Liverpool John Moores University

SCHOOL OF ENGINEERING

CHARACTERIZATION OF HIGH-K LAYERS AS THE GATE DIELECTRIC FOR MOSFETS

MOHAMMED B ZAHID

A thesis submitted in partial fulfillment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

This research programme was carried out in collaboration with the Interuniversity Micro-Electronics Center (IMEC)

DECEMBER 2006

ACKNOWLEDGMENTS

Here I am writing the last page of my thesis. I know that trying to explicitly thank everybody who helped me, directly or indirectly to accomplish this work during these 3 years is by far not simple. So this is the moment and the part where I would like to express my gratitude to everyone which was involved.

I would like to acknowledge my supervisor, Prof. J. F. Zhang for all the help and continuous guidance, advice, encouragement and fruitful discussions throughout the project and for the opportunity he gave me to do my PhD project at Liverpool John Moores University and in Collaboration with Interuniversity Micro-Electronic Center (IMEC) in Belgium.

I also deeply appreciate Dr. C. Z. Zhao for his advice and support throughout the project. Many thanks to other members of our Microelectronic research group, Mr. Mo Huai Chang, Mr. Y. G. Wang for their discussions during the time that I worked on my project at Liverpool John Moore University.

I would like to acknowledge Prof. Guido Groeseneken for giving me the opportunity to carry out half of my PhD project at IMEC and my supervisor, Dr. Robin Degraeve, for all the help and continuous guidance and fruitful discussions throughout the project during my stay at IMEC. Many thanks to Florence, Thomas, Luigi, Lionel, Marc, Guillaume, Etienne, Bastien, Rachid, Adelina, Maria, Barry, Sahar, Christian, Rozbeh and Laurent. Without them, my stay at IMEC wouldn't have been so wonderful. Working with all of them has entirely been, and will always be, my pleasure and privilege.

Many thanks to Yves, Michael, Muddassar, Sabri, Mohammed, Issac, Rumbi, Rizwana, Lokesh, Mathiew, Reinaldo, Heera, Surita, Yasmeen for the good time and laugh we had together in Liverpool.

Last but not the least, I owe deep gratitude to my parents and to all my family, for their love and support, for their patience with me and for the education I received during all these years.

Zahid Mohammed B,

Liverpool, December, 2006

The best way to predict the future is to invent it.

Alan Kay

You have an idea today, a better tomorrow, but the best of all ...never

Sir Robert Watson Watt

ABSTRACT

As the gate oxide thickness of the Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) is continuously scaled down with lateral device dimensions, the gate leakage current during operation increases exponentially. This increase in leakage current raises concerns regarding power consumption and device reliability. Alternative dielectrics with higher dielectric constant (high-k) than that of SiO_2 have been searched. High-k layers allow the use of physically thicker gate dielectrics, so that the gate leakage current is controlled. The intensive world-wide research has identified the Hf-dielectric as the lead candidate for future CMOS technologies. However, the commercial application of Hf-dielectrics as the gate oxide has been held back by a number of issues, including process integration, low carrier mobility, and high instability.

This project focuses on characterizing the defect responsible for the instability of Hf-dielectrics. The thesis consists of six chapters. After an introduction in Chapter 1, the characterization techniques used are described in Chapter 2. Two main contributions are: setting up the pulse transfer characteristic technique and developing a newly improved charge pumping technique called Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP).

The research results are presented in Chapters 3, 4 and 5. Chapter 3 characterizes as-grown electron traps in HfO_2/SiO_2 stacks. The issues addressed include the impact of measurement technique on electron trapping, contribution of different current components to trapping, trap location, and the capture cross section and trapping kinetics. It is shown that the use of pulse transfer characteristic technique is essential for measuring electron trapping, since the traditional quasi-dc transfer characteristic is too

slow and the loss of charges is significant. The trap assisted tunneling and the thermally enhanced conduction contributes little to trapping. The trapping does not pile up at the interfaces and the region near to one or both ends of HfO_2 has little trapping, when compared with the trapping in the bulk. To evaluate the electron fluency through the gate stack, efforts are made to estimate the trapping-induced transient gate current through simulation. This allows the determination of two capture cross sections: one in the order of $10^{-14}cm^2$ and the other in the order of $10^{-16}cm^2$.

Chapter 4 concentrates on the characterization of generated electron traps and the time dependent dielectric breakdown (TDDB). Amplitude charge pumping and frequency sweep charge pumping are used to investigate the impact of gate electrodes and channel length on charging and discharging of the bulk defects. As channel length increases, it is found that bulk trapping increases and TDDB time shortens. Efforts are made to show that there is a quantitative correlation between the trapping and TDDB data. The newly improved VT^2CP is used to separate trapping in the interfacial SiO_2 from that in HfO_2. The results show that new traps are generated in both layers and the generation follows a power law with similar power factors. Investigation is also carried out to assess the dependence of trap generation on process and deposition conditions. Finally, it is found that Hf-dielectric with metal gate always suffers hard-breakdown.

In Chapter 5, attention is turned to positive charging in Hf-dielectric. It is shown that the use of metal gate enhances the positive charging, when stressed under a positive gate bias. This is explained by assuming that there is a large number of hydrogenous species within the metal gate or at its interface with gate dielectric. Two types of threshold voltage instabilities have been identified for pMOSFETs. The first one results in a loop in the transfer characteristics when a pulse is applied to the gate. The second one is caused by the generation of new positive charge. Both are enhanced by

nitridation. For sub-2nm Hf-dielectric, the threshold voltage instability of pMOSFETs can be more severe than that of nMOSFETs and it can be a limiting factor for the operation voltage.

Finally, the project is summarized in Chapter 6 and the future work is discussed.

LIST OF PUBLICATIONS

CONFERENCES PAPERS

1. **M.B.Zahid**, L.Pantisano, R.Degraeve, M.Aoulaiche, I.Ferain, L.Trojman, E.San Andres, S.Severi, M.Toledano, B.O'Sullivan S.De Gendt, M.Heyns, G.Groeseneken, "Optimization of hole trapping in HfSiON causing V_T-instability in 1nm EOT pMOS," **Accepted and to be Presented at VLSI 2007**

2. **M.B.Zahid**, R. Degraeve, L. Pantisano, J. F. Zhang, G. Groeseneken, "Defects Generation in SiO_2/HfO_2 Studied With Variable t_{charge} - $t_{discharge}$ Charge Pumping (VT^2CP)," **Accepted and to be Presented at IRPS 2007**

3. **M.B.Zahid**, R. Degraeve, J. F. Zhang, G. Groeseneken," Impact of Process Conditions on Interface and High-κ Trap Density Studied by Variable Tcharge-Tdischarge Charge Pumping (VT^2CP)," **Accepted and to be Presented at INFOS 2007**

4. **M.B.Zahid,** R.Degraeve, T.Kauerauf, G.Groeseneken, J.F.Zhang, "Investigation of Channel-length dependent Time-to-Breakdown (t_{BD}) with variable frequency charge pumping," **in 36th IEEE SISC, Arlington, USA, November 2005**

5. C.Z.Zhao, **M.B.Zahid**, J. F. Zhang, G. Groeseneken, R. Degraeve, and S. De Gendt, "Properties and dynamic behavior of electron traps in HfO_2/SiO_2 stacks," **Microelectronic Eng., vol.80, pp.366-369, 2005**.

6. M. Toledano-Luque, L. Pantisano, R. Degraeve, **M.B.Zahid**, I. Ferrain, E. San Andrés, G. Groeseneken, S. De Gent," Charge Pumping HfSiON defect spectroscopy after Substrate Hot Electron Injection," **Accepted and to be Presented at INFOS 2007**

7. R. Degraeve, T. Kauerauf, M. Cho, **M. B. Zahid,** L-Å. Ragnarsson, D.P. Brunco, B. Kaczer, Ph. Roussel, S De Gendt, G. Groeseneken, "Degradation and breakdown of 0.9 nm EOT SiO_2/ALD HfO_2/metal gate stacks under positive Constant Voltage Stress", **IEDM Technical Digest, pp. 419-422, 2005**

JOURNAL PAPERS

1. C. Z. Zhao, **M.B.Zahid,** and J. F. Zhang," Threshold voltage instability of pMOSFETs with Hf-based dielectrics," **Applied Physics Letters 2006, Accepted**

2. Z. Zhao, J. F. Zhang, **M.B.Zahid,** B. Govoreanu, G. Groeseneken, and S. De Gendt," Determination of capture cross sections for as-grown electron traps in HfO_2/HfSiO stacks", **J. Appl. Phys. 100, 093716 (2006)**

3. J.F.Zhang, C.Z.Zhao, **M.B.Zahid,** G. Groseneken, R. Degraeve, and S. De Gendt," An assessment of the location of as-grown electron traps in HfO_2 /HfSiO stacks", **IEEE Electron Dev. Lett., 27:1010, 817-820, 2006.**

4. T. Kauerauf, R. Degraeve, **M.B.Zahid,** M. Cho, B. Kaczer, Ph. Roussel, G. Groeseneken, H. Maes and S. De Gendt," Abrupt breakdown in dielectric/metal gate stacks: A potential reliability limitation?," **Electron Devices Letters Vol.26, Issue:10, pp 773- 775 , 2005**

5. C.Z.Zhao, J.F.Zhang, and **M.B.Zahid,** "Impact of gate materials on positive charge formation in HfO_2/SiO_2 stacks," **Appl. Phys. Lett. 89, 023507 , 2006**

6. M. Toledano-Luque, L. Pantisano, R. Degraeve, **M.B.Zahid**, E. San Andrés, G. Groeseneken, S. De Gent," New Developments In Charge Pumping Measurements On Thin Stacked Dielectrics," **In Preparation and to be submitted to TED**

LIST OF ABBREVIATIONS

Abbreviations	*Signification*
ALCVD	Atomic Layer Chemical Vapor Deposition
ALD	Atomic Layer Deposition
APC	Anomalous Positive charges
BD	BreakDown
CP	Charge Pumping
CVS	Constant Voltage Stress
DPN	Decoupled Plasma Nitridation
EOT	Equivalent Oxide Thickness
FGA	Forming Gas Anneal
FUSI	Fully Silicide
HBD	Hard BreakDown
HfO_2	Hafnium dioxide
HfSiON	Hafnium Silicate Oxide Nitride
IL	Interfacial Layer
MOCVD	Metal Organic Chemical Vapor Deposition
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
NBTI	Negative Bias Temperature Instability
N_2	Nitrogen
NH_3	Ammonia (Nitrogen bounded to hydrogen)
O_2	Dioxygen
PDA	Post Deposition Anneal
PVD	Physical Vapor Deposition
SBD	Soft BreakDown
SiO_2	Silicon Oxide
SILC	Stress Induced Leakage Current
SiON	Silicon Oxide Nitride
TaN	Tantalum Nitride
TDDB	Time Dependent Dielectric Breakdown
TiN	Titanium Nitride
WKB	Wentzel-Kramers-Brillouin
ZrO_2	Zirconium dioxide

LIST OF SYMBOLS

Symbols	*Description*	*Unit*
Δ	Differential operator	
σ	Capture cross section	
β	Beta	
μ	Micro	
$\hbar$	h is the Plank's constant	$\hbar=h/(2\pi)$
ε_{SiO2}	Dielectric Constant of SiO_2	t
ε_o	Electric permittivity of vacuum	F/cm
ε_{iL}	Dielectric constant of the interfacial layer	
ε_{HfO}	Dielectric constant of HfO_2	
ψ_{ms}	Work function difference	V
$\Phi(x)$	Bottom edge of conduction band in the oxide	eV
ϕ_B	Energy barrier height at the IL / Substrate L	eV
ϕ_K	Conduction band offset between HfO_2 and Silicon	eV
ρ	Volume density	cm^3
Δ_I	Current Step	A
A	Area	cm^2
Cox	Oxide Capacitance	$F\ cm^{-2}$
D_{IT}	Density of Interface Traps	cm^{-2}
D_{HFO}	HfO_2 trap density per decade frequency	$cm^{-2}\ dec^{-1}$
D_{HfO2}	Trap generation in HfO_2	$cm^{-2}\ dec^{-1}$
D_{OT}	Frequency dependent HfO_2 bulk traps density	cm^{-2}
D_{SiO2}	Trap generation in SiO_2	$cm^{-2}\ dec^{-1}$
E_{ij}	Sub-band energy level	eV
E_F	Fermi level	eV
EOT	Equivalent Oxide Thickness	nm
f	frequency	Hz
F_{eff}	Effective electrical filed strength	eV
f_{ij}	Frequency at the dielectric /substrate interface	Hz
g_i	Degeneracy of the i^{th} valley	
I_{CP}	Charge pumping current	A
I_D	Drain current	A
I_G	Gate current	A
J_{de}	Gate current density caused by charge trapping	$Area/cm^2$
J_{ge}	Current density caused by electron flowing through the gate	$Area/cm^2$
J_{gm}	Gate current density per unit area	$Area/cm^2$
J_g	Gate current density	$Area/cm^2$
k	Boltzmann constant	J/K
L	Length	µm
m_{di}	The density of states effective mass	kg
m_{zi}	the effective mass in the i^{th} valley	kg
m_{iL}	Electron mass in IL	kg
m_o	Free electron mass	kg
m_{ox}	Effective electron mass	kg
m_{SiO2}	Electron mass in SiO_2	kg

m_{HfO2}	Electron mass in HfO_2	kg
N_A	Substrate doping density	cm^{-3}
N_{depl}	Charged dopant per unit area in the depletion layer	
N_{ic}	Electron fluency calculated with the transient gate current	cm^{-2}
N_{ij}	Density of charge carriers available for tunnelling	cm^{-2}
N_{im}	Electron fluency calculated without the transient gate current	cm^{-2}
N_{inj}	Electron fluency	cm^{-2}
N_{inv}	Total carrier density in the inversion layer	cm^{-2}
q	Elementary charge	C
Q	Charge	C
Q_C	Pumped charge per cycle	C
T	Temperature	K
t_{BD}	Time to breakdown	Sec
t_f	Fall time of the pulse	µs
t_g	Negligible trapping region located near the gate	nm
t_{iL}	Thickness of the interfacial layer	nm
t_{HfO2}	Thickness of the HfO_2	nm
t_{pw}	Pulse width	µs
t_{top}	Peak period of the pulse	µs
t_s	Negligible trapping region from the HfSiO interface	nm
t_r	Rise time of the pulse	µs
T_{Rij}	Correction factor	
$TWKB_{ij}$	WKB tunnelling probability	
V_{TH}	Threshold Voltage	V
V_{THC}	Created Threshold Voltage	V
V_G	Gate Voltage	V
V_{HfO2}	Voltage drop across HfO_2	V
V_{iL}	Voltage drop across the Interface Layer	V
V_{OX}	Voltage drop across the oxide	V
W	Width	µm
X_{Hf}	**HfO_2 thickness**	nm

CONTENTS

Acknowledgments ii

Abstract iv

List of publications vii

List of Abbreviations ix

List of Symbols x

Contents xii

Chapter 1 Introduction 15

1.1 Historical evolution 15

1.2 Challenges for further downscaling 16

1.3 Requirements of high-k gate dielectrics 18

1.4 Outline of the thesis 19

Chapter 2 Characterization Techniques for High-k Dielectrics 23

2.1 Conventional techniques 23

2.1.1 Conventional C-V 24

2.1.2 Conventional I_D ~V_G 25

2.1.3 Stress and sense 26

2.2 Times resolved techniques 28

2.2.1 Pulsed C-V 28

2.2.2 Pulsed I_D~V_G 30

2.3 Formulae for pulsed I_D~V_G technique 33

2.3.1 Normalization of drain current 33

2.3.2 Typical test pulse and I_D~V_G 34

2.3.3 Calculation of the electron fluency 36

2.4 Charge Pumping (CP) 39

2.4.1 Introduction 40

2.4.2 Conventional Charge Pumping 42

2.4.3 Variable Frequency Charge Pumping (VFCP) 46

2.4.4 Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP) 48

2.5 Time Dependent Dielectric Breakdown (TDDB) 51

2.5.1 Physics of breakdown 52

2.5.1.1 Fowler-Nordeim Tunneling 54

2.5.1.2 Direct Tunneling 55

2.5.1.3 Trap Assisted Tunneling 56

2.5.1.4 Poole-Frenkel mechanism 56

2.5.2 Techniques of Breakdown Tests 58

2.5.3 Soft Breakdown and Hard Breakdown 61

2.5.3.1 Oxide Breakdown 63

2.5.4 Advanced measurement sequences 68

2.6 Instrument and technical overview 70
2.6.1 Measurement set-up 70
2.6.1.1 Pulse I_D~V_G 70
2.6.1.2 Charge Pumping 71
2.6.1.3 Time dependent dielectric breakdown 72
2.6.2 Instruments characteristics 73
2.6.2.1 Pure DC Voltage Supply characteristics 73
2.6.2.2 Pulse Generator characteristics 78
2.6.2.3 Oscilloscope characteristics 79
2.6.2.4 Keithley K4200 79

2.7 Conclusion 80

Chapter 3 Characterization of as-grown electron traps 81

3.1 Introduction 81

3.2 Properties and dynamic behaviour of electron traps in HfO_2/SiO_2 stacks .. 83
3.2.1 Devices and Techniques 83
3.2.2 Impacts of measurement techniques on trapping 83
3.2.3 Dependence on conduction mechanism 87

3.3 Location of as grown electron traps in HfO_2/HfSiO stacks 91
3.3.1 Devices and Techniques 91
3.3.2 Test Condition 93
3.3.3 Location of as grown electron traps 96

3.4 Determination of capture cross section for as-grown electron traps in HfO_2/HfSiO stacks 99
3.4.1 Devices 99
3.4.2 Test condition and measurement 100
3.4.3 Transient gate current estimation 103
3.4.3.1 Numerical modeling of gate current without trapping 104
3.4.3.2 Numerical modeling of gate current with trapping 111
3.4.4 Capture cross section extraction 113
3.4.4.1 Electron fluency estimation 113
3.4.4.2 Capture cross section extraction and trapping kinetics 115
3.4.5 Dependence on fabrication processes and techniques 120

3.5 Conclusions 123

Chapter 4 Generated electron traps and breakdown 124

4.1 Introduction 124

4.2 Conventional Charge Pumping (CP) 125
4.2.1 Base level and Amplitude sweeps charge pumping 125
4.2.1.1 Effects of charging and discharging of HfO_2 bulk. 126
4.2.1.2 Influence of the channel length on charging and discharging 130

4.3 Improved Charge Pumping (CP) 134
4.3.1 Variable Frequency CP 134
4.3.2 Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP) 142
4.3.3 Trap generation in SiO_2/HfO_2 stacks measured by VT^2CP 150

4.3.4 Trap generation in different thickness of HfO_2 and different thickness and composition of HfSiO characterized by using VT^2CP 156
4.3.5 Trap Generation evaluated by TDDB and waver level uniformity 160
4.3.5.1 Evaluation of 2nm HfO_2 160
4.3.5.2 Evaluation of 2nm 50% HfSiON with Plasma Nitridation 164
4.3.5.3 Evaluation of 3nm 80% HfSiON with Plasma Nitridation 166
4.3.5.4 Evaluation of 3nm 50% HfSiON with Thermal Nitridation 168

4.4 Time Dependent-Dielectric Breakdown 170
4.4.1Hard breakdown in dielectric/metal gate stack 172

4.5 Conclusion 174

Chapter 5 Positive charges and V_{TH} instability 177

5.1 Introduction 177

5.2 Impact of the gate material on positive charge 178

5.3 Threshold voltage instability of pMOSFETs with Hf-based dielectrics 184
5.3.1 Comparisons of V_{TH} instability between nMOS and pMOS 184

5.4 Separation of as-grown traps from generated traps 190
5.4.1 Generated V_{TH} instability 191

5.5 Impact on operation Voltage 196
5.5.1 NBTI generation in HfO_2 196
5.5.2 NBTI generation in Hf-silicate 201

5.6 Impact of Nitridation on V_{TH} instability: Comparison between ALD vs MOCVD 204
5.6.1 Impact of process conditions on V_{TH} instability 210

5.7 Conclusions 212

Chapter 6 Conclusions and future work 215

6.1 Conclusion 215
6.1.1 Conclusions on the characterizations of as grown electrons traps 215
6.1.2 Conclusions on the generated electron traps and breakdown 216
6.1.3 Conclusion on positive charges and V_{TH} instability 218

6.2 Future work 221

References 224

CHAPTER 1 INTRODUCTION

1.1 HISTORICAL EVOLUTION

In 1947 W. Shockley, J. Bardeen and W.H. Brattain invented the bipolar transistor at Bell Laboratories, Murray Hill, New Jersey. The invention of the transistor, a solid-state amplifier, resulted in great efforts in the field of semiconductor devices. The integration of semiconductor devices on a single chip was one of the consequences of the combined efforts. J. Kilby at Texas Instruments first demonstrated the concept of Integrated Circuits (IC) in 1959.

