國 立 交 通 大 學

電子工程學系電子研究所

博 士 論 文

互補式金氧半電晶體之金屬閘極材料與技術研究

A Study on the Metal Gate Materials and Technology for the CMOS Application

研 究 生 ：黃誌鋒

指導教授 ：崔秉鉞 博士

中華民國 九十五 年 十 月

互補式金氧半電晶體之金屬閘極材料與技術研究

A Study on the Metal Gate Materials and Technology for the CMOS Application

研 究 生 ：黃誌鋒　　Student : Chih-Feng Huang

指導教授 ：崔秉鉞 博士　　Advisor : Dr. Bing-Yue Tsui

國立交通大學

電子工程學系　電子研究所

博士論文

A Dissertation

Submitted to Department of Electronics Engineering
and Institute of Electronics
College of Electrical Engineering and Computer Science
National Chiao-Tung University
In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in Electronics Engineering
2006
Hsinchu, Taiwan, Republic of China

中華民國 九十五 年 十 月

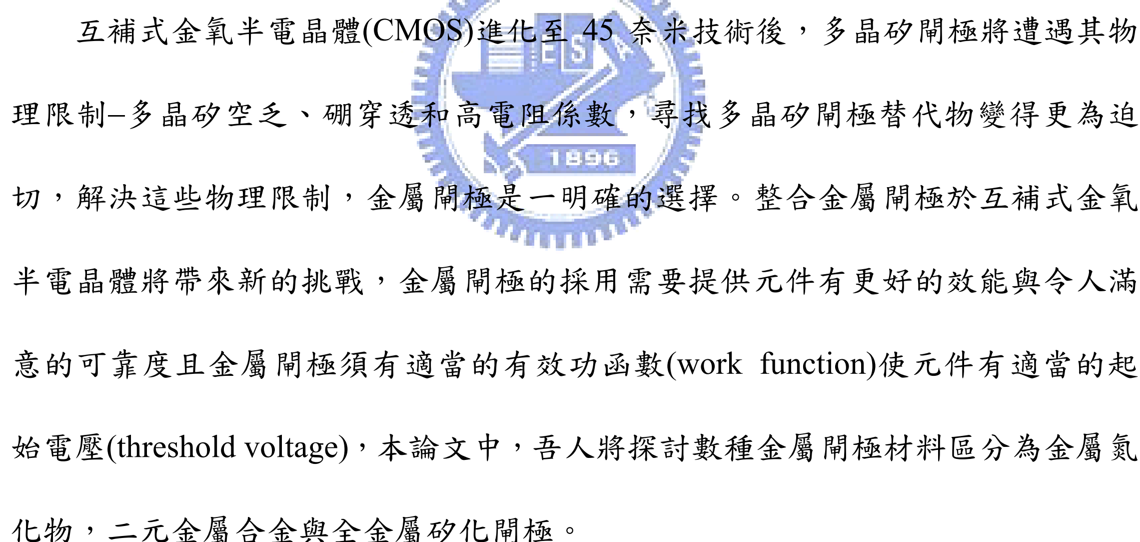
互補式金氧半電晶體之金屬閘極材料與技術研究

研究生：黃誌鋒　　　　　　　　　　　　指導教授：崔秉鉞 博士

國立交通大學

電子工程學系電子研究所

摘要



互補式金氧半電晶體(CMOS)進化至 45 奈米技術後，多晶矽閘極將遭遇其物理限制–多晶矽空乏、硼穿透和高電阻係數，尋找多晶矽閘極替代物變得更為迫切，解決這些物理限制，金屬閘極是一明確的選擇。整合金屬閘極於互補式金氧半電晶體將帶來新的挑戰，金屬閘極的採用需要提供元件有更好的效能與令人滿意的可靠度且金屬閘極須有適當的有效功函數(work function)使元件有適當的起始電壓(threshold voltage)，本論文中，吾人將探討數種金屬閘極材料區分為金屬氮化物，二元金屬合金與全金屬矽化閘極。

首先，吾人將探討氮化鉬(MoN_x)閘極，利用反應式濺鍍沉積氮化鉬薄膜，其氮含量原子百分比為 46-59，主要晶相為 MoN，隨著氮含量的增加，其微結構趨近非結晶且阻值增加，氮化鉬有效功函數亦隨氮含量增加且逐漸飽和於矽價帶，在氧化鉿(HfO_2)上沒有費米栓(Fermi-level pinning)效應，結果顯示氮化鉬閘極是 p-型金氧半場效電晶體很好的選擇。

其次，另一耐火金屬鎢經氮化後調變有效功函數，氮化鎢含氮元素高於原子百分比 44 其晶相為 WN 且此晶相穩定度高達 800 ^{o}C，縱使氮濃度高達原子百分比 61，並無更高晶相產生。氮化鎢中過多的氮元素在溫度低於 766 ^{o}C 便會析出，且過多的氮會導致氮化鎢有效功函數下降，在氧化鉿上會有輕微的費米栓效應，這種情形下，氮化鎢可應用於 p-型全空乏絕緣層上矽(SOI)金氧半場效電晶體元件，但不適合當塊材(bulk) p-型金氧半場效電晶體閘極。

金屬氮化物的高阻值迫使吾人思考堆疊(stack)結構當閘極，探討銅/氮化鉭(Cu/TaN_x)堆疊，其中低電阻值層(銅層)適合當主要的導電材料，底層(氮化鉭層)用以當起始電壓控制材料。氮化鉭中氮含量原子百分比為 23-39，且薄膜幾乎為非晶結構，熱穩定度高達 800 ^{o}C，氮化鉭的有效功函數約為 4.31-4.38 eV，氮成分調變有效功函數能力低於 70 meV，元件平帶電壓均值隨著溫度增高而下降且平帶電壓標準差增大，雖然在高溫時晶相變化、晶粒成長與銅污染會貢獻平帶電壓不穩定性，在 600 ^{o}C 熱應力致氧化層電荷主導平帶電壓下降與標準差，銅/氮化鉭堆疊閘極可使用於 n-型金氧半場效電晶體銅/氮化鉭，其製程最高溫度需控制在 500 ^{o}C.

雖然金屬氮化物具高溫熱穩定性，但其有效功函數範圍並不夠大，吾人展示一種有效的功函數調變方法，利用二元金屬合金而非金屬半化物，選擇適當的金屬成分與組成，可獲得任意功函數之二元金屬合金。調變鉭基材合金(鉭-鈦和鉭-鉑)其有效功函數可連續分部於 4.16-5.2 eV， 如此大範圍功函數使得該金屬合金適用於互補式金氧半場效電晶體，鉭與鉑經混合後，其核心電子束縛能改變，佐證合金功函數可依金屬合金系統調整，而且鉭膜中混合鉑元素會造成結晶變差且形成合金晶相，經參透式電子顯微鏡確認鉭-鉑合金薄膜中晶粒為奈米尺寸。

最後，一種新方法(雜質植入矽化物)用以實現全金屬矽化閘極有效功函數調

變，全金屬矽化閘極容易整合於傳統金氧半場效電晶體製程，且利用雜質推積於全矽化物與底層絕緣材料介面，擁有大範圍的有效功函數調變。半導體矽經全金屬矽化後，植入 BF^{+}與 P^{+}雜質接著作退火可有效調變全金屬矽化閘極功函數。但當底層絕緣層材料為氧化鉿，會產生費米栓效應，限制功函數調變，一層薄的二氧化矽介面層可減緩費米栓效應。雖然矽化鎳全矽化閘極在二氧化矽尚可穩定到 600 ^{o}C，但長時間處於 600 ^{o}C，熱應力與雜質擴散會導致氧化層劣化，全矽化後製程溫度建議低於 500 ^{o}C 以緩和氧化層劣化。

總結，金屬氮化物由於功函數調變範圍過小，僅能使用於雙金屬閘極組合，儘管金屬合金擁有大範圍的功函數，但很難利用單一金屬閘極製程整合不同成份合金於互補式金氧半電晶體，經 N-型與 P-型參雜全金屬矽化閘極製程最有可能符合單金屬雙功函數組合，但金屬矽化前多晶矽與高介電材料反應產生費米栓效應限制了運用全金屬矽化閘極，到目前為止並沒有完美的金屬閘極材料與製程供選擇，標準互補式金氧半電晶體金屬閘極製程尚須詳細考慮雙金屬閘極或金屬全矽化閘極。

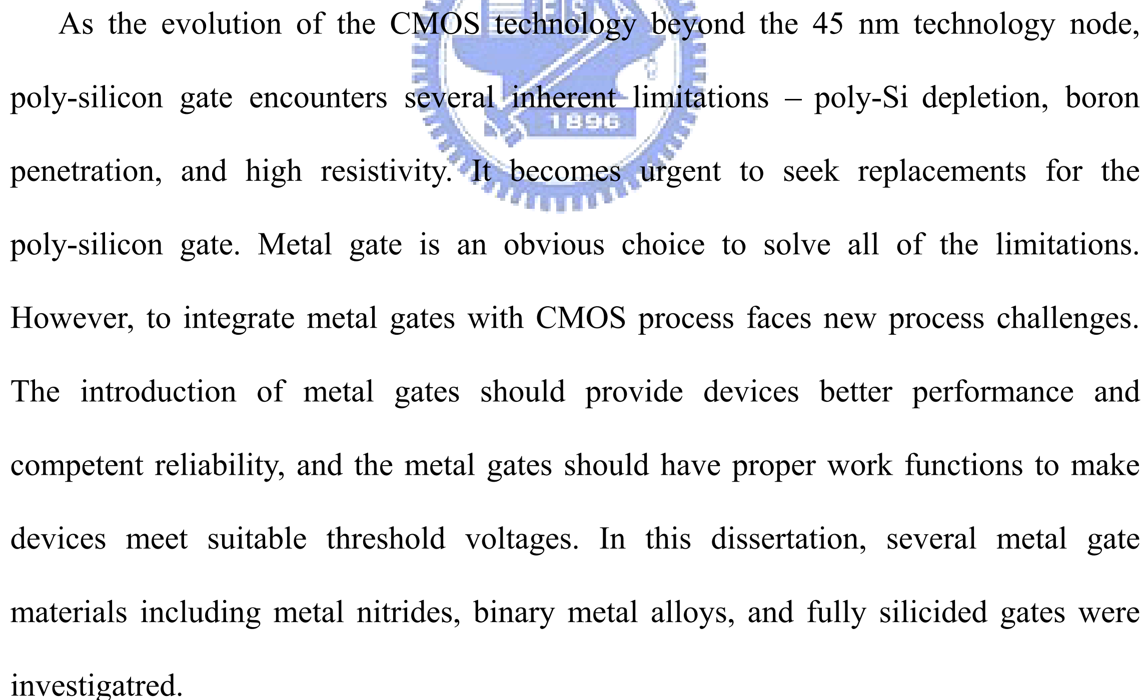
A study on the Metal Gate Materials and Technology for the CMOS Application

Student: Chih-Feng Huang **Advisor: Dr. Bing-Yue Tsui**

Department of Electronics Engineering & Institute of Electronics

National Chiao-Tung University

Abstract



As the evolution of the CMOS technology beyond the 45 nm technology node, poly-silicon gate encounters several inherent limitations – poly-Si depletion, boron penetration, and high resistivity. It becomes urgent to seek replacements for the poly-silicon gate. Metal gate is an obvious choice to solve all of the limitations. However, to integrate metal gates with CMOS process faces new process challenges. The introduction of metal gates should provide devices better performance and competent reliability, and the metal gates should have proper work functions to make devices meet suitable threshold voltages. In this dissertation, several metal gate materials including metal nitrides, binary metal alloys, and fully silicided gates were investigatred.

The MoN_x gates were studied at first. The main phase of the MoN_x films with the 46 – 59 at. % of nitrogen is MoN. As the N/Mo ratio increases, the microstructure of MoN_x film tends to be amorphous-like and the resistivity increases. The work function of MoN_x

increases with the increase of nitrogen content and tends to saturate at the valence band of Si. No Fermi-level pinning effect is observed on HfO_2 film. These results indicate that MoN is a good candidate of gate electrode for PMOSFETs.

The WN_x films with nitrogen higher than 44 % atomic ratio has a main phase of WN, and the WN phase is stable up to 800 °C. The higher order WN_x phase does not form even if the nitrogen concentration is as high as 61%. The excess nitrogen atoms in WN_x films desorbs at temperatures below 766 °C. The excess nitrogen in WN_x films can cause the effective work function lowering. Weak Fermi-level pinning effect is observed on HfO_2 film. In this case, WN_x/HfO_2 gate stack can be applied to p-type fully-depletion SOI devices but WN_x is not suitable to be meal gate of bulk p-type MOSFTEs.

The high resistance of metal nitride forces us to consider the stack structure as gate electrodes. We investigate the Cu/TaN_x stack of witch the Cu layer serves as the major conduction material and the TaN_x layer serves as the threshold voltage control material. As the nitrogen content of TaN_x varies in the range 23 – 39 at. %, TaN_x films are almost amorphous and are thermally stable up to 800 ℃. The work function of TaN_x is about 4.31-4.38 eV and the modulation range is less than 70meV. The mean value of flatband voltage decreases and the deviation of the flatband voltage increases with the increase of the annealing temperature. Although phase change, grain growth, and Cu contamination contribute the instability of flat band voltage at high temperature, thermal stress-induced oxide charges dominate these thermal instability at 600 ℃. The Therefore, Cu/TaN_x stack gate electrode can be used for the nMOSFETs, and the maximum process temperature following gate electrode deposition should be 500 ℃.

Although metal nitrides are thermally stable, the work function modulation range is not large enough. We demonstrate an effective method of work function modulation by the binary metal alloys instead of metal nitrides. Any work function can be obtained by

the binary metallic alloy system with a proper composition of high and low work function metals. Binding energies of core level electrons of Ta and Pt change due to the intermixing of Ta and Pt, which is the mechanism that work function of alloys are adjusted in the metallic alloy systems. The work function of the Ta-based metal alloys (Ta-Ti and Ta-Pt) can be modulated from 4.16 eV to 5.2 eV continuously. Such a wide range work function modulation makes them suitable for CMOS applications. Moreover, incorporation of Pt in Ta film induces poor crystallization and a compound phase of Ta-Pt alloys. Transmission electron microscopy analysis confirmed that grain sizes of Ta-Pt alloys were nano-scale.

Finally, a new method, implant-to-silicide (ITS), is used to realize effective work function ($\Phi_{m,eff}$) adjustments of fully silicided (FUSI) gates. FUSI gate is easily to integrate with the conventional CMOS process and have a wide range modulation of effective work function by impurities pile-up at the interface between silicide and gate dielectric. The $\Phi_{m,eff}$ of NiSi FUSI gates on SiO_2 can be tuned by incorporating BF_2^+ or P^+ dopants after silicidation. Nevertheless, the Fermi-level pinning effect is observed in the NiSi/HfO_2 gate which limits the $\Phi_{m,eff}$ adjustment. A thin SiO_2 interfacial layer between HfO_2 and FUSI gates can reduce the Fermi-level pinning effect. A NiSi FUSI gate on SiO_2 is thermally stable up to 600 °C. The thermal stress and impurity diffusion after a prolonged 600 °C annealing degrade the oxide integrity. The temperature of the post-silicidation process should be as low as 500 ^{o}C.

In summary, the small work function range of single metal nitrides restricts that the metal nitrides can only been used for the dual metal scheme. Although the metallic alloys reveal wide range of work function, the integration of different composition of metallic alloys in CMOS application is difficult by the single metal process. Fully silicided gates doped by the N-type and P-type impurities are the most promising process for the single

metal dual work function scheme. However, the interaction between poly-Si and high-k materials before silicidation induces Fermi-level pinning and will retard the application of fully silicided gates. Up to date, there are still no perfect choice of metal gate materials and processes. The standard metal gates for the CMOS are still under consideration to the dual metal scheme or fully silicided gates.

誌謝

本論文能夠順利完成，首先要特別感謝我的指導教授 崔秉鉞博士。感謝老師在這幾年的博士生涯中給予實驗、研究與論文上充分且詳實的指導與教誨，並時時的鼓勵與支持，尤其在實驗上遭遇困難時老師的關懷讓學生在困境中生信心，老師認真的研究態度與親切的處事風範，令學生受益良多。

其次感謝國家奈米實驗室與交通大學奈米中心提供優良研究設備，特別感謝奈米中心的 林素珠、徐秀鑾、黃月美與陳聯珠小姐在實驗上的幫忙，並感謝奈米實驗室中的工作同仁不厭其煩的給予幫助，使我得以順利完成實驗。諸位口試委員的指正與建議，也在此表示特別感謝。

衷心的感謝實驗室的學長以及一起參與實驗的學長們，特別是徐偉成、林成利、吳振誠等學長於實驗上的指導及幫忙，另外要感謝實驗室同學們:國龍、家彬、偉豪、修維、敏晟、鎮吉、志民、天才、旻達、堅立、建翔、智勛、逸璿、柏智、立峰、宜澤、季霈、孟漢等。由於有了你們相伴，使我渡過此漫長研究生活。再次感謝所有幫助過我、關心過我的學長及朋友們……。

最後，我特別要感謝我的父母 黃資良先生與 李玉鳳女士以及我的兄弟姊妹吉泉、水蓮、麗真、振卿，感謝你們多年來的栽培，默默付出與支持，使我得以全心全力地投入學業。另外特別感謝女友 宜樺在研究生涯的最後年度裡，給我莫大的照顧與關心。

Contents

Abstract (Chinese) i

Abstract (English) iv

Acknowledgement (Chinese) viii

Contents ix

Table Captions xiii

Figure Captions xiv

Chapter 1

Introduction..1

1.1 Historical Perspective 1

1.2 Metal Gate Electrodes..5

1.2.1 Metal Gate Materials..5

1.2.2 Resistivity 7

1.2.3 Work Function..7

1.2.4 Metal-Gate Structures 8

1.3 Work Function Engineering..9

1.3.1 Structure Modification 9

1.3.2 Interface Treatment 10

1.4 Fermi-Level Pinning 11

1.5 Process Integration..12

1.6 Overview of Dissertation 13

References..15

Chapter 2

Experimental Methods 33

2.1 Introduction 33

2.2 Electrical Characterizations 34

2.2.1 Basic Electrical Characterizations and Parameters 34

2.2.2 Work Function Extraction 35

2.3 Inelastic Electron Tunneling Spectroscopy 39

2.3.1 Basic Principle of IETS 39

2.3.2 Measurement Method 42

2.3.3 Signal Analysis 43

References 45

Chapter 3

Molybdenum Nitride Gates 57

3.1 Introduction 57

3.2 Process Flow 58

3.3 Results and Discussions 59

3.3.1 Physical Properties of MoN_x films 59

3.3.2 Effective Work Function on SiO_2 and HfO_2 60

3.3.3 Electrical Effect of Thermal Annealing 63

3.4 Summaries and Conclusions 64

References 65

Chapter 4

Tungsten Nitride Gates 76

4.1 Introduction 76

4.2 Process Flow 77

4.3 Results and Discussions 78

4.3.1 Physical Properties of N-rich WN_x Films 78

4.3.2 Effective Work Function on SiO_2 and HfO_2 81

4.3.3 Electrical Effect of Thermal Annealing 85

4.4 Summaries and Conclusions 86

References 88

Chapter 5

Cu/Tantalum Nitride Stack Gates 102

5.1 Introduction 102

5.2 Process Flow 103

5.3 Results and Discussions 104

5.3.1 Physical Properties of TaN_x Films 104

5.3.2 Effective Work Function on SiO_2 106

5.3.3 Electrical Effect of Thermal Annealing 108

5.4 Summaries and Conclusions 109

References 111

Chapter 6

Ta-Pt Metallic Alloy Gates 122

6.1 Introduction 122

6.2 Process Flow 123

6.3 Results and Discussions 124

6.3.1 Work Function Modulation 124

6.3.2 Physical Characteristic of Ta-rich Ta-Pt Alloys 125

6.3.3 Thermal Stability of the Ta-Pt Alloy on SiO_2 and HfO_2 128

6.3.4 Effect of Impurity Incorporation to the Effective Work Function 131

6.4 Summaries and Conclusions 133

References 135

Chapter 7

Fully Ni-Silicided Gates 154

7.1 Introduction 154

7.2 Process Flow 155

7.3 Results and Discussions 156

7.3.1 Thermal Stability of the Fully-Silicided NiSi Gates 156

7.3.2 Effect of Work Function due to Impurity Redistribution 158

7.3.3 Effects of Silicidation and Post Annealing on Electrical Characteristics 159

7.4 Summaries and Conclusions 161

References 163

Chapter 8

Conclusions and Recommendations 178

8.1 Conclusions 178

8.2 Recommendation for Future Works 184

References 188

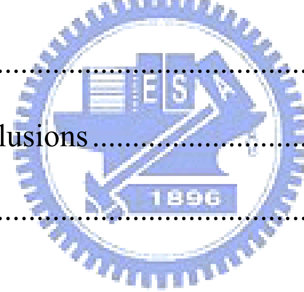


Table Captions

Chapter 1

Table 1-1 Desired properties of the metal gates for CMOS..22

Table 1-2 The sheet resistance and thickness of gate electrodes predicted in ITRS roadmap..23

Table 1-3 Charge neutral level and S parameter for several dielecitrics......................24

Chapter 3

Table 3-1 MoN_x sample ID and some basic characteristics. 68

Chapter 5

Table 5-1 Mean values and standard deviations of retention time of Cu/TaN-2 MOS capacitors under transient capacitance measurement. 113

Chapter 6

Table 6-1 Some conditions of binary metal alloy samples and their compositions...139

Table 6-2 Material characteristics of Ta, Ta-Pt (A2), and Pt films. 140

Chapter 6

Table 7-1 Some conditions of the FUSI test samples 167

Chapter 8

Table 8-1 Comparison of metal gates. 190

Figure Captions

Chapter 1

Fig. 1-1 The evolution of MOSFET gate electrodes with the innovation of technologies. 25

Fig. 1-2 A schematic cross-section of a sub-45 nm metal gate CMOS structure, where the novel materials and process-integration solution are indicated. 26

Fig. 1-3 The common metal elements in periodic table. The elements were marked as p, n, and m according to its work function near the valence band, conduction band, and mid-gap of silicon energy band, respectively. 27

Fig. 1-4 The threshold voltages of (a) bulk devices and (b) SOI devices versus channel concentrations. The estimation of threshold voltage was evaluated with the simulation of long channel and uniform charnel dopant devices by Avant! MEDICI. The dependence of oxide thickness or thickness of silicon body was also taken an account. 28

Fig. 1-5 A structure of the stack metal-gate electrode. The bottom layer serves as threshold control and dielectric-contact layer; the top layer severs as conduction layer connected with metallic interconnect. 29

Fig. 1-6 Schematic methods of work function modulation divided into two categories: structure modification and interface treatment. 30

Fig. 1-7 The critical Fermi level of metal gate on dielectric. The effective work function of metal on dielectric is the same as in vacuum while the interface is perfect. The effective work function is fixed to the delectric-charge-neutral level while the interface-state density is high. The interface states are caused by both

intrinsic states (MIGS, ViGS) and extrinsic states (defect levels)...............31

Fig. 1-8 Schematic process flows of metal gate MOSFETs including (a) gate-first process; (b) gate-last process, and (c) fully-silicided process......................32

Chapter 2

Fig. 2-1 High-frequency (HF) and quasi-frequency (QS) capacitance-voltage (CV) characteristic curves of metal-oxide-semiconductoe capacitance as well as the $1/C^2$-V plot corresponding to the HFCV. Some basic parameters are extracted from these plots according to the described formula. 47

Fig. 2-2 (a)The I-V characteristic curve, and (b) the relative Fowler-Nordheim plot as the electrons are injected from the gate electrode. The inset in (b) shows the band diagram. The abrupt jump at V_{BD} is referred to oxide breakdown voltage. 48

Fig. 2-3 A example for the extraction of the effective work function by fitting the HFCV with a theoretical curve...49

Fig. 2-4 An example for the extraction of the effective work function by a V_{fb}-EOT plot extrapolated at zero of EOT. 50

Fig. 2-5 A schematic plot of charge distribution in the gate/HfO_2/SiO_2/Si structure. The bulk charges in SiO_2 and HfO_2 are assumed to be a uniform distribution....51

Fig. 2-6 (a) A schematic energy diagram for tunneling between normal metals A and B at a voltage bias V which is the threshold voltage to trigger the inelastic tunneling due to the interaction of electron and vibration mode of energy eV. (b) The relative tunneling current of metal A/insulator/metal B structure at 0 K and the corresponding first derivative and second derivative plots. 52

Fig. 2-7 A schematic diagram of the coordinates used to determined interaction energies of a tunneling and a molecule positioned near an insulator-metal A interface.

....53

Fig. 2-8 (a) A circuit schematic diagram of mixer in the IETS system; (b) the schematic structure of the IETS test sample. 54

Fig. 2-9 (a) Schematic diagram and (b) the configuration of the IETS system consisted of a DC voltage source (Agilent 4140), a lock-in amplifier with an AC source (SR 830), DC voltage meter (Agilent 34420A), a mixer circuit, a cryostat and controller. 55

Fig. 2-10 (a) The primary and smoothed 2ω signal measured by the lock-in amplifier; (b) the IETS signal extracted from the smoothed 2ω signal after subtraction of a polynomial background. 56

Chapter 3

Fig. 3-1 Rutherford-back-scattering (RBS) spectra of MoN_x films on graphite substrate. 69

Fig. 3-2 X-ray diffraction (XRD) spectra of (a) MoN-1, (b) MoN-2, and (c) MoN-3 films after annealing at different temperatures 70

Fig. 3-3 Flat-band voltage (V_{fb}) versus equivalent oxide thickness (EOT) plot of (a) $MoN_x/SiO_2/Si$ structure and (b) $MoN_x/HfO_2/SiO_2/Si$ structure after annealing at 500 °C. 71

Fig. 3-4 (a) Effective work function of MoN_x film on SiO_2 layer and (b) quasi-effective work function of MoN_x film on HfO_2 layer versus annealing temperature. . 72

Fig. 3-5 The electron diffraction pattern of the MoN-0 film after annealing at 600 °C. 73

Fig. 3-6 Capacitance-Voltage curves of MoN_x (60 nm)/SiO_2 (40 nm)/Si structure of (a)MoN-0 and (b) MoN-3 films after annealing at different temperatures...74

Fig. 3-7 Capacitance-voltage curves of MoN_x (60 nm)/HfO_2 (5 nm)/SiO_2 (40 nm)/Si

structure of (a) MoN-0 and (b) MoN-3 films after annealing at different temperatures. 75

Chapter 4

Fig. 4-1 Nitrogen concentration and relative resistivity of the WN_x films as a function of the N_2/Ar gas flow ratio during the sputtering deposition. 92

Fig. 4-2 Plane view SEM micrographs of the WN-2 films after annealing at (a) 500 °C and (b) 800 °C in a rapid thermal annealing system. The temperature rising rate is 90 °C/sec. 93

Fig. 4-3 Thermal desorption spectrum of the as-deposited WN-2 film. 94

Fig. 4-4 X-ray photoelectron spectroscopic spectra of the (a) O 1s, (b) Hf 4f, and (c) N 1s of the WN-1 film after annealing at 500 °C for 30 min. in N_2 ambient. The de-convolution of O 1s is also given. 95

Fig. 4-5 X-ray diffraction patterns of the WN_x films after annealing at different temperatures. All of the as-deposited films are amorphous. 96

Fig. 4-6 Auger electron spectrometer depth profiles of the as-deposited and 800 °C annealed WN-2 film. 97

Fig. 4-7 V_{fb} vs EOT plots of (a) $WN_x/SiO_2/Si$ structure and (b) $WN_x/HfO_2/SiO_2/Si$ structure after annealing at 600 ^oC. 98

Fig. 4-8 (a) Effective work function of WN_x film on SiO_2 layer and (b) quasi-effective work function of WN_x film on HfO_2 layer versus annealing temperature. ...99

Fig. 4-9 Cumulative failure time of the (a) WN_x/SiO_2 (5 nm)/Si capacitors under constant voltage stress at Vg=-6 V and (b) WN_x/HfO_2 (2 nm)/SiO_2 (2 nm)/Si capacitors under constant voltage stress at Vg=-4.5 V as a function of anneal temperature. 100

Fig. 4-10 The current-voltage characteristics of (a) WN_x/SiO_2 (5 nm)/Si capacitor and

(b) WN_x/HfO_2 (2 nm)/SiO_2 (2 nm)/Si capacitors. 101

Chapter 5

Fig. 5-1 Wild angle X-Ray Diffraction (XRD) spectrums of TaN-1, TaN-2, and TaN-3 films after annealing at various temperatures. 114

Fig. 5-2 Plane view Transmitted Electron Microscopy (TEM) micrographs of (a) TaN-1 film annealed at 400 ℃, (b) TaN-1 film annealed at 700 ℃, (c) TaN-2 film annealed at 400 ℃, and (d) TaN-2 film annealed at 700 ℃. 115

Fig. 5-3 Diffraction patters of (a) TaN-1 film annealed at 400 ℃, (b) TaN-1 film annealed at 700 ℃, (c) TaN-2 film annealed at 400 ℃, and (d) TaN-2 film annealed at 700 ℃. 116

Fig. 5-4 Variation of sheet resistance of the three-kinds of TaN_x films versus annealing temperature. 117

Fig. 5-5 Effective work function ($\Phi_{m,eff}$) of (a) Cu/TaN-1 sample, (b) Cu/TaN-2 sample, and (c) Cu/TaN-3 sample versus the annealing temperature. The effective work function was extracted from the C-V curves by a theoretical curve fitting. 118

Fig. 5-6 Capacitance-Voltage curves before and after bias-temperature stress (BTS) test at +1MV/cm and 150 ℃ for 60 min. of Cu/TaN-2 sample annealing at (a) 500 ℃, (b) 550 ℃, and (c) 600 ℃. 119

Fig. 5-7 Secondary ion mass spectroscopy (SIMS) depth profile of Cu of the 600 ℃ annealed Cu/TaN-2 sample. No BTS was performed before SIMS analysis and the Cu layer was removed with dilute HNO_3 solution. 120

Fig. 5-8 Interface state density (D_{it}) of (a) Cu/TaN-1 sample, (b) Cu/TaN-2 sample, and (c) Cu/TaN-3 sample versus the annealing temperature. 121

Chapter 6

Fig. 6-1 (a) The C-V characteristics of alloy samples annealed at 400 ^{o}C N_2 ambient for 30 min, as well as the N^+ poly, pure Ta, and Pt samples. The (b) Effective work function ($\Phi_{m,eff}$) of alloys extracted from the C-V curves by a theoretical C-V curve fitting. The $\Phi_{m,eff}$ approximates to $\Phi_{m,B} - x(\Phi_{m,A}-\Phi_{m,B})$ where x parameter counters for the concentration of element B in alloys and has the value between 0 and 1. 141

Fig. 6-2 Binding energies of (a) Ta atoms and (b) Pt atoms of the 800 °C annealed A2 samples detected by x-ray photoelectron spectroscopy (XPS) after Ar sputtering for various times. The binding energies of Ta and Pt in pure single metal films were also revealed as references. 142

Fig. 6-3 Binding energies of Ta atoms of the 400 °C and 800 °C annealed A2 samples detected by x-ray photoelectron spectroscopy (XPS) after Ar sputtering for 200 sec. 143

Fig. 6-4 Depth distribution of the Ta and Pt atoms of the 800 °C annealed A2 samples detected by Auger electron spectroscopy (AES). 144

Fig. 6-5 X-ray diffraction (XRD) patterns of the A2 samples after 400 °C, 500 °C, 600 °C, 700 °C, and 800 °C annealing for 30 min. The XRD patterns of pure Ta and Pt films annealed for 30min are also given as references. 145

Fig. 6-6 TEM plane view of the A1 sample annealed at 800 ^{o}C for 30 min. The inset is its relative TEM diffraction pattern. 146

Fig. 6-7 The inelastic electron tunneling spectra of Ta-Pt (A5)/SiO_2/degenerated Si samples annealed at 500 °C and 800 °C. 147

Fig. 6-8 Fowler-Nordheim plot of the Ta-Pt (A5)/SiO_2 (5.5 nm)/Si structure after annealing at 500 ^{o}C and 800 ^{o}C for 30 sec.The m^*/m_o=0.26 is used to

calculate the barrier height (Φ_B). The inset shows the relative C-V curves. 148

Fig. 6-9 The inelastic electron tunneling spectra of Ta-Pt (A5)/HfO_2/degenerated Si samples annealed at 500 °C and 800 °C. 149

Fig. 6-10 TEM photos of Ta-Pt (A5)/HfO_2/Si after annealing at 800 °C RTA which confirms that Ta-Pt on HfO_2 is chemically stable. 150

Fig. 6-11 The flat-band voltage shift (ΔV_{fb}) of (a) A6 samples and (b) A7 samples with various implantation conditions after 500 °C and 800 °C annealing for 30 sec. The ΔV_{fb} value is the difference between the flat-band voltage of the un-implanted and implanted samples. 151

Fig. 6-12 Depth distributions of the (a) arsenic and (b) boron implanted A7 sample before and after 800 °C annealing for 30 sec detected by SIMS. The energies of implantation are 50 KeV and 30 KeV for As^+ and BF_2^+, respectively, and both dosages are 5×10^{15} cm^{-2}. 152

Fig. 6-13 Capacitance-Voltage characteristics of the 800 °C annealed A7 samples with and without ion implantation. 153

Chapter 7

Fig. 7-1 X-ray diffraction patterns of nickel silicide. Curve A is recorded after 300 °C silicidation in a furnace (FA) for 1 hr; curve B is recorded after condition A plus 600 °C RTA for 30 sec; curve C is recorded after condition B plus 600 °C annealing for 4 hr in a furnace. 168

Fig. 7-2 Sheet resistance of NiSi film vs annealing time at 600 °C. The measured NiSi film is on the SiO_2 layer. 169

Fig. 7-3 TEM cross-sectional views of (a) ITS sample and (b) pre-doped sample. 170

Fig. 7-4 (a) Scheme of detection by X-ray photoelectron spectrometer. Area A is

exposed after removing NiSi, and B is exposed after removing the interfacial layer. The electron binding energies of detected areas A and B of (b) Si 2p orbital, (c) O 1s orbital, and (d) Hf 4f orbital are shown. 171

Fig. 7-5 TEM cross-sectional views of (a) as-deposited amorphous Si/ $HfO_2/SiO_2/Si$, and (b) sample (P+ poly-Si sample) implanted with BF_2 and annealed at 900°C for 30 sec in N2 atmosphere. 172

Fig. 7-6 Effective work function ($\Phi_{m,eff}$) of doped and undoped NiSi FUSI gate on SiO_2 and HfO_2. $\Phi_{m,eff}$ of the P^+ polysilicon gate on HfO_2 is also included. 173

Fig. 7-7 Depth profiles of (a) boron and (b) phosphorous dopants in NiSi FUSI gates detected by secondary ion mass spectrometry. 174

Fig. 7-8 C-V characteristics of NiSi FUSI gate capacitor with SiO_2 gate dielectric after bias-temperature stress of ±2 MV/cm at 200 °C for 30 min. 175

Fig. 7-9 (a) Cumulative distribution of interface state density of NiSi FUSI gate capacitors; NiSi gates were formed at 500 or 600 °C. (b) The relative quasi-C-V curves of NiSi FUSI gate capacitance and the high frequency curve of sample C. 176

Fig. 7-10 Cumulative distribution of (a) flat-band voltage (V_{fb}) and (b) hysteresis of NiSi FUSI gate capacitors after prolonged 600 °C annealing. 177

Chapter 8

Fig. 8-1 Work function of Ta, W, Mo, Ni, their silicides and nitrides. The χ is the electronegativity. 191

Chapter 1

Introduction

1.1 Historical Perspective

In December of 1947, the first point contact transistor was constructed by John Bardeen and Walter Brattain of Bell Telephone Laboratories [1]. In 1959, Jack S. Kilby of Texas Instrument Incorporate and Robert N. Noyce of Fairchlid Semiconductor Corporation invented the solid integrated circuit [2, 3]. In 1960, a practical metal-oxide-semiconductor field-effect transistor (MOSFET) was first announced [4]. From this moment, the MOS technology speedily grows along the Moore's law. To understand the emergence and revolution of modern metal gate technology, the historical perspective of gate electrodes is introduced. In 1960s, gate electrodes of MOS transistors were made from aluminum deposited after the source and drain region were doped. The aluminum had to overlap onto the source/drain to insure that a continuous channel from source to drain was formed when the gate was biased to turn on the transistor. The overlap between gate and source/drain regions causes a significant capacitance to slow the circuit speed. To reduce the overlap capacitance, a self-align gate (SAG) process was proposed by Robert Bower in 1966, where the gate electrode was used as the implant mask. Dopants were implanted to form self-aligned source and drain regions with minimum overlap to gate electrodes [5]. This method drastically reduces the overlap capacitance to speed the device turn-on. However, crystalline defects were induced during implantation. A high temperature annealing, higher than

1000ºC, is the only way to repair these defects. However, it is impossible for aluminum to sustain the high temperature annealing. A refractory material for gate electrodes was emerged from the demand of self-align Source/Drain process and was thought to replace aluminum. In 1967, poly-silicon gate was first used to replace the aluminum gates at Bell Labs [6]. Poly-Si with the capability of high temperature allows the integration of complex process, and its work function is easily modulated by the concentration and types of dopant to match the work function of Si substrate, hence, threshold voltage of transistor is significantly lower. The resistance of poly-Si is also lowered due to be heavily doped. The successful integration of poly-Si gate ruled out the refractory metal gate and disclosed the age of silicon gate technology. Although the structure of poly-Si/oxide/Si replaced the structure of metal (aluminum)/oxide/Si, the acronym of MOS is still continued to be used up to date.

In the 1970s, nMOS transistors in integration circuits (ICs) gradually substituted for bipolar transistors and pMOS transistors, based on the successful planar process and the high quality gate oxide (SiO_2) [7]. N^+ ploy-Si was used as the gate material due to its higher conductivity than p^+ poly-Si gate and easier integration in nMOS. Further, improvement and innovation of lithography technology led to vertical and horizontal device scaling access to large scale integration (LSI) from small Scale integration (SSI). Circuit performance became more and more improved. To ensure growth of very large scale integration (VLSI) circuits, MOS transistors were continuously scaled to improve the performance. Unfortunately, the resistivity of poly-Si, several hundred μΩ-cm, is almost two orders of magnificent higher than that of typical metals. The high resistance became to limit the signal transmission and circuit performance.

In the 1980s, refractory metals and polycides were proposed to solve the issue of high gate electrode resistance. Refractive metals, such as W and Mo, and their silicides

were the most potential materials to replace poly-Si due to its low resistivity and the highly thermal stability [8-11]. At that time, CMOS integrated circuits gradually superseded nMOS integrated circuits because of its low power consumption. W and Mo have mid-gap work function (mid of silicon energy band) to offer symmetric threshold voltage of nMOS to pMOS, which is a benefit to simplify circuit design. This single metal gate process once was thought to be a good choice. However, the refractory metal gates cannot be passivated by oxidation at high temperature oxidizing ambient. Interaction occurred at the interface between metal and SiO_2, seriously degraded the reliability of SiO_2 as SiO_2 thickness is decreased with device scaling. These issues retarded the usage of metal gates. On the other hand, polycide consisting of a layer of doped poly-Si covered by a metal silicide combines the advantages of low resistivity of silicide with the well poly-Si/SiO_2 interface [12, 13]. Polycide gate MOSFETs perform the same characteristic of conventional MOSFETs, and do not degrade the reliability of SiO_2, and subthreshold leakage. In addition, a self-aligned silicide process (salicide process) without an additional mask effectively simplified the fabrication of polycide gate as well as formation of the silicide on source/drain region for improving contact characteristics [14]. Silicide was confined to the exposed window of gate, source, and drain region to dramatically minimize pattern size. This approach not only solves the issue of high resistance of poly-Si gate but also lowers the contact resistance hence poly-Si gate is lasted [15].

In the 1990s, deep submicron CMOSFETs were the most promising devices for the ultra large scale integration (ULSI) circuits. Surface channel device is superior to buried channel device in the short channel characteristics, such as threshold voltage roll-off and subthreshold leakage [16, 17]. It is not easy for surface channel devices to have a suitable threshold voltage by just only controlling the channel doping. The

threshold voltage should be also adjusted by the work function of gate. Dual poly-Si gate was used for CMOS, where an n^+ poly-Si gate has a work function near silicon conduction band for nMOS and a p^+ poly-Si gate has a work function near the silicon valance band for pMOS [18].

In 2003, MOSFET gate lengths were scaled down to sub-0.1μm (nano dimension regime) while the thickness of gate dielectric was scaled down to thinner than 3nm. Various issues were encountered by poly-Si gates when attempting to realize high yield and high performance ULSI circuits. These are high gate resistance, poly-Si gate depletion, boron penetration, and interaction between poly-Si gate and high dielectric constant (high-k) gate dielectrics (the next generation gate dielectrics). Poly-Si gate depletion increases the equivalent gate dielectric thickness by about 0.3nm and degrades the capability of channel current drive [19]. High gate resistance increases the resistance-capacitance (RC) delay time to degrade the high frequency performance. Although silicide technology can reduce the gate resistance, it is difficult to maintain a proper aspect ratio for gate stake and silicidation of fine line poly-Si is critical [20]. Boron penetration from the p^+ poly-Si gate induces the issues of threshold voltage control and gate oxide reliability [21]. High-k materials are expected to replace SiO_2 as the physical thickness of SiO_2 is below 1.5nm, where the direct tunneling current becomes too high to be acceptable. Interface reaction between poly-Si and high-k induced undesirable threshold voltage due to the Fermi-level pinning (FLP) [22]. Metal gate is considered to solve high resistance, poly-Si depletion, boron penetration, and integration issues of high-k dielectric in gate stake simultaneously and to continue CMOS scaling down below the 45nm technology node. The evolution of MOSFET gate electrodes is schematically shown in Fig. 1-1.

1.2 Metal Gate Electrodes

The demand of high performance integrated circuits (ICs) drives the CMOS device to scale down. As the device scales down, it is important to maintain or improve the switch speed, power consumption, and reliability so that new materials are included. The incorporation of new material in ICs is crucial since it requires consideration of numerous issues: contamination, process integration, thermal stability, performance, and reliability. In order to highlight the issue arising while metal materials are integrated into the advance CMOS, it is helpful to consider the basic design of this device. Figure 1-2 is a schematic cross-section of a sub-45nm metal gate CMOS structure which indicates the area where novel materials and process-integration solutions are necessary. For the aspect of metal gates, candidate metals must have high melting point and high thermal stability on gate dielectrics. These metals must have high conductivity and must have thermal expansion coefficients close to Si to prevent thermal stress during high temperature processes. The work function of metal gates must be suitable to obtain low and symmetric threshold voltages on the n- and pMOSFETs. The desired properties of metal gate for CMOS are listed in table 1-1.

1.2.1 Metal Gate Materials

Metal gate materials are generally classified into four categories including elementary metals, metal nitrides, metal silicides, and metallic alloys. In addition, there are some conducting metal oxides, metal carbines and ternary alloys such as In_2O_3, RuO_2, IrO_2, ZnO_2, ITO, ReO_2, TaC, TaTiN, TaAlN, and TaSiN.

i. Elementary metals:

The commonly used metals in IC industry are located in the group B of

periodic table, as shown in Fig. 1-3. The work functions periodically ranges from conduction band (E_c) to valence band (E_v) or Si energy band relative to locations of elements [23, 24]. Elements including Ti, Zr, La, V, Nb and Ta in column IVB and VB have the n-type work function; elements including Mo and W in column VI B have mid-gap work function and elements including Re, Ru, Co, Rh, Ir, Ni, Pd and Pt in VIIB and VIIIB have p-type work function. N-type elements are chemically reactive and are not suitable for the conventional CMOS process due to the interaction with the gate dielectric. P-type elements are relatively inert and can sustain the high temperature process. However, the chemical inertia reflects difficulty in patterning and poorness in adhesion.

ii. Metallic alloys:

The common metallic alloys are binary alloys consists of an n-type metal and a p-type metal, such as Ta-Pt alloys, Ru-Ta alloys, and Hf-Mo alloys [25-27]. The actual work function is determined by the atomic composition. Alloys with higher n-type metal content have a lower work function. On contrary, alloys with higher p-type metal content have a higher work function.

iii. Metal nitrides:

Metals react with nitrogen to form metal nitrides which are more chemically stable on dielectrics than pure metals. The commonly considered metal nitrides are tantalum nitride, titanium nitride, tungsten nitride, molybdenum nitride, and hafnium nitride [28-32]. The work function can be adjusted by the nitrogen composition and the nitride phase, but the tunable range is not wide enough to be used for both nMOSFETs and pMOSFETs. In

addition, an obvious drawback is the high resistivity.

iv. Metal silicides:

Metals react with silicon to form metal silicides. The commonly considered metal silicides are molybdenum silicide, tungaten silicide, nickel silicide, cobalt silicide, titanium silicide, platinum silicide and hafnium silicide. Silicide gate is compatible with conventional CMOS process and its effective work function is adjusted to suit both nMOSFETs and pMOSFETs by the IIIA and VA impurity [33-36]. Therefore, silicides are thought to be the mostly possible materials.

1.2.2 Resistivity

As the devices scale down, the line width, gate length, and contact hole get narrow and small. The resistance contribution to the RC delay increases not only at the metallic interconnects and contacts but also at the gate lines The sheet resistance of gate line should be less than 5-7 Ω/□ to prevent offsetting the advantage of device scaling by the interconnect in the gate level. In other words, the resistivity of gate material should be less than 12.5-25 μΩ-cm in 2007. Table II lists the gate electrode thickness and sheet resistance predicted in the International Technology Roadmap for Semiconductor (ITRS) [37].

1.2.3 Work Function

To evaluate the suitable work functions of gate electrodes, the threshold voltage of metal gate MOSFETs are considered. The threshold voltages (V_t) are determined by [38]:

$$V_{Tn} = (\Phi_{ms} - \frac{Q_f}{C_{ox}}) + 2\Psi_B + \frac{\sqrt{4\varepsilon_{Si} q N_A \Psi_B}}{C_{ox}}$$

$$V_{Tp} = (\Phi_{ms} - \frac{Q_f}{C_{ox}}) - 2\Psi_B - \frac{\sqrt{4\varepsilon_{Si} q N_A \Psi_B}}{C_{ox}}$$

, where Φ_{ms} is the work function difference between gate and Si channel, C_{ox} is the capacitance of gate dielectic, Q_f is the effective oxide charge, Ψ_B is the potential energy difference between intrinsic Fermi level and Fermi level of silicon channel, q is the elementary charge, and ε_{Si} is silicon permittivity, N_A is the substrate dopant concentration. As the gate dielectric thinning, the C_{ox} becomes large enough to neglect the effect of charges in dielectric and channel so that the threshold voltage is determined by theΦ_{ms} and Ψ_B relative to the work functions of gate and silicon channel. Figure 1-4 shows the simulated threshold voltage of bulk and SOI MOSFETs [39, 40]. In order to obtain V_t between 0.2-0.4 V ($\sim V_{dd}/4$), the ideal work function is 4.1eV-4.3 eV for bulk nMOSFETs and is 4.9-5.1 eV for pMOSFETs, where the channel concentration is about 1×10^{18} cm^{-3} in order to suppress the short channel effect [39]. The ultra-thin body or double-gate SOI MOSFETs with undoped channel require the work function of 4.3-4.5eV for n-type devices and 4.7-4.9eV for p-type devices [40].

1.2.4 Metal-Gate Structures

It is not easy that single metal layer can satisfied all of the above criteria. The most promising gate structure is stack of two metal layers, as shown in Fig. 1-5. The bottom layer serves as a threshold control layer. It is a thin barrier layer which needs to have proper work function, good adhesion, and chemical stability on gate dielectric. The top layer serves as a conduction layer. It is a thick layer which needs to have low resistivity, low stress, to protect the bottom layer from ion implantation and to passivate the bottom layer against oxidation and chemical solution attack during oxide passivation and clean process, respectively. The

proper top layer is suggested to be polycide due to its low thermal stress and compatibility with conventional process.

1.3 Work Function Engineering

The definition of work function (Φ) is the least amount of energy required to remove an electron from the surface of a conducting material. The work function is definitely dependent upon the surface quality. In other words, interface property is the main factor to determine the effective metal work function as the metal contact with the dielectric. The interface property is affected by the texture, composition, orientation and interfacial layers. Based on the concept of the interface modification, the work function engineering has been undertaken. The approaches are divided into two categories: one is structure modification and another is interface treatment, as shown in Fig. 1-6.

1.3.1 Structure Modification

Changes to the structural morphology of the thin metal films can be further classified into phase, orientation, and texture. The structure modification is usually relied on the deposition technique and post deposition anneal. On the other hand, chemical reactions are used to modify the film texture after film deposition, such as silicidation, metal alloying, nitridation and oxidation. Silicidation and metal alloying can be formed by deposition a two-layer stack followed by a thermal process [33, 34]. Reactive NH_3 (O_2) gas or N (O) plasma was used to nitrify (oxidize) the deposited film [41]. As the textures are changed, the intrinsic physical and chemical properties differ from the as-deposited films. This implies the work function should be changed after the texture modification. The phase

transition often contains the changes of microstructure in the films as well as the texture. For example, the Ni_3Si transits to NiSi and $NiSi_2$ at 400 °C and 650 °C, respectively [42]. The work function changes from 4.38 to 4.8 eV. The work function related to orientation is origin of the anisotropic atomic packing. Along the various crystalline planes, the periodic atoms have different lattice constant related to the orientation. This implies the free electron will have anisotropic conduction band due to asymmetric lattice structure. For example, the work function of Mo (110) is 4.95 eV while the work function of Mo (100) is 4.53 eV [23].

