國立交通大學

電子工程學系 電子研究所碩士班

碩 士 論 文

二氧化鉿奈米晶體快閃式記憶元件可靠度分析

Investigation of Reliability Issues in HfO_2 Nanocrystal Flash Memory Cells

研 究 生：李智雄

指導教授：汪大暉 博士

中華民國 九十六 年 六 月

二氧化鉿奈米晶體快閃式記憶元件可靠度分析

Investigation of Reliability Issues in HfO_2 Nanocrystal Flash Memory Cells

研 究 生：李智雄　　Student：Chih-Hsiung Li

指導教授：汪大暉 博士　　Advisor：Dr. Tahui Wang

國 立 交 通 大 學

電子工程學系 電子研究所碩士班

碩 士 論 文

A Thesis

Submitted to Department of Electronics Engineering & Institute of Electronics

College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of Master

in

Electronic Engineering

July 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年六月

二氧化鉿奈米晶體快閃式記憶元件可靠度分析

學生：李智雄　　指導教授：汪大暉 博士

國立交通大學　電子工程學系　電子研究所

摘要

本篇論文主要是著重於探討二氧化鉿奈米晶體快閃式記憶元件中電荷流失的機制，並建立一個資料保存模型(retention model)。在這個研究當中，我們在大面積的元件中量測電荷流失導致閘極漏電流對溫度及時間的關係，結果發現電荷流失的過程符合阿瑞尼士圖(Arrehenius plot)的溫度關係，跟在 SONOS 快閃式記憶元件中的 Frenkel-Poole 模型有所不同。一個經由熱活化的穿隧模型(thermally activated tunneling front model)被提出來解釋二氧化鉿奈米晶體快閃式記憶元件中電荷流失的行為。

在小面積的元件中，我們藉由量測次啟始電流(subthreshold current)對時間的暫態行為來監控二氧化鉿奈米晶體快閃式記憶元件中電荷流失的行為。溫度的關係圖也被量測。一個兩階段變化的行為被觀察到，我們推論第一階段的行為與電荷流失導致閘極漏電流的行為互相呼應，而第二階段的行為的研究仍在進行當中。最後我們經過理論計算將大小面積的元件之行為做連結。

Investigation of Reliability Issues in HfO_2 Nanocrystal Flash Memory Cells

Student: Chih-Hsiung Li Advisor: Dr. Tahui Wang

Institute of Electronics, National Chiao Tung University

Hsinchu, Taiwan, R.O.C.

Abstract

The main concern of this dissertation will focus on the charge loss mechanism of HfO_2 nanocrystal flash memory cells and develop a retention model. In this study, we measure the temperature and time dependence of a charge loss induced gate leakage current in a large area cell directly. We find that the stored charge emission process exhibits an Arrehenius relationship with temperature, as opposed to linear temperature dependence in a SONOS flash memory. A thermally activated tunneling front model is proposed to account for the charge loss behavior in a HfO_2 dot flash memory.

In a small area device, a subthreshold transient current technique is adopted to monitor charge loss behavior. The temperature dependence is also investigated. A two-stage behavior of subthreshold current transient is observed. The first-stage behavior might correspond to the charge loss induced gate leakage current while the second-stage behavior remains under further investigation. Finally, a connection between large and small area devices is made.

謝誌

首先向指導教授汪大暉教授表達深摯的謝意，對學生總是從平常的話語、態度中給予教誨，以言教及身教指導學生做研究該有的嚴謹態度，使我在碩士兩年的時間獲益良多。

在碩士兩年的生活中，感謝各位學長在論文研究上以及做人處事上的指導；尤其是小馬學長以及小兔學長不厭其煩的給予我實驗上的指導，並使我能順利完成這份論文。也感謝志昌學長以及俊榮學長平時給予知識上及生活規劃上的指導。

在這兩年中也感謝本研究群的小鴨、阿福、ADorken、markwu 以及各位學弟妹(元元、彥君、子華、佑亮、柏凱、勛廷)的陪伴鼓勵，另外也感謝莊老師研究群的各位夥伴在生活上、研究上的幫忙，讓我的實驗室生活充滿樂趣，給我難忘的兩年。

最後感謝我的家人以及我的女朋友靖妮，每當我遭遇挫折、灰心失意時，總是能諒解並體量給我關懷、支持與鼓勵，讓我的碩士生涯充滿溫馨與甜蜜，少了你們任何一個都將不會有今日的我。

Contents

Figure Captions

Chapter 1

Introduction

Flash memory has become a kind of nonvolatile memory widely used in personal computers and portable electronic products. It can be storing, reading and erasing data for many times and the data will not lose even after the supplied power is cut off. There are two kinds of charge storage devices: “Floating gate devices [1,2]” and “Charge trapping devices”. The charges in floating gate devices are stored in a thin conducting layer sandwiched between insulators, while the charges in charge trapping devices are stored in the traps at the interfaces of a multi-layer gate structure and/or in the bulk of insulator.

Flash memory cells employing discrete charge storage nodes have received much interest for their better cell scalability [3-10]. To improve program/erase (P/E) speed and data retention properties in these cells, a lot of efforts have been made with regard to charge storage media. For instance, various trapping materials have been studied in a SONOS-type flash memory [4]. Another category of discrete charge storage flash memories are to use nano-crystals as storage nodes. Many different types of nano-crystals from semiconductors (Si, Ge) to metals (W, Au) have been proposed [5-8]. The nano-crystal memories should have better reliability because the discrete storage node will minimize the impact of weak spots in the tunnel oxide [11]. Therefore, the supplied voltage for programming and erasing can be lowered due to reducing the tunnel oxide thickness. Recently, a HfO_2 dielectric dot flash memory was presented with superior characteristics in terms of a large memory window, fast P/E speed, and long charge retention time [9]. As compared to semiconductor/metal dots, electrons in a dielectric dot are stored in trap states rather than conduction states. Because trapped electrons have a very sharp wave-function distribution in space, size

quantization effect is not expected for the programmed electrons in a dielectric dot.

There are five chapters in this thesis. Chapter 1 is Introduction and we will introduce the evolutions of the nanocrystal flash memory. In Chapter 2, the charge loss induced gate leakage current of nanocrystal flash memory is measured and a thermally activated tunneling front model is developed. In Chapter 3, we describe the subthreshold current transient model. Next, the charge loss in nanocrystal flash memory is investigated by using a subthreshold transient current technique in Chapter 4. Finally, summary and conclusion will be given in chapter 5.

Chapter 2

Charge Retention Loss in a HfO_2 Dot Flash Memory

2.1 Introduction

Nanocrystal flash memory has recently received much attention for the application in the next-generation nonvolatile memories because their better data retention. Much research with respect to quantum dot memory has been conduced to study the data retention of nanocrystal flash memories. Compagnoni et al. analyzed the data retention of Si nanocrystal flash memories [11]. Yang et al. reported that the retention time of memory based on Ge/Si hetero-nanocrystals increases due to the stair-like potential barrier [10]. All the above studies, however, are restricted to conductive quantum dot flash memory. The dielectric dot flash memory has not been understood clearly so far. In this chapter, we will explore the charge retention loss mechanism in a HfO_2 dielectric dot cell. We fabricate a large area cell to measure a charge loss induced gate leakage current directly. The temperature and the retention time dependence of the gate leakage current is characterized. In addition, we measure a high-voltage stress induced gate leakage current in a SONOS cell for comparison. A thermally activated tunneling front model is developed for charge retention loss in a HfO_2 dot flash memory.

2.2 Charge Loss Induced Gate Leakage Current Characterization

In this chapter, charge loss induced leakage current is measured to monitor storage charge loss in a flash memory because it is directly related to the charges escaped from the storage layer. This section aims to investigate the correlation between gated leakage current and escaped charges.

Fig. 2.1 shows the storage charges stored in the storage layer and equivalent circuit of capacitor part. Two capacitors will form after a flash memory is programmed. The transient behavior of stored charges is shown in Fig. 2.2. These two capacitors are connected in parallel due to the small-signal variation. The charge distribution is

$$Q_{image1} = \frac{C_1}{C_1 + C_2} Q_{storage} \text{ and } Q_{image2} = \frac{C_2}{C_1 + C_2} Q_{storage} \quad (1)$$

If some storage charges (ΔQ_{loss}) escape through the bottom oxide, the remaining charges will be redistributed. This redistribution induces charge loss both in Q_{image1} and Q_{image2}. The arrowheads shown in Fig. 2.2 (b) represent the direction which charges escape. Current flows, including I_{loss}, I_{loss1} and I_{loss2}, result from these charge loss mentioned above (Fig. 2.2 (b)). Their relationships are expressed as follows:

$$I_{loss1} = \frac{C_1}{C_1 + C_2} I_{loss} \text{ and } I_{loss2} = \frac{C_2}{C_1 + C_2} I_{loss} \quad (2)$$

Inspection of charge loss directions presented in Fig. 2.2(b) shows that I_{loss1} is the only current flow we can sense from gate terminal, i.e.

$$I_g(\text{transient}) = I_{loss1} = \frac{C_1}{C_1 + C_2} I_{loss} \quad (3)$$

$I_g(\text{transient})$ is negative because it flows out gate terminal. Based on the explanation above, the total charge loss from storage layer is

$$\Delta Q_{loss} = \int_0^t \frac{C_1 + C_2}{C_1} \cdot Ig(\text{transient})\, dt \quad (4)$$

The similar behavior can be observed if the charges escape through the top oxide as shown in Fig 2.3. The only difference is

$$I_g\left(\text{transient}\right) = I_{loss} - I_{loss1} = \frac{C_2}{C_1 + C_2} I_{loss} \tag{5}$$

The measured gate current is positive since the net current flows into the gate terminal.

2.3 Measurement Results

The HfO_2 nanocrystal flash memory used in this work has a 8nm top oxide, a 10nm intermediate oxide layer with embedded HfO_2 dots, and a 6nm bottom oxide. The dot size is 5~8nm. The device area is 500×500μm^2 (Fig. 2.4). The fabrication process and device characteristics were published in [9]. A Fowler-Norheim (FN) stress is performed at V_g=-19V for 2000s. Uniform negative FN injection is employed for programming. The program V_t window is 3V. The program-state gate leakage current versus retention time at different temperatures is shown in Fig. 2.5 . On the other side, we also measure the program-state gate leakage current in a SONOS cell (Fig. 2.6). The SONOS cell used for comparison has a 9 nm top oxide, a 6 nm silicon nitride, and a 6 nm bottom oxide. The capacitor area is also 500×500 μm^2. Uniform FN programming is performed after a FN stress at V_g=-20V for 2500s.