This concept together with the fabrication of the first Metal Oxide Semiconductor Field Effect Transistor (MOSFET) by D. Kahng and M.M. Attala in 1960 provided the basis for the evolution of the microelectronics industry. Lilienfeld and Heil already proposed the principle of a surface field effect transistor in the early 1930's. The experimental verification of the surface field effect, however, could not be demonstrated for more than 30 years. As an interesting fact, the basic MOSFET structure and the principle of the operation, however, did not change.

In 1960s, Gordon Moore predicted that the number of transistors per chip wills double every 18-24 months. Since then, most of the research efforts have been spent on solving problems occurred when down-scaling the transistor sizes. Today, 90nm CMOS technology has been widely available and chips can have tens of millions of transistors. Although roadmap has been proposed to downscaling devices further in the foreseeable future, this will not be as straightforward as it was in the past. The difficulties encountered by the industry are summarized in the next section

1.2 CHALLENGES FOR FURTHER DOWNSCALING

They may be roughly divided into two groups:

Problems due to a reduction in gate oxide thickness: As the channel length shrinks, the gate oxide thickness must be reduced as well, to maintain the gate control. For the 90nm CMOS technology, the gate oxide has become as thin as 1.5nm. This leads to many difficulties, as listed below:

The oxide reliability: Time-Dependent Dielectric Breakdown (TDDB) is strongly dependent on the oxide thickness, as the operation voltage reduces for smaller MOSFETs, it is at a slower rate than the decrease of oxide thickness. This leads to an increase of oxide field strength and TDDB is a crucial issue for the current CMOS industry.

Hot-carrier effects in silicon: the avalanche multiplication, bombardment of SiON/Si interface, and hot-carrier injection into the oxide.

The mobility reduction: due to the high field perpendicular to the interface and defect creation. This leads to the degradation of the driving capabilities of the devices and the speed of the circuits.

When a MOSFET is in its off-state: the voltage drop between the gate and the drain over the thin gate dielectric gives rise a high field. This can cause electron tunneling from the valence band into the conduction band of silicon and form the gate-induced drain leakage current (GIDL).

Power consumption: because of the leakage current by tunneling through the gate oxide. For thin SiON, direct tunneling substantially increases the leakage current.

Problems related to a reduction of channel length and doping: The downscaling of channel length leads to a number of short channel effects, as described below:

The drain-induced-barrier-lowering (DIBL) effect: When a MOSFET is in its off-state, the bias applied on the drain can increasingly influence the source/body junction for smaller channel length. This lowers the potential barrier at the source/body junction and increases the leakage current. The threshold voltage reduction due to the charge-sharing effect.

The punch-through effect and the parasitic bipolar-transistor action (between source, bulk and drain): With reducing effective channel length the breakdown voltage reduces because the diffusion length of the minority carriers does not scale down.

Threshold voltage variations: Statistical variations increase with reduction of dimensions; we have to take care of the spatial tolerances and the variations in the dopant concentration.

The errors caused by ionized irradiation: With shrinking dimensions the total charge in devices reduces with a square law. Under some critical charge the error probability due to ionized irradiation increases rapidly. The ionization can be produced by the α-particles from the chip-package or by the cosmic rays.

In addition, other important issues include polysilicon gate depletion, increasing resistances of the contacts and the interconnection lines, the isolation between the devices and between the devices and the substrate, as well as the latchup effect in CMOS circuits.

1.3 REQUIREMENTS OF HIGH-K GATE DIELECTRICS

From the problems listed in the previous section, the primary aim of introducing high-k gate dielectrics is to reduce the leakage current. The main factors that determine the leakage current through an insulator are the barrier height and the physical thickness of the layer. To obtain low leakage currents in high-k dielectrics barrier heights of 1.5 eV or higher combined with a significantly higher dielectric constant compared to SiO_2 are required. Theoretical calculations and experimental data show that significant benefit in gate leakage can be obtained for high-k dielectrics like Al_2O_3, ZrO_2 and HfO_2.

Fixed charge and threshold voltage control are another important aspect when considering high-k gate dielectrics. For a standard CMOS process, n-and p- degenerated poly-Si gate electrodes with work functions of 4 eV and 5 eV are used to control the threshold voltage, V_T, respectively. In MOS devices with high-k dielectrics and poly-Si electrodes V_T is often found to deviate significantly from the values measured in devices with a SiO_2 control oxide. The observed V_T shifts are commonly attributed to fixed charge in the dielectric either located at the interface or distributed throughout the film. In order to control V_T a low fixed charge ($<10^{11}cm^{-2}$) is of importance.

Achieving high carrier mobility is considered to be essential. Again, SiO_2 devices serve as a baseline for devices with high-k materials. From literature it is known that most high-k devices suffer from severe mobility degradation. The cause for the degradation is still under debate. Scattering due to fixed charge in the dielectric or remote phonon scattering are proposed as origin for the mobility reduction. The experimental verification of the origin of the low carrier mobility, however, is still missing. In any case, control of fixed charge and understanding of the impact of remote phonons are essential for improving the carrier mobility in high-k devices.

Finally, reliability is considered as a further challenge for high-k materials. Initial experiments suggest that high gate reliability may be obtainable with these materials. However, detailed studies are required to investigate the impact of the atomic structure (amorphous versus crystalline), the interfacial layer, the stress polarity and the gate electrode on the reliability. Early learning is important to understand the reliability limitations of high-k gate dielectrics, which is the objective of this project.

1.4 OUTLINE OF THE THESIS

CHAPTER 2 INTRODUCES THE DIFFERENT TECHNIQUES TO CHARACTERIZE HIGH-K DIELECTRICS:

Some techniques used for classical SiO_2-based dielectrics can also be applied to high-k dielectrics. However, due to the specific properties of high-k layers, dedicated measurement methodologies have been introduced recently. Early in the development of high-k gate stacks it was found that they suffer from rather large DC hysteresis and threshold voltage instability effects **[YoungC05]**. Indeed, it was found that, when combined with a poly-Si gate, the dielectric stack exhibits a defect band of bulk traps that can very efficiently trap electrons. This causes hysteresis in the C-V curve and transistor characteristics. Furthermore, the charge trapping also affects the extraction of important transistor parameters like transconductance, mobility, and threshold voltage.

In this project, various measurement methods, such as Pulsed I_D-V_G, Charge Pumping (CP), and Time Dependent Dielectric Breakdown (TDDB) will be modified and used to study these defects and V_T-instabilities. The details will be given in chapter 2.

CHAPTER 3 CHARACTERIZES AS GROWN ELECTRON TRAPS:

When SiON is replace by HfO_2, the presence of as-grown electron traps becomes important as they can induce threshold voltage instability by electron trapping, reduce electron mobility through coulombic scattering **[Groes04]**, **[Degra04]** and resulting in early breakdown by forming conduction path and reduce the yield substantially **[Groes04], [Degra04] [Degra03]**. In this chapter as grown electron traps are systematically studied. Their properties and dynamic behaviour in HfO_2/SiO_2 stacks are carefully analyzed, based on the measurements techniques given in Chapter 2.

After selecting test conditions to ensure that the energy level of filling electrons was above the energy level of traps and measurement time was sufficiently fast that detrapping is negligible and using samples with a progressive reduction of HfO_2 thickness, it is ruled out that the traps are piled up at the HfO_2/SiO_2 interface. A uniform distribution throughout HfO_2 layer does not agree with the test data, either. The results support that there are trapping-free regions around 1.3~1.8nm near to either HfO_2/SiO_2 or gate/HfO_2

The determination of trap capture cross section is carried out by evaluating the gate current and the electron fluency for filling the trap. This is achieved by a numerical simulation to estimate the trapping-induced transient gate current following the application of a pulse to the gate. It is found that trapping can reduce the gate current by two orders of magnitude and the gate current can drop substantially within microseconds. The results show the presence of two distinctive capture cross sections in the order of $10^{-14}cm^2$ and $10^{-16}cm^2$, respectively, which most likely originate from two different types of as-grown electron traps in HfO_2. These capture cross sections are insensitive to fabrication and processing techniques.

CHAPTER 4 INVESTIGATES THE GENERATION OF ELECTRON TRAPS AND TIME DEPENDENT DIELECTRIC BREAKDOWN (TDDB):

Charge pumping techniques introduce in chapter 2 will be used to investigate the generation of electron traps in SiO_2/HfO_2 stacks. Specifically, the variable frequency charge pumping will be used to investigate the bulk trap density on both fresh and degraded device with different gate electrodes and an investigation on why time-to breakdown (t_{BD}) depends on the channel length will be given. The variable t_{charge} $t_{discharge}$ charge pumping (VT^2CP) is found to be more suitable than the variable frequency charge pumping to investigate the states in the interfacial layer and the bulk states at very low frequency (~90 Hz). It is able to clearly separate the traps in the interfacial SiO_2 from the traps in the HfO_2 and observe the creation of new traps in both constituent layers. During degradation at a constant positive voltage, the increase of traps, both in the SiO_2 as well as in the HfO_2, follows a power law behavior as a function of time with an exponent ~0.32-0.34. The voltage acceleration of creation of HfO_2 traps matches that of the TDDB

The VT^2CP will be used to compare traps in high-k layers of different HfO_2 thickness, different % Hf in HfSiO, various Post Deposition Anneal (PDA), and different plasma and thermal nitridation.

Furthermore, TDDB tests are carried out. TDDB shows that when poly-Si gate is replaced with a metal gate, and TDDB lifetime is consistently limited by an abrupt large current increase of several hundred μA. The occurrence of those large ΔI is a potential limitation for the reliability of metal gate devices.

CHAPTER 5 FOCUSES ON POSITIVE CHARGE GENERATION AND V_{TH} INSTABILITIES ON NMOSFETS AND PMOSFETS

The use of metal gates removes gate depletion, allows aggressive downscaling of equivalent oxide thickness. In this chapter it is demonstrated that changing poly-si gate to metal gate (TaN) affects the positive charge formation in HfO_2/SiO_2 stacks. Under positive gate bias stresses, the results show that positive charge formation in metal-gated samples is significantly higher than in poly-si gated samples. However, positive charge formation is similar in these two types of samples, when stressed under negative gate bias. The results are explained by assuming that there are more hydrogenous species at the metal/dielectric interface and the hydrogen release from the anode dominates the positive charge formation. The behaviour of positive charges in the stack is also compared with that in a single SiO_2 layer. When compared with nMOSFETs, the V_{TH} instability of pMOSFETs using Hf-silicates for sub-2nm nitrided layer is shown to be insensitive to measurement time, does not saturate as stress voltage increases and is not controlled by carrier fluency. Impacts of different process conditions, such as plasma nitridation, thermal nitridation, HfO_2 thickness and Hf content, is carried out. The results show that NH_3 nitridation can give more V_{TH} instability for pMOSFETS than nMOSFETs with high Hf-content (80%) but similar or less with 50% Hf-content and can reduce the operation voltage by 1V. Also it is shown that nitridation introduces two different types of defects, resulting in two types of instabilities with distinctive characteristics.

CHAPTER 6 CONCLUDES THE NEW FINDINGS AND DISCUSSES THE FUTURE WORK.

CHAPTER 2 CHARACTERIZATION TECHNIQUES FOR HIGH-K DIELECTRICS

Charge trapping in conventional SiO_2 gate dielectrics is commonly studied using either hysteresis measurements or 'stress and sense' techniques. In the past, the same methodology was also been applied to study charge trapping in high-k dielectrics. In a recent study it has been demonstrated that transient charging effects in high-k materials (e.g. HfO_2) make a reliable assessment of the charge trapping more complicated than in SiO_2. Therefore fast measurement techniques have been developed to capture the fast transient effects previously not reported. **[Ker04]**, **[Ker03]**, **[Cart04]**, **[Gus04]**. In this chapter, different techniques used for the characterization of high-k dielectrics will be described. In this chapter we will discuss on the different techniques available for the characterization of high-k dielectrics. Conventional, pulsed and advanced techniques will be described also the test condition to calculate the capture cross section will be given.

2.1 CONVENTIONAL TECHNIQUES

In a conventional hysteresis measurement a quasi DC ramp is applied to the MOS device using ramp rates ranging from 0.1V/s to 10V/s. Charge trapping is monitored either using the I_D-V_G or the Capacitance-Voltage characteristic. In both cases, when charge trapping / detrapping is present, the initial trace will deviate from the final trace due to build up or loss of charges. Fast transient effects, however, are not captured by the conventional hysteresis measurement due to the inherently slow ramp

rates. A schematic drawing of the gate voltage during a hysteresis measurement is shown in Fig. 2.1. **[Ker04]**, **[Ker03]**, **[Cart04]**, **[Gus04]**

2.1.1 Conventional C-V

The C-V hysteresis measurement provides a fast screening method with respect to instabilities in MOS devices. Additional information on the voltage dependence of the instability can be obtained when the conventional hysteresis measurement is extended using a bias sequence as shown in Fig. 2.1. Furthermore, the use of multiple traces should also allow to address the reversibility of the observed effects. **[Cart04]**, **[Gus04]**

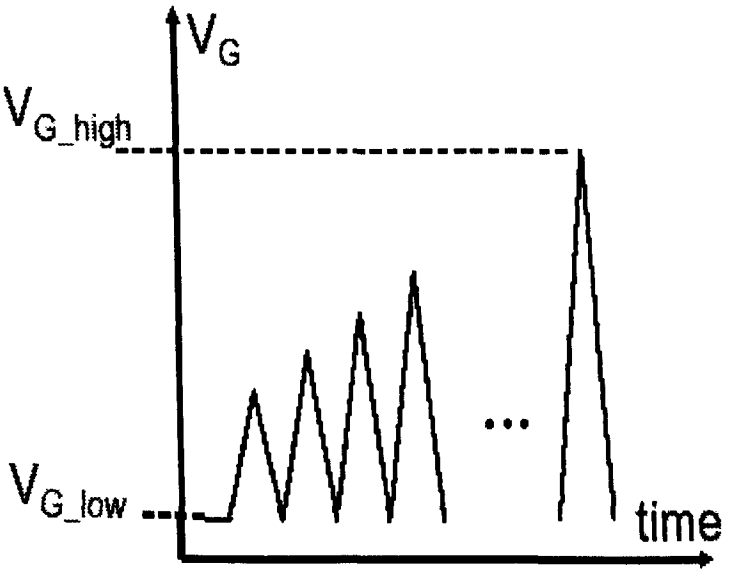


Fig.2.1 Schematic drawing of the gate bias during a hysteresis measurement applied to n-channel MOSFETs. An appropriate negative gate bias (V_G -low) is chosen to allow for complete discharging, whereas the high level of the gate bias (V_G -high) is continuously increased to monitor charge trapping in inversion [Ker03], [Cart04]

Previous works shows **[Ker03]**, **[Cart04]** when the gate bias is swept from negative to positive identical C-V traces are obtained. From this it could be concluded that the gate stack does not suffer from instability. However, when the sweep direction

is reversed from positive to negative the instability become evident depending on the maximum gate bias, C-V shifts of ~ 100mV and more can easily be measured. When the device is swept further into accumulation (negative gate bias) the C-V trace merge again indicating that the instability completely recovers. The major drawback of the C-V hysteresis or the multiple C-V trace technique is the poor control over the amount of injected charge during such an experiment. Therefore the magnitude of the measured instability can significantly vary depending on the ramp rates of the sweep.

2.1.2 Conventional I_D~V_G

A complementary technique to the C-V hysteresis measurement is the use of multiple I_D~V_G traces on MOSFETs. Again sequences of sweeps are taken starting from accumulation, a negative bias for n-channel devices or a positive gate bias for p-channel devices, and sweeping the MOSFET towards inversion successively increasing the maximum voltage. In this case the transfer characteristic is used to monitor the instability as shown in Fig.2.2.

Using this technique a shift of ~100mV and more in the I_D~V_G characteristic can easily be detected depending on the maximum positive gate bias applied to the gate. The use of the linear I_D~V_G characteristic to monitor the instability also allows to study the recovery during the down ramp of the sweep. When going to large positive gate biases, where larger shifts are observed, a stretch-out on the reverse trace is evident. This indicates that a fraction of the instability already dynamically recovers when the high gate bias is reduced.

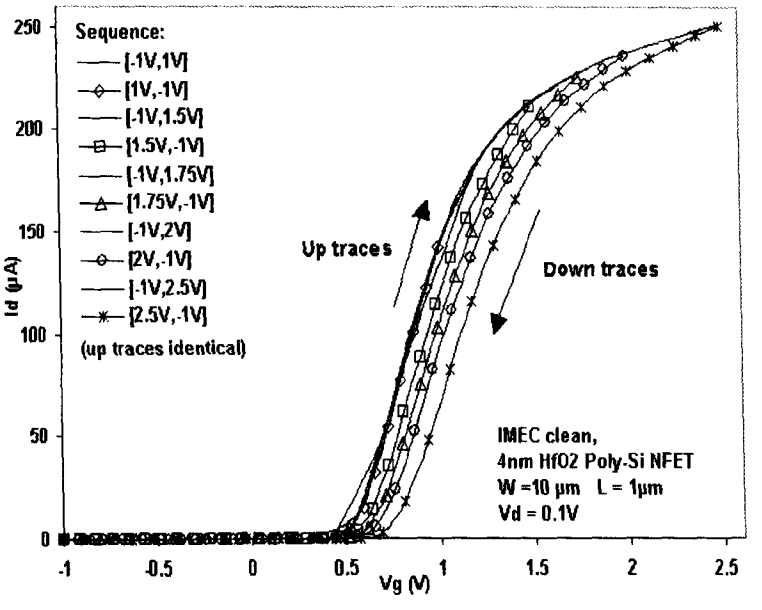


Fig.2.2 I_D~V_G sweeps sequence from accumulation to inversion using conventional technique to monitor instability for n-channel MOSFET.A large shift could be observed when sweep to high voltage.

The use of the conventional I_D~V_G characteristic limits the monitoring capability to the operation mode of the MOS transistor, which is anyway the focus of interest from application point of view. In the case of I_D~V_G instabilities, when applying a sufficient negative gate bias complete recovery are also obtained. **[Cart03]**

2.1.3 Stress and sense

A further procedure, which was used extensively in the literature to study the instabilities in conventional SiO_2 based gate dielectrics, is known as "stress and sense" method.