1.3.2 Interface Treatment

The interface treatment can be performed with the surface treatment before metal deposition, or with the ion implantation after metal deposition. The surface treatment can be plasma treatment or reactive gas annealing. The interface treatment locally change the texture, bonding and even defects (dangling bonds) at the interface between metal gate and dielectric as well as incorporation of extra impurities. The changes of interface and incorporation of impurities will upset a balance within the primary interface. Chemical thermodynamic equilibrium occurred and lead to a new static state. This self-reaction will regulate the charge distribution at interface and produce a new effective work function of gate electrode. For example, the work function of Mo film charges from 5.0 eV to 4.4 eV after N implantation and post annealing because as Mo film is implanted with N^+ ions, the texture structure of Mo film is destroyed [43].The structure incorporated N impurities is rearranged after annealing. Mo_2N phase forms and N also piles-up at interface between Mo film and gate dielectric. The local texture is thoroughly changed by implantation and annealing [43].

1.4 Fermi-Level Pinning

Fermi-level pinning (FLP) means that $\Phi_{m,eff}$ differs from $\Phi_{m,vac}$, such that the V_{fb} (V_t) is almost independent upon metal gates. Fermi-level pinning is a consequence of interfacial reaction or interfacial charge exchange between gate electrode and gate dielectric. The reasons of Fermi-level pinning could be categorized as intrinsic and extrinsic charge exchange. The former, intrinsic charge exchange, comes from the concept of the metal induced gap states (MIGS) and virtue gap states (ViGS) [44]. The models of MIGS and ViGS are constructed from the comprehensive study of metal-semiconductor contact (Schottky barrier).The origin of the ViGS is from the dangling bonds of uncoordinated surface atoms. These dangling bonds produce surface states which continuously disperse in the energy gap of dielectrics (semiconductors). The surface states are confined at the surface to exponentially decay into vacuum and into solid [45]. MIGS are derived from the ViGS as a metal is laid on a dielectric (semiconductor). Free electrons in the metal penetrate into the dielectric at the interface and cause the interface charges. The amount of transferred charges is dependent upon the charge neutrality level (E_{CNL}) and metal Fermi level (E_{Fm}). E_{CNL} is a branch point of surface energy level, which is usually near the mid-gap of dielectric energy band. No charge will be transferred across the interface when the E_{Fm} is coincided with E_{CNL}. With the surface level above or below E_{CNL}, the net charge of the surface is negative and positive on the dielectric side, respectively. In the case where E_{Fm} is above E_{CNL}, the dipole layer created at the interface will be negative on the dielectric side. This dipoles will tend to drive the E_{Fm} close to E_{CNL}, and hence the effective work function ($\Phi_{m,eff}$) will differ from the work function in vacuum ($\Phi_{m,vac}$). The effective work function takes account the effects of interfacial

dipoles and is determined by

$$\Phi_{m,eff} = \Phi_{CNL} + S(\Phi_{m,vac} - \Phi_{CNL})$$

where S is a slop parameter that accounts for interfacial dipoles and depends on the electronic component of the dielectric constant, ε_{∞}, [46]

$$S = [1 + 0.1(\varepsilon_{\infty} - 1)^2]^{-1}$$

Figure 1-7 shows two critical conditions, Fermi Level Free S=1, and Fermi Level Fixed S=0. A high-k dielectric will lower the slope parameter S and highly pin the $E_{m,eff}$ in the E_{CNL} since the electronic component of the dielectric constant is high. This indicates that high-k materials will suffer from serious Fermi-level pinning effect, and induce effective work function close to Φ_{CNL}. To obtain suitable effective work function, metal gates on high-k materials for pMOSFETs (nMOSFETs) should have higher (lower) work function to compensate the effects of Fermi-level pinning. Table III lists the S values and relative E_{CNL} values of the common dielectrics including SiO_2, Si_3N_4, HfO_2 and ZrO_2 [47]. On the other hand, the external charge exchange comes from interfacial reaction to cause excessive layer, interfacial defects and interfacial bonds often occurs within the reactive metal/oxide gate stack or poly-Si/high-k stack. These external charge exchanges form a dipole layer to reach a new static equilibrium at the interface. The experimental results reveal that Hf-Si bonds will pin the metal Fermi level (E_{Fm}) to silicon conduction band and Al-O-Si bonds will pin the E_{Fm} to silicon valence band [22].

1.5 Process Integration

Metal gate devices can be realized by three basic approaches: gate-first, gate-last, and fully-silicided, as shown in Fig. 1-8 [48-50]. The gate-first approach is superior due to its simple fabrication, similar to the conventional poly-gate process.

However, its main disadvantages are the issues of contamination of front-end equipment, difficult gate etching, and integrity of gate stack during high temperature annealing. The gate-last, also called replacement gate, uses a dummy gate, usually a poly-Si gate, which is etched and replaced by metals after front-end process. The process is complicated since two gate processes are needed. In addition, it is critical to remove the dummy gate without damaging gate dielectric and channel surface. Furthermore, both gate-first and gate-last approaches will be very complicated if two kinds of metal gates are needed to meet the dual work function requirement of CMOSFETs. Recently metal-like fully-silicided (FUSI) gate has been proposed. It is compatible with the conventional fabrication process and its work function can be modified by the incorporation of dopants. The processes are the same as those for convention poly-Si gate devices' until the step of source/drain activation, and only a fully silicided process is added to transfer the poly-Si gate into a silicided gate.

1.6 Overview of Dissertation

This dissertation focuses on the characterization of alternative metal gates for the sub-45 nm technology node. In this chapter, the history of gate electrode evolution and the key criteria of metal-gate materials have been reviewed. Chapter 2 provides the experimental techniques of metal-film deposition, physical characterization, and electrical characterization. Following is the main materials to evaluate various types of metal gates. Chapter 3, 4, and 5 deeply study the tantalum nitride gate, molybdenum nitride gate, and tungsten nitride gate on SiO_2 and HfO_2, respectively. Chapter 6 presents the study of Ta-Pt alloy gates. It is focused on the work function modulation and the effects of II and V column impurities incorporation, as well as the thermal stability. Chapter 7 studies the FUSI gate process integration in CMOS. A

summary and suggestion for future work are present in chapter 8.

References

[1] J. Bardeen and W. Brattain, ”The Transistor, A Semiconductor Triode,” *Phys. Rev.* vol. 74, pp.130-231,1948.

[2] J. S. Kilby, ”Miniaturized Electronic Circuits,” U.S. Patent 3138743, June 23, 1964 (filed February 6, 1959).

[3] R. N. Noyce, “Semiconductor Device-and-Lead structure,” U.S. Patent 2918877, April 25,1961 (filed July 30, 1959).

[4] D. Kahng and M.M. Atalla,” Silicon-Silicon Dioxide Field Induced Surface Device,” paper presented at *IRE-AIEE Solid-State Device Conference*, Pittsburg, 1960.

[5] R. W. Bower, H. G. Dill, K. G. Aubuchon, and S.A. Thompson, “MOS Field Effect Transistors Formed by Gate Masked Ion Implantation,” *IEEE Trans. Electron Dev.,* vol. ED-51, pp. 757-761, 1968.

[6] J. C. Sarace *et al.*,”Metal-nitride-oxide-silicon Field EffectTransistors with Self-aligned Gate,” *J. Solid-State Electron.*, vol. 11, pp.653-660, 1968.

[7] J. T. Clemens, “Silicon Microelectronics Technology,” *Bell Labs Technical Journal*, pp.76-102, 1997.

[8] P. L. Shah, “Refractory Metal Gate Processes for VLSI Application,” *IEEE Trans. on Electron Dev.* vol. ED-26, no.4 pp.631-640, 1979.

[9] S. Iwata, N. Yamamoto, N. Kobayashi, T. Terada, and T. Mizutani, “A New Tungsten Gate Process for VLSI Application,” *IEEE Trans. Electron Dev.* vol. ED-31, no. 9, pp.1174-1179, 1984.

[10] T. Mochizuki, K. Shibata, T. Inoue, and K. Ohuchi,”A New MOS Process Using $MOSi_2$ as a Gate Material,” *Japan. J. Appl. Phys.*, vol. 17-1, pp.37-42,

Oct. 1977.

[11] K. C. Saraswat, F. Mohammadi, and J. D. Meindl, “WSi_2 Gate MOS Device,” in *1979 Int. Electron. Devices Meet. Tech. Dig.*, pp.462-464.

[12] B. L. Crowder, and S. Zirinsky, “One Micro MOSFET VLSI Technology: Part VII-Metal silicide interconnection technology-A future perspective,” *IEEE Trans. Electron Dev.* vol. ED-26, no. 4, pp.369-371, 1979.

[13] H. J. Geipel, N. Hsieh, M. H. Ishaq, C. W. Koburger, and F. R. White, “Composite Silicide Gate Electrodes- Interconnection for VLSI Device Technologies,” *IEEE Trans. Electron Dev.* vol. ED-27, p.1417, 1980.

[14] T. Shibata, K. Hieda, M. Sato, M. Konka, R.L.M. Dang, and H. Iizuka, “An Optimally Design Process for Submicron MOSFETs,” in *1981 Int. Electron. Devices Meet. Tech. Dig.*, pp.647-650

[15] C. Y. Ting, “Silicide for Contactsand Interconnection,” in *1984 Int. Electron. Devices Meet. Tech. Dig.*, pp. 110-113.

[16] K. Tanaka and M. Fukuma, “Design Methodology for Deep Submicron CMOS,” in *1987 Int. Electron. Devices Meet. Tech. Dig.,* pp.628-631.

[17] B. Davari, W. H. Chang, M. R. Wordeman, C. S. Oh, Y. Taur, K.E. Petrillo, D. Moy, J.J. Buchignano, H. Y. Ng, M. G. Rosenfield, F. J. Hohn, M.D. Rodriguez, ”A High Performance 0.25 μm CMOS Technology,” in *1988 Int. Electron. Devices Meet. Tech. Dig.,* pp.56-59.

[18] C. Y. Wong, J.Y.-C. Sun, Y. Tour, C.S. Oh, R.Angelucci, and B. Davari,”Doping of N^+ and P^+ Poly-Si in a Dual-Gate CMOS Process,” in *1988 Int. Electron. Devices Meet. Tech. Dig.,* pp.238-241.

[19] N. D. Arora, R. Rios, and C. L. Huang,” Modeling the Poly-Si Depletion Effect and its Impact on Submicrometer CMOS Circuit Performance,” *IEEE*

Trans. Electron Dev. vol. ED-42, no. 5, pp.935-943, 1995.

[20] T. Ohguro, S. Nakamura, M. Koike, T. Morimoto, A. Nishiyama, Y. Ushiku, T. Yoshitomi, M. Ono, M. Saito, H. Iwai, "Analysis of Resistance Behavior in Ti- and Ni-salicided Poly-Si Films," *IEEE Trans. Electron Dev.* vol. ED-41, no. 12, pp.2305-2317, 1994.

[21] J. R. Pfiester, F. K. Baker, T. C. Mele, H. H. Tseng, P. J. Tobin, J. D. Hayden, J. W. Miller, C. D. Gunderson, and L. C. Parrillo, "The Effects of Boron Penetration on p+ Polysilicon Gated PMOS Devices," *IEEE Trans. Electron Dev.* vol. ED-37, no. 8, pp.1842-1851, 1990.

[22] C. C. Hobbs, L. R. C. Fonseca, A. Knizhnik, V. Dhandapani, S. B. Samavedam, W.J. Taylor, J.M. Grant, L. G. Dip, D. H. Triyoso, R. I. Hegde, D.C. Gilmer, R. Garcia, D. Roan, M.L. Lovejoy, R.S. Rai, E. A. Hebert, H.-H. Tseng, S. G. H. Anderson, B.E. white and P. J. Tobin,"Fermi-Level Pinning at the Ploysilicon/Metal-oxide Interface- Part II," *IEEE Trans. Electron Dev.* vol. ED-51, no. 6, pp.978-984, 2004.

[23] H. B. Michaelson," The Work Function of the Elements and its Periodicity," J. Appl. Phys., 48, pp.4729-4733, 1977.

[24] C. Cabral, Jr., C. Lavoic, A.S. Ozcan, R.S. Amos, V. Narayanan, E. P. Gusev, J. L. Jordan-Sweet and J. M. E. Harper, "Evaluation of Thermal Stability for CMOS Gate metal materials," *J. Electrochem. Soc.* 151 p.F283-F287, 2004.

[25] B. Y. Tsui, C. F. Huang, "Wide Range Work Function Modulation of Binary Alloys for MOSFET Application," *IEEE Electron Device Lett.* vol. 24 no.3 pp.153-155, March, 2003.

[26] V. Misra, H. Zhong; H. Lazar," Electrical properties of Ru-based Alloy Gate Electrodes for Dual Metal Gate Si-CMOS," *IEEE Electron Device Lett.* vol.

23 no.6 pp.354-356, June 2002.

[27] T.-L. Li, C.-H. Hu, W.-L. Ho, H.C.-H. Wang, C.-Y. Chang," Continuous and Precise Work Function Adjustment for Integratable Dual Metal Gate CMOS Technology Using Hf-Mo Binary Alloys," *IEEE Trans. Electron Dev.* vol. ED-52, no. 6, pp.1172-1179, 2005.

[28] H. Shimada, I. Ohshima, T. Ushiki, S. Sugawa, T. Ohmi, "Tantalum Nitride Metal Gate FD-SOI CMOS FETs Using Low Resistivity Self-grown bcc-tantalum Layer," *IEEE Trans. Electron Dev.* vol. ED-48, no. 8, pp.1619-1626, 2001.

[29] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, T. Kunio, "A Dual-metal Gate CMOS Technology Using Nitrogen-concentration-controlled TiNx film," *IEEE Trans. Electron Dev.* vol. ED-48, no. 10, pp.2363-2369, 2001.

[30] B.-Y. Tsui, C.-F. Huang and C.-H. Lu," Investigation of Molybdenum Nitride Gate on SiO_2 and HfO_2 for MOSFET Application," *J. Electrochemical Society* vol. 153 p.G197-G201, 2006.

[31] J. Pan, C. Woo, M.-C. Ngo, C.-Y. Yang, P. Besser, P. King, J. Bernard, E. Adem, B. Tracy, J. Pellerin, Qi Xiang, M.-R. Lin," Self-aligned nickel, cobalt/tantalum nitride stacked-gate pMOSFETs fabricated with a low temperature process after metal electrode deposition," *IEEE Trans. Electron Dev.* vol. ED-50, no. 12, pp.2456-2460, 2003.

[32] C. Ren, D.S.H. Chan, M.-F. Li, W.-Y. Loh, S. Balakumar, C.H. Tung, N. Balasubramanian, D.-L. Kwong, "Work Function Tuning and Material Characteristics of Lanthanide-Incorporated Metal Nitride Gate Electrodes for NMOS Device Applications," *IEEE Trans. Electron Dev.* vol. ED-53, no. 8, pp.1877-1884, 2006.

[33] J. H. Sim, H. C. Wen, J. P. Lu, D. L. Kwong, "Dual Work Function Metal Gates Using Full Nickel Silicidation of Doped Poly-Si," *IEEE Electron Device Lett.* vol. 24 no.10 pp.631-633, Oct. 2003.

[34] K. Hosaka, T. Kurahashi, K. Kawamura, T. Aoyama, Y. Mishima, K. Suzuki, S. Sato, "A Comprehensive Study of Fully-silicided Gates to Achieve Wide-range Work Function Differences (0.91 eV) for High-performance CMOS Devices," in *2005 Proc. Symp. VLSI Tech.,* pp.66-67.

[35] S. P. Chang, B. J. Cho, D.-L. Kwong," Thermally stable fully silicided Hf-silicide metal-gate electrode," *IEEE Electron Device Lett.* vol. 25 no.6 pp.372-374, June 2004.

[36] T.-L. Li; W.-L. Ho; H.-B. Chen, H.C.-H. Wang, C.-Y. Chang, Chenming Hu, "Novel dual-metal gate technology using Mo-MoSi/sub x/ combination," *IEEE Trans. Electron Dev.* vol. ED-53, no. 6, pp.1420-1426, 2006.

[37] International Technology Roadmap Semiconductors 2002 update, p.58.

[38] S. M. Sze, Physics of Semiconductor Devices, 2nd edition, 1985, John Wiley & Sons, p.442.

[39] De, D. Johri, A. Srivastava, C. M. Osburn, "Impact of Gate Work Function on Device Performance at the 50 nm Technology Node," *Solid-State Electronics,* vol. 44, no.6, p.1077, June,2000.

[40] L. Chang, S. Tang, T. –J. King, J. Bokor and C. Hu, "Gate Length Scaling and Threshold Voltage Control of Double-Gate MOSFETs," in *2000 Int. Electron. Devices Meet. Tech. Dig.,* pp.719-723.

[41] G. Qi, W. M. Small and T. Debray," Thermochemistry and Diffusion of Nitrogen in Solid Molybdenum," *Met. Trans. B.* vol.22 p.219, April 1991.

[42] K. Takahashi, K. Manabe, T. Ikarashi, N. Ikarashi, T. Hase, T. Yoshihara, H.

Watanabe, T.Tatsumi and Y. Mochizuki, "Dual Workfunction Ni-Silicide/HfSiON Gate Stacks by Phase-Controlled Full-Silicidation (PC-FUSI) Technique for 45nm-node LSTP and LOP Devices" in *2004 Int. Electron. Devices Meet. Tech. Dig.*, pp.91-94.

[43] Qiang Lu, Ronald Lin, Pushkar Ranade, Tsu-Jae King, Chenming Hu" Metal Gate Work Function Adjustment for Future CMOS Technology" *in 2001 Proc. Symp. VLSI Tech.*, 5A-1.

[44] W. Mönch, Semiconductor Surfaces and Interfaces, 2nd edition, chapter 3. Springer-Verlag, 1995.

[45] V. Heine, "Theory of Surface States," *Phys. Rev.* vol. 138, pp.A1689-A1696, June 1965.

[46] W. Mönch,"Electronic Properties of Ideal and Interface-Modified Metal-Semiconductor Interfaces," *J. Vac. Sci. Techno. B*, vol. 14, p.2985, July/Aug. 1996.

[47] Y.-C. Yeo, P. Ranade, T.-J. King and C. Hu, "Effects of High-k Gate Dielectric Materials on Metal and Silicon Gate Workfunctions," *IEEE Electron Device Lett.* vol. 23 no.6 pp.342-344, June 2002.

[48] J. Chen, B. Maiti, D. Connelly, M. Mendicino, F. Huang, O. Adetutu, Y. Yu,D. Weddington, W.Wu, J. Candelaria, D. Dow, P. Tobin, J. Mogab," 0.18μm Metal Gate Fully-Depleted SOI MOSFETs for Advanced CMOS Applications," in *1999 Proc. Symp. VLSI Tech.*, T3B-1.

[49] Kouji Matsuo, Tomohiro Saito, Atsushi Yagishita, Toshihiko Iinuma, Atsushi IWn-akoshi, Kazuaki Nakajima, Seiichi Omoto and Kyoichi Suguro," Damascene Metal Gate MOSFETS with Co Silicided Source/Drain and High-k Gate Dielectrics" in *2000 Proc. Symp. VLSI Tech.*, T8-1.

[50] B. Tavel1, T. Skotnicki2, G. Pares2, N. Carrière2, M. Rivoire2, F. Leverd2, C. Julien2, J. Torres2, R.Pante, "Totally Silicided (CoSi2) Poly-Si: a novel approach to very low-resistive gate without metal CMP nor etching," in *2001 IEDM Tech. Dig.*, p.825.

Table 1-1 Desired properties of the metal gates for CMOS.

Suitable work function
Low resistivity
Stable on gate dielectric, high melting point
Easy pattern
Mechanical stability: good adhesion, low stress
Stability throughout processing, including S/D activation, post annealing, silicide contact formation, passivation, metallization,
Surface smooth
Should not contaminate devices, wafers, and working apparatus
Should not diffuse into gate oxide

Table 1-2 The sheet resistance and thickness of gate electrodes predicted in ITRS roadmap.

Year of performance	2005	2006	2007	2010	2013	2016
Technology node (nm)	80	70	65	45	32	22
Gate electrode thickness (nm)	32-64	30-60	25-50	18-36	13-26	9-18
Gate electrode sheet Rs (Ω/□)	5	5	5	5	6	7

Table 1-3 Charge neutral level and S parameter for several dielecitrics.

			Theory		Experiment	
	E_G	ε_∞	S	$E_{CNL}-E_v$	S	$E_{CNL}-E_v$
SiO_2	9	2.25	0.86	4.5	0.95	5.04
Si_3N_4	5.3	3.8	0.56	2.6	0.59	2.79
HfO_2	6	4	0.53	3.7	0.52	3.64
ZrO_2	5.8	4.8	0.41	3.6	0.52	3.82

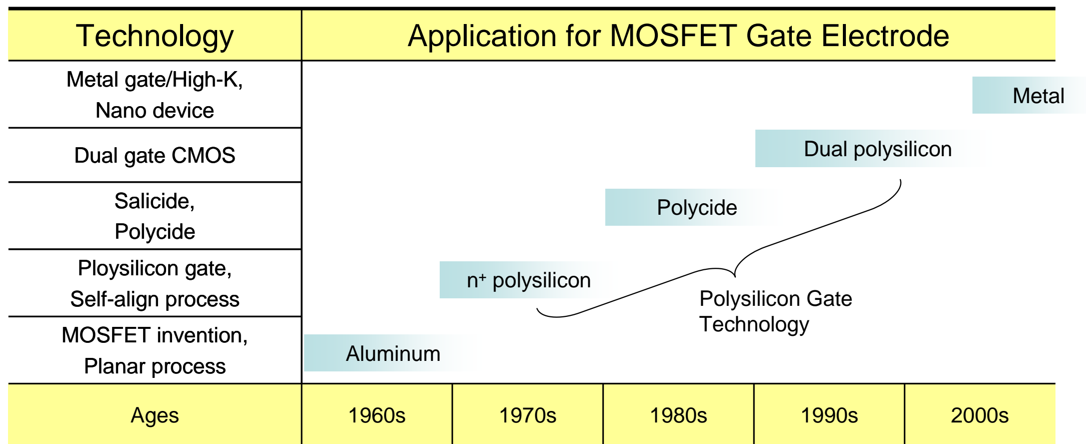


Fig. 1-1 The evolution of MOSFET gate electrodes with the innovation of technologies.

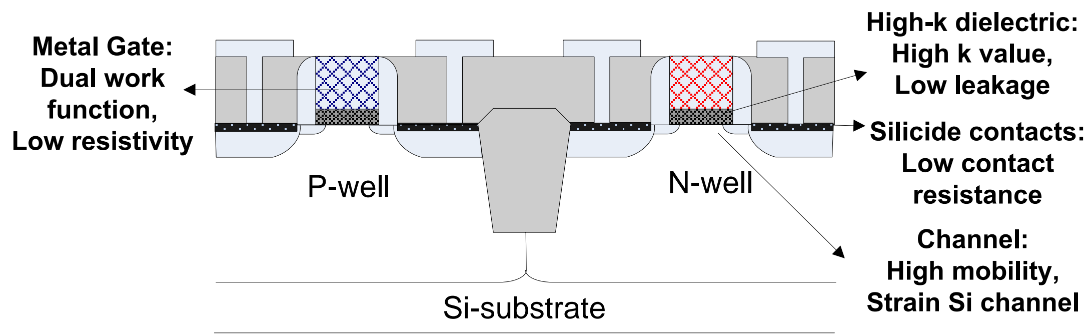


Fig. 1-2 A schematic cross-section of a sub-45 nm metal gate CMOS structure, where the novel materials and process-integration solution are indicated.

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg	IIIB	IVB	VB	VIB	VIIB		VIIIB		IB	IIB	Al^{n}	Si	P	S	Cl	Ar
K	Ca	Sc	Ti^{n}	V^{n}	Cr^{m}	Mn^{n}	Fe	Co^{p}	Ni^{p}	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr^{n}	Nb^{n}	Mo^{m}	Tc	Ru^{p}	Rh	Pd^{p}	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La^{n}	Hf^{n}	Ta^{n}	W^{m}	Re^{p}	Os	Ir^{p}	Pt^{p}	Au^{p}	Hg	Tl	Pb	Bi	Po	At	rn
Fr	Ra	Ac															

Fig. 1-3 The common metal elements in periodic table. The elements were marked as p, n, and m according to its work function near the valence band, conduction band, and mid-gap of silicon energy band, respectively.

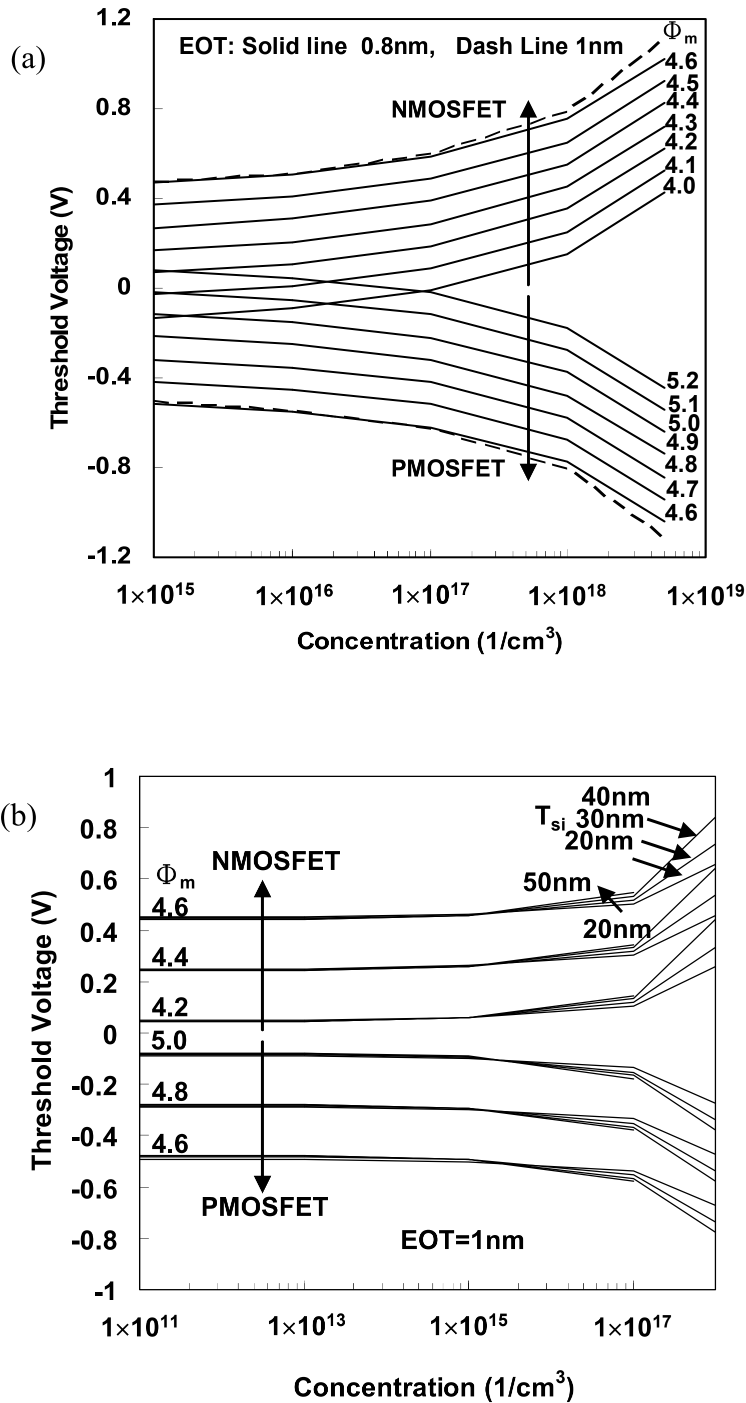


Fig. 1-4 The threshold voltages of (a) bulk devices and (b) SOI devices versus channel concentrations. The estimation of threshold voltage was evaluated with the simulation of long channel and uniform charnel dopant devices by Avant! MEDICI. The dependence of oxide thickness or thickness of silicon body was also taken an account.

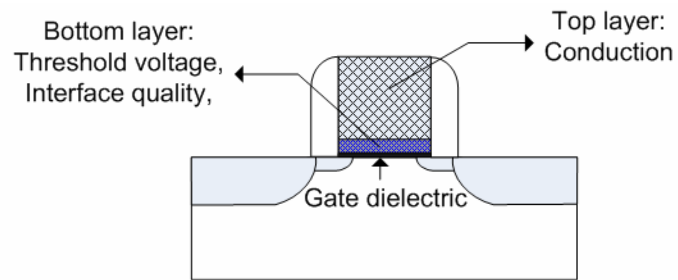
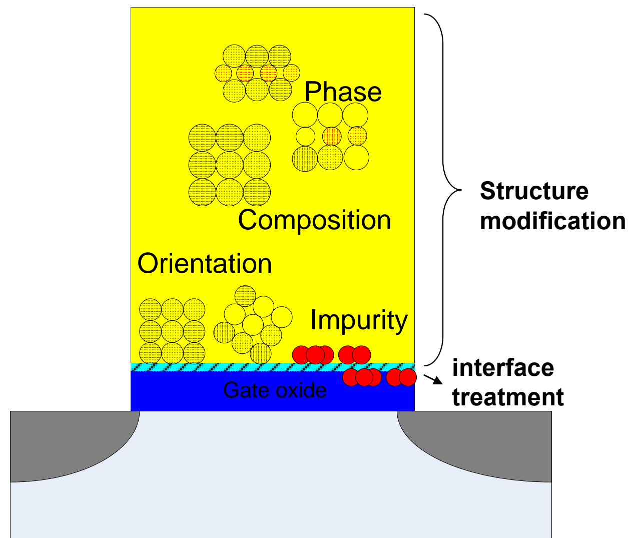


Fig. 1-5 A structure of the stack metal-gate electrode. The bottom layer serves as threshold control and dielectric-contact layer; the top layer severs as conduction layer connected with metallic interconnect.



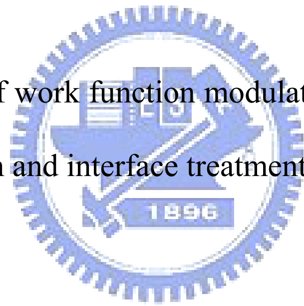


Fig. 1-6 Schematic methods of work function modulation divided into two categories: structure modification and interface treatment.

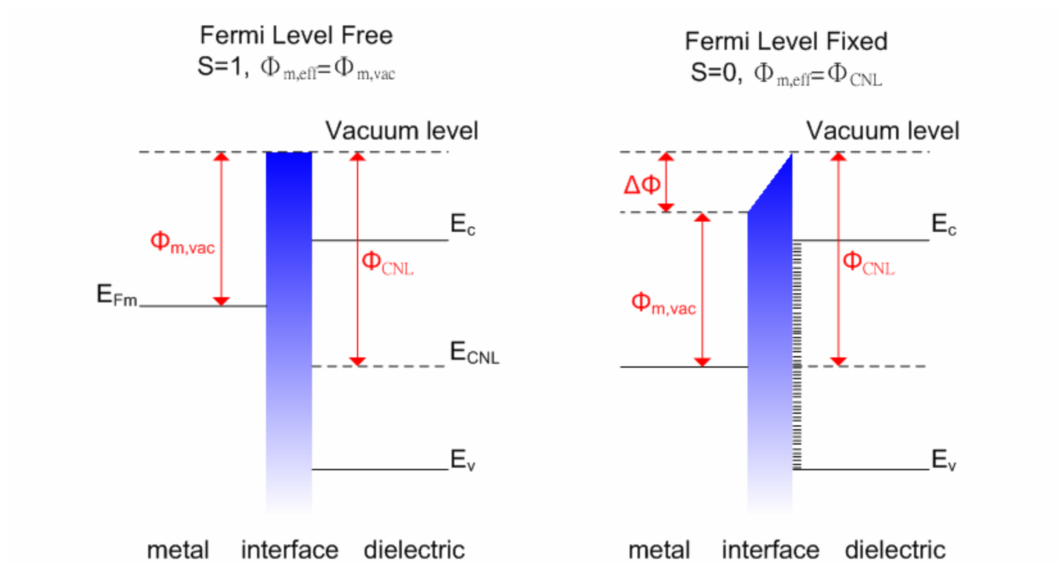
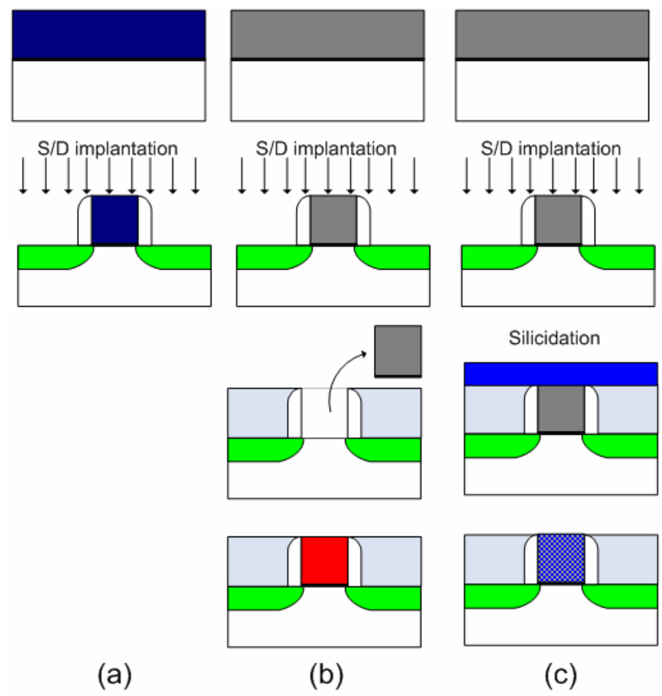


Fig. 1-7 The critical Fermi level of metal gate on dielectric. The effective work function of metal on dielectric is the same as in vacuum while the interface is perfect. The effective work function is fixed to the delectric-charge-neutral level while the interface-state density is high. The interface states are caused by both intrinsic states (MIGS, ViGS) and extrinsic states (defect levels).



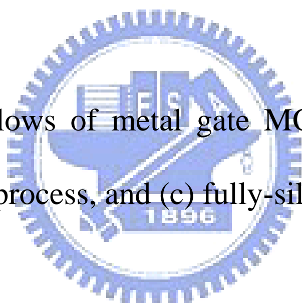

1896

Fig. 1-8 Schematic process flows of metal gate MOSFETs including (a) gate-first process; (b) gate-last process, and (c) fully-silicided process.

Chapter 2

Experimental Methods

2.1 Introduction

This research utilized varieties of semiconductor processes, electrical characterizations and material characterizations. The common experimental techniques are briefly described in this chapter. The detail please refer to the handbooks of thin film deposition, device characterization, and surface analysis [1-3]. Several metal gate electrodes of this research were deposited by a sputtering deposition (PVD) system. The sheet resistance was measured by a four-point probe. High frequency capacitance-voltage characteristics were measured by a precise impendence meter (Agilent 4284A). Quasi capacitance-voltage characteristic was measured by the Agilent 4140B. Current-voltage characteristics were measured by semiconductor parameter analyzer (Agilent 4156C). We will briefly discuss the basic electrical characterization (parameters), and clearly illustrate work function extraction in section 2.2. The composition depth profile was analyzed by the Auger electron spectroscopy (AES). The dopant (impurity) depth profile was analyzed by the secondary ion mass spectroscopy (SIMS). The composition of thin film was analyzed by the Rutherford backscattering spectroscopy (RBS). Surface bonds ware analyzed by the X-ray photoelectron spectrometer (XPS). Interface morphology was analyzed by the transmission electron microscopy (TEM). Microstructure was analyzed by the X-ray diffractometer (XRD). In section 2.3, we will clearly discuss

an interfacial analysis technique- inelastic electron tunneling spectroscopy (IETS). IETS can detect interfacial quality of the metal-insulator-metal or metal-insulator-semiconductor structures without destroying structures, which is different from surface analysis.

2.2 Electrical Characterizations

2.2.1 Basic Electrical Characterizations and Parameters

The basic C-V characteristic of a metal-oxide-semiconductor capacitor and the relative $1/C^2$-V curve are shown in Fig. 2-1. Both the high-frequency C-V (HFCV) and quasi-static C-V (QSCV) curves are included. The capacitance equivalent thickness (CET) was calculated from the capacitance at accumulation mode. If the physical thickness of oxide layer is thick enough to suppress the quantum phenomenon, the CET is equal to the equivalent oxide thickness (EOT) and can be obtained from $C = \varepsilon_{ox} / EOT$ at accumulation mode. Flat-band voltage (V_{fb}) is extracted as the gate bias to achieve the flat-band capacitance (C_{fb}) which is determined by :

$$\frac{1}{C_{fb}} = \frac{1}{C_{ox}} + \frac{1}{C_D}, \qquad \text{......... (1)}$$

where $C_D = \frac{\varepsilon_{Si}}{L_D}$ and $L_D = \sqrt{\frac{KT\varepsilon_{Si}}{N_A q^2}}$.

C_{ox} is the oxide capacitance and C_D is the capacitance contributed by the Si substrate as biased at V_{fb}. ε_{Si} is the silicon permittivity, K is the Boltzman constant, T is the absolute temperature, and q is electron charge. L_D is Debye length dependent upon substrate concentration N_A, which is evaluated from the slop of $1/C^2$ -V_g plot at the depletion region. The interface states (D_{it}) between oxide and

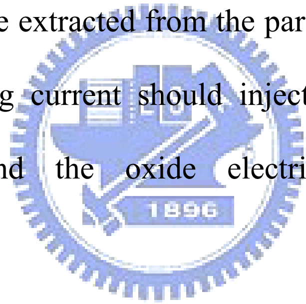
gate dielectric can be calculated from the capacitance difference between HFCV and QSCV, and is determined by :

$$D_{it} = \frac{C_{ox}}{q}\left(\frac{C_{QS}}{C_{ox} - C_{QC}} - \frac{C_{HF}}{C_{ox} - C_{HF}}\right). \qquad \text{……... (2)}$$

The J_g-V_g characteristic is shown in Fig. 2-2 (a). The corresponding Fowler-Nordheim (FN) plot is shown in Fig. 2-2 (b). As the conduction through the oxide is due to FN tunneling, the FN plot is a straight line. The slope is determined by:

$$S = \frac{8\pi\sqrt{2m^{*}\Phi_B^3}}{3qh} \ , \qquad \text{……….. (3)}$$

, where h is Planck constant and m* is effective mass of electron in dielectric. The barrier height (Φ_B) can be extracted from the parameter S. To get accurate Φ_B of metal gate, the tunneling current should injected from the gate electrode at accumulation mode and the oxide electric field is calculated from $(V_g - V_{fb})/EOT$.



2.2.2 Work Function Extraction

The work function extracted from the electrical characteristic is always called effective work function (Φ_{eff}). Effective work function is electrically extracted commonly using the V_{fb} versus EOT plot and sometime using the HFCV curves and FN plots. In some cases, effective work function can be extracted from single C-V measurement as the theoretical C-V curve is known. We can calculate the theoretical C-V curve of a MOS capacitor from the following equations [4]:

$$V_g = V_{ox} + V_{Si}$$
$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{1}{C_{Si}} \ , \qquad \text{……….. (4)}$$

where V_{ox} is the voltage drop accross oxide and V_{Si} is the voltage drop accross Si substrate. C_{Si} is the capacitance contributed by the Si substrate. V_{Si} also represents for the surface potential ψ_s. V_{ox} is the product of applied electrical field (F_{ox}) and the oxide thickness (EOT). To evaluate the F_{ox} and C_{Si} regardless of quantum mechanics, we should first introduce an abbreviation:

$$F(\beta\psi\frac{n_{po}}{p_{po}}) \equiv [(e^{-\beta\psi}+\beta\psi-1)+\frac{n_{po}}{p_{po}}(e^{\beta\psi}-\beta\psi-1)]^{1/2} \quad , \qquad \text{..........(5)}$$

The electrical field is hence determined by:

$$F_{ox} = \pm\frac{\varepsilon_{Si}}{\varepsilon_{ox}}\frac{\sqrt{2}KT}{qL_D}F(\beta\psi,\frac{n_{po}}{p_{po}}) \ , \qquad \text{.......... (6)}$$

with positive sign for $\psi_s>0$ and negative sign for $\psi_s<0$, where n_{po} and p_{po} are the equilibrium densities of electrons and holes, respectively, in the bulk of semiconductor, and $\beta = KT/q$ as the substrate dopant is p-type. The C_{Si} is determined by:

$$C_{Si} = \frac{\varepsilon_{Si}}{\sqrt{2}L_D}\frac{[1-e^{-\beta\psi_s}+(n_{po}/p_{po})(e^{\beta\psi_s}-1)]}{F(\beta\psi,n_{po}/p_{po})} , \qquad \text{.......... (7)}$$

As the substrate concentration and oxide thickness are known and given a value of ψ_s, the relative C and V_g are obtained from equations (4)-(7). A whole theoretical curve with a V_{fb}=0 can be obtained as all ψ_s are assigned.

To extract effective work function from the measured C-V curves, we first obtained the theoretical curve and then shift the theoretical curve to fit the measured curve, as shown in Fig. 2-3. Regarding the offset (ΔV_{fb}), at V_{fb} the band structure of silicon substrate is flat for both theoretical curve and measured curve; the ΔV_{fb} is hence contributed by the effective work function difference between the gate electrode and Si substrate. We can obtain the effective work function of the gate electrode from the value of the sum of the offset and the substrate work

function which is obtained according to the substrate concentration. If the oxide is charged, the effective work function is further minus Q_{eff}/C_{ox} of which Q_{eff} is the effective oxide charges.

The extraction of effective work function from the FN plot is based on the extraction of Φ_B as described in the last section. The effective work function is then determined by:

$$\Phi_{m,eff} = \Phi_B + \chi_{ox}, \qquad \text{.......... (8)}$$

where χ_{ox} is the electron affinity of oxide.

The extraction of effective work function from the V_{fb}-EOT plot is based on the equation:

$$V_{fb} = \Phi_{m,eff} - \Phi_s - Q_{eff}/C_{ox} = \Phi_m - \Phi_s - Q_{eff} EOT/\varepsilon_{ox}\ , \qquad \text{.......... (9)}$$

The extrapolation of V_{fb} at EOT=0 give the value of $\Phi_{m,eff}$-Φ_{Si}, as shown in Fig. 2-4. If the Φ_{Si} is known, we can obtain the effective work function. The hypothesis is based on that the substrate concentration is constant and the oxide charges are the same and fixed at the interface for all capacitors with different EOT so that the linear equation is assured. A general formula of V_{fb} should be express as:

$$V_{fb} = \Phi_{m,eff} - \Phi_{Si} - \frac{1}{\varepsilon_{ox}}\left[\int_0^{EOT} x\rho(x)dx\right], \qquad \text{.......... (10)}$$

where ρ(x) is the charge distribution in gate oxide. As the bulk charge is not low enough, eq. (9) cannot be applied to extraction $\Phi_{m,eff}$ because of the nonlinear behavior. For example, high dielectric-constant (high-k) material such as HfO_2 has a high density of oxide charges in bulk and a thick interfacial layer at the interface between HfO_2 and Si substrate. These result in the failure of the conventional work-function extraction from the V_{fb}-EOT plot. To solve this issue, an intentional, thick silicon-dioxide is grown before high-k film deposition [5]. The intentional

oxide layer can suppress the interfacial layer growth between high-k film and Si substrate during the high-temperature annealing such that the structure is thermally stable. We can easily fabricate a set of stable dielectric stacks with fixed HfO_2 layer and varied SiO_2 layer or otherwise. If the charge distribution is uniform in bulk and the interface charge is a sheet function as shown in Fig. 2-5, the V_{fb} could be rewritten as

$$V_{fb} = \Phi_{m,eff} - \Phi_{Si} - \frac{1}{\varepsilon_{ox}}[\frac{1}{2}\rho_{h.k.}EOT_{h.k.}^2 + Q_{h.k.}EOT_{h.k.} - \frac{1}{2}\rho_{ox}EOT_{h.k.}^2] - \frac{1}{\varepsilon_{ox}}[Q_{ox}(EOT_{h.k.} + EOT_{ox})] - \frac{1}{\varepsilon_{ox}}[\frac{1}{2}\rho_{ox}(EOT_{h.k.} + EOT_{ox})^2], \quad \ldots(11)$$

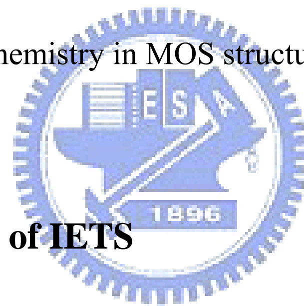
where $\rho_{h.k.}$ and ρ_{ox} are the bulk charge density in high-k layer and SiO_2 layer, respectively. $Q_{h.k.}$ and Q_{ox} are the interface charges. In general, the amount of bulk oxide charges of SiO_2 is small enough to be ignored. The terms of ρ_{ox} hence are eliminated from eq. (11). With a fixed high-k thickness, the first bracket term is constant and eq. (11) becomes a linear function of EOT_{tot} ($EOT_{tot} = EOT_{h.k.} + EOT_{ox}$). Eq. (11) is simplifies to be

$$V_{fb} = \Phi_{m,eff} - \Phi_{Si} - \frac{Q_o}{\varepsilon_{ox}} - \frac{1}{\varepsilon_{ox}}[Q_{ox}EOT_{tot}], \quad \ldots\ldots\ldots\ldots (12)$$

where Q_o is a constant and represents for $(1/2)\rho_{h.k.}EOT_{h.k.}{}^2 + Q_{h.k.}EOT_{h.k.}$. According to eq. (12), the effective work function extracted from linear extrapolation is counting for an offset due to charges in high-k layer. As the $EOT_{h.k.}$ is small enough, the offset can be ignored. Hence eq. (12) is the same as eq. (9), and the evaluated extracted work function on high-k layer is more accurate.

2.3 Inelastic Electron Tunneling Spectroscopy

I Inelastic electron tunneling spectroscopy (IETS), discovered in 1966 by Jacklevic and Lambe, is a very sensitive spectroscopy to detect the interface nature of metal-insulator-metal junctions [6, 7]. Electrons tunneling through the insulator excite the vibration modes corresponding to the type of molecular species in the insulator with energy loss [7, 8]. The loss energy levels represent the particular vibration modes of molecular species and reveal peak features in the IETS spectra. This specific peak features can be used to analyze the quality of barrier, adsobates, and interfacial chemistry in the analysis of adhesion, corrosion, and catalysis [9, 10]. The IETS is also applied to investigate the intrinsic dielectric quality, defect formation, and interfacial chemistry in MOS structures as the oxide thickness scales down below 2nm [11-15].



2.3.1 Basic Principle of IETS

The process of tunneling can be visualized in a rather simple picture as shown in Fig. 2-6(a). Electrons tunnel through the potential barrier in the ways of elastic tunneling and inelastic tunneling. Elastic tunneling occurs without energy loss while inelastic tunneling occurs with energy loss to excite the vibration modes. The quantum mechanical treatment of tunneling indicates that tunneling electrons exciting the vibration mode only takes place at the electron energy higher than the vibration energy and the vibration mode is quantized. Inelastic tunneling process proves electrons an additional way through the potential barrier as the voltage bias offers electron energy higher than vibration energy [7, 8]. This results in an increase of the slope of the tunneling current-voltage curve. By taking the first and second derivative of the I-V curve, a step and a pulse show up at where the slop

changes in the I-V curve as shown in Fig. 2-6(b).

The tunneling current is determined by the tunneling probability, γ, the number of filled states in metal A, N_A, and the empty states in metal B, N_B. Using the Wentzel-Kramers-Brillouin approximation, the γ is determined by

$$\gamma = exp(-2\int\{\frac{2m^*[U(z)-E_z]}{\hbar^2}\}^{1/2})dz \ , \qquad \text{.......... (13)}$$

where m* is electron effective mass, ħ is reduced Plank constant, U is the potential barrier, and E_z is the perpendicular part of electron energy. The elastic tunneling current, I_e, at voltage bias V is determined by

$$I_e = \int \gamma \times N_A(E-eV)F(E-eV)N_B(E)[1-F(E)]dE, \qquad \text{.......... (14)}$$

where F is the Fermi distribution function. At absolute zero temperature, the Fermi distribution is a step function. For small bias, the potential barrier becomes constant which results in a constant γ. For small bias and at absolute zero temperature, the equation (14) becomes

$$I_e = \gamma \times N_A(0)N_B(0)\int_0^{eV} 1dE = \sigma_o V, \qquad \text{.......... (15)}$$

I_e becomes a linear function of voltage if N_A and N_B are constant.