Some distinguished features in the gate current of the two cells are observed. First, the gate leakage current is positive in a HfO_2 dot flash (flowing into the gate) but is negative in a SONOS cell. The direction of the gate current flow indicates the escape of programmed electrons through a top oxide in the HfO_2 cell. Second, the gate current in both SONOS [12] and HfO_2 dot flash cells exhibits 1/t time dependence. The 1/t characteristic can be derived either from a tunneling front model [13] or from a Frenkel-Poole (FP) emission model [14]. In order to distinguish these two models,

we compare the temperature dependence of the gate current in HfO_2 dot flash (Fig. 2.7) and in SONOS cell (Fig. 2.8) . The charge detrapping current in the SONOS cell obeys a linear dependence on temperature. As Gu noted that the nitride charge leakage current exhibits two stage [12] : the leakage current of the first stage is limited by the amount of stress created oxide traps while the leakage current of the second stage is dictated by Frenkel-Poole emission. The measured result in Fig. 2.6 is obviously in the second stage due to the heavily FN stress and which is expected from the FP emission model, i.e, $I_g \propto kT/t$ [14]. Fig. 2.9 Illustrates the stored charge loss process limited by FP emission in a SONOS cell. The gate leakage current in the HfO_2 cell, however, deviates from a linear relationship apparently. In Fig. 2.10, we replot the temperature dependence of the gate current in the HfO_2 cell in Arrehenius coordinates at t=0.5s and 5s. An Arrehenius relationship is obtained with activation energy of 0.19eV. The Arrehenius dependence excludes the possibility of the FP emission. Instead, the observed temperature and the time dependence in the HfO_2 cell can be well accounted for by thermally activated tunneling through traps in a top oxide.

2.4 Thermally Activated Tunneling Front Model

Fig. 2.11 illustrates the stored charge loss process in a HfO_2 cell. Based on the WKB approximation, the tunneling time for an electron from a HfO_2 trap to a top oxide trap (assumed to be a limiting step in a charge loss process) can be formulated as

$$t^{-1} = N_t \nu_{th} \sigma \exp(-\alpha_{ox} x) \tag{6}$$

and

$$\alpha_{ox} = \frac{2\sqrt{2m_e q\phi_b}}{\hbar}$$

where N_t is the trap density in the top oxide, x is the distance between a HfO_2 trap and

a top oxide trap, σ is a trap cross-section and other variables have their usual definition. Since we do not observe significant dependence of the gate leakage current on stress time, N_t is suspected to be pre-existing traps. Assuming the stored charges have a uniform distribution in x, we can derive the time and the temperature dependence of the gate current based on a tunneling front model,

$$I_g = AQ_s \exp(\frac{-E_a}{kT})\frac{dx}{dt}$$
$$= \frac{AQ_s}{\alpha_{ox} t}\exp(\frac{-E_a}{kT}) \quad (7)$$

where $Q_s\exp(-E_a/kT)$ represents the activated charge density in HfO_2 for tunneling. E_a is the activation energy of trapped charges and A is the cell area.

2.5 Discussion on Different Programming Condition

As mentioned in section 2.3, the gate leakage current is positive in a HfO_2 dot flash. On the other hand, a negative gate leakage is measured when positive FN programming is performed. One deduces that the different leakage direction may depend on what kind of programming approach is applied. The most likely explanation is that negative FN injection will fill the traps in the upper part of dots that the stored charges will escape through the top oxide. In contrast, the traps in the bottom part of dots will be programmed during positive FN injection. Therefore, the leakage path of stored charges will pass through the bottom oxide due to the shorter distance (Fig. 2.12).

2.6 Conclusion

In this section, we investigate the charge loss mechanism in a HfO_2 dot flash cell by characterizing a charge loss induced gate leakage current. The Frenkel-Poole emission model is not suitable for charge loss in the cell. A thermally activated

tunneling front model is proposed. Our model can well explain the measured temperature and the retention time dependence of a gate leakage current.

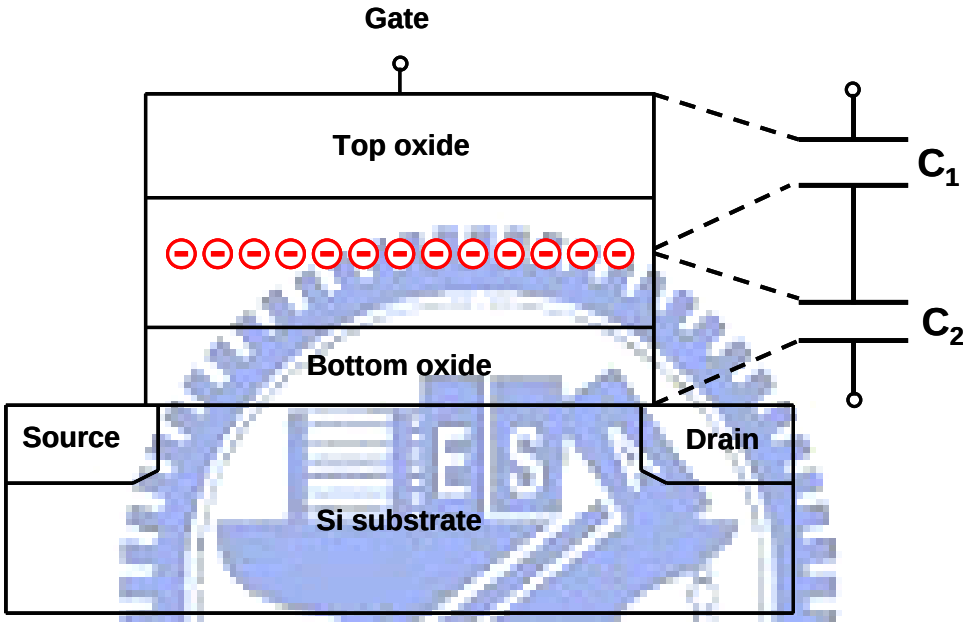


Fig. 2.1 Schematic diagram of storage charges in storage layer and equivalent circuit of capacitor part.

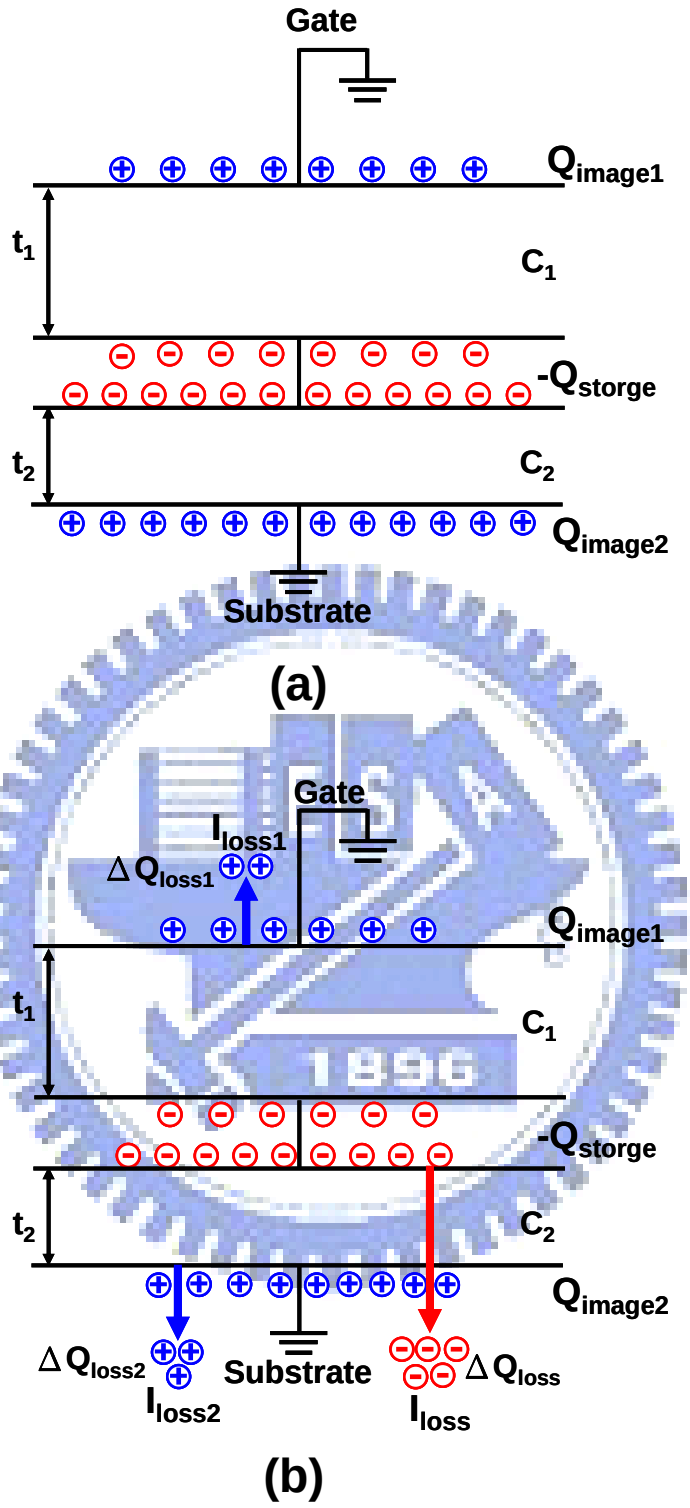


Fig. 2.2(a) Distribution of storage charges while the trapped position is near the bottom oxide.

Fig. 2.2(b) Transient behavior of charge loss. The storage charges escape through the bottom oxide.

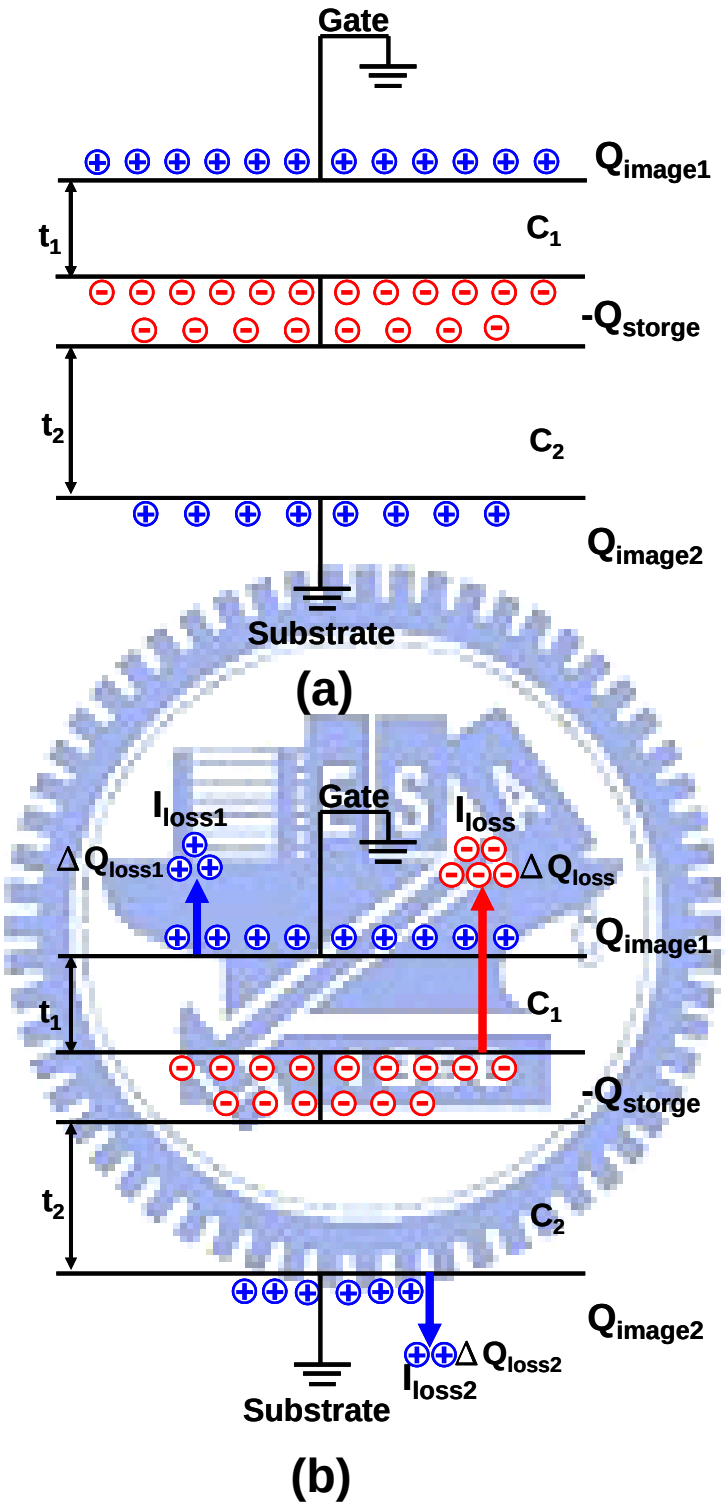


Fig. 2.3(a) Distribution of storage charges while the trapped position is near the top oxide.