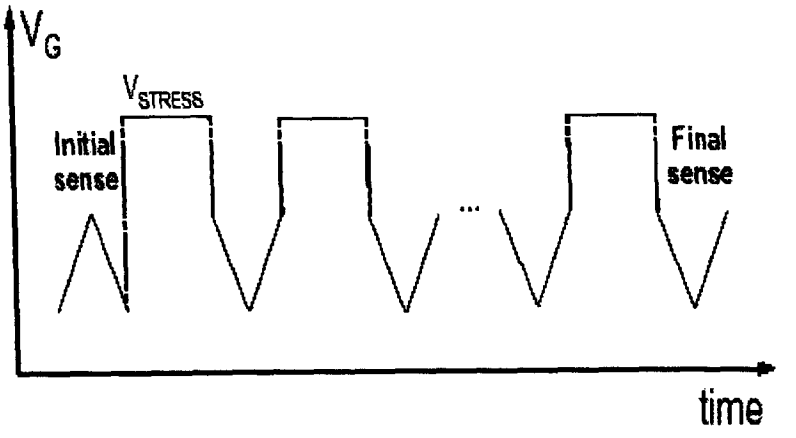


Fig.2.3 Schematic drawing of a "stress and sense" experiment. The stress applied to the device is periodically interrupted to sense the charge state by comparing the V_{FB} or V_T to the initial trace.

In this case, first an initial sense measurement is carried out prior to stressing the devices. The stress is then interrupted periodically and a sense measurement is performed. A schematic drawing of the bias sequence in the conventional stress and sense procedure is shown in Fig. 2.3. **[ZhaoY]**

The instability of the device is extracted by comparing the sense measurement after stress with the initial device characteristic. Predictions to operation conditions are usually made when combining the time dependence with the voltage dependence of the instability. One of the known issues of the "stress and sense" procedure is the inherent time delay between stressing and sensing. In case some recovery of the instability occurs at time periods of the order of ~10 to 100 ms this procedure will not capture its full extent. Furthermore, the selection of the sense condition requires special attention. In any case, the use of this procedure will help to understand the stability of MOSFETs with high-k gate dielectrics.

2.2 TIMES RESOLVED TECHNIQUES

As briefly mentioned in the previous section due to the presence of fast recovery of the instabilities measurement techniques significantly faster than the semiconductor parameter analyzers are required for a correct quantification. Therefore, alternative techniques were introduced which will be discussed in the following section and some of them will be extensively applied to high-k gate stacks in this project.

2.2.1 Pulsed C-V

The pulsed C-V technique is similar to the quasi-static C-V measurement using a linear voltage ramp **[Nico82]**. A linear voltage ramp is applied to the Device Under Test (DUT) while the capacitor displacement current is recorded using a current meter. The CV characteristic can be readily obtained from the voltage and current values, as it will be explained further below.

In conventional quasi-static measurement often the DUT consists of an MOS capacitor that has no means to provide minority carriers. As a consequence, non-equilibrium carrier concentrations can exist within the device for time spans well above 100μs. Therefore, typical quasi-static CV measurements use a voltage ramp rate well below 1 V/s. The pulsed CV technique employs DUTs that feature a pn-junction at the perimeter of the gate electrode to provide minority carriers. Hence, fast ramp rates of 1kV/s and above can employed without pulling the MOS capacitor out of equilibrium. The resulting CV characteristic is identical to that obtained from a conventional high-frequency CV measurement.

The pulsed CV technique has been used by Weinberg and Fischetti **[Wein86]** in a study of the SiO_2- induced substrate current in field-effect transistors. A fast ramp rate in the order of 1kV/s has been used to capture V_{FB} shifts immediately after application

of stress pulse. A similar set-up has been used by Agarwal et al. **[Agar02]** to study charge transport, however using a slow ramp rate below 1V/s.

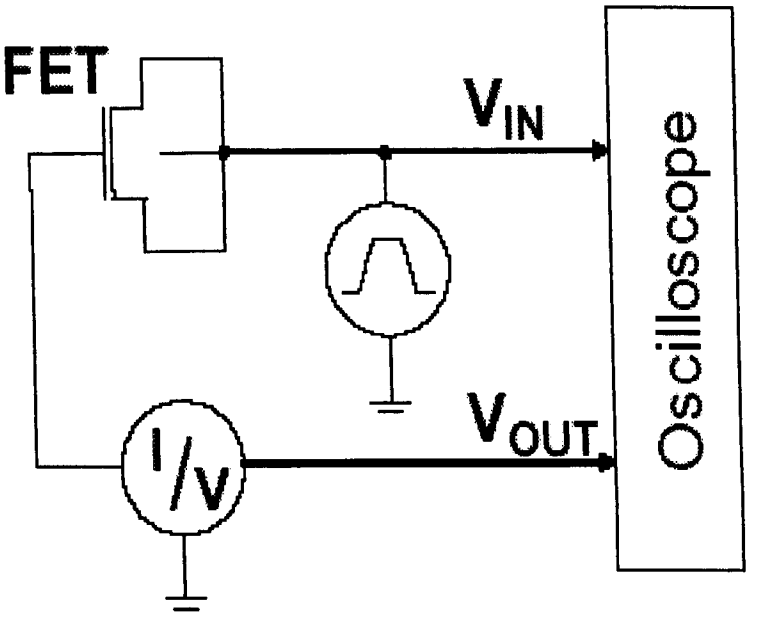


Fig.2.4 Schematic setup used to obtain pulsed C-V characteristics. The voltage pulse applied to the source, drain and substrate (V_{IN}) is recorded using a digital oscilloscope together with the output voltage (V_{OUT}) of the current-voltage amplifier.[Ker03A]

In this pulsed CV technique (Fig.2.4) the ramp rates can vary from ~ 0.1V/s to >10 kV/s and as a result the corresponding displacement current is several orders of magnitude higher than in the conventional quasi static C-V measurement. The displacement current is converted into a voltage trace using a current-voltage amplifier and recorded with a digital oscilloscope. The voltage ramp with fixed ramp rate is provided at the edges of a pulse. From the voltage trace the C-V characteristic can be extracted using the equation 2.1 **[Ker03A]**

$$C(V_{IN}) = \frac{V_{OUT} - V_{OFFSET}}{Z_{AMP}.dV/dt} \qquad (2.1)$$

Where dV/dt is the voltage ramp, V_{OFFSET} is the offset voltage of the current-voltage amplifier and Z_{AMP} the current-voltage gain. The input (V_{IN}) and output (V_{OUT})

voltage correspond to the terminals of the circuit as illustrated in the schematic measurement setup shown in Fig. 2.4.

2.2.2 Pulsed I_D~V_G

The pulsed I_D~V_G measurement technique uses a pulse generator to provide a pulse with sloped edges. During the pulse edge, the current that flows through the transistor is monitored. Using a sampling oscilloscope, both the pulse bias and transistor current are recorded, which allows for the determination of the I_D~V_G characteristics. The technique allows for determination of the V_T immediately after application of a pulse by capturing the I_D~V_G during the returning edge of the stress pulse. This short read delay time is essential for determining the fast initial charge loss in charge trapping devices. Details of the instruments and technical overview is given in the section 2.7.The set-up for pulsed I_D~V_G measurements is obtained by a slight modification of the set-up for pulsed CV. Figure 2.5 shows the schematic of the set up and Figure 2.6 the V_G-t and V_D-t traces recorded by the oscilloscope.

The major advantage of this technique over the pulsed C-V method is the reduced sensitivity to the parasitic gate leakage current. Unless the I_D~V_G characteristic is distorted by the leakage current, this procedure can be used to monitor V_T shifts in the μs time range. **[Ker03], [Cart04]**

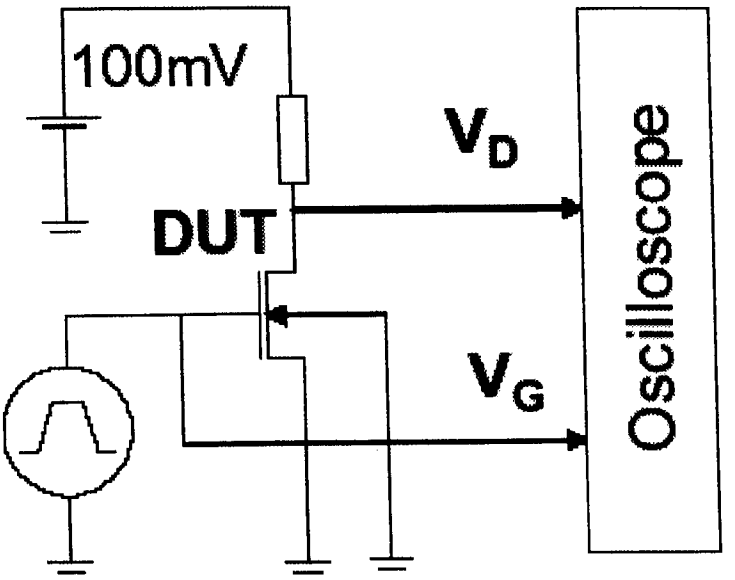


Fig.2.5 Schematic of setup used for pulsed I_D~V_G measurements. The FET is used in an inverter circuit with a resistive load (R_l). From the voltage - time traces (V_G (t), V_D (t)) the I_D~V_G characteristic is extracted.

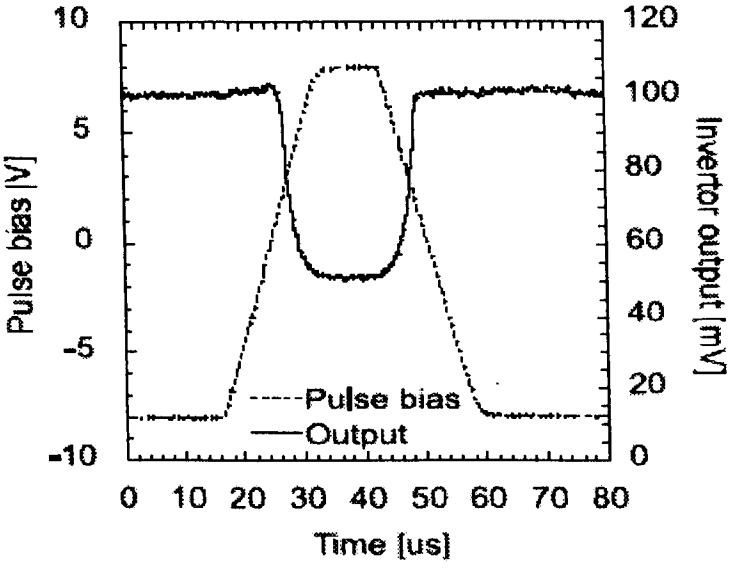


Fig.2.6 The V_G-t and V_D-t traces recorded by the oscilloscope.

In the setup in Fig 2.5, the MOSFET is used in an inverter circuit with a resistive load (R_L). A small DC bias is applied to the resistor, which acts together with the channel resistor as a voltage divider. The I_D~V_G characteristic is obtained by applying a trapezoidal (triangular) pulse to the gate and recording the drain voltage using a digital oscilloscope as shown in Fig 2.6. From the measured voltage traces the I_D~V_G characteristic can be constructed using the Eq. 2.2

$$I_D = \frac{100mV}{V_D}\left(\frac{100mV - V_D}{R_L}\right) \tag{2.2}$$

where V_D is the measured drain voltage and R_L the resistive load of the inverter **[Ker03]**. Due to the use of a voltage divider the drain voltage dynamically changes during the measurement. However, this effect can simply be eliminated by normalizing the extracted drain current to a constant drain voltage, which is given by the term $100mV/V_D$ in Eq. 2.2.

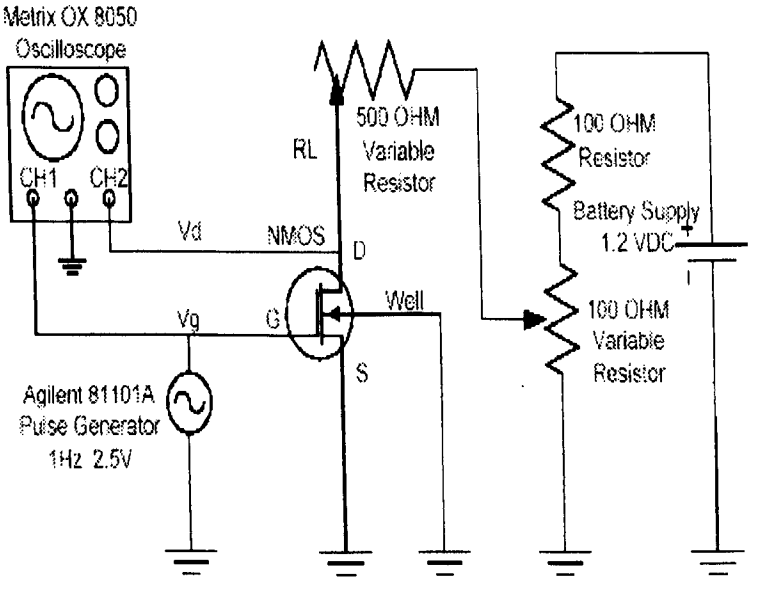


Fig.2.7 Schematic drawing of measurement setup used for pulsed I_D~V_G experiment in the µs range slightly different to the one presented in literature and in Fig.2.5.

It is noted that this normalization is only correct, when the MOSFET is operated in the linear regime, which limits the range for the DC bias applied to the resistor **[Ker03B]**, **[Cart04]**. The derivation of Eq. (2.2) is given in section 2.3.1. When replacing the resistor with a current-voltage amplifier connected to the source terminal this problem can be circumvented and a constant bias can directly be applied to the drain terminal of the device.

In the work presented in chapters 3 and 5, the pulsed I_D~V_G was modified to the schematic setup shown in Fig. 2.7 by adding a voltage divider to apply a small DC bias (100mV) to the drain instead of using a power supply. This change is explained in more details in section 2.6.

2.3 FORMULAE FOR PULSED I_D~V_G TECHNIQUE

2.3.1 Normalization of drain current

For the conventional DC measurement, the drain bias, V_D, is fixed at the supply voltage, $V_D=V_{DD}$. The I_D equation is:

$$I_D = \frac{C_{OX}\mu_n W}{L}\left[(V_G - V_T)V_{DD} - \frac{1}{2}V_{DD}^2\right] \approx \beta(V_G - V_T)V_{DD} \tag{2.3}$$

where

$$\beta = \frac{C_{OX}\mu_n W}{L} \tag{2.4}$$

For Pulsed I_D~V_G, the existence of the resistor, R_L, connected to the drain in Figs. 2.5 and 2.7 results in $V_D<V_{DD}$, when the MOSFET is switched on. The measured drain current, I_{Dm}, is:

$$I_{Dm} = \frac{V_{DD} - V_D}{R_L} \tag{2.5}$$

and

$$I_{Dm} = \beta(V_G - V_T)V_D \tag{2.6}$$

To find the drain current at a V_{DD}, the measured drain current, I_{Dm}, must be normalized. From Eq.(2.3), one has,

$$\beta(V_G - V_T) = \frac{I_D}{V_{DD}} \tag{2.7}$$

Substituting Eqs.(2.5) and (2.7) into the left and right hand side of Eq. (2.6), one can obtain,

$$I_D = \frac{V_{DD}}{V_D}\left(\frac{V_{DD} - V_D}{R_L}\right) \tag{2.8}$$

Typically, V_{DD} is set at 100mV and Eq. (2.8) is the same as Eq. (2.2).

2.3.2 Typical test pulse and I_D~V_G

In this section the typical test gate pulse and the resultant I_D-V_G will be presented. They are used to calculate the capture cross section in chapter 3. As mentioned earlier, to extract the capture cross section, detrapping should be suppressed. This requires using the pulsed I_D~V_G technique.

Fig.2.8a shows the typical pulse applied to the gate. A -1V was first applied for a sufficiently long time (~1sec) to 'reset' the sample by emptying traps through detrapping. During the rising edge of the pulse, drain current was monitored to give the I_D~V_G curve on the left in Fig.2.8b. Traps were filled with electrons during the period of 't_{top}'. The negative trapped charges lead to a positive shift of the I_D~V_G during the falling edge of the pulse, as shown by the curve on the right in Fig.2.8b. The trapping level is measured from this shift, ΔV_G. The measurement time is within the falling period, t_f, and their impacts on the result will be evaluated in chapter 3 section 3.4.4.

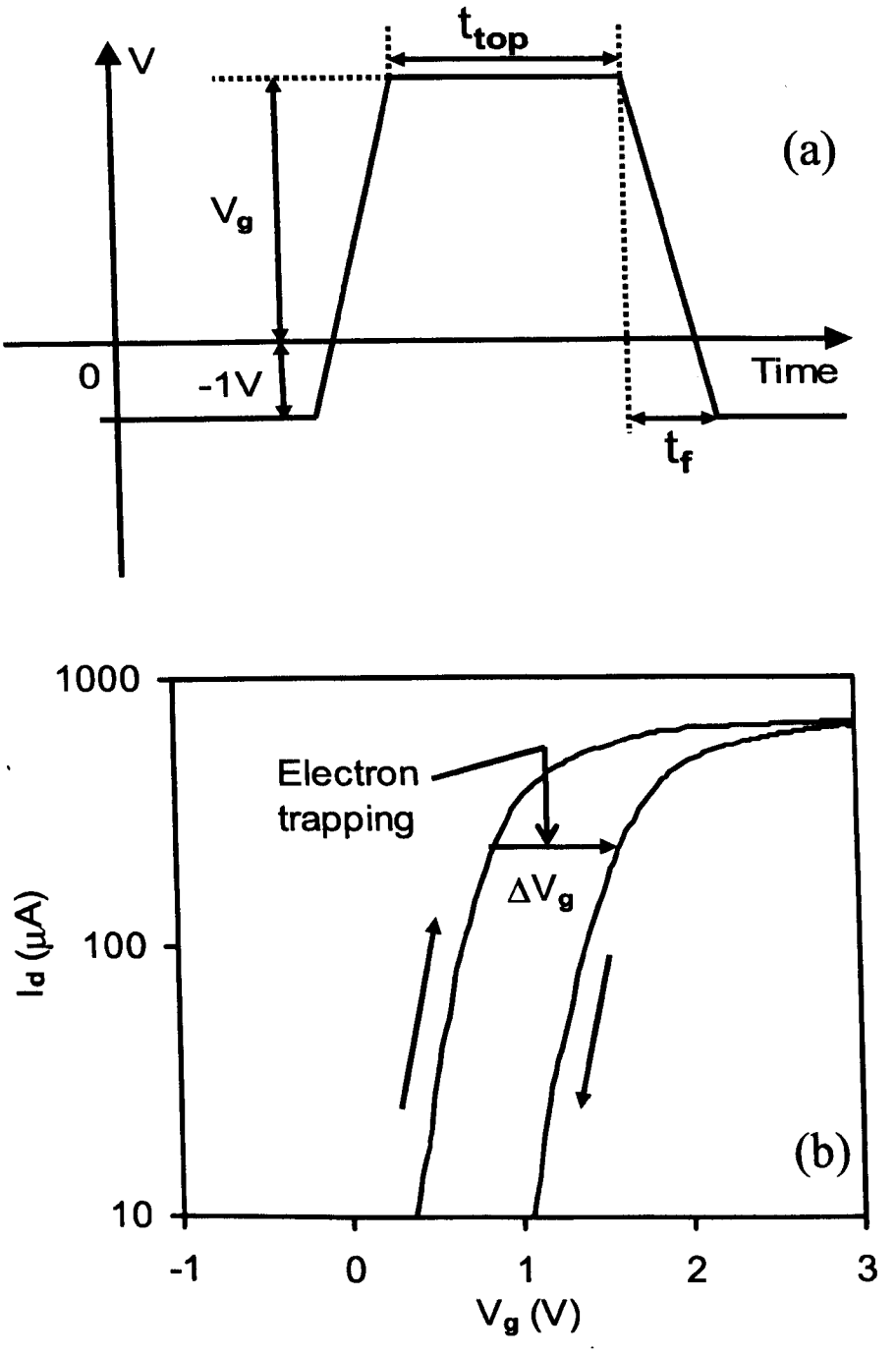


Fig.2.8 The pulse applied to the gate (a) and the pulsed transfer characteristics (b). A -1V was first applied to the gate to empty any trapped electrons. During the rising edge of the pulse, trapping is negligible and the corresponding I_D~V_G is the left curve in (b). Traps were filled during the period of t_{top}. This leads to a shift of gate bias, ΔV_G, when the I_D~V_G was measured again during the falling edge of the pulse. The delay between trap filling and measurement is less than the falling time, t_f.