The inelastic tunneling current was first calculated by D. J. Scalapino using the mechanism whereby coupling occurs between tunneling electron and vibration molecules can be thought of in terms of a small perturbation of barrier height [16]. If a molecular dipole (p) as a source of perturbation is located very near one of metal electrodes, the image dipole must be included. The small perturbation potential, $U_{int,}$ is then given as

$$U_{int} = 2ep_z z/(z^2 + r_\perp{}^2)^{3/2}, \qquad \text{.......... (16)}$$

where $r_\perp$ is defined by Fig. 2-7. The tunneling matrix becomes

$$M = exp(-\int \{ \frac{2m^* [U(z) + U_{int}(z) - E_z]}{\hbar^2} \}^{1/2})dz, \qquad \text{.......... (17)}$$

Equation (17) can be expanded by the trick of $(1+x)^{1/2} = 1 + (1/2)x$ as x is very small. The z integral is evaluated as

$$M = [(\frac{2m^*}{\Phi_B})^{1/2} \frac{ep_z}{\hbar l} g(\frac{r_\perp}{l}) + 1] exp([-2m^* \Phi_B / \hbar^2]^{1/2} l), \qquad \text{.......... (18)}$$

$$g(x) = 1/x - 1/(1+x)^{1/2},$$

where Φ_B is the maximum barrier height, and $U(Z)-E_z$ is approximated to Φ_B. The second term in the bracket is contributed by the elastic tunneling, while the first term is due to the inelastic tunneling. If we neglect the correlation between the tunneling electrons and the molecule dipole at a distance $r_\perp$, the increase in current density from the molecule dipole due to the first term in eq. (18) is expressed as

$$I_i(\omega\,,\ V) = (\frac{dj}{dV})_0 |<1/p_z/0>|^2 [\frac{4\pi\ m^* e}{\hbar^2 \Phi}] \times ln(\frac{l}{r_o}) \int_{-\infty}^{\infty} N_A(E) F(E) N_B(E + eV - \hbar\omega)[1 - F(E + eV - \hbar\omega)] dE \qquad \text{.......... (19)}$$

where $(dj/dV)_0$ is the normal junction elastic conductance per unit area, and $|<1/p_z/0>|^2$ is the dipole matrix element. A term $2\pi l^2 ln(l/r_0)$ results form the spatial integral of g^2, where r_0 is the cutoff distance in the vicinity of the dipole. The total inelastic current due to all dipoles is give as

$$I_i(V) = N(\frac{dj}{dV})_0 \sum_m |<m/p_z/0>|^2 [\frac{4\pi\ m^* e}{\hbar^2 \Phi}] \times ln(\frac{l}{r_o}) \int_{-\infty}^{\infty} N_A(E) F(E) N_B(E + eV - \hbar\varpi_m)[1 - F(E + eV - \hbar\omega_m)] dE \qquad \text{.....(20)}$$

where N is the total number of molecular dipoles and $<m/p_z/0>$ is the strength of electron-dipole interaction. At low temperature, the first and second derivation are evaluated as

$$dI_i / dV = N(\frac{dj}{dV})_0 \sum_m |< m / p_z / 0 >|^2 [\frac{4\pi\ m^* e}{\hbar^2 \Phi}] ln(\frac{l}{r_o}) S(eV - \hbar\omega), \quad \cdots\cdots\cdots(21)$$

$$d^2 I_i / d^2 V = N(\frac{dj}{dV})_0 \sum_m |< m / p_z / 0 >|^2 [\frac{4\pi\ m^* e}{\hbar^2 \Phi}] ln(\frac{l}{r_o}) \delta(eV - \hbar\omega), \quad \cdots\cdots\cdots(22)$$

where S is a step function and δ is a delta function.

2.3.2 Measurement Method

It is not practical to measure I-V characteristic and to obtain the IETS spectrum by numerical derivation because derivation usually generates great noise. Modulation technique is often used to obtain the derivative of the I-V curve. Considering a Tayler series expansion of current function at voltage of V_b which is defined as

$$I(V) = I(V_b) + (\frac{dI}{dV})(V - V_b) + \frac{1}{2}(\frac{d^2 I}{dV^2})(V - V_b)^2 + \ldots.., \quad \ldots\ldots\ldots. (23)$$

1896

if the voltage is the sum of DC voltage and small sine-wave modulation, i.e, $V = V_b + \delta cos(\omega t)$, equation (22) becomes

$$\begin{aligned} I(V) &= I(V_b) + (\frac{dI}{dV})\delta\ cos(\omega t) + \frac{1}{2}(\frac{d^2 I}{dV^2})[\delta\ cos(\omega t)]^2 + \ldots.. \\ &= I(V_b) + (\frac{dI}{dV})\delta\ cos(\omega t) + \frac{1}{4}(\frac{d^2 I}{dV^2})\delta^2 [1 + cos(2\omega t)] + \ldots.. \end{aligned}, \quad \ldots\ldots..(24)$$

Therefore the amplitude of the second harmonic is proportional to the second derivative term, which can be detected by the lock-in amplifier (LIA). The d^2I/dV^2 can be converted to d^2V/dI^2 through the relationship:

$$\frac{d^2 I}{dV^2} = -(\frac{dI}{dV})^3 \frac{d^2 V}{dI^2} = -\sigma^3 \frac{d^2 V}{dI^2}, \quad \ldots\ldots\ldots. (25)$$

where σ is the conductance. According to eq. (25), either the current measurement or voltage measurement can be used to obtain IET spectra if the conductance varies slowly. Otherwise the conductance will cause the incorrect quantity by a

multiplied term of the varied conductance.

There is no commercial IET spectrometer, so that we set up a home-made system. The system is composite of a DC voltage source (Agilent 4140B), a sine wave oscillator operating at frequency from 0 to 50 KHz (SR 830 internal Osc.), a high input impendence DC digital voltmeter (Agilent 34420A DVM), a lock-in amplifier (SR 830 DSP LIA), and a home-made circuit as shown in Fig. 2-8. The system was controlled by a HP VEE program. The IETS system configuration is shown in Fig. 2-9. We can get the IET spectrum which comes from a plot of the voltage amplitude of 2ω modulation measured by the LIA versus the voltage bias measured by the DVM.

2.3.3 Signal Analysis

A measurement of 2ω amplitude versus bias voltage curve as well as the corresponding smooth curve due to the Ta-Pt/SiO_2/degenerate Si structure is shown in Fig. 2-10. The 2ω signal is not only composite of peaks but also smooth background which was fitted by a least square polynomial curve as shown in Fig. 2-10 (b). It is believed that the peak signals come from the inelastic tunneling current while the smooth background comes from elastic tunneling [17]. After the subtraction of background signal, we can obtain the IET spectrum. Another background subtraction method proposed by W.-K. Lye uses two-temperature measurement of IETS. The IET signal (ΔG) is determined by:

$$\Delta G_{4.2K} = \frac{(\Delta G_{4.2K} - \varepsilon) G_{77K} - \Delta G_{77K} G_{4.2K}}{G_{77K} - 0.0545 G_{4.2K}} \quad \text{……….(26)}$$

where G is the conductance and ε is an error term [12].

The ideal measurement temperature of IETS is at absolute zero. In fact, we can not reach absolute zero temperature. A small temperature can cause the electron

distribution broadening and hence smearing the IET signal. This is the reason why the IETS measurement should be performed at very low temperature. J. Lambe et al. have calculated half-width of peak signal given by 5.4 KT due to the thermal broadening [7]. In addition, J. Klein et al. have calculated the broadening effect of modulation amplitude to be $(1.22eV_{\omega})^2$ of which V_{ω} is the amplitude of modulation [18]. To combine these two effects including thermal broadening and modulation broadening, the peak signal of half-height width appropriate to IETS with a normal-state electrodes can be taken as $[(1.22eV_{\omega})^2+(5.4KT)^2]^{1/2}$.

References

[1] K. Seshan, *Handbook of Thin Film Deposition Process and Techniques*. 2nd ed. New York: Noyes Publications, 2002.

[2] D. K. Schroder, *Semiconductor Material and Device Characterization*, 2nd ed. New York: John Wiley & Sons, 1998.

[3] D. Briggs and M. P. Seah, *Practical Surface analysis by Auger and X-ray Photoelectron Spectroscopy*. New York: John Wiley & Sons,,1983.

[4] S. M. Sze: *Physics of Semiconductor Devices*, 2nd ed. New York: John Wiley & Sons, 1981, p.369.

[5] R. Jha, J. Gurganos, Y. H. Kim, R. Choi, J. Lee, "A Capacitanc-Based Methodology for Work Function Extraction of Metals on High-k," *IEEE Electron Device Lett.* vol. 25 no. 6 pp.420-422, Jun. 2004.

[6] R. C. Jaklevic and J. Lambe, Molecular Vibration Spectra by Electron Tunneling," *Phys. Rev. Lett.* vol. 17, pp. 1139-1140, Nov. 1966.

[7] J. Lambe and R. C. Jaklevicm, "Molecular Vibration Spectra by Inelastic Electron Tunneling," *Phy. Rev.* vol. 165, no.3 pp.821-832, Jan. 1968.

[8] S. Ewert, "Excitation of Molecules in Inelastic Electron Tunneling Spectroscopy," *Appl. Phys.* vol. A26, pp.63-82, 1981.

[9] M. Higo, H. Hayashi and S. Kamata," Hydride Formation of Evaporated Silicon Film Observed by Inelastic Electron Tunneling Spectroscopy," *Appl. Surf. Sci*. vol. 32, pp. 338-341, 1988.

[10] K. W. Hipps and U. Mazur, " Unoccupied Orbital Mediated Tunneling: Resonance-like Structure in the Tunneling Spectra of Polyacenes, " *J. phys. Chem.* vol. 98 pp. 5824-5829, 1994.

[11] G. Salace and J. M. Patat, "Tunneling Spectroscopy Possibilities in Metal-Oxide-Semiconductor devices with a Very Thin Oxide Barrier," *Thin Solid Films,* vol. 207 pp.213-219, 1992.

[12] W.-K. Lye, E. Hasegawa, T.-P. Ma, and R. C. Barker, "Quantitative Inelastic Tunneling Spectroscopy in the Silicon Metal-Oxide-Semiconductor System," *Appl. Phys. Lett.* vol. 71, no. 7 pp.2523-2525, 1997.

[13] G. Salace and C. Petit, "Inelastic Electron Tunneling Spectroscopy: Capabilities and Limitations in Metal-Oxide-semiconductor Devices," *J. Appl. Phys*. vol. 91, no.9, pp.5896-5901, 2002.

[14] W. He and T.P. Ma, "Inelastic Electron Tunneling Spectroscopy Study of Trap in Ultrathin High-k Gate dielectrics" *Appl. Phys. Lett*. vol. 83, no. 26, pp. 5461-5463, 2003.

[15] W. He and T.P. Ma, "Inelastic Electron Tunneling Spectroscopy Study of Ultrthin HfO_2 and HfAlO," *Appl. Phys. Lett.* vol. 83, no. 13, pp. 2605-2607, 2003.

[16] D. J. Scalapino and S. M. Marcus, "Theory of Inelastic Electron-Molecule Interactions in Tunnel Junction," *Phys. Rev. Lett.* vol. 18, no. 12 pp.459-461, 1967.

[17] G. Salace, and C. Petit, "Inelastic Electron Tunneling Spectroscopy: Capability and Limitations in Metal-Oxide-Semiconductor Devices," *J. Appl. Phys.* vol. 91, pp.5896-5901, 2002.

[18] J. Klein, A. Léger, and M. Belin, "Inelastic-Electron-Tunneling Spectroscopy of Metal-insulator-metal junctions," *Phys. Rev. B,* no. 7, pp.2336-2348, 1968.

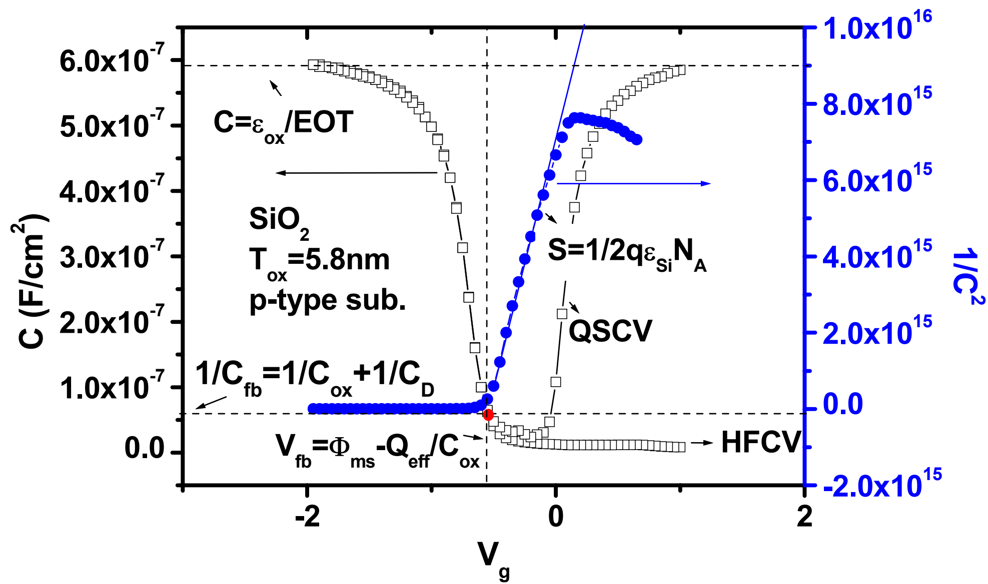


Fig. 2-1 High-frequency (HF) and quasi-frequency (QS) capacitance-voltage (CV) characteristic curves of metal-oxide-semiconductoe capacitance as well as the $1/C^2$-V plot corresponding to the HFCV. Some basic parameters are extracted from these plots according to the described formula.

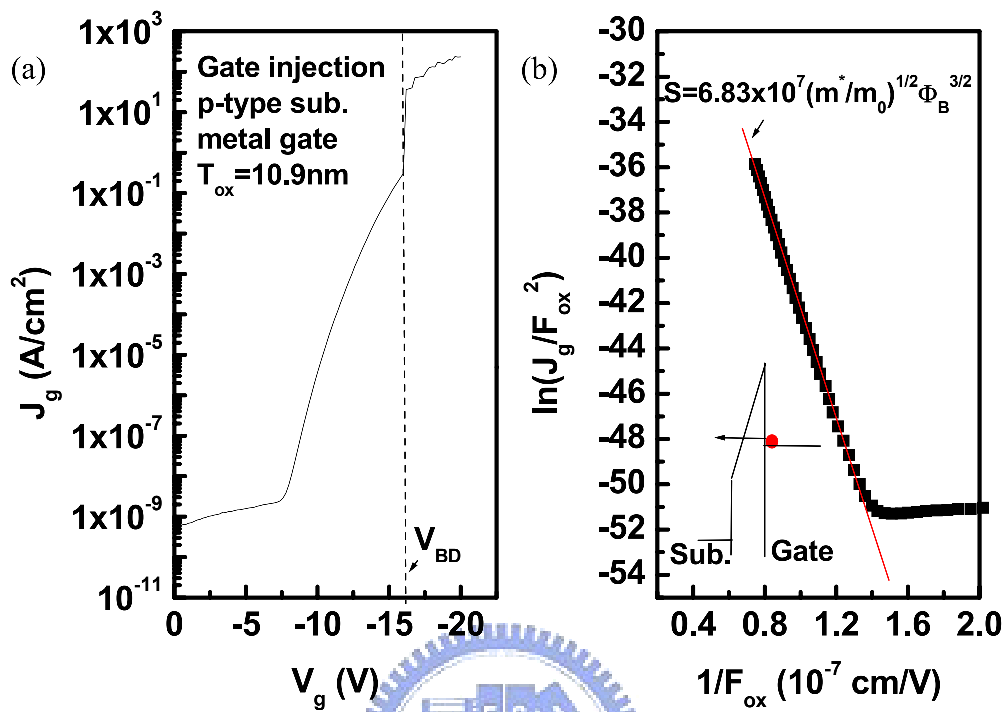


Fig. 2-2 (a)The I-V characteristic curve, and (b) the relative Fowler-Nordheim plot as the electrons are injected from the gate electrode. The inset in (b) shows the band diagram. The abrupt jump at V_{BD} is referred to oxide breakdown voltage.

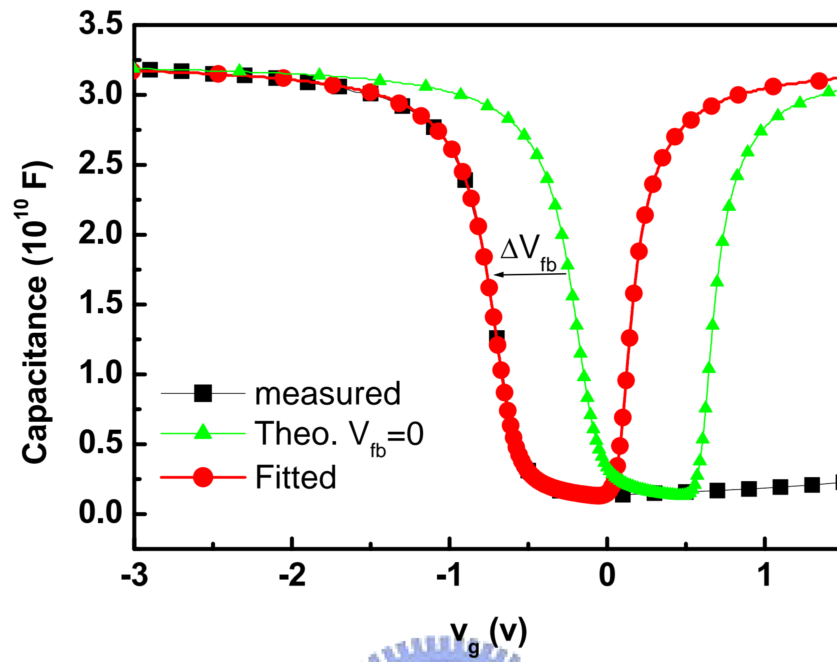
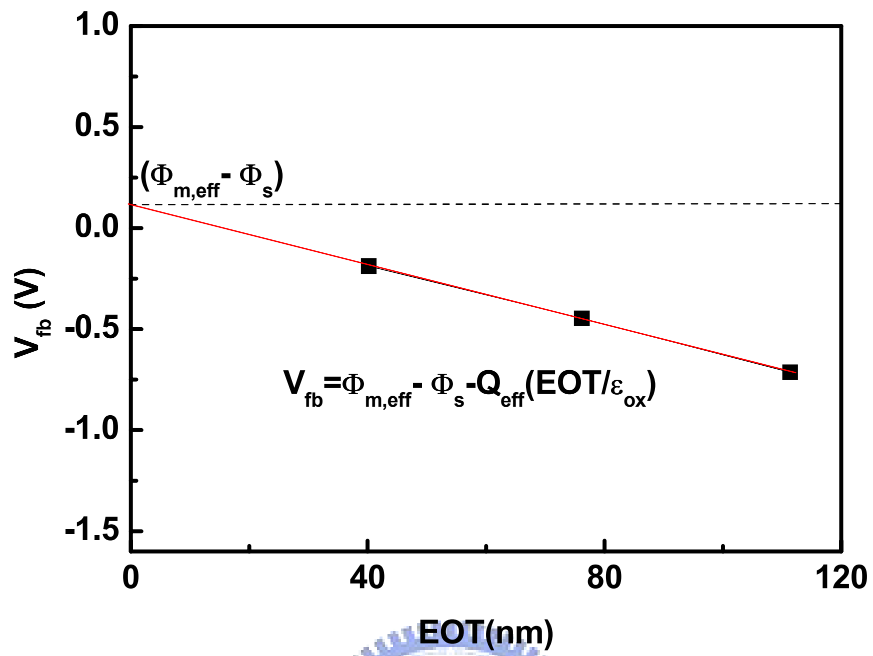


Fig. 2-3 A example for the extraction of the effective work function by fitting the HFCV with a theoretical curve.



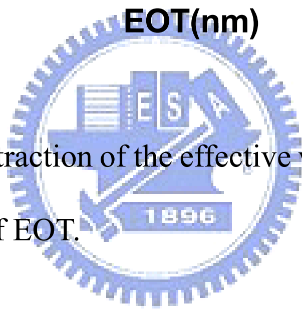


Fig. 2-4 An example for the extraction of the effective work function by a V_{fb}-EOT plot extrapolated at zero of EOT.

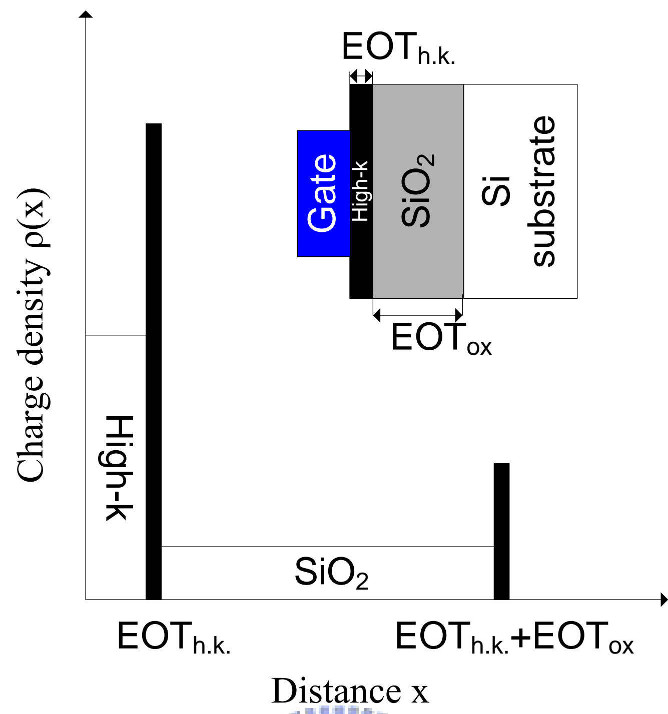


Fig. 2-5 A schematic plot of charge distribution in the gate/HfO_2/SiO_2/Si structure. The bulk charges in SiO_2 and HfO_2 are assumed to be a uniform distribution.

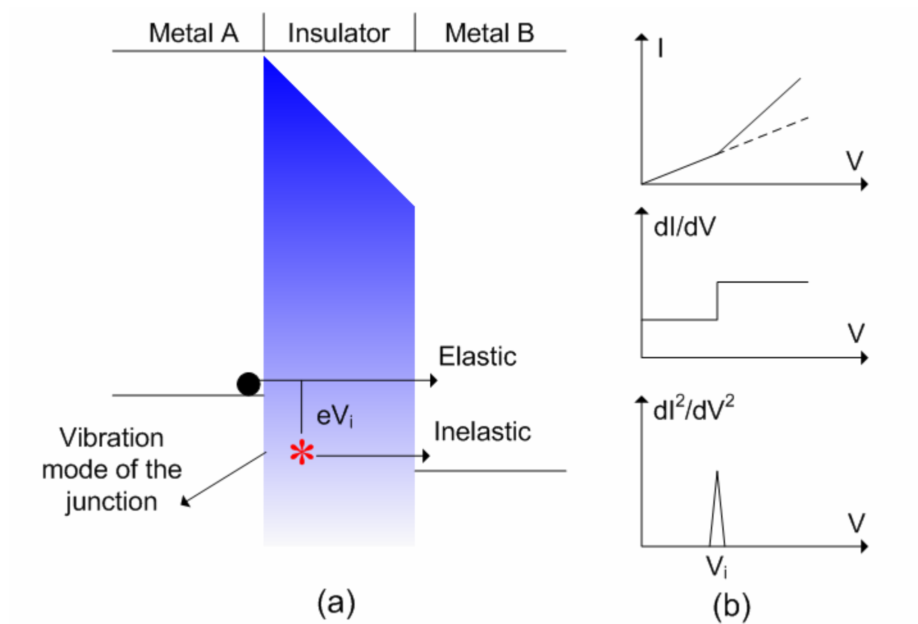


Fig. 2-6 (a) A schematic energy diagram for tunneling between normal metals A and B at a voltage bias V which is the threshold voltage to trigger the inelastic tunneling due to the interaction of electron and vibration mode of energy eV. (b) The relative tunneling current of metal A/insulator/metal B structure at 0 K and the corresponding first derivative and second derivative plots.

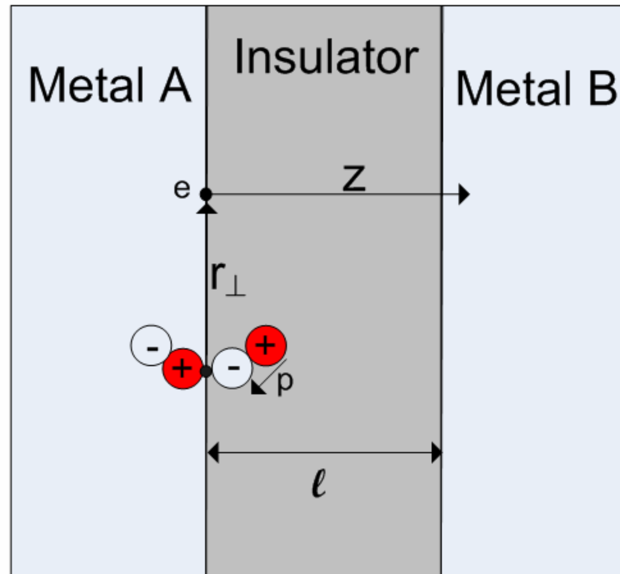
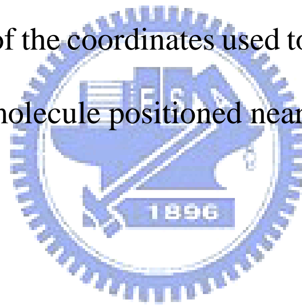
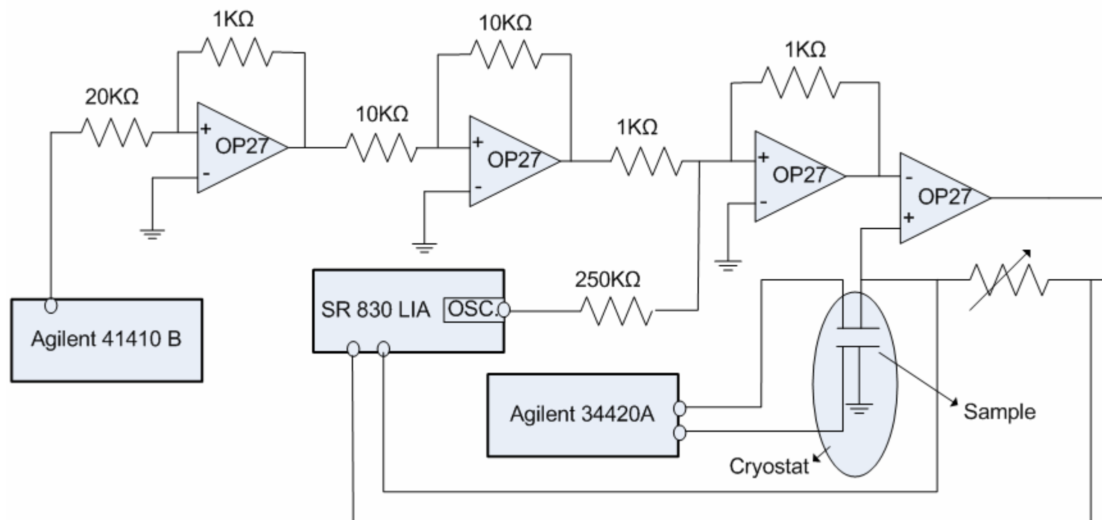
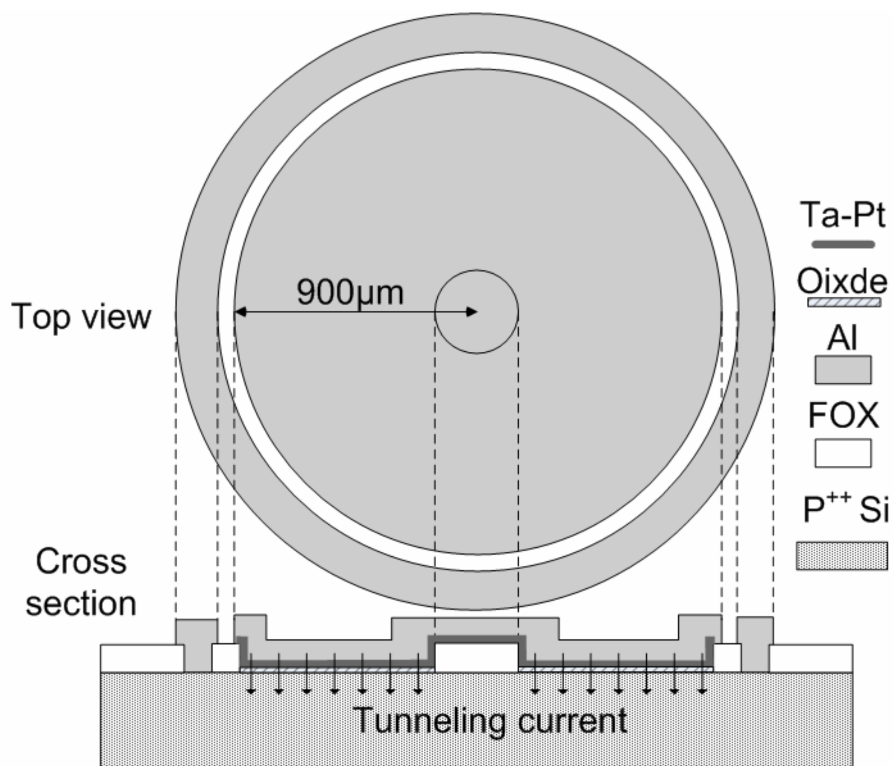


Fig. 2-7 A schematic diagram of the coordinates used to determined interaction energies of a tunneling and a molecule positioned near an insulator-metal A interface.



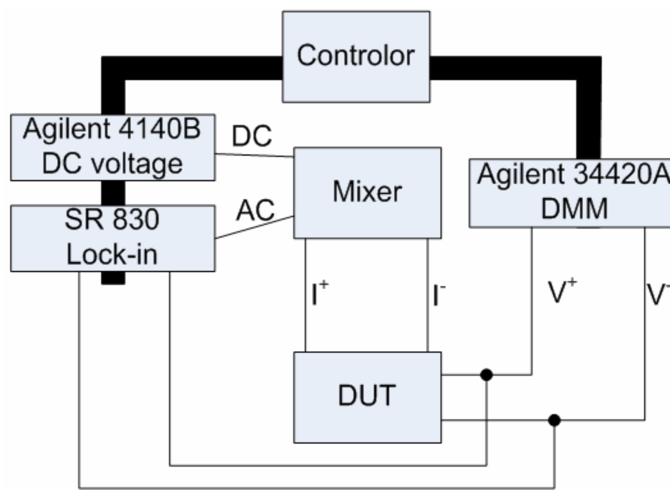


(a)

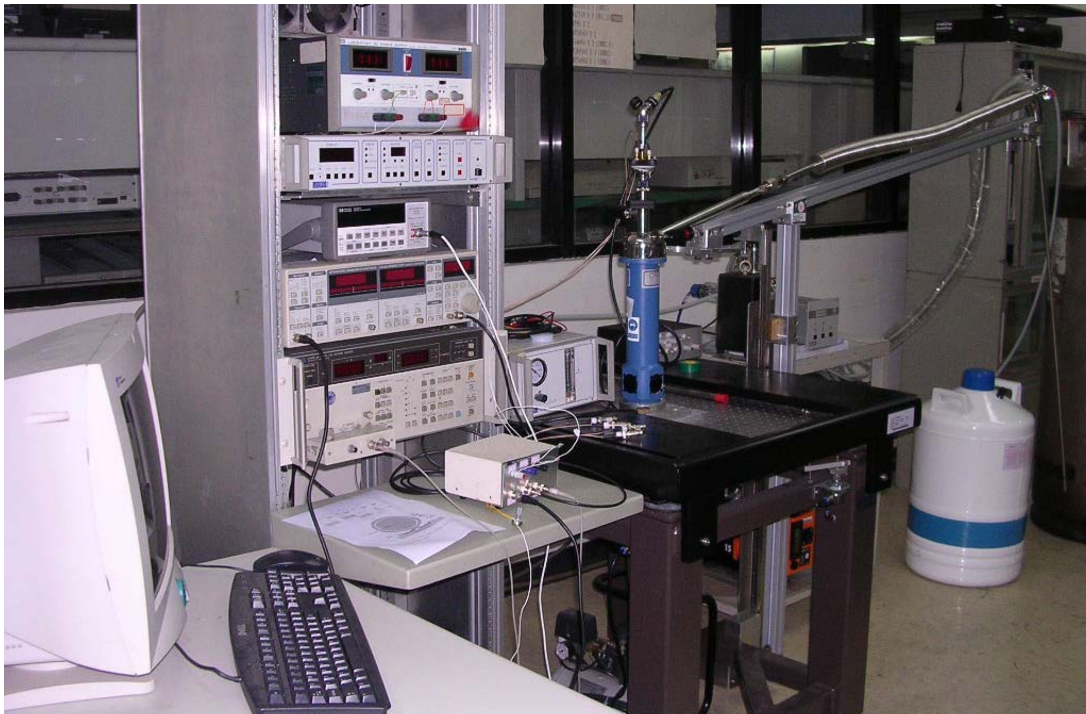


(b)

Fig. 2-8 (a) A circuit schematic diagram of mixer in the IETS system; (b) the schematic structure of the IETS test sample.

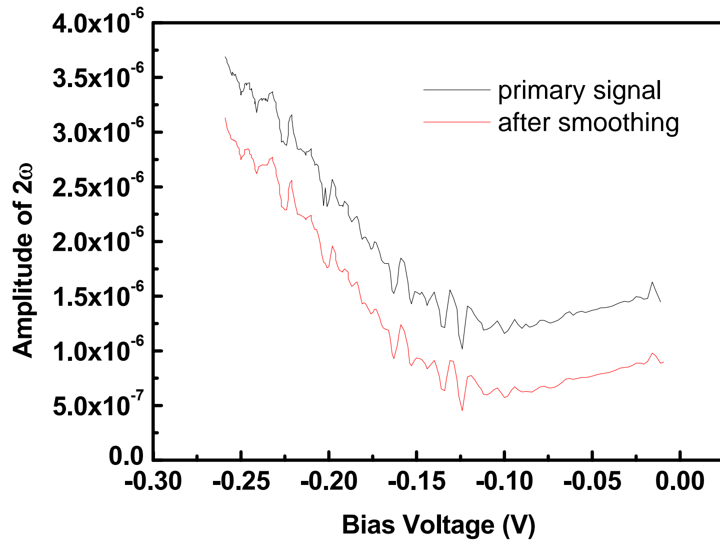


(a)

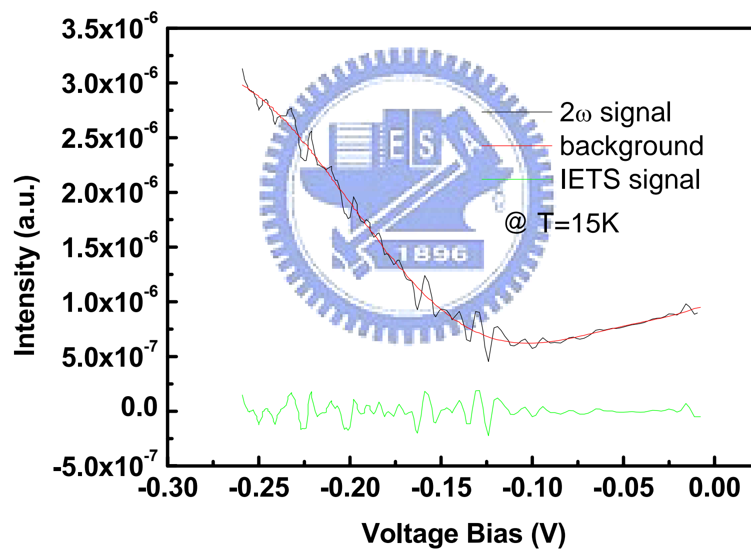


(b)

Fig. 2-9 (a) Schematic diagram and (b) the configuration of the IETS system consisted of a DC voltage source (Agilent 4140B), a lock-in amplifier with an AC source (SR 830), DC voltage meter (Agilent 34420A), a mixer circuit, a cryostat and controller.



(a)



(b)

Fig. 2-10 (a) The primary and smoothed 2ω signal measured by the lock-in amplifier; (b) the IETS signal extracted from the smoothed 2ω signal after subtraction of a polynomial background.

Chapter 3

Molybdenum Nitride Gates

3.1 Introduction

The refectory metal metals are first considered for the metal gate application due to its high thermal stability. Molybdenum (Mo) as a refractory metal is a silvery-white and hard transition metal which has drawn a lot of attention as a single metal gate electrodes of CMOS devices since the work function of Mo on SiO_2 can be adjusted from 4.36 to 4.95 eV [1-4]. Moreover, molybdenum has low resistivity, good thermal stability, and high density which is benefit for ion implantation mask [2, 3]. The work function of molybdenum varies with the bulk microstructure which depends on the deposition and annealing conditions [3, 4]. It can also be modulated by the implantation of nitrogen or argon [4-8]. The work function of the (110)-oriented Mo film is 4.95 eV. After nitrogen implantation to produce Mo_2N phase, the work function can be lowered [4, 8]. However, the work function of Mo film on high-k dielectric, such as $ZrSiO_4$ or Si_3N_4, is quite different from that on SiO_2 and it is also thermally unstable [1, 8]. This observation was explained by the Fermi-level pinning effect due to interfacial dipole formation or minimum interfacial reaction [9-13].

In this chapter, the molybdenum nitride films were deposited by the reactive sputtering. To adjust the nitrogen content, the N_2/Ar gas flow rates were set to various conditions. The work function variation to the phase and nitrogen

composition of molybdenum nitrides films were evaluated, and the Fermi-level pinning phenomenon on SiO_2 and HfO_2 were also investigated after annealing at different annealing temperatures.

3.2 Process Flow

Simple MOS capacitors were used as the test devices. The starting material was 6-inches, (100)-oriented, and boron doped p-type silicon wafer with a resistivity of 15-25 Ω-cm. Oxide layers with various thickness of 40 nm, 70 nm, or 100 nm were thermally grown in dry O_2 ambient. Some wafers were followed to be deposited a 5nm thick HfO_2 layer in a reactively sputtering deposition system with an Ar/O_2 gas-flow ratio of 24 sccm/3 sccm. During deposition, the chamber pressure was at 7.6 mtorr, the DC sputtering power was 100 watts, and neither substrate bias nor heating was intentionally applied. Since the bottom oxide is thick enough, additional oxidation during HfO_2 deposition will not occur. After the preparation of gate dielectrics including SiO_2 single layer and HfO_2/SiO_2 stack, all wafers were annealed at 900 °C in N_2 ambient for 30 sec in a rapid thermal annealing (RTA) system in order to reduce the oxide charges and make the microstructure of HfO_2 stable during the post metal anneal. Lift-off process was then used to define the metal gate electrodes. The MoN_x films with various Mo/N atomic ratios were reactively sputtered at 4.5 mtorr in a DC sputtering system to a thickness of 60nm, with Ar/N_2 gas-flow ratio of 20/0, 20/5, 20/10, and 20/20. During deposition, the DC sputtering power was set to 25 watts and neither substrate bias nor substrate heating was intentionally applied. After lift-off process, wafers were cut into small pieces and were thermally annealed at 400 °C, 500 °C, 600 °C, 700 °C, and 800 °C for 30 sec in N_2 ambient. The gate electrode is circular shape with diameter of 400 μm. Finally,

aluminum film was deposited on wafer backside by thermal evaporation to establish good ohmic contact.

3.3 Results and Discussions

3.3.1 Physical Properties of MoN_x Films

Fig. 3-1 shows the RBS spectrum of MoN_x films deposited on graphite substrate with various Ar/N_2 gas-flow ratios of 20/5, 20/10 and 20/20. Due to the use of graphite substrate instead of Si substrate, the nitrogen signals can be resolved clearly and the N/Mo atomic ratio can be determined precisely. The respective nitrogen contents are 46, 50, and 59 in atomic percentage (at. %) for MON-1, MoN-2 and MoN-3.

Fig. 3-2 shows the XRD spectra of MoN_x films annealed at different temperatures in N_2 ambient for 30sec. Strong MoN (200) phase was detected over all samples. With the increase of annealing temperature, the intensity of MoN (200) phase increases. The increasing intensity of MoN (200) signal implies grain growth during high temperature annealing. Looking into the XRD spectra, the higher the nitrogen content is, the weaker the MoN (200) signal intensity is. This phenomenon implies that the excess nitrogen can make the MoN_x films tend to be amorphous-like. However, the drawback of high nitrogen content in the MoN_x is the high resistivity of the film. As the annealing temperature increases, the sheet resistance decreases due to grain growth but the higher nitrogen content samples still show higher resistivity. Table 3-1 lists the MoN_x films ID and summarizes the above material characteristics.

The Mo-N system phase diagram was reported by Jehn *et al.* in 1978 [14]. When

the nitrogen content is less than 28 at. %, the main phase is Mo. With the increase of nitrogen content ranged from 28.7 at. % to 35 at. %, the main phase is Mo_2N including β-Mo_2N and γ-Mo_2N. The β-Mo_2N will transfer into the γ-Mo_2N phase at higher temperature anneal and higher nitrogen content. The nitrogen content about 40 at. %, the main phase is Mo_3N_2. Furthermore, molybdenum nitride with nitrogen content higher than 40 has the main phase of MoN [14].

3.3.2 Effective Work Function on SiO_2 and HfO_2

Fig. 3-3(a) and (b) show the flat-band voltage (V_{fb}) versus equivalent oxide thickness (EOT) plot of the 500 °C annealed samples with SiO_2 and HfO_2/SiO_2 stack as gate dielectric, respectively, where the V_{fb} values are the average from more than ten samples. The good linearity reflects the validity of effective work function extraction by the linear extrapolation of EOT vs V_{fb} plot. According to the discussion of chapter 2, the Y-axis interception in the V_{fb}-EOT plot of SiO_2 samples represents the work function difference between metal gates and Si substrate ($\Phi_{ms}=\Phi_{m,eff}-\Phi_{Si}$). While the Y-axis interception in the V_{fb}-EOT plot of HfO_2 samples represents the work function difference (Φ_{ms}) plus the effect of the HfO_2/SiO_2 interface charges ($Q_{high\text{-}k}$). With known Si substrate concentration, the work function of metal gate on SiO_2 layer can be extracted. However, on HfO_2 layer, an offset due to the $Q_{high\text{-}k}$ are inevitable. We denote the effective work function of HfO_2 samples as $\Phi_{m,eff}'$ counting for the oxide charge effect.

The $\Phi_{m,eff}$ of MoN_x films on SiO_2 and $\Phi_{m,eff}'$ of MoN_x films on HfO_2 extracted from the V_{fb}-EOT plots are shown in Fig. 3-4 (a) and (b), respectively. The $\Phi_{m,eff}$ and $\Phi_{m,eff}'$ of MoN_x films are almost independent of the annealing temperatures but are dependent on the nitrogen content. This is one of the advantages of MoN_x because the work function of several metals change after thermal annealing [11].

The MoN_x films on SiO_2 with nitrogen contents of 0, 46, 50, and 59 at. % have work function of 4.60 eV, 4.97 eV, 5.03 eV, and 5.11 eV, respectively, after 500 °C annealing. Compared with the MoN-0 sample, the increase of work function ($\Delta\Phi_m$) is 0.37 eV, 0.43 eV, and 0.51 eV as the nitrogen contents are 46, 50, and 59 at. %, respectively. The increase of work function is not linearly dependent on the nitrogen composition. It increases rapidly with the addition of nitrogen and then gradually saturates at high nitrogen concentration. It is concluded that the work function of MoN_x film on SiO_2 can be adjusted by nitrogen incorporation from mid-gap to close to valance band of Si and can be applied as gate electrode of bulk PMOSFETs or fully-depleted SOI PMOSFETs.

The $\Phi_{m,eff}'$ values of MoN_x films on HfO_2 layers with nitrogen contents of 0, 46, 50, and 59 at. % are 4.89eV, 5.31eV, 5.41eV, and 5.37eV, respectively after 500°C annealing. Since the HfO_2/SiO_2 stacks were prepared simultaneously and HfO_2 film is immune to sputtering damage [15], it is reasonable to assume that the Q_{high-k} values are independent of gate electrode. The increase of work function, excluding the effect of Q_{high-k}, is 0.42 eV, 0.52 eV, and 0.48 eV as the nitrogen contents are 46 at. %, 50 at. %, and 59 at. %, respectively. It is important that the magnitude of work function adjustment ($\Delta\Phi_{m,eff}$) on HfO_2 film is nearly the same as that on SiO_2 film.

The $\Phi_{m,eff}$ and $\Phi_{m,eff}'$ of samples with the same MoN_x film and annealing conditions has a difference of about 0.3 eV over all samples. It is indeed necessary to concern if the interface dipole layer at the MoN_x/HfO_2 interface induces the Fermi-level pinning effect. The formation of interface dipole is due to the interaction of gate material and gate dielectric. The pinned Fermi energy should depend on gate material. Furthermore, the reported Fermi-level pinning effects

always force the work function of metal gate tends to the mid-gap of silicon. Both phenomena are not observed in this work. The $\Phi_{m,eff}'$ of MoN_x films on HfO_2 shifts toward the valence band of silicon. Moreover, compared with MoN-0 film, the work function increase of MoN_x films incorporated with the same nitrogen on both SiO_2 and HfO_2 is almost the same. Therefore, it is believed that Fermi-level pinning does not occur in the MoN_x/HfO_2 stack. This is an important advantage for MoN films as gate electrode because it is quite possible that metal gate will integrate with high-k dielectric while Hf-based dielectrics are very promising. The 0.3 eV shift between $\Phi_{m,eff}$ and $\Phi_{m,eff}'$ is attributed to $Q_{high\text{-}k}$ now. After a simple calculation, the $Q_{high\text{-}k}$ is -1×10^{13} cm^{-2}, which is consistent with that reported in literature [16].

The extracted work function of pure Mo film (MoN-0) on SiO_2 is 4.6 eV in this work. This value is different from the results of some literatures. It should be noted that the work function of Mo depends on its crystalline orientation. There are several possible orientations of Mo film deposited by sputtering technique. Among them, only the (110) orientation has work function close to valence band of Si. The other orientations show work function close to mid-gap of Si [3]. Fig. 3-5 shows the electron diffraction pattern of the 600 °C annealed Mo film. Although (110) is the dominant orientation, several crystalline orientations including (200), (211), and (220) are also identified. Therefore, the work function of the MoN-0 sample of 4.6eV is reasonable and is consistent with some other works [2, 3]. Although some recent papers reported that the work function is reduce by implanting nitrogen into Mo(110) film. It should be noted that the nitrogen concentrations are not high enough so that the main phase is Mo_2N in those works [4-8]. In this work, Mo-nitride films were prepared by reactive

sputtering technique and the N/Mo ratio is around unity. As shown in Fig. 3-2, the main phase of our samples is MoN(200) but not Mo_2N. Therefore, the different observation between our work and the references could be explained as follows. As the orientation of Mo film is pure (110), the work function is close to the valance band of Si. Nitrogen implantation destroys the (110) crystalline and the Mo film becomes amorphous. After annealing, the phase is Mo_2N due to insufficient nitrogen concentration. The work function of Mo_2N is close to the mid-gap of Si. Therefore, the higher the nitrogen implantation dose is, the lower the work function is. Once the nitrogen concentration is high enough and the phase changes to be MoN, the work function increases. The work function of MoN is close to the valence band of Si while the work function of Mo_2N is close to the mid-gap of Si.

3.3.3 Electrical Effect of Thermal Annealing

Fig. 3-6 (a) and (b) shows the C-V curves of MoN-0 and MoN-3 samples, respectively, after annealing at different temperatures. The gate dielectric is a 40 nm thick SiO_2 film. Slight distortion of the C-V curves at weak accumulation mode is observed for MoN-0 sample annealed at temperatures below 600°C and the distortion can be recovered after annealing at higher temperature. The distortion should be attributed to the interface states between gate dielectric and Si substrate. Similarly, the C-V curves of HfO_2 samples with HfO_2 (5 nm)/SiO_2 (40 nm) stack gate dielectric were shown in Fig. 3-7. No distortion of C-V curves is observed. This observation confirms that HfO_2 is more immure to sputtering damage than SiO_2 previously [15]. The accumulation capacitance of all SiO_2 and HfO_2 samples is almost the same. This result confirms the thermodynamic stable of HfO_2/thick SiO_2 stack [17]. All of these results indicate that MoN_x films on both

SiO_2 and HfO_2 layers are thermally stable up to 800 °C. This result is better than several metals such as TiN, TaN, and TaSiN [11,12].

3.4 Summaries and Conclusions

We investigated the work function modulation and thermal stability of MoN_x films with different nitrogen contents. The nitrogen contents of MoN_x films analyzed by the RBS are 46 at. %, 50 at. % and 59 at. %. Several important observations are summarized as follows:

i The main phase of the MoN_x films is MoN(200). As the nitrogen content increases, the microstructure of MoN_x film tends to be amorphous-like and the resistivity increases. After high temperature annealing, the phase remains stable and grain size increases slightly.

ii The work function of MoN_x increases with the increase of nitrogen content and tends to saturate at the valence band of Si. No Fermi-level pinning effect is observed on HfO_2 film.

iii The work function and thermal stability of MoN_x show good thermal stability on both SiO_2 and HfO_2 films up to 800°C at least.

According to these results, MoN_x is a promising gate material for PMOSFETs. The high resistivity of the metal nitride could be solved by stack structure which consists of a low resistivity top layer and a metal nitride bottom layer. The top layer is used to reduce the sheet resistance and the bottom layer is used to control the threshold voltage. In this scheme, Mo/MoN stack might be a good choice.

References

[1] Q. Lu, R. Lin, P. Ranade, Y. C. Yeo, X. Meng, H. Takeuchi, T.J. King, C. Hu, H. Luan, S. Lee, W. Bai, C.-H. Lee, D.-L. Kwong, X. Gau, X. Wang, and T.-P. Ma, "Molybdenum Metal Gate MOS Technology for Post-SiO_2 Gate Dielectrics," in *2000 Int. Electron Devices Meet. Tech. Dig.*, pp. 641-644.

[2] M. J. Kim, and D. M. Brown," Mo_2N/Mo Gate MOSFET's," *IEEE Trans. Electron Devices*, ED-30, No. 6, pp. 598-602, 1983.

[3] H. B. Michaelson, , "The Work Function of the Elements and its Periodicity," *J. Appl. Phys.* vol.48, pp. 4729-4733, 1977.

[4] P. Ranade, Y.-K.Choi, Daewon Ha, A. Agarwal, M. Ameen, and T.-J. King, "Tunable Work Function Molybdenum Gate Technology for FDSOI-CMOS," in *2000 Int. Electron Devices Meet. Tech. Dig*, pp.363-366.