Fig. 2.3(b) Transient behavior of charge loss. The storage charges escape through the top oxide.

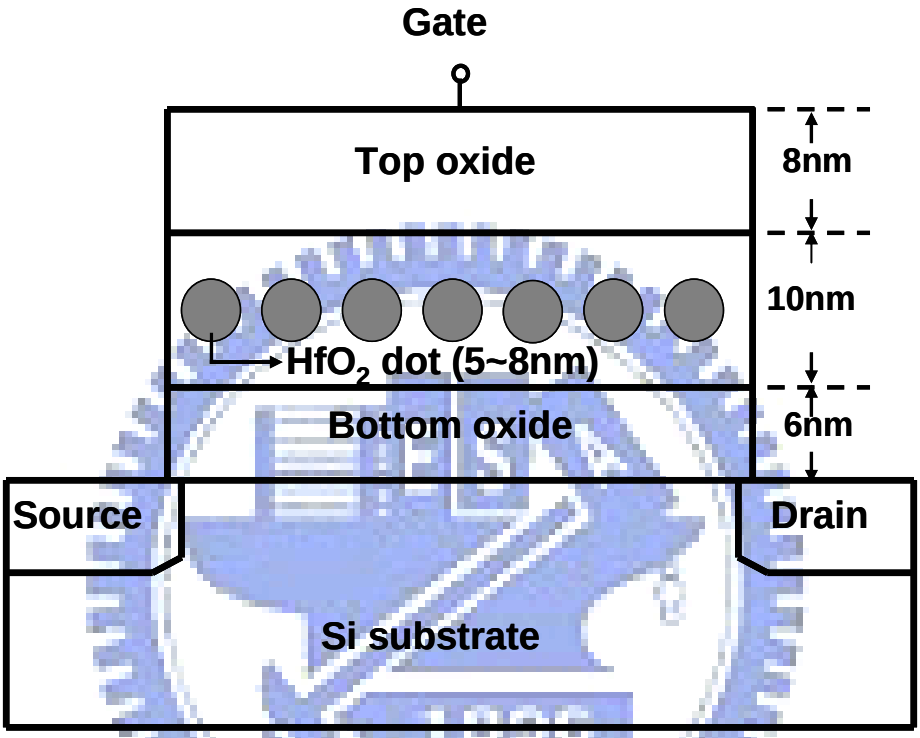


Fig. 2.4 Schematic representation of the HfO_2 nanocrystal flash memory cell structure.

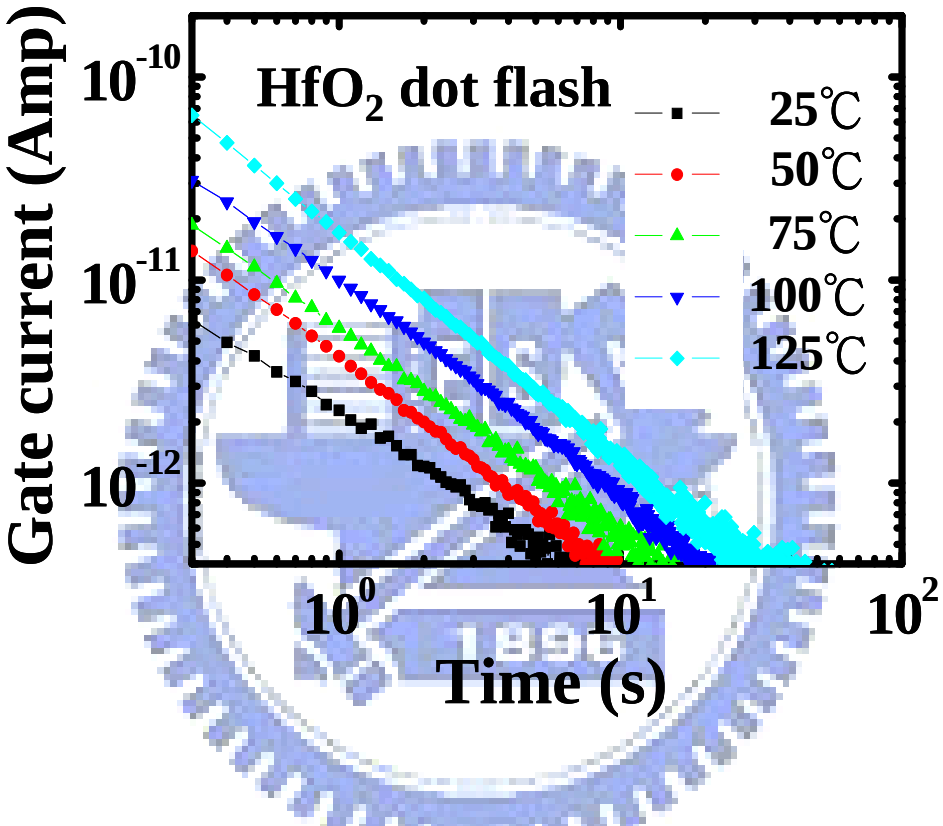


Fig. 2.5 Measured gate leakage current at V_g =0V versus time in a HfO_2 dot flash memory cell (500μm×500μm). The temperature is form 25 ◦ C to 125 ◦ C.
The device is programmed to a threshold window of 3V.

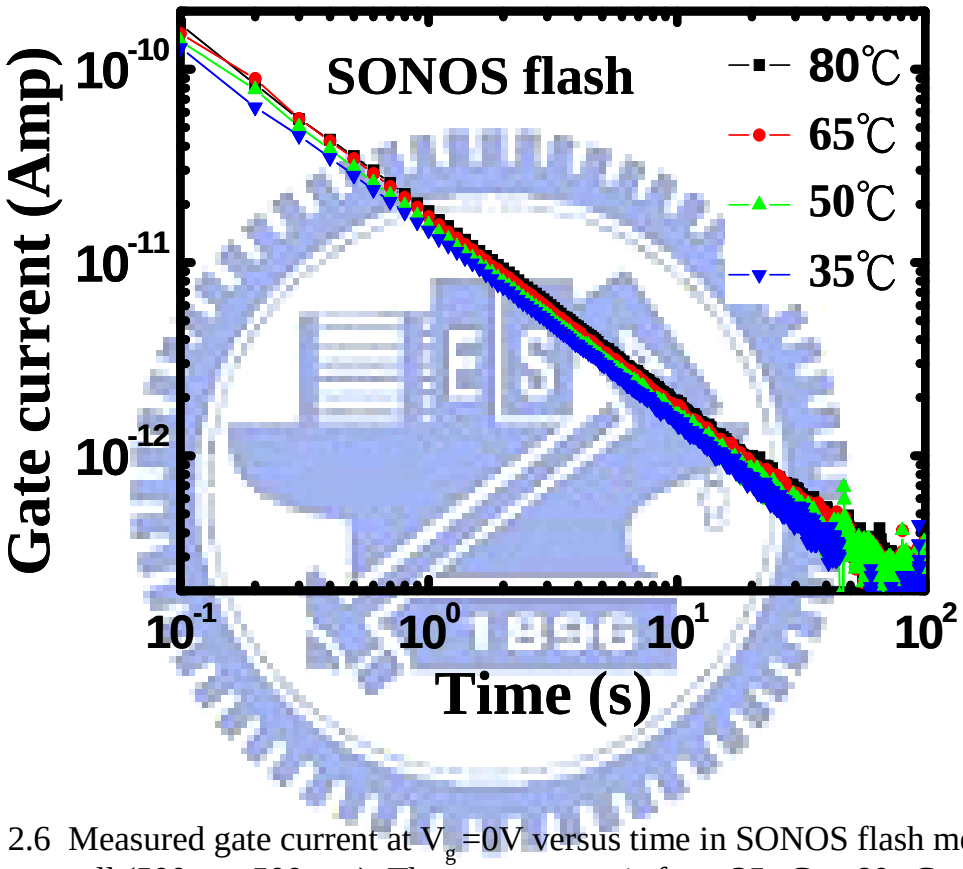


Fig. 2.6 Measured gate current at V_g=0V versus time in SONOS flash memory cell (500μm×500μm). The temperature is form 35 ◦ C to 80 ◦ C. The device is programmed to a threshold window of 3V.

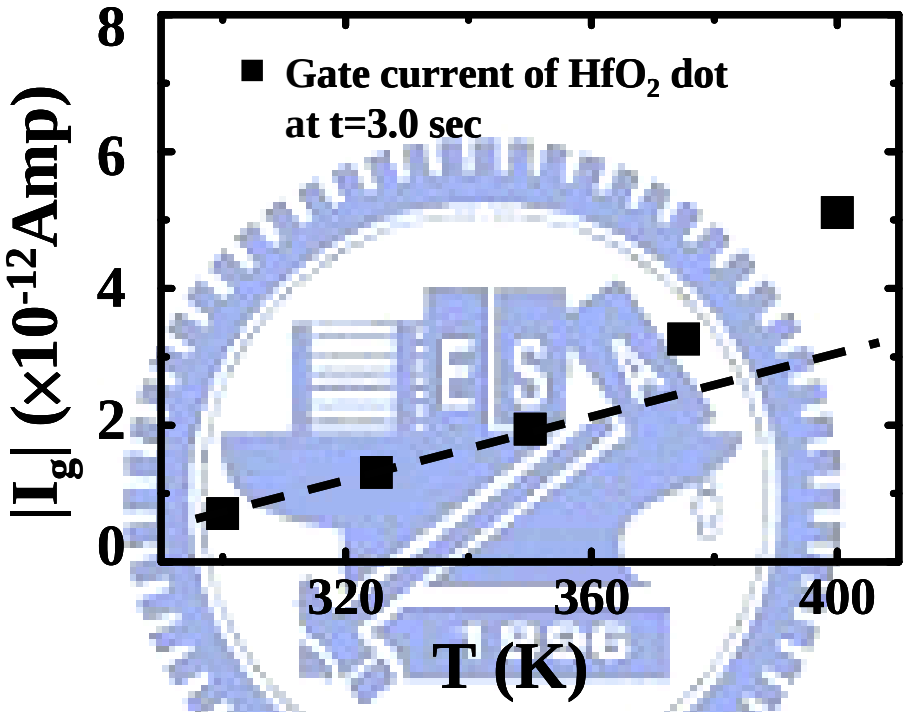


Fig. 2.7 Temperature dependence of the gate current in a HfO_2 dot flash memory cell at t=3sec.

