroelectric layer that opposes the remanent polarization (P_r) and leads to eventual back–switching of domains during the retention state [14]-[15]. One possible alternative to mitigate the effect of depolarization is to store nonvolatile charge on floating nodes adjacent to the ferroelectric thin film [16]. This was demonstrated in a ferroelectric and charge hybrid memory device that utilized gate–injection of electrons on storage nodes above the

ferroelectric film during program operation. The stored electrons benefit the memory performance in two distinctive ways: additive memory window (ΔV_{TH}) to the remanent polarization and partial compensation of E_{dp} to improve the overall retention in these hybrid devices [17].

The integration of these two memory mechanisms can be further extended to exploit the short polarization response time in ferroelectrics along with the long reliable retention in Flash–like storage. The previous chapter has shown that in the case of fast switching ferroelectrics like PZT, the program operation in the hybrid device naturally evolves into a two–step mechanism [18]. At the start of the program operation, large electric field in the ferroelectric film accelerates the process of dipole alignment. As program time (t_{PROG}) increases, the dipole polarization adds to the total displacement in the ferroelectric, enhancing the electric field in the adjoining dielectrics. This phenomenon establishes a dual–speed attribute to the program transients which was experimentally demonstrated in the program characteristics of PZT–Au nanocrystal (NC) hybrid memory that showed fast DRAM–like response from the PZT as well as slow Flash–like electron tunneling to the storage nodes.

This chapter details the generalized methodology to model program and retention characteristics in ferroelectric and charge hybrid memories. We have used a statistical region–by–region switching model to describe ferroelectric polarization evolution to the applied program bias while solving it consistently with 1–D electrostatics as well as tunneling current calculations in the memory gate stack [19]-[21]. The switching model is further calibrated by actual measurements from the PZT

thin film capacitors and PZT–Au NC hybrid memory. The hybrid model is shown to verify the observed two–step program mechanism and also provide useful insights into the design for DRAM and Flash mode operation.

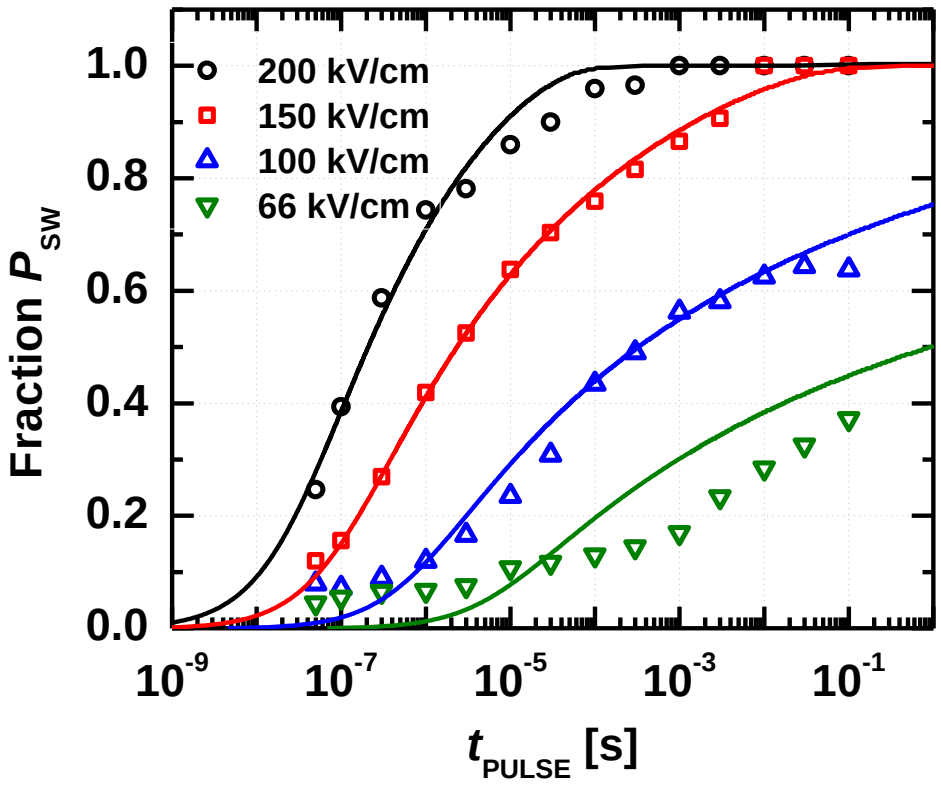


Figure 7-1 Polarization reversal data for Pt/PZT/Pt capacitors for pulse widths (t_{PULSE}) ranging from 50 ns to 100 ms at different applied fields fitted to the statistical part–by–part switching model

7.3 Hybrid Device Transient Modeling

Evolution of polarization under the program bias is critical to high speed operation in the hybrid device. It is known that the domain switching rate is extremely sensitive to the instantaneous electric field experienced by the ferroelectric thin film. Unlike the ferroelectric capacitor, for which the instantaneous field equals the applied electric field (E_{appl}) for the entire pulse duration, the field inside the ferroelectric in series with a dielectric (with comparable or higher effective oxide thickness) continuously changes along the duration of the applied voltage pulse. As discussed earlier, this field is higher at the start of the program operation and gradually diminishes as the polarization aligns in the direction of the applied bias. To capture this phenomenon and estimate the two–step program process accurately, we need to model the kinetics of ferroelectric switching consistently with the electrostatics in the memory gate stack as well as the electron tunneling flux to the storage node.

Over the last four decades, modeling of polarization switching in ferroelectric thin films has been investigated in great detail [22]-[24]. Kolmogorov–Avrami–Ishibashi model [25]-[27] is frequently used to describe domain switching in single crystal films. However, in the case of sputtered polycrystalline films like PZT, the rate of domain switching is restricted to the grain boundaries and often limited by the domain wall friction between perpendicularly aligned neighbors [28]-[29]. These effects cannot be captured in such idealized models that assume unrestricted domain growth after initial nucleation. Therefore, theories that consider the limitation of domain reversal by nucleation in polycrystalline films as well as empirical models that

include a wide spread in switching times have been investigated [30]-[31]. We have used a simple 1–D region–by–region switching model that is based on statistical distribution in switching delays for finitely many domains in the ferroelectric thin film [19]. This model is also shown to yield excellent matching with PZT polarization switching data as well as precisely predict the domain reversal to depolarization during retention condition in short and moderately long time scales [20].

For the maximum value of spontaneous polarization (P_s), the model assumes the film to be divided uniformly into M_0 parts or regions of equal polarization strength ($=P_s/M_0$). The model starts with all regions M_0 aligned opposite to the applied field. Then, assuming that N parts have already switched in the direction of the field (where N = 0, 1, 2,..., M_0–1), time of switching for the $(N+1)^{th}$ domain (t_{N+1}) is evaluated solving the following Eq. (1),

$$\frac{M_0 - N - 1}{M_0 - N} = \exp\left(-\frac{t_{N+1}}{t_\infty} \exp\left(-\frac{\alpha}{E_{FE}} \right) \right) \tag{1}$$

where t_∞ is the infinite-field switching time, α is the activation field for switching and E_{FE} is the total electric field in the ferroelectric. Deposition of the ferroelectric film on a metal or dielectric often leads to formation of a low–k dead layer that modifies the applied electric field (E_{appl}). If d_{DL} and d_{FE} are the effective oxide thickness (EOT) of the dead layer and physical thickness of the ferroelectric, ε_0 is the permittivity of free space and ε_{ox} is the dielectric constant of SiO_2,

$$E_{FE} = E_{appl} + \frac{d_{DL} P_{SW}}{d_{FE} \varepsilon_{ox} \varepsilon_0} \tag{2}$$

When N domains have switched, the net switched polarization ($P_{SW} = P_{M_0 - N}$) is given by,

$$P_{M_0 - N} = \frac{M_0 - 2N}{M_0} P_s \tag{3}$$

t_∞, d_{DL} and α can be extracted from the polarization switching data of a ferroelectric capacitor. These parameters can then be used in modeling switching dynamics of the hybrid device.

At the start of the program operation, we assume that the initial ferroelectric domain configuration is completely randomized (i.e. $N = M_0/2$). Then, for any instantaneous N, known d_{DL}, fixed applied program voltage (V_{PROG}) and given geometry, electrostatics in the entire gate stack (including E_{FE}) is solved and $t_{\mathrm{N+1}}$ may be uniquely determined from Eq. (1). The instantaneous electron tunneling current from the gate electrode is calculated based on the electric field in the tunnel oxide by the Tsu–Esaki tunneling model [21]. The switching kinetics in ferroelectrics as well as electronic charge injected into the storage nodes are then evolved for this discrete time step of $t_{\mathrm{N+1}}$. At the end of $t_{\mathrm{N+1}}$, P_{SW} is updated, electrostatic values are recalculated and the method is used iteratively.