[5] R. Lin, Q. Lu, P. Ranade, T.-J. King, and C. Hu," An Adjustable Work Function Technology Using Mo Gate for CMOS Devices," *IEEE Electron Device Lett.*, vol. 23, No. 1, pp. 49-51, 2002.

[6] Y.-K. Choi, L. Chang, P. Ranade, J.-S. Lee, D. Ha, S. Balasubramanian, A. Agarwal, M. Ameen, T.-J. King, and J. Bokor," FinFET Process Refinements for Improved Mobility and Gate Work Function Engineering," in *2002 Int. Electron Devices Meet. Tech. Dig.,* pp. 259-262.

[7] Q. Lu, R. Lin, P. Ranade, T.-J. King, C. Hu, "Metal Gate Work Function Adjustment for Future CMOS Technology," in *Proceedings of the 2001 Symposium on VLSI Technology*, pp. 45-46.

[8] P. Ranade, H. Takeuchi, T.-J. King, and C. Hu, "Work Function Engineering of Molybdenum Gate Electrodes by Nitrogen Implantation," *Electrochemical and*

Solid-State Lett., vol. 4, No. 11, pp. G85-G87, 2001.

[9] Y.-C. Yeo, P. Ranade, T.-J. King, and C. Hu, "Effects of High-κ Gate Dielectric Materials on Metal and Silicon Gate Workfunctions," *IEEE Electron Device Lett.* vol. 23, No.6, pp. 342-344, 2002.

[10] Y.-C. Yeo, T.-J. King, and C. Hu," Metal-dielectric Band Alignment and its Implications for Metal Gate Complementary Metal-Oxide-Semiconductor Technology," *J. Appl. Phys.,* vol. 92, No.12 pp.7266-7271, 2002.

[11] H. Y. Yu, C. Ren, Y.-C. Yeo, J.F. Kang, X. P. Wang, H. H. H. Ma, M.-F. Li, D.S.H. Chan, D.-L. Kwong, "Fermi Pinning-Induced Thermal Instability of Metal-Gate Work Functions," *IEEE Electron Device Lett.* vol. 25, No. 5, pp.337-339, 2004.

[12] C. Ren, H. Y. Yu, J. F. Kang, Y. T. Hou, M.-F. Li, W. D. Wang, D. S. H. Chan, and D.-L. Kwong, "Fermi-level Pinning Induced Thermal Instability in the Effective Work Function of TaN in TaN/SiO_2 Gate Stack," *IEEE Electron Device Lett.*, vol. 25, No. 3, pp. 123-125, 2004.

[13] C. C. Hobbs, L.R.C. Fonseca, A. Knizhnik, V. Dhandapani, S. B. Samavedam, W. J. Taylor, J. M. Grant, L. G. Dip, D. Roan, M. L. Lovejoy, R. S. Rai, E. A. Hebert, H.-H. Tseng, S. G. H. Anderson, B. E. white, P. J. Tobin," Fermi-level Pinning at the Polysilicon/Metal Oxide Interface-Part I," *IEEE Trans. Electron Devices*, vol. ED-51, No.6 pp. 971-977, 2004.

[14] H. Jehn and P. Ettmayer, "The Molybdenum-Nitrogen Phase Diagram," *J. Less-Common Met.*, vol.58, No. 1, pp.85-98, 1978.

[15] Bing-Yue Tsui, Chih-Feng Huang, and Chih-Hsun Lu, "Investigation of Molybdenum Nitride Gate Deposited by Reactive Sputtering on SiO2 and HfO2 for MOSFETs Application," *J. Electrochemical Soc.*, vol.153, No.3,

pp.G197-G202, 2006.

[16] R. Jha, J. Gurganos, Y. H. Kim, R. Choi, J. Lee," A Capacitance-Based Methodology for Work Function Extraction of Metals on high-K" *IEEE Electron Device Lett.* vol.25, No.6, pp.420-422, 2004.

[17] C. F. Huang and B. Y. Tsui," Analysis of NiSi Fully-silicided Gate on SiO_2 and HfO_2 for CMOS Application" in Proceedings of the *2005 Int. Conf. on Solid State Devices and Materials* (SSDM 2005), pp.506..

Table 3-1 MoN_x sample ID and some basic characteristics.

ID	MoN-0	MoN-1	MoN-2	MoN-3
Nitrogen content (at. %)	0	46	50	59
Phase	(110), (200), (211), (220)	MoN (200)	MoN (200)	MoN (200)
Resistivity (μΩ-cm) (As-deposited)	178	650	1158	1736
Resistivity (μΩ-cm) (After 600 °C annealing)	86.7	464	832	1135

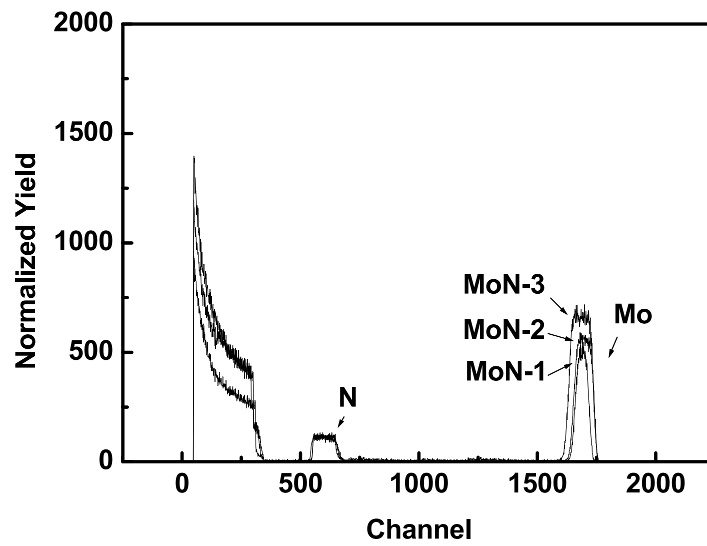


Fig. 3-1 Rutherford-back-scattering (RBS) spectra of MoN_x films on graphite substrate.

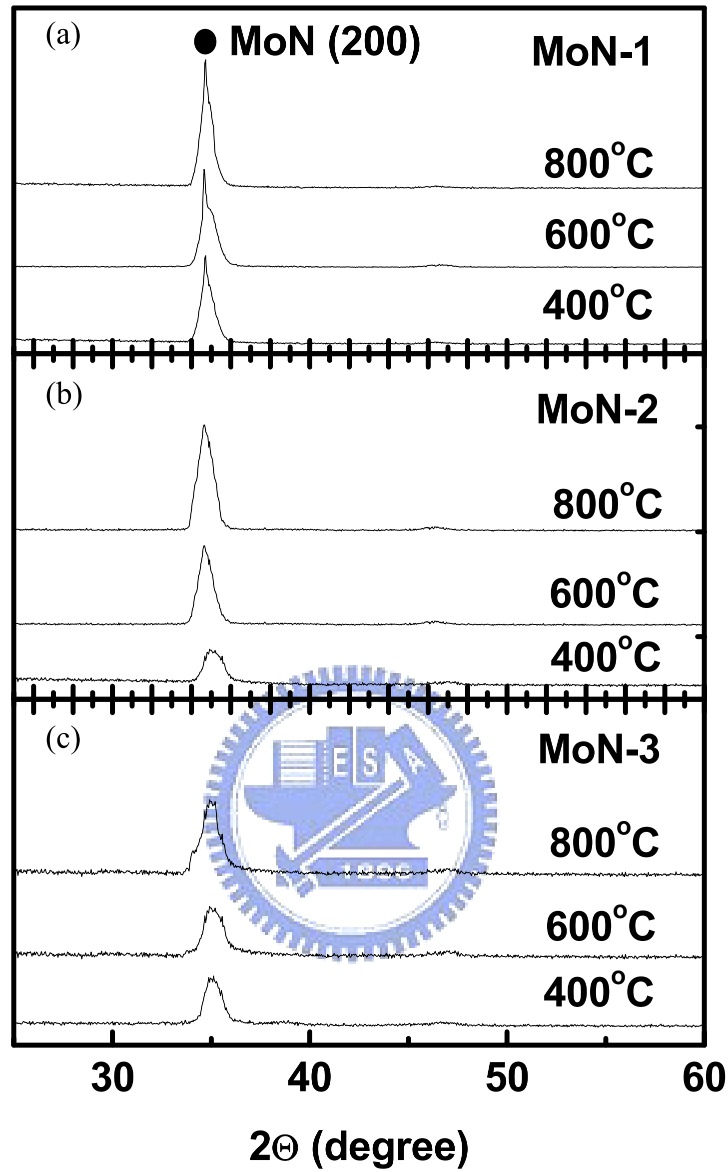
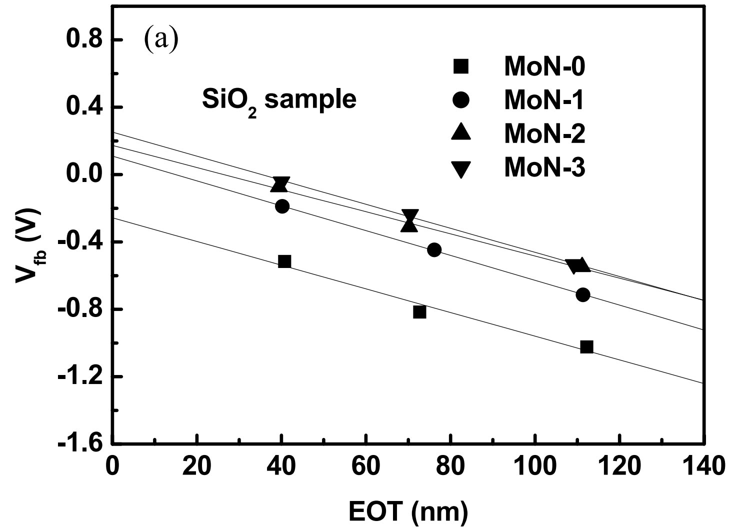


Fig. 3-2 X-ray diffraction (XRD) spectra of (a) MoN-1, (b) MoN-2, and (c) MoN-3 films after annealing at different temperatures



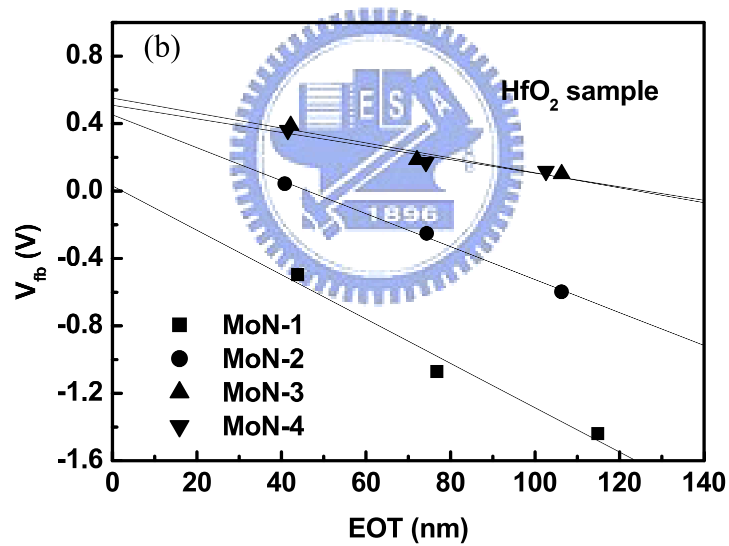
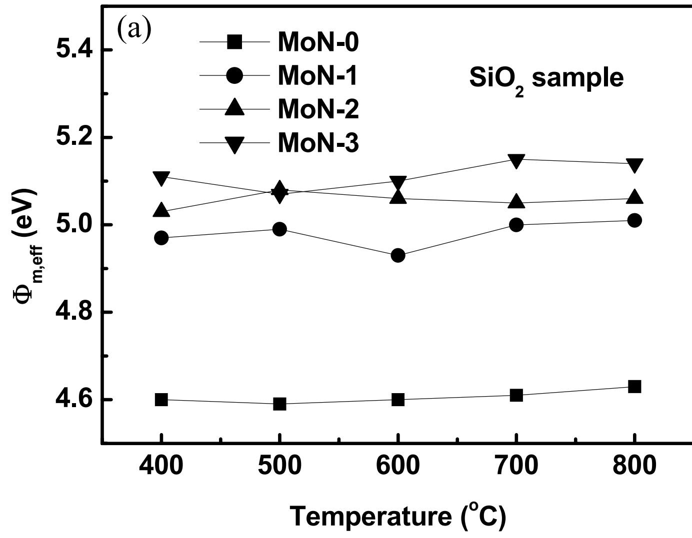


Fig. 3-3 Flat-band voltage (V_{fb}) versus equivalent oxide thickness (EOT) plot of (a) $MoN_x/SiO_2/Si$ structure and (b) $MoN_x/HfO_2/SiO_2/Si$ structure after annealing at 500 °C.



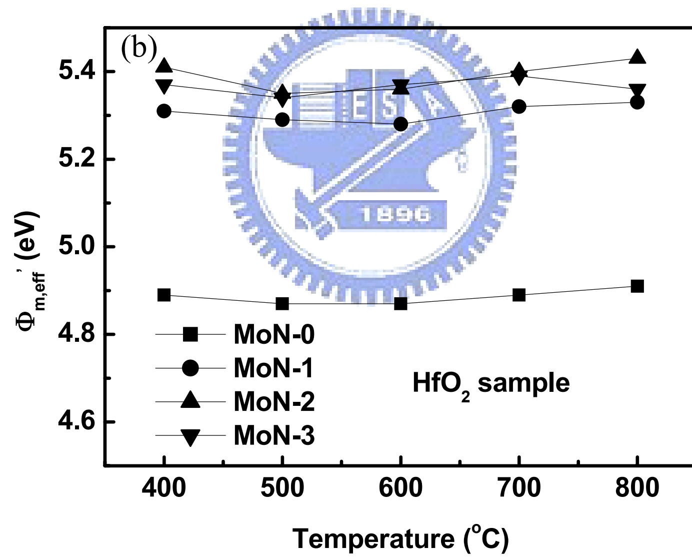


Fig. 3-4 (a) Effective work function of MoN_x film on SiO_2 layer and (b) quasi-effective work function of MoN_x film on HfO_2 layer versus annealing temperature.

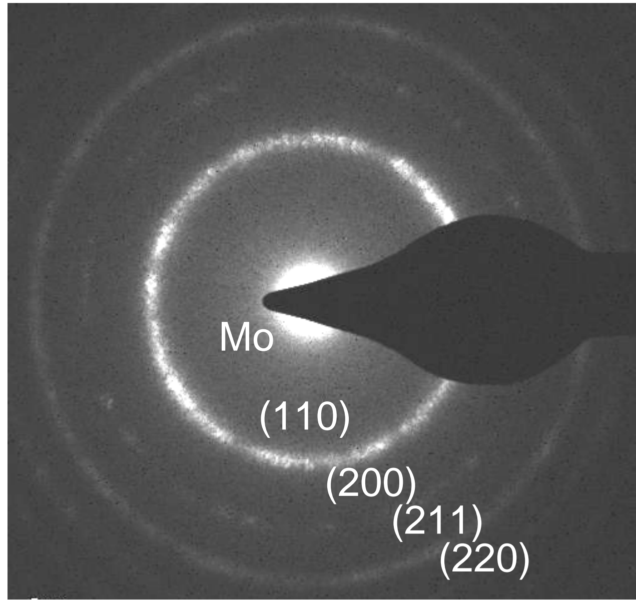
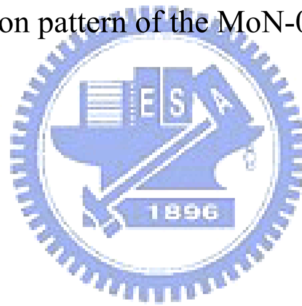
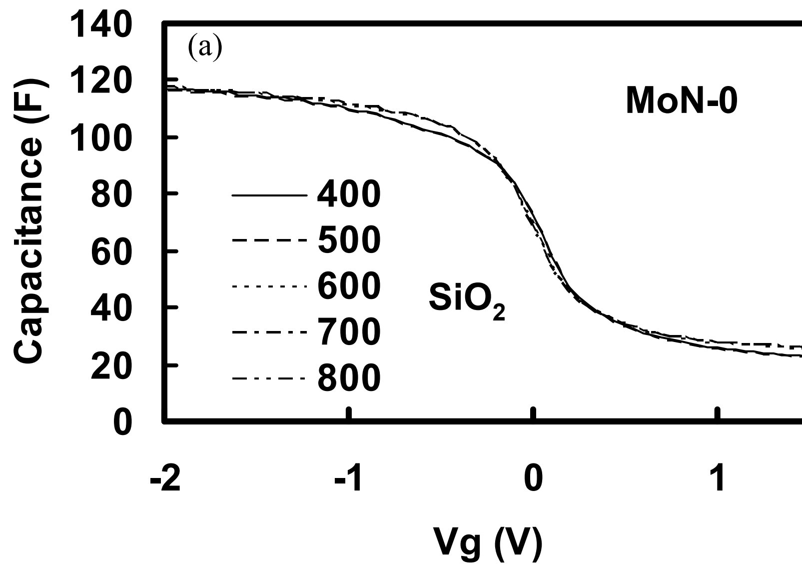


Fig. 3-5 The electron diffraction pattern of the MoN-0 film after annealing at 600 °C.





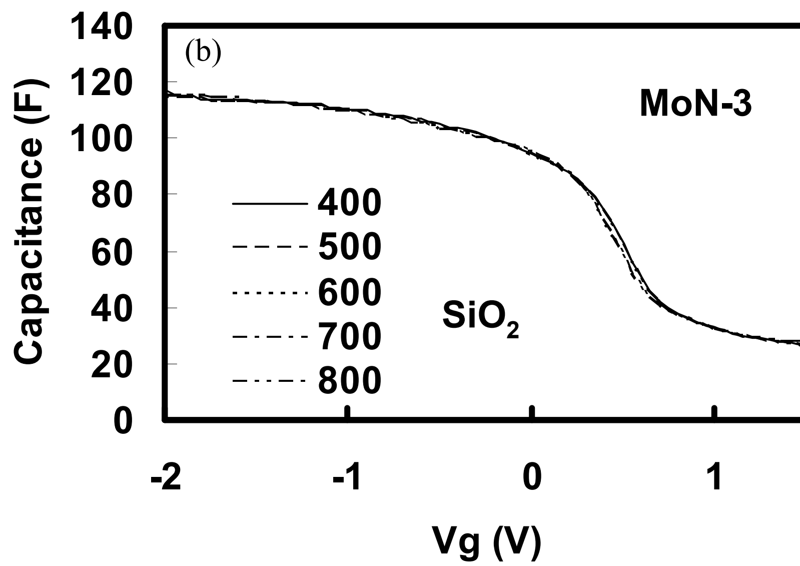
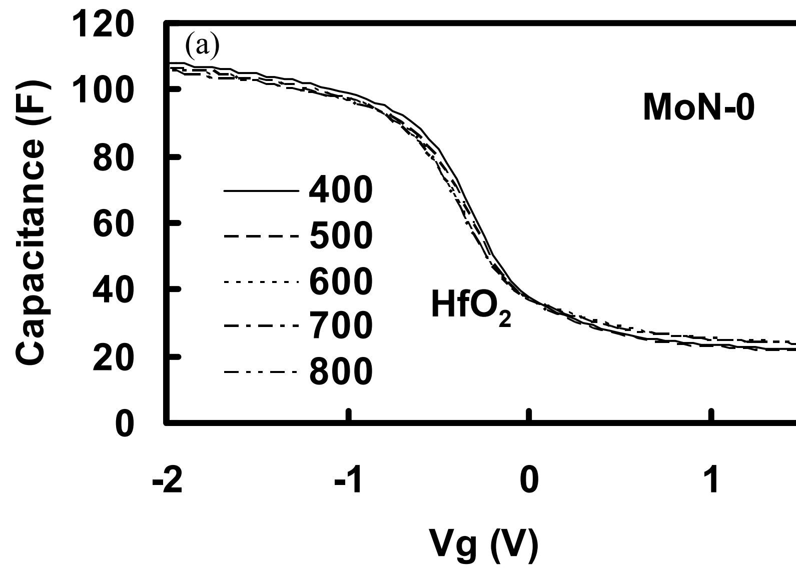


Fig. 3-6 Capacitance-Voltage curves of MoN_x (60 nm)/SiO_2 (40 nm)/Si structure of (a)MoN-0 and (b) MoN-3 films after annealing at different temperatures.



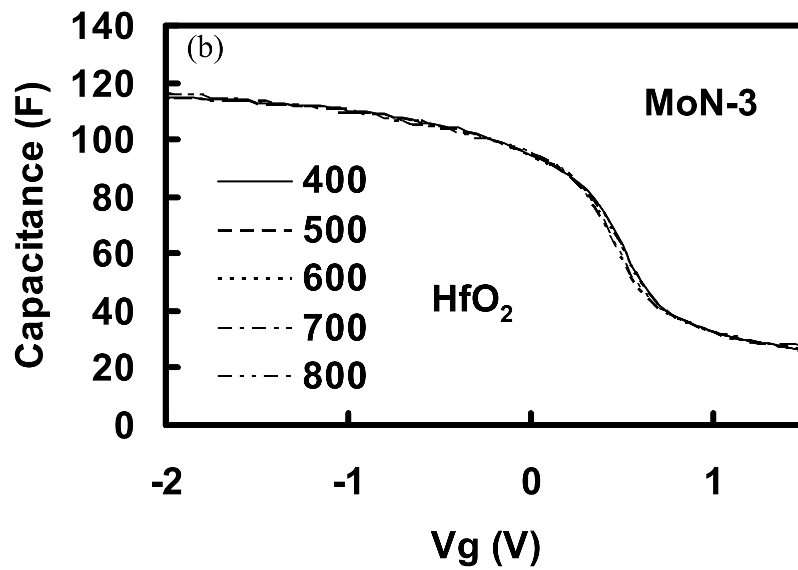


Fig. 3-7 Capacitance-voltage curves of MoN_x (60 nm)/HfO_2 (5 nm)/SiO_2 (40 nm)/Si structure of (a) MoN-0 and (b) MoN-3 films after annealing at different temperatures.

Chapter 4

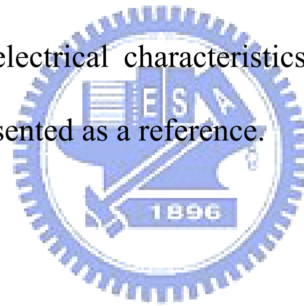
Tungsten Nitride Gates

4.1 Introduction

In chapter 3, we have investigated the work function modulation and thermal stability of molybdenum by the incorporation of nitrogen. Molybdenum has a mid-gap work function and its work function is adjusted to the silicon valence band by the nitrogen composition. Another mid-gap, refractory material, tungsten, also have drawn lots of attentions in the modern very large scale integrated circuits (VLSI) because of low resistivity, easy patterning, and compatible to conventional CMOS technology [1, 2]. In addition, tungsten plug has been used to eclectically connect the multilevel interconnects in mass production. Tungsten nitride also emerged as a promising candidate for Cu diffusion barrier [3]. Recently, tungsten and its nitride (WN_x) were again considered as possible materials of metal gates since metal gates have been considered to replace the poly-Si gates in the later 1970s on account of high-resistivity [1, 2]. Tungsten with a mid-gap work function was considered as a single metal gate of fully depleted silicon-on-insulator (FDSOI) CMOS devices to provide a symmetric threshold voltage for both N-type and P-type devices [1, 4]. WN_x has a work function near the silicon valence band was suggested as the metal gate of p-channel bulk MOSFETs [5, 6]. There are lots of literatures reporting the basic material properties of WN_x films, including the crystalline structure, effects of nitrogen composition, resistivity, patterning, residual stress, and thermal stability

[7-10]. Moreover, there are also some literatures reporting the investigation of electrical characteristics and reliability of MOS devices with tungsten nitride gates [11-14]. An obvious phenomenon of nitrogen desorption from tungsten nitride films was noticed. However, electrical effects of nitrogen desorption in previous studies were not well explored.

In this chapter, the N-rich tungsten nitride films deposited by reactive sputtering were used as electrodes of MOS capacitors with SiO_2 and HfO_2/SiO_2 stack gate dielectrics. The thermal stability and electrical stability were investigated after prolonged (30 min) post-metal annealing at 400-800 °C in N_2 ambient. The effective work function of WN_x gates on SiO_2 and HfO_2 was first studied with the view of incorporated nitrogen concentration and interface bonds between WN_x electrodes and gate dielectrics. The electrical characteristics of MOS capacitors with pure tungsten gate were also presented as a reference.



4.2 Process Flow

Simple MOS capacitors were used as the test devices. The starting material was 6-inch, (100)-oriented, and boron-doped p-type silicon wafers with a resistivity of 15-25 Ω-cm. After wafer clean, oxide layer with various thicknesses of 40nm, 70nm, and 100nm was thermally grown in dry O_2 ambient. Some wafers were then deposited a 5-nm-thick HfO_2 in a reactively sputtered deposition system with an Ar/O_2 gas ratio of 24:3. During the deposition, the chamber pressure was at 7.6m torr, the DC power was 100 Watt and neither substrate bias nor heating was intentionally applied. Because the bottom oxide was thick enough, addition oxidation during HfO_2 deposition did not occur [15]. After the preparation of gate dielectric including SiO_2 single layer and HfO_2/SiO_2 stack, all wafers were annealed at 900 °C in N_2 ambience

for 30 sec in a rapid thermal annealing (RTA) system in order to reduce the oxide charges and make the microstructure of HfO_2 stable during the post-metal annealing. A lift-off process was used to define the metal gate electrodes. The WN_x films with various N/W atomic ratios were reactively sputtered at 4.5mtorr in a dc sputtering system to a thickness of 60nm, with N_2/Ar gas flow ratios of 0:20, 5:20, 10:20 and 20:20, respectively. During deposition, the dc bias was set to 25 Watt and neither substrate bias nor substrate heating was applied. After the lift-off process, wafers were diced and annealed at 400, 500, 600, 700, and 800 °C in N_2 ambient furnace for 30min. Finally, the aluminum film was deposited on the wafer back side by thermal evaporation to establish good ohmic contact. In addition, some samples with thin SiO_2 and HfO_2/SiO_2 stack were also prepared in order to investigate electrical reliability of WN_x metal gate MOS capacitors.

4.3 Results and Discussions

4.3.1 Physical Properties of N-rich WN_x Films

Since the WN_x films were deposited by reactive sputtering method with a mixture of Ar and N_2 gases, the increase of the N_2 gas flow rate makes the increases of the ratio of N composition in WNx films. Figure 4-1 shows the N concentration and relative resistivity of the as-deposited WNx films. The N concentrations in WN_x films analyzed by the RBS are 0, 44, 56 and 61 in atomic percentage for WN-0, WN-1, WN-2, and WN-3, respectively. The N concentration increases rapidly and then tends to saturation. The relative resistivities are 94.09, 1322, 1735, 2438 μΩ-cm. The resisitivities of the WN_x films are much higher than that of W_2N (220 μΩ-cm) and is considered to related

to high WN_x phase [7, 10]. The common tungsten nitride phases exist as W_2N or WN with 33.3 % and 50 % nitrogen stoichiometry. Because high WN_x phases are unstable, any further incorporation of N in WN_x films is difficult [7, 10]. It is believed that the stability of WN_x phase is dependent on its heat of formation. The heat of formation for W_2N is -5.3 Kcal/mol and is slightly more exothermic than that for WN, -3.6 Kcal/mol [16]. The W_2N is the most stable tungsten nitride phase with a cubic crystal lattice.

To investigate thermal properties of the N-rich WN_x films, high temperature annealing was performed. After rapid thermal annealing with a temperature rising rate of 90 °C/sec the WN-2 and WN-3 films were bubbled or cracked. As shown in Fig. 4-2, the WN-2 film was bubbled at 500 °C and cracked at 800 °C. The phenomena of bubbling and then cracking indicate gas explosion in the N-rich films. The thermal desorption spectroscopy (TDS) analysis was used to investigate the temperature dependent thermal desorption of nitrogen. Figure 4-3 shows the background signal and sample signal of the WN-2/SiO_2 sample detected by the TDS spectrometer. The temperature rising rate was set at 20 °C/min. The intensity of background signal is at least one order lower than that of the sample signal and can be neglected. The signal increases rapidly with the temperature up to T=600 °C and slowly reaches a maximum at T=766 °C. At 800°C annealing for 10 min, the signal becomes week. Most nitrogen unwell bounded with tungsten desorbs at temperatures below 766 °C and the phase of WN-2 exists as WN as analyzed by XRD. After the detection of TDS spectrometer, the sample was not bubbled and cracked. The temperature rising rate was hence used as for the furnace anneal in order to prevent films cracking.

The thermally desorption of nitrogen has been detected by the TDS analysis

and then the XPS analysis is used to detect if nitrogen diffuses into the under-layer dielectric. After the tungsten nitride film was removed by the H_2O_2 solution, the exposed surface of dielectric was analyzed. T. Yamada et al. has analyzed the nitrogen distribution in the WN_x/SiO_2 structure [14]. The nitrogen is incorporated at the interface between WN_x and SiO_2 during the reactive sputtering deposition of WN_x films and the concentration is the same after annealing for 30min at 750 °C in N_2 ambience. The binding energy of N core level reflects major $N{\equiv}Si_3$ bond and minor $Si\text{-}N{=}O_2$ bond [14]. It is thus concluded that nitrogen does not diffuse into SiO_2 during annealing. In this work, we focused on the WN_x/HfO_2 structure. Figures 4-4 shows the core level binding energies of O, Hf, and N for the 500 °C annealed WN-1 sample. The binding energy of the Hf core level only reflects Hf-O bonds. The signal of O binding energy is a composition of two signals, one for N-O bonds and one for Hf-O bonds. The signal for the binding energy of N core level is very week and represents N-O bonds. After 800 °C annealing, the binding energies of O and Hf are all the same with those of the 500 °C annealed sample, and the area ratio of N-O and Hf-O is 1/4, which indicates that the high temperature annealing does not enhance the incorporation of N into HfO_2 film.

The phases in the WN_x films were analyzed by the XRD diffraction patterns. Figure 4-5 shows the diffraction patterns of WNx films annealed at the different temperatures for 30 min in N_2 ambient. The as-deposited WN_x films are all amorphous. After 600 °C annealing, the WN-1 film obviously reveals a WN phase. Regarding the WN-2, the film is crystallized at 400 °C, and the main phase is still WN. The WN-3 sample is also crystallized at 400 °C and peak intensity grows with annealing temperature. The increase of nitrogen content to above 50 % does not produce the higher WN_x phase, which is confirmed that the stable highest WN_x

phase is WN after 800 °C anneal [7, 10]. The wider half maximum width in WN-2 and WN-3 XRD spectra of the high temperature annealed samples implies that the excess nitrogen can suppress the grain growth. The relatively weak WO_3 phase is presumably due to the slightly surface oxidation during annealing. The WN phase is stable up to 800 °C while it was reported that WN and W_2N phases can transfer into W phase at 900 °C in N_2 ambient, and the temperature for phase transition is lower in the mixture of H_2 and N_2 ambience [17]. Since the phase of the annealed WN_x film is WN, lots of N atoms in the as-deposited WN_x films are not well bounded to W. Figure 4-6 shows the nitrogen depth profile of WN-2 detected by AES. The intensity of nitrogen signal for the 800°C annealed WN-2/SiO_2 smaple is much lower than that for as-deposited sample and the intensities of W signal for both samples are the same. These physical analyses obviously confirm the nitrogen desorption after annealing, and the desorption occurs in the whole film.

4.3.2 Effective Work Function on SiO_2 and HfO_2

The effective work function was extrapolated from the V_{fb} - EOT plot at the zero EOT. The V_{fb} - EOT plots of 600 ^{o}C annealed samples with SiO_2 and HfO_2/SiO_2 stack as gate dielectric were shown in Fig. 4-7. The good linearity reflects the validity of the effective work function extraction. For the single layer SiO_2 samples, the y-axis intersection in the V_{fb} – EOT plot represents the work function difference between metal gate and Si substrate ($\Phi_{ms} = \Phi_m - \Phi_s$). It should be noted that the HfO_2 sample has a HfO_2/SiO_2 stack which will cause the value of y-axis intersection including not only the work function difference (Φ_{ms}) but also the effect of the HfO_2/SiO_2 interface charges ($Q_{high\text{-}k/SiO2}$). With known Si substrate concentration, the effective work function of metal gate on SiO_2 can be

extracted. However, on the HfO_2, an offset due to the $Q_{high\text{-}k/SiO2}$ is inevitable. The evaluated work function of HfO_2 sample will be denoted as $\Phi'_{m,eff}$ counting for the offset. The detail extraction procedure of effective work function has been discussed in chpater 2. The slopes of the SiO_2 curves are almost the same and reflect the positive oxide charges of about 1.8×10^{11} cm^{-2}. The slopes of HfO_2 curves changes from -0.0166 to -0.0064 and reflect the positive oxide charges decrease from 3.6×10^{11} to 1.5×10^{11} cm^{-2}. The reduction of oxide charges on the HfO_2 samples is reasonable due to the passivation by nitrogen plasma during the reactive sputtering [15]. The effective work functions ($\Phi_{m,eff}$ and $\Phi'_{m,eff}$) of WN_x films on SiO_2 and HfO_2 after annealing are extracted and shown in Fig. 4-8. Regarding the WN_x on SiO_2, the $\Phi_{m,eff}$ of pure W (WN-0) is stable at 4.60 eV below 800 °C, and the $\Phi_{m,eff}$ of tungsten nitrides slightly increase with annealing temperatures and tend to a stable value at 600 °C or 700 °C. The WN-1, WN-2, and WN-3 samples annealed above 700 °C shows a $\Phi_{m,eff}$ of 5.12 eV, 5.05, and 5.0 eV, respectively. The $\Phi_{m,eff}$ are slightly lower below 700 °C. The effective work function of metal gate is influenced by their composition, orientation, phase, and interface interaction [18, 19]. To explain the extracted work functions of WN_x films, these factors must be considered. First, WN_x films do not chemically react with SiO_2 and HfO_2. Although the $N\equiv Si_3$ bonds and N-O bonds were observed at the interface between WN_x and under-layer dielectric, these bonds are stable without further interaction below 800°C [14]. The phase of W film is W (110) below 900°C [10]. The main phase of N-rich WN_x is WN (100) and stable up to 800 °C. Therefore, Nitrogen incorporation in tungsten films forms the tungsten nitride compound (WN), which is the main reason why the effective work function changes from 4.60 eV of W phase to 5.12 eV of WN phase. The only change

during the high temperature annealing is N content in WN_x. It suggests that the thermal instability of effective work function for WN_x comes from the excess nitrogen contents. This result is similar to that of MoN films [20, 21]. The excess nitrogen atoms (pile-up nitrogen atoms) at the interface between MoN and SiO_2 reduce the $\Phi_{m,eff}$ value of MoN. Moreover, the nitrogen concentration at interface determined the reduction of $\Phi_{m,eff}$ [20, 21]. Therefore, it is reasonable that WN-2 and WN-3 with more excess nitrogen have the relative lower $\Phi_{m,eff}$ to WN-1. The WN_x films annealed at low temperature have high contents of excess nitrogen to lower the $\Phi_{m,eff}$. After high temperature annealing, most excess nitrogen desorbs so that the $\Phi_{m,eff}$ reaches a high and stable value. Similarly, the results of WN_x/HfO_2 samples are determined by the nitrogen contents. The $\Phi'_{m,eff}$ of pure W and WN-1 on HfO_2 are 4.75 eV and 5.18 eV after annealing at 800 °C.

It is important to note that the magnitude of work function adjustment ($\Delta\Phi'_{m,eff}$) on HfO_2 film is smaller than the $\Delta\Phi_{m,eff}$ on SiO_2 film. It means that the incorporation of nitrogen content by reactive sputtering will cause the change of $\Phi_{m,eff}$ and $\Phi'_{m,eff}$ is different on different underlying dielectric. The difference may be affected by the work function offset due to interface charge ($Q_{high\text{-}k/SiO2}$) between HfO_2 and SiO_2 and Fermi-level pinning effect [22]. First, the effect of the interface charges is discussed. For the pure W samples, the 0.15 eV shift between $\Phi_{m,eff}$ and $\Phi'_{m,eff}$ for pure W on SiO_2 and HfO_2 samples could be mainly attributed to the $Q_{high\text{-}k/SiO2}$ between HfO_2 and SiO_2. Since the pure W has a work function at mid-gap of silicon energy band, it is free of Fermi-level level pinning [22]. After a simple calculation, the $Q_{high\text{-}k/SiO2}$ is $-5x10^{12}$ cm^{-2}, which is consistent with that reported in literature [23]. No matter how the nitrogen plasma have passivated or created the interface charges during the reactive sputtering, the smaller $\Delta\Phi'_{m,eff}$ for

HfO_2 samples should be induced by a relative small $Q_{high-k/SiO2}$ for the tungsten nitride samples in comparison with the pure tungsten sample. Actually, the nitrogen plasma may passivate the interface charges mentioned in the last paragraph. However, it is impossible to completely passivate all the interface states. The difference of $\Phi'_{m,eff}$ and $\Phi_{m,eff}$ between HfO_2 and SiO_2 samples due to the $Q_{high-k/SiO2}$, hence, should be smaller than 0.15 eV. However, the difference of $\Phi'_{m,eff}$ and $\Phi_{m,eff}$ between HfO_2 and SiO_2 samples is larger than 0.15 eV. Comparing pure W with WN-1 samples, the $\Delta\Phi_{m,eff}$ on HfO_2 (SiO_2) film after annealing at 400 °C and 800 °C is 0.31 eV (0.50 eV) and 0.37 eV (0.52 eV), respectively. Moreover, comparing pure W with WN-3 samples, the $\Delta\Phi_{m,eff}$ on HfO_2 (SiO_2) film after annealing at 400°C and 800°C is 0.05 eV (0.3 eV) and 0.21 eV (0.4 eV), respectively. The difference of $\Phi'_{m,eff}$ and $\Phi_{m,eff}$ between HfO_2 and SiO_2 sample are more and more large with the increase of nitrogen content. It is indeed necessary to concern if the interface dipole layer at the WN_x/HfO_2 interface induces the Fermi-level pinning effect. The formation of interface dipole is due to the interaction of gate material and gate dielectric. The pinned Fermi energy should depend on gate material. Furthermore, the reported Fermi-level pinning effects always force the work function of metal gate tends to the mid-gap of silicon. Although the $\Phi_{m,eff}$ of WN_x films on HfO_2 shifts toward the valence band of silicon. Compared with WN-0 film, the work function increase of WN_x films incorporated with the same nitrogen on HfO_2 is much smaller than that on SiO_2. Therefore, it is believed that Fermi-level pinning does occur in the WN_x/HfO_2 stack. Because of the Fermi-level pinning effect, WN_x films is not suitable for bulk PMOSFET as integrating with HfO_2. However, since the Fermi-level pinning effect only shifts the work function by 0.15-0.2 eV, WN_x could be applied on p-type fully depleted

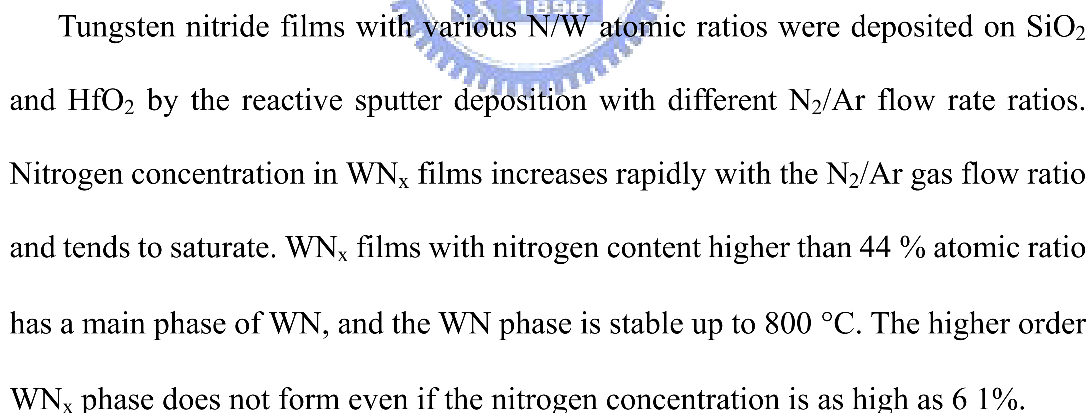
SOI devices.

4.3.3 Electrical Effect of Thermal Annealing

Figures 4-9 (a) and (b) show the statistic distributions of time-to-breakdown (T_{bd}) of the WN_x/SiO_2 (5 nm)/Si and WN_x/HfO_2 (2 nm)/SiO_2 (2 nm)/Si MOS capacitors with different post-metal annealing temperatures. The capacitance equivalent thickness (CET) of the HfO_2/SiO_2 stack is 3.5 nm. Samples were stressed at constant voltage of Vg=-6 V for WN_x/SiO_2 capacitors and at Vg=-4.5 V for $WN_x/HfO_2/SiO_2$ capacitors because these samples have different thickness of gate dielectric. Both are in gate injection mode and the electric field within the SiO_2 layer of HfO_2 and SiO_2 samples is -12 MV/cm. It is found that the integrity of both SiO_2 and HfO_2/SiO_2 dielectric with pure W gate is not degraded during the post-metal annealing. On the other hand, the SiO_2 dielectric with WN-1 gate is seriously degraded during the 700 °C post-metal annealing while the HfO_2/SiO_2 dielectric with WN_x gates are not degraded. Compared with pure W SiO_2 samples, the 500 ^oC and 600 ^oC WN-1 SiO_2 samples performed better time-dependent dielectric breakdown (TDDB) characteristics. To better understand the TDDB characteristics, the I-V curves as shown in Fig. 4-10 were used to evaluate the effect of work function. Work functions of pure W is lower than that of WN-1 and the same bias stress will cause relative low leakage current for the WN-1 due to high potential barrier for electron on both SiO_2 and HfO_2 sample. This means that the stress condition at the same bias for the pure W is stronger than that of WN-1 sample. This is the reason why the SiO_2 samples with higher work function have better TDDB characteristics [24]. The degradation of 700 ^oC SiO_2 sample should be due to the creation of oxide defects, which is confirmed by the increased leakage current at low bias as shown in the I-V curves. Although the stress

condition is stronger for the pure W HfO_2 samples, the TDDB characteristics of pure W samples are almost the same as that of WN-1 HfO_2 samples. It is reported that the time-to-breakdown (or charge-to-breakdown) for thin oxide (thickness less than 5 nm) is difficult to be distinguished at high electrical stress [25]. This is the possible reason why that the TDDB characteristics are undistinguishable when comparing these pure W with WN-1 HfO_2 samples. Moreover, the leakage current of the 700 ^{o}C HfO_2 are not degraded. The degradation of the 700 ºC annealed WN-1/SiO_2 sample is suspected to be due to the stress exerted by the nitrogen desorption. Since the single SiO_2 layer can not sustain the post metal annealing, the HfO_2 layer of HfO_2/SiO_2 stack acts as an effective buffer layer to avoid SiO_2 degradation from WN_x gate during post-metal annealing.

4.4 Summaries and Conclusions



Tungsten nitride films with various N/W atomic ratios were deposited on SiO_2 and HfO_2 by the reactive sputter deposition with different N_2/Ar flow rate ratios. Nitrogen concentration in WN_x films increases rapidly with the N_2/Ar gas flow ratio and tends to saturate. WN_x films with nitrogen content higher than 44 % atomic ratio has a main phase of WN, and the WN phase is stable up to 800 °C. The higher order WN_x phase does not form even if the nitrogen concentration is as high as 6 1%.

Tungsten nitride has an effective work function near the valance band of silicon so that it satisfies the metal-gate demand of p-type MOSFETs to provide suitable threshold voltage. However, the excess nitrogen will make the work function unstable during the high temperature annealing, and may cause the WN_x film crack during the high temperature annealing with a rapid temperature rising rate. The post-metal thermal annealing must be carefully controlled to avoid bubbling and

cracking. The excess nitrogen in WN_x films can cause the effective work function lowering. Weak Fermi-level pinning effect is observed on HfO_2 film. In this case, WN_x is not suitable to be meal gate of bulk p-type MOSFTEs. Fully depleted SOI devices require work function lower than Si valence band for pMOSFET and higher than Si conduction band for nMOSFET. Therefore, WN_x/HfO_2 gate stack can be applied to p-type FD-SOI devices. The good integrity of the WN_x/HfO_2 gate stack also suggests WN_x as a promising gate material.

References

[1] P. L. Shah, "Refractory metal gate processes for VLSI applications," *IEEE Trans. Electron Devices*, vol. ED-26, No. 4 pp. 631-640, 1979.

[2] K. Sinha, T. E. Smith, T. T. Sheng, and N.N. Axelrod, "Control of Resistivity, Microstructure, and Stress in Electron Beam Evaporated Tungsten Films," *J. Vac. Sci. Technol.*, vol.10, No. 3, pp. 436-443, 1973.

[3] H. Lee and K. Yong, "Diffusion barrier properties of metalorganic chemical vapor deposition -WN_x compared with other barrier materials," *J. Vac. Sci. Technol. B*. vol. 22, No.5, pp.2375-2379, 2004.

[4] H. Noda, H. Sakiyama, Y. Goto, T. Kure and S. Kimura," Tungsten Gate Technology for Quarter-Micron Application," *Jpn. J. Appl. Phys.* vol. 35, pp.807-811, 1996.

[5] D.-G. Park, Z.J. Luo, N. Edleman, W. Zhu, P. Nguyen, K. Wong, C. Cabral, P. Jamison, B. H. Lee, A. Chou, M. Chudzik, J. Bruley, O. Gluschenkov, P. Ronsheim, A. Chakravarti, R. Mitchell, V. Ku, H. Kim, E. Duch, P. Kozlowski, C. D. Emic, V. Narayanan, A. Steegen, R. Wise, R. Rengarajan, H. Ng, A. Sekiguchi, and C. H. Wann, "Thermally Robust Dual-Work Function ALD-MNx MOSFETs Using Conventional CMOS Process Flow," in *Proceedings of the 2004 Symposium on VLSI Technology*, pp.186-187.

[6] B. Claflin, M. Binger, and G. Lucovsky, "Interface studies of tungsten nitride and titanium nitride composite metal gate electrodes with thin dielectric layers," *J. Vac. Sci. Technol. A*. vol.16, No.3, pp.1757-1761, 1998.

[7] J. Lin, A. Tsukune, T. Suzuki, and M. Yamada, Different effects of annealing temperature on resistivity for stoichiometric," Different effect of annealing

temperature on resistivity for stoichiometric, W rich, and N rich tungsten nitride films," *J. Vac. Sci. Technol. A* , vol.17 No. 3, pp.936-938, 1999.

[8] H.-T. Chiu, and S.-H. Chuang, "Tungsten Nitride Thin Film Prepared by MOCVD," *J. Mater. Res.*, vol. 8, No. 6, pp.1353-1360, 1993.

[9] Y. G. Shen, and Y. W. Mai," Composition, residual stress, and structural properties of thin tungsten nitride films deposited by reactive magnetron sputtering," *J. Appl. Phys.* vol. 88, No.3, pp.1380-1388, 2000.

[10] C. C. Baker and S. Ismat Shah, "Reactive sputter deposition of tungsten nitride thin films," *J. Vac. Sci. Technol. A,* Vol. 20, No.5, pp.1699-1703, 2002.

[11] P.-C. Jiang and J. S. Chen, "Effects of Post-Metal Annealing on Electrical Characteristics and Thermal Stability of $W_2N/Ta_2O_5/Si$ MOS Capacitors," *J. Electrochem. Soc.* vol. 151, No. 11, pp.G751-G755, 2004.

[12] H.-J. Cho, T.-H. Cha, K.-Y. Lim, D.-G. Park, J.-Y. Kim, J.-J. Kim, S. Heo, I.-S. Yeo, and J. W. Park, "Reliability Characteristics of $W/WN/TaO_xN_y/SiO_2/Si$ Metal Oxide Semiconductor Capacitors," *J. Electrochem. Soc.* vol. 149, No. 7, pp.G403-G407, 2002.

[13] J. W. Lee, C. H. Han, J.-S. Park, and J. W. Park, "Electrical Characteristics and Thermal Stability of W, WN_x, and TiN Barriers in $Metal/Ta_2O_5/Si$ Gate Devices," *J. Electrochem. Soc.* vol. 148, No. 3, pp.G95-G98, 2001.

[14] T. Yamada, M. Moriwaki, Y. Harada, S. Fujii, and K. Eriguchi, "Effects of the sputtering Deposition Process of Metal Gate Electrode on the Gate Dielectric Characteristics," *Microelectronics reliability* vol. 41, pp.697-704, 2001.

[15] B. Y. Tsui, C. F. Huang and C. H. Lu, "Investigation of Molybdenum Nitride Gate on SiO_2 and HfO_2 for MOSFET Application," *J. Electrochem. Soc.* vol. 153, No.3, pp. G197-G202, 2006.

[16] J. W. Klaus, S. J. Ferro, S. M. George, "Atomically Control Growth of Tungsten and Tungsten Nitride Using Sequential Surface reaction," *Appl. Surf. Sci.* vol. 162-163 pp.479-491, 2000.

[17] Y. Z. Hu and S.-P. Tay, "Applications of Rapid Thermal Process to Nitridation of Tungsten and Denudation of WNx for Poly-Si/Metal Gates," *10th IEEE International Conf. on Adv. Thermal Processing of Semiconductors* (RTP's 2002) pp.125-130.

[18] K. Nakajima, Y.Akasaka, M. Kaneko, M. Tamaoki, Y. Yamada, T. Shimizu, Y. Ozawa and K. Suguro, "Work Function Controlled Metal Gate Electrode on Ultrathin Gate Insulators," in *Proceedings of the 1999 Symposium on VLSI Technology*, pp. 95-96.