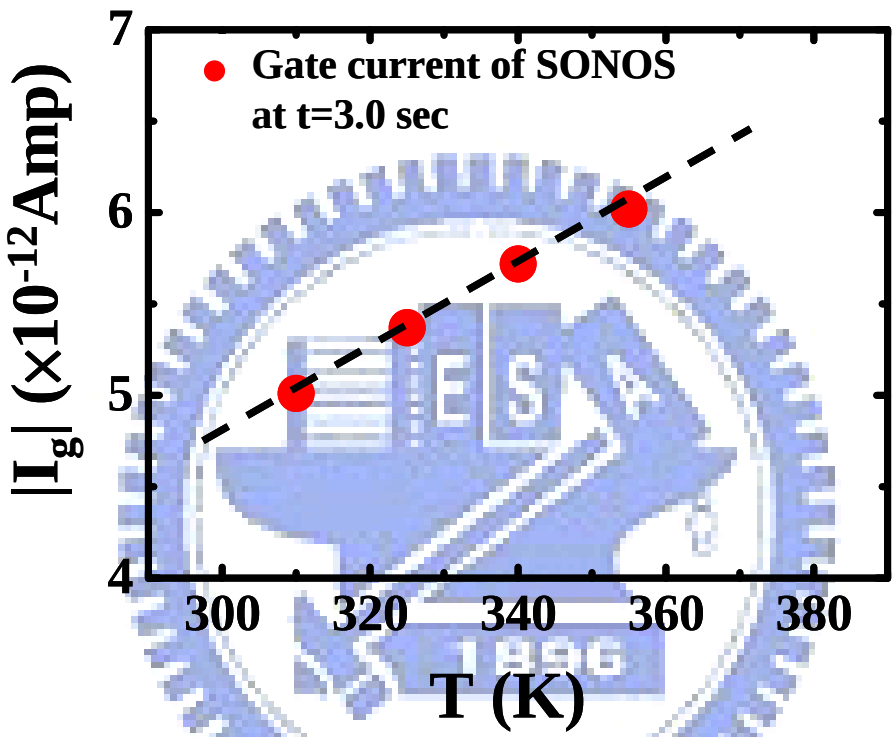


Fig. 2.8 Temperature dependence of the gate current in a SONOS flash memory cell at t=3sec.

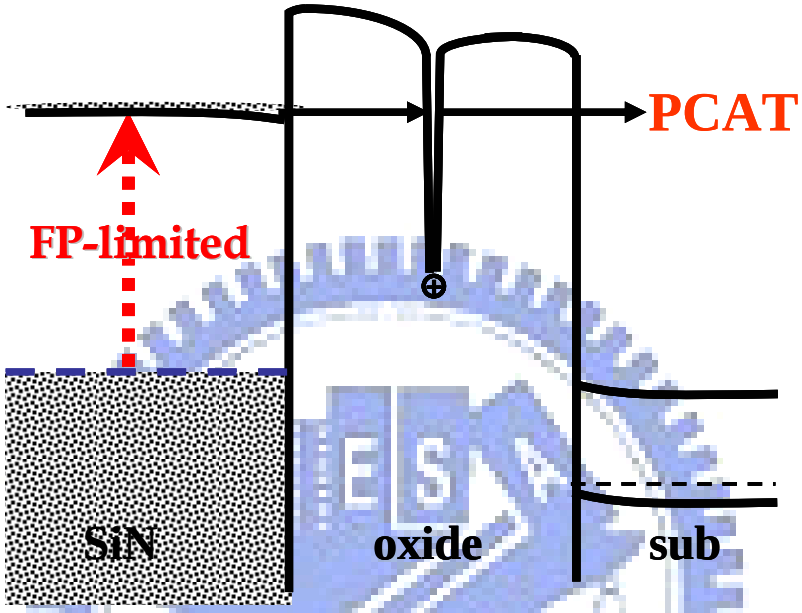


Fig. 2.9 Illustration of the stored charge loss process limited by FP emission in a SONOS cell.

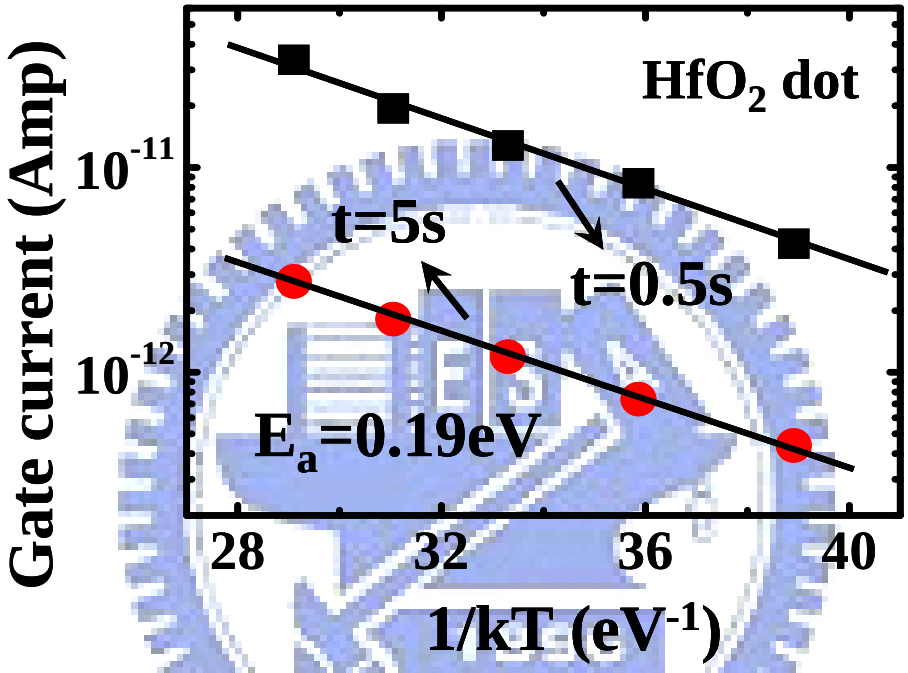
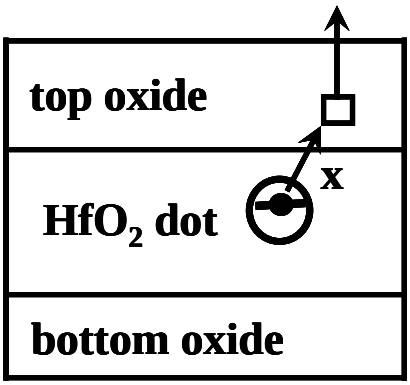


Fig. 2.10 Arrehenius plot of the gate leakage current in a HfO_2 dot flash memory cell.
The retention time is 0.5sec and 5sec.
The extracted activation energy is 0.19eV.



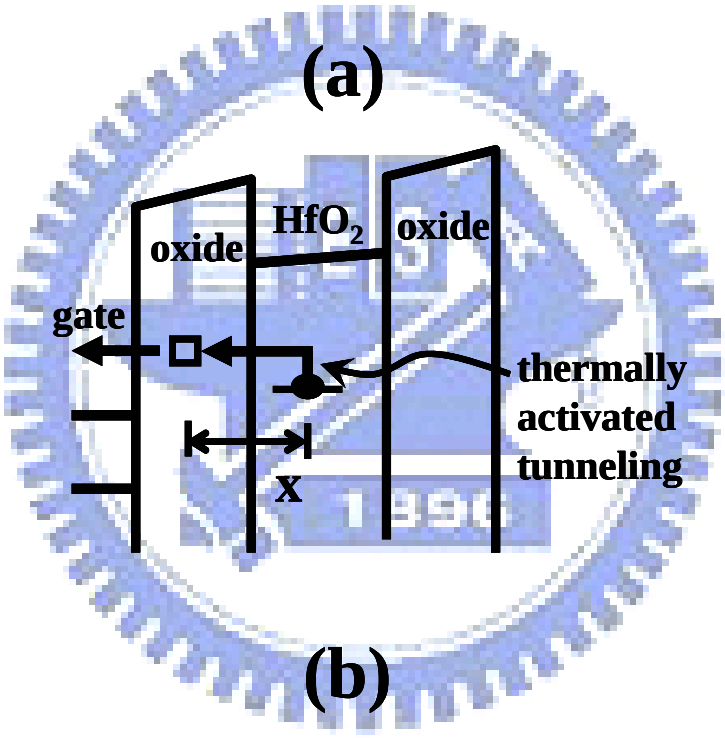


Fig. 2.11 (a) Illustration of charge loss through a top oxide trap in a HfO_2 dot flash memory cell.

Fig. 2.11 (b) Energy band diagram and thermally assisted tunneling of a trapped charge.

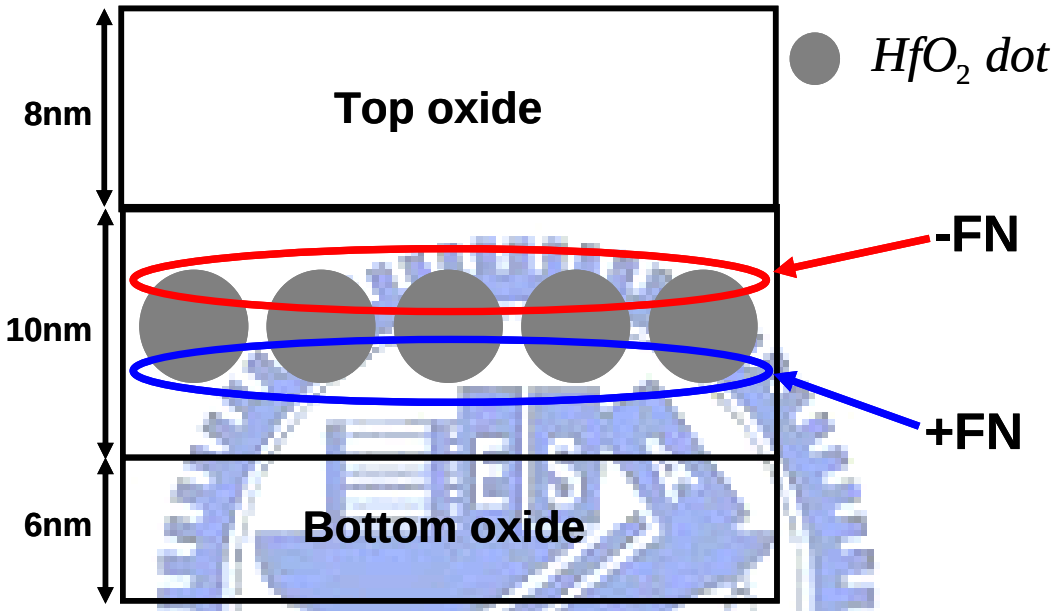


Fig. 2.12 Schematic representation of the possible charge distribution in the vertical direction of the HfO_2 dots after positive/negative FN programming.

Chapter 3

Model of Subthreshold Current Transient

3.1 Introduction

In this chapter, model of subthreshold current transient is developed.

Device instabilities resulting from oxide charge trapping/detrapping such as subthreshold current [15] at a DC bias have been reported. Similarly, storage charge loss in a flash memory also have the same effect on subthreshold current. The principle of the mechanism is illustrated in Fig. 3.1. Fig. 3.1(a) shows a structure of a nonvolatile flash memory cell. Negative charges are stored in the storage layer and the charges may escape vertically through either the top or bottom oxides. The potential barrier in the channel direction will reduce from solid line to dashed line (Fig. 3.1(b)) due to the charge loss. As a result of the potential barrier lowering, subthreshold current at the same gate/drain bias increases. Consequently, the behavior of the escaping stored charge can be deduced by monitoring the evolution of the subthreshold current with cumulative time.