The total accumulated t_{PROG} in the hybrid device (or equivalently the switching time up to N+1 domains in a ferroelectric film) is then given by,

$$t_{PROG} = \sum_{i}^{N+1} t_i \tag{4}$$

Nonlinear capacitance of the ferroelectric film to the applied field gives rise to gate voltage dependence to the total EOT of the gate stack. This effect is also incorporated in the simulation based on the experimental results from the last chapter to give a realistic estimate of threshold voltage shift (ΔV_{TH}) under read disturb. Flat band voltage at $P_{SW} = 0$ and no stored charge is assumed to be 0 V. Electron tunneling current out of the storage layer during program operation is not considered and charge centroid movement during all memory operations is ignored. Simulation of retention dynamics follows the same methodology under short circuit condition ($V_{PROG} = 0$ V).

7.4 Simulation Results

7.4.1. Ferroelectric Parameter Extraction

Fig. 7-1 shows the polarization switching data for a Pt/PZT/Pt capacitor measured by a modified PUND method [18], [32]. The figure also depicts the switching model fitted to all applied fields. The fraction of domains switched is determined from an initial condition where all the polarization dipoles are opposing the applied electric field. In other words, P_{SW} is normalized to $2P_S$ (= 32 μC/cm^2 here). The extracted parameters of t_∞ = 140 ps, α = 1.7 MV/cm and d_{DL} = 0.2 nm are seen to be consistent with values reported previously [19], [20], [33]. For high applied fields (> 100 kV/cm), the model shows excellent matching to the actual measurements. Such fields are experienced by the ferroelectric at the start of program/erase (P/E) operation as well as during short (< 10 s) retention time scales in

the hybrid device.

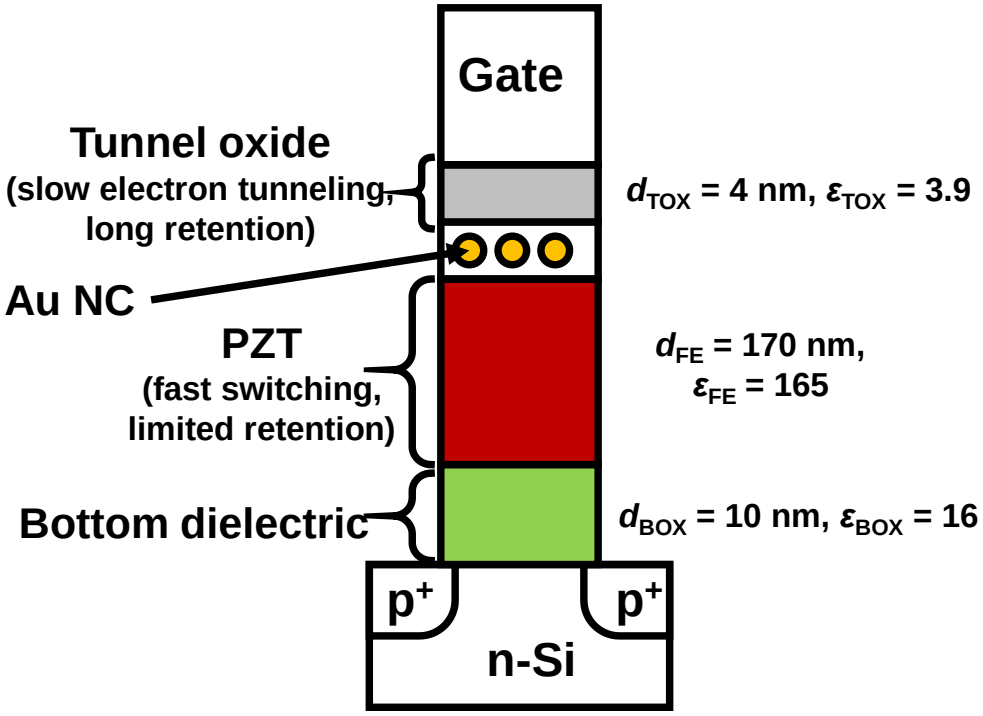


Figure 7-2 Device schematic of dual–speed ferroelectric and charge hybrid nonvolatile memory including the geometry parameters used for simulations

However, for low electric fields (< 70 kV/cm), which are important for depolarization assisted domain reversal at long retention time scales (> 100 s), the model is seen to deviate from the switching data. Polarization back–switching under such low fields is often dominated by other slow processes like internal charge compensation in the bulk and hole leakage from the channel. Such effects have not been considered in this model which leads to deviation from the observed data [20].

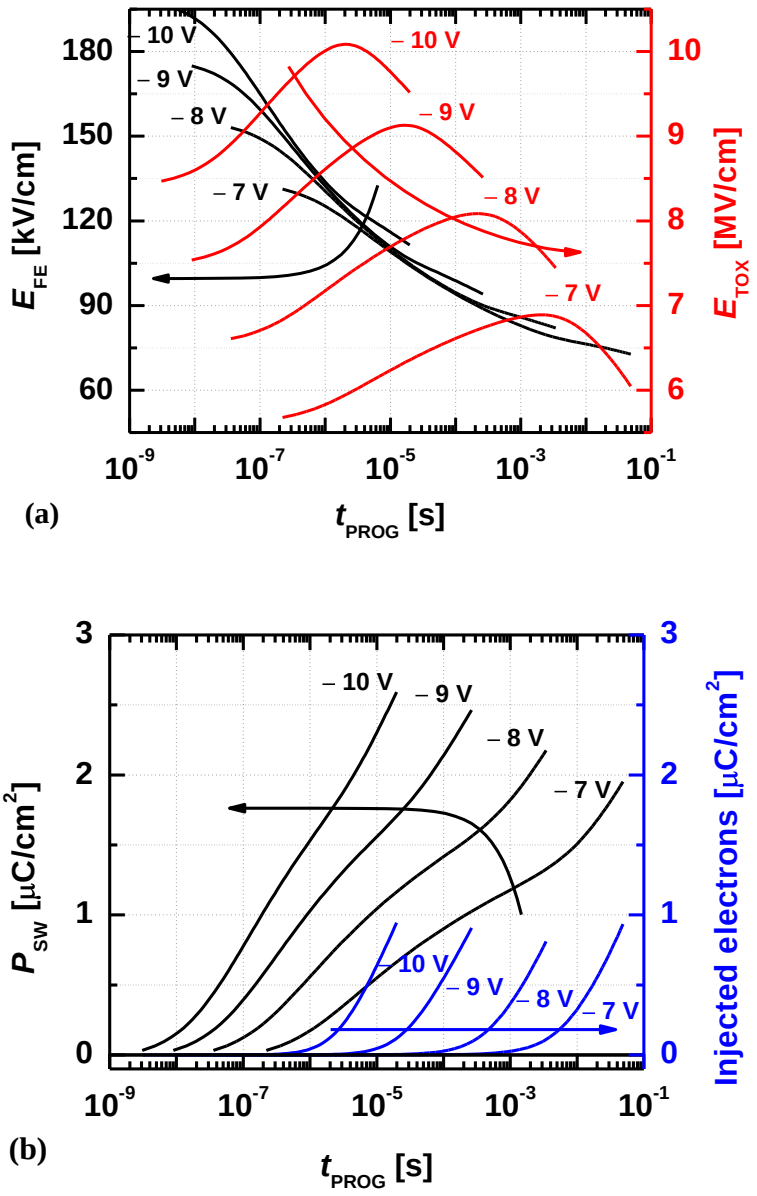


Figure 7-3 (a) Evolution of electric field in the ferroelectric (E_{FE}) and tunnel oxide (E_{TOX}) as a function of t_{PROG} for V_{PROG} = –7 to –10 V. The two–step program process is evident from the transition of peak fields from E_{FE} to E_{TOX} with increasing t_{PROG}. (b) Switched polarization in PZT (P_{SW}) and density on the captured electrons on the storage node as a function of t_{PROG}.

As stated previously, hybrid memory program transients were simulated with the ferroelectric in a completely randomized state. This corresponds to P_{SW} fraction equal to 0.5 at the start of program pulse. Further, ferroelectric film in a FE–FET or hybrid memory gate stack cannot attain complete polarization while still maintaining adjoining oxide reliability. For example, magnitude of P_{SW} above 3 μC/cm^2 generates an electric field over 1 V/nm in the adjoining SiO_2 layer that is often detrimental to its integrity over repeated cycling. Therefore, P_{SW} fraction in hybrid memory P/E transients is restricted to the region of ~ 0.5 ± 0.1 in Fig. 7-1.

7.4.2. Program Transients in Hybrid Device

We have simulated the program characteristics for the dual–speed PZT–Au NC hybrid device fabricated in the last chapter. The bottom dielectric and tunnel oxide comprised of ALD HfO_2 and SiO_2 respectively. Fig. 7-2 shows the device schematic with the geometry parameters. d_{TOX}, d_{FE} and d_{BOX} are the thicknesses and ε_{TOX}, ε_{FE} and ε_{BOX} are the dielectric constants of the tunnel oxide, ferroelectric and the bottom dielectric, respectively. Tsu–Esaki tunneling model parameters of electron effective mass (m_{eff}) and gate emission barrier height (φ_b) were calibrated to program transients of gate–inject Flash devices fabricated in the same process flow.