[19] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, and T. Kunio," A novel W/TiNx metal gate CMOS technology using nitrogen-concentration-controlled TiNx," in *1999 Int. Electron. Devices Meet. Tech. Dig*, pp.253-256.

[20] P. Ranade, Y.-K.Choi, Daewon Ha, A. Agarwal, M. Ameen, and T.-J. King, "Tunable work function molybdenum gate technology for FDSOI-CMOS," in *2000 Int. Electron. Devices Meet. Tech. Dig,* pp.363-366.

[21] P. Ranade, H. Takeuchi, T.-J. King, and C. Hu, "Work Function Engineering of Molybdenum Gate Electrodes by Nitrogen Implantation," *Electrochemical and Solid-State Lett.,* vol. 4, No. 11, pp. G85-G87, 2001.

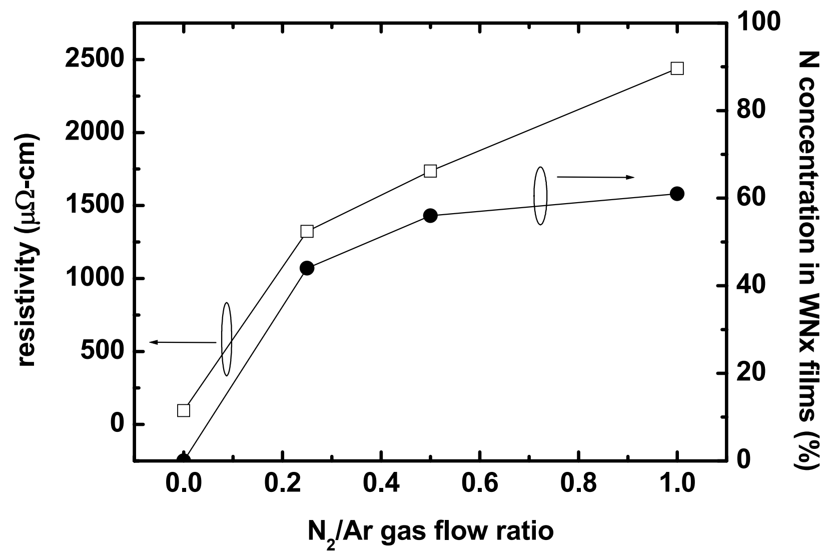
[22] Y.-C. Yeo, P. Ranade, T.-J. King and C. Hu, "Effects of High-k Gate Dielectric Materials on Metal and Silicon Gate Workfunctions," *IEEE Electron Device Lett.* vol. 23 no.6 pp.342-344, June 2002.

[23] R. Jha, J. Gurganos, Y. H. Kim, R. Choi, J. Lee," A capacitance-based methodology for work function extraction of metals on high-k," *IEEE Electron*

Device Lett. vol. 25, No. 6, pp.420-422, 2004.

[24] D. J. DiMaria, " Dependence of Gate Work Function of Oxide Charging, Defect Generation, and Hole Currents in Metal-Oxide-Semiconductor Structures," *J. Appl. Phys.*, vol. 81, no. 7, pp.3220-3226, 1997.

[25] C. Chen, S. Holland, and C. Hu, "Oxide Breakdown dependence on Thickness and Hole Current-enhanced Reliability of Ultra Thin Oxide,"in *1986 Int. Electron. Devices Meet. Tech. Dig*, pp.660-663.



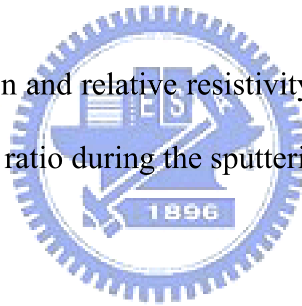


Fig. 4-1 Nitrogen concentration and relative resistivity of the WN_x films as a function of the N_2/Ar gas flow ratio during the sputtering deposition.

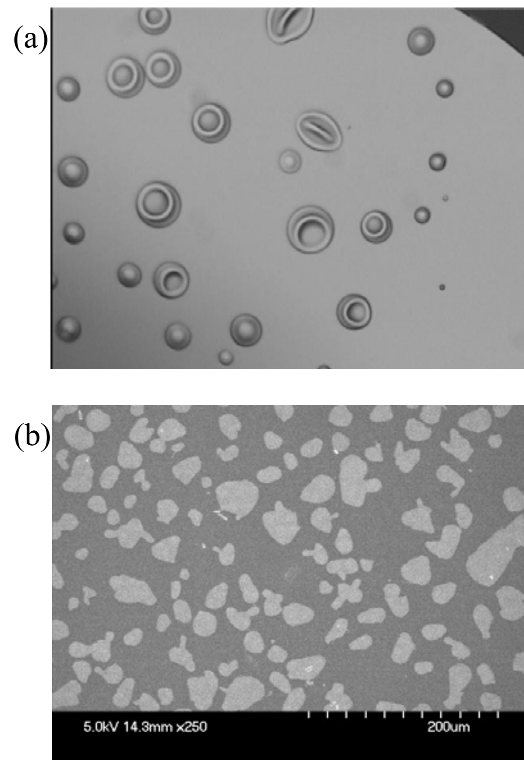


Fig. 4-2 Plane view SEM micrographs of the WN-2 films after annealing at (a) 500 °C and (b) 800 °C in a rapid thermal annealing system. The temperature rising rate is 90 °C/sec.

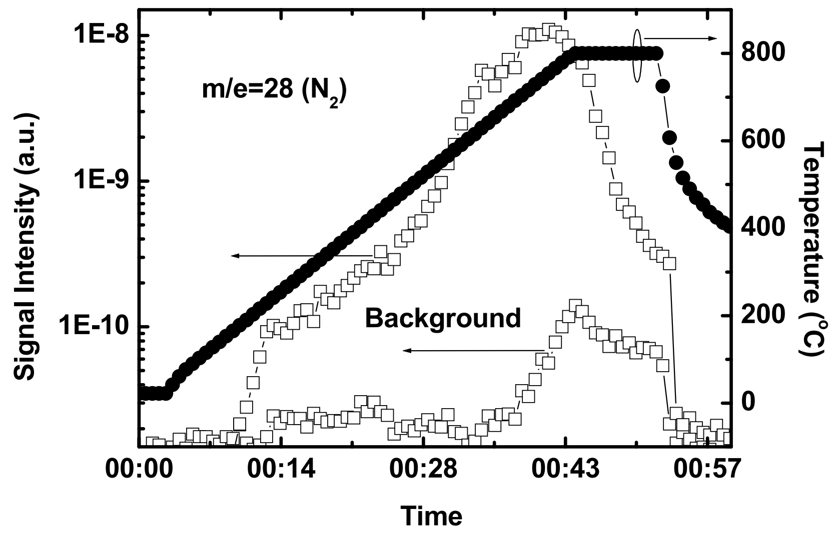


Fig. 4-3 Thermal desorption spectrum of the as-deposited WN-2 film.

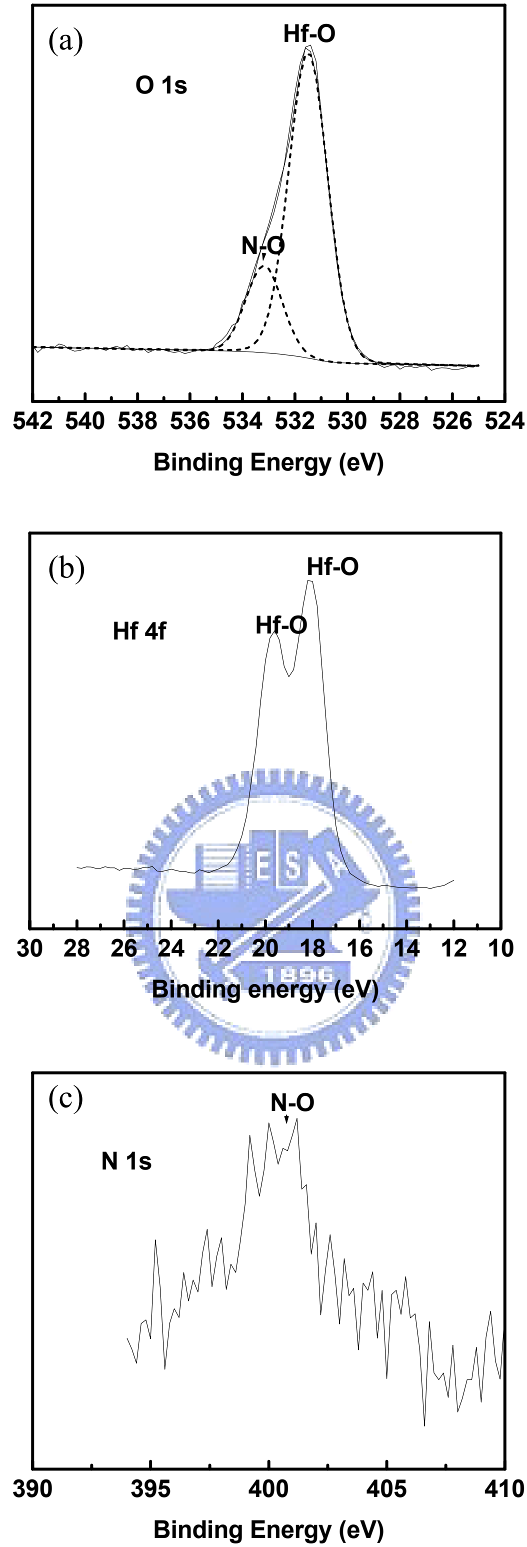


Fig. 4-4 X-ray photoelectron spectroscopic spectra of the (a) O 1s, (b) Hf 4f, and (c) N 1s of the WN-1 film after annealing at 500 °C for 30 min. in N_2 ambient. The de-convolution of O 1s is also given.

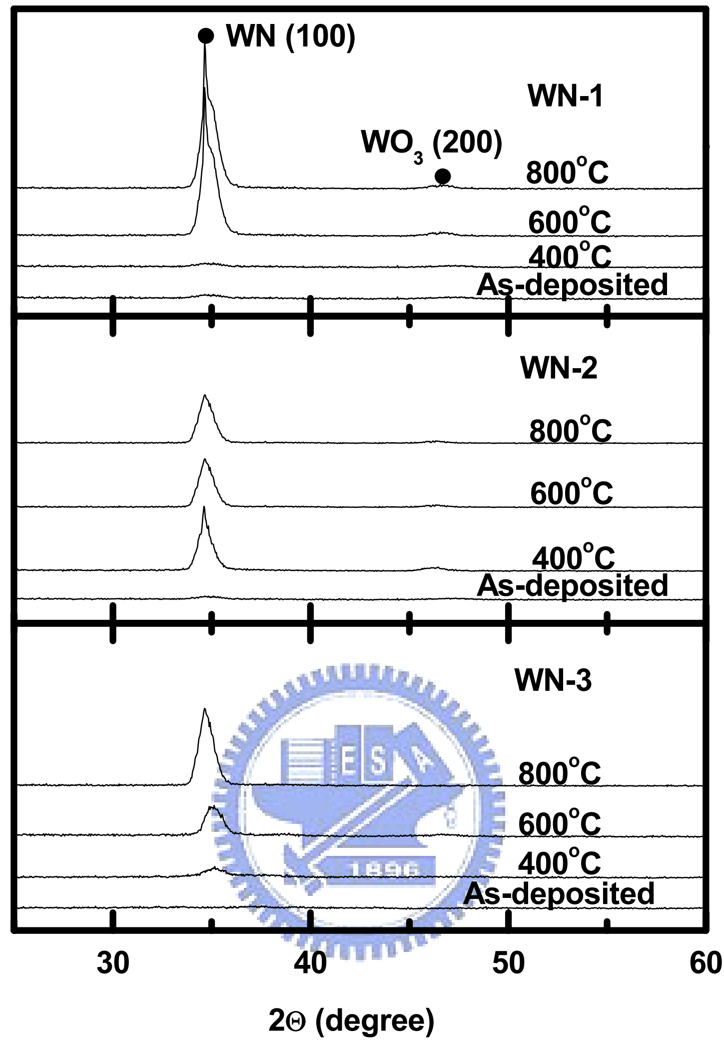
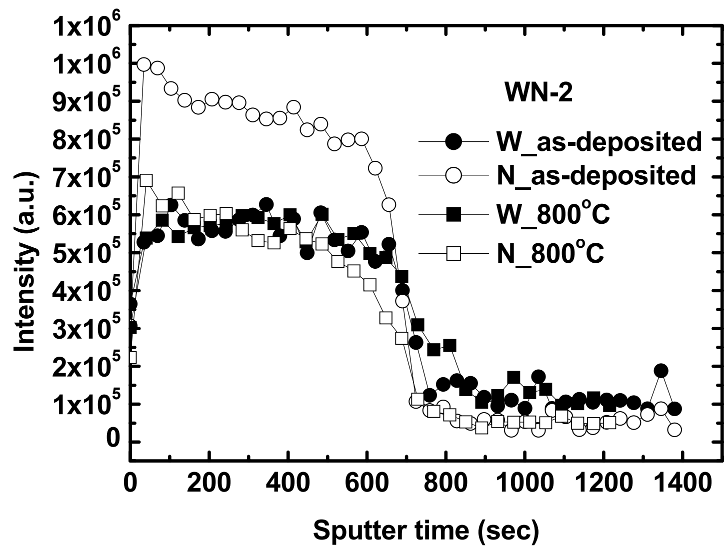


Fig. 4-5 X-ray diffraction patterns of the WN_x films after annealing at different temperatures. All of the as-deposited films are amorphous.



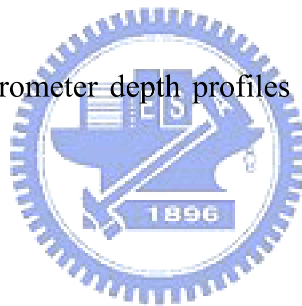
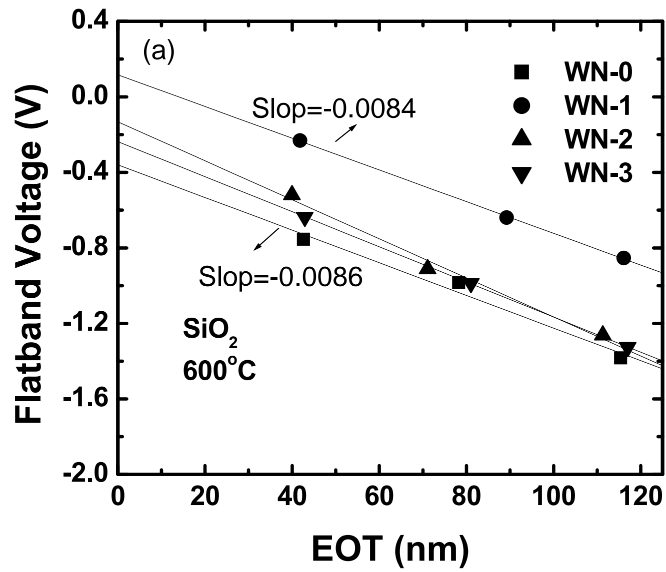


Fig. 4-6 Auger electron spectrometer depth profiles of the as-deposited and 800 °C annealed WN-2 film.



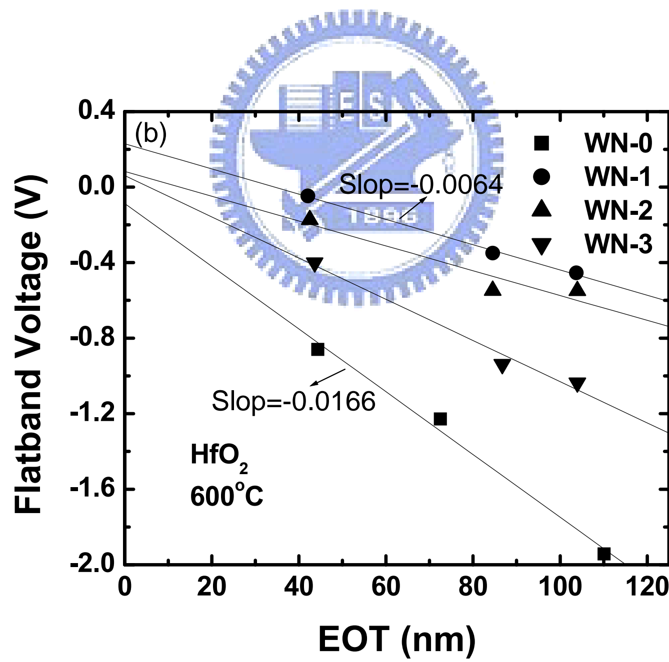
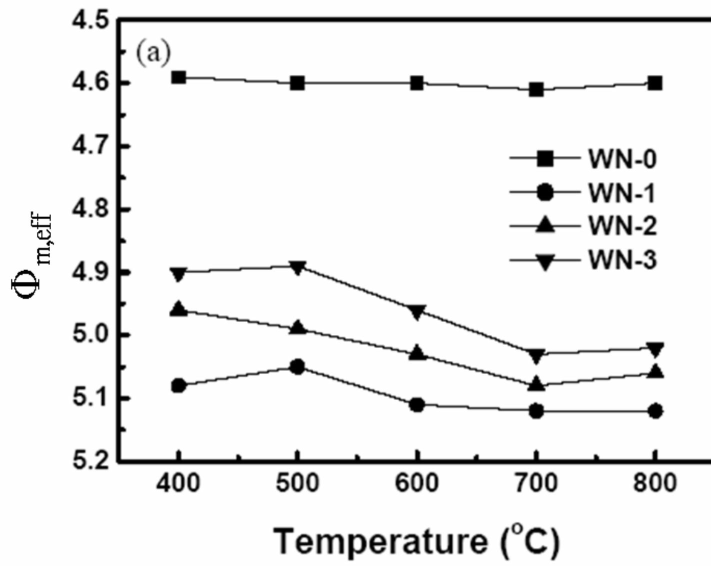


Fig. 4-7 V_{fb} vs EOT plots of (a) $WN_x/SiO_2/Si$ structure and (b) $WN_x/HfO_2/SiO_2/Si$ structure after annealing at 600 ^oC.



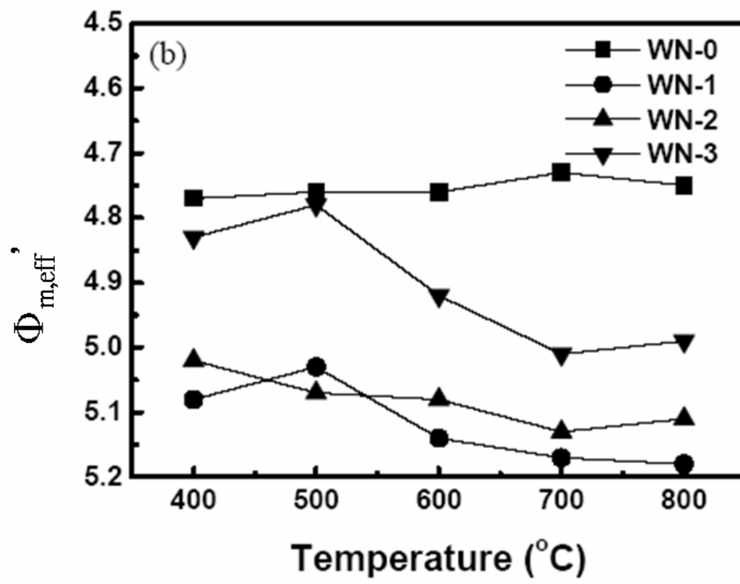
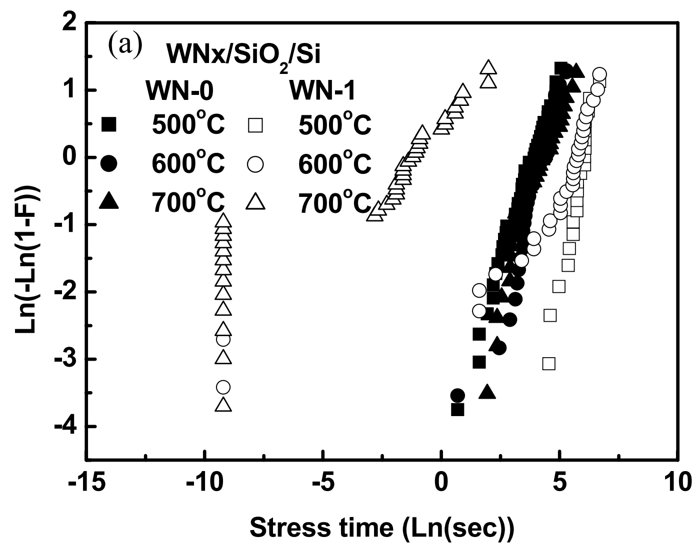


Fig. 4-8 (a) Effective work function of WN_x film on SiO_2 layer and (b) quasi-effective work function of WN_x film on HfO_2 layer versus annealing temperature.



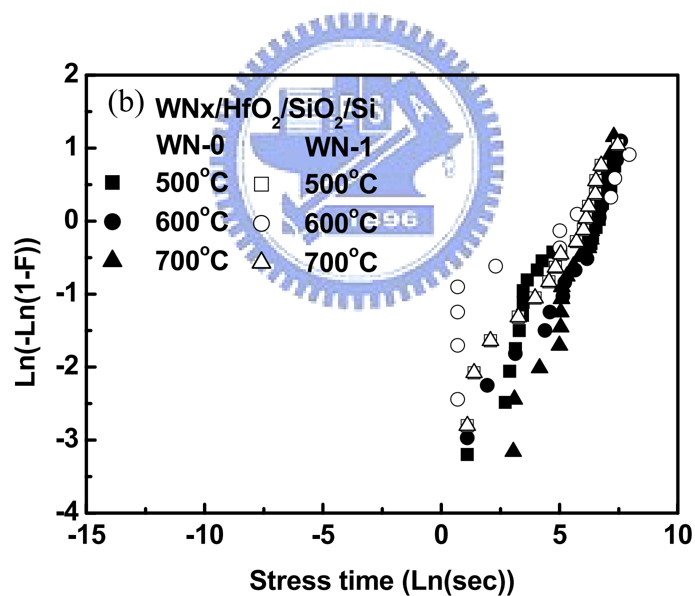
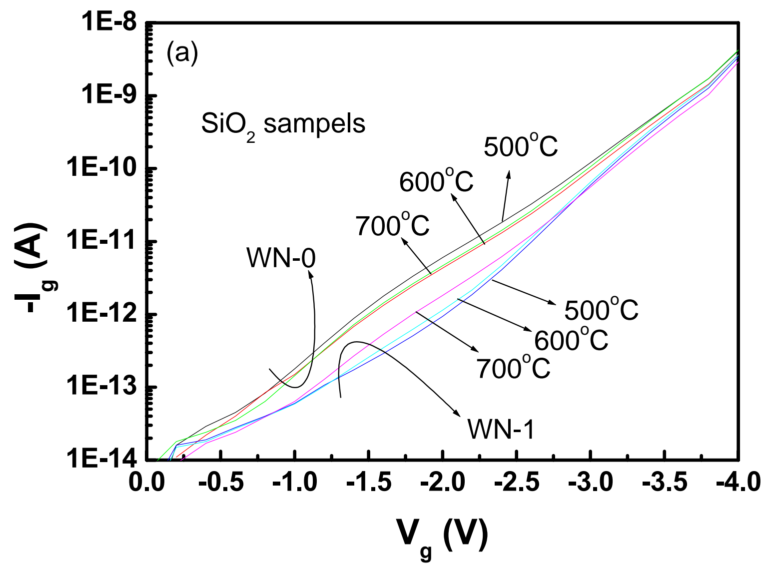


Fig. 4-9 Cumulative failure time of the (a) WN_x/SiO_2 (5 nm)/Si capacitors under constant voltage stress at Vg=-6 V and (b) WN_x/HfO_2 (2 nm)/SiO_2 (2 nm)/Si capacitors under constant voltage stress at Vg=-4.5 V as a function of anneal temperature.



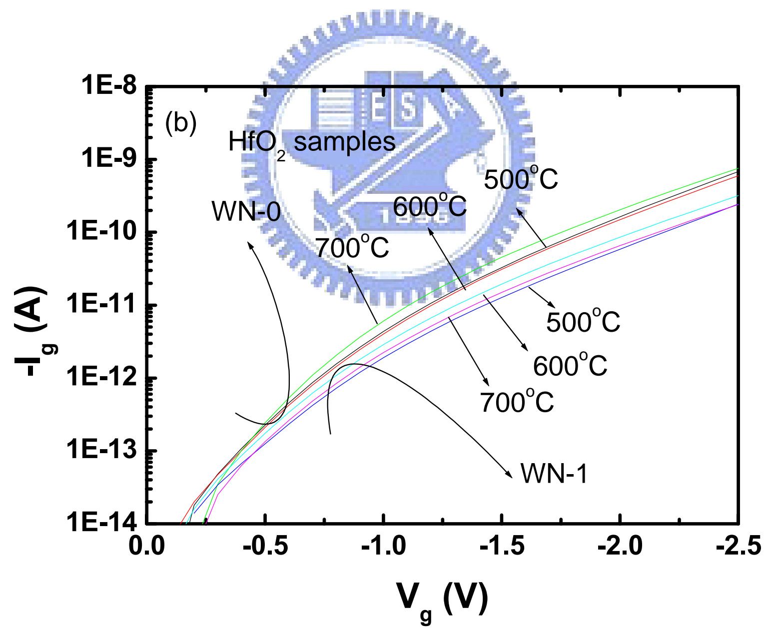


Fig. 4-10 The current-voltage characteristics of (a) WN_x/SiO_2 (5 nm)/Si capacitor and (b) WN_x/HfO_2 (2 nm)/SiO_2 (2 nm)/Si capacitors.

Chapter 5

Cu/Tantalum Nitride Stack Gates

5.1 Introduction

We have investigated the work function and thermal stability of single layer refractory metals, W and Mo, and their nitrides in the chapters 3 and 4. However, the nitrogen content increases the resistivity of the metals. According to the international technology roadmap for semiconductors of 2005 edition, the high resistivity of the material must be lowered with a stake structure [1]. This implies that several metal nitrides are not suitable to be used as a single layer gate electrode. Consequently, the gate electrode should be formed with stack structure that consists of a low resistivity metal and refractory metal nitride layers [2-4]. Wherever it is possible, the low resistivity layer serves as the major conduction material and metal nitride layer serves as the threshold voltage control material.

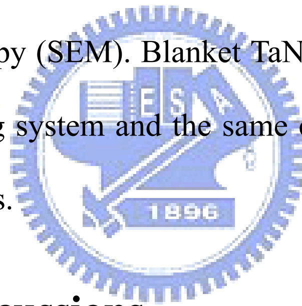
In this chapter, we will study the Cu/TaNx stack as a gate electrode. The widely accepted barrier metal, TaN_x, was reported as the gate electrode of fully depleted silicon-on-insulator devices and showed that its work function is closer to the mid-gap of silicon than that of β-tantalum [4]. The Cu interconnect has become an alternative to conventional Al interconnect, because of its low resistivity, low via resistance, and high electron-migration resistance [5]. Investigating the feasibility of using Cu/TaN_x at the front-end-of-line (FEOL), as the gate electrode, becomes natural [3]. The use of Cu/TaN_x at both the FEOL and back-end-of-line (BEOL) can

simplify the management of production line and reduce the cost-of-ownership because the equipment for BEOL can share with FEOL. Therefore, Cu/TaN_x gate is one of the desirable gate structures. We will investigate the work function modulation of TaN_x films and the thermal stability of Cu/TaN_x MOS devices The criteria and the optimum thermal process conditions for Cu/TaN_x gate MOS devices were proposed.

5.2 Process Flow

Simple metal oxide semiconductor (MOS) capacitors were fabricated with a single damascene process on p-type (100)-oriented Si wafers. After cleaning, a 500 nm thick thermal oxide was thermally grown and a 1 μm thick oxide was deposited in a high-density plasma chemical vapor deposition (HDPCVD) system. A conventional photo-lithography process was then used to define an active region with an area of 300 x 300 μm^2. The oxide in the active region was etched by a two-step (dry by wet) process to prevent Si surface from being damaged. Following pre-oxidation cleaning, a dry oxide, 10 nm thick, was thermally grown at 900 °C as the gate dielectric. The purpose of using such a thick gate oxide is to prevent the gate stacks from the ultrathin oxide issues. TaN_x films with various Ta/N ratios were reactively sputtered in a physical vapor deposition (PVD) system to a thickness of 50 nm, with various gas-flow ratios of Ar to N_2. Before deposition, the process chamber was pumped down to 10^{-7} Torr. During deposition, neither bias nor heating was applied to the wafers, and the chamber pressure was kept at 10^{-4} Torr. It is known that the PVD process may introduce physical bombardment damage and radiation damage to the substrate [6, 7]. The physical bombardment damage can be reduced with lower deposition power and zero substrate bias. Therefore, the DC power and

radio frequency power were set to 15 KW and 350 W, respectively. Radiation-induced oxide charges can be neutralized after a post-deposition annealing at 400 °C. The Ar/N_2 gas flow ratio was set to 60/12, 60/20, and 60/28 for TaN-1, TaN-2, and TaN-3 films, respectively. A 200 nm thick Cu film was continuously deposited on TaN_x films as a seed layer in another chamber of the same PVD system, without exposing to air. Then a 1 μm thick electrochemically deposited copper film was deposited. Finally, the chemical mechanical polishing separated MOS capacitors and a 30 nm thick silicon nitride film were deposited to passivate the Cu surface. The completed structure is Si_3N_4 (30 nm)/Cu (500 nm)/TaN_x (50 nm)/SiO_2 (10 nm)/Si substrate. The thickness of dielectric layers were measured with the ellipsometry method and the thickness of metal layers were inspected with cross-sectional scanning electron microscopy (SEM). Blanket TaN_x films deposited on a thin SiO_2 layer in the same sputtering system and the same conditions were used to analyze their physical characteristics.



5.3 Results and Discussions

5.3.1 Physical Properties of TaN_x Films

The nitrogen content of TaN_x films determined by the analysis of Rutherford backscattering spectra are 23, 33, and 39 at. % for TaN-1, TaN-2, and TaN-3 films, respectively. Figures 5-1 (a)-(c) show the wide angle XRD spectra of TaN-1, TaN-2, and TaN-3 films after annealing at various temperatures. The phases of all TaN_x are quite thermally stable. For the TaN-1 film, an extremely weak Ta_2N (101) signal was detected after annealing at all temperatures. No significant phase change was observed with the increase of annealing temperature except that a

weak Ta_3N_5 signal appeared after 800 °C annealing. No XRD signals of tantalum nitrides were detected for the TaN-2 films after annealing at different temperatures. This implies that the morphology of TaN-2 films is almost amorphous. For TaN-3 film, the Ta_3N_5 (320) phase is observed after 400 °C annealing and the signal intensity increases with the increase in the annealing temperature. Because all of the XRD signal intensity is relatively weak, grain size and phases are further identified with plane view inspection and diffraction pattern of TEM, respectively. Figure 5-2 shows the plane view TEM micrographs of TaN-1 and TaN-2 films after annealing at 400 and 700 °C. Since there is no apparent grain boundary but only some microcrystals, the films are almost amorphous. Despite 700 °C annealing treatment, the grain is small and the grain sizes are in the nanometer range. It is sure that TaN_x film is dense and has no serious grain growth, but some microcrystals are precipitated. The transmitted electron diffraction patterns identify the phases of microcrystals in TaN-1 and TaN-2 films more precisely than those identified with XRD. As shown in Fig. 5-3, Ta_2N and TaN phases were observed for TaN-1 and TaN-2 films, respectively, after annealing at 700 °C. The sheet resistance of TaN_x films was also measured to understand the effect of grain growth and phase transformation. The resistivities of the as deposited TaN-1, TaN-2, and TaN-3 films were 463, 480, and 1290 mV cm, respectively. Figure 5-4 shows the variation of sheet resistance of the three kinds of TaN_x films vs. annealing temperature. These values were normalized to the sheet resistance of the corresponding films annealed at 400 °C. A slight decrease of sheet resistance is observed in the TaN-1 films upon increasing the annealing temperature from 400 to 700 °C. Referring to the TEM pictures in Fig. 5-2, the decrease of sheet resistance is attributed to the slight grain growth with the increase of annealing

temperature. The sheet resistance of TaN-1 film increased suddenly at 800 °C which can be attributed to the formation of Ta_3N_5 phase, as identified with the XRD spectra. According to Radhakrishnan et al., the Ta_3N_5 phase has a very high resistivity [8]. The resistance of TaN-2 is almost constant after 800 ^{o}C annealing which reflects that no grain growth and phase transition occur in the TaN-2 film as analyzed by the XRD and TEM. Unlike TaN-1 and TaN-2 films, the sheet resistance of TaN-3 films increases with the annealing temperature. The XRD spectra reveal that the increase is due to the growth of high resistivity phase (Ta_3N_5 phase) after high temperature annealing.

5.3.2 Effective Work Function on SiO_2

Figure 5-5 shows the effective work function ($\Phi_{m,eff}$') of Cu/TaN-1, Cu/TaN-2, and Cu/TaN-3 samples vs. the annealing temperature. The $\Phi_{m,eff}$' was extracted from the C-V curves by using a theoretical C-V curve fitting, which has an error counting for the oxide charges as discussed in chapter 2. All the $\Phi_{m,eff}$' difference of Cu/TaNx samples is less than 70 mV below or at 600 °C. This implies that the modulation of work function of TaN_x film with the usage of various N/Ta ratios is very limited. In other words, the work function of TaN_x film is quite stable against variation induced by process disturbance. According to the ITRS road-map, the allowed deviation of threshold voltage (V_{th}) is 25 mV at the sub-70 nm technology node [9]. For 400 °C annealing condition, the 3-sigma (3σ) deviation of $\Phi_{m,eff}$' of all samples satisfied the demand of 25 mV. Particularly, the 3σ deviation of $\Phi_{m,eff}$' of Cu/TaN-2 sample is less than 5 eV. However, the deviation of $\Phi_{m,eff}$' continuously increases with the increase of annealing temperature and is out of the demand after annealing at and beyond 500°C. The reasons that cause the deviation of $\Phi_{m,eff}$' include the characteristic of gate

material, substrate doping concentration, oxide thickness, and oxide charges. The factors of oxide thickness and substrate deviation of 400 and 500°C annealed samples are ruled out using the $\Phi_{m,eff}$' extracting procedure of the theoretical curve fitting. The characteristics of gate electrode resulting in $\Phi_{m,eff}$' variation are phase change and grain growth. The XRD spectra show no apparent difference among the TaN_x films annealed below 700°C. The primary phase after annealing at 400°C changed from Ta_2N to Ta_3N_5 as the nitrogen content increased from 23 to 39 at. % but the difference of $\Phi_{m,eff}$' is only 70 mV. In addition, the $\Phi_{m,eff}$' differences of Cu/TaN_x samples annealed at the same temperature below 700°C are almost the same. It is thus clear that the phase transformation cannot dominate the $\Phi_{m,eff}$' variation. In other words, the main reason that caused the variation of $\Phi_{m,eff}$' should be the variation of effectives oxide charges. Copper contamination is the first possible factor to be examined. Figures 5-6 (a), (b), and (c) show the C-V curves before and after bias temperature stress (BTS) test at +1 MV/cm and 150°C for 60 min of Cu/TaN-2 sample annealed at 500, 550, and 600 °C, respectively. The C-V curves before and after BTS test of the 500 °C annealed sample are identical. A slight shift of C-V curve toward the negative voltage axis after the BTS test is observed on the 550 °C annealed sample. Furthermore, the appearance of a hump at the inversion region (about +0.5 V) reveals that the carrier lifetime at the Si substrate is shortened. Both the observations indicate that the device was contaminated, while the most possible contamination source is Cu. Transient capacitance analysis was performed to detect lifetime degradation. The transient capacitance technique is sensitive to small quantity of deep-level atoms, such as copper [10, 11]. Once the copper reaches silicon, it affects the silicon/oxide interface and the minority carrier lifetime in the bulk. Therefore, The MOS

capacitor was biased at the accumulation mode of -4 V and was switched to deep depletion mode of +5 V. The retention time is defined as the time required for capacitance to recover 90% of the stable value. Table 5-1 lists the retention time and its deviation of Cu/TaN-2 samples annealed at 400, 500, 550, and 600 °C. The retention time of 550 and 600 °C annealed samples decreases to shorter than 0.1 s, which indicates that copper had penetrated into the Si substrate. A slight decrease in the retention time of the 500 °C annealed sample is discussed later. Although SIMS was not detected Cu signal in the substrate of the 550 °C annealed sample, Fig. 5-7 shows the SIMS depth profile of Cu in the 600 °C annealed Cu/TaN-2 sample before BTS test. Before SIMS analysis, the Cu film was removed with dilute HNO_3 solution. It is thus concluded that the decrease of $\Phi_{m,eff}$' and increase in the deviation of $\Phi_{m,eff}$' beyond 550 °C annealing is related to Cu contamination. However, the decrease of $\Phi_{m,eff}$' and increase of $\Phi_{m,eff}$' deviation occur after annealing at 500 °C even if no Cu contamination would occur.

5.3.3 Electrical Effect of Thermal Annealing

Figure 5-8 shows the interface state density (D_{it}) at the mid-gap of Cu/TaN-1, Cu/TaN-2, and Cu/TaN-3 samples versus the annealing temperature. The magnitude and deviation of D_{it} increase with the increase of annealing temperature. It is reasonable to expect that the effective oxide charges (Q_{eff}) also increase with the increase of annealing temperature. A deviation in Q_{eff} of $1x10^{11}$ cm^{-2} results in a deviation in flat band voltage of 46mV. Therefore, $\Phi_{m,eff}$' deviation can be reasonably attributed to the deviation of Q_{eff} induced by the deviation of thermal stress. For 500 ℃ annealed samples, the drop in the $\Phi_{m,eff}$' compared to the corresponding 400 ℃ annealed samples is less than 60 mV, and the 3σ deviation is about 50 mV which is higher than the value of the demand. It

is postulated that the deviation of $\Phi_{m,eff}$' comes from the large difference of thermal expansion coefficient between Cu and Si. By scaling down the gate dielectric thickness, the effect of Q_{eff} on threshold voltage can also be scaled down. However, the Q_{eff} generated by thermal stress may increase. The slight decrease in the retention time of 500 ℃ annealed sample, listed in Table 5-1, is thus explained by the increase of surface generation rate due to the increase of D_{it}. Replacing Cu by another low resistivity material with thermal expansion coefficient close to Si, for example W and Mo, is recommended.

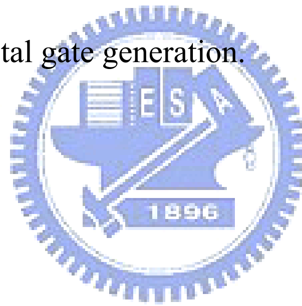
Above 600 ℃ annealing, the $\Phi_{m,eff}$' rises with the annealing temperature. It is presumable that the interface dipoles caused by the interaction between oxide and metal gate induce effective work function offsets since effective oxide charges are increasing [12]. The interaction between TaN metal gate and gate dielectric is still under investigation.

5.4 Summaries and Conclusions

This chapter investigates the work function modulation of TaN_x film and the thermal stability of Cu/TaN_x stack as the gate electrode. The nitrogen content was varied in the range 23-39 at. % by using the reactive sputtering with various Ar/N_2 gas flow ratios. The main phases of the TaN-1, TaN-2, and TaN-3 films are Ta_2N, TaN, and Ta_3N_5, respectively. The TaN_x films are thermally stable up to 800 ℃. However, the formation of the Ta_3N_5 phase in a TaN_x film annealed at high temperature or with a high N/Ta ratio increases the effective resistivity. The effective work function of TaN_x is about 4.31-4.38 eV and the range is less than 70 mV. Such a weak work function modulation implies that the work function of TaN_x film is quite stable to avoid the variation induced by process disturbance. On the

other hand, TaN_x is a suitable to be a gate electrode only for the surface channel NMOSFETs. The flat band voltage decreases with an increase in the annealing temperature. In addition, the deviation of the flat band voltage increases with the annealing temperature. Although phase change, grain growth and Cu contamination contribute at high temperature, thermal stress-induced oxide charges dominate the decrease and deviation of the flat band voltage at temperature below 500 ℃.

In conclusion, according to the material and electrical analysis, the Cu/TaN_x stack can be used as a gate electrode for the surface channel NMOSFETs, and the maximum process temperature following gate electrode deposition should be lower than 500 ℃. The thermal stress-induced oxide charges are additional sources of deviation in the threshold voltage. This result must be considered in controlling the threshold voltage during metal gate generation.



References

[1] International Technology Roadmap Semiconductors 2005 edition, Front end Process, p.33.

[2] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, and T. Kunio," A Novel W/TiNx Metal Gate CMOS Technology Using Nitrogen-Concentration-Controlled TiNx Film," in *1999 Int. Electron Devices Meet. Tech. Dig.*, pp.253-256.

[3] T. Matsuki, K. Kishimoto, K. Fujii, N. Itoh, K. Yoshida, K. Ohto, S. Yamasaki, T. Shinmura, and N. Kasai, "Cu/poly-Si Damascene Gate Structured MOSFET with Ta and TaN Stacked Barrier," in *1999 Int. Electron Devices Meet. Tech. Dig.*, pp.261-264.

[4] H. Shimada, I. Ohshima, T. Ushiki, S. Sugawa, and T. Ohmi," Tantalum Nitride Metal Gate FD-SOI CMOS FETs Using Low Resistivity Self-grown bcc-tantalum Layer," *IEEE Trans. Electron Devices*, vol. ED-48, No.8, pp.1619-1626, 2001.

[5] S.Venkatesan, A.V.Gelatos, V.Misra, B. Smith, R. Islam, J.Cope, B, Wilson, D.Tuttle, R. Cardwell, S. Anderson, M. Angyal, R. Bajaj , C. Capasso, P.Crabtree, S.Das, J.Farkas, S.Filipiak, B. Fiordalice, M. Freeman, P.V. Gilbert, M. Herrick, A. Jain, H. Kawasaki, C. King, J.Klein, T.Lii, K. Reid, T. Saaranen, C. Simpson, T. Sparks, P.Tsui, R.Venkatraman, D. Watts, E. J.Weitzman, R. Woodruff, I. Yang, N. Bhat, G, Hamilton and Y.Yu, "A High Performance 1.8V, 0.20μm CMOS Technology with Copper Metallization," in *1997 Int. Electron Devices Meet.,* pp.769-773.

[6] K. Nakajima, Y. Akasaka, M. Kaneko, M. Tamaoki, Y. Yamada, T. Shimizu,

and K. Suguro," Work Function Controlled Metal Gate Electrode on Ultrathin Gate Insulators," in Procee*dings of the 1999 Symposium on VLSI Technology*, pp.95-96.

[7] B. Y. Tsui, S. H. Liu, G. L. Lin, J. H. Ho, and C. H. Chang, "A Comprehensive Study on Radiation Damage in Plasma System," in *Proceedings of Symposium on Plasma Process-Induced Damage*, pp.148-150, 1996.

[8] K. Radhakrishnan, N. G. Ing, and R. Gopalakrishnan," Reactive Sputter Deposition and Characterization of Tantalum Nitride Thin Films," *Material Science and Engineering*, B vol. 57, No.3, pp.224-227, 1998.

[9] 2000 International Technology Roadmap for Semiconductors - Process Integration, Devices, & Structures, p.42, Semiconductor Industry Association, San Jose.

[10] C.N. Berlung, "Surface States at Steam-grown Silicon-silicon Dioxide Interfaces," *IEEE Trans. Electron Devices*, vol. ED-13, No. 10, pp.701-705.

[11] Yosi Shacham-Diamand, Baral Israel, Yelena Sverdlov," The electrical and material properties of MOS capacitors with electrolessly deposited integrated copper gate," *Microelectronic Engineering*, vol. 55, No.1-4, pp.313-322, 2001.

[12] J. Robertson, "Band offsets of wide-band-gap oxides and implications for future electronic devices, "*J. Vac. Sci. Technol.*, B vol. 18, No. 3, pp.1785-1791, 2000.

Table 5-1 Mean values and standard deviations of retention time of Cu/TaN-2 MOS capacitors under transient capacitance measurement.

sample	400 °C	500 °C	550 °C	600 °C
Retention time (s)	33.4	14.1	<0.1	<0.1
Deviation (s)	16.2	7.4	0	0

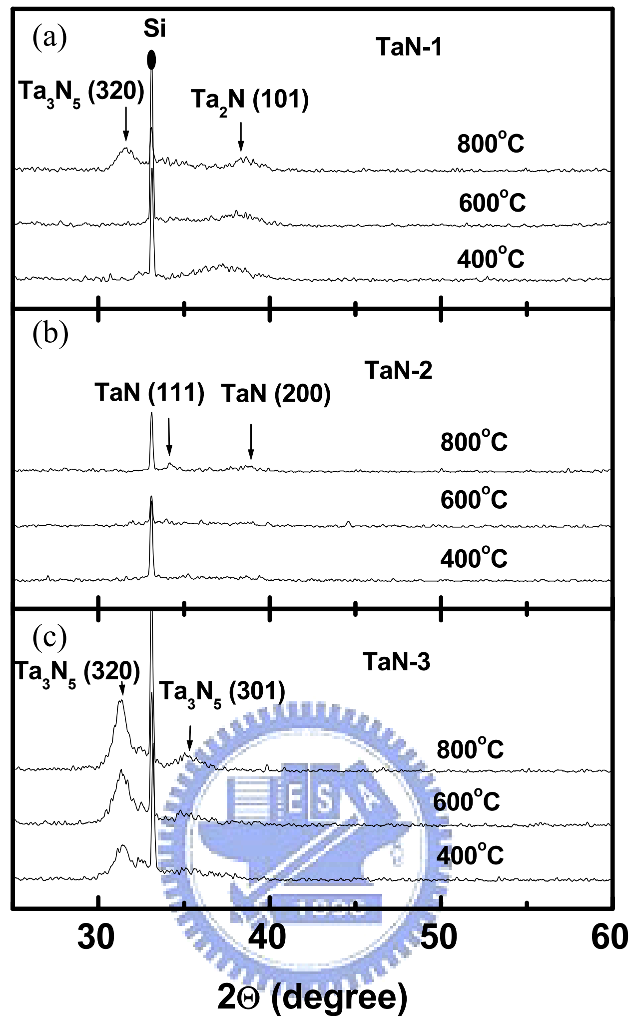


Fig. 5-1 Wild angle X-Ray Diffraction (XRD) spectrums of TaN-1, TaN-2, and TaN-3 films after annealing at various temperatures.

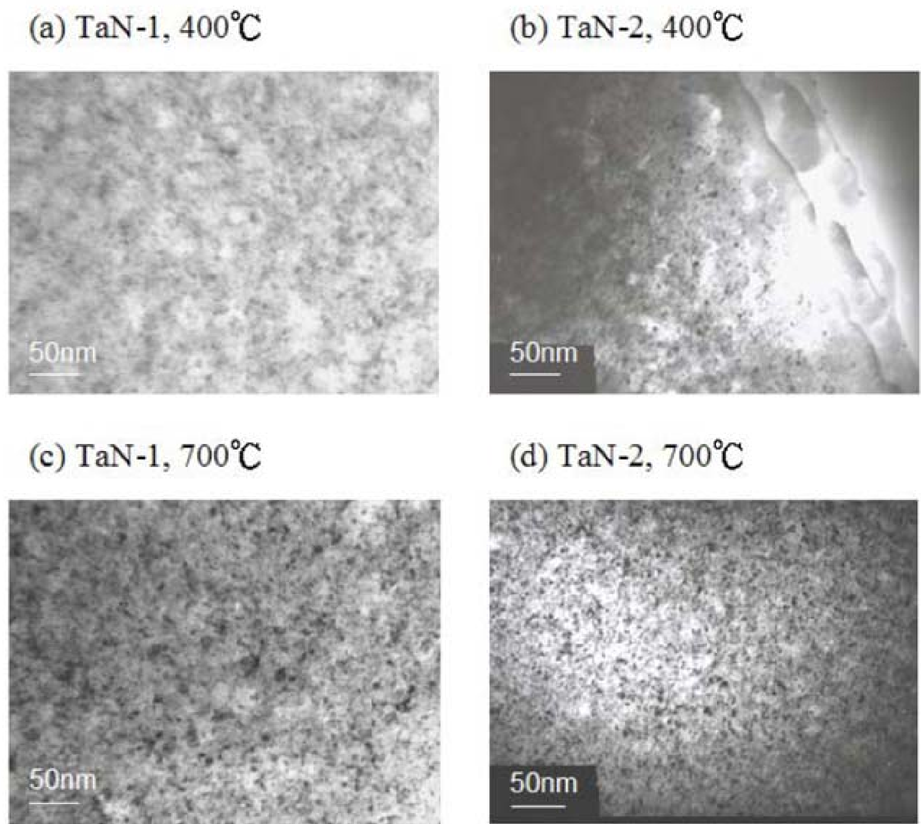


Fig. 5-2 Plane view Transmitted Electron Microscopy (TEM) micrographs of (a) TaN-1 film annealed at 400 ℃, (b) TaN-1 film annealed at 700 ℃, (c) TaN-2 film annealed at 400 ℃, and (d) TaN-2 film annealed at 700 ℃.