2.3.3 Calculation of the electron fluency

To evaluate the capture cross section of electron traps, one must know the number of electrons injected from the substrate into the gate dielectric. The formulae used in Chapter 3, Eq.(3.1), is

$$N_{im}=J_{gm}\times t_{top}/q + \Delta V_G \times C_{ox}/q. \quad (3.1)$$

There has been some confusion on whether the 2nd term should be included in the above formulae. In this section, details will be given on how this equation is derived.

Since Eq.(3.1) is applicable for both a single layer and a stack of dielectrics, to simplify the derivation, the dielectric stack is considered as a single layer of SiO_2 with a thickness of EOT. As illustrated in Fig.2.9.

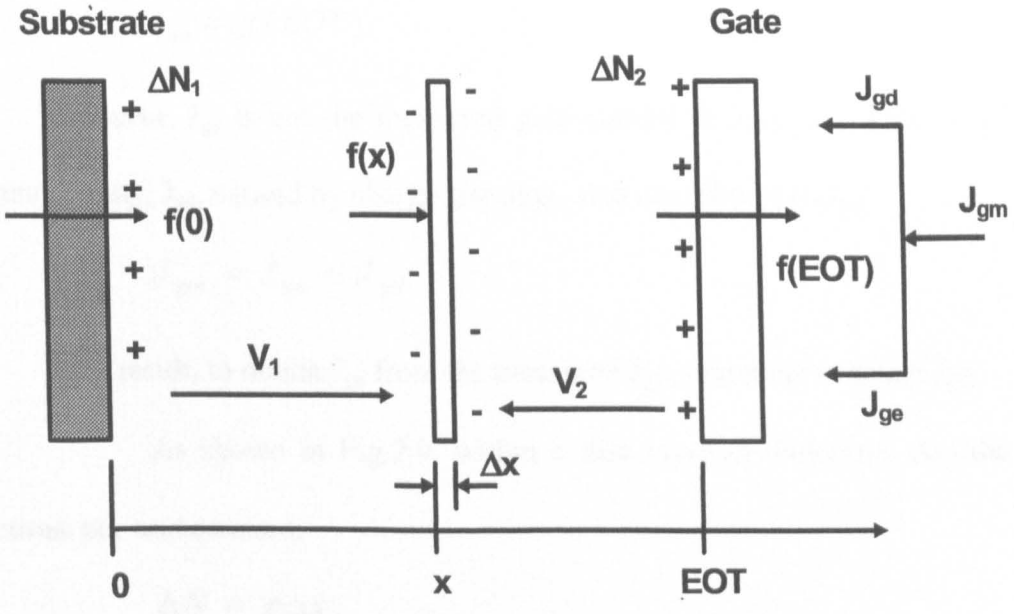


Fig.2.9 A schematic illustrations of electron injection, trapping, and the relevant current components used for calculating electron fluency.

If we define f(x) as the number of electrons passing through a unit area per unit time at a distance of 'x' from the substrate/dielectric interface, the electron conservation equation is,

$$\frac{\partial f}{\partial x} = -\frac{\partial \rho}{\partial t},$$

where ρ is the number of trapped electrons per unit volume. An integration against 'x' leads to,

$$f(0) = f(EOT) + \int_0^{EOT} \frac{\partial \rho}{\partial t} dx. \tag{2.9}$$

The second term on the right hand side represents the loss of electrons through trapping in dielectric. By definition, N_{im} is related to f(0) through,

$$N_{im} = \int_0^t f(0) dt. \tag{2.10}$$

The current density caused by electron flowing through the gate/dielectric interface, J_{ge}, is related to f(EOT) by,

$$J_{ge} = qf(EOT). \tag{2.11}$$

However, J_{ge} is not the measured gate current density J_{gm}, since displacement current density, J_{gd}, caused by charge trapping, also contributes to J_{gm},

$$J_{gm} = J_{ge} + J_{gd}. \tag{2.12}$$

As a result, to obtain J_{ge} from the measured J_{gm}, one must evaluate J_{gd}.

As shown in Fig.2.9, within a thin layer of dielectric, Δx, the trapped electrons per unit area are,

$$\Delta N = \rho \Delta x. \tag{2.13}$$

The ΔN induces a positive charge of ΔN_1 at the substrate interface and ΔN_2 at the gate interface with,

$$\Delta N = \Delta N_1 + \Delta N_2. \tag{2.14}$$

The corresponding potential drop between the substrate and the layer at 'x' is,

$$V_1 = \frac{q \Delta N_1 x}{\varepsilon_0 \varepsilon_{SiO2}}, \tag{2.15}$$

and similarly,

$$V_2 = \frac{q\Delta N_2 (EOT - x)}{\varepsilon_0 \varepsilon_{SiO2}}. \tag{2.16}$$

Since the test is carried out with a constant voltage over the dielectric, it requires,

$$V_1 = V_2 . \tag{2.17}$$

Solving Eq. (2.14-2.17), we have,

$$\begin{aligned} \Delta N_2 &= \frac{x}{EOT}\Delta N \\ &= \frac{x}{EOT}\rho \Delta x . \end{aligned} \tag{2.18}$$

Integrating Eq. (2.18), the total trapping induced positive charges on the gate is,

$$N_g = \int_0^{EOT} \frac{x}{EOT}\rho dx . \tag{2.19}$$

The displacement current is,

$$\begin{aligned} J_{gd} &= q\frac{\partial N_g}{\partial t} \\ &= q\frac{\partial}{\partial t}\left(\int_0^{EOT} \frac{x}{EOT}\rho dx\right). \end{aligned} \tag{2.20}$$

By using Eqs. (2.11), (2.12) and (2.20), the f(EOT) can now be evaluated through,

$$\begin{aligned} f(EOT) &= \frac{1}{q}(J_{gm} - J_{gd}) \\ &= \frac{J_{gm}}{q} - \frac{\partial}{\partial t}\left(\int_0^{EOT} \frac{x}{EOT}\rho dx\right). \end{aligned} \tag{2.21}$$

From Eq. (2.9), we have,

$$f(0) = f(EOT) + \int_0^{EOT} \frac{\partial \rho}{\partial t} dx$$
$$= \frac{J_{gm}}{q} - \frac{\partial}{\partial t}\left(\int_0^{EOT} \frac{x}{EOT} \rho dx \right) + \int_0^{EOT} \frac{\partial \rho}{\partial t} dx$$
$$= \frac{J_{gm}}{q} + \frac{\partial}{\partial t}\left[\int_0^{EOT} (1 - \frac{x}{EOT}) \rho dx \right]. \quad (2.22)$$

Using Eq. (2.10), the electron fluency at the substrate/dielectric interface is,

$$N_{im} = \int_0^t f(0) dt$$
$$= \frac{1}{q} \int_0^t J_{gm} dt + \int_0^{EOT} (1 - \frac{x}{EOT}) \rho dx . \quad (2.23)$$

For a fixed drain current, the trapping induced gate voltage shift is,

$$\Delta V_g = \frac{q}{C_{ox}} \int_0^{EOT} (1 - \frac{x}{EOT}) \rho dx . \quad (2.24)$$

Combining Eqs. (2.23) and (2.24), we have,

$$N_{im} = \frac{1}{q} \int_0^t J_{gm} dt + \frac{\Delta V_g C_{ox}}{q}. \quad (2.25)$$

If one assumes J_{gm} is a constant over a period of t_{top}, Eq. (2.25) becomes the Eq.(3.1),

$$N_{im} = J_{gm} \times t_{top}/q + \Delta V_g \times C_{ox}/q. \quad (3.1)$$

2.4 CHARGE PUMPING (CP)

Charge pumping (CP) is a transient effect in MOS devices. It represents the transfer of charge from the source and drain terminals to the bulk by transient pulses at the terminals. A DC component of the terminal currents can be measured. In this section conventional and advanced use of the charge-pumping are presented. Moreover, an introduction on the conventional charge pumping is given. The applicability of the

conventional CP combine with the Variable Frequency Charge-Pumping (VFCP) is introduced. Finally we develop the technique of the variable frequency charge pumping by independently controlling the pulse low and high level timings. This technique is called as Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP). In chapter 4, VFCP and VT^2CP are used to analyze the degradation of MOS devices due to defect generation in SiO_2 and HfO_2 layers.

2.4.1 Introduction

The charge-pumping effect in MOS devices has been first reported by J.S.Brugler and P.G.A.Jespers in 1969 **[Brug69]** and by A.Goetzberger and E.H.Nicollian **[Goet70]**. They have measured a DC component of the bulk current when periodic voltage pulses are applied to the gate in the circuit. This current was in the opposite direction and much larger than the leakage current of the reversely biased source and drain junctions. It was proportional to the pulse frequency and the gate area. This current is called the charge-pumping current I_{CP}

Neglecting the junction leakage current, the DC component of the bulk current is caused by two effects (for n-channel MOSFETs):

1. By varying the gate pulse sufficiently so that the condition at the interface changes from inversion to accumulation, traps at the interface (and some bulk traps) are successively filled by electrons and holes. This produces a positive net bulk hole current and a negative net source/drain electron current. The effect originates due to the longer emission times for traps compared to the duration of the gate-pulse rise and fall times. The charge-pumping current due to this effect contains information on the traps in MOSFETs.

2. For slow turn-off of MOSFETs (long fall time of the gate pulse) electrons in the channel have sufficient time to arrive at the source and drain junctions. The

inversion layer vanishes before the accumulation occurs at the interface. However, for fast turn-off, when the fall time of the gate pulse is shorter than the time necessary to remove the inversion layer, a significant amount of electrons remain at the interface when the hole accumulation takes place. These electrons are either recombined by holes or transferred to the bulk, which both produce a positive net bulk current of holes and electrons, respectively. The charge-pumping current due to this fast-switching effect does not contain any information on the traps in the device.

The first effect has been extensively used in the last ten years to extract the amount and the distribution (in both, energy and position space) of traps in MOS devices. In these measurements, the second effect introduces a parasitic undesired component which can be removed in most cases. The reasons for an increasing popularity of charge pumping are its simplicity, high sensitivity, good accuracy and its direct applicability on small devices.

Charge pumping is mostly applied to analyze the localized degradation after hot-carrier stress in MOSFETs **[Shaw89]**, **[Poor84]**, **[Mahn88]**, **[Maes82]**, **[Here89]**, **[Here88]**, **[Berg92]**, **[Ncona88]**. Its ability to be performed on small devices (with real design dimensions) is particularly useful in studying the nonuniform degradation of memory MOS devices (EPROMs and EEPROMs) under working conditions **[Witt87]**, **[Witt87A]**, **[Houd90]**, **[Here88]**. Uniform degradation caused by E-beam **[Here88]** and γ-rays **[Wach86]** irradiation and by Fowler-Nordheim injection **[Wach86]**, **[Here88]**, **[ChenW92]**, has been studied by the charge pumping as well. A particular problem is to study traps in SOI devices, where both, front and back interface play an important role in determining the properties of those devices.

2.4.2 Conventional Charge Pumping

In this section a review of the charge-pumping techniques used to characterize the interface traps is given. If not explicitly said, we assume n-channel devices.

The gate of the MOSFET is connected to a pulse generator as shown in Fig.2.10, which is repeatedly switched from accumulation to inversion and vice versa, while keeping the source, drain and body contacts grounded or slightly reverse biased. During accumulation, some of the majority carriers provided by the body are trapped on interfaces states. During the rising edge of the gate pulse, the mobile majority carriers are collected rapidly from the accumulation layer by the body, and then the trapped majority carriers recombine with the minority carriers provided by the source and drain. Similarly, during the falling edge of the gate pulse, when the gate surface is pulsed from inversion to accumulation, the trapped minority carriers recombine with majority carriers.

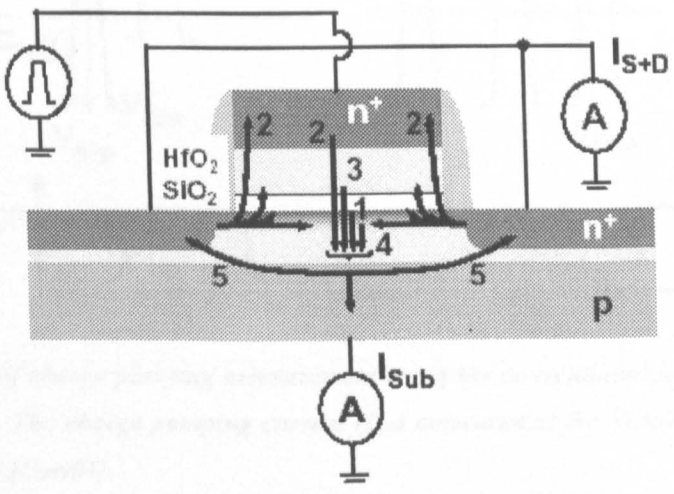


Fig.2.10 Possible current contributions in a charge pumping measurement with alternative gate dielectrics. Beside the recombination current due to interface states (1), the gate current contribution (2), charging and discharging of bulk defects (3), recombination of inversion carriers (4) and minority carrier diffusion (5) due to electron injection from the HfO_2 layer need to be considered.[Ker03],[Cart04]

All charge-pumping techniques have one common property: they are direct methods to sense the interface states. In addition, they can be applied directly on small devices.

In a conventional Base level CP measurement (left in Fig.2.11) the amplitude of the gate pulse is fixed and the base level is swept from either accumulation to inversion or vice versa. An alternative charge pumping procedure is given when the base level of the pulse is kept constant and the amplitude is varied (right in Fig.2.11). This procedure is called amplitude sweep charge pumping. In case of conventional gate dielectrics both techniques yield similar results, whereas for high-k dielectrics significant differences can be observed, result and experiment will shown in chapter 4 section 4.2.1.

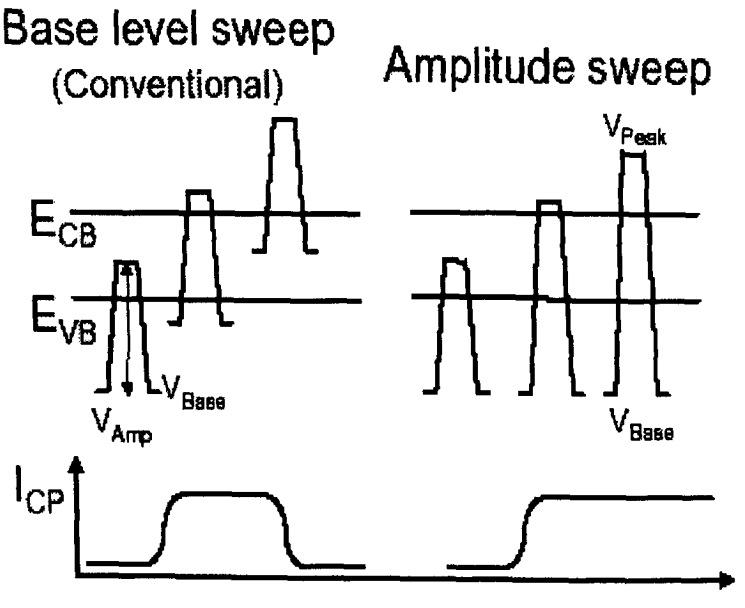


Fig.2.11 Schematic of charge pumping measurement using the conventional base level sweep and an amplitude sweep. The charge pumping current (I_{CP}) measured at the Si substrate is shown for comparison.[Ker03],[Cart04]

To understand the cause of the discrepancy between the two charges pumping modes we first discuss the different current contributions in a charge pumping measurement as illustrated in Fig. 2.10.

The possible current contributions in a charge pumping measurement with alternative gate dielectrics include the recombination current due to interface states (1), the gate current contribution (2), charging and discharging of bulk defects (3), recombination of inversion carriers (4) and minority carrier diffusion (5). The recombination process gives rise to a DC charge pumping current (I_{CP}) in the body, which flows in the opposite direction of the normal drain and source to substrate leakage currents and is given by Eq.2.26

$$I_{CP} = qfAD_T \tag{2.26}$$

where q is the electron charge, A the device area, f the frequency and D_T the total trap density [D_{IT} (interfaces traps density) + D_{OT} (Bulk traps density)]

The measured charge pumping current depends on the rise (t_r) and fall (t_f) time of the gate pulse due to thermal emission of carriers and the capture cross-section for electrons σ_e and holes σ_h. For the Si / SiO_2 interface with a uniform surface states density throughout the bandgap the charge pumping current follows Eq. 2.27 where $\overline{D_{it}}$ is the average interface state density, V_{TH} the thermal velocity of the carriers, n_i the intrinsic carrier concentration and ΔV_G the amplitude of the gate pulse.

$$I_{CP} = 2qfA\overline{D_{it}}\left[\ln\left(v_{th} n_i \sqrt{\sigma_\varepsilon \sigma_h}\,\frac{V_{FB} - V_T}{\Delta V_G}\sqrt{t_r t_f}\right)\right] \tag{2.27}$$

When rise and fall times or ramp rates (V/s) are kept constant the normalized charge per cycle N_{CP} is frequency independent.

$$N_{CP} = \frac{I_{CP}}{qfA} \tag{2.28}$$

Beside the contribution from recombination of interface states, gate leakage is known as a source of additional substrate current. Especially in ultra-thin gate dielectrics the contribution exponentially increases with decreasing oxide thickness.

Therefore, experimental procedures have been proposed which can efficiently suppress this effect **[ChungS02]**. In addition to interface states, the charging and discharging of bulk defect states in the high-k gate dielectrics also contribute to the substrate current. Other than for conventional gate dielectrics this component can add significantly to the substrate current, due to the presence of a high trap density in the high-k material. In particular a strong voltage and frequency dependence is expected due to the spatial and energy distribution of the bulk defects.

Furthermore, the recombination of inversion carriers with bulk majority carriers can also contribute to the substrate current. This effect is known in the literature as the geometrical component **[Bosch93]**. Therefore, short channel devices (< 1 μm) are commonly used for charge pumping measurements, where this effect is minimized. When considering the charging and discharging process during a charge pumping cycle in more detail, minority carriers injected from the inversion layer get emitted when majority carriers in the Si substrate are accumulated again. In order to contribute to the charge pumping current these carriers have to recombine in the substrate. However, the minority carriers in the Si have a certain lifetime and can diffuse back to the source and drain junctions. Therefore, only a fraction of the emitted carriers will effectively contribute to the substrate current. This effect is expected to significantly depend on the device geometry.

Conventional Base Level Sweep Charge Pumping and Amplitude Charge Pumping are used in chapter 4 section 4.2.1 to study interface and bulk defect in SiO_2 / HfO_2 with different gate electrode and different channel length. The conventional base level sweep with rather small amplitude is used to sense interfaces state, whereas amplitude sweep is applied to investigate in more details the interface and bulk defect in HfO_2.

2.4.3 Variable Frequency Charge Pumping (VFCP)

CP is commonly used to measure the substrate/dielectric interface trap density, but it was demonstrated in **[Ker03A]**, **[Ker02]** that it can also sense traps in the HfO_2 defect band. By varying the CP frequency we can sense different fractions of the trap density **[Degra03].** At high frequency (i.e. MHz ,Fig 2.12), only Si/SiO_2 interface trap are pumped, but at lower frequency (i.e. kHz, Fig 2.12) traps deeper in the stack can also respond to the signal.