Fig. 7-3 (a) depicts the evolution of the electric fields in PZT (E_{FE}) and the tunnel oxide (E_{TOX}) with t_{PROG} for V_{PROG} ranging from –7 V to –10 V. As shown, E_{FE} decreases linearly with every decade increase in t_{PROG}. For V_{PROG} = –10 V, E_{FE} can be above 200 kV/cm at the start of the program operation (when P_{SW} = 0 μC/cm^2). Polarization switching rate ($=\partial P_{SW}/\partial t$) at such fields can be as high as 1 C/cm^2–s

(observed from Fig. 7-1), which assists rapid alignment of domains in the direction of the applied bias. However, E_{FE} reduces linearly with increase in P_{SW}, and this process gradually begins to decelerate as P_{SW} becomes comparable to the total displacement (D_{FE}) in the ferroelectric film ($D_{FE} = \varepsilon_0 \varepsilon_{FE} E_{FE} + P_{SW}$).

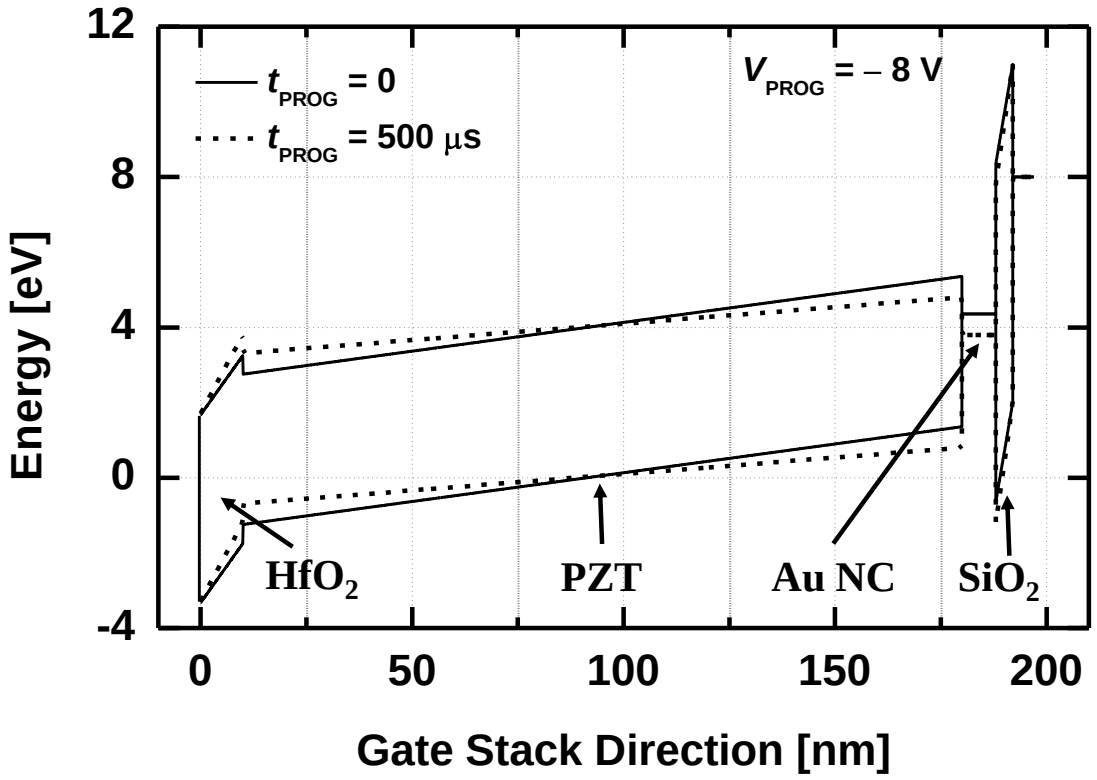


Figure 7-4 Band diagram for the hybrid gate stack at the beginning of the program operation and after its transition into Flash mode

This initial boost in E_{FE} is crucial to DRAM–like switching for the hybrid device. It can be shown that such effect is more prominent at lower ε_{FE} and smaller d_{FE}. The transient evolution of P_{SW} in t_{PROG} is depicted in Fig. 7-3 (b) for various V_{PROG}. As P_{SW} increases, the polarization switching rate diminishes with increasing t_{PROG}. Further, this gradual saturation in P_{SW} is seen to be more prominent for lower program voltages ($|V_{PROG}| < 9$ V).

E_{TOX} is significantly enhanced when P_{SW} becomes higher than ~1.5 $\mu C/cm^2$ and the hybrid device subsequently enters the tunneling regime. As seen from Fig. 7-3 (a), improvement in E_{TOX} can be 20 – 30 % over the initial value and exhibits a strong dependence on V_{PROG}. The transition from the DRAM to the Flash mode occurs when the gate–injected electrons begin to influence the overall electrostatics in the memory gate stack. This may also be characteristically distinguished by the maxima attained in the evolution of E_{TOX}.

As electrons begin to accumulate in the storage nodes, they repel the incoming electrons injected from the gate by reducing E_{TOX}. The stored electrons however, have an opposite influence on E_{FE}. They diminish the contribution of effective surface polarization to the total electric displacement at the ferroelectric interface. This leads to a 'secondary' domain reversal process in the ferroelectric at considerably diminished E_{FE} and exhibits switching time scales comparable to electron tunneling. This switching is primarily aided by the accumulation of stored electrons and therefore depicts a weaker dependence on E_{FE}. It also means that in an ideal scenario, with negligible hole tunneling from the substrate, this process may be practically sustained

over the remainder of the program time. Such phenomenon is more evident in low program voltages, as observed in Fig. 7-3 (b).

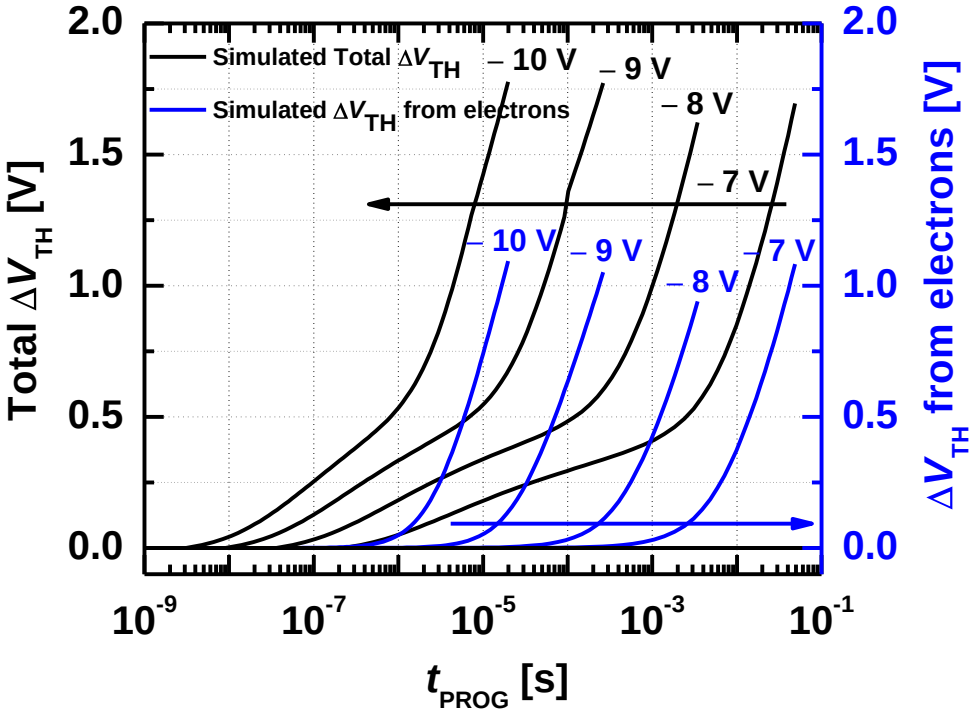


Figure 7-5 Total ΔV_{TH} transient with contribution from injected electrons evolving with t_{PROG}. In spite of a lower surface charge density, electrons show higher contribution to the total ΔV_{TH} compared to ferroelectric polarization.

The injected electronic surface charge density can become comparable to the polarization surface charge for longer t_{PROG}. This is seen for $V_{PROG} = -7$ V and $t_{PROG} >$ 10 ms, when injected charge density is ~ 50 % of the ferroelectric polarization. Fig. 7-4 depicts the band diagram in the vertical direction at the beginning of program

operation (V_{PROG} = –8 V) and after its transition into the Flash mode (t_{PROG} = 0.5 ms) when P_{SW} accumulates above 1.5 μC/cm^2. E_{FE} is observed to diminish by ~30 % and E_{TOX} is enhanced over 25 % in the Flash mode. The electric field in the bottom dielectric (E_{BOX}) also increases by the same proportion.

Fig. 7-5 shows the predicted total ΔV_{TH} as well as the contribution from injected electrons. The two–step mechanism is clearly discernible in fast responding domain switching and slower electron tunneling. The contribution from stored electrons to the total ΔV_{TH} is only significant after P_{SW} rises above 1.5 μC/cm^2. It is possible to achieve ΔV_{TH} > 0.2 V with less than 100 ns program pulses in DRAM mode as well as ΔV_{TH} > 1 V with long retention in Flash mode with sub–10 V P/E operation. ΔV_{TH} from electrons in the Flash mode can contribute towards more than 50 % to the total ΔV_{TH} of the device in spite of electron surface charge density being half of the ferroelectric polarization.