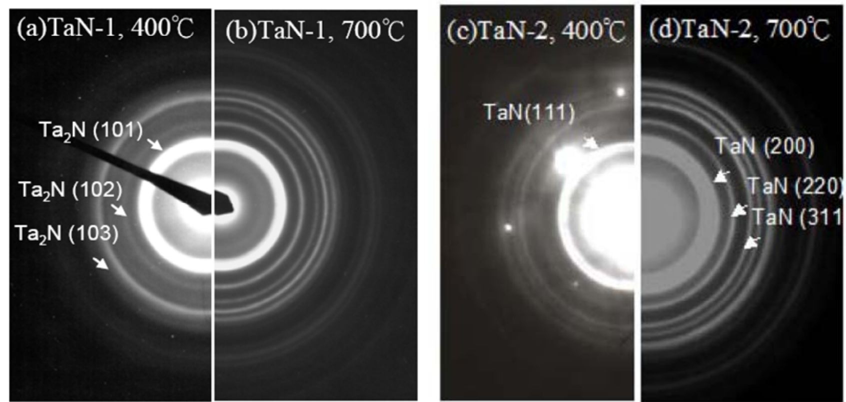


Fig. 5-3 Diffraction patters of (a) TaN-1 film annealed at 400 ℃, (b) TaN-1 film annealed at 700 ℃, (c) TaN-2 film annealed at 400 ℃, and (d) TaN-2 film annealed at 700 ℃.

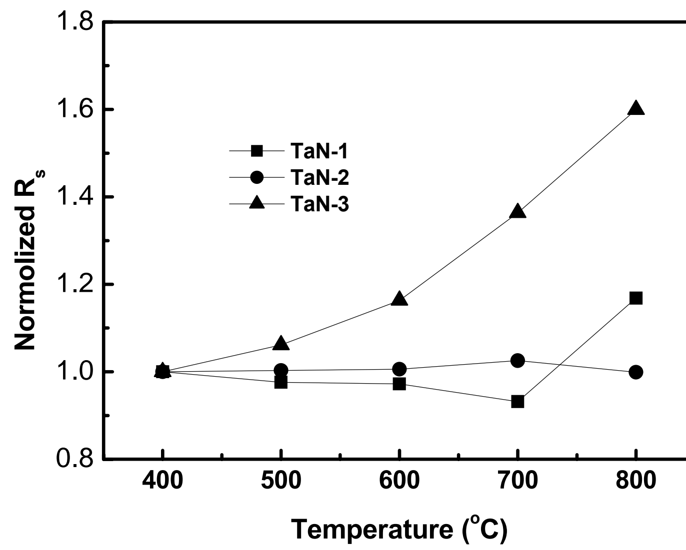


Fig. 5-4 Variation of sheet resistance of the three-kinds of TaN_x films versus annealing temperature.

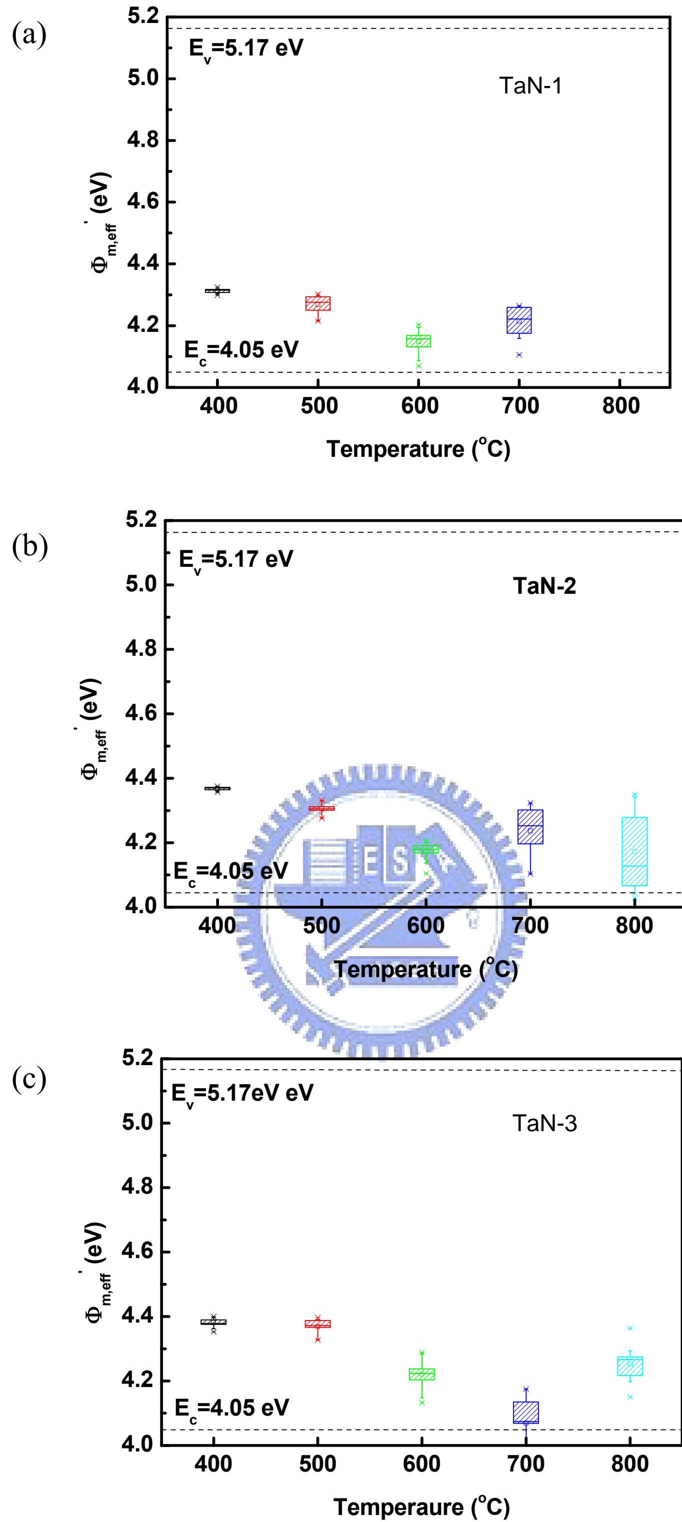


Fig. 5-5 Effective work function ($\Phi_{m,eff'}$) of (a) Cu/TaN-1 sample, (b) Cu/TaN-2 sample, and (c) Cu/TaN-3 sample versus the annealing temperature. The effective work function was extracted from the C-V curves by a theoretical curve fitting.

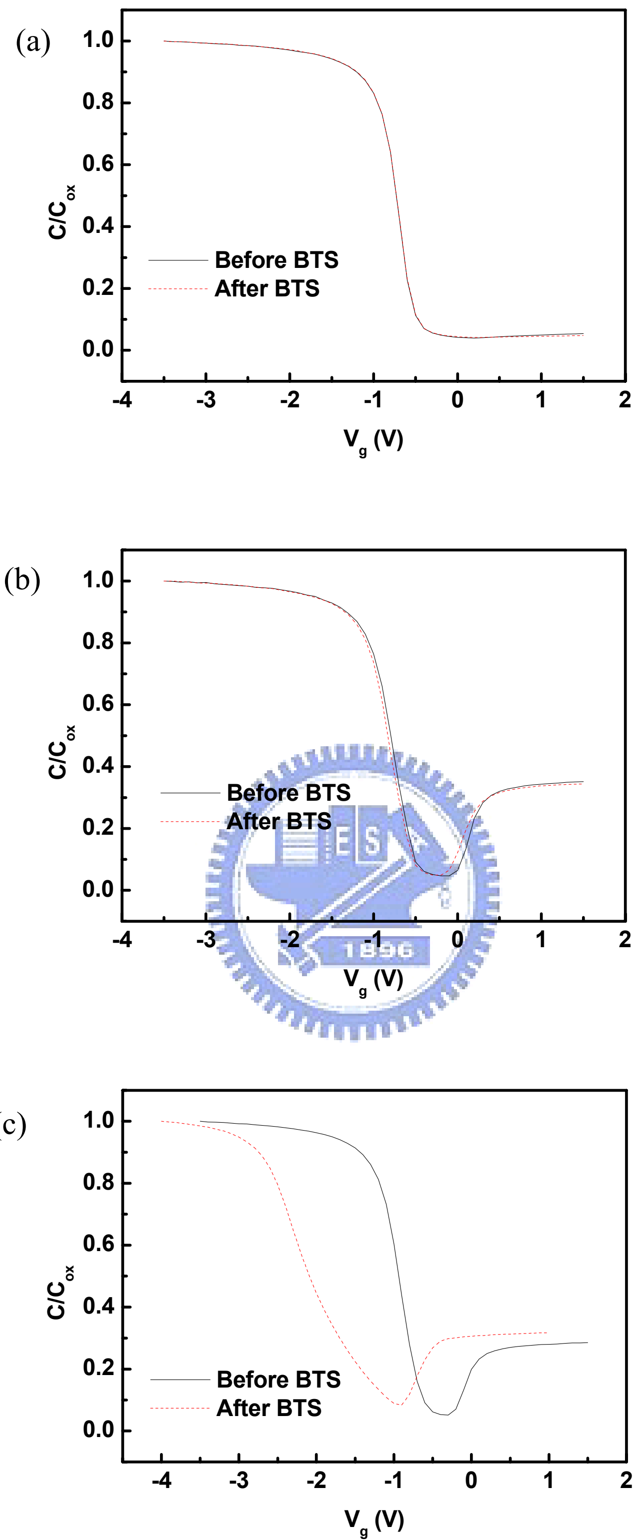


Fig. 5-6 Capacitance-Voltage curves before and after bias-temperature stress (BTS) test at +1MV/cm and 150 ℃ for 60 min. of Cu/TaN-2 sample annealing at (a) 500 ℃, (b) 550 ℃, and (c) 600 ℃.

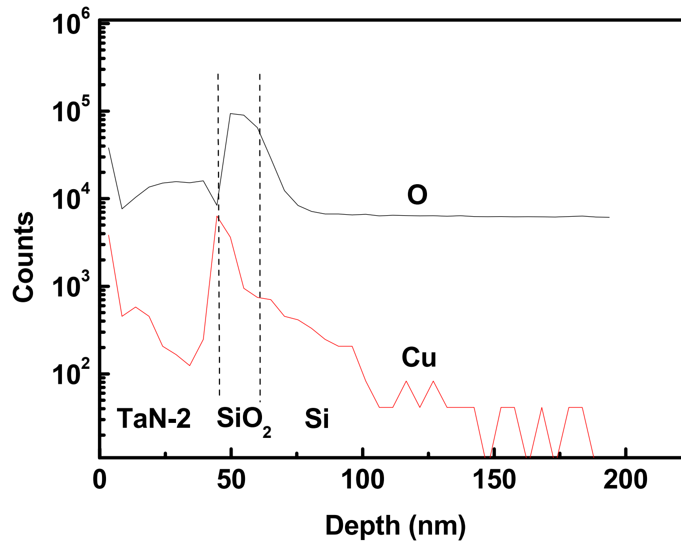


Fig. 5-7 Secondary ion mass spectroscopy (SIMS) depth profile of Cu of the 600 ℃ annealed Cu/TaN-2 sample. No BTS was performed before SIMS analysis and the Cu layer was removed with dilute HNO_3 solution.

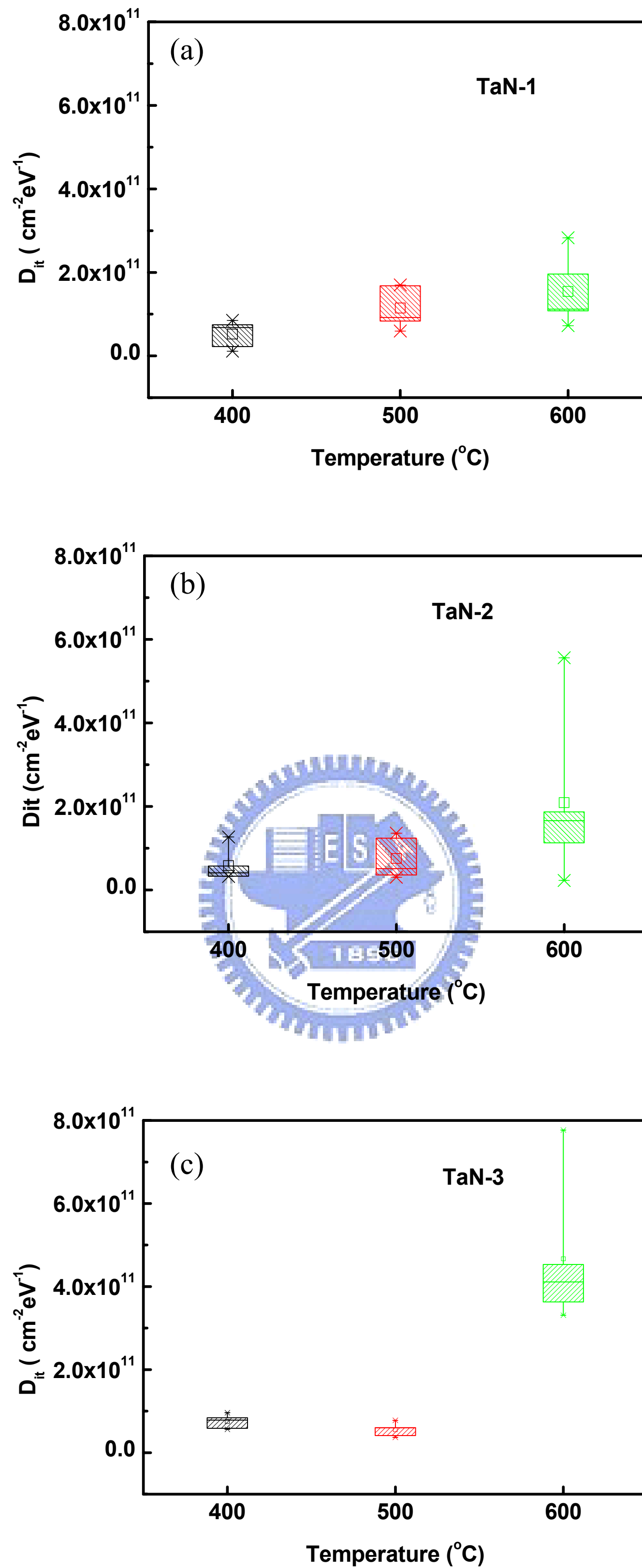


Fig. 5-8 Interface state density (D_{it}) of (a) Cu/TaN-1 sample, (b) Cu/TaN-2 sample, and (c) Cu/TaN-3 sample versus the annealing temperature.

Chapter 6

Ta-Pt Metallic Alloy Gates

6.1 Introduction

Several metal nitrides have been discussed in the previous chapters. Although metal nitrides are thermal stable under high-temperature annealing, the work function range is not large enough for both NMOSFETs and PMOSFETs. A principle of work function tuning by the metallic alloy system is adapted to achieve full range of work function over the silicon energy band gap. It is difficult to theoretically evaluate the work function of alloy since the components in alloy is incoherent and the structure of alloy is always disordered. Therefore, there are few works in discussion of work function of alloy. In 1974, Gelatt and Ehrenreich have first theoretically solved the Hamiltonian equation by coherent-potential-approximation (CPA) to obtain the work function of the alloy [1]. The work function of NiCu and PtRh alloys is recently measured to be linearly dependent upon the concentrations [2]. Moreover, the proposed application of binary metallic alloys in MOSFETs included Ru-Ta, Ta-Pt and Mo-Hf alloys [3-5]. Work function modulation used by the binary alloys seems a easily controlled method. In this chapter, Ta-base alloys (Ta-Ti and Ta-Pt) were applied as gate electrodes of MOSFETs due to the dense morphology and chemical inert of Ta which is completely immune to chemical attack at temperature lower 150 ^{o}C and is attacked only by hydrofluoric acid [6]. The Ta element has a high melting point exceeded only by tungsten and rhenium. The

thermal stability of Ta-rich Ta-Pt alloy on SiO_2 and HfO_2 is investigated. The basic physical characteristic of Ta-Pt alloy is also discussed.

6.2 Process Flow

Simple MOS capacitors with gate electrodes of Ta-Pt and Ta-Ti alloys were fabricated on (100)-oriented Si wafers. After standard wafer cleaning, a gate oxide was thermally grown to 10 nm thick in a dry oxygen atmosphere at 900°C. After gate oxide deposition, the alloy gate was patterned by lift-off process. The 55nm thick Ta-Pt alloys were deposited in a co-sputtering system. The sputtering power of Ta and Pt (Ti) targets were set to several split conditions in order to control the concentration in alloys. Before lift-off, some samples with 6 nm gate oxide were implanted with arsenic ions (As^+), phosphorous ions (P^+), or boron di-fluoride ions (BF_2^+) to a dose of 5×10^{15} cm^{-2}. Then, those samples as well as some undoped control samples were capped by a 60nm thick silicon nitride film which was deposited at 300 °C by a plasma-enhanced chemical vapor deposition (PECVD) system. The nitride film is used to avoid dopant escape during high temperature annealing. Samples doped by As, P and BF_2 were annealed at temperatures from 400 °C to 800 °C in a rapid thermal annealing (RTA) system for 30sec in N_2 atmosphere. The samples uncapped with nitrides were annealed in furnace for 30 min in N_2. Contact-hole patterning and Al-metallization processes were used to complete the device fabrication. Finally, backside contact was formed with an evaporated aluminum film.

Ta-Pt alloy for physical analysis was deposited on a thin SiO_2 layer in the same co-sputtering system with sputtering power of DC 100 watts and DC 30 watts for

Ta and Pt targets, respectively, to form a Ta-rich alloy film. It is well known that alloys with high ratio of low work function components are much chemically active. The annealing conditions were from 400 °C to 800 °C for 30 min in a horizontal furnace. Table 6-1 lists all of the samples and process conditions used in this work.

6.3 Results and Discussions

6.3.1 Work Function Modulation

According to result of Gelatt and Ehrenreich, the work function of A_xB_{1-x} alloys can be approximately expressed as

$$\Phi_m = x\Phi_{m,A} + (1-x)\Phi_{m,B} + x(1-x)[\frac{(\Phi_{m,A}-\Phi_{m,B})(\rho_A-\rho_B)}{x\rho_A+(1-x)\rho_B}]$$

$$= x\Phi_{m,A} + (1-x)\Phi_{m,B} + x(1+x)[\frac{(\Phi_{m,A}-\Phi_{m,B})(\frac{\rho_A}{\rho_B}-1)}{x\frac{\rho_A}{\rho_B}+(1-x)}]$$

, where $\Phi_{m,A}$ and $\Phi_{m,B}$ are the pure constituent work functions of A and B, respectively, and ρ_A and ρ_B are the pure constituent total densities of states [1]. The density of states at Fermi-level energy $\rho(E_F)$ is proportional to the electronic specific heat constant $C_e = (1/3)\pi^2\rho(E_F)K_B{}^2T$, where K_B is Boltzmen constant and T is the temperature. In the case of $C_{e,A}/C_{e,B} \sim 1$, the Φ_m changes with x linearly as $\Phi_m(x) = \Phi_{m,B} + x(\Phi_{m,A}-\Phi_{m,B})$ [2]. This means that we can get any work function between the $\Phi_{m,A}$ and $\Phi_{m,B}$ we want if we properly control the content and composition of the alloys. The C_e values of Pt, Ta, and Ti are 6.8, 5.9, and 3.35 mJ/mole/K^2, respectively [7]. Since the C_e values of Ta and Pt are very close, a linearly correlation can be expected.

The effective work functions of the 400 °C annealed samples are extracted by

comparing the measured C-V curves with the theoretical C-V curves, as shown in Fig. 6-1 [8]. The effective work functions count for the effect of oxide charges. As the gate oxide thickness is 11nm, an amount of oxide charge (Q_{ss}) of 1×10^{11} cm^{-2} results in a Φ_m deviation of 0.05 eV only. The Q_{ss} of the poly-Si gate is around 5×10^{10} cm^{-2} estimated from the interface states. The Q_{ss} of metal gate devices might be higher than that of poly-Si ones. However, since the interface density of metal gate devices is similar to that of poly-Si one, it is believed that the Q_{ss} of metal gate devices will not be much higher than those of poly-Si ones. It is reasonable that the poly-Si gate has an effective work function near the Silicon conduction band edge. Therefore, the effect of oxide charges can be ignored. With the increasing contents of high Φ_m element, Pt in Ta-Pt alloys or Ta in Ta-Ti alloys, the $\Phi_{m,eff}$ shifts toward higher value. The $Ta_{0.63}Ti_{0.37}$ (A1 alloy) and $Ta_{0.58}Pt_{0.42}$ (A4 alloy) show $\Phi_{m,eff}$ of about 4.16 eV and 5.05 eV and is suitable for NMOSFETs and PMOSFETs, respectively. As expected, the $\Phi_{m,eff}$ can be modulated continuously by adjusting the atomic composition of the alloy. This property allows us to get precise work function for CMOSFETs.

6.3.2 Physical Characteristic of Ta-rich Ta-Pt Alloys

The Ta-Pt phase diagram was reported by Waterstrate in 1981, who identified the main phases to be Ta, ε, σ, PtTa, Pt_2Ta, Pt_3Ta, and Pt [9]. The phase changes from Ta-rich phase to Pt-rich phase with the increase of Pt content [10]. The phases of Ta, ε, σ, PtTa, Pt_2Ta, Pt_3Ta, and Pt are present in the alloys with Ta atomic concentration of 90.5-100 %, 84.5-86.5 %, 65-85 %, 48.5-50 %, 32-33.3 %, 24-25.5 % and 0-19 %, respectively. The compositions of the alloys in this work are listed in Table 6-1, which are analyzed by the Rutherford backscattering spectroscopy.

Figures 6-2(a) and (b) show the Ta and Pt XPS spectra of the A2 sample with a Ta-Pt (55 nm)/SiO_2 (6 nm)/Si structure after 800 °C annealing for 30 min. The XPS spectra of pure Ta and Pt films were also shown for comparison. The signals of pure Ta and Pt films have been rescaled by factors of 0.5 and 0.22, respectively. At the sample surface, the binding energy $4f_{7/2}$ electron of Ta shifts 3.9 eV from the binding energy of the pure Ta $4f_{7/2}$ electron (22.1eV) to 26.0 eV. The large shift in the Ta $4f_{7/2}$ electron binding energy is due to the presence of surface oxide (Ta-O bond) [11]. The binding energy of Pt at the surface is almost identical to that of pure Pt. This is due to the fact that Pt is chemically more inert than Ta. The presence of oxidized Ta on the surface is due to the sample exposure to air. After different sputtering periods, the peaks of Ta $4f_{7/2}$ are all at 22.6 eV (0.5 eV shift from pure Ta). The peaks of Pt $4f_{7/2}$ in the TaPt alloys are located at 71.9 eV indicating a 0.7 eV shift from the pure Pt. No obvious O 1s signal is observed. The lack of oxidized signal in the bulk of Ta-Pt alloy indicates that the Ta-rich alloys are almost oxygen free and stable on the SiO_2 film. The obvious shifts of binding energies of Ta and Pt of the Ta-Pt alloy with respect to pure Ta and Pt films imply that Ta and Pt interact with each other to influence the final energy states of the Ta-Pt alloys [12]. It is postulated that this interaction changes the work function of Ta-Pt alloys with different Ta/Pt atomic ratios [4]. The binding energies of Ta atoms within the alloy of the 400 and 800 °C annealed A2 samples are shown in Fig. 6-3, where the signal of 800 °C sample is the 200 sec-sputtered signal in Fig. 6-2. The almost identical spectra imply that the alloy is stable up to 800 °C.

The composition profile of the A2 alloy detected by the AES is shown in Fig. 6-4. Since there is no standard for Ta-Pt alloys, the atomic ratio can not be accurately determined by the AES analysis. When the average atomic ratios have

been determined by the RBS, the AES is used to detect the composition uniformity of the alloy film. The AES depth profile shows that the atomic composition of both Ta and Pt of the 800 °C annealed samples is almost constant, except at the sample surface. The AES signals at the surface are disturbed by the surface oxide, as confirmed by the XPS results.

Figure 6-5 shows the X-ray diffraction pattern of pure Pt, pure Ta and A2 alloy [13]. The intensity of pure Pt film was divided by a factor of 150. The peak positions of A2 samples after annealing at 400°C to 800°C for 30min are the same. XRD patterns of pure Ta and Pt metal films were used to distinguish σ phase from phases of pure Ta and Pt films. The single peak of pure Ta film is at 33.24° and the single peak of pure Pt film is at 39.76°. The peak position of Ta-Pt alloy is different from the peak positions of pure Ta and Pt films. This peak is considered to belong to the σ phase of the alloy system according to the Ta-Pt binary phase diagram. Furthermore, the peak position is also different from the positions of Pt_3Ta and PtTa [14].

The calculation of d-spaces (the distance between adjacent planes of atoms) according to Bragg's Law shows the d-spacing of the peak $2\theta=38.86°$ is 0.232 nm. The peaks of pure Ta and Pt occurred at $2\theta=33.24°$ and 39.76°, respectively. The corresponding d-spaces are 0.269 nm, and 0.226 nm, respectively. The diffraction peak of A2 samples is very broad, implying that the grain size of A2 alloy is very small. The average grain size can be quantified by the Scherrer equation [15]. The calculated average grain size of A1 alloy for the peak at $2\theta=38.86°$ is 4.37 nm - 8.95 nm as the annealing temperature increases from 400 to 800 °C.

TEM diffraction pattern was used to double check the results of XRD analysis. As shown in Fig. 6-6, only a broadened diffraction ring was clearly observed for

the A1 sample annealed at 800°C and the corresponded d-space is 0.236 nm, which is similar to the value from XRD analysis. The diffraction points come from the Si substrate. The broad XRD patterns and absence of lattice structure in the TEM also suggest a poorly ordered crystal. The fact that XRD intensity increases and peak width decreases would suggest anneal induced crystalline ordering. The peak at 45.56° is only observed in the 800 °C XRD spectrum, while it is not observed in the TEM diffraction pattern. The major difference between the XRD sample and the TEM sample is the surface of the TEM sample sputtered during the preparation of TEM sample. Therefore, the peak at 45.56° in the 800 °C XRD pattern presumably comes from the sample surface. The sample surface of XRD sample is oxidized during exposure to air and the XPS analysis indicates the tantalum oxide and pure Pt compounds in Fig. 6-2. Considering the diffraction patterns of tantalum oxide and Pt, this peak is close to that of Pt (200) which is the secondly most intense phase in Pt powder.[16] Therefore, the peak at 45.56° in the 800°C XRD spectrum is presumed due to Pt (200).The material characteristics of the A2 film are summarized in table 6-2.

6.3.3 Thermal Stability of the Ta-Pt Alloy on SiO_2 and HfO_2

The thermal stability of the Ta-Pt alloys on SiO_2 and HfO_2 was analyzed by the bias polarity-dependence inelastic electron tunneling spectroscopy (IETS) where the Ta-Pt alloy is also a Ta-rich material [17]. The IETS spectra of 500 °C and 800 °C annealed Ta-Pt (A5)/SiO_2 (1.5 nm)/degenerated Si samples are shown in Fig. 6-7. The asymmetric IETS was caused by different bias polarities. The negative gate bias qualitatively reflects the interaction between electrons and phonons near the bottom interface [18]. The signals of Si substrate and SiO_2 phonon modes have

been well identified and analyzed by previous researchers [18, 19]. The signals of Si phonons were at 20 mV and 59 mV and the signals of SiO_2 phonons ranged broadly from 130 to 170 mV [18, 20]. Several additional peaks occurred on the 800 °C annealed sample (800 °C/-Vg in Fig. 6-7). The peak at 77 mV (equal to 640 cm^{-1}) was attributed to amorphous Ta_2O_5 [21, 22]. The observation proved the reaction between Ta of Ta-Pt alloy and SiO_2. The obvious and intense peaks between 90 and 120 mV were qualitatively attributed to deteriorated SiO_2 instead of Ta_2O_5 since the main peaks of Ta-O was located only at 77 mV. Those peaks between Ta-O modes and Si-O modes represent Ta-O-Si (Metal-O-Si) bonds or imperfect SiO_2 (SiO_x) due to the chemical reaction between Ta-Pt gate and SiO_2 [23, 24]. The spectra at positive gate bias were more sensitive to the phonon interaction near the Ta-Pt alloy gate. In addition to the Ta-O peak at 77 mV, the other Ta-O peaks at 38 mV and 49 mV were also observed [21, 22]. Both the 500 °C and 800 °C samples revealed Ta-O bond and the higher and sharper peaks of the 800 °C annealed sample indicated the growth of tantalum oxide.

To investigate the electrical effects of chemical reaction at the interface between Ta-Pt and SiO_2, the work function of Ta-Pt gate after thermal annealing was analyzed. In Fig. 6-8, the barrier height of Ta-Pt electrode extracted by the Fowler-Nordheim tunneling plot increases from 3.18 eV to 3.65 eV as the annealing temperature increases from 500 ^{o}C to 800 ^{o}C. The inset shows that the equivalent oxide thickness is almost unchanged after 800 ^{o}C annealing while the flatband voltage (V_{fb}) changes from the -0.69 V to -0.48 V. These observations imply that the effective work function is increased after 800 ^{o}C annealing and is consistent with the published data that low work function (near the silicon conduction band) of metal gate tends to mid-gap of silicon energy band after

high-temperature annealing [25]. It is believed that the Fermi-level pinning (work function offset) originates from the chemical reaction between metal gate and SiO_2 gate dielectric.

Figure 6-9 shows the IETS spectra of the Ta-Pt (A5)/HfO_2 (3 nm)/degenerated Si samples annealed at 500 °C and 800 °C. Both the negative and positive biased IETS spectra were measured. The peaks were identified based on the literature values [18, 23, 24]. The peaks of Si-O and Hf-O-Si vibration modes of the 500 °C annealed sample were weaker than those of the 800 °C annealed sample, and those of the negative gate-bias IETS spectra were stronger than those of the positive ones. These observations imply the growth of silicon oxide at the lower interface during high-temperature annealing, which has been clearly detected by the transmission electron microscope (TEM) inspection. The shift of the main peak of Hf-O vibration modes after high-temperature annealing was reasonably caused by the crystallization, where the main peak of the 500 °C annealed sample was located at 32 mV and was lower than the main peak of 800 °C annealed sample located at 43 mV [19]. A special peak of the Hf-O vibration mode at 71mV observed on both the 500 °C and 800 °C annealed samples was attributed to the hafnium-silicate but not to pure HfO_2 [23, 26]. This implied that the interfacial layer contained the mixture of Hf-O and O-Si bonds, i.e. Hf-silicate. Another special peak at 38 mV was suspected to the Ta-O vibration mode and was only observed on the positive gate-bias IETS spectra. The peak was small and did not grow even after high-temperature annealing. The Ta-Pt/HfO_2 interface was further inspected by high resolution transmission electron microscopy, as shown in Fig. 6-10. An intermittent monolayer material was observed at the top interface which corresponded to the weak IETS peaks of both 500 °C and 800 °C samples. The

monolayer Ta-O bond might be formed during the Ta-Pt alloy deposition.

The anomaly peak at 7 mV appears in all the SiO_2 and HfO_2 IETS spectra. It is report that metallic electrode will cause the peak near zero bias due to low energy phonon caused by the strong metallic bonds.[27] However the positive bias IETS spectra do not have a stronger peak than that of negative bias IETS spectra. The intensities of this peak do not reasonably consist with the fact of bias polarity-dependence. The intensity of this peak is dependent on the individual sample instead of bias polarity, and it is though to be possible due to charge accumulation while the sample was measured at the beginning [28].

6.3.4 Effect of Impurity Incorporation to the Effective Work Function

The effect of III or V impurity incorporation on the work function of alloy is investigated by the flat band voltage (V_{fb}) change. The V_{fb} is defined as:

$$V_{fb} = \Phi_{ms} - \frac{Q_{eff}}{C_{ox}},$$

where Φ_{ms} is the work function difference between metal gate and silicon substrate, and Q_{eff} and C_{ox} are the effective oxide charges and oxide capacitance, respectively. The samples were fabricated with the same processes and materials so that they are assumed to have the same Q_{eff}. The C_{ox} deviation is less than ±5%, and could only have caused a V_{fb} shift of ±25 mV under the condition of very high oxide charges density of 10^{12} cm^{-2}. Neglecting the V_{fb} deviation caused by the Q_{eff} and C_{ox}, the variation of V_{fb} can reflect the change of work function of metal gate since the silicon substrate is the same for all samples. Figure 6-11 shows the flat band voltage difference (ΔV_{fb}) as a function of implantation conditions of the 500 °C and 800°C annealed samples. The ΔV_{fb} is defined as the difference of flat band

voltage between implanted samples and un-implanted samples. In Fig. 6-11 (a), almost all the ΔV_{fb} of the A6 samples are about -0.05V, except for the 25 KeV BF_2^+ implanted and 800 °C annealed sample. Under all the implantation conditions, the difference of ΔV_{fb} is generally less than 0.05 V. Furthermore, the difference of the ΔV_{fb} of all the 500 °C and 800 °C samples with the same implantation condition is less than 0.02 V.

A condition similar to the ΔV_{fb} independence of implantation conditions was observed on A7 samples, as shown in Fig. 6-11(b). The ΔV_{fb} of A7 samples are approximately at -0.1V for 500 °C annealed samples and at -0.05 V for 800 °C annealed samples. For all the implantation conditions, the difference of the ΔV_{fb} is generally less than 0.05 V. The difference of the ΔV_{fb} for all the 500 °C and 800 °C samples with the same implantation conditions is also about 0.05V. These results indicate that the ΔV_{fb} is almost independent of the implantation conditions and annealing temperatures even for the Ta-rich alloy.

To verify the effect of impurity, the impurity depth distribution in the alloy was analyzed by SIMS. Figure 6-12(a) shows that the depth distribution of arsenic of the A7 samples before and after 800 °C annealing are almost the same. No arsenic is found near the interface between metal gate and silicon dioxide, which implies that the ΔV_{fb} of the implanted sample should not come from the work function change but mainly from oxide charges. For the BF_2^+-implanted samples, boron impurities are distributed near the surface of the alloy before annealing, and boron impurities are diffused and redistributed throughout the alloy film after 800 °C annealing as shown in Fig. 6-12 (b). The depth distributions of P^+ implanted samples before and after annealing are also similar to those of boron in the BF_2^+-implanted samples. Although the impurities of boron can be diffused

throughout the alloy films at 800 °C, the ΔV_{fb} of both A6 samples and A7 samples are still very small. This further reflects that impurities in the alloy do not significantly change the work function of the alloy.

The amount of implanted dose is 5×10^{15} cm^{-2}. The concentration of impurity could be higher than 8×10^{20} cm^{-3} as the impurities are uniformly distributed. Such a high impurity concentration can only cause a ΔV_{fb} of about 0.05 V at most, which should be mostly caused by oxide charges. The implantation dose of MOSFET fabrication is generally less than 5×10^{15} cm^{-2}. The work function variation of the alloy gate affected by the impurities should be much less than 0.05 V and could be neglected. However, it is considered that boron and phosphorus can easily diffuse in the Ta-Pt alloy during the high temperature process, which may cause the impurity penetration into the device channel. To avoid the boron penetration issue, a hard mask and low thermal budget activation process, such as spike annealing, are suggested during implantation and activation, respectively. [29] The C-V characteristics of the 800 °C annealed A7 samples are shown in Fig. 6-13. These C-V curves are almost identical. Curves of the implanted samples are slightly smoother than that of the un-implanted one. The distortion of curves comes from the interface states. The interface states and oxide charges are suspected to be caused by the ion implantation process which is the reason why the flat band voltage of the implanted sample is slightly shifted from that of the un-implanted one [30, 31].

6.4 Summaries and Conclusions

This chapter investigates the work function modulation of Ta-based binary alloys and the physical properties of Ta-Pt alloys. Any work function can be

obtained by the binary metallic alloy system with a proper composition of high and low work function metals. The work function of the Ta-based metal alloys (Ta-Ti and Ta-Pt) can be modulated from 4.16 eV to 5.2 eV continuously. Such a wide range work function modulation makes them suitable for use in CMOS applications.

IETS spectroscopy provides a direct inspection of thermal instability at interface of Ta-Pt/SiO_2/Si and Ta-Pt/HfO_2/Si structures. The analysis of results provides a direct evidence that the Fermi-level pinning of Ta-Pt metal gates on SiO_2 is counted for the generation of extrinsic states due to the formation of the tantalum oxide. The Ta-Pt/HfO_2 interface is more thermally stable than the HfO_2/Si interface.

The work function would not be changed by the incorporation of the implantation impurities such as arsenic ions, boron ions and phosphorus ions. Therefore, the control of threshold voltage of the alloy-gate MOSFET could be easier. However, the boron impurity still can easily diffuse in the Ta-Pt alloy gates. To prevent the boron penetration of the Ta-Pt alloy gate, a hard mask and low thermal budget activation process, such spike annealing, are suggested to be used during implantation and activation.

In conclusion, an easy and practical method of wok function modulation is demonstrated by the binary metallic alloy system. The Ta-Pt alloys with tunable work function are potential gate materials for SOI CMOS if the underlying gate dielectric is Hf-based high-k dielectric. It could be integrated in the conventional MOSFET process with high-K material (HfO_2).

References

[1] C. D. Gelatt, Jr. and H. Ehrenreich, “Charge Transfer in Alloys: AgAu,” *Phys. Rev. B* vol. 10 no.2 pp.398-415, 1974.

[2] R. Ishii, K. Matsumura, A. Sakai, T. Sakata, ”Work Function of Binary Alloys,” Appl. Surf. Sci. vol.169-170 pp.658-661, 2001.

[3] H. Zhong, S. N. Hong, Y. S. Suh, H. Lazar, G. Heuss and V. Misra, “Properties of Ru-Ta alloys as gate electrodes for NMOS and PMOS silicon devices,” in *2001 Int. Electron. Devices Meet. Tech. Dig*., pp.432-435.

[4] B. Y. Tsui and C. F. Huang, “Wide Range Work Function Modulation of Binary Alloys for MOSFET Application,” *IEEE Electron Device Lett. vol.* 24 no.3 pp.153-155, March 2003

[5] T.-L. Li, C.-H. Hu, W.-L. Ho, H.C.-H. Wang, C.-Y. Chang,” Continuous and Precise Work Function Adjustment for Integratable Dual Metal Gate CMOS Technology Using Hf-Mo Binary Alloys,” *IEEE Trans. Electron Dev.* vol. ED-52, no. 6, pp.1172-1179, 2005.

[6] The University of Sheffield and WebElements Ltd, UK: WebElements Periodic table. Available: http://www.webelements.com/.

[7] C. Kittle, *Introduction to Solid State Physics*, 5th ed. Wiley, New York 1976, p.167.

[8] B. Y. Tsui, and C. F. Huang, “Investigation of Cu/TaNx Metal Gate for Metal-Oxide-Silicon Devices,” *J. Electrochem. Soc*., vol. 150, pp.G22-G27, 2003.

[9] R. M. Waterstrat, “Analysis of Selected Alloys in the Systems Cr-Pd, Cr-Ru, V-Pd and Ta-Pt,” *Journal of the Less-Common Metals,* vol.80, pp.31-36, 1981.

[10] T. B. Massalski, *Binary Alloy Phase Diagrams 2nd ed*., Metals Park, American Society for Metals, Ohio, 1990, p.3133

[11] M. L. Steigerwald, R. M. Fleming, R.L. Opila, D.V. Lanf, R.B. Van Dover, and C. D. W. Jones," Material and electrical characterization of carbon-doped Ta_2O_5 films for embedded dynamic random access memory applications" *J. Appl. Phys.* 91, 308, 2002.

[12] N J Shevchik and D Bloch, "XPS d Band and Core level of Pt-Ni alloys," *J. Phys. F: Metal Phys.* vol. 7, pp.543-550, 1977.

[13] X. Chen, H. L. Frisch and A. E. Kaloyeros, "Low Temperature Plasma-Promoted Chemical Vapor Deposition of Tantalum from Tantalum Pentabromide for Copper Metallization, "*J. Vac. Sci. Technol. B*, 16, pp.2887-2990, 1998.

[14] B. C. Giessen, R. H. Kane, and N. J. Grant, "On the Constitution Diagram Ta-Pt between 50-100 At. Pct Pt," *Trans. Met. Soc. AIME,* vol.223, pp.855-864, 1965.

[15] H. P. Klug and L. E. Alexander, *X-ray diffraction Procedure for Polycrystalline and Amorphous Materials 2nd ed.*, John Wiely & Sons, New York, 1974, p.656.

[16] D.-Z. Xie, D.-Z. Zhu, and P.-X. Cao, "Damage and thermal annealing of dual-ions (Pt and S) implanted (100)YSZ material," *J. phys: Condens. Matter*, vol.9, pp.4377-4383, 1997.

[17] C. F. Huang, B. Y. Tsui, P. J. Tzeng, L. S. Lee, and M. J. Tsai, "A Study on the Interface Thermal Stability of Metal-Oxide-Semiconductor Structure by Inelastic Electron Tunneling Spectroscopy," *Appl. Phys. Lett.* vol.88, pp.262909-1-262909-3, 2006.

[18] W. He and T.P. Ma, "Inelastic Electron Tunneling Spectroscopy Study of Ultrthin HfO_2 and HfAlO," *Appl. Phys. Lett.* vol. 83, no. 13, pp. 2605-2607, 2003.

[19] G. Chynoweth, R. A. Logan, and D.E. Thomas," Phonon-Assisted Tunneling in Silicon and Germanium Esaki Junctions" *Phys. Rev.,* vol. 125, pp.877-881,1962.

[20] G. Salace, C. Petit, and D. Vuilaume, "Inelastic electron tunneling spectroscopy:

Capabilities and limitations in metal–oxide–semiconductor devices," *J. Appl. Phys.*, vol. 91, pp.5896-5901, 2002.

[21] R. A. B. Devine," Nondestructive measurement of interfacial SiO2 films formed during deposition and annealing of Ta2O5," *Appl. Phys. Lett.,* vol. 68, pp.1924-1926, 1996.

[22] H. Ono, and K.-I. Koyanagi, "Infrared absorption peak due to Ta = O bonds in Ta_2O_5 thin films," *Appl. Phys. Lett.*, vol. 77, pp.1431-1433, 2000.

[23] K. T. Queeney, M. K. Weldon, J. P. Chang, Y. J. Chabal, A. B. Gurevich, J. Sapjeta, and R. L. Opila, "Infrared spectroscopic analysis of the Si/SiO2 interface structure of thermally oxidized silicon," *J. Appl. Phys.*, vol.87, pp.1322-1330, 2000.

[24] D. A. Neumayer and E. Cartier, "Materials characterization of ZrO2–SiO2 and HfO2–SiO2 binary oxides deposited by chemical solution deposition," *J. Appl. Phys.*, vol. 90, pp.1801-1808, 2001.

[25] H. Y. Yu, C. Ren, Y.-C. Yeo, J. F. Kang, X. P. Wang, H. H. H. Ma, M.-F. Li, D.S.H. Chan, and D.-L. Kwong," Fermi pinning-induced thermal instability of metal-gate work functions," *IEEE Electron Device Lett.* vol. 25, pp.337-339, 2004.

[26] M. A. Quevedo-Lopez, J. J. Chambers, M. R. Visokay, A. Shanware, and L. Colombo," Thermal stability of hafnium–silicate and plasma-nitrided hafnium silicate films studied by Fourier transform infrared spectroscopy," *Appl. Phys. Lett.,* vol. 87, p.12902, 2005.

[27] R. Magno and J. G. Adler, "An Electron Tunneling Study of the Formic Acid on Different Surface," *J. Appl. Phys.*, vol. 49, pp.4465-4467, 1978.

[28] W.-K. Lye, E. Hasegawa, T.-P. Ma, and R. C. Barker, "Quantitative Inelastic Tunneling Spectroscopy in the Silicon Metal-Oxide-Semiconductor System," *Appl. Phys. Lett.* vol. 71, no. 7 pp.2523-2525, 1997.

[29] S. Saito, S. Shishguchi, A. Mineji, and T. Matsuda, "Ultra Shallow Junction Formation by. RTA at High Temperature for Short Heating Cycle Time," *Mater. Res. Soc. Symp*. 1998 Proc. 532, p.3.

[30] M. J. Goeckner, S. B. Felch, J. Weeman, S. Mehta, and J. S. Reedholm," Evaluation of charging damage test structures for ion implantation processes" *J. Vac. Sci. Technol. A,* vol.17, pp.1501-1509, 1999.

[31] M. Current, W. Lukaszek, and M. Vella, "Control of Wafer Charging During Ion Implantation: Issues, Monitors and Models" in the *5th International Symposium on Plasma Process-Induced Damage*, Santa Clara, CA, May 22-24, 2000.

Table 6-1 Some conditions of binary metal alloy samples and their compositions.

Sample ID	Sputtering power (W)			Implant	anneal	Atomic ratio (at. %)		
	Ta Target	Pt Target	Ti Target			Ta	Pt	Ti
A1	DC/50	0	DC/50	No	FA	63	0	37
A2	DC/100	DC/30	0	No	FA	74	26	0
A3	DC/50	DC/30	0	No	FA	65	35	0
A4	DC/30	DC/30	0	No	FA	58	42	0
A5	DC/30	RF/30	0	No	FA	85	15	0
A6	DC/35	RF/35	0	Yes; No	RTA	82	18	0
A7	DC/20	RF/40	0	Yes; No	RTA	65	35	0

FA : Furnace annealing; RTA: rapid thermal annealing

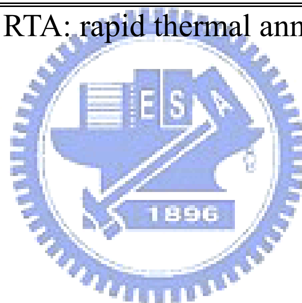


Table 6-2 Material characteristics of Ta, Ta-Pt (A2), and Pt films.

Thin film	Phase(orientation)	d-spacing (nm)	Grain size (nm)	BE of Ta $4f_{7/2}$ (eV)	BE of Pt $4f_{7/2}$ (eV)
Ta/400°C	βTa (002)	0.269	21.11	22.1	
Ta-Pt(A2)/400°C	σ	0.232	4.37	22.6	71.9
Ta-Pt(A2)/800°C	σ	0.232	8.95	22.6	71.9
Pt/800°C	Pt (111)	0.226	39.77		71.2

BE: binding energy

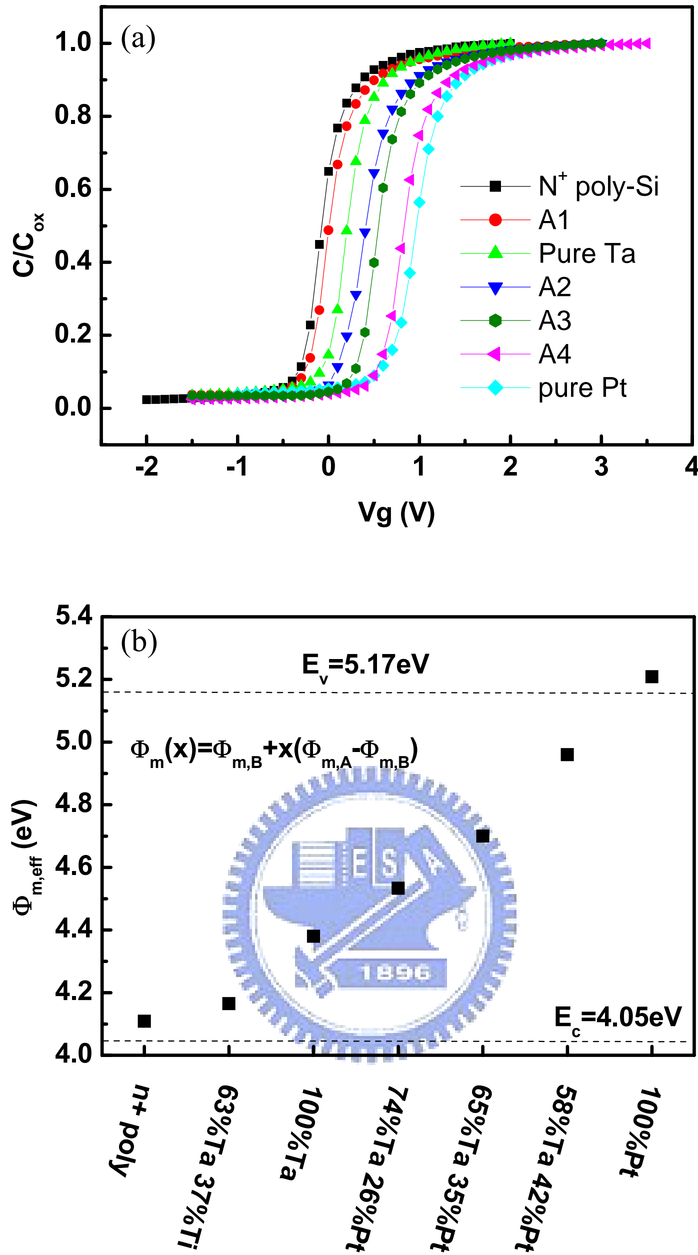


Fig. 6-1 (a) The C-V characteristics of alloy samples annealed at 400 °C N_2 ambient for 30 min, as well as the N^+ poly, pure Ta, and Pt samples. The (b) Effective work function ($\Phi_{m,eff}$) of alloys extracted from the C-V curves by a theoretical C-V curve fitting. The $\Phi_{m,eff}$ approximates to $\Phi_{m,B} - x(\Phi_{m,A}-\Phi_{m,B})$ where x parameter counters for the concentration of element B in alloys and has the value between 0 and 1.