3.2 Modeling of Subthreshold Current Transient

The dependence of subthreshold current on a gate bias is expressed as follows [16]：

$$I_d = I_0 \exp\left(S\left(V_{gs} - V_{fb}\right)\right) \tag{1}$$

where V_{fb} is flat-band voltage and S is the slope of the subthreshold current. The relationship between S and temperature is approximated by [16]

$$S = \frac{q}{nkT} \tag{2}$$

where the parameter n can be evaluated directly from the $I_{ds} - V_{gs}$ in a measured device.

The programmed storage charges are assumed to have a distribution $Q_s(x)$ in the vertical direction in the storage layer (Fig. 2.1) and corresponding detrapping time-constant is $\tau(x)$. x is the distance from the storage charge to the interface between storage layer and bottom oxide. The storage charge detrapping rate is

$$\frac{dQ_S(x,t)}{dt} = -\frac{Q_S(x,t)}{\tau(x)} \tag{3}$$

and

$$Q_S(x,t) = Q_S(x,0)\exp\left(\frac{-t}{\tau(x)}\right) \tag{4}$$

As a result, the change of flat-band voltage due to storage charge detrapping is

$$V_{fb}(t) = V_{fb}(0) + \int_0^{t_s} \frac{\frac{\varepsilon_S}{\varepsilon_{OX}} t_{TOX} + t_S - x}{\varepsilon_S} \times Q_S(x,0)\left(1 - \exp\left(\frac{-t}{\tau(x)}\right)\right) dx \tag{5}$$

where t_{TOX} is the top oxide thickness, t_S is the storage layer thickness, ε_{OX} is the permittivity of oxide and ε_S is the permittivity of the storage layer. Substituting (2) and (5) into (1), we arrive at

$$\ln\left(I_d(t)\right)=\ln\left(I_d(0)\right)-\int_0^{t_S}\frac{q}{nkT}\frac{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{TOX}+t_S-x}{\varepsilon_S}\times Q_s(x,0)\left(1-\exp\left(\frac{-t}{\tau(x)}\right)\right)dx \tag{6}$$

According to the WKB approximation, the trapped charge emission time is formulated as [17]

$$\tau(x)=\tau_0\exp\left(\alpha_{OX}t_{BOX}\right)\exp\left(\alpha_S x\right) \tag{7}$$

with

$$\alpha_{OX}=\frac{2\sqrt{2m_{OX}^*\left(E_t+\Phi_B\right)}}{\hbar}\ ;\ \alpha_S=\frac{2\sqrt{2m_S^*E_t}}{\hbar}$$

where τ_0 is the tunneling characteristic time, E_t is the electron trap energy measured from the conduction band of the storage layer, Φ_B is the conduction band difference between storage layer and oxide, and other variables have their usual definitions. Then, Eq. (6) becomes

$$\ln\left(I_d(t)\right)=\ln\left(I_d(0)\right)-\int_0^{t_S}\frac{q}{nkT}\frac{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{TOX}+t_S-x}{\varepsilon_S}\times Q_s(x,0)\left(1-\exp\left(\frac{-t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\exp\left(-\alpha_S x\right)\right)\right)dx \tag{8}$$

Since the double exponential function in the integrand changes abruptly from 0 to 1 around $x=\alpha_S^{-1}\ln\left(\frac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right)$, it can be reasonably approximated by a step-function as follows：

$$\exp\left(\frac{-t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\exp\left(-\alpha_S x\right)\right)$$

$$= \begin{cases} 0 \quad \text{for } x \le \alpha_S^{-1}\ln\left(\dfrac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right) \\ \\ 1 \quad \text{for } x \ge \alpha_S^{-1}\ln\left(\dfrac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right) \end{cases} \tag{9}$$

This step function approximation is actually the same as the tunneling front model in [13], [18]. Therefore, (8) can be further simplified

$$\begin{aligned} \ln\left(I_d(t)\right) &= \ln\left(I_d(0)\right) - \int_0^{\alpha_S^{-1}\ln\left(\frac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right)} \frac{q}{nkT}\frac{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S-x}{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S}Q_S(x,0)\,dx \\ &= \ln\left(I_d(0)\right) - \frac{q}{nkT}\left(\frac{t_{BOX}}{\varepsilon_{OX}}+\frac{t_S}{\varepsilon_S}\right) \\ &\quad \times \int_0^{\alpha_S^{-1}\ln\left(\frac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right)} \frac{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S-x}{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S}Q_S(x,0)\,dx \end{aligned} \tag{10}$$

By neglecting the term $\dfrac{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S-x}{\frac{\varepsilon_S}{\varepsilon_{OX}}t_{BOX}+t_S}$ and assuming a uniform storage charge distribution, then

$$\ln\left(I_d(t)\right) = \ln\left(I_d(0)\right) - \frac{q}{nkT}\left(\frac{t_{BOX}}{\varepsilon_{OX}}+\frac{t_S}{\varepsilon_S}\right)Q_S\alpha_S^{-1}\ln\left(\frac{t}{\tau_0}\exp\left(-\alpha_{OX}t_{BOX}\right)\right) \tag{11}$$

Finally, a power-law dependence of the subthreshold current transient is readily

obtained, i.e.,

$$I_d(t) \propto t^p \tag{12}$$

with $p = -\frac{q}{nkT}\frac{Q_S}{\alpha_S}\left(\frac{t_{BOX}}{\varepsilon_{OX}} + \frac{t_S}{\varepsilon_S}\right)$

The power factor p is positive for electron detrapping.

The case mentioned above is that charges escape through the bottom oxide. On the other hand, if the stored charges escape through top oxide, a similar derivation is performed. A power-law dependence is also obtained. The only difference is the slope of the subthreshold current transient.

$$p = -\frac{q}{nkT}\frac{Q_S}{\alpha_S}\left(\frac{t_{TOX}}{\varepsilon_{OX}}\right) \tag{13}$$

Therefore, the time-dependent of a subthreshold current transient on a log-log scale is a straight line without respect to the direction that stored charges escape.

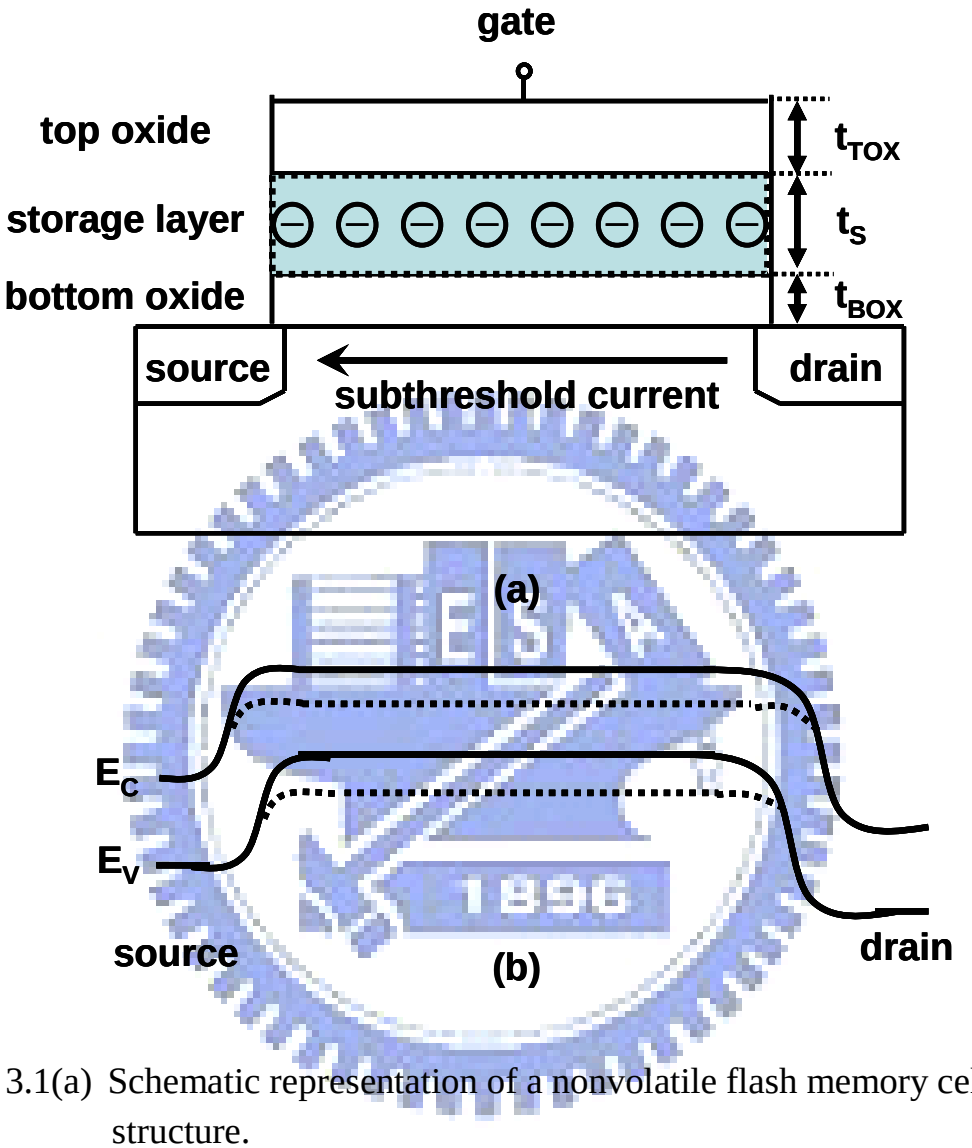


Fig. 3.1(a) Schematic representation of a nonvolatile flash memory cell structure.

Fig. 3.1(b) Potential barrier lowering in the channel direction due to storage charge loss.

Chapter 4

Monitor Storage Charge Loss by Using a Subthreshold Transient Current Technique

4.1 Introduction

Data retention is an important issue in non-volatile flash memory cells. Charge loss through either the bottom or top oxides is considered for the data retention loss. Threshold voltage shift measurement was widely adopted to monitor the storage charge loss in a flash memory. However, threshold voltage is only linearly dependent on storage charge and thus may not be sufficiently sensitive to storage charge variation. In this chapter, we used a subthreshold transient current technique to monitor storage charge variation. Unlike threshold voltage, subthreshold current exhibits an exponential dependence on storage charges. As shown in Fig. 4.1, storage charge loss results in I-V curve shift which is corresponding to threshold voltage shift. Fig. 4.1 also demonstrates that as threshold voltage decreases linearly, subthreshold current increases exponentially when the device is biased at a fixed gate voltage in subthreshold region. Therefore, we can have a better measurement resolution by monitoring the temporal evolution of subthreshold current.

4.2 Time-Dependent of a Subthreshold Transient

4.2.1 Experimental

The devices used in this work are HfO_2 nanocrystal flash memories which have a 8nm top oxide, a 10nm intermediate oxide layer with embedded HfO_2 dots, and a 6nm bottom oxide. Detail fabrication process and device characteristics can be found in [9]. The device area is $10\mu m \times 1\mu m (W \times L)$.

The device is first subjected to an either uniform positive or negative injection for programming. The program V_t window is about 3V. After programming, the device is immediately biased to "subthreshold current measurement phase". In subthreshold current measurement phase, as an appropriate gate voltage is chose at V_d=0.1V, the drain current can be determined in the subthreshold region.