It should however be noted that, as the Flash mode t_{PROG} advances, both ferroelectric polarization and stored electrons generate massive enhancement in E_{BOX}. The highest ΔV_{TH} that the hybrid device can reliably display would depend on the maximum reliable E_{BOX} that the bottom dielectric can support before substantial hole injection from the channel negates the intentional ΔV_{TH}. These effects have not been included in the simulations for the purpose of simplicity. For the current device, the highest observed ΔV_{TH} without channel hole–injection was ~0.8 V [Rajwade13]. This may however be improved by engineering the optimal bottom dielectric material and d_{BOX} [34]-[35].

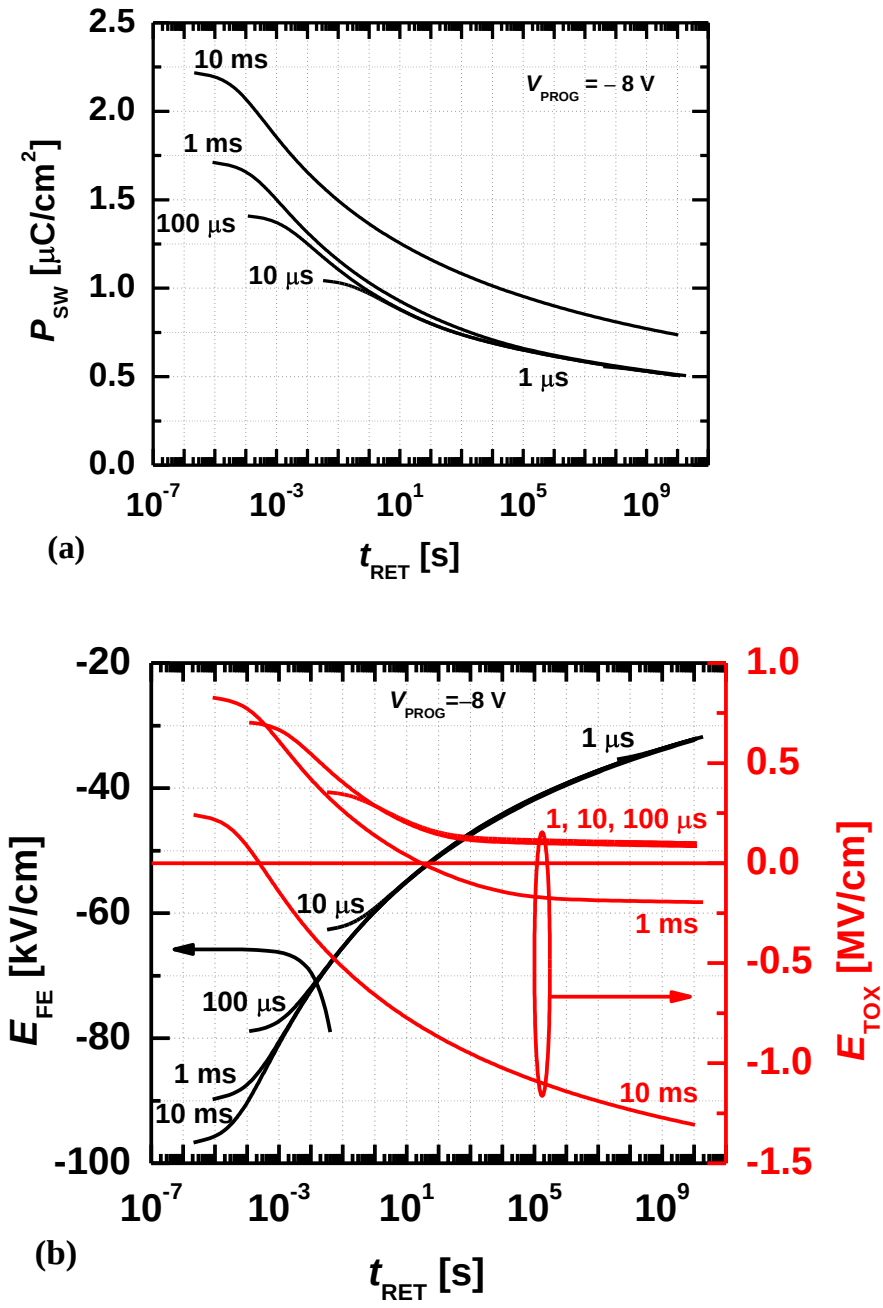


Figure 7-6 (a) Evolution of P_{SW} under retention condition in the hybrid device for various t_{PROG}. (b) Evolution of E_{FE} and E_{TOX} as a function of t_{RET} for varying t_{PROG}. When sufficient electrons are stored at the floating nodes (t_{PROG} = 1 ms ans 10 ms), E_{TOX} undergoes polarity as t_{RET} increases as a result of reducing residual P_{SW}.

7.4.3. Dipole Relaxation Effect

Dipole relaxation practically follows the same switching behavior as dipole alignment according to the description in the statistical model. Initial back–switching in domains is rapid aided by high E_{dp} generated after program operation. This process determines the retention time ($t_{RET,DRAM}$) in the DRAM mode. Fig. 7-6 (a) depicts the retention characteristics of P_{SW} following the program operation with varying t_{PROG}.

Dipole randomization process is faster for thicker adjoining dielectrics and ferroelectrics with smaller activation field. Secondly, with the exception of their initial separation immediately after program, all P_{SW} retention curves practically merge into one another. This can be explained from Eq. (4). For every increasing step in t_{PROG}, and consequent increase in P_{SW}, the residual P_{SW} retention time increases by the back–switching time for that incrementally switched domain. This suggests that every additional t_{PROG} spent on increasing P_{SW} offers diminishing returns in the retention state.

However, with sufficient contribution from the injected electrons, P_{SW} retention characteristics for the Flash mode display a lateral right shift in t_{RET} from those for the DRAM mode (see Fig. 7-6 (a) P_{SW} plot for t_{PROG} = 10 ms). For the same magnitude of residual P_{SW}, the ferroelectric experiences smaller E_{dp} with stored electrons, which increases the back–switching time in the Flash mode.

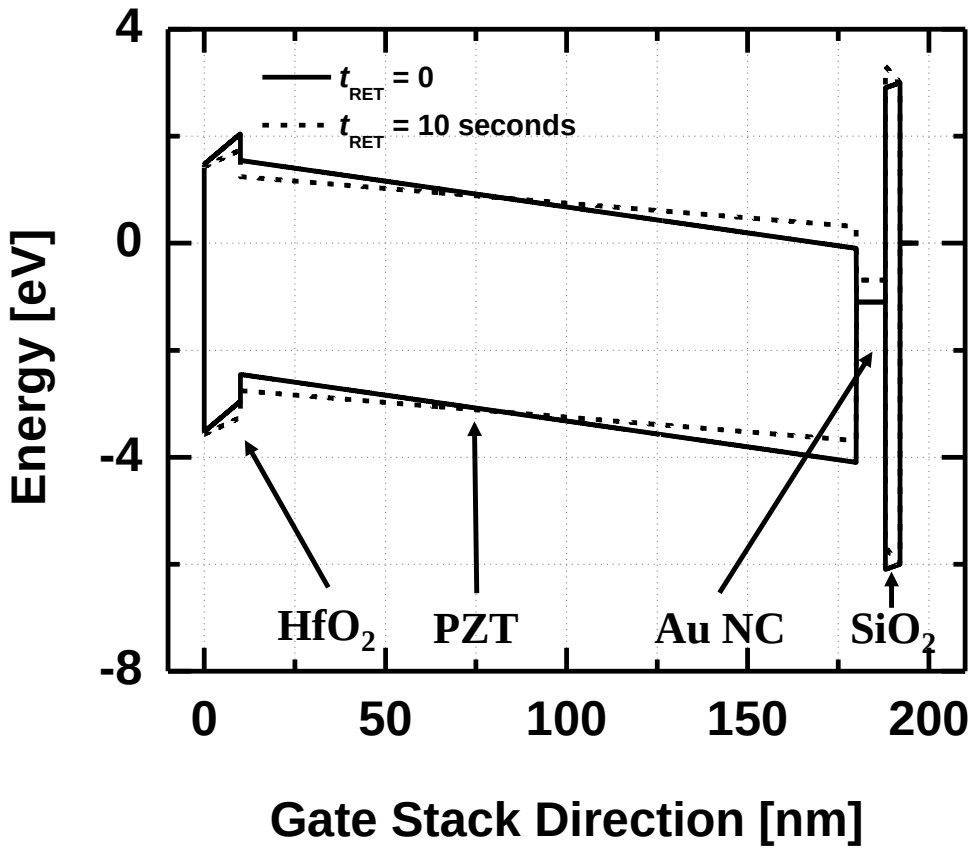


Figure 7-7 Band diagram for the hybrid gate stack in retention condition at t_{RET} = 0 seconds and 10 seconds, showing ~ 50 % reduction in E_{FE} and reversal of polarity in E_{TOX}. The device is previously programmed at V_{PROG} = –8 V for 10 ms

Fig. 7-6 (b) shows the evolution of E_{FE} and E_{TOX} with t_{RET} after different t_{PROG}. All E_{dp} plots eventually merge into one another, as expected from the model. It should be noticed that retention E_{FE} immediately after a long program operation (t_{PROG} > 1 ms) is ~50 % of the E_{FE} at the beginning of the program operation and can be even higher than that at the end of program pulse. Let us compare this case to the E_{TOX} in conventional Flash memory. In Flash, the retention E_{TOX} is an order of magnitude lower than E_{TOX} during program operation This is the primary reason why Flash

demonstrates t_{RET}/t_{PROG} ratios in excess of 10^{12}. For the hybrid device, the asymmetry between E_{FE} at $t_{PROG} = 0^+$ and $t_{RET} = 0^+$ is inherently low (~ 2) compared to the same asymmetry for E_{TOX} in Flash memory (~10). This diminished field asymmetry (high E_{FE} at $t_{RET} = 0^+$) justifies the accelerated randomization of aligned polarization and limits the t_{RET}/t_{PROG} ratio in the DRAM mode.