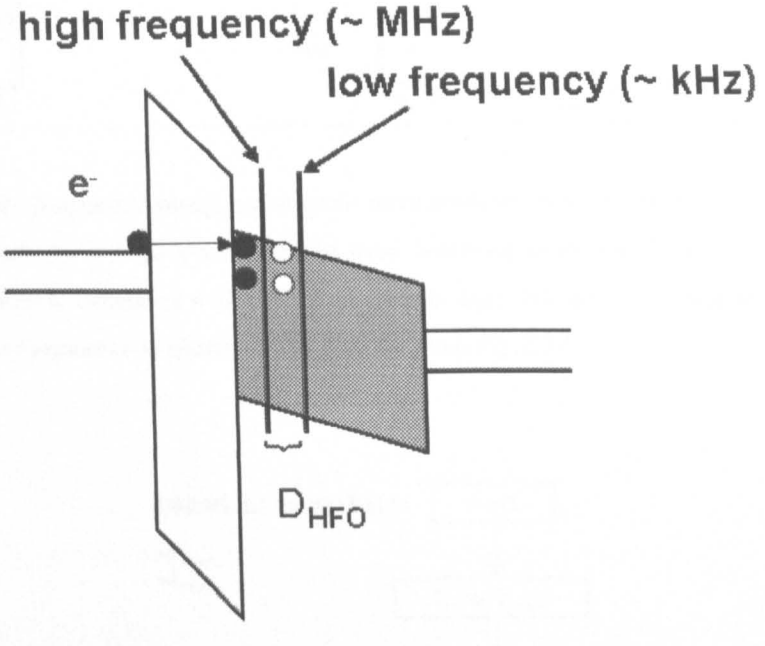
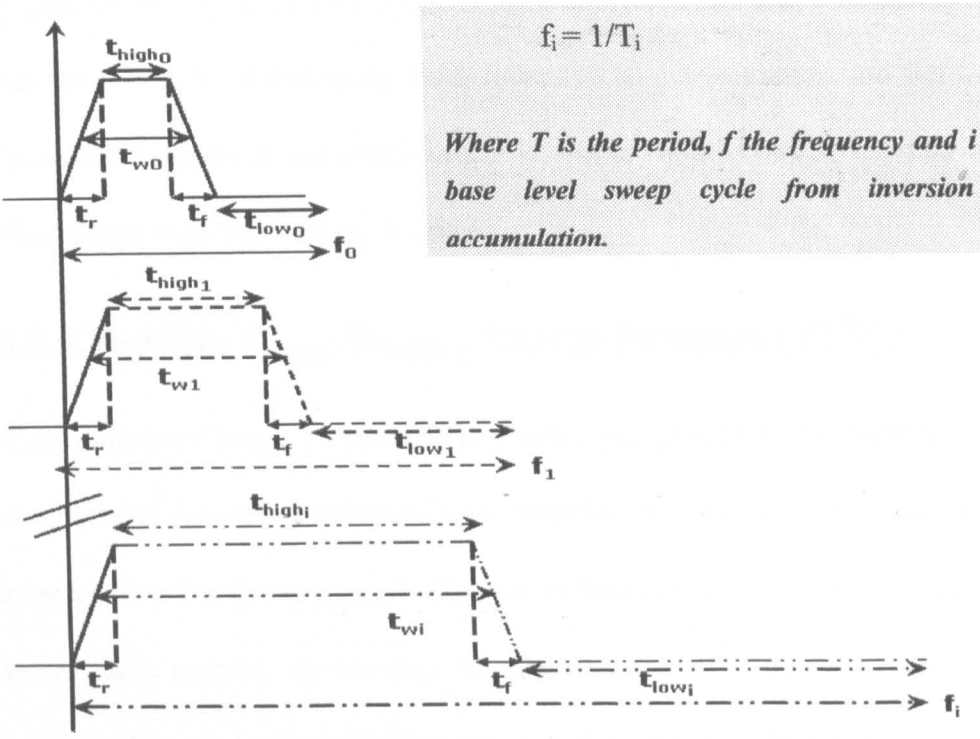


Fig.2.12 Band diagram showing that high-frequency only Si/SiO_2 interface trap are pumped, but at lower frequency traps deeper in the stack can also respond to the signal

Therefore, if the trap density sensed at high frequency is subtracted from the trap density measured at low frequency, a measure for the bulk trap density, "D_{HFO}", is obtained and this quantity is a measure of the bulk HfO_2 trap density. The measurement sequence is described in detail in Fig.2.13 and 2.14 and used in chapter 4 section 4.3.1 to investigate the channel length dependent time to breakdown.



$$f_i = 1/T_i$$

Where T is the period, f the frequency and i the base level sweep cycle from inversion to accumulation.

Fig.2.13 Variable frequency sweep pulse cycle measurement. A pulse with a fix duty cycle of 50% is applied at the gate for a base level sweeping from inversion to accumulation and the I_{CPmax} is taken, then the frequency is increased while the duty cycle is kept 50% and new base level sweep is taken. A full measurement sequence is shown in the flow diagram Fig.2.14

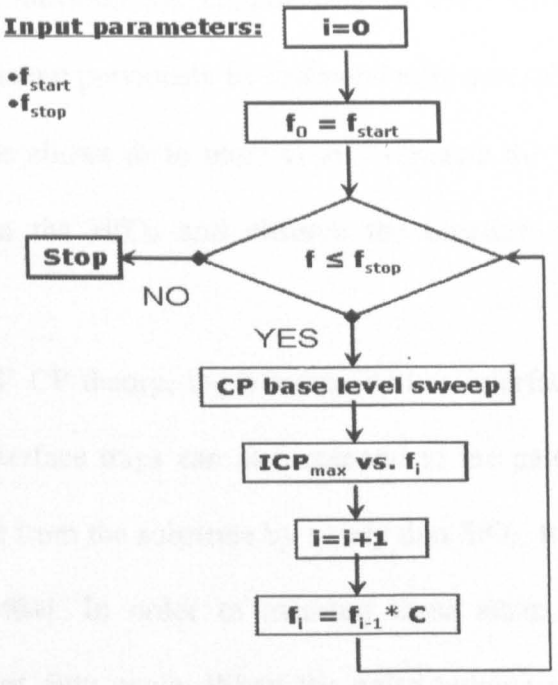


Fig.2.14 Flow diagram of the variable frequency charge pumping. C is a multiplication factor to increase the frequency after each base level sweep cycle from inversion to accumulation.

A gate pulse with a fix duty cycle of 50% and fixed rise and fall time is applied at the gate for a base level sweeping from inversion to accumulation and the I_{CPmax} is taken. Then the frequency is increased (i.e. $f_0 \rightarrow f_1 \rightarrow f_i$ in fig 2.13) while the duty cycle is kept 50% and new base level sweep is taken.

2.4.4 Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP)

Conventional Charge pumping (CP) technique **[Brug69]**, **[Groes84]** is a very sensitive method to determine substrate/gate dielectric interface state parameters such as the average density D_{it}, the energy distribution within the Si band-gap, the electron and hole capture cross section. By varying the charge pumping frequency, this technique also allows to characterize bulk defects close to the interface **[Bauza94]**, and traps in nitride layers separated from the substrate by a thin SiO_2 layer **[Paul94]**, **[Arrw05]**. In some recent papers, charge pumping was used in a similar way for characterizing HfO_2 bulk traps in SiO_2/HfO_2 stacks **[ZahM05]**, **[Ker04]**, **[Cart04]**, **[Degra03]**, **[Degra05]**.

Here we further develop the applicability of the variable frequency charge pumping technique presented previously by independently controlling the pulse low and high level timings. This allows us to more clearly separate the traps in the interfacial SiO_2 from the traps in the HfO_2 and observe the creation of new traps in both constituent layers.

In 'conventional' CP theory, D_T is interpreted as interface state density, but at low frequency, near-interface traps can also respond to the gate pulse. In particular, when HfO_2 is separated from the substrate by a very thin SiO_2, traps in the HfO_2 can be sensed **[Ker04]**, **[Cart04]**. In order to measure these states, a frequency scan is performed with constant duty cycle. When the pulse voltage is high (inversion, Fig. 2.15), traps are filled through tunneling in both the SiO_2 interface layer and the HfO_2 bulk. We will refer to the time that the pulse is in inversion as the "charging time".

When the pulse voltage is low (accumulation, Fig. 2.15), traps in the interface layer and HfO_2 bulk traps are emptied. We will refer to the time that the pulse is in accumulation as the "discharge time".

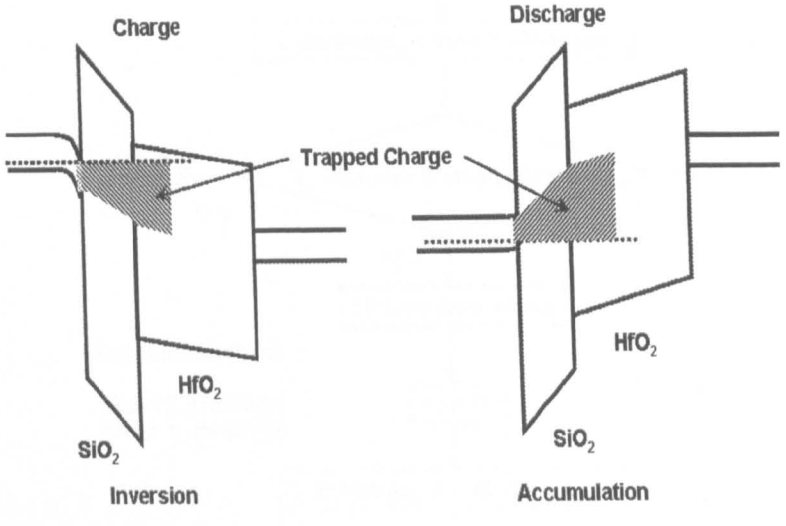


Fig.2.15.Schematic band diagram showing how charging and discharging occurs when VT^2CP is used for nMOSFETs.

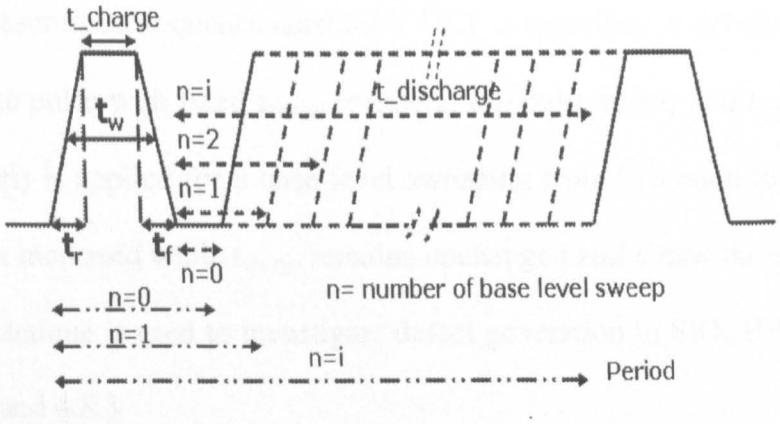


Fig.2.16 A gate pulse with fixed t_{charge} (=time at top pulse level) and $t_{discharge}$ (=time at base pulse level) is applied for a base level sweeping from inversion to accumulation. Then $t_{discharge}$ is increased while t_{charge} remains unchanged and a new base level sweep is taken.

With a simple frequency sweep at constant duty cycle, both the charging and discharging time of the bulk defects is changed simultaneously. Using VT^2CP, we will show that an independent control of charging and discharging time allows for an easier interpretation of the data in chapter 4 section 4.3.2. Note that in conventional CP theory,

all the interface states are sensed during the pulse voltage transients only. States deeper in the oxide are accessed during the constant high and low voltage parts of the pulse.

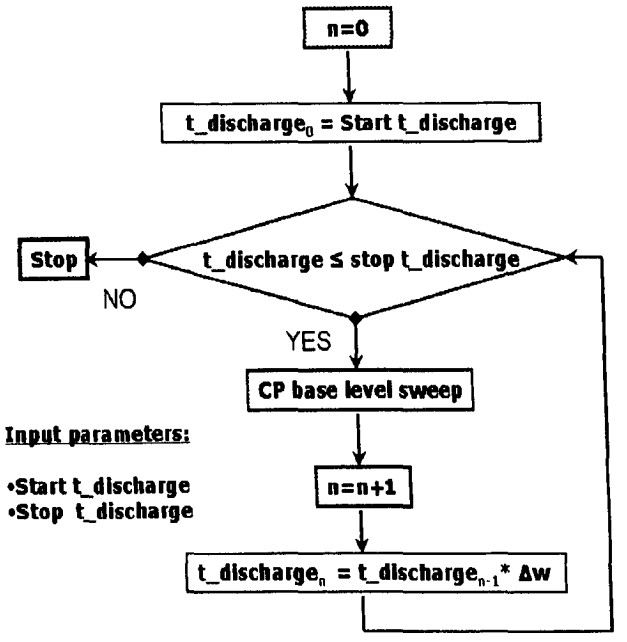


Fig.2.17 The measurement flow diagram used to monitor the I_{CP} current in SiO_2 / HfO_2 stack. Δw is the multiplication factor for increasing $t_{discharge}$ time after each loop until Stop t_discharge is reach.

The measurement sequence used for VT^2CP is described in detailed in Figs. 2.16 and 2.17: a gate pulse with fixed t_{charge} (=time at top pulse level) and $t_{discharge}$ (=time at base pulse level) is applied for a base level sweeping from inversion to accumulation. Then $t_{discharge}$ is increased while t_{charge} remains unchanged and a new base level sweep is taken. This technique is used to investigate defect generation in SiO_2/HfO_2 in chapter 4 sections 4.3.2 and 4.3.3

2.5 TIME DEPENDENT DIELECTRIC BREAKDOWN (TDDB)

Time Dependent Dielectric Breakdown, or TDDB, is one of the most heavily researched failure mechanisms in the semiconductor reliability community; it is a wear-out of the insulating properties of silicon dioxide in the CMOS gate leading to the formation of a conducting path through the oxide to the substrate. TDDB lifetime is strongly affected by the number of defects in the gate oxide produced during wafer fabrication. When a voltage is applied to the gate of a transistor, electrons and holes tunnel through the dielectric. Depending on the fluence, the applied gate voltage and the electric field, defects are generated in the insulating layer and act as electron traps. If a critical amount of defects is reached, these defects can form a percolation path **[Degra98]**. In the best case, such a percolation path leads to a small trap-assisted localized current and the performance of the transistor is almost not affected **[Kacz00]**. In the worst case, the dielectric completely loses its insulating properties and the transistor fails.

TDDB and the extrapolation of the time-to-breakdown to low operating conditions as a method to determine the gate oxide reliability is already known and used for many years. Lots of effort was spent in understanding and modeling the degradation, but ultimately the intrinsic reliability of thin SiO_2/SiON is expected to be safe **[Wuy02]**, **[Kacz04]**. The scaling of ultra-thin gate oxides below 2 nm EOT and the introduction of high-k gate dielectrics, however, raised the question of fundamental reliability limits. **[Gree01]**, **[Wilk01]**. In the past 5 years, major improvements in the understanding and the adaptation of measurement and analysis techniques were made to answer this question.

In this section we discuss the multiple aspects associated with Constant Voltage Stress (CVS) and Constant Current Stress (CCS) measurements. An overview of the various conduction mechanisms and initial device testing and the experimental procedure are described and we motivate the implementation of complementary measurements around CVS

2.5.1 Physics of breakdown

To understand the physics of breakdown, we must first examine the metal oxide semiconductor system. In a basic MOS transistor, the channel beneath the gate region is controlled by the voltage on the gate. In an n-channel transistor, a positive voltage on the gate causes the channel to invert, permitting charge flow from source to drain. In a p-channel transistor, a negative voltage on the gate causes the channel to invert, permitting charge flow from drain to source.

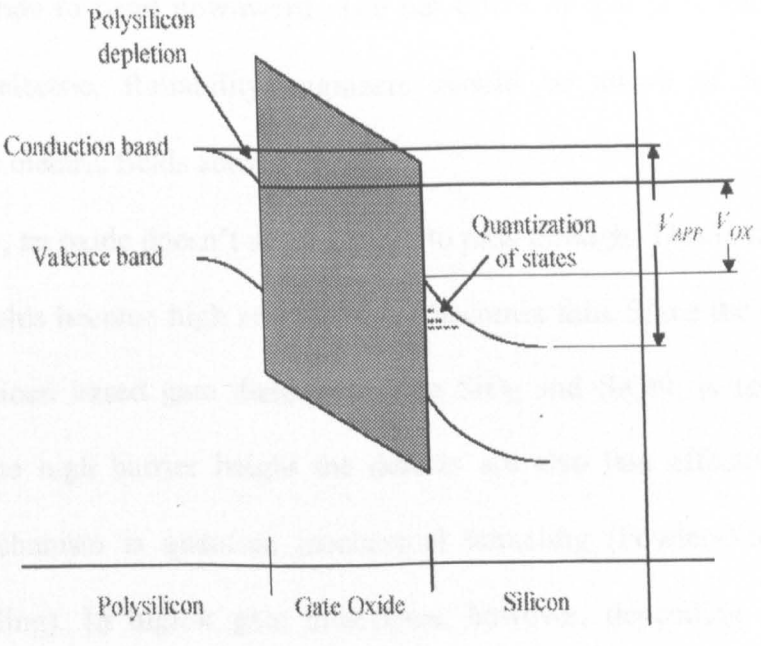


Fig.2.18 Energy band diagram for a gate oxide with a large applied electric field. Note that the voltage across the oxide (V_{OX}) is less than the applied voltage V_{APP}.

The voltage across the gate produces an electric field across the dielectric. It is this electric field that we need to understand to determine the reliability of the dielectric. Fig 2.18 shows an energy band diagram depicting the voltage across the dielectric. The electric field across the oxide is given by the simple equation $E = \frac{V_{ox}}{t_{ox}}$. Because today's semiconductor processes use ultra thin gate oxides, the electric field across the dielectric can be quite high. For high electric fields, there is a voltage drop due to the quantization of states in the silicon, and there is a voltage drop due to depletion of the polysilicon material (Fig 2.18). In a high electric field, the channel region becomes inverted. This means that individual electrons become caught in states immediately adjacent to the interface between the silicon and the oxide. As the states become fully occupied, the valence band and the conduction bands bend. As the field becomes even higher, the states in the polysilicon become depleted. These causes the valence and conduction bands to bend downward. The net effect of this is to reduce the voltage across the dielectric. Reliability engineers should be aware of this effect when calculating the electric fields across oxides.

Ideally, an oxide doesn't allow charge to pass through. This is not the case when the electric fields become high and the oxide becomes thin. Since the defect density of unstressed silicon based gate dielectrics, like SiO_2 and SiON, is relatively low and because of the high barrier height the defects are also less effective, the dominant transport mechanism is quantum mechanical tunneling (Fowler-Nordeim tunneling, Direct tunneling). In high-k gate dielectrics, however, depending on the dielectric composition, the material properties and the processing conditions, initial defect densities are expected to be higher than for SiO_2 and due to the significantly lower barrier height trap-assisted conduction is, especially at elevated temperatures, more

likely to occur. However, independent of the material, during electrical stress defects are created in the dielectric. An increase of the defect density leads to SILC **[Olivo88]**, discrete local conduction paths **[Kau05]** and when a critical defect density is reached, it leads to ultimate breakdown of the dielectric **[Degra98]**.There are several mechanisms that allow charge to pass through the oxide: Fowler-Nordeim tunneling, Direct tunneling, and Trap Assisted tunneling and Poole-Frenkel mechanism.

2.5.1.1 Fowler-Nordeim Tunneling

Fowler Nordeim tunneling is a quantum mechanical tunneling process where the electrons can penetrate through the oxide barrier into the conduction band of the oxide (Fig.2.19) which has been studied extensively in Metal-Oxide-Semiconductor structures where it has been shown to be the dominant current mechanism especially for thick oxides.