4.2.2 Measurement Results

The subthreshold current transient in a device $(10\mu m \times 1\mu m)$ is measured at T=100°C after uniform positive FN programming. The measured result is plotted on Fig. 3.2 with a log-log scale and the subthreshold transient increases with increasing cumulative time due to storage charge loss. However, instead of a straight line, the result shows a two-stage behavior with a corner time at about 100 seconds. The two-stage behavior might imply that there are two kinds of charge loss mechanism competing inside the storage layer. At beginning, the first kind of mechanism dominates the loss, while the second kind of mechanism will show up at about 100 seconds. At the same time, slope of the subthreshold transient will change to a larger value. We consider that the mechanism of first stage is thermally activated tunneling front model as shown in Chapter 2. The mechanism of the second stage is still unknown. This work is focused on the first stage behavior while the second stage mechanism remains under further investigation.

4.2.3 Comparison Between Positive and Negative FN Programming

Fig. 4.3 shows the subthreshold current transient of a device ($10\mu m \times 1\mu m$) which is programmed by uniform negative FN injection at T=100°C. The two-stage behavior is also observed in Fig. 4.3. Based on the gate leakage current direction in chapter 2, the programmed electrons will escape through the top oxide if negative FN programming is performed. On the other hand, the programmed electrons will escape through the bottom oxide on condition that positive FN programming is applied. The subthreshold current transients after normalized with respect to positive and negative FN programming are replotted on Fig. 4.4. As Fig. 4.4 shown, the charge loss behaviors are the same through either top or bottom oxide.

4.3 Temperature-Dependent of a Subthreshold Transient

Fig. 4.5 illustrates the threshold voltage shift over time of a device ($10\mu m \times 1\mu m$) which is programmed by uniform negative FN injection at various temperatures including 85℃, 100℃ and 115℃. At the temperature at 115℃, we observed a more significant charge loss than that at 85℃ and 100℃. This result indicates the retention characteristic of HfO_2 nanocrystal flash memory has strong temperature dependence which is consistent with the model developed in chapter2. However, Vt variation over time doesn't present the two-stage behavior as clear as subthreshold current transient. The subthreshold current behavior over time versus temperature is plotted on Fig. 4.6 and normalized subthreshold current transients are shown on Fig. 3.7. It is clear from Fig. 3.7 that the slope in first stage of measured transients raises slightly with increasing temperature. The increasing slope supports the fact that raising temperature will speed up the charge loss.

By introducing the gate leakage current results from Chapter 2, theoretical slope values of subthreshold current transient can be worked out. The dependence of subthreshold on a gate bias is expressed as follows [16].

$$I_d = I_0 \exp\left(S\left(V_{gs} - V_{fb}\right)\right) \tag{1}$$

where S is the slope of the subthreshold current which can be evaluated directly from the $I_{ds} - V_{gs}$ curve in a measured device. Charge loss will result in the change of flat-band voltage:

$$V_{fb}(t) = V_{fb}(0) + \frac{\int_0^t I_g dt}{C_1} \tag{2}$$

where C_1 is defined in Fig. 2.1 and a analytical equation of I_g is developed in Chapter 2:

$$I_g = \frac{AQ_s}{\alpha_{ox} t} \exp\left(\frac{-E_a}{kT}\right) \tag{3}$$

Substituting (2) and (3) into (1) and the integration of Ig is modified according to Equation (4) in chapter2, then simplify the result, we arrive at

$$\ln\left(I_d(t)\right) = \ln\left(I_{d0}\right) + \frac{S}{C_1} \cdot \frac{AQ_S'}{\alpha_{OX}} \exp\left(\frac{-E_a}{kT}\right) \cdot \ln(t) \tag{4}$$

The slope of subthreshold current transient is obtained, i.e.,

$$Slope = \frac{S}{C_1} \cdot \frac{AQ_S'}{\alpha_{OX}} \exp\left(\frac{-E_a}{kT}\right) \tag{5}$$

Fig. 4.8 shows theoretical value and measured data of slope at different temperature. The line bar represents the 20% variation based on theoretical value. The variation between theoretical value and measured data might result from the device fluctuation in small area devices [19].

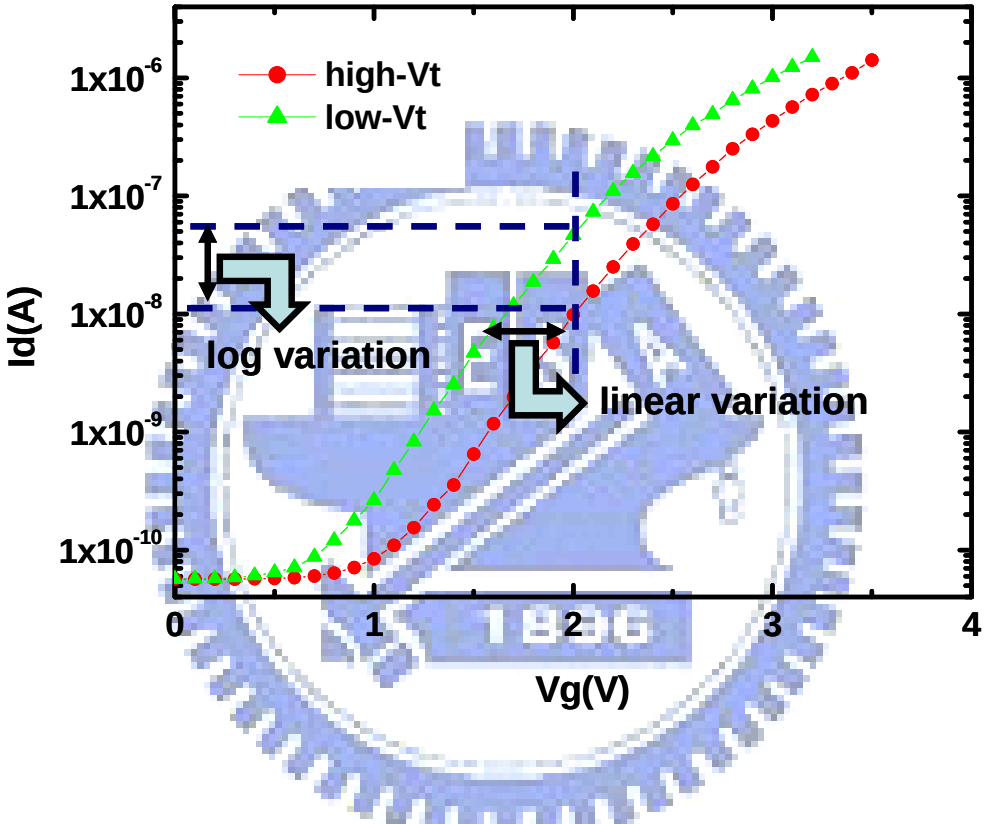


Fig. 4.1 Illustration of linear V_t shift will result in log I_d variation when device biased in subthreshold region.

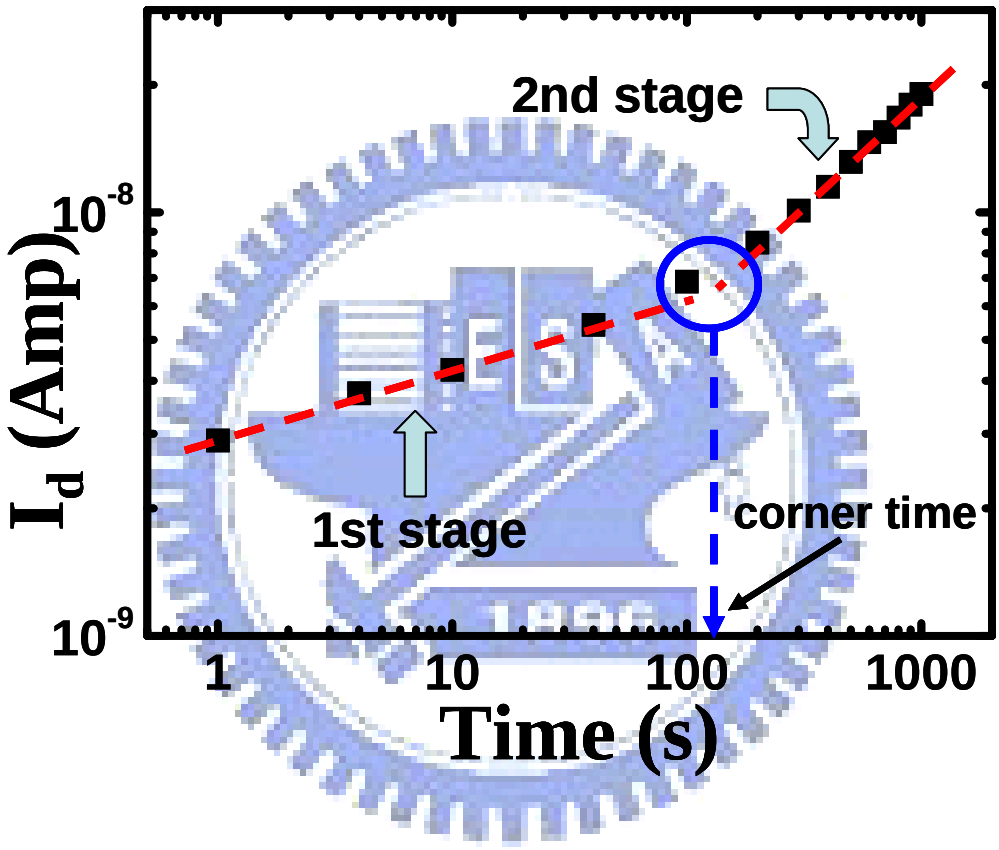


Fig. 4.2 Subthreshold current transient measured at T=100℃. The device is programmed by positive uniform FN injection.

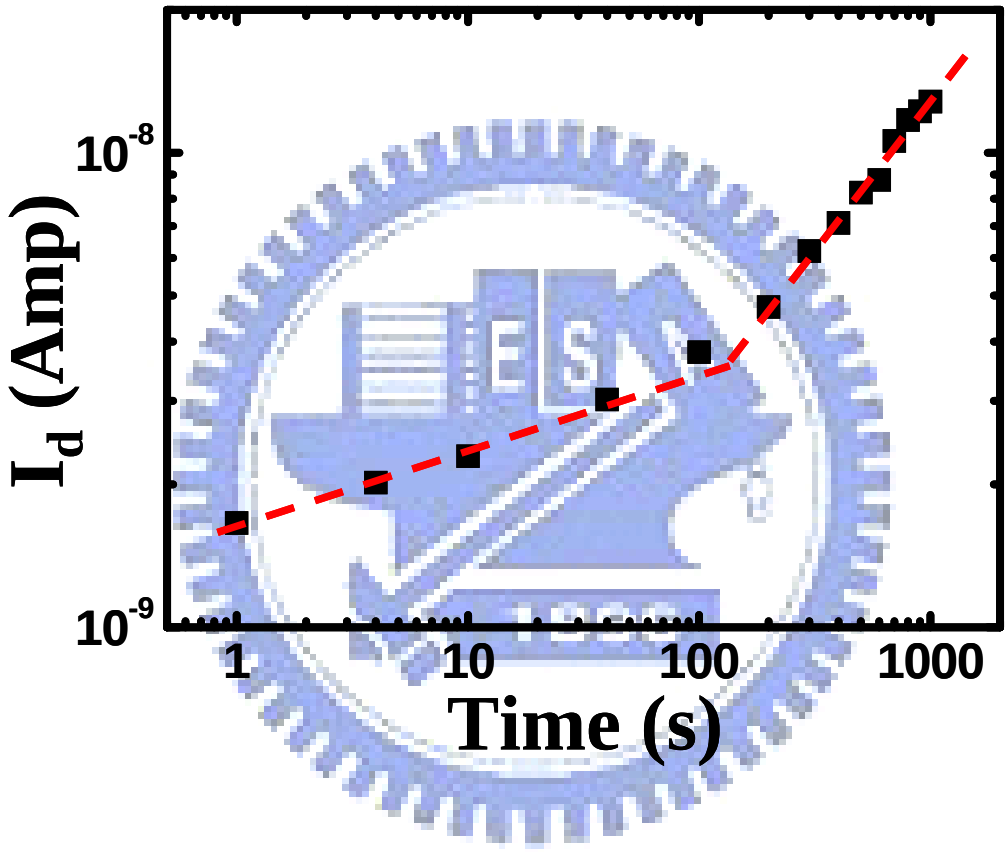


Fig. 4.3 Subthreshold current transient measured at T=100℃. The device is programmed by negative uniform FN injection.