Fig. 7-6 (b) also shows that E_{TOX} in the hybrid device aids electron injection from the gate during retention state in the DRAM mode. However, E_{TOX} undergoes interesting dynamics in the Flash mode retention characteristics. Immediately after program, due to high P_{SW}, E_{TOX} still opposes the escape of stored electrons to the gate, as shown in Fig. 7-6 (b). This confirms that ΔV_{TH} loss in short retention times for Flash mode strictly arises from back–switching domains. With gradual reduction in residual P_{SW}, E_{TOX} is observed to change polarity. Apart from the geometry parameters, the E_{TOX} polarity–reversal–t_{RET} depends on the total P_{SW} and stored electron density. Fig. 7-7 displays the band diagram of the hybrid device in retention condition immediately after program and after substantial loss in residual P_{SW} that generates E_{TOX} polarity reversal. The device is initially programmed in the Flash mode by $V_{PROG} = -8$ V for 10 ms.

Depolarization and read disturb have a vital consequence on the measured memory window in ferroelectric devices. Every measurement setup constitutes an inherent delay between program and read operation that should be accounted in making accurate predictions on ΔV_{TH}. Measurements on the hybrid devices were

performed by Keithley SCS 4200 source–measure unit and this delay was ~ 1 second for the program–read setup [18].

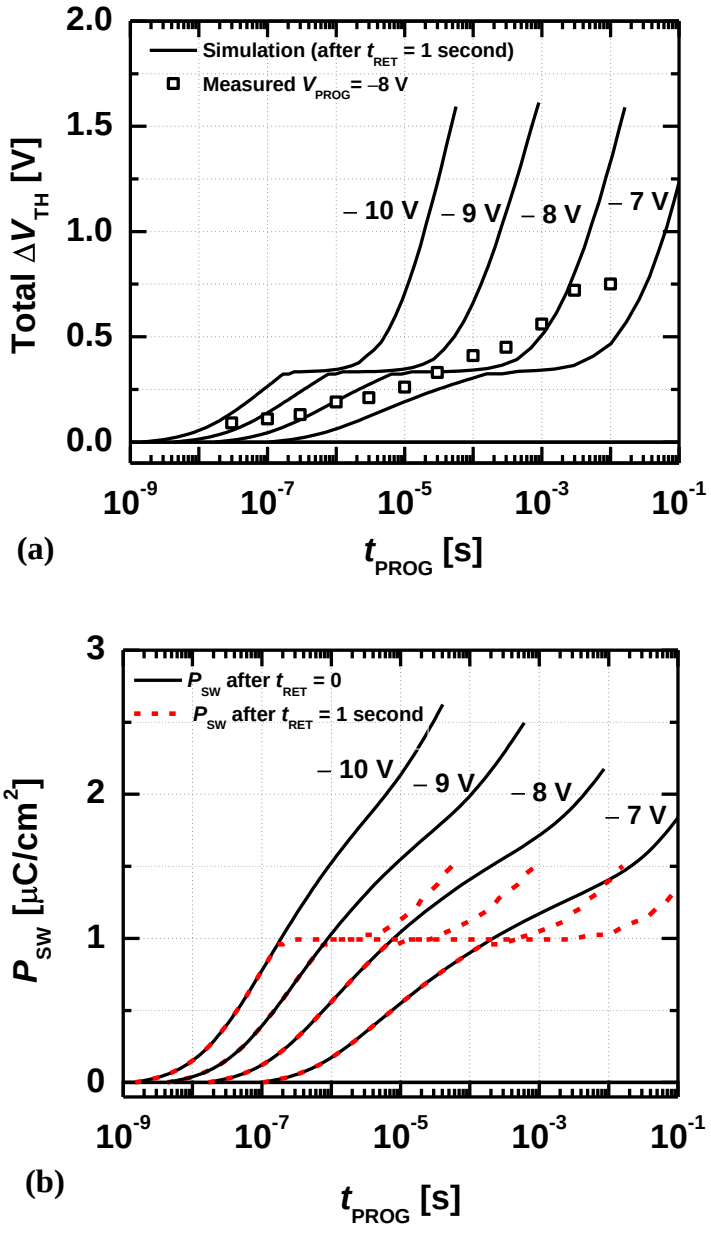


Figure 7-8 (a) Total ΔV_{TH} retained after 1 second retention plotted as a function of t_{PROG}. The measured data for the hybrid device at V_{PROG} = –8 V is also shown for comparison. (b) Difference between P_{SW} immediately after program and residual P_{SW} retained after t_{RET} = 1 second

Fig. 7-8 (a) shows the simulated total ΔV_{TH} after 1 second dipole relaxation in reasonable agreement with the measured results (for V_{PROG} = –8 V) on the hybrid device. Fig. 7-8 (b) draws a comparison between P_{SW} immediately after program (t_{RET} = 0) and residual P_{SW} after 1 second as a function of t_{PROG}. For the present geometry, residual P_{SW} in DRAM mode is seen to saturate at 1 μC/cm^2 for 1 s relaxation time. This behavior is also reflected in total ΔV_{TH} saturation at ~0.3 V in Fig. 7-8 (a). For the specific case of V_{PROG} = –8 V, the value of P_{SW} = 1 μC/cm^2 is obtained after t_{PROG} = 9 μs . As discussed previously, any further increase in t_{PROG} may yield no improvement in retained P_{SW} after 1 second retention. In practical applications of the hybrid memory in DRAM mode, device geometry and material parameters need to be optimized to yield sufficient residual P_{SW} at the end of $t_{RET,DRAM}$.

As t_{PROG} advances, there is a subsequent rise in this residual P_{SW} when the device transitions into the Flash mode. The injected electrons favor the retention of an increased fraction of the aligned dipoles for any t_{RET}. This transition occurs at t_{PROG} = 0.1 ms for V_{PROG} = –8 V. The simulation results for the Flash mode provide reasonable concurrence to measured ΔV_{TH} in moderate t_{PROG} (30 μs – 3 ms). However, this matching is seen to deviate for t_{PROG} > 3 ms. The small PZT conduction band barrier height observed from the metal NC Fermi level likely promotes high charge loss and early saturation in electron storage on Au NCs during program operation. The present simulation does not consider this electron–out–tunneling effect and therefore predicts a continued rise in the total ΔV_{TH}.