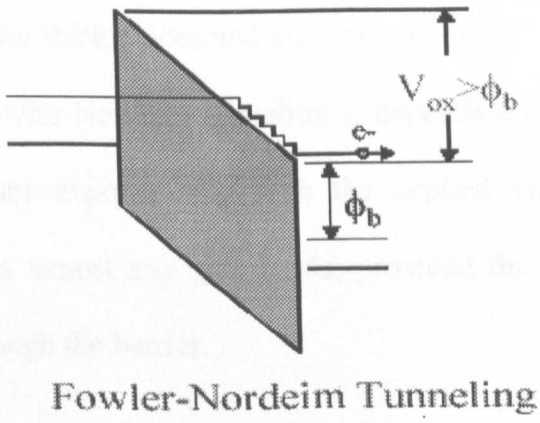


Fig.2.19 Fowler-Nordeim Tunneling diagram

The basic idea is that quantum mechanical tunneling from the adjacent conductor into the insulator limits the current through the structure. Once the carriers have tunneled into the insulator they are free to move within the valence or conduction band of the insulator. The calculation of the current is based on the WKB approximation

yielding the following relation between the current density, J_{FN}, and the electric field in the oxide, $\mathcal{E}_{ox}$:

$$J_{FN} = C_{FN}\mathcal{E}_{ox}^2 \exp_{ox}\left(-\frac{4}{3}\frac{\sqrt{2m_{ox}^*}}{q\hbar}\frac{(q\phi_B)^{3/2}}{\mathcal{E}_{ox}}\right) \tag{2.29}$$

where ϕ_B is the barrier height at the conductor/insulator interface.

To check for this current mechanism, experimental *I-V* characteristics are typically plotted as $\ln(J_{FN}/\mathcal{E}_{ox}^2)$ versus $1/\mathcal{E}_{ox}$, a so-called Fowler-Nordheim plot. Provided the effective mass of the insulator is known (for SiO2, $m_{ox}^* = 0.42\ m_0$) one can then fit the experimental data to a straight line yielding a value for the barrier height.

It is this type of measurement, which has yielded experimental values for the conduction band difference between different metals and silicon dioxide. It is important to note that carriers must tunnel through the insulator, which requires:

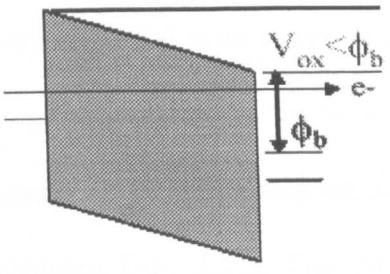
$$\mathcal{E}_{ox}d \geq \phi_B \tag{2.30}$$

as is typically the case for thick oxides and high electric fields.

In summary Fowler-Nordeim tunneling is dependent on the voltage across the gate oxide; it increases exponentially with the applied voltage. Fowler Nordeim tunneling can occur in almost any gate oxide, provided the voltage is sufficient for electrons to tunnel through the barrier.

2.5.1.2 Direct Tunneling

Direct tunneling is also a quantum mechanical tunneling process (Fig.2.20). Direct tunneling is a phenomenon that is important to understand in ultra thin oxides. It occurs when electrons tunnel through the gate oxide region directly from the gate to the channel region.



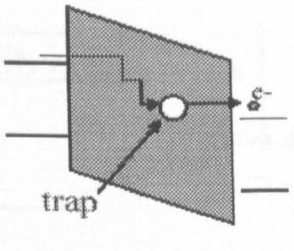
Direct Tunneling

Fig.2.20 Direct Tunneling diagram

Direct tunneling is dependent on the thickness of the gate region; it increases exponentially as the thickness of the oxide decreases. Direct tunneling is relatively less sensitive to the electric field across the gate oxide

2.5.1.3 Trap Assisted Tunneling

The other tunneling mechanism that can occur in a gate oxide region is trap assisted tunneling. Trap assisted tunneling occurs when electrons tunnel through the oxide into traps (empty bonding sites on silicon dioxide molecules) and then from the traps into the silicon, as illustrated in Fig.2.21.



Trap Assisted Tunneling

Fig.2.21 Trap Assisted Tunneling diagram

2.5.1.4 Poole-Frenkel mechanism

The expression for Fowler-Nordheim tunneling implies that carriers are free to move through the insulator. Whereas this is indeed the case in thermally grown silicon

dioxide it is frequently not so in deposited insulators, which contain a high density of structural defects. Silicon nitride (Si_3N_4) is an example of such material. The structural defects cause additional energy states close to the band edge, called traps. These traps restrict the current flow because of a capture and emission process, thereby becoming the dominant current mechanism Fig. 2.21a The current is a simple drift current described by

$$J = qn\mu\mathcal{E}_N \tag{2.31}$$

while the carrier density depends exponentially on the depth of the trap, which is corrected for the electric field

$$n = n_0 \exp\left[-\frac{q}{kT}\left(\phi_B - \sqrt{\frac{q\mathcal{E}_N}{\pi\varepsilon_N}}\right)\right] \tag{2.32}$$

The total current then equals:

$$J_{PF} = qn_0\mu\mathcal{E}_N \exp\left[-\frac{q}{kT}\left(\phi_B - \sqrt{\frac{q\mathcal{E}_N}{\pi\varepsilon_N}}\right)\right] \tag{2.33}$$

The existence of a large density of shallow traps in CVD silicon nitride makes Poole-Frenkel emission a frequently observed and well-characterized mechanism.

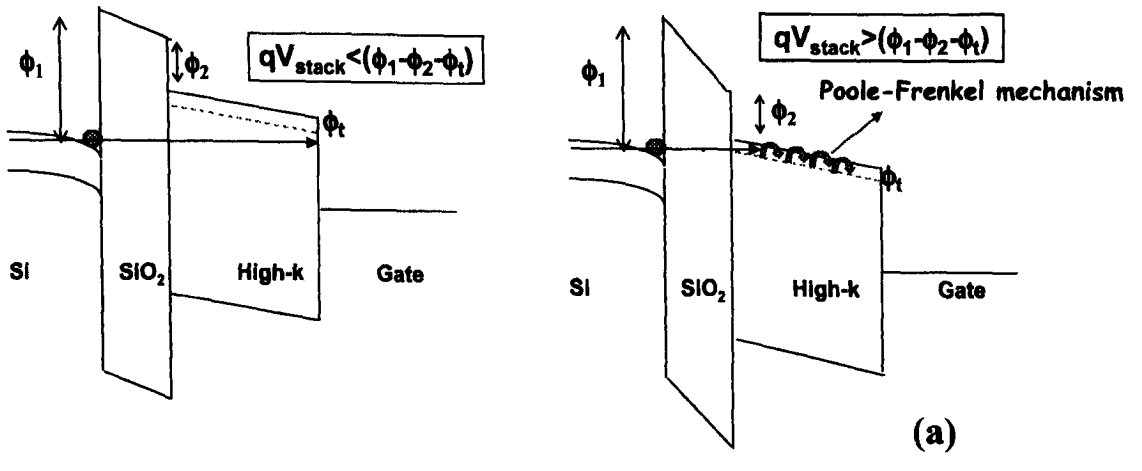


Fig 2.21a. Poole-Frenkel mechanism

2.5.2 Techniques of Breakdown Tests

Generated defects affect a number of device characteristics, such as transconductance, threshold voltage, carrier mobility, leakage current and noise level, resulting in long-term parameter drift and eventually device failure. Accurate and reliable measurements are the foundation of any kind of electrical characterization.

There are two commonly used methods to stress a MOS structure and to determine the time-to-breakdown (t_{BD}).

Constant current stress (CCS): a certain fixed current density is forced through the gate and the gate voltage is measured versus time (Fig. 2.22).

This technique is mainly used for SiO_2/SiON gate dielectrics with oxide thickness larger than ~6 nm **[Nig98]**. During FN-stress, the electrons enter the conduction band of the dielectric and move non-ballistically towards the anode.

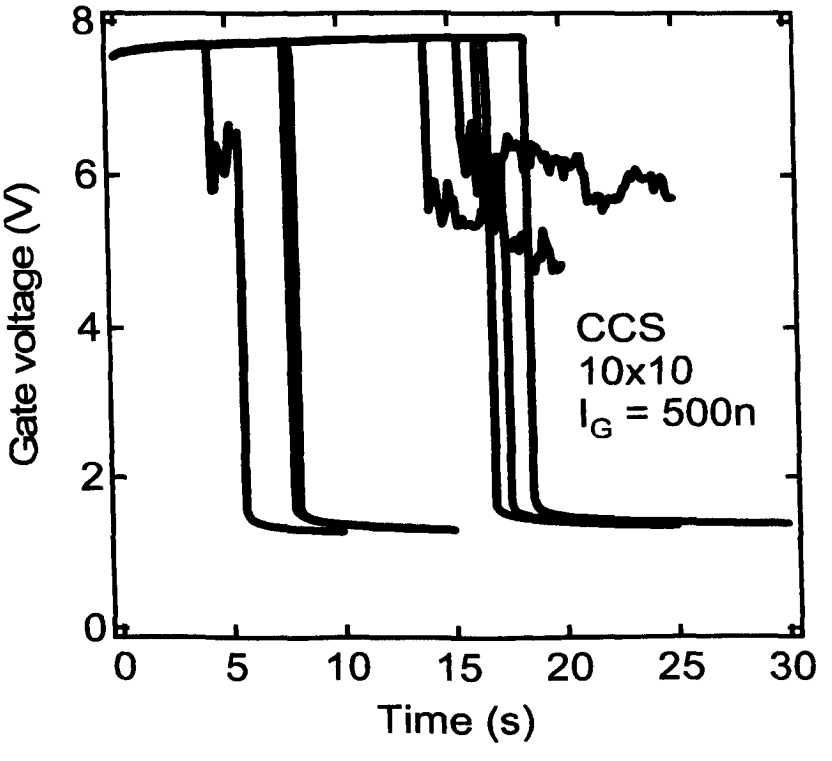


Fig.2.22 During CCS the gate voltage is recorded versus the time. The field over the dielectric remains constant and at breakdown the gate voltage collapses.

The degradation is therefore determined by the field in the oxide and not so much by the applied gate voltage. Throughout CCS the cathode field stays constant, and in case of charge trapping only the gate voltage will vary.

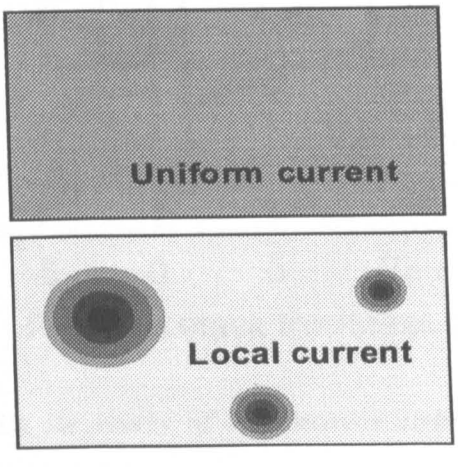


Fig.2.23 A defect free dielectric will be stressed uniformly during CCS (top). The formation of conduction paths leads to high local currents, while the remaining area is not stressed anymore (bottom).

For sub-6 nm SiO_2/SiON dielectrics, however, the use of CCS is not applicable anymore. Electrons tunnel ballistically, and the energy at the anode determines degradation **[DiMa93]**, **[Nig98]**. Therefore, it is more meaningful to fix the gate voltage instead of the oxide field. Furthermore, in very thin SiO_2/SiON layers and in high-k dielectrics, traps can create localized conduction paths. The current is flowing through a number of small spots (Fig.2.23) and this leads to a reduction of the gate voltage during CCS.

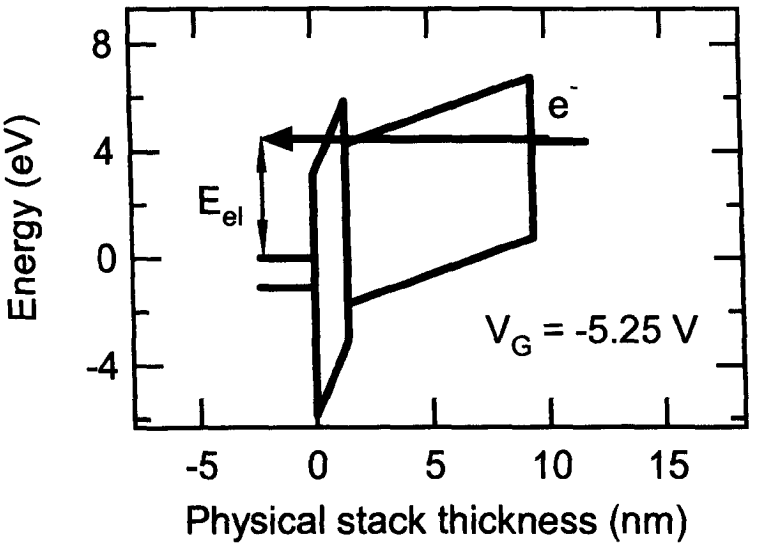


Fig.2.24 For thinner dielectrics the energy of the electrons arriving at the anode determines degradation.

Constant voltage stress **(CVS):** a constant voltage is applied to the gate and the gate current is measured versus time. Created traps result in an increase of the gate current and after breakdown typically compliance of the instrument is reached. CVS is applied when the degradation is determined by the applied gate voltage, rather than the applied field (Fig.2.24). This is the case for direct tunneling stress at low gate voltages and FN-stress with quasi-ballistic transport. Note that even in the presence of localized conduction paths the stress of the remaining area is unaffected.

Also, detecting breakdown is much easier when using CVS. The current depends exponentially on the gate voltage and at breakdown significant current increase is observed (Fig.2.25).

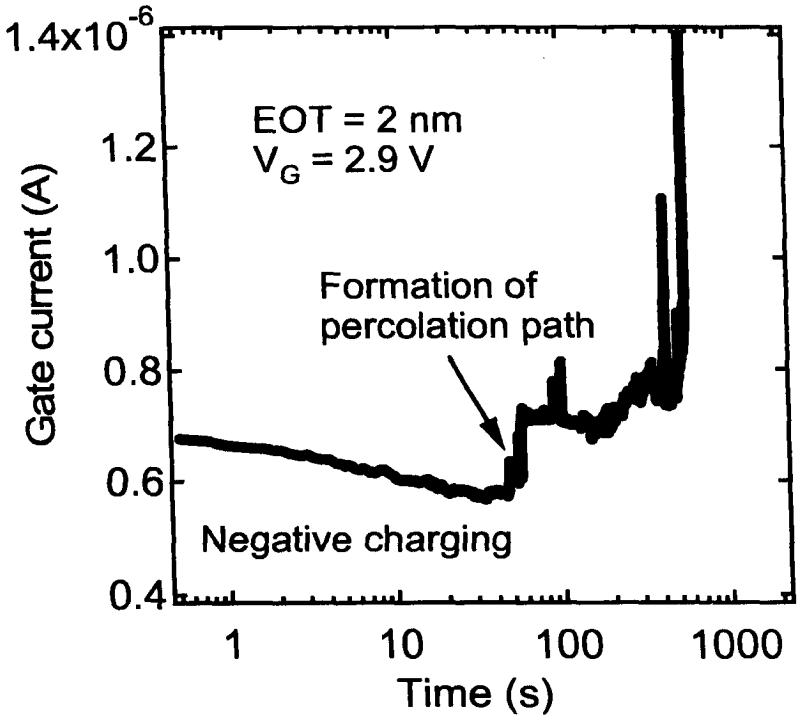


Fig.2.25 Charging during CVS leads to slight increase/decrease of the gate current and after the formation of a percolation path significant current increase is observed.

Even the creation of single traps can be detected, when studying the small current increase in the range of pA to μA due to the formation of localized conduction paths **[Cell02]**, **[Kacz04]**, **[Sue04]**. For CCS on the contrary, breakdown of the dielectric causes a reduction of the gate voltage of only a few mV. High-k gate dielectrics usually contain large amounts of initial defects **[Ker03A]**, and after applying a voltage to the gate these defects will trap charges. Depending on the dielectric material and the gate bias polarity, this results in a buildup of either positive or negative net charge.

2.5.3 Soft Breakdown and Hard Breakdown

Traditionally, TDDB measurements on thick (> 5nm) SiO_2/SiON gate dielectrics were performed using CVS (CCS) and monitoring the gate current (voltage) versus time (Fig.2.26).

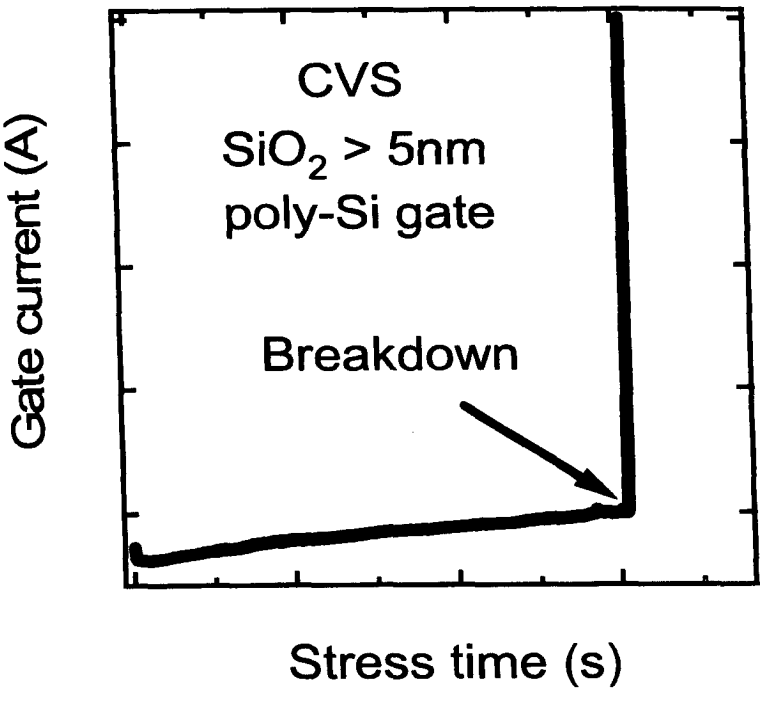


Fig.2.26 For thicker SiO_2 dielectrics under CVS measurements clear hard breakdown is observed and the determination of t_{BD} is simple.

As soon as a percolation path is created, the current (voltage) had increased (decreased), the measurement was stopped immediately and the time-to-breakdown (t_{BD}) recorded. This kind of measurement was ideally suited for real-time trigger and automatic data acquisition. The t_{BD} distributions of a statistically relevant number of devices were then used to extrapolate the lifetime down to operation conditions. The first complications appeared on thinner SiO_2/SiON dielectrics below 5 nm, when during CVS a "soft" breakdown before the final hard breakdown was observed (Fig.2.27) **[Sak98]**, **[Sue04]**.Typically the t_{BD} was triggered then at the appearance of SBD, although transistors with soft breakdown would still be functioning.

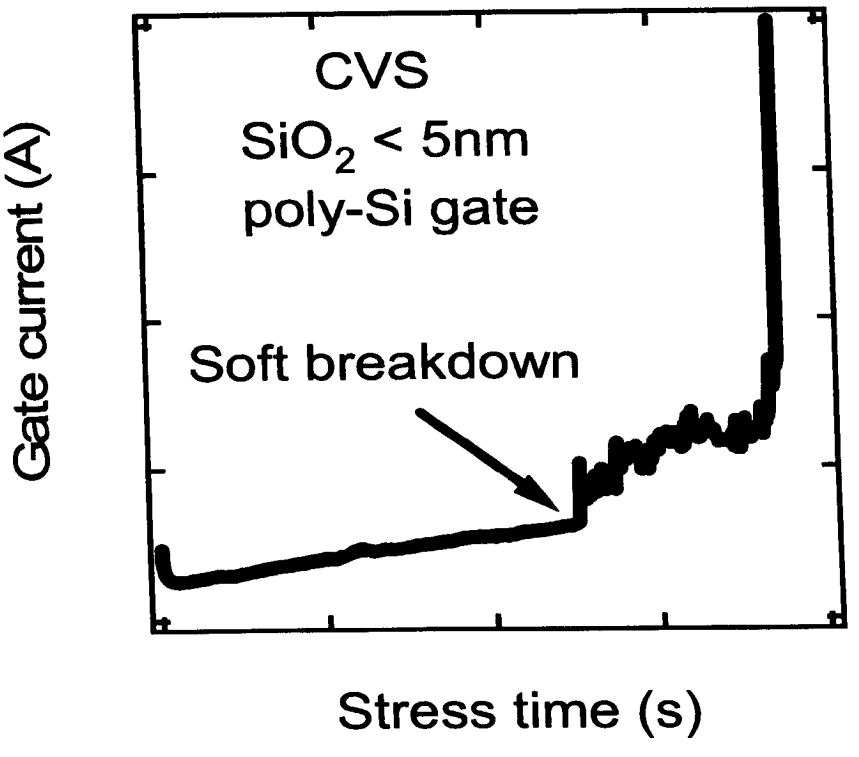


Fig.2.27 For SiO_2 dielectrics below 5 nm current fluctuations and noise increase can be observed before the final hard breakdown.