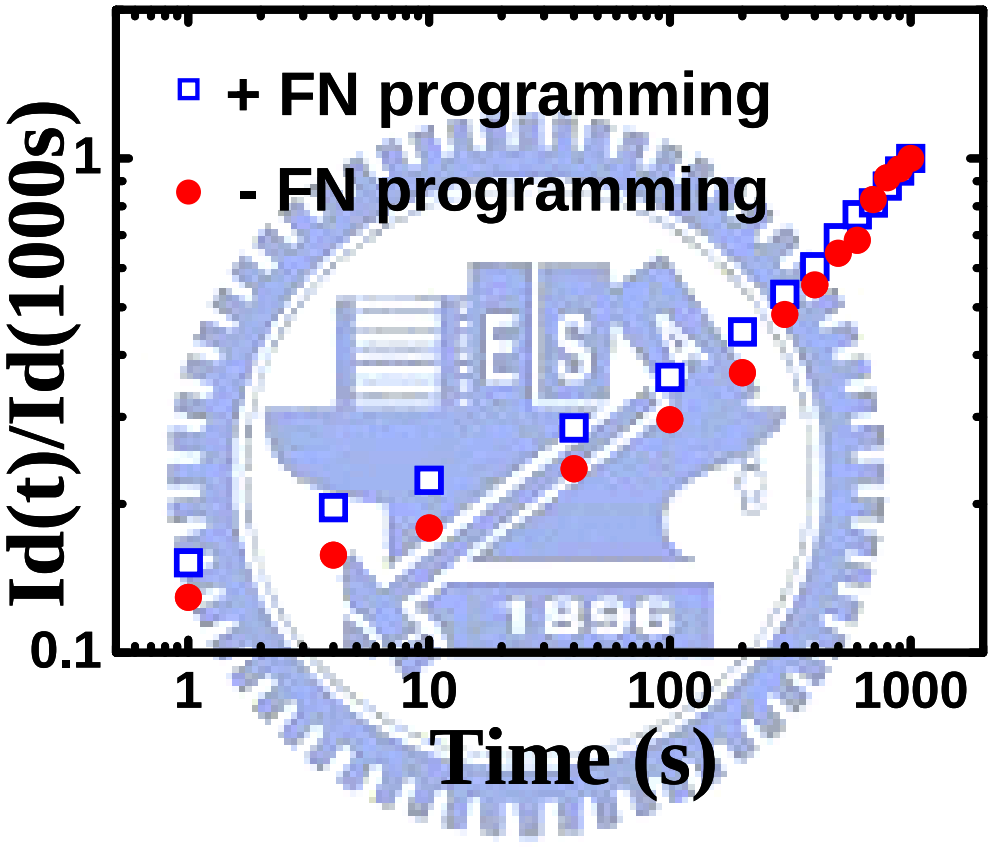


Fig. 4.4 Normalized subthreshold current transients of positive and negative uniform FN programming.

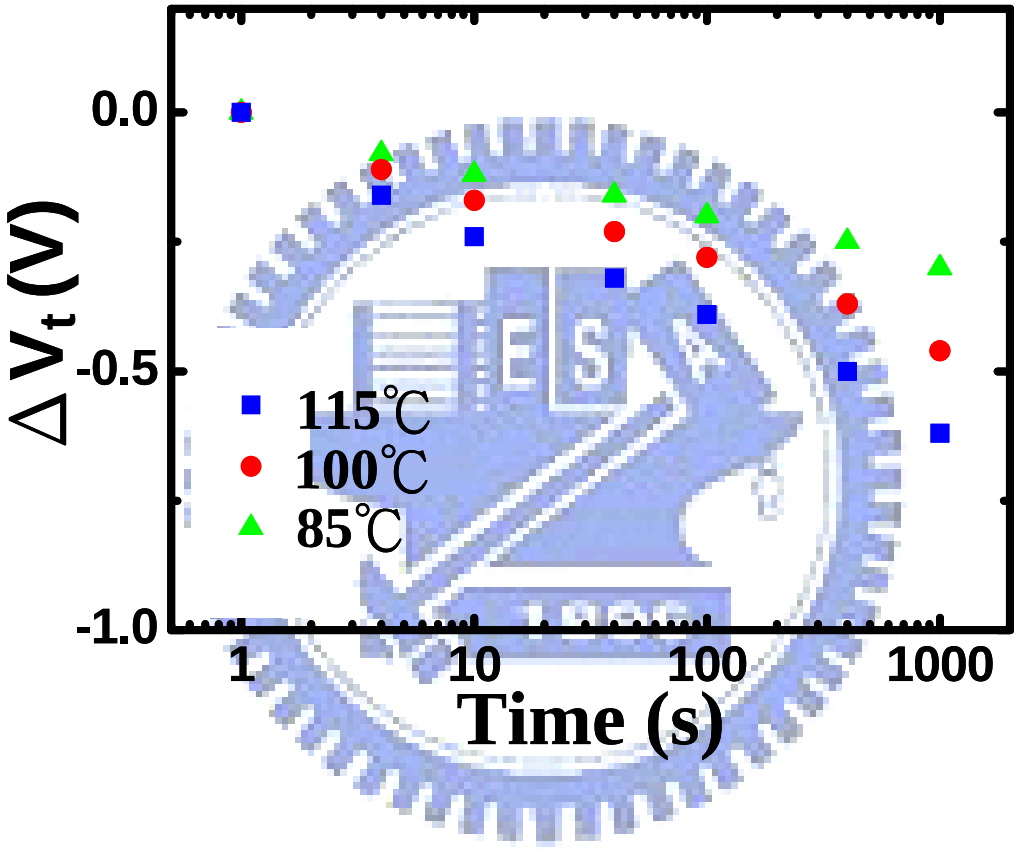


Fig. 4.5 Threshold voltage shift of HfO_2 nanocrystal flash memory measured at T=85℃, T=100℃ and T=115℃.

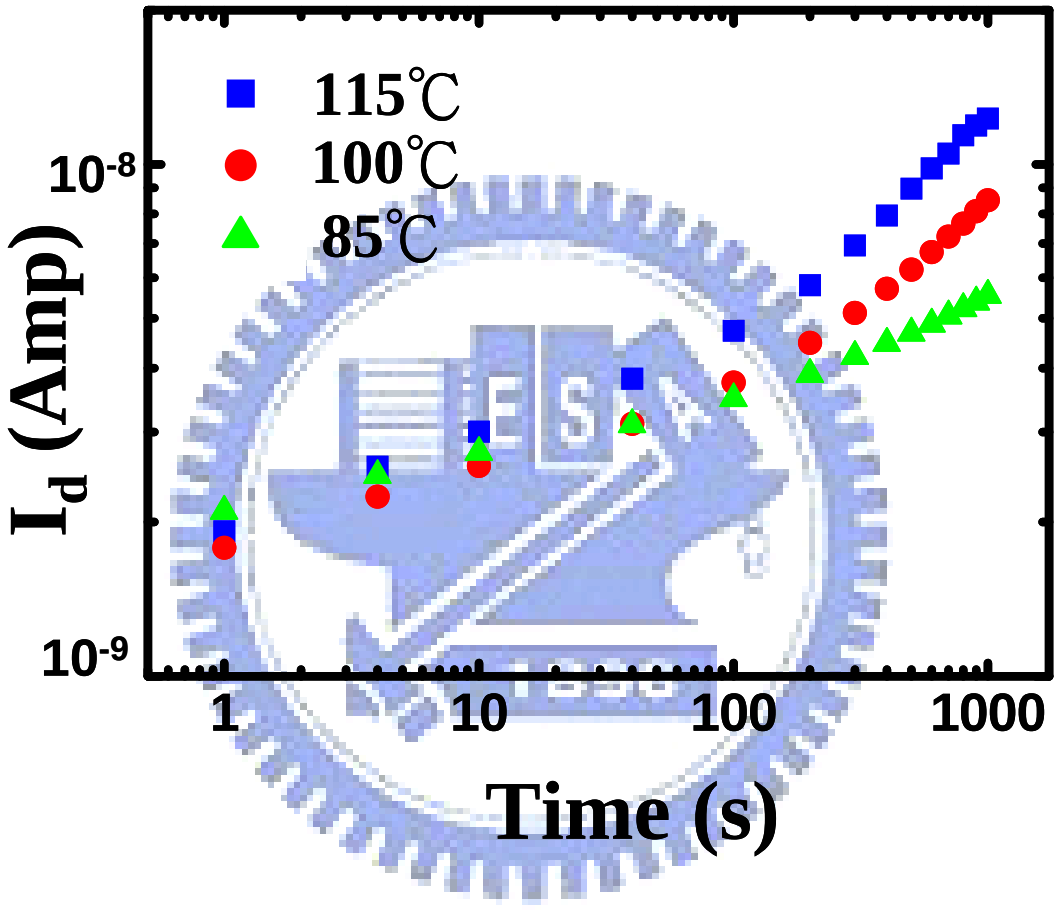


Fig. 4.6 Subthreshold current transients measured at T=85℃, T=100℃, and T=115℃.