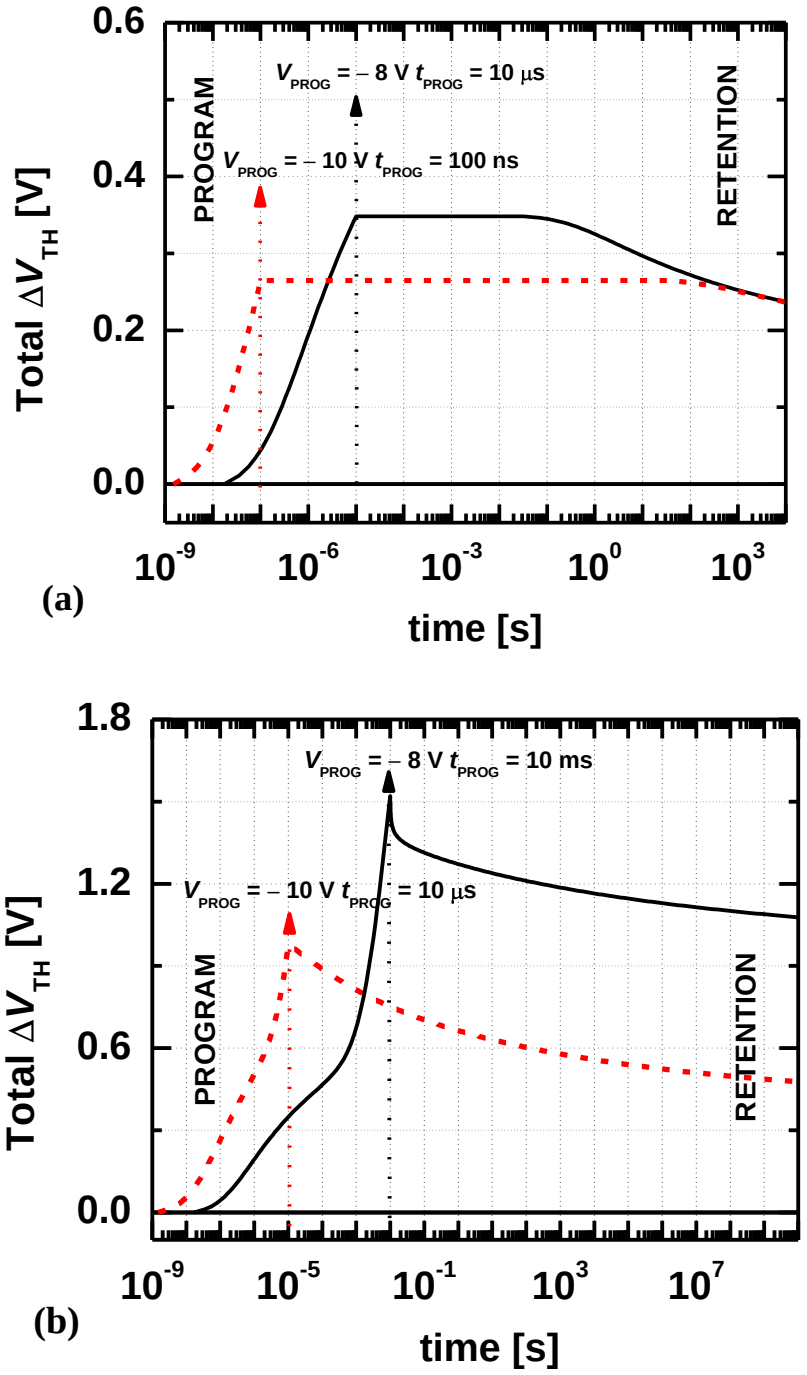


Figure 7-9 (a) Simulation of ΔV_{TH} transient during program and subsequent retention in the DRAM mode for two different V_{PROG} and t_{PROG}. (b) ΔV_{TH} transient during program and retention in the Flash mode for the same values of V_{PROG}.

Fig. 7-9 (a) depicts the ΔV_{TH} transient for program and subsequent retention condition in the DRAM mode. A short 100 ns program pulse at V_{PROG} = –10 V is shown to generate ΔV_{TH} > 0.25 V. As the resulting P_{SW} is less than 1 μC/cm^2, there is negligible polarization lost in the first 10^2 seconds of retention. For a longer program operation at V_{PROG} = –8 V, the graph depicts considerable P_{SW} loss in moderate retention time scales (0.1 – 10^2 seconds).

Program and retention transients in the Flash mode for the same values of V_{PROG} are depicted in Fig. 7-9 (b). For continued program at V_{PROG} = –8 V, secondary switching in P_{SW} assisted by electron storage is visible. The sudden dip in the total ΔV_{TH} at the beginning of retention transient is the effect of accelerated dipole back–switching induced by high E_{dp}. As seen from the figure, this initial relaxation time scales (~10–100 μs) can be shorter than those for electron tunneling and secondary dipole switching.

As discussed previously, ferroelectric switching for E_{FE} below 70 kV/cm cannot be fitted accurately to the simplified switching model. Further, the retention simulations only capture the electron leakage from Au NCs by direct tunneling to the gate. In practice, other leakage mechanisms like inter–NC charge sharing and trap–assisted tunneling from the tunnel oxide ultimately limit the retention time in discrete NC–based memories [36]. All these effects need to be modeled with careful calibration to precisely predict the retention dynamics for the Flash mode in longer time scales (10^3–10^8 seconds). Lastly, the simulations presented in this chapter, especially those on high–speed program and short–time retention transients in the

DRAM mode, are primarily dependent on the switching characteristics of the ferroelectric film. These material properties are often strong functions of the film thickness, elemental composition, and the deposition and anneal conditions [36]-[37]. Therefore, optimization of these material parameters is essential to determine an ideal design space for the hybrid device under DRAM and Flash mode operational specifications.

7.5 Conclusion

This chapter has presented a generalized method to simulate program and retention characteristics in ferroelectric and charge hybrid memories. The approach solves for the evolution of ferroelectric polarization in continually varying applied field by a statistical switching model and couples these dynamics with the electrostatics and tunneling current calculations in the hybrid gate stack. Polarization switching data from PZT capacitors was used to extract the model parameters and the resulting predictions on PZT–Au NC hybrid memory were shown to be in reasonable agreement to the measured results. The model also verified the dual–speed program process and the factors influencing this phenomenon. Effect of depolarization in short time scales was shown to be crucial for retention in the DRAM mode.

REFERENCES

[1] S. Mathews, R. Ramesh, T. Venkatesan and J. Benedetto, "Ferroelectric field effect transistor based on epitaxial perovskite heterostructures," *Science,* vol. 266, no. 5310, pp. 238–240, Apr. 1997.

[2] T. P. Juan, C. Chang and J. Y. Lee, "A new metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (Dy_2O_3)–semiconductor (MFIS) FET for nonvolatile memory applications," *IEEE Elec. Dev. Lett.*, vol. 27, no. 4, pp. 217-220, Apr. 2006.

[3] Y. Watanabe, M. Tanamura and Y. Matsumoto, "memory retention and switching speed of ferroelectric field effect in (Pb,La)(Ti,Zr)O3/La2CuO4:Sr hetereostructure," *Jpn. J. Appl. Phys*., vol. 35, pp. 1564–1568, Feb. 1996.

[4] G. A. Salvatore, D. Bouvet, I. Stolitchnov, N. Setter and A. N. Ionescu, "Low voltage ferroelectric FET with sub-100 nm copolymer P(VDF-TrFE) gate dielectric for nonvolatile 1T memory", in *Proc. Euro. Solid State Dev. Res. Conf.,* pp. 162-165, Sept. 2008.

[5] R. Moazzami, C. Hu and W. H. Shepherd, "Electrical characteristics of ferroelectric PZT thin films for DRAM applications," *IEEE Tran, Elec. Dev.*, vol. 39, no. 9, pp. 2044–2049, Sept. 1992.

[6] S. Sakai, M. Takahashi, K. Takeuchi, Q. –H. Li, T. Horiuchi, S. Wang, K. –Y. Yun, M. Takamiya and T. Sakurai, "Highly scalable Fe (ferroelectric)-NAND cell with MFIS (metal-ferroelectric-insulator-semiconductor) structure for for sub-10 nm terabit capacity NAND Flash memories" International Mem.

Workshop pp.103 (2008).

[7] M. H. Tang, Z. H. Sun, Y. C. Zhou, Y. Sugiyama and H. Ishiwara, "Capacitance-voltage and retention characteristics of Pt/$SrBi_2Ta_2O_9$/HfO_2/Si structures with various buffer layer thicknesses," *Appl. Phys. Lett.*, vol. 94, p. 212907, May 2009.

[8] M. Tang, X. Xu, Z. Ye, Y. Sugiyama and H. Ishiwara, "Impact of HfTaO buffer layer on data retention characteristics of ferroelectric-gate FET for nonvolatile memory applications," *IEEE Trans. Electron Devices*, vol. 58, no. 2, pp. 370−375, Feb. 2011.

[9] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Appl. Phys. Lett.*, vol. 99, 102903, Sept. 2011.

[10] J. Müller, U. Schröder, T. S. Böscke, I. Müller, U. Böttger, L. Wilde, J. Sandqvist, M. Lemberger, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectricity in yttrium–doped hafnium oxide," *J. Appl. Phys.*, vol. 110, 114113, Dec. 2011.

[11] J. Müller, T. S. Böscke, D. Bräuhaus, U. Schröder, U. Böttger, J. Sandqvist, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectric $Zr_{0.5}Hf_{0.5}O_2$ thin films for nonvolatile memory applications," *Appl. Phys. Lett.*, vol. 99, 112901, Sept. 2011.

[12] J. Müller, T. S. Böscke, U. Schröder, R. Hoffmann, T. Mikolajick and L. Frey, "Nanosecond polarization switching and long retention in a novel MFIS–FET based on ferroelectric HfO_2," *IEEE Elec. Dev. Lett.*, vol. 33, no. 2, pp. 185–

187, Feb. 2012.

[13] R. F. Freitas and W. W. Wilcke, "Storage–class memory: The next storage system technology," *IBM J. Res. Dev*. vol. 52, no. 4/5, pp. 439–447 July/Sept. 2008.

[14] R. R. Mehta, B. D. Silverman and J. T. Jacobs, "Depolarization fields in thin ferroelectric films," *J. Appl. Phys.*, vol. 44, no. 8, pp. 3379−3385, Aug. 1973.

[15] T. P. Ma and J. –P. Han, "Why is nonvolatile ferroelectric memory field-effect transistor still elusive?," *IEEE Electron Device Lett*., vol. 23, no. 7, pp.386−388, Jul. 2002.

[16] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part I: device concept and modeling," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 441–449, Feb. 2012.

[17] S. R. Rajwade, K. Auluck, J. B. Phelps, K. G. Lyon, J. T. Shaw and E. C. Kan, "A ferroelectric and charge hybrid nonvolatile memory—Part II: experimental validation and analysis," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 450–458, Feb. 2012.