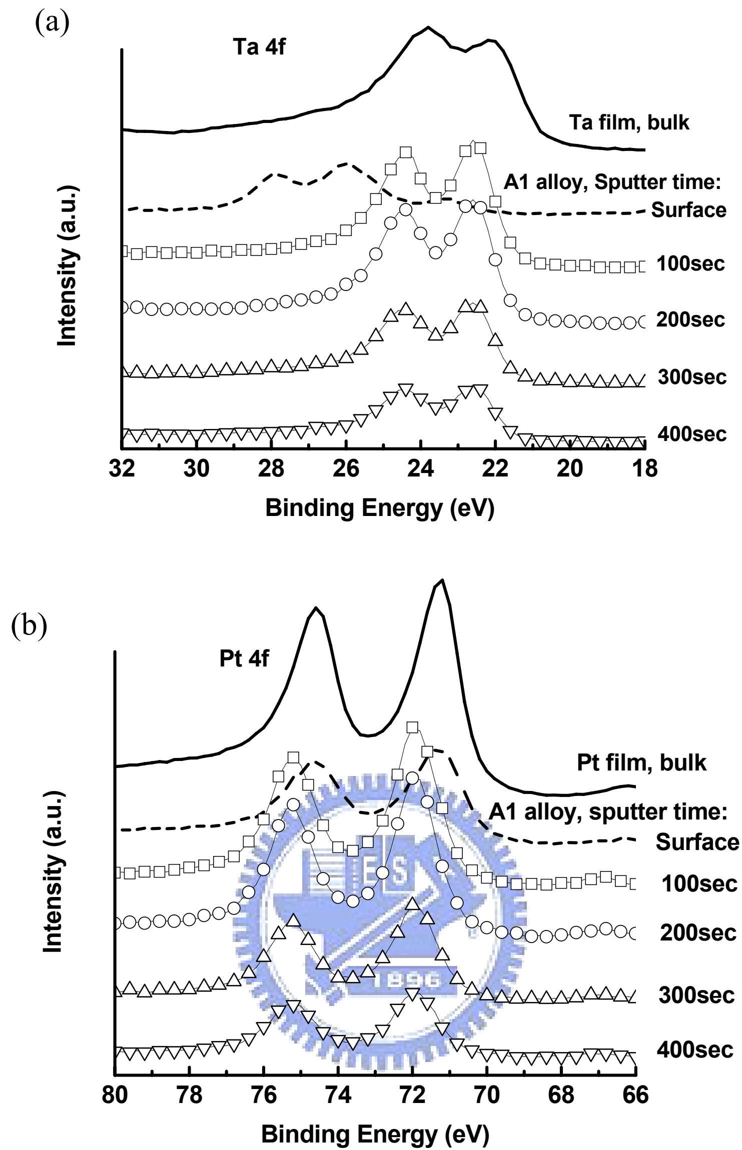


Fig. 6-2 Binding energies of (a) Ta atoms and (b) Pt atoms of the 800 °C annealed A2 samples detected by x-ray photoelectron spectroscopy (XPS) after Ar sputtering for various times. The binding energies of Ta and Pt in pure single metal films were also revealed as references.

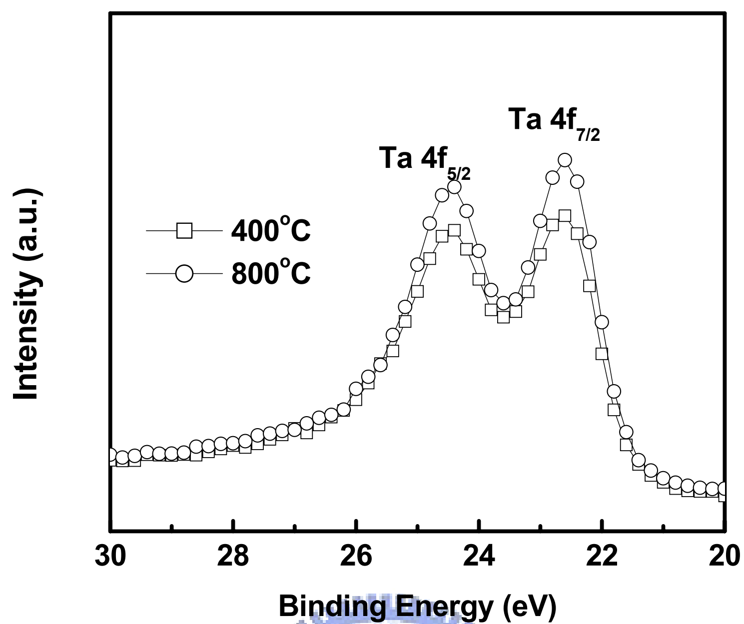


Fig. 6-3 Binding energies of Ta atoms of the 400 °C and 800 °C annealed A2 samples detected by x-ray photoelectron spectroscopy (XPS) after Ar sputtering for 200 sec.

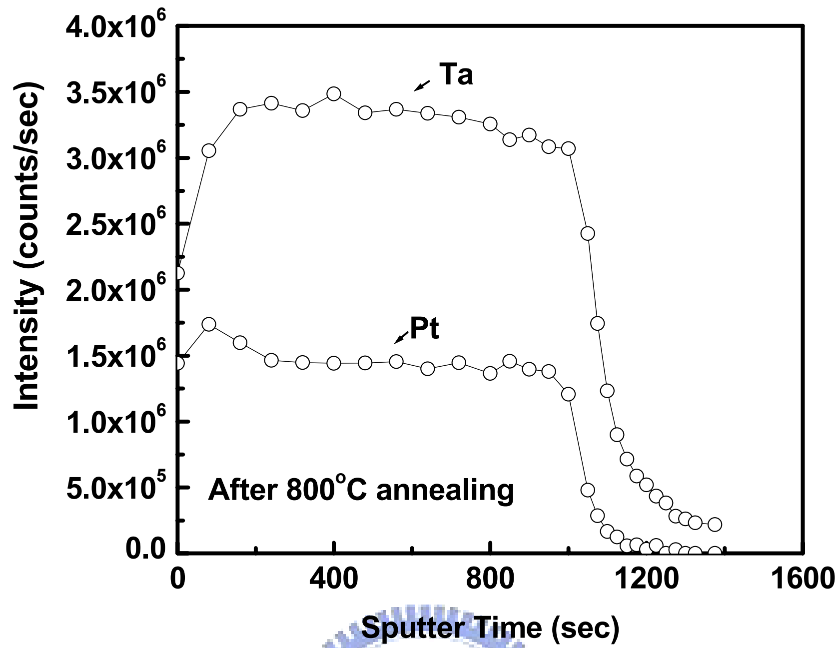


Fig. 6-4 Depth distribution of the Ta and Pt atoms of the 800 °C annealed A2 samples detected by Auger electron spectroscopy (AES).

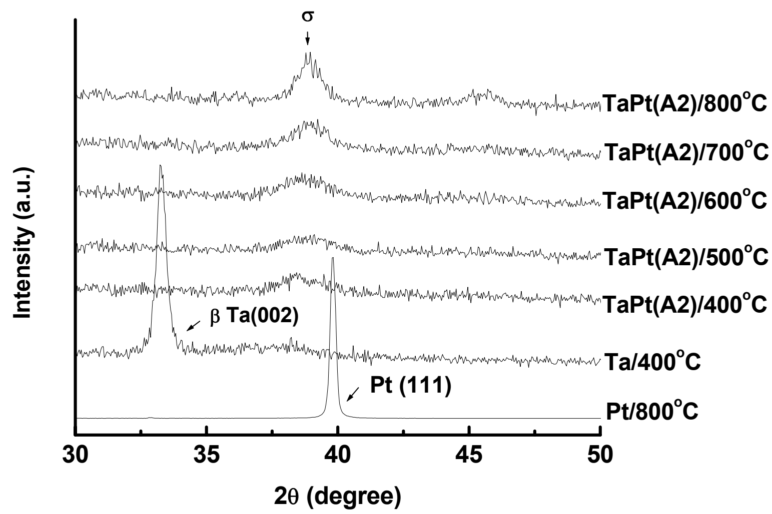


Fig. 6-5 X-ray diffraction (XRD) patterns of the A2 samples after 400 °C, 500 °C, 600 °C, 700 °C, and 800 °C annealing for 30 min. The XRD patterns of pure Ta and Pt films annealed for 30min are also given as references.

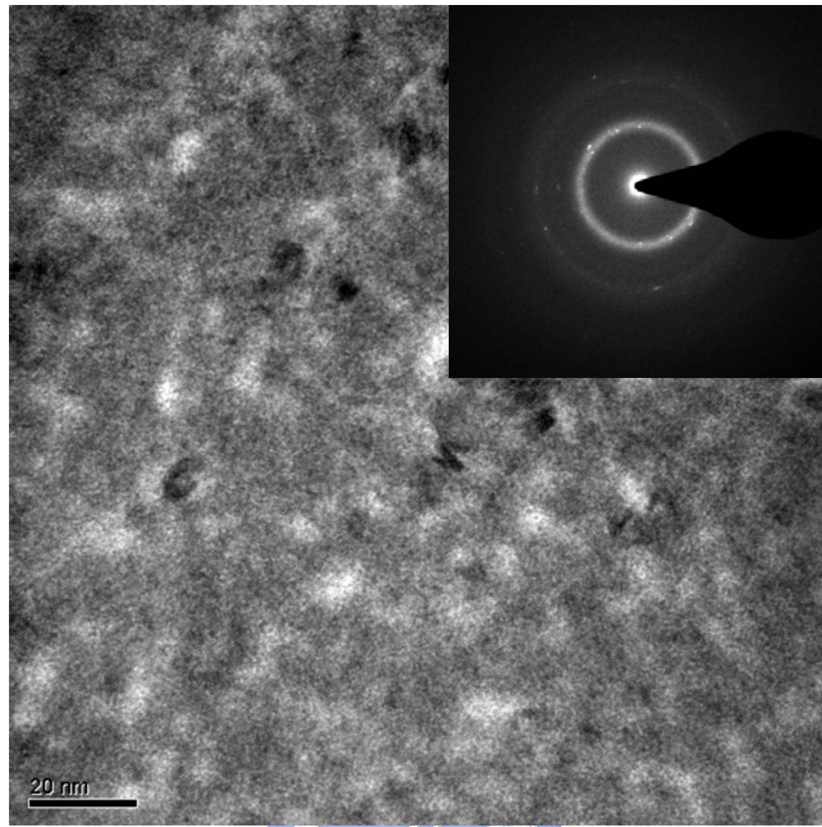


Fig. 6-6 TEM plane view of the A1 sample annealed at 800 ^{o}C for 30 min. The inset is its relative TEM diffraction pattern.

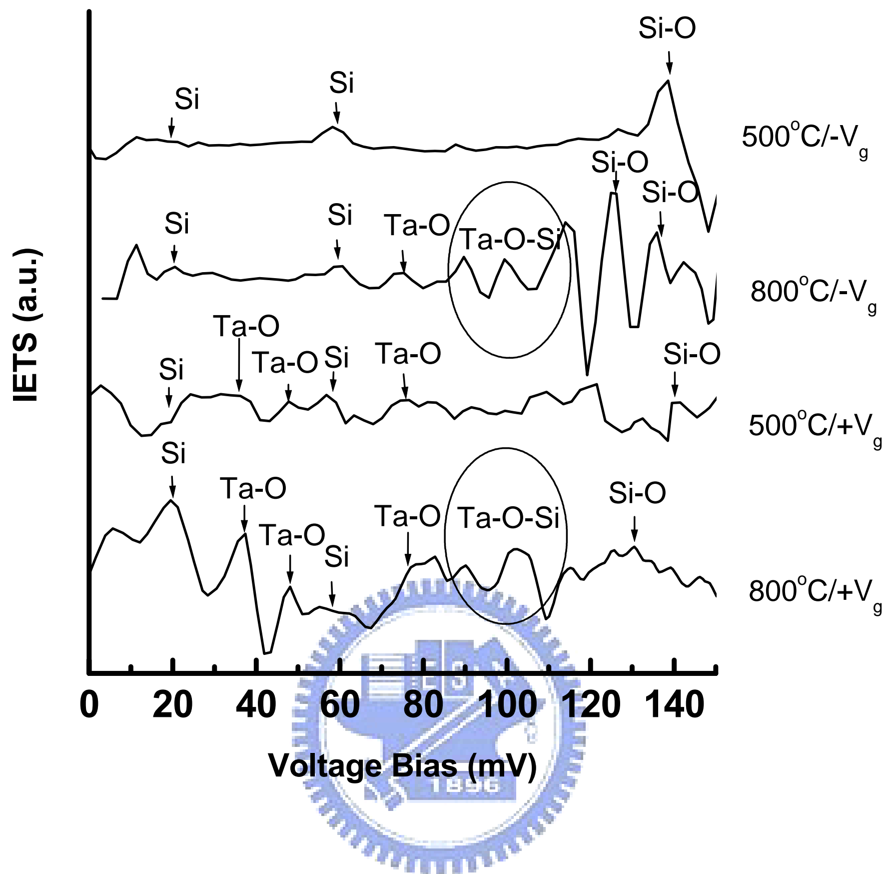


Fig. 6-7 The inelastic electron tunneling spectra of Ta-Pt (A5)/SiO_2/degenerated Si samples annealed at 500 °C and 800 °C.

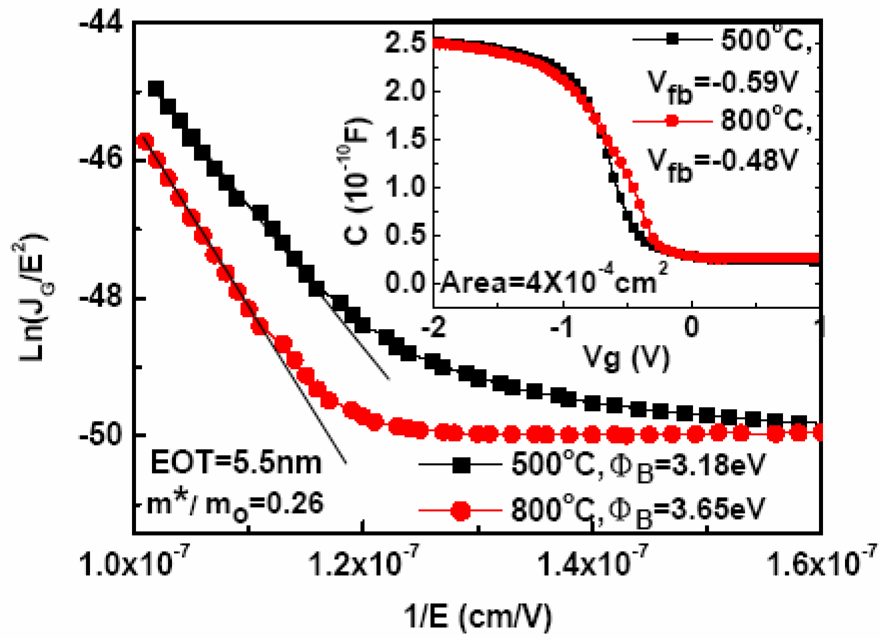


Fig. 6-8 Fowler-Nordheim plot of the Ta-Pt (A5)/SiO_2 (5.5 nm)/Si structure after annealing at 500 °C and 800 °C for 30 sec.The m^*/m_o=0.26 is used to calculate the barrier height (Φ_B). The inset shows the relative C-V curves.

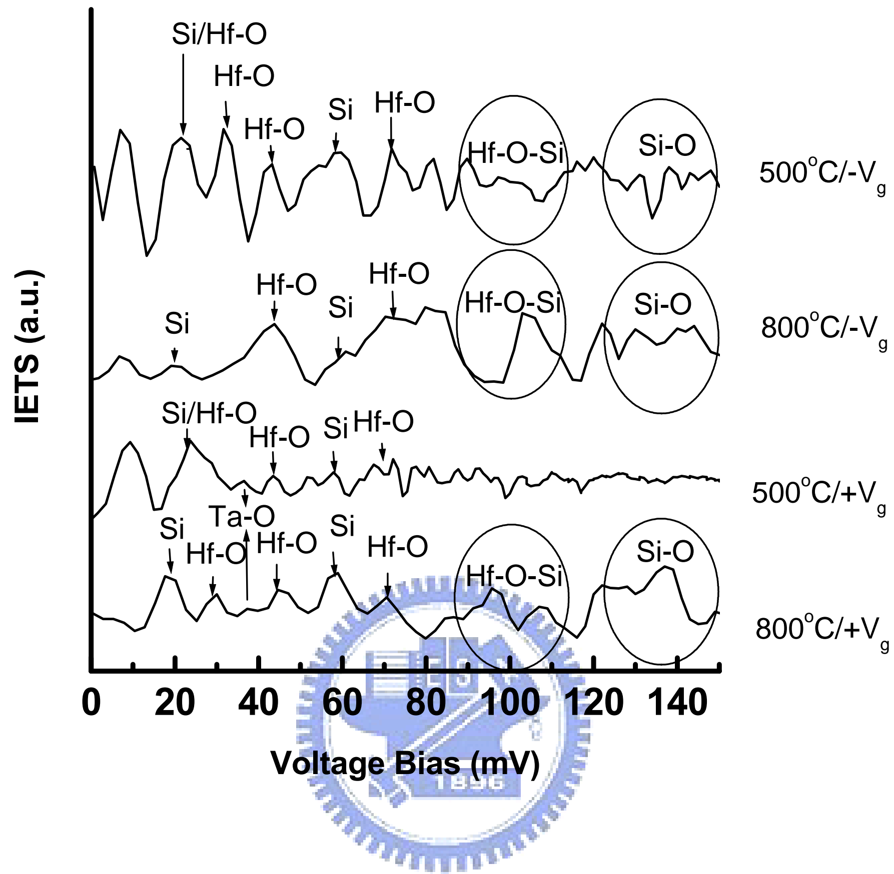


Fig. 6-9 The inelastic electron tunneling spectra of Ta-Pt (A5)/HfO_2/degenerated Si samples annealed at 500 °C and 800 °C.

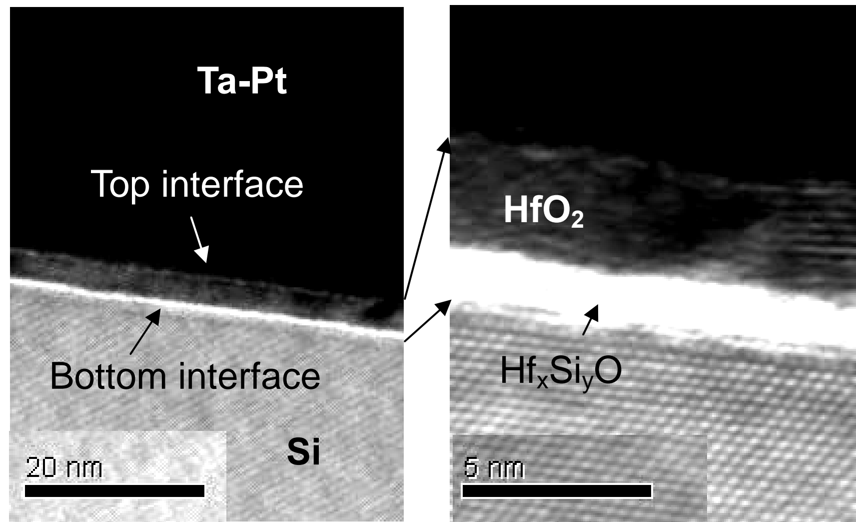
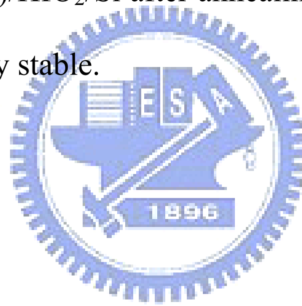


Fig. 6-10 TEM photos of Ta-Pt (A5)/HfO_2/Si after annealing at 800 ^{o}C RTA which confirms that Ta-Pt on HfO_2 is chemically stable.


ES A 1896

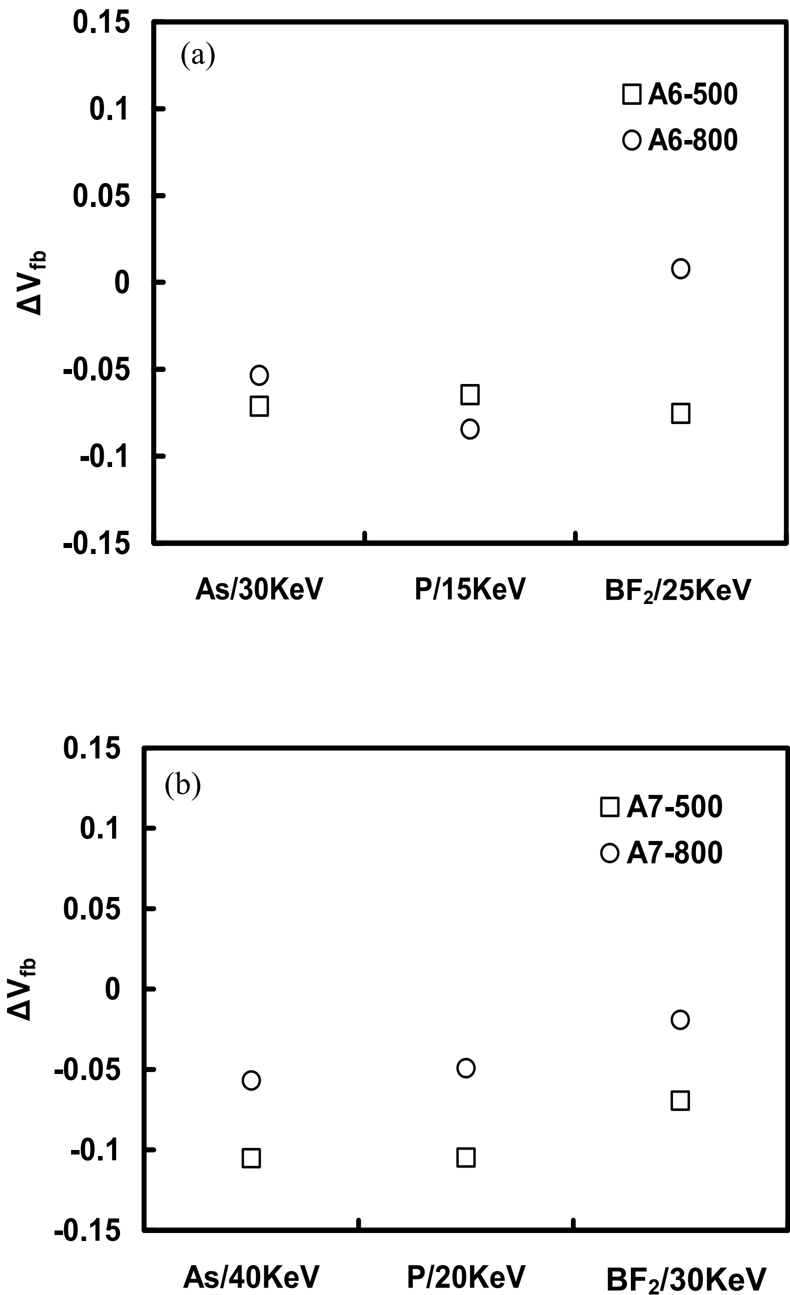


Fig. 6-11 The flat-band voltage shift (ΔV_{fb}) of (a) A6 samples and (b) A7 samples with various implantation conditions after 500 °C and 800 °C annealing for 30 sec. The ΔV_{fb} value is the difference between the flat-band voltage of the un-implanted and implanted samples.

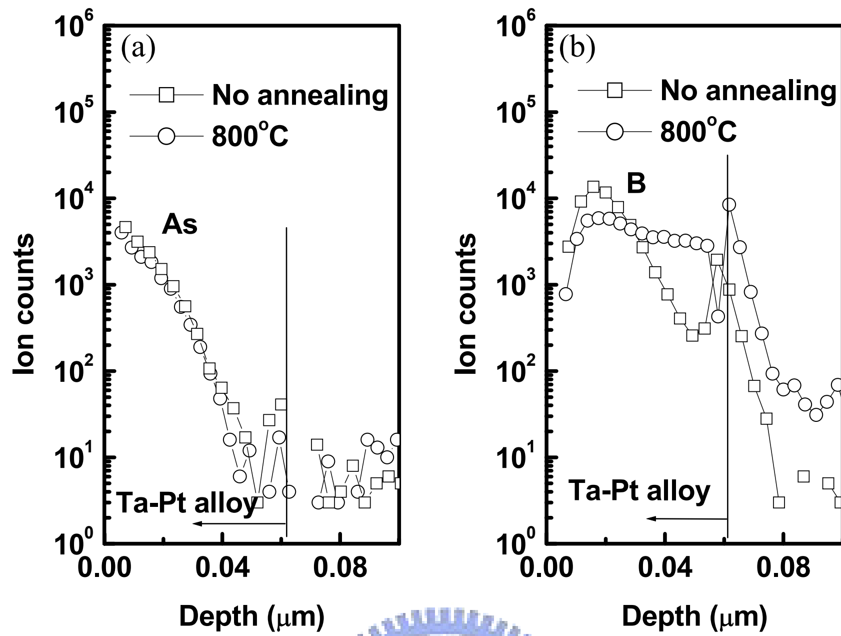


Fig. 6-12 Depth distributions of the (a) arsenic and (b) boron implanted A7 sample before and after 800 °C annealing for 30 sec detected by SIMS. The energies of implantation are 50 KeV and 30 KeV for As^+ and BF_2^+, respectively, and both dosages are 5×10^{15} cm^{-2}.

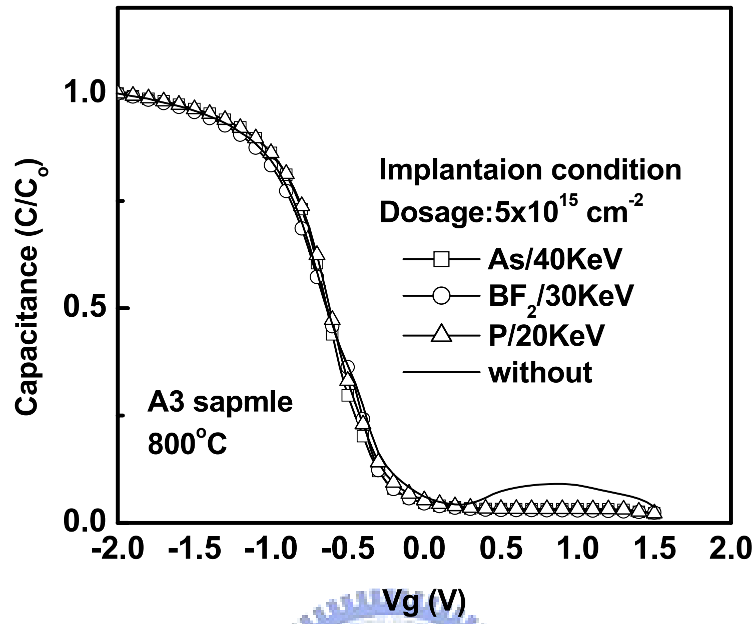


Fig. 6-13 Capacitance-Voltage characteristics of the 800 °C annealed A7 samples with and without ion implantation.

Chapter 7

Fully Ni-Silicided Gates

7.1 Introduction

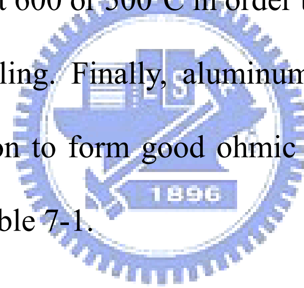
Fully silicided (FUSI) gates have drawn lots of attention due to the easy integration with the conventional MOSFETs fabrication and the wide range of effective work function adjusted by the impurities pile-up at the interface between silicide layer and gate dielectric layer [1- 5]. The work function shift of FUSI gate is dependent on the type and concentration of impurities and the underlying dielectric as well as the material of FUSI gates. NiSi FUSI gate has a work function near the mid-gap of silicon energy band which can be tuned toward conduction band of Si by As, Sb and P impurities and toward valence band of Si by B, and Al impurities as the underlying dielectric is SiO_2 [1, 6]. It is believed that the impurities piled up at interface will induce the electrical dipoles to affect effective work function [6]. The manner of impurity pile-up at the interface is attributed to snow-plow phenomenon - the impurities are pushed toward the interface of FUSI gates during the silicidation since the solid state solubility of impurities in silicide is low [7, 8]. However, Fermi-level pining effect occurs on the NiSi/Hf-based dielectric stack, which is unfavorable to the work function tuning by impurities. Another method of work-function adjustment controlled by the phases of nickel silicides including Ni_3Si, NiSi, and $NiSi_2$ was proposed to be better as the dielectric is Hf-based dielectric [9]. In this chapter, impurities (P and BF_2) were implanted to silicides (ITS) after NiSi full-silicidation and samples were annealed to redistribute impurities at 600

^{o}C. We will investigate the effect of work function on both NiSi/SiO_2 and NiSi/HfO_2 stack due to the impurity redistribution. The electrical effects of silicidation and post silicidation annealing are also investigated.

7.2 Process Flow

MOS capacitors were fabricated as test devices. The initial material was 6-in. p-type (100)-oriented single crystalline wafers with a resistivity of 15-25 Ω-cm. After wafer cleaning, wet oxidation was performed to grow a silicon dioxide layer 550 nm thick patterned to create active regions of MOS capacitors by conventional lithography and wet etching. This thick oxide layer was used to isolate the MOS capacitors from each other. After active region patterning, two sets of gate dielectric layers were grown or deposited. The first set, the SiO_2 sample, was a single SiO_2 layer of 5.5 nm, 10.5 nm and 17 nm thick thermally grown in dry oxide at 900 ^{o}C. The second set, the HfO_2 sample, had an HfO_2/SiO_2 stack structure with combined thicknesses of 2/5.5 nm, 2/10.5 nm, and 2/17 nm, where the SiO_2 layers were grown with the single-layer ones at the same time, and HfO_2 was deposited by the metal organic chemical vapor deposition (MOCVD) system at 500 °C. In order to minimize oxide charges in the HfO_2 layer, HfO_2 samples were annealed at 1000 °C in N_2 ambient for 30 sec after HfO_2 deposition. A 50 nm thick amorphous silicon layer was deposited on both SiO_2 and HfO_2 samples with SiH_4 gas as precursor at 550 °C in a low pressure chemical vapor deposition furnace. Following deposition of an amorphous silicon layer, the silicon gate electrode was patterned by conventional lithography and dry etching techniques. Some HfO_2 samples were doped by BF_2^+ implantation at 10 KeV with dosages of 1×10^{15} cm^{-2} and were activated at 900 °C for 30 sec in N_2 ambient, as were the pre-doped

ones. All wafers were dipped in dilute HF solution, and then a 25 nm-thick nickel layer was deposited by sputter deposition. Silicidation was performed by a two-step annealing scheme [10]. First, nickel and silicon were thermally reacted at 300°C in N_2 ambient for 1 hr in a furnace, and then the unreacted nickel was selectively etched by a solution of sulfuric acid. A second annealing at 600 °C for 30 sec was performed in a rapid thermal annealing (RTA) system to transform Ni_2Si to NiSi. The undoped samples were implanted by BF_2^+ or P^+ ions at 10 KeV with dosages of 1×10^{15} cm^{-2} and then annealed at 600 °C for 30 sec in an RTA system. These were referred to as implant-to-siliicde (ITS) samples. In addition, some undoped samples received either one-step silicidation at 500 °C RTA for 30 sec or two-step silicidation at 300 °C for 1 h and 500 °C RTA for 30 sec. Prolonged high-temperature annealing was also performed at 600 or 500°C in order to investigate the electrical effects of thermal silicidation and annealing. Finally, aluminum film was deposited on the wafer backside by thermal evaporation to form good ohmic contact. The process conditions for MOS capacitors are listed in Table 7-1.



7.3 Results and Discussions

7.3.1 Thermal Stability of the Fully-Silicided NiSi Gates

The phase transitions of nickel silicide analyzed from the X-ray diffraction pattern are shown in Fig. 7-1, where Ni_2Si is the main silicide phase after the first step annealing at 300 °C in N_2 ambient. Following the second step annealing at 600 °C for 30 sec, the main phase is transformed from Ni_2Si to NiSi. With further annealing at 600 °C, the main phase is not changed and remains NiSi, which implies that the NiSi phase is stable at 600 °C. The Ni_2Si phase was also observed after a prolonged period annealing at 600 °C, reflecting that

the nickel silicide layer is nickel-rich. This is confirmed by electron dispersion analysis. The atomic ratio of Ni/Si is about 1.15/1.

The sheet resistance of NiSi on SiO_2 versus annealing times was measured to examine the thermal stability of NiSi on SiO_2 film at 600 °C. As shown in Fig. 7-2, the resistance decreases and approaches the limit of 6.4 Ω/sqr after a 240-min annealing. This saturated value indicates that NiSi on silicon dioxide is stable up to 600 °C without reaction with the SiO_2. In addition, the cross-sectional structure of the NiSi/HfO_2 (HfO_2 sample) was inspected by transmission electron microscope (TEM). The NiSi gate was almost fully silicided, but an interfacial layer (IL) between NiSi and the HfO_2 layer was observed, as shown in Fig. 7-3. The IL of ITS sample is smeared with the NiSi layer and is not continuous, whereas for the pre-doped sample the interface between NiSi and HfO_2 is clear and sharp, as shown in Fig. 7-3(b). The thin and light colored IL revealed in the pre-doped sample is about 1.1 nm thick, and is almost continuous. The compound types of the IL in ITS samples are further identified by X-ray photoelectron spectrometer (XPS) analysis after the NiSi film been selectively removed by a mixture of 95 % IPA and 5 % HF. The IL is also etched by the mixture with a relatively low etching rate in order to detect the underlying HfO_2 by XPS. Electron binding energies of Hf 4f, Si 2p, and O 1s core levels are shown in Fig. 7-4. Hf-O signals relative to a binding energy of $4f_{7/2}$ (19 eV) and $4f_{5/2}$ (20.6 eV) are detected on both surfaces due to the thin and discontinuous interfacial layer. The main O 1s binding energy shifts from the Si-O type with a binding energy of 534.1 eV to the Hf-O type with a binding energy of 532.3 eV as the analyzed surface changes from IL to HfO_2. It is worthy to note that the observed Si binding energy is equal to 104.8 eV, which is related to the Si-O type [25]. These results indicate that the IL is a SiO_2-like material, not a HfO_2-like material, since the Si-O signal is stronger than that of

the Hf-O signal. The IL is similar to SiO_2, and no Hf-Si was observed. Hence the NiSi on HfO_2 film should be stable under 600 °C thermal annealing. To further consider the reason why the interfacial layer was formed, the TEM cross-sectional views of the as-deposited amorphous Si/HfO_2/SiO_2/Si structure and the sample (P-type poly-Si sample) implanted with BF_2 and annealed at 900 °C for 30 sec in N_2 atmosphere were analyzed and is shown in Fig. 7-5. No obvious interfacial layer was observed between Si gates and HfO_2 layers. The interfacial layer only existed in the FUSI samples after silicidation, which indicates that the interfacial layer was formed during the Ni silicidation. This phenomenon has been previously noted but the formation mechanism is not yet clear [11].

7.3.2 Effect of Work Function due to Impurity Redistribution

The effective work function ($\Phi_{m,eff}$) of the NiSi gate is shown in Fig. 7-6 for both SiO_2 and HfO_2 samples. The $\Phi_{m,eff}$ values of NiSi on SiO_2 are 4.36, 4.58, and 5.14 eV for P^+ doped, undoped, and BF_2^+ doped samples, respectively, while the relative values of HfO_2 samples are 4.44, 4.49, 4.62 eV. In addition, the BF_2^+ pre-doped HfO_2 sample has a $\Phi_{m,eff}$ value of 4.93 eV. Compared with the undoped NiSi gate, the $\Phi_{m,eff}$ of P^+ doped SiO_2 sample shift -0.22 eV while the $\Phi_{m,eff}$ of the P^+ doped HfO_2 sample shift -0.05 eV. The effective work function difference ($\Delta\Phi_{m,eff}$) between undoped and BF_2^+ doped NiSi SiO_2 samples is 0.56 eV while the$\Delta\Phi_{m,eff}$ between undoped and BF_2^+ doped NiSi HfO_2 samples is 0.13 eV. For the BF_2^+ pre-doped HfO_2 sample, the shift of effective work function is as large as 0.44 eV. The shift of $\Phi_{m,eff}$ implies that the work function of pure NiSi can be adjusted by the incorporation of BF_2^+ or P^+ species, and the work function change is dependent upon the dopant species and the underlying dielectric. To investigate the dopant effect on $\Phi_{m,eff}$, the doping profile in NiSi film analyzed by secondary ion mass spectrometry is shown in Fig. 7-7. The boron and phosphorus profiles of the as-implanted

samples reveal only one peak near the surface of the NiSi film. After a 600 °C annealing for the ITS sample and NiSi silicidation for the pre-doped sample, dopant pile-up at the interface between the NiSi layer and the gate dielectric is observed. The accumulated dopant at the interface between NiSi and gate dielectric plays the main role in determining the effective work function [1, 12, 13]. The dopant profiles of P^+ in NiSi for SiO_2 and HfO_2 samples are almost identical. This implies that the concentration is not the only factor affecting the $\Phi_{m,eff}$.

Although the $\Phi_{m,eff}$ of pure NiSi on HfO_2 is close to that on SO_2, a large difference in the $\Phi_{m,eff}$ between the SiO_2 and HfO_2 are presented after doping. The Fermi-level pinning effect usually causes the $\Phi_{m,eff}$ to move toward the mid-gap of the silicon energy band, i.e. 4.61 eV [14]. The pure NiSi has a work function near the Φ_{midgap}, so that the $\Phi_{m,eff}$ of NiSi on SiO_2 and on HfO_2 are close. The results from the pure NiSi gate on HfO_2 do not rule out the Fermi-level pinning effect. On the other hand, the doped NiSi gate reflects obvious effects of the under-layer dielectric for the ITS samples. Therefore, the small difference in $\Phi_{m,eff}$ in the HfO_2 sample may reasonably be explained as the Fermi-level pinning effect, which is very significant on a high-dielectric constant dielectric [14, 15]. Whereas the pre-doped HfO_2 sample has a continuous SiO_2 interlayer between the NiSi film and the HfO_2 film, the interlayer screens the effect of HfO_2 so that the $\Phi_{m,eff}$ value is close to that of the SiO_2 sample. Although the interfacial layer can suppress Fermi-level pinning, the increase in the effective oxide thickness is adverse to device scaling [16]. The $\Phi_{m,eff}$ of a P-type poly-Si gate on HfO_2 is also shown in Fig. 7-6. The value is 4.87 eV and is 0.22 eV lower than work function of poly-Si. This data implies that the Fermi-level pining effect occurres [16, 17].

7.3.3 Effects of Silicidation and Post Annealing on Electrical

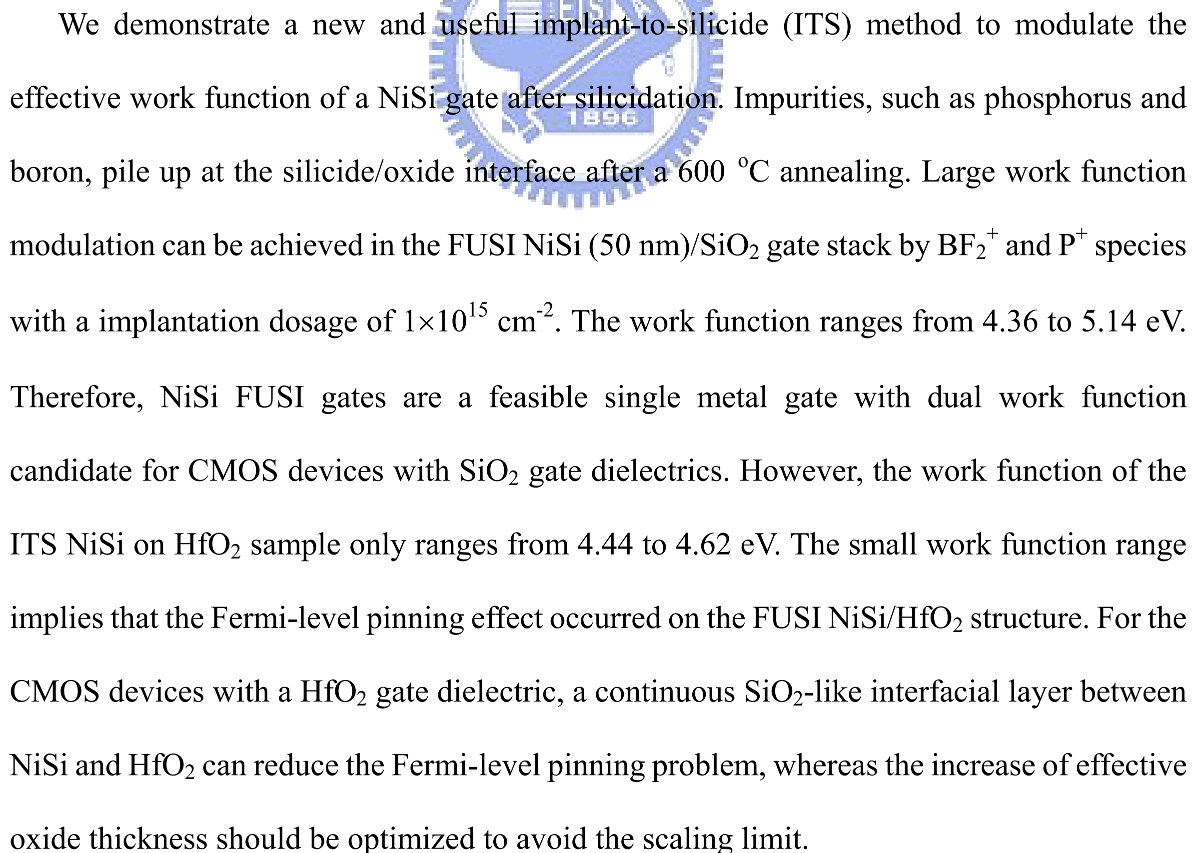
Characteristics

Ni impurities in silicon can form a deep level energy state of silicon band gap to induce increased leakage current [18]. Bias temperature stress (BTS) was applied to 600°C-silicided undoped NiSi/SiO_2 (10.5 nm)/Si samples to examine the possibility of Ni diffusion from the NiSi gate into SiO_2 and the silicon substrate. As shown in Fig. 7-8, the high frequency C-V curves are almost identical after BTS under the typical condition of ±2 MV/cm at 200 °C for 30 min [19]. There are no distortion in the C-V curves and no increase in the capacitance at inversion mode, which is very sensitive to the defects caused by the substitutional impurities in Si substrate. Thus it is presumed that under the BTS condition employed, the SiO_2 grown in dry oxygen ambient is a good diffusion barrier to Ni atoms [20-23].

The SiO_2/Si interface quality of the NiSi gate was analyzed from the interface state densities. As shown in Fig. 7-9(a), the interface state density is 4×10^{11} cm^{-2} eV^{-1} for the 600°C silicided sample and is 8×10^{10} cm^{-2} eV^{-1} for the 500 °C silicided samples. Figure 7-9(b) shows the relative quasi-static C-V curves. No V_{fb}-shift occurs, but curve distortion is observed near Vg=-0.5 V. Particularly for the prolonged 500 °C annealed sample, the quasi-static C-V curve is almost identical to the original 500 °C silicidation curve. If Ni diffuses into SiO_2, the oxide charge will increase and C-V curve will separate from the low temperature annealed samples. Since C-V curve distortion occurs near inversion, stress-induced interface state is the main reason for damage. Moreover, the distortion tends toward the positive voltage which reflects that the interface states are negative charged. The mechanical stress is confirmed by the measurement of the thin film stress, which indicates that the 600 °C sample has a 0.96 GPa tensile stress while the 500 °C sample has a 0.75 GPa tensile stress.

To further investigate the effects of thermal budge, prolonged thermal annealing at 600°C was performed on the undoped NiSi samples. The V_{fb} and hystereses are degraded with annealing time, while they are almost unchanged within 15 min annealing, as shown in Fig. 7-10. The negative shift of V_{fb} is caused by positive charges, which is different from the negative charged surface states caused by thermal mechanical stress. In addition, the device can sustain thermal stress for 15min. Therefore, it is presumed that degradation is induced not only by the stress but also by Ni diffusion. From the thermal degradation point of view, the temperature of Ni-silicidation process should be at or below 500°C and the prolonged thermal process should be below 500 °C.

7.4 Summaries and Conclusions



We demonstrate a new and useful implant-to-silicide (ITS) method to modulate the effective work function of a NiSi gate after silicidation. Impurities, such as phosphorus and boron, pile up at the silicide/oxide interface after a 600 °C annealing. Large work function modulation can be achieved in the FUSI NiSi (50 nm)/SiO_2 gate stack by BF_2^+ and P^+ species with a implantation dosage of 1×10^{15} cm^{-2}. The work function ranges from 4.36 to 5.14 eV. Therefore, NiSi FUSI gates are a feasible single metal gate with dual work function candidate for CMOS devices with SiO_2 gate dielectrics. However, the work function of the ITS NiSi on HfO_2 sample only ranges from 4.44 to 4.62 eV. The small work function range implies that the Fermi-level pinning effect occurred on the FUSI NiSi/HfO_2 structure. For the CMOS devices with a HfO_2 gate dielectric, a continuous SiO_2-like interfacial layer between NiSi and HfO_2 can reduce the Fermi-level pinning problem, whereas the increase of effective oxide thickness should be optimized to avoid the scaling limit.

Regarding the thermal stability, although the NiSi on both SiO_2 and HfO_2 are stable up to

600 °C, thermal stress can degrade the device performance during the silicidation and post-silicide annealing. A silicidation temperature and post silicidation annealing temperature at 500 ^{o}C only causes a slight increase of interface state of less than 1×10^{11} $cm^{-2}eV^{-1}$, while the siliciadtion at 600 ^{o}C will cause interface states higher than 3×10^{11} $cm^{-2}eV^{-1}$. The thin film stress increases from 0.75 GPa to 0.96GPa as the silicidation temperature increases from 500 ^{o}C to 600 ^{o}C. The preferred temperature for Ni-silicidation is suggested at or below 500 °C and the prolonged thermal process should be below 500 °C.

References

[1] J. Kedzierski, D. Boyd, C. Cabral, Jr., P. Ronsheim, S. Zafar, P. M. Kozlowski, J. A. Ott, and M. Ieong "Threshold Voltage Control in NiSi Gated MOSFETs Through SIIS," *IEEE Trans. Electron Dev.* vol. ED-52, no. 1, pp.39-46, 2005.

[2] C. S. Park, B. J. Cho, and D. L. Kwong, "Thermally Stable Fully Silicided Hf-Silicide Metal-Gate Electrode" *IEEE Electron Device Lett.* vol. 25 no.6 pp.372-374, June 2004.

[3] J. Liu, H. C. Wen, J. P. Lu and D.-L. Kwong, "Dual-Work-Function Metal Gates by Full Silicidation of Poly-Si With Co-Ni Bi-Layer," *IEEE Electron Device Lett.* vol. 26 no.4 pp.228-230, 2005.

[4] M. J. H. van Dal, G. Pourtois, J. Cunniffe, A. Veloso, A. Lauwers, K. Maex, and J. A. Kittl, "Effect of SIIS on Work Function of Self-Align PtSi FUSI Metal-Gated Capacitors," *IEEE Trans. Electron Dev.* vol. ED-53, no. 5, pp.1180-1185, 2006.

[5] J. Kedzierski, M. Ieong, T. Kanarsky, Y. Zhang, and H.-S. P. Wong, "Fabrication of Metal Gated FinFets Through Complete Gate Silicidation with Ni," *IEEE Trans. Electron Dev.* vol. ED-53, no. 5, pp.1180-1185, 2006.

[6] C. Cabral, Jr., J. Kedzierski, E. Gusev, B. Linder, S. Zafar, M. Copel, V. Narayanan, S. Fang, P. Kozlowski, R. Carruthers, and R. Jammy, "Dual Workfunction Fully Silicided Metal Gates," in *2004 VLSI Symp. technology digest* pp.184-185.

[7] C. Zaring, H. Jiang, B. G. Svensson, and M. Ostling, "Boron redistribution during formation of nickel silicides," *Appl. Surf. Sci.*, vol. 53, pp.147–152, 1991.

[8] L. R. Zheng, L. S. Hung, and J.W. Mayer, "Redistribution of dopant arsenic during silicide formation," *J. Appl. Phys.*, vol. 58, pp. 1505–1514, 1985.

[9] J. A. Kittl, M. A. Pawlak, A. Lauwers, C. Demeurisse, K. Opsomer, K. G. Anil, C. Vrancken, M. J. H. van Dal, A. Veloso, S. Kubicek, P. Absil, K. Maex, and S. Biesemans, “Work Function of Ni Silicide Phases on HfSiON and SiO_2: NiSi, Ni_2Si, $Ni_{31}Si_{12}$ and Ni_3Si Fully Silicided Gates,” *IEEE Electron Device Lett.* vol. 27 no.1 pp.34-36, 2006.

[10] B. Y. Tsui and C.P. Lin, “Process and Characteristics of Modified Schottky Barrier (MSB) p-channel FinFETs,” *IEEE Trans. Electron Dev.* vol. ED-52, no. 12, pp.2455-2662, 2005.

[11] Y. H. Kim, C. Cabral. Jr. P. Gusev, R. Carruthers, L. Gignac, M. Gribelyuk, E. Cartier, S. Zafar, M. Copel, V. Narayanan, J. Newbury, B. Price, J. Acevedo, P. Jamison, B. Linder, W. Natzle, J. Cai, R. Jammy and M. Ieong, ” Systematic Study of Workfunction Engineering and Scavenging Effect Using NiSi Alloy FUSI Metal Gates with Advanced Gate Stacks” in *2005 IEDM Tech. Dig*., pp.657-660.

[12] P. Xuan, and J. Bokor,” Investigation of NiSi and TiSi as CMOS gate materials,” *IEEE Electron Device Lett.,* vol. 24, No. 10, pp.634-636 2003.

[13] Y. Tsuchiya, M. Yoshiko, M. Koyama, A. Kinoshita, and J. Koga,” Physical Mechanism of Work Function Modulation due to Impurity Pileup at Ni-FUSI/SiO(N) Interface” in *2005 IEDM Tech. Dig*. pp.637-640.