Nowadays, there exist numerous gate dielectrics and gate electrodes. Depending on the targeted application, ultra-thin SiON, Hf-Silicates and several high-k's are combined with poly-Si, FUSI or metal gates. Even the substrate is not necessarily Si anymore but can be replaced by SiGe or Ge. For each gate stack the degradation behavior under CVS varies, making the determination of t_{BD} an art on its own.

2.5.3.1 Oxide Breakdown

In Fig.2.28/2.29 we show two typical I~t traces recorded on nMOS transistors for substrate injection during CVS stress. Fig.2.28 is for an HfSiON/poly-Si gate stack of 1.38 nm EOT **[Kau05]**, whereas data in Fig.2.29 were measured on a 1.9 nm EOT SiON/HfO_2/TaN stack.

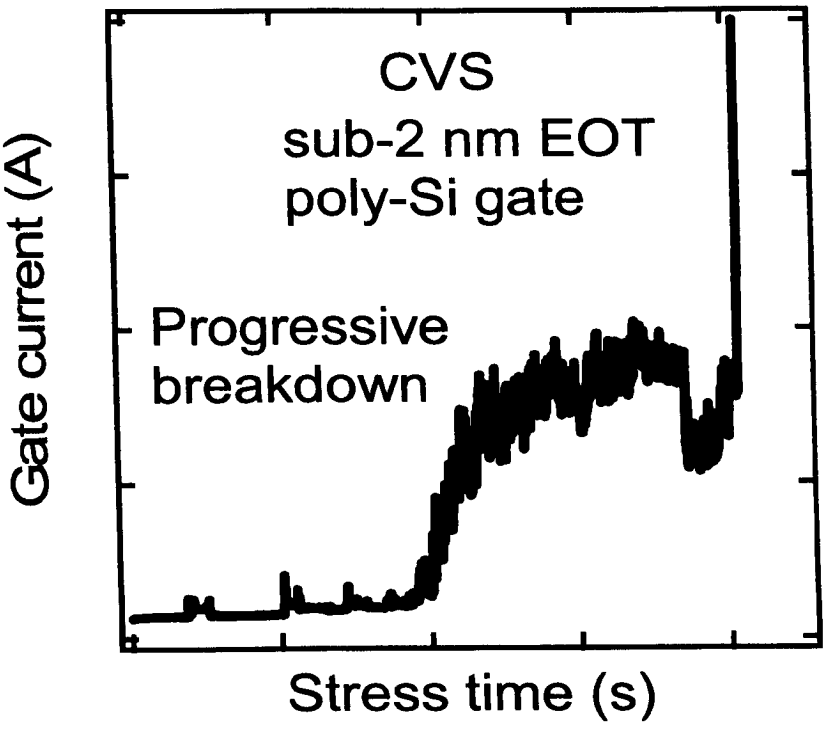


Fig.2.28 For ultra-thin dielectrics with poly-Si electrodes the formation of a percolation path is followed by a wearout phase.

Obviously, depending on the gate stack and the experimental design (device dimensions, temperature, injection polarity), the appearance of the measured gate current can strongly vary. The degradation can be classified into various phases, and each phase can be pronounced to a different degree in terms of magnitude and duration. Moreover it is possible to directly link each stage with an event in the dielectric (Fig. 2.28-2.29 and 2.30):

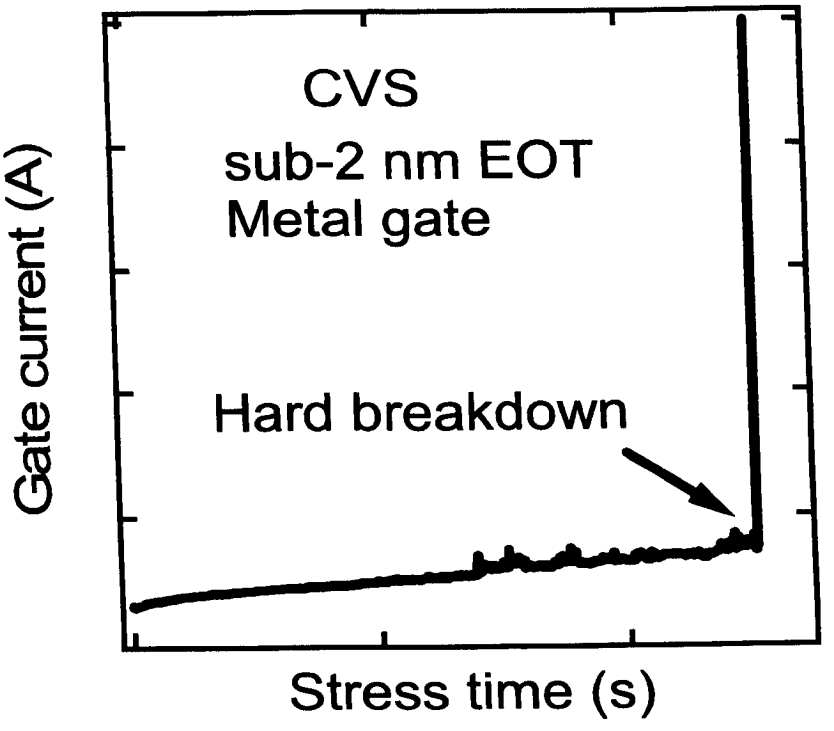


Fig.2.29 With metal gate electrodes the time of progressive wearout is strongly reduced and hard breakdown is the main failure mechanism.

As shown in Fig.2.30, initially, in the unstressed fresh device no defects are created yet (A) and the dominating conduction mechanism is direct or Fowler-Nordheim tunneling (assuming no initial defect densities). After continued stress, single defects are created in the oxide (B). These defects can act as traps and give rise to a small, localized current, typically referred to as classical SILC. Depending on the physical thickness of the dielectric, the barrier height and the position of the trap, the current through a *single trap* conduction path strongly varies.

It was shown that for gate dielectrics with $t_{ox} < 2$ nm, currents through single trap conduction paths can be as large as several hundred nA. Moreover, by capturing an electron the trap can be neutralized, switching "off" the conduction path again. **[Degra05A]**.With an increasing number of defects, the probability of creating a percolation path with 2 or more traps is raising (C). If conductive, for the same gate

stack, such a *multiple trap* conduction path can carry currents much larger than a single trap path. Since the amount strongly depends on the physical thickness and the appearance on the device area, temperature and gate voltage, in literature multiple terms can be found to describe multiple trap conduction paths: anomalous SILC **[Ielmi02]**, micro-breakdown **[Cell02]**, **[Degra05A]**, pre-breakdown **[Degra01]** and B-mode SILC **[Oka94]**.

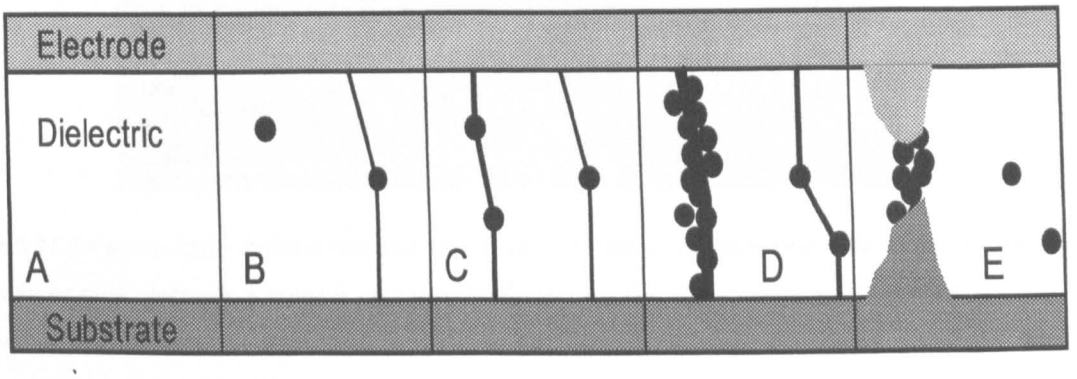


Fig.2.30 The degradation of gate oxides under CVS can be classified into several stages. It however depends strongly on the materials and thicknesses to which extends each stage is pronounced.

Large local currents along conduction paths whether single trap paths in ultra-thin or multiple trap paths in thicker dielectrics, lead to accelerated degradation. More defects are created along the existing conduction path (D), resulting in a growth and *wearout* of the breakdown spot **[Kacz04]**. This phase of progressive wearout is commonly referred to as *digital soft breakdown* **[Sak98]** and characteristic is the rapid switching between multiple current levels. With continued stress the positive feedback mechanism leads to further trap generation (E) and temperature increase, resulting in significant current and noise increase.

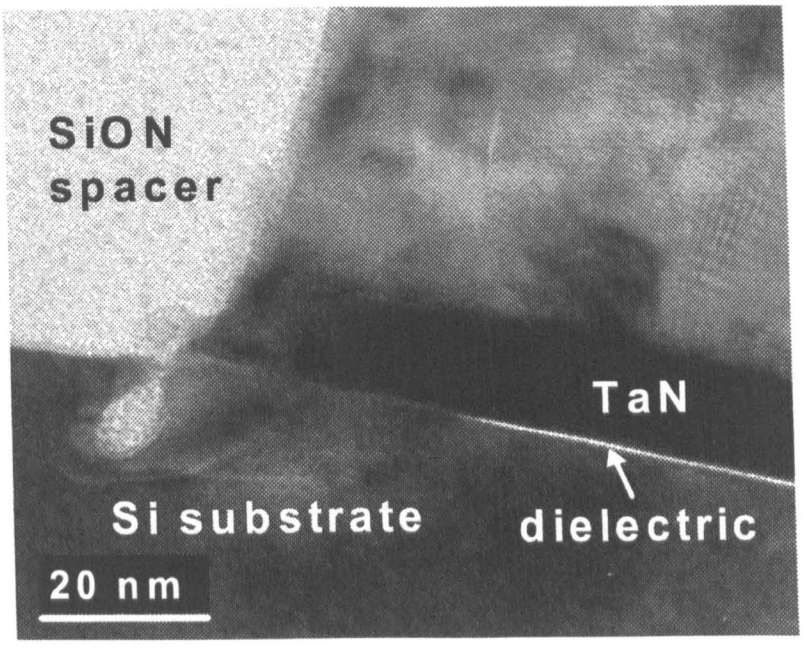


Fig.2.31 In some cases breakdown can create large physical damage and lead to migration of silicon or gate electrode materials through the dielectric. The size of such a breakdown spot can reach up to few hundred nm.

This analog soft breakdown can even lead to the migration of channel or gate electrode material into the dielectric (Fig.2.31) **[PeyK04]**. The current can rise up to several mA and is eventually limited by series resistance **[Kacz04]**. If hard breakdown occurs, the dielectric looses its insulating properties and the current is increasing up to compliance. For completeness is has to be added, that in some rare cases a full deactivation of a breakdown path can be observed. In Figure 2.32 we show such an example, where a conduction path is wearing out, until a current level of more than 500 μA is reached.

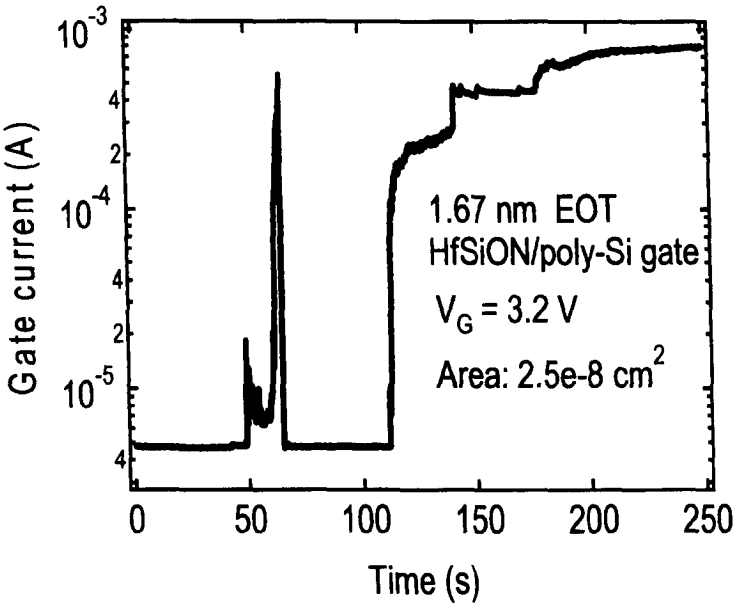


Fig.2.32 High current densities through a percolation path can lead to electro migration of the gate electrode and a shutdown of the breakdown path. The remaining area is unaffected from this event and the gate current returns to the original level.

Suddenly the current returns to the initial value of the unstressed device with exact the same noise level. This can be explained by the melting of the gate electrode due to the large local current and the creation of a void, isolating the BD path.

2.5.4 Advanced measurement sequences

Combining CVS with complementary measurements can drastically increase the accuracy of the experiment and provide additional learning. Especially when studying new high-k gate dielectrics these supplemental findings can sometimes be more important than the actual CVS stress. In the following sections we introduce those measurements, which can be either implemented directly before and after the CVS or executed in loops by interrupting CVS in regular intervals.(Fig.2.33)

The experimental conditions of the complementary measurements have to be chosen such that no additional degradation occurs in the dielectric and the determined t_{BD} is still correct. This can be achieved if any applied voltage is significantly smaller

than the actual stress condition. Short measurements with maximum applied voltage 0.5 V below the stress condition do not impact the stress or the t_{BD}

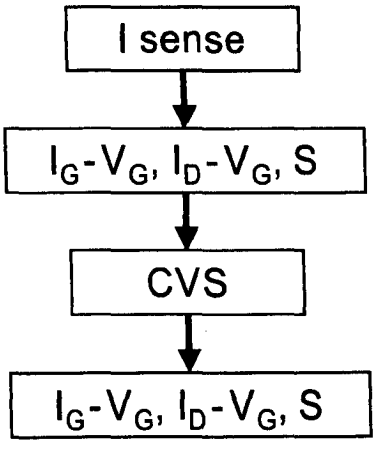


Fig.2.33 To gather information about the initial device properties, the degradation process and the breakdown spot location (S), we embed the CVS in a sequence of measurements.

On the other hand, partially severe degradation and breakdown occurred when switching to another instrument (e.g. pulse generator, CV-meter) or reconnecting the device after for example varying the temperature of the thermo chuck. Those effects are due to electrostatic discharge and physical damage of the contacts respectively and have to be avoided. This can be done using a switching matrix, but requires additional hardware, complicates the experimental setup and reduces the current resolution. Therefore we limit ourselves to measurements that can be done with the same standard parameter analyzer, which performs the actual stress.

In summary, it can be stated that during degradation of the dielectric two main mechanisms proceed simultaneously: 1) the random creation of defects and thus the formation of percolating paths and 2) the wearout of those localized paths. The appearance of these events during an actual measurement, however, strongly depends on the gate stack and the experimental design itself. In a small area device with low background current a 10 μA current increase through a conduction path may be

misinterpreted as a HBD, whereas on a larger area of the same gate stack at high temperatures it would nearly disappear in the background noise and can be interpreted as SILC instead of breakdown.

In Chapter 4 section 4.4 an abrupt hard breakdown where a sudden current increase of several orders of magnitude up to the current compliance of the measurement instrument, or until the current is limited by series resistance is reported for thin high-k gate stacks. Hard breakdown can represent the final destruction after the occurrence of soft breakdown or progressive breakdown; and it is also the dominating failure mode when using metal gate electrodes **[Kau05A].**

2.6 INSTRUMENT AND TECHNICAL OVERVIEW

In this section an overview of the instrument available, the measurement set-up and schematic of the pulsed I_D~V_G, Charge Pumping and Time Dependent Dielectric Breakdown will be given. Also briefly discussed in section 2.2.2 the change in the schematic of the pulse I_D~V_G will be explained.

2.6.1 Measurement set-up

2.6.1.1 Pulse I_D~V_G

An HP81101A pulse generator is used to apply pulses to the gate of the transistor and pure DC Supply voltage is used to apply a small voltage at the drain (100mV)

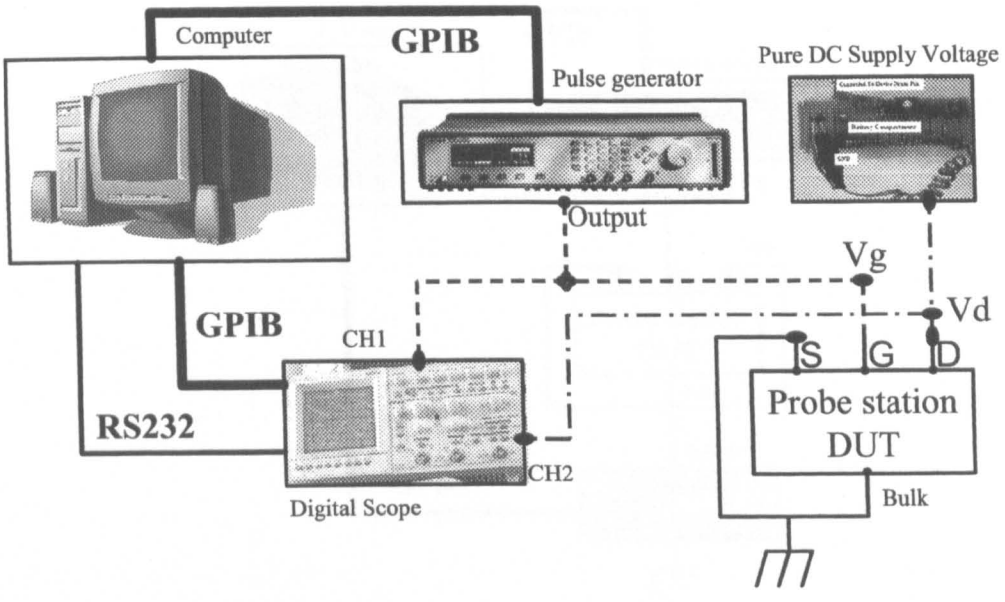


Fig.2.34 Schematic setup showing the instruments used to carry out the pulsed I_D~V_G and their connection.