[18] S. R. Rajwade, T. A. Naoi, K. Auluck, K. Jayant, R. B. van Dover and E. C. Kan, "Ferroelectric–assisted dual switching speed DRAM–Flash hybrid nonvolatile memory," *under review IEEE Tran. Elec. Dev.*

[19] X. J. Lou, "Statistical switching kinetics of ferroelectrics," *J. Phys.: Condens. Matter*, vol. 21, 012207, 2009.

[20] X. J. Lou, "Polarization retention on short, intermediate, and long time scales

in ferroelectric thin films," *J. Appl. Phys.*, vol. 105 094107, May 2009.

[21] R. Tsu and L. Esaki, "Tunneling in a finite superlattice," *Appl. Phys. Lett.*, vol. 22 no. 11, pp. 562–564, June 1973.

[22] W. J. Merz, "Domain formation and domain wall motions in ferroelectric $BaTiO_3$ single crystals," *Phys. Rev.*, vol. 95, no. 3, pp. 690–698, Aug. 1954.

[23] J. F. Scott, L. Kammerdiner, M. Parris, S. Traynor, V. Ottenbacher, A. Shawabkeh and W. F. Oliver, "Switching kinetics of lead zirconium titanate submicron thin film memories," *J. Appl. Phys.*, vol. 64, no. 2, pp. 787–782, Mar. 1988.

[24] Y. W. So, D. J. Kim, T. W. Noh, J.–G. Yoon and T. K. Song, "Polarization switching kinetics of epitaxial $Pb(Zr_{0.4}Ti_{0.6})O_3$ thin films," *Appl. Phys. Lett.*, vol. 86, no. 9, 092905, Feb. 2005.

[25] A. N. Kolmogorov, "On the static theory of metal crystallization," *Izv. Akad. Nauk. Ser. Math.*, vol. 3, pp. 355, 1937.

[26] M. Avrami, "Transformation–time relationsfor random distribution of nuclei," *J. Chem. Phys.*, vol. 8, pp. 212, 1940.

[27] Y. Ishibashi and Y. Tagaki, "Note on ferroelectric domain switching," *J. Phys. Soc. Jap.*, vol. 31, pp. 506–510, 1971.

[28] I. Stolichnov, L. Malin, E. Colla, A. K. Tagantsev and N. Setter, "Microscopic aspects of the region–by–region polarization reversal kinetics of polycrystalline ferroelectric $Pb(Zr,Ti)O_3$ films," *Appl. Phys. Lett.*, vol. 86, no. 1, 012902, 2005.

[29] W. Li and M. Alexe, "Investigation on switching kinetics in epitaxial

$Pb(Zr_{0.2}Ti_{0.8})O_3$ ferroelectric thin films: Role of 90° domain walls," *Appl. Phys. Lett.,* vol. 91, no. 26, 262903, Dec. 2007.

[30] A. K. Tagantsev, I. Stolichnov and N. Setter, "Non–Kolmogorov–Avrami switching kinetics in ferroelectric thin films," *Phys. Rev. B*, vol. 66, 214109, Dec. 2002.

[31] O. Lohse, M. Grossmann, U. Boettger, D. Bolten and R. Waser, "Relaxation mechanism of ferroelectric switching in $Pb(Zr,Ti)O_3$ thin films," *J. Appl. Phys.*, vol. 89, no. 4, pp. 2332–2336, Feb. 2001.

[32] A. Sheikholeslami and P. G. Gulak, "Transient modeling of ferroelectric capacitors for nonvolatile memories," *IEEE Tran. Ultrason. Ferroelectr. Freq. Control*, vol. 43, no. 3, pp. 450–456, May 1996.

[33] J. Y. Son, C. S. Park, S.–K. Kim, Y.–H. Shin, "Writing ferroelectric domain bits on the $PbZr_{0.48}Ti_{0.52}O_3$ thin film," *Appl. Phys. Lett.*, vol. 104, 064101, Sept. 2008.

[34] T. P. Juan, C. Chang and J. Y. Lee, "A new metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (Dy_2O_3)–semiconductor (MFIS) FET for nonvolatile memory applications," *IEEE Elec. Dev. Lett.*, vol. 27, no. 4, pp. 217-220, Apr. 2006.

[35] T. P. Juan, C. –L. Lin, W. –C. Shih, C. –C. Yang, J. Y. Lee, D. –C. Shye and J. –H. Lu, "Fabrication and characterization of metal–ferroelectric ($PbZr_{0.53}Ti_{0.47}O_3$)–insulator (La_2O_3)–semiconductor capacitors for nonvolatile memory applications," *J. Appl. Phys.*, vol. 105, 061625, Mar. 2009.

[36] C. Compagnoni, D. Ielmini, A. S. Spinelli, A. L. Lacaita, C. Previtali and C.

Gerardi, "Study of data retention for nanocrystal Flash memories," *Proc. Int. Rel. Phys. Symp.* pp. 506-512, 2003.

[37] J. Y. Jo, S. M. Yang, H. S. Han, D. J. Kim, W. S. Choi, T. K. Song, J.–G. Yoon, C.–Y. Koo, J.–H. Cheon, S.–H. Kim and T. W. Noh, "Composition–dependent polarization switching behaviors of (111)–preferred polycrystalline $Pb(Zr_xTi_{1-x})O_3$ thin films," *Appl. Phys. Lett.*, vol. 92, 012917, Jan. 2008.

CHAPTER 8 SUMMARY AND FUTURE WORK

8.1 Summary of Major Contributions

The major contributions accomplished towards working for this dissertation are briefly summarized as follows:

1. Developed a novel 10-T nonvolatile-SRAM memory by integration of low-voltage metal-nanocrystal PMOS Flash transistors in the conventional 6-T SRAM design. The circuit was designed to perform store/erase operations without interfering with the regular SRAM operation and not consuming any bias current. Spice-level simulations performed after calibration of the Flash cell to measured results depicted substantial performance enhancement in check-pointing operations with immunity to process and supply voltage variations at massively parallel scales compared to competing designs based on PC-RAM and FeRAM [1]-[2].
2. Proposed the integration of self-assembled ordered nanopores in the tunneling oxide and storage layer of charge-trap Flash memory. This work was the first investigation into studying the effects of intentional as well as unintentional porosity in Flash memory device. The results clearly suggested that nanopores altered the program and retention characteristics providing higher tunneling efficiencies and longer retention at the cost of smaller memory window.
3. Introduced a novel hybrid memory design that integrated a complementary switching mechanism between ferroelectric polarization and nonvolatile charge injection. The two mechanisms were shown to aid each other in providing

better memory performance. The gate-injection of electrons into floating nodes situated above the ferroelectric added to the total memory window of the hybrid device. During retention, these stored electrons partially compensated the ferroelectric polarization that reduced the depolarization in ferroelectric and electron escape from the tunnel oxide simultaneously. The first generation of such devices made from PVDF polymer showed larger memory window and longer retention over conventional FE-FETs.

4. Discovered the unique characteristics of evolution of electric field in a ferroelectric-dielectric compound stack under constant voltage stress. At the beginning of the voltage pulse, the field inside the ferroelectric is high that promotes accelerated domain alignment. The aligned dipoles add to the total displacement in the ferroelectric and in turn diminish the electric field with increase in polarization. This leads to field enhancement in the adjoining oxide. This unique two-step mechanism was proposed to be exploited in the hybrid device to promote dual-switching speed hybrid memory with integration of fast switching oxide-based ferroelectric material.
5. Demonstrated a low-voltage dual-speed hybrid memory based on the proposed two-step mechanism by integration of PZT as the ferroelectric and Au nanocrystals as charge storage nodes. The devices displayed two modes of program with sub-10-V operation: 1) fast DRAM-like mode with sub-100 ns switching in ferroelectric depicting ~0.2 V memory window and short retention and 2) slower Flash-like mode arising from electron tunneling into Au NCs which resulted in higher memory window (~0.8 V) and longer

retention. The two memory mechanisms were shown to be clearly discernibly at all program pulses.

6. Proposed a generalized method to model the program and retention dynamics in the dual-speed hybrid memory by integration of a part-by-part ferroelectric switching model into a conventional Flash memory dynamics simulation. The model correctly predicted the enhancement of fields in the ferroelectric and the tunnel oxide in the two-step program mechanism. The model predictions were shown to be in reasonable agreement with the measurements performed on the PZT-Au NC devices.

8.2 Suggestions for future work

1. Performance evaluation of the hybrid memory in DRAM and Flash mode is necessary for understanding its applications to embedded low-power applications. There is a design trade-off involved in optimizing the DRAM mode by improving the thickness of the ferroelectric layer with the Flash mode that benefits from a thicker tunnel oxide. Further, there are different ferroelectric materials available for integration in the hybrid device today and yet there is no clear understanding into the switching properties of these materials and their impact on performance. For example, a fast switching ferroelectric also depolarizes rapidly during retention and therefore there is a need for investigation of new and existing ferroelectric materials that optimize the fast-programming in polarization, yet meeting the targeted retention in the DRAM mode.