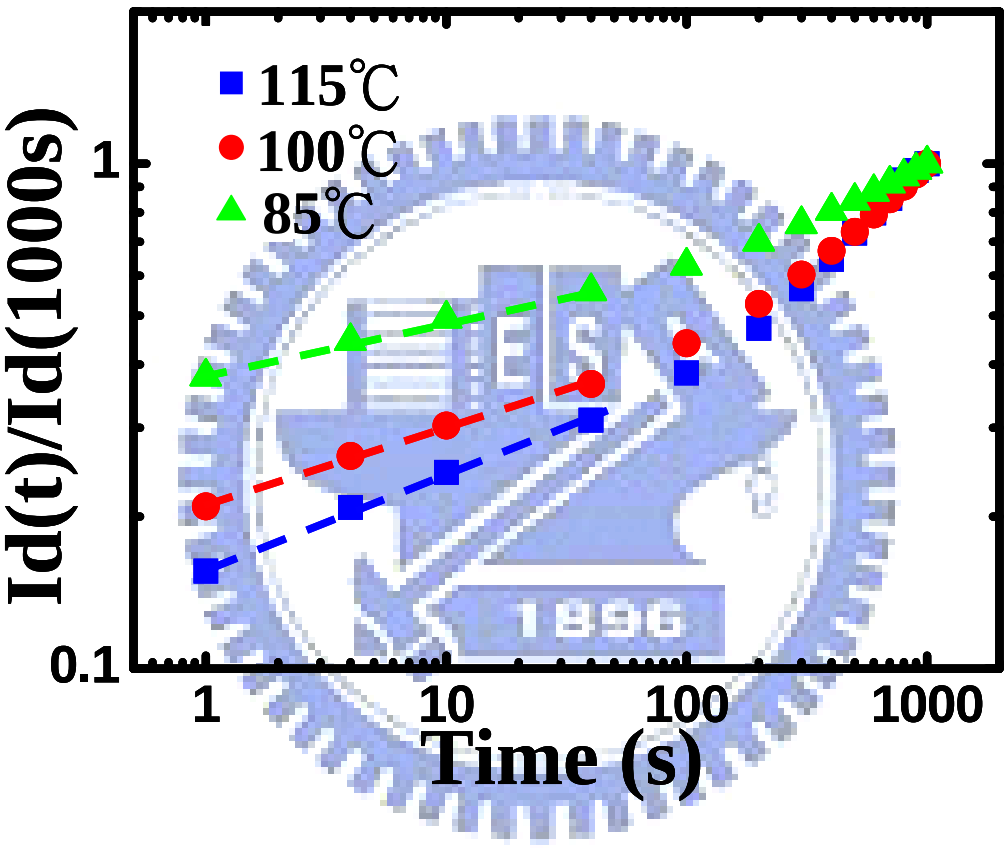


Fig. 4.7 Normalized subthreshold current transients measured at T=85℃, T=100℃, and T=115℃.

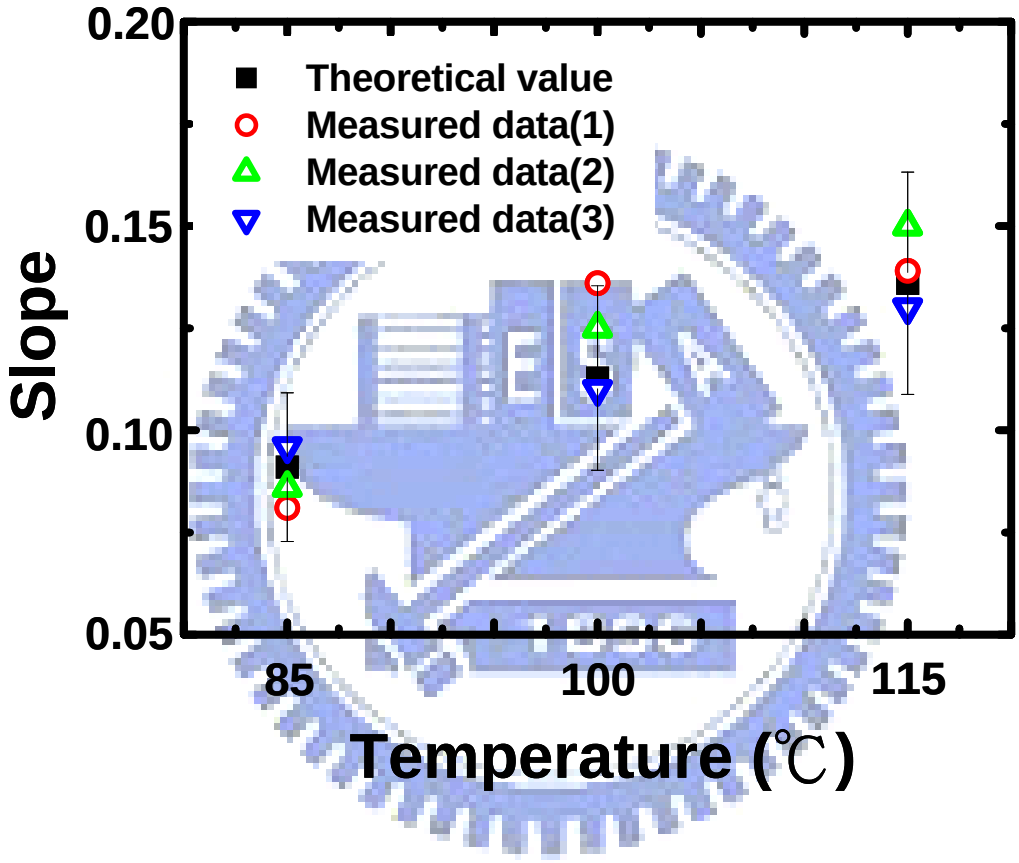
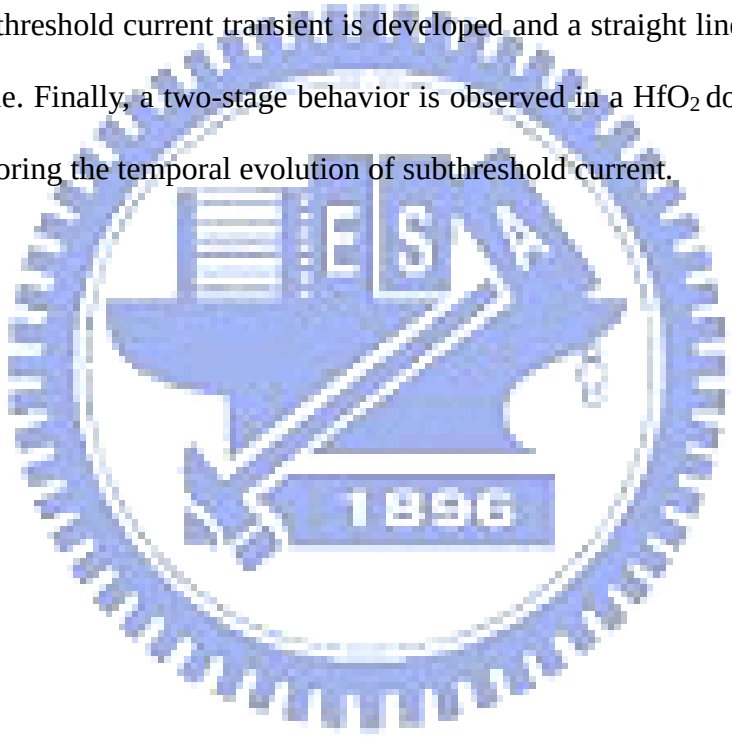


Fig. 4.8 Theoretical value and measured data of subthreshold current transient slope at T=85℃, T=100℃ and T=115℃.

Chapter 5

Conclusions

In this thesis, the data retention model of a HfO_2 dot flash memory cell is reported for the first time. The charge loss mechanism in a HfO_2 dot flash memory cell is investigated by characterizing the charge loss induced gate leakage current. The Frenkel-Poole emission model is not suitable for charge loss in the cell. A thermally activated tunneling front model is developed for a HfO_2 dot flash memory cell. Next, model of subthreshold current transient is developed and a straight line is obtained on a log-log scale. Finally, a two-stage behavior is observed in a HfO_2 dot flash memory cell by monitoring the temporal evolution of subthreshold current.



Reference

[1] R. Bez, E. Camerlenghi, A. Modelli, and A. Visconti, "Introduction to Flash memories," *Proc. IEEE,* vol. 91, pp. 489-502, 2003

[2] P. Pavan, R. Bez, P. Olivo, and E. Zanoni, "Flash memory cells – An overview," *Proc. IEEE,* vol. 85, pp. 1248-1271, 1997

[3] M.H. White, D. Adams and J. Bu, "On the Go with SONOS," *IEEE Circuits and Devices Magazine*, Vol. 16, pp.22-31, 2000.

[4] Y.N. Tan, W.K. Chim, W.K. Choi, M.S. Joo, T.H. Ng and B.J. Cho, "High-k HfAlO charge trapping layer in SONOS-type nonvolatile memory device for high speed operation," IEDM Tech. Dig., pp.889-892, 2004

[5] S. Tiwari, F. Rana, K. Chan, H. Hanafi, W. Chen, and D. Buchanan, "Volatile and nonvolatile memories in silicon with nano-crystal storage," IEDM Tech. Digest, pp. 521, 1995.

[6] Ya-Chin King, Tsu-Jae King, and Chenming Hu, "MOS Memory Using Germanium Nanocrystals Formed by Thermal Oxidation of $Si_{1-x}Ge_x$," *IEDM Tech. Dig.*, pp. 115-118, 1998.

[7] J. J. Lee, X. Wang, W. Bai, N. Lu, J. Liu, and D. L. Kwong, "Theoretical and Experimental Investigation of Si Nanocrystal Memory Device with HfO_2 High-k Tunneling Dielectric," *VLSI Tech. Dig.*, 2003.

[8] Jong Jin Lee, and Dim-Lee Kwong, "Metal Nanocrystal Memory With High-k Tunneling Barrier for Improved Data Retention," *IEEE Trans. Electron Devices*, vol. 52, pp. 507-511, April, 2005.

[9] Y. H. Lin, C. H. Chien, C. T. Lin, C. W. Chen, C. Y. Chang, T. F. Lei, "High performance multi-bit nonvolatile HfO_2 nanocrystal memory using spinodal phase separation of hafnium silicate," *IEDM Tech. Dig.*, pp. 1080-1082, 2004.

[10] H. G. Yang, Y. Shi, L. Pu, R. Zhang, B. Shen, P. Han, S. L. Gu, Y. D. Zheng, "Nonvolatile memory based on Ge/Si hetero-nanocrystals," Applied Surface Science 224, pp. 394-398, 2004

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

[12] S.H. Gu, C.W. Hsu, Tahui Wang, W.P. Lu, Y.H. Joseph Ku, and Chih-Yuan Lu, "Numerical Simulation of Bottom Oxide Thickness Effect on Charge Retention in SONOS Flash Memory Cells," *IEEE Trans. Electron Devices*, vol. 54, pp. 90-97, 2007.

[13] D. J. Dumin and J. Maddux, "Correlation of stress-induced leakage current in thin oxides with trap generation inside the oxides," *IEEE Trans. Electron Devices*, vol. 40, pp. 986-993, 1993.

[14] T. Wang, W. J. Tsai, M. I. Liu, S. H. Gu, C. T. Chan, C. C. Yen, N. K. Zous, T. C. Lu, S. Pan, and C. Y. Lu, "Reliability models of data retention and read-disturb in 2-bit nitride storage flash memory cells," *IEDM Tech. Digest*, pp. 169-172, 2003.

[15] M. Kato, N. Miyamoto, H. Kume, A. Satoh, T. Adachi, M. Ushiyama, and K. Kimura, "Read-disturb degradation mechanisms due to electron detrapping in the tunnel oxide for low-voltage flash memories," in *IEDM Tech. Dig.*, 1994, pp. 45-48.

[16] B. El-Kareh and R. J. Bombard, *Introduction to VLSI Silicon Devices.* Boston, MA: Kluwer, 1986, p. 461.

[17] I. Lundsorm and C. Svensson, "Tunneling to traps in insulator," *J. Appl. Phys.*, vol. 43, pp. 5045-5047, 1972

[18] R. S. Scott, N. A. Dumin, T. W. Hughes, D. J. Dumin, and B. T. Moore, "Properties of high voltage stress generated traps in thin silicon oxides," in *Proc. Int. Reliab. Phys. Symp*., 1995, pp. 131-141.

[19] Ya-Chin King, Tsu-Jae King, Chenming Hu, "Charge-Trap Memory Device Fabricated by Oxidation of $Si_{1-x}Ge_x$," *IEEE Trans. Electron Devices*, vol. 48, pp. 696-700, April 2001.