[14] Y. C. Yeo, P. Ranado, T.-J. King, and C. Hu,” Effects of high-κ gate dielectric materials on metal and silicon gate workfunctions,” *IEEE Electron Device Lett.,* vol.23, No. 6, pp.342-344, 2002.

[15] T. Nabatame, M. Kadoshima, K. Iwamoto, N. Mise, S. Migita, M. Ohno, H. Ota, N. Yusuda, A. Ogawa, K. Tominaga, H. Stake and A. Toriumi,” Partial silicides technology for tunable work function electrodes on high-k gate dielectrics—

Fermi-level pinning controlled PtSix for HfOx(N) pMOSFET," in *2005 IEDM Tech. Dig.*, pp. 83-86.

[16] K. SHIRAISHI1, Y. AKASAKA, S. Miyazaki, T. Nakayama, T. Nakaoka, G. Nakamura, K. Torii, H. Furutou, A. Ohta, P. Ahmet, K. Ohmori, H. Watanabe, T. Chikyow, M. L. Green, Y. Nara, and Yamada, "A trench-sidewall single-wafer-MEMS technology and its typical application in high-performance accelerometers," in *2005 IEDM Tech. Dig.*, pp.43-47.

[17] K. Shiraishi, K. Yamada, K. Torii, Y. Akasaka, K. Nakajima, M. Konno, T.Chikyow, H. Kitajima and T. Arikado, "Oxygen Vacancy Induced Substantial Threshold Voltage Shifts in the Hf-based High-*K* MISFET with p+poly-Si Gates -A Theoretical Approach," *Jpn. J. Appl. Phys.* vol.43, No. 11A, pp. L1413-L1415, 2004.

[18] S. M. Sze, *Physics of Semiconductor Device,* 2nd ed.,John Wiley & Sons, 1981, p. 21.

[19] D. K. Schroder, *Semiconductor Material and Device Characterization*, 2nd ed., John Wiley & Sons, 1998, p.365.

[20] B. Y. Tsui, and C. F. Hung," Investigation of Cu/TaN Metal Gate for Metal-Oxide-Silicon Devices," *J. Electrochem. Soc.* vol. 50, No. 1, pp.G22-G27, 2003.

[21] J. H. Sim, H.C. Wen, J.P. Lu, and D. L. Wong**,"** Work function tuning of fully silicided NiSi metal gates using a TiN capping layer**,"** *IEEE Electron Device Lett.*, vol.25, No. 9, pp. 610-612, 2004

[22] C. Cabral, Jr., J. Kedzierski, B. Linder, S. Zafar, V. Narayanan, S. Fang, A. Teegen, P. Kozlowski, R. Carrathers, and R. Jammy," Dual Workfunction Fully Silicided Metal Gates," *Proc. Symp. VLSI Tech. Dig.*, 2004, p. 184.

[23] J. Liu, H. C. Wen, J. P. Lu and D. L. Kwong," Improving gate-oxide reliability by TiN capping layer on NiSi FUSI metal gate," *IEEE Electron Device Lett.*, vol. 26 No. 7, pp. 458-490, 2005.

Table 7-1 Some conditions of the FUSI test samples

	SiO_2 sample	HfO_2 sample
Dielectric structure	Single layer of SiO_2	Stack of HfO_2/SiO_2
Dielectric Thickness (nm)	5.5 10.5 17	2/5.5 2/10.5 2/17
NiSi doping method	ITS	ITS Pre-doped
Implantation conditions	BF_2^+/1×10^{15}/10 KeV P^+/1×10^{15}/10 KeV	BF_2^+/1×10^{15}/10 KeV P^+/1×10^{15}/10 KeV

ITS: implant-to-silicide

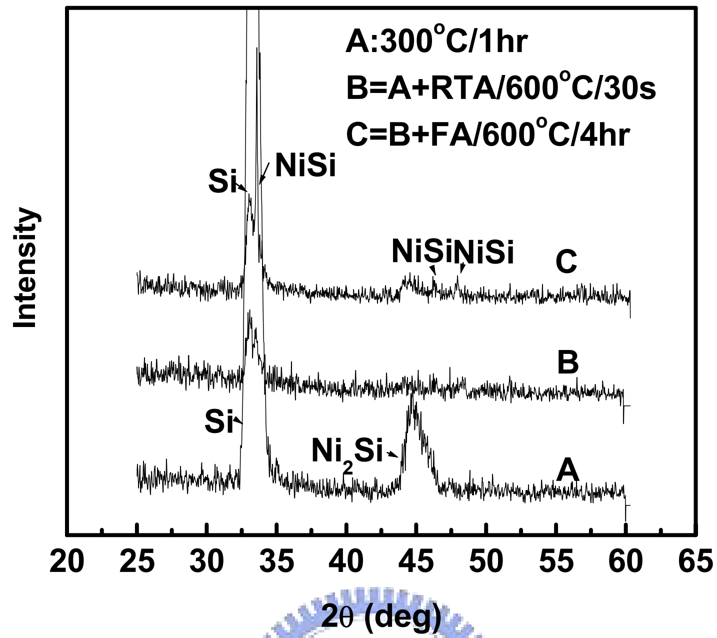


Fig. 7-1 X-ray diffraction patterns of nickel silicide. Curve A is recorded after 300 °C silicidation in a furnace (FA) for 1 hr; curve B is recorded after condition A plus 600 °C RTA for 30 sec; curve C is recorded after condition B plus 600 °C annealing for 4 hr in a furnace.

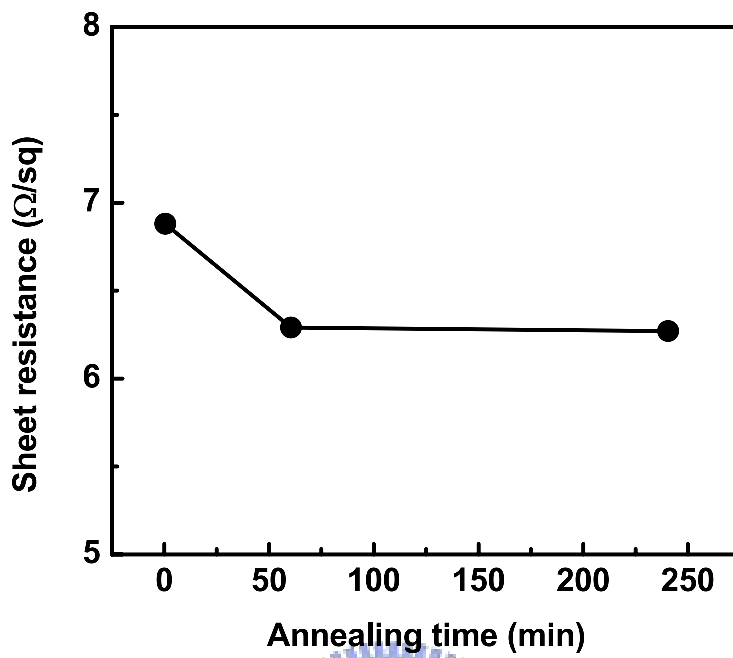


Fig. 7-2 Sheet resistance of NiSi film vs annealing time at 600 °C. The measured NiSi film is on the SiO_2 layer.

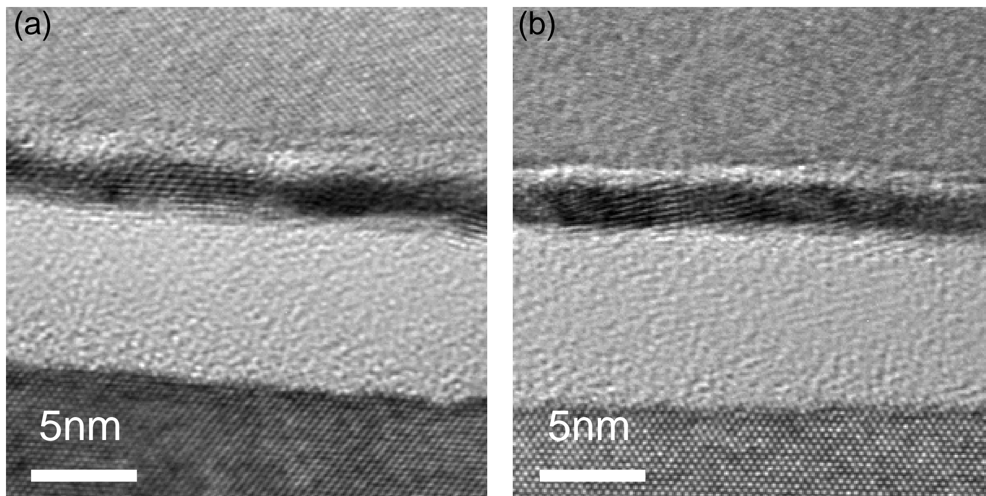


Fig. 7-3 TEM cross-sectional views of (a) ITS sample and (b) pre-doped sample.

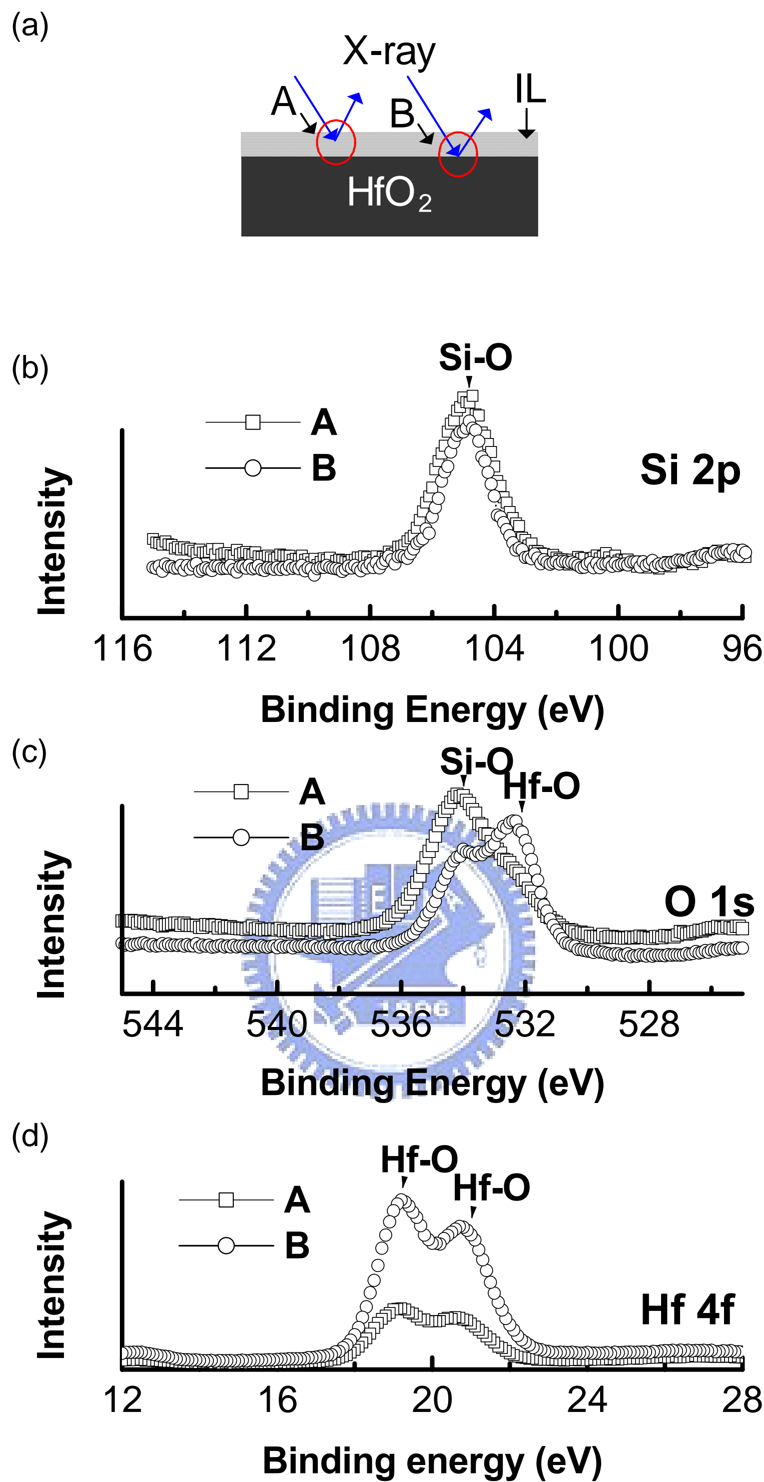


Fig. 7-4 (a) Scheme of detection by X-ray photoelectron spectrometer. Area A is exposed after removing NiSi, and B is exposed after removing the interfacial layer. The electron binding energies of detected areas A and B of (b) Si 2p orbital, (c) O 1s orbital, and (d) Hf 4f orbital are shown.

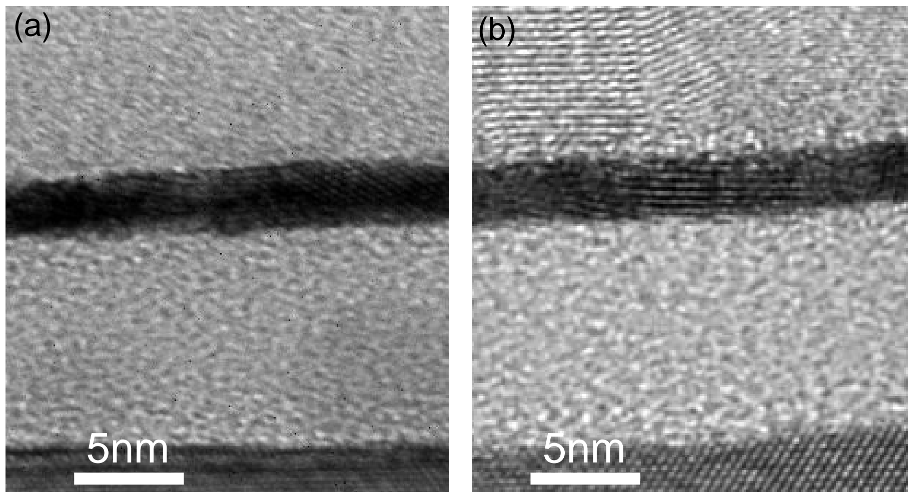
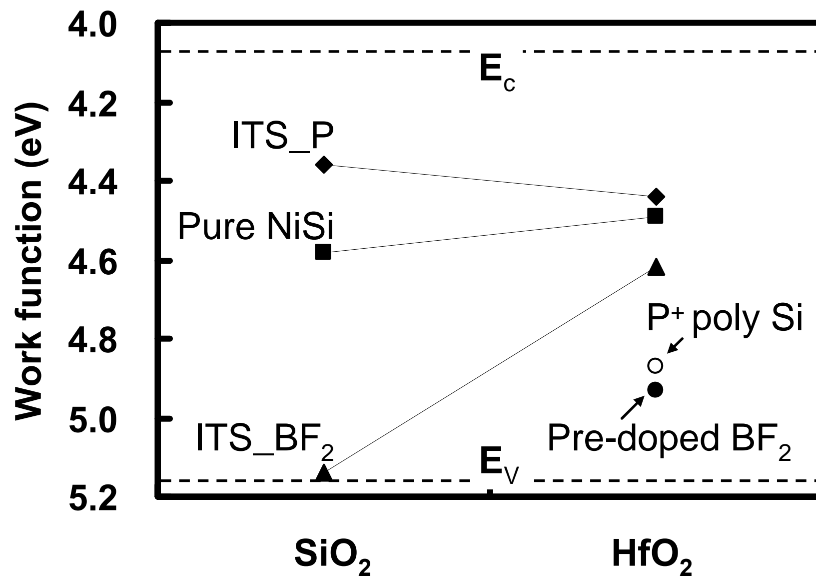


Fig. 7-5 TEM cross-sectional views of (a) as-deposited amorphous Si/ $HfO_2/SiO_2/Si$, and (b) sample (P+ poly-Si sample) implanted with BF_2 and annealed at 900^oC for 30 sec in N2 atmosphere.



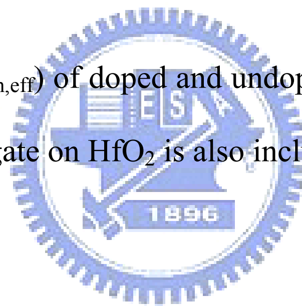


Fig. 7-6 Effective work function ($\Phi_{m,eff}$) of doped and undoped NiSi FUSI gate on SiO_2 and HfO_2. $\Phi_{m,eff}$ of the P^+ polysilicon gate on HfO_2 is also included.

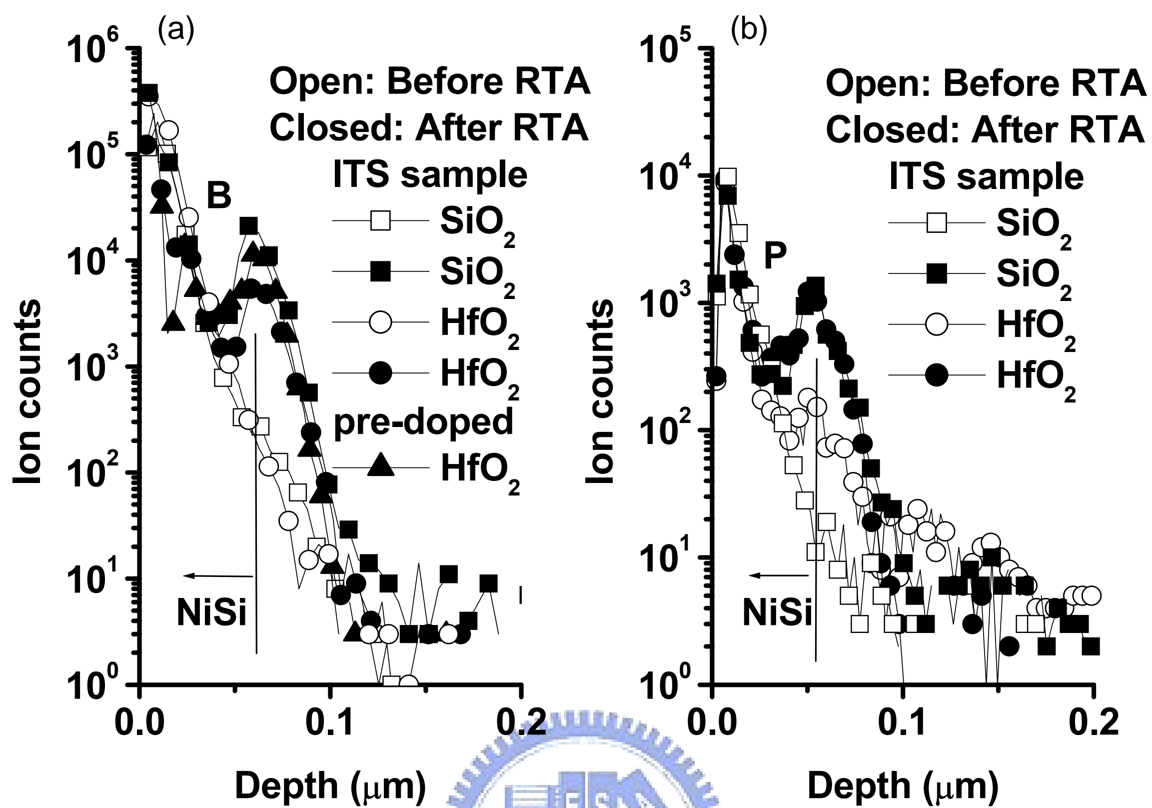


Fig. 7-7 Depth profiles of (a) boron and (b) phosphorous dopants in NiSi FUSI gates detected by secondary ion mass spectrometry.

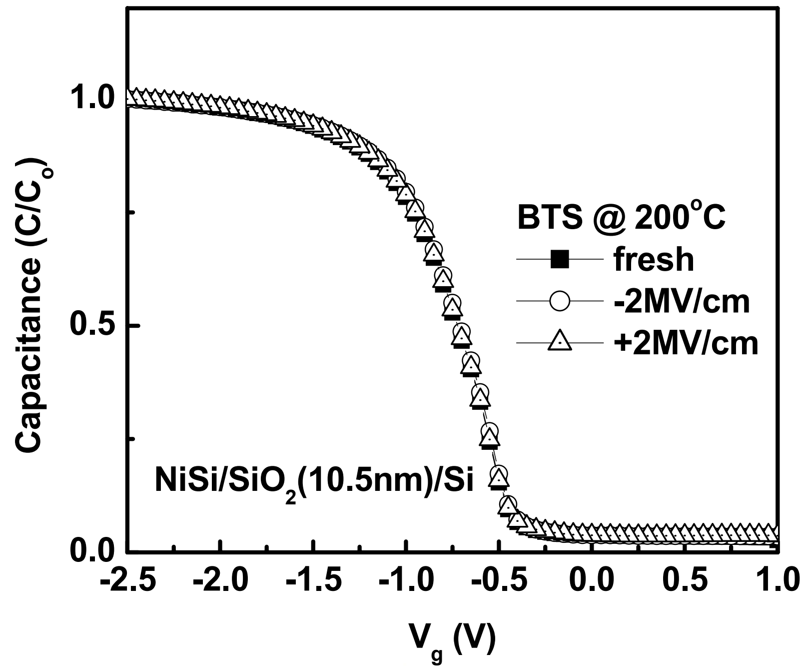


Fig. 7-8 C-V characteristics of NiSi FUSI gate capacitor with SiO_2 gate dielectric after bias-temperature stress of ±2 MV/cm at 200 °C for 30 min.

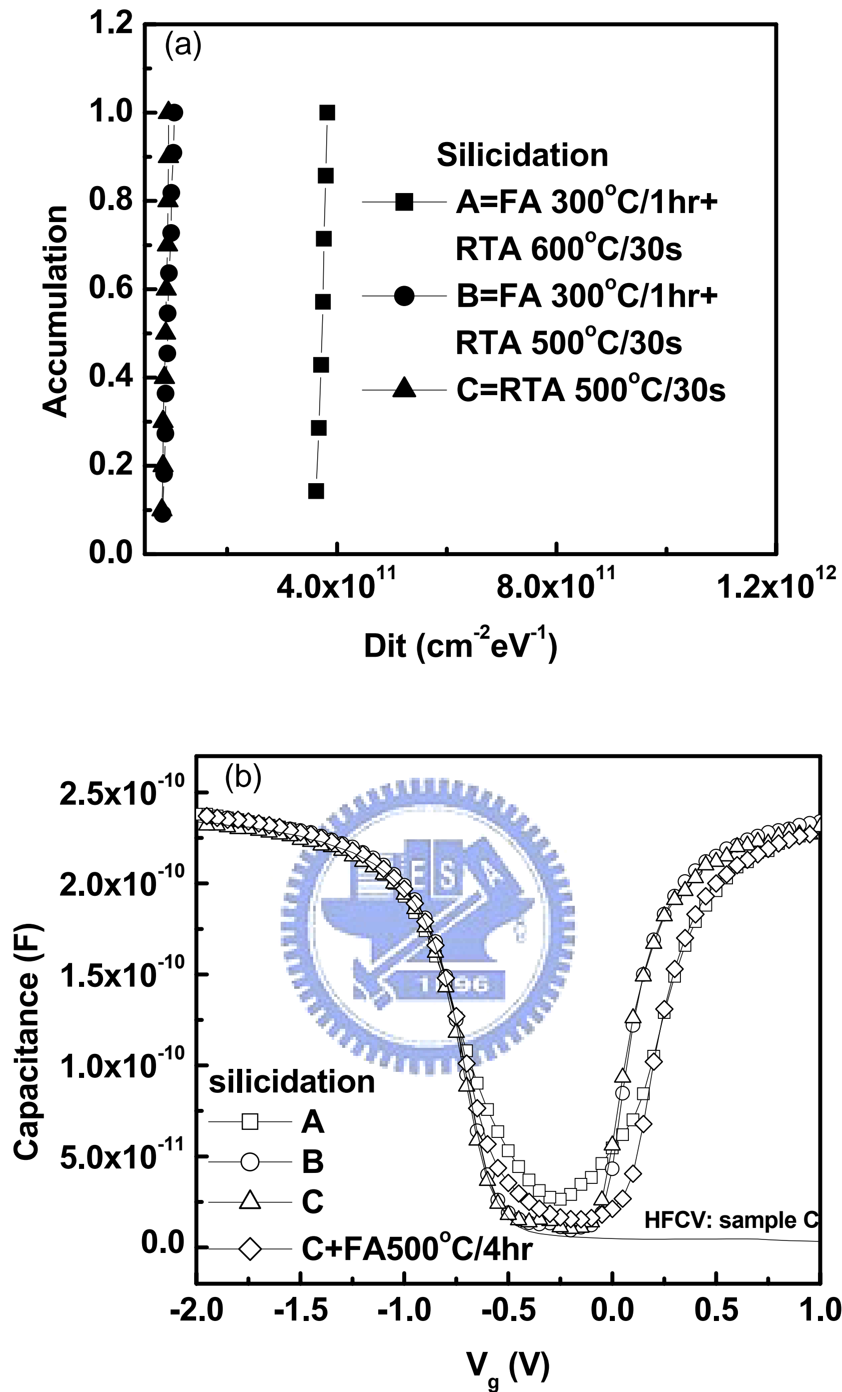


Fig. 7-9 (a) Cumulative distribution of interface state density of NiSi FUSI gate capacitors; NiSi gates were formed at 500 or 600 °C. (b) The relative quasi-C-V curves of NiSi FUSI gate capacitance and the high frequency curve of sample C.

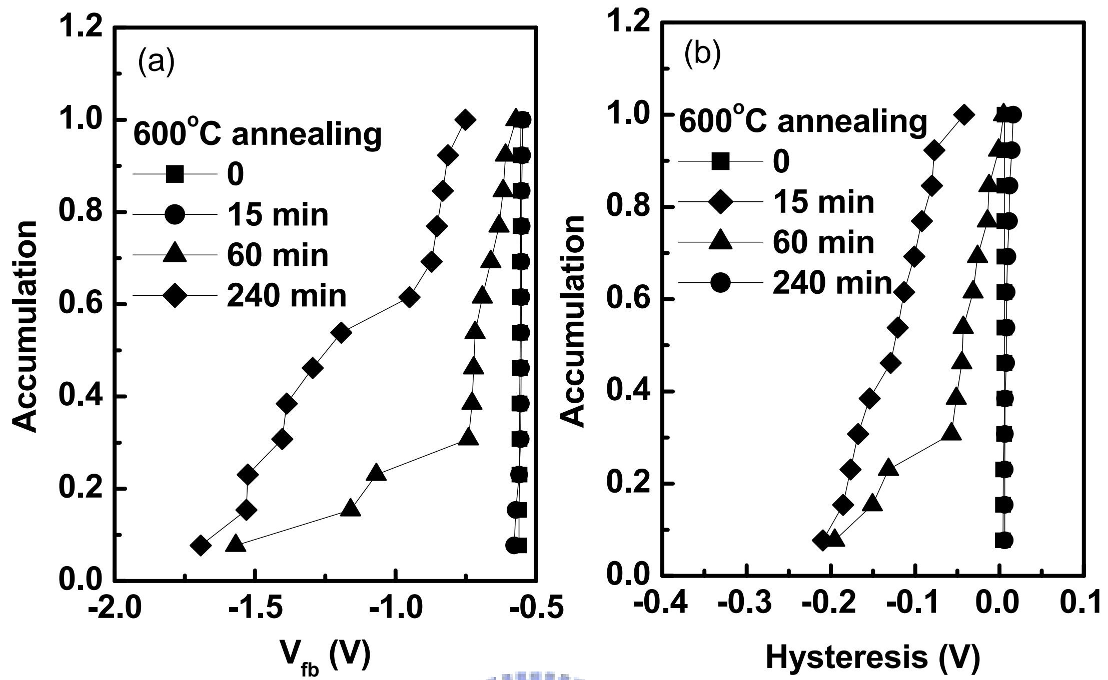


Fig. 7-10 Cumulative distribution of (a) flat-band voltage (V_{fb}) and (b) hysteresis of NiSi FUSI gate capacitors after prolonged 600 °C annealing.

Chapter 8

Conclusions and Recommendations

8.1 Conclusions

In this dissertation, several metal gate materials were studied to replace poly silicon gate for the MOSFET applications. The first study is an attempt to control the work function of the refractory metal nitrides including molybdenum nitrides, tungsten nitrides and tantalum nitrides by the incorporation of nitrogen. The second study is to attain any work function of the binary metallic alloys of which work function is located between silicon conduction band and valence band by the selection of the elements and their concentration in binary metallic alloys. The third study is focused on the fully silicided metal gates which is favorite for dual-gate CMOS device fabrication. The work function of fully silicided gate was modulated by the dopants piling-up at the interface between silicide and underlying dielectric.

In chapter 3, we study the work function modulation and thermal stability of MoN_x films on SiO_2 and HfO_2. Molybdenum nitride films with various N/W atomic ratios were deposited by the reactive sputter deposition with different N_2/Ar flow rate ratios. The nitrogen concentrations of MoN_x films analyzed by the RBS are 46, 50 and 59 at. %. The main phase of the MoN_x films is MoN (200). As the nitrogen content increases, the microstructure of MoN_x film tends to be amorphous-like and the resistivity increases. After high temperature annealing up to 800°C, the phase

remains stable and grain size increases slightly. The work function of MoN_x increases with the increase of nitrogen content and tends to saturate at the valence band of Si. No Fermi-pinning effect is observed on HfO_2 film. The work function and thermal stability of MoN_x show good thermal stability on both SiO_2 and HfO_2 films up to 800 °C at least. According to these results, MoN_x is a promising gate material for PMOSFETs. The high resistivity of the molybdenum nitride could be solved by stack structure which consists of a low resistivity top layer and a metal nitride bottom layer. The top layer is used to reduce the sheet resistance and the bottom layer is used to control the threshold voltage.

In chapter 4, we study the work function modulation and thermal stability of WN_x films on SiO_2 and HfO_2. Tungsten nitride films with various N/W atomic ratios were deposited by the reactive sputter deposition with different N_2/Ar flow rate ratios. Nitrogen concentration in WN_x films increases rapidly with the N_2/Ar gas flow ratio and tends to saturate. The nitrogen concentrations of MoN_x films analyzed by the RBS are 44, 56 and 61 at. %. WN_x films with nitrogen content higher than 44% atomic ratio has a main phase of WN, and the WN phase is stable up to 800°C. The higher order WN_x phase does not form even if the nitrogen concentration is as high as 61 at. %. Tungsten nitride has an effective work function near the valance band of silicon so that it satisfies the metal-gate demand of p-type MOSFETs to provide suitable threshold voltage. However, the excess nitrogen will make the work function unstable during the high temperature annealing, and may cause the WN_x film crack during the high temperature annealing with a rapid temperature rising rate. The post-metal thermal annealing must be carefully controlled to avoid bubbling and cracking. The excess nitrogen in WN_x films can cause the effective work function lowering as the N content is higher than 50 at. %. Weak Fermi-level pinning effect is

observed on HfO_2 film. In this case, WN_x is not suitable to be meal gate of bulk p-type MOSFTEs. Fully depleted SOI devices require work function lower than Si valence band for pMOSFET and higher than Si conduction band for nMOSFET. Therefore, WN_x/HfO_2 gate stack can be applied to p-type FD-SOI devices. The good integrity of the WN_x/HfO_2 gate stack also suggests WN_x as a promising gate material. The high resistivity of the tungsten nitride could be solved by stack structure which consists of a low resistivity top layer and a tungsten nitride bottom layer.

In chapter 5, we studied the work function modulation of TaN_x film and the thermal stability of Cu/TaN_x stack as the gate electrode of which the Cu layer severs as the conduction layer to lower the resistance of the gate stack. Tantalum nitride films with various N/W atomic ratios were deposited by the reactive sputter deposition with different N_2/Ar flow rate ratios. The nitrogen concentrations of MoN_x films analyzed by the RBS are 23, 33 and 39 at. %. The main phases of the TaN_x films range from Ta_2N, TaN, and Ta_3N_5, respectively. The TaN_x films are thermally stable up to 800 ℃. However, the formation of the Ta_3N_5 phase in a TaN_x film annealed at high temperature or with a high N/Ta ratio increases the effective resistivity. The effective work function of TaN_x is about 4.31-4.38eV and the range is less than 70 meV. Such a weak work function modulation implies that the work function of TaN_x film is quite stable to avoid the variation induced by process disturbance. On the other hand, TaN_x is suitable to be a gate electrode only for the NMOSFETs. The flat band voltage decreases with an increase in the annealing temperature. In addition, the deviation of the flat band voltage increases with the annealing temperature. Although phase change, grain growth and Cu contamination contribute at high temperature, thermal stress-induced oxide charges dominate the decrease and deviation of the flat band voltage at temperature below 500 ℃. The

Cu/TaN_x stack can be used as a gate electrode for the surface channel NMOSFETs, and the maximum process temperature following gate electrode deposition should be lower than 500 ℃. The thermal stress-induced oxide charges are additional sources of deviation in the threshold voltage. This result must be considered in controlling the threshold voltage during metal gate generation.

In chapter 6, we studied the work function modulation of Ta-based binary alloys and the physical properties of Ta-Pt alloys. A principle of work function modulation is first demonstrated by the binary metallic alloys. Any work function can be obtained by the binary metallic alloys with a proper composition and concentration in the alloy. The work function of the Ta-based metal alloys (Ta-Ti and Ta-Pt) can be modulated from 4.16 eV to 5.2 eV continuously. Such a wide range work function modulation makes them suitable for use in CMOS applications. The thermal stability of alloy is a main concern for the gate first process. IETS is used to provide a direct inspection of thermal instability at interface of Ta-Pt/SiO_2/Si and Ta-Pt/HfO_2/Si structures. The analysis of results provides a direct evidence that the Fermi-level pinning of Ta-Pt metal gates on SiO_2 is counted for the generation of extrinsic states due to the formation of the tantalum oxide. The Ta-Pt/HfO_2 interface is more thermally stable than the HfO_2/Si interface. The work function of Ta-Pt alloys would not be changed by the incorporation of the implantation impurities such as arsenic ions, boron ions and phosphorus ions with a concentration of more than 8×10^{20} cm^{-3}. Therefore, the control of threshold voltage of the alloy-gate MOSFET could be easier. However, the boron impurity still can easily diffuse in the Ta-Pt alloy gates. To prevent the boron penetration of the Ta-Pt alloy gate, a hard mask and low thermal budget activation process, such spike annealing, are suggested to be used during implantation and activation. The Ta-Pt alloys with tunable work function are

potential gate materials for SOI CMOS if the underlying gate dielectric is Hf-based high-k dielectric. It could be integrated in the conventional MOSFET process with high-K material (HfO_2).

In chapter 7, we studied the work function modulation and thermal stability of fully Ni-silicided gates on SiO_2 and HfO_2. A new, useful implant-to-silicide (ITS) method is used to modulate the effective work function of a NiSi gate after silicidation. Impurities, such as phosphorus and boron, pile up at the silicide/oxide interface after a 600°C annealing. Large work function modulation can be achieved in the FUSI NiSi (50 nm)/SiO_2 gate stack by BF_2^+ and P^+ species with a implantation dosage of 1×10^{15} cm^{-2}. The work function ranges from 4.36 to 5.14 eV. Therefore, NiSi FUSI gates are a feasible single metal gate with dual work function candidate for CMOS devices with SiO_2 gate dielectrics. However, the work function of the ITS NiSi on HfO_2 sample only ranges from 4.44 to 4.62 eV. The small work function range implies that the Fermi-level pinning effect occurred on the FUSI NiSi/HfO_2 structure. For the CMOS devices with a HfO_2 gate dielectric, a continuous SiO_2-like interfacial layer between NiSi and HfO_2 can reduce the Fermi-level pinning problem, whereas the increase of effective oxide thickness should be optimized to avoid the scaling limit. Regarding the thermal stability, although the NiSi on both SiO_2 and HfO_2 are stable up to 600 °C, thermal stress can degrade the device performance during the silicidation and post-silicide annealing. A silicidation temperature and post silicidation annealing temperature at 500 °C only causes a slight increase of interface state of less than 1×10^{11} $cm^{-2}eV^{-1}$, while the siliciadtion at 600 °C will cause interface states higher than 3×10^{11} $cm^{-2}eV^{-1}$. The thin film stress increases from 0.75 GPa to 0.96 GPa as the silicidation temperature increases from 500 °C to 600 °C. The preferred temperature for Ni-silicidation is suggested at or below 500 °C and the

prolonged thermal process should be below 500 °C.

We have studied several metal compounds such as tungsten nitrides, molybdenum nitrides, tantalum nitrides and nickel silicides. A plot of work function relative to the metal elements and its compounds, nitrides and silicides, is shown in Fig. 8-1. A rule of work function adjustment by the added nonmetallic elements is figured out. If we add in higher electronegativity elements, the element will attract electrons from the metal elements and cause the work function of compound higher than that of the metal element. On the other hand, if we add the low electronegativeity elements within metal elements, we can get low work function compounds. For example, tungsten has an electronegativity of 2.36. Nitrogen has an electronegativity of 3.04. As tungsten is incorporated with nitrogen to form tungsten nitride, the work function changes from 4.5 eV to 5 eV. We conclude that to change to high work function of metal elements is to add higher electronnegativity elements. To add lower electronegativeity elements in compounds is to get lower work function.

In summary, the common materials for the metal gates, metal nitrides, metallic alloys, and silicides, are all investigated. The comparisons of these metal gates are listed in table 8-1. Metal nitrides have better thermal stability while the small work function range of single metal nitrides restricts that the metal nitrides can only been used for the dual metal scheme. Although the metallic alloys reveal wide range of work function, the integration of different composition of metallic alloys in CMOS application is difficult by the single metal process. Fully silicided gates doped by the N-type and P-type impurities are the most promising process for the single metal dual work function scheme. However, the interaction between poly-Si and high-k materials before silicidation induces Fermi-Level pinning and will retard the

application of fully silicided gates. On contrary, dual metal gates can provide a suitable work function for CMOS but its process will be more complicated. There are still no perfect choice of metal gate materials and processes. Up to date, the standard metal gates for the CMOS are still under consideration to the dual metal scheme or fully silicided gates.

8.2 Recommendation for Future Works

Although lots of works and efforts have been devoted to the metal gate technology, there are still significant rooms to reach the goal of mass production. As mentioned in chapter 1, metal gates should satisfy several criteria before replacement of poly-Si gates. Lots of gate materials such as TaSiN, TaN, Ta-rich Ta-Ru alloy, HfSi and TaC are suggested for the NMOSFETs. Materials such as WN, MoN, Mo, TiN, and Ta-Pt alloys are suggested for PMOSFETs. However, the gate materials are still not decided for both NMOSFETs and PMOSFETs. There are several reasons. First, all these metal gates should be integrated by the dual metal gate process (DMG) which is complicated and high cost. There are literatures demonstrating a conventional gate first DMG process and a dual high-k and dual metal gate (DHDMG) process [1, 2]. The former can not prevent the gate dielectric from the chemical attack during removal of the first metal gate and the intermixing of the first metal and second metal possibly occur. The later can not prevent the channel surface from chemical attack during the first high-k layer removal. To think over the dual gate process, it seems that we can not prevent the gate dielectric or channel from the chemical attack which possibly cause low mobility and poor reliability. The way to minimize the effect should focus on the innovation of the selectively chemical removal process of high-k layer and metal gate layer.

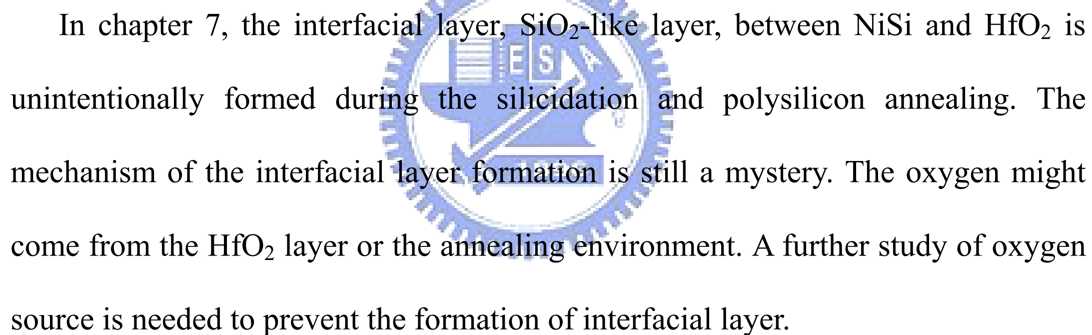
Second, the work function of metal gates is thermally unstable due to the interfacial reaction between metal gate and the underlying dielectric. As discussed in chapter 6, the Ta-Pt alloy is unstable on SiO_2, and its work function is changed due to the formation of tantalum oxide at interface, while Ta-Pt is stable on HfO_2. The instability of work function is strongly dependent on the underlying dielectric at high temperature. In addition, the Fermi-level level pinning effect coming from the metal induced gap states and virtual gap states will pin the effect the effective toward the mid-gap of silicon energy band. As we cannot make sure the gate dielectric for the next generation device, we can not guarantee the materials of metal gates. Regardless of the gate dielectric, the study of interfacial material characteristic at the interface between metal gate and gate dielectric is helpful to understand work function stability and to improve its thermal stability. To understand the season why work function is dependent on the gate dielectric, we should know interfacial product caused by the interfacial chemical reaction. To understand how to eliminate Fermi-level pinning, we should know what can cause the gap states at the interface to disappear. As discussed in chapter 3, MoN_x is free of Fermi-level pinning. In chapter 4, WN_x is slightly suffered from Fermi-level pinning. The way to reveal the root-cause of Fermi-level pining is to identify the material (chemical) characteristic at the interface between metal gate and underlying dielectric. The identification of interfacial material (chemical) characteristics relies on the development of new material analytic tools. The typical analysis techniques such as Fourier transform infrared spectroscopy, X-ray photoelectron spectroscopy, secondary ion mass spectroscopy, and electron spin resonance spectroscopy may not be suitable for interface analysis on a real device structure. As mentioned in the chapter 6, the polarity-dependent inelastic electron

tunneling spectroscopy is a possible choice to identify the interfacial reaction. On the other hand, the material characteristic of interface between gate electrodes and the underlying dielectric plays a critical role to determine effective work function. The understanding of interface characteristics can apply surface treatment to control the effective function before gate electrode deposition.

In chapter 7, fully silicided gates provide a low cost dual work function and single metal gate process. The effective work function is modulated by the dopants piling up at the interface between silicide gates and underlying dielectric. The work function of phosphorus doped NiSi sample do not provide suitable work function for NMOSFETS. Some literature suggests using the antimony (Sb) [3]. The work function modulation caused by the dopants pile-up is believed by the generation of interface dipoles which will effect the equilibrium of chemical potential (Fermi level) of the NiSi gates at the interface. However, change of effective work function caused by the same kinds of piling-up dopants is quite different in the previous literatures. It is reported that the work function modulation is dependent on not only the concentration but also the underlying dielectric as well as the kinds of elements. There are still some works to clearly understand somehow interface dipoles are generated by the piling up dopant. A first principle simulation proposed by Mark J. H. van Dal have evaluate the interface dipole and work function offset caused by the dopants piling-up at the interface between PtSi and SiO_2 with a surface concentration of 5×10^{14} cm^{-2} [4]. A further atomic-scale measurement of the surface potential profile near the interface is needed to evidence the band offset due to the dipoles caused by impurities.

In chapter 7, Fermi-level pinning occurs on the HfO_2 samples to induce that the work function of NiSi is not proper for the CMOS application. The usage of SiO_2 to

screen the effect of Fermi-level pinning can limit the device scaling due to its low permittivity. An effective substitution with higher k value dielectric will be needed to replace SiO_2. It is suggested to use the plasma treatments such as NH_3 plasma, N_2O plasma, and SiH_4 plasma which possibly do not increase the physical thickness of gate dielectric before the gate electrode deposition. The selection of category of plasma treatments should rely on the further study of the surface chemical reaction. One the other hand, somehow, we can use two type of capping layers to pin the Fermi level to the Si conduction band and Si valence band if we can overcome the process integration. A lanthanum oxide layer has demonstrated to perform a pining level near Si conduction band and the Al_2O_3 is performed to pin to the Si valence band [5, 6].



In chapter 7, the interfacial layer, SiO_2-like layer, between NiSi and HfO_2 is unintentionally formed during the silicidation and polysilicon annealing. The mechanism of the interfacial layer formation is still a mystery. The oxygen might come from the HfO_2 layer or the annealing environment. A further study of oxygen source is needed to prevent the formation of interfacial layer.

A significant amount of efforts remains to be devoted on the metal gate technology. We have point out the possible future directions for this work. Emphasis should be placed on understanding the physics of how the effective work function is determined by interface between metal gates and underlying high-k dielectric, obtaining a better understanding of the interfacial-layer formation between metal gates and underlying dielectrics, and finally achieving a standard process integration of dual metal gates CMOS.

References

[1] S. B. Samavedam, L. B. La, J. Smith, S. Dakshina-Murthy, E. Luckowski, J. Schaeffer, M. Zavala, R. Martin, V. Dhandapani, D. Triyoso, H. H. Tseng, P. J. Tobin, D. C. Gilmer, C. Hobbs, W. J. Taylor, J. M. Grant, R. I. Hegde, J. Mogab, C. Thomas, P. Abramowitz, M. Moosa, J. Conner, J. Jiang, V. Arunachalam, M. Sadd, B.-Y. Nguyen, and B. White, "Dual Metal Gate CMOS with HfO_2 Gate Dielectric," in *2002 IEDM Tech. Dig*., pp. 433-436.

[2] S. Song, Z. Zhang, M. Hussain, C. Huffman, J. Barnett, S. Bae, H. Li, P. Majhi, C. Park, B. Ju, H. Park, C. Kang, R. Choi, P. Zeitzoff, H. Tseng, B. Lee and R. Jammy, " Highly Manufacturable 45nm LSTP CMOSFETs Using Novel Dual High-k and Dual Metal Gate CMOS Integration," in *2000 Proc. Symp. VLSI Tech.*, T2-4.

[3] J. Kedzierski, D. Boyd, C. Cabral, Jr., P. Ronsheim, S. Zafar, P. M. Kozlowski, J. A. Ott, and M. Ieong, "Threshold voltage control in NiSi-gated MOSFETs through SIIS," *IEEE Trans. Electron Devices,* vol. 52, no. 1, pp. 39-46, 2005.

[4] M. J. H. van Dal, G.. Pourtois, J. Cunniffe, A. Veloso, A. Lauwers, K. Maex, J. A. Kittl, "Effect of SIIS on Work Function of Self-Alignd PtSi FUSI Metal-Gated Capacitors," *IEEE Trans. Electron Devices,* vol. 53, no. 5, pp. 1180-1185, 2006.

[5] C. C. Hobbs, L. R. C. Fonseca, A. Knizhnik, V. Dhandapani, S. B. Samavedam, W.J. Taylor, J.M. Grant, L. G. Dip, D. H. Triyoso, R. I. Hegde, D.C. Gilmer, R. Garcia, D. Roan, M.L. Lovejoy, R.S. Rai, E. A. Hebert, H.-H. Tseng, S. G. H. Anderson, B.E. white and P. J. Tobin,"Fermi-Level Pinning at the Ploysilicon/Metal-oxide Interface- Part II," *IEEE Trans. Electron Dev.* vol. ED-51, no. 6, pp.978-984, 2004.

[6] H. Alshareef, H. Harris, H. Wen, C. Park, C. Huffman, K. Choi, H. Luan, P. Majhi, B. Lee,R. Jammy, D. Lichtenwalner, J. Jur and A. Kingon, " Thermally Stable N-Metal Gate MOSFETs Using La-Incorporated HfSiO Dielectric," in *2000 Proc. Symp. VLSI Tech.*, T2-1.

Table 8-1 Comparison of metal gates.

	$\Phi_{m,eff}$	Thermal stability	ρ	FLP	CMOS Process
MoN_x	☹	☺	☹	☺	☹
WN_x	☹	☺	☹	😐	☹
TaN_x	☹	☺	☹	😐	☹
Ta-Pt alloys	☺	😐	😐	☹	☹
Ni-FUSI gates	☺	😐	☺	☹	☺

$\Phi_{m,eff}$: effective work function; ρ : resistivity; FLP: Fermi-level pinning

Better: ☺; good: 😐; poor: ☹

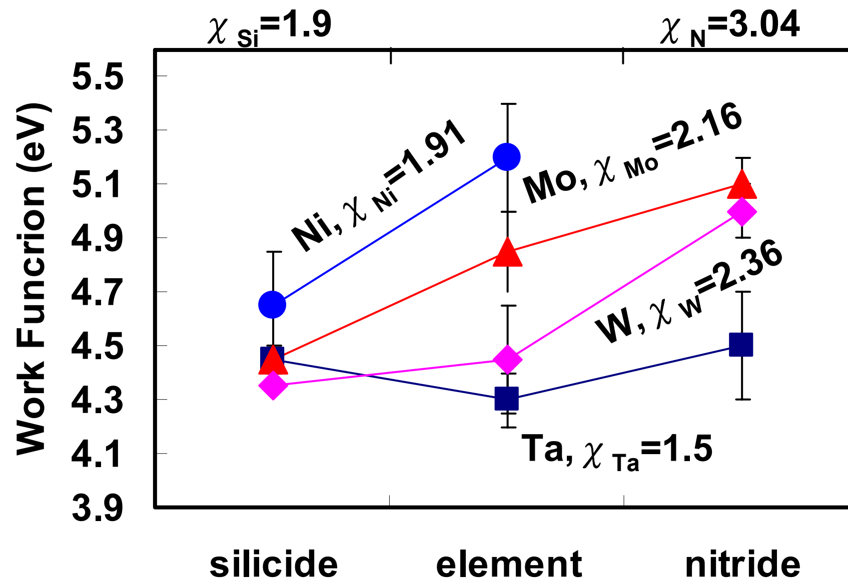


Fig. 8-1 Work function of Ta, W, Mo, Ni, their silicides and nitrides. The χ is the electronegativity.