2. With the possibility of thin-film ferroelectric HfO_2 [3]-[5], hybrid devices may be further pushed into the realms of sub-5 V nonvolatile memory devices. Program and erase mechanism in the hybrid device does not involve high current specifications like resistance switching devices, making them truly viable in NAND or NOR like architectures in storage class memory applications.
3. Cell-to-cell interference is the primary bottleneck for scaling of Flash memory devices today. We need to investigate such interference in ferroelectric transistors by studying the dynamics of polarization-coupling in off-axis configurations. Detailed 3-D simulations must be undertaken after evaluating the complete dielectric tensor properties of the ferroelectric thin film to estimate the cross-talk between adjacent cells and the impact of charge storage on the interference.

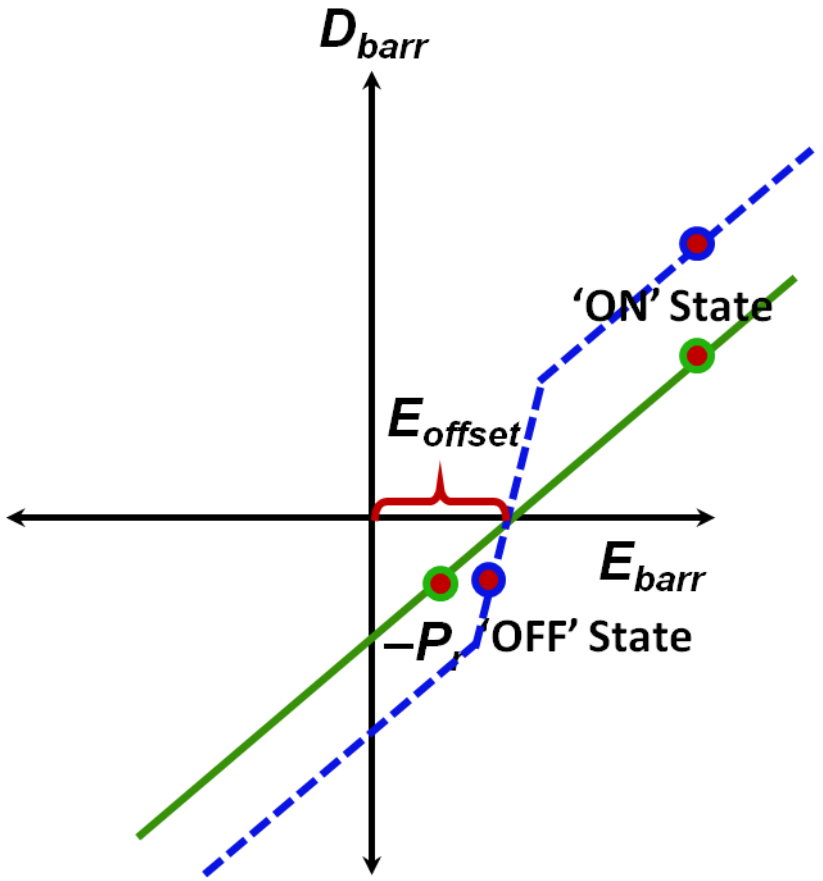


Figure 8-1 The concept of Paraelectric Tunnel Junction as an effective tunnel barrier

4. The idea of nonlinear dielectric response in the ferroelectric can be extended to bring about steep switching in tunnel diodes. This is shown in Fig. 8-1. The forward switching curve in the ferroelectric can be utilized to create sharper switching. The properties of such a tunnel barrier are analogous to a paraelectric response that is right-shifted in the *D-E* characteristics. Right-shift along the *D-E* characteristics can be obtained by placing fixed charge at the thin film interface by chemical or electrical doping. Let us compare this barrier to a conventional right-shifted dielectric barrier (Fig. 8-1). The electric field

and the net displacement in the paraelectric and the competing dielectric during 'OFF' and 'ON' state are shown in the figure. At the offset field (E_{OFFSET}), the paraelectric tunnel barrier goes through displacement reversal due to the polarization reversal in the atomic dipoles. This creates a huge resulting change in displacement at the source injection barrier, increasing the carrier concentration instantaneously. The dielectric barrier on the other hand, goes through a slow transition in displacement that is only aided by electronic polarization. In effect, for the same magnitude of change in the voltage applied across the two barriers, the paraelectric barrier creates larger change in carrier concentration compared to dielectric barrier. This leads to steeper change in the tunneling current across the paraelectric barrier.

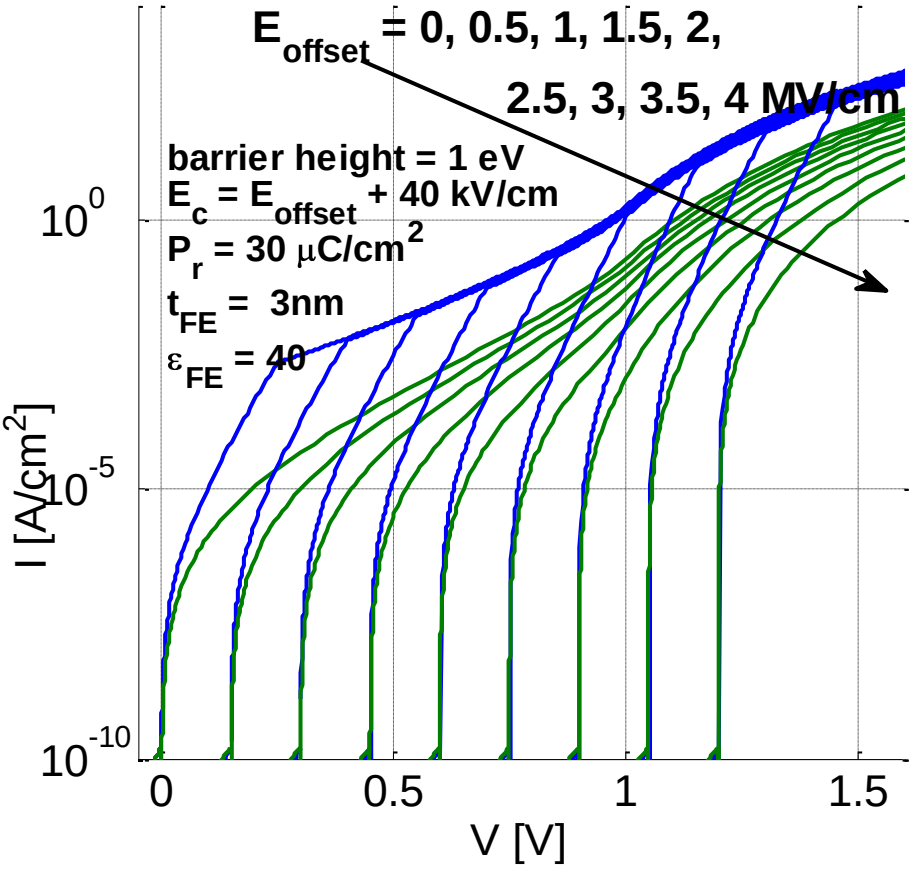


Figure 8-2 I-V characteristics comparison between a right-shifted paraelectric (blue) and a dielectric (green) tunnel barrier

Fig. 8-2 shows the simulated I-V characteristics for the paraelectric barrier compared to a dielectric barrier for different E_{OFFSET}. The paraelectric material and electrical properties are also mentioned in the figure. We can see an effective improvement of ~ 50 X in the ON current characteristics with the introduction of the paraelectric for the same OFF-state leakage in both structures. As a proof-of-concept, a forward curve in the ferroelectric

hysteresis may be used to demonstrate such sharper switching in ferroelectric tunnel diodes from thin film ferroelectric materials like HfO_2 [3]-[5].

REFERENCES

[1] M. Takata, K. Nakayama, T. Izumi, T. Shinmura, J. Akita and A. Kitagawa, "Nonvolatile SRAM based on phase change," in *Proc. Nonvolatile Semicond. Mem. Workshop*,Feb. 2006, pp. 95-96.

[2] T. Miwa, J. Yamada, H. Koike, H. Toyoshima, K. Amanuma, S. Kobayashi, T. Tatsumi, Y. Maejima, H. Hada and T. Kunio "NV-SRAM: A Nonvolatile SRAM with backup ferroelectric capacitors," *IEEE J. Solid-State Circuits*, vol. 36, no. 3, pp. 522-527, Mar. 2001.

[3] T. S. Böscke, J. Müller, D. Bräuhaus, U. Schröder and U. Böttger, "Ferroelectricity in hafnium oxide thin films," *Appl. Phys. Lett.*, vol. 99, 102903, Sept. 2011.

[4] J. Müller, U. Schröder, T. S. Böscke, I. Müller, U. Böttger, L. Wilde, J. Sandqvist, M. Lemberger, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectricity in yttrium–doped hafnium oxide," *J. Appl. Phys.*, vol. 110, 114113, Dec. 2011.

[5] J. Müller, T. S. Böscke, D. Bräuhaus, U. Schröder, U. Böttger, J. Sandqvist, P. Kücher, T. Mikolajick and L. Frey, "Ferroelectric $Zr_{0.5}Hf_{0.5}O_2$ thin films for nonvolatile memory applications," *Appl. Phys. Lett.*, vol. 99, 112901, Sept. 2011.