All instruments apart from the DC voltage are controlled by a workstation via a GPIB connection and the data are collected via a RS232 connection from the oscilloscope as shown in Fig.2.34. Measurements are performed using a probe station; standard coaxial cables are used to connect the Device Under Test (DUT) to the pulse generator, pure DC voltage source, and the oscilloscope. Since the shields of all oscilloscope inputs are connected internally, all instruments are referenced to a common ground. In order to avoid ground loops, no extra ground connection are provided between the various instruments

2.6.1.2 Charge Pumping

An HP81101A pulse generator is used to apply pulses to the gate of the transistor and the current is measured by a Keithley K4200 analyzer which provide a very high resolution and current down to fento amp can be measured, an illustration of the setup is shown in Fig 2.35

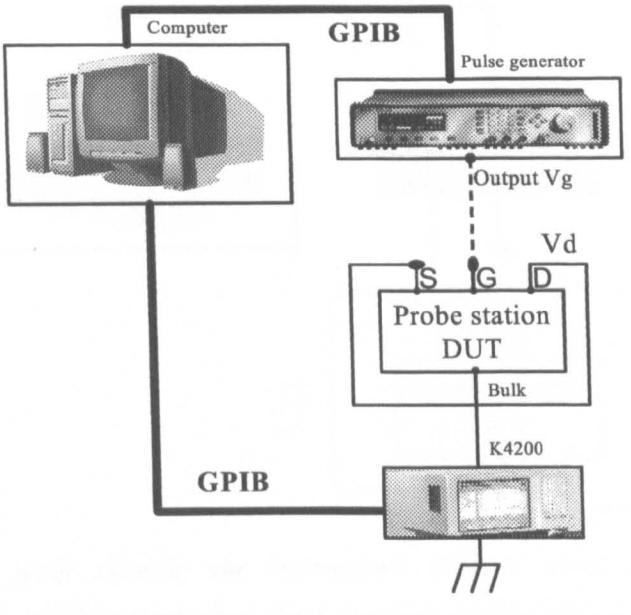


Fig.2.35 Schematic setup showing the instruments used to carry out the Charge Pumping and their connection.

The characteristics of the various instruments are given in more details in section 2.6.2. Good knowledge of the instrument's capabilities is essential to the experimentalist in order to maximize the performance of the measurement set-up and guarantee the reliability of the obtained result. Moreover, it allows for estimation of the accuracy on the measurement outcome, which is an essential aspect of any experiment technique

2.6.1.3 Time dependent dielectric breakdown

K4200 is used to monitor the current and apply the required constant voltage to stress the DUT and data are collected via a GPIB connection and a software running under UNIX as shown in Fig.2.36

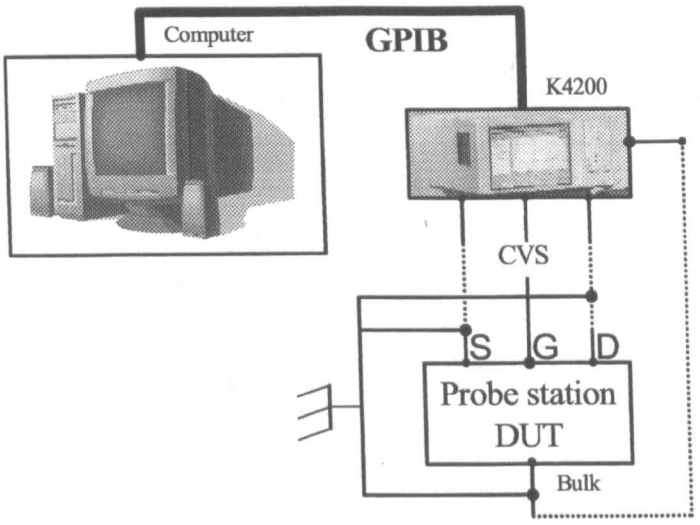


Fig.2.36 Schematic setup showing the instruments used to carry out TDDB and their connection.K4200 monitor the current from the 4 terminals and the apply a Constant Volatge Stress (CVS) at the gate.

2.6.2 Instruments characteristics

2.6.2.1 Pure DC Voltage Supply characteristics

As briefly discussed previously a pure DC Supply voltage is used instead of a normal power supply, in this section we will explain why we use dry cell battery as a pure DC Supply voltage and its characteristics.

The circuit of the pulsed I_D~V_G introduced in section 2.2.2 from the literature (Fig.2.5) was changed to the setup shown in Fig. 2.7 by adding a voltage divider to apply a small DC bias to the drain.

If the circuit in the literature is used with a normal power supply it was found that the output V_D trace has an overshoot, although the pulse generator gives a near waveform as shown in Fig.2.37.

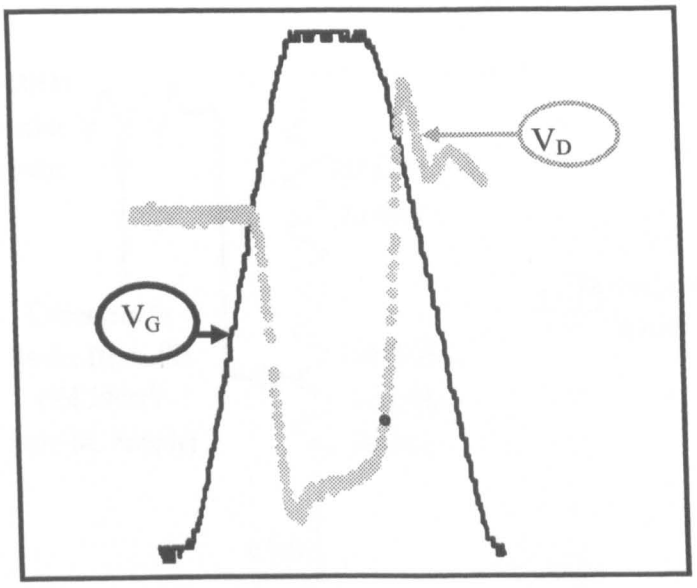


Fig.2.37 The input pulse V_G and output waveform V_D with overshoot.

Further investigation found that the DC power supply Keithley 487contain some ripple about 25mV peak-to-peak, this amount of voltage is 25% of test supply voltage (V_D) and with such a high ripple involved, the test result is significantly affected. Two solutions have been proposed: building a filtering or smoothing circuit to remove the ripple voltage or using a pure DC power supply. Building a filtering or smoothing circuit for 100mV power supply is very complicated and time consuming.

A simple solution is using a dry cell battery. The dry cell batteries are able to generate pure DC power. However, most of the batteries available in the market are rated at 1.5V DC for primary battery and 1.2V DC for secondary battery (rechargeable). In order to get the required voltage (100mV), a voltage divider has been used. This method is far easier and less time consuming. Fig.2.38 shows the circuit and the required 100mV can be obtained by adjusting 100Ω variable resistor. The 500Ω variable resistor can be used to adjust the load resistance. With the test setup shown in Fig 2.7, the overshoot in V_D has been removed successfully as shown in Fig 2.39

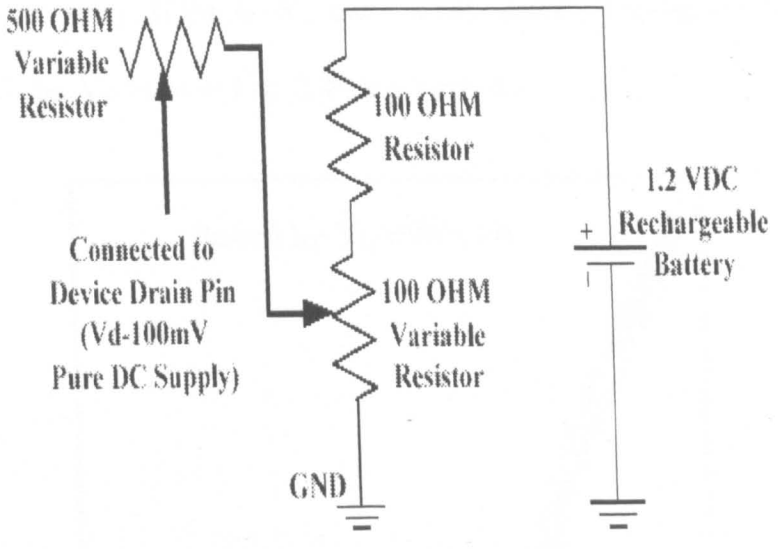


Fig.2.38 Divider schema using a rechargeable battery of 1.2 V

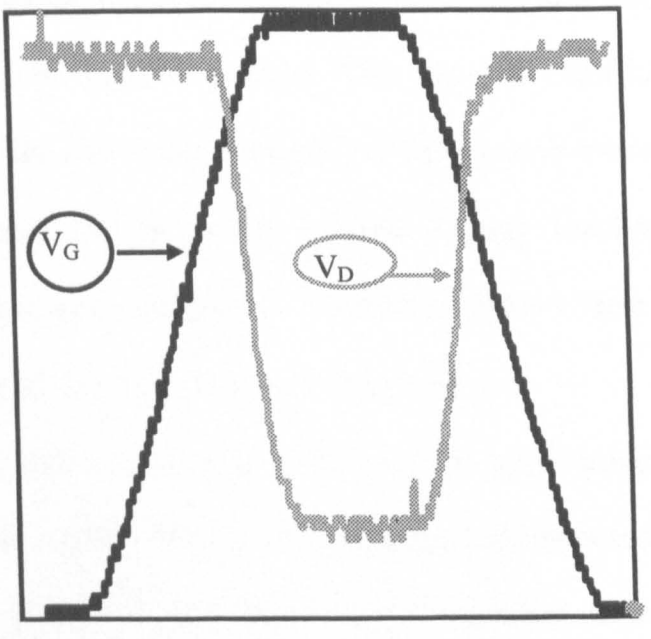


Fig.2.39 Input pulse V_G and output waveform V_D without overshoot and containing some noise

Noise was a problem to a successful experiment. In Fig.2.39, we can observe that V_G and V_D are noisy. If the I_D~V_G trace is calculated by using the data in Fig.2.39 and using Eq.2.2, I_D~V_G trace in Fig. 2.40 is obtained.

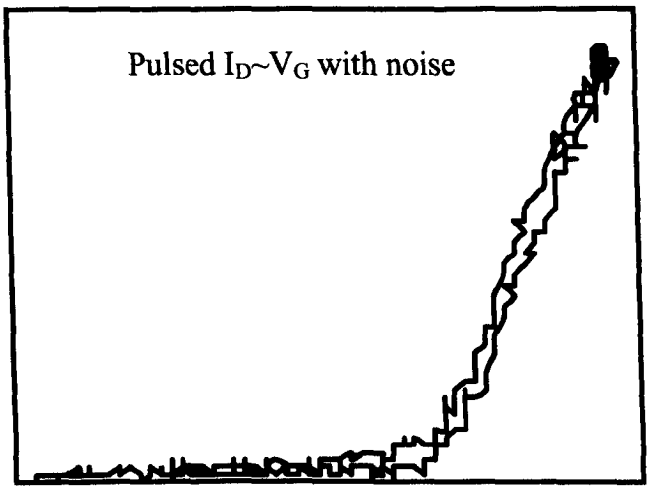


Fig.2.40 Pulsed I_D~V_G with noise.

The noise in the measurement may come from electrical line, Radio Frequency noise, white noise, and scope digitization errors. This makes the data hard to analyze and leads to inaccuracy. The noise can be suppressed by a sample averaging function, which is built-in for most of the modern Digital Storage Oscilloscope (DSO). Unfortunately the oscilloscope used in the experiment doesn't have an averaging function built in. We will discuss how data can be averaged next.

It is possible to extract the signal from random noise using an averaging technique. This technique exploits the fact that the noise superimposed on a signal is random, and that over a period of time, the average of a random noise signal is zero. Some modern DSO fitted with averaging function in the front panel where the user can utilized this function on the field. This will give a good result immediately **[Ross94]**. For those DSO without averaging function built in. The user can download the data to Microsoft Excel or any other software for further processing. The averaging signal can be obtain from the averaging function built in Excel. This analysis tool and its formula

project values in the forecast period, based on the average value of the variable over a specific number of preceding periods. Each forecast value is based on the following Eq. (2.34)

$$F_{(t+1)}=\frac{1}{N}\sum_{j=1}^{N} A_{t-j+1} \tag{2.34}$$

where

- N is the number of prior periods to include in the moving average
- Aj is the actual value at time j
- Fj is the forecasted value at time j and t the time.

A moving average provides trend information that a simple average of all historical data would mask. When the Eq.(2.34) is used, the curve can be smoother as shown in Fig.2.41 compared with Fig.2.40. The averaging of the data suppresses the noise.

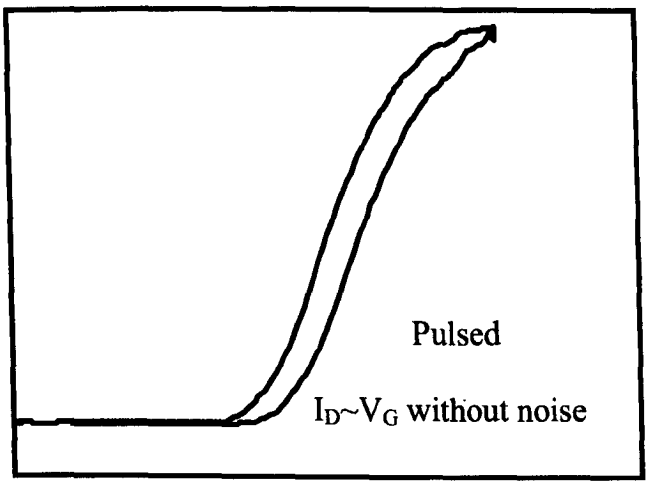


Fig.2.41 I_D vs V_G curve with noise suppressed.

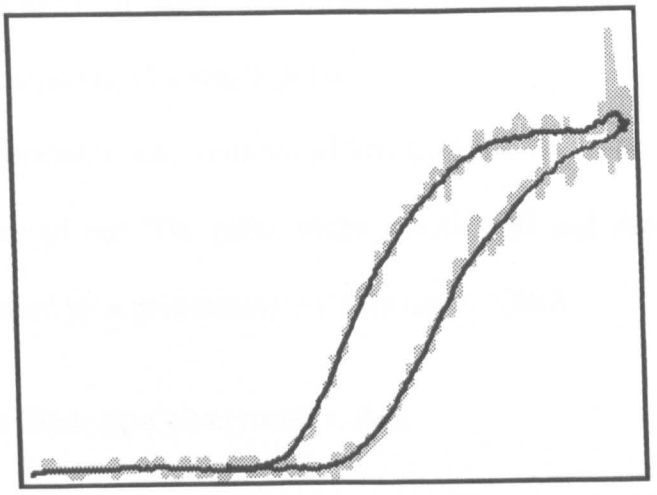


Fig.2.42 Graph showing I_D vs V_G with and without average. 'Black' average data, 'grey' row data.

We would like to point out that the average is only carried out on the input V_G pulse and the output V_D. No average is performed to the calculated the I_D~V_G curve. Figure 2.42 compares the I_D~V_G before and after applying the averaging technique. It can be seen that the feature of the I_D~V_G curve is not affected by the average.

2.6.2.2 Pulse Generator characteristics

The HP81101A pulse generator is used to generate pulses with varying pulse width, voltage levels and linearly sloped edges. A minimum pulse width of 10ns can be applied. The linearity of the slopes is specified to be +/-3% within the 10% - 90% part of the slope. Internally, the linear slope is generated by charging a capacitor with a fixed current. This is in contrast with arbitrary function generators that construct a slope by small incremental steps. The importance of this specification are from the fact that the capacitance is directly affected by inaccuracies in the slope of the pulse edge. Maximum output current is not considered to be important since, commonly; 50 Ω ballast is used at the output, leading to currents of up to 160mA. This value is well above the current levels that are used in our experiments. In the present set-up, high impedance ballast is used. Ballast values can be selected at the input of the oscilloscope. Selecting 50 Ω

could damage the oscilloscope input. The pulse generator is used in 'TRIG'mode, this allows for software triggering of a single pulse.

The Pulse generator can output different shape of pulse according to the experiment to be carried out. The pulse shape, width, fall and rise time, period, and amplitude are controlled by a programme running under UNIX.

2.6.2.3 Oscilloscope characteristics

The Metrix OX8050 has a maximum sample rate of 100 MSamples/s. The oscilloscope is used in the digital acquisition mode. The ADC outputs 8 bits, leading to a dynamic range of 10.24 divisions. The sensitivity ranges from 1 mV/div to 20V/div. The input impedance of the channels is set to 1MΩ. Selection of 50Ω input impedance limits the allowed input signal to 5Vrms; higher voltages can damage the instrument. The oscilloscope is triggered by the signal coming from the pulse generator. The 'Mode&Holdoff' option is set to 'Normal', disabling auto-triggering, to capture one Vg pulse from the pulse generator and the V_D.

2.6.2.4 Keithley K4200

The K4200 have many features which are not used in the charge pumping experiment as we only need measuring the substrate current.

The easy-to-use K4200 performs laboratory grade DC and pulse device characterization, real-time plotting, and analysis with high precision and sub-fentoamp resolution. It is the best tool available for interactive parametric analysis and device characterization. It offers the most advanced capabilities available in a fully integrated characterization system.

2.7 CONCLUSION

In this chapter, the principle of different characterization techniques available for High-k devices are described. The fast measurement technique was implemented and the difficulties encountered, such as voltage overshoot and noises, were overcome. A new charge pumping, VT^2CP, is proposed. The details for TDDB tests and results are presented. The instruments and their limitations are given. Good knowledge of instruments and measurement technique lay a solid foundation for systematic tests and interpretation of results in the following chapters.

CHAPTER 3 CHARACTERIZATION OF AS-GROWN ELECTRON TRAPS

3.1 INTRODUCTION

As silicon dioxides and oxynitrides approach their thickness limit (~1nm), a major challenge for the current complementary metal-oxide-semiconductor (CMOS) industry is the rapid increase of gate leakage current. To reduce this leakage, intensive worldwide efforts have been made to find an alternative gate dielectric of higher dielectric constant (high-k) than SiON **[Groes04], [Ker03], [Lerou04], [Reim05]**. For the same equivalent electrical oxide thickness (EOT), high-k layers are physically thicker and it has been demonstrated that gate leakage can be reduced by several orders of magnitude, when Hf-based dielectric was used **[Groes04]**. At present, there are a number of difficulties holding back their commercial applications due to electron trapping, which can reduce mobility through Coulombic scattering **[Groes04],[Degra03],** lower yield **[Groes04],[Degra03]**, shift threshold voltage (Vth) and cause breakdown by forming a conduction path **[Groes04],[Degra03]**.As a result, understanding electron traps becomes a pressing issue.

For SiO_2, there are little as-grown electron traps **[ZhangW02]**. In contrast, it has been reported that the effective density of as-grown electron traps in HfO_2 can be in the order of $10^{13}cm^{-2}$ **[ZhaoC05]**. There is only limited information on the spatial distribution of these traps **[Reim05], [Felnh05]** and agreement on trap location has not been reached. Early work **[Felnh05]** showed that these traps have an energy band of 0.5-0.8eV below the bottom edge of HfO_2 conduction band. The electron trapping is

highly dynamic and trap levels can be substantially underestimated when measured by conventional methods based on the recording of the shift of transistor's dc transfer characteristics **[Groes04], [Ker03], [ZhaoC05], [Reim05]**. Despite the early efforts and their importance, our understanding of electron trap properties, behaviour, kinetics, capture cross section in Hf-based dielectrics is still limited and many issues remain to be resolved.

In this chapter, three important issues will be addressed:

1. The properties and dynamic behaviour of electron traps in HfO_2/SiO_2: attention will be paid to the impact of measurement techniques on trapping and the dependence of trapping on conduction mechanism in section 3.2.

2. The trap location in HfO_2/HfSiO stacks: it will be assessed by selecting test samples, measurement techniques, and test conditions carefully in section 3.3.

3. The capture cross section (σ) and trapping kinetics: the capture cross section is one of the most important properties for traps. Agreement has not been reached on whether σ has a discrete or continuously distributed value **[Zafar03]**, **[ZhaoC05]**, **[Bers04]**, **[AKang03]**.The reported capture cross sections spread over a range as large as 10^{-12}~$10^{-19}cm^2$ **[Zafar03]**, **[ZhaoC05]**, **[Reim05]**, **[Bers04]**, **[KangA03]**. At the large end of this range, $\sigma=3\times10^{-13}$ cm^2 was obtained based on an irradiation experiments, where the electron fluency was limited to 3×10^{13} cm^{-2} **[KangA03]**. This is too low for detecting traps of smaller cross sections **[ZhangJ92]**. At the small end of this range, capture cross sections in the order of $10^{-19}cm^2$ was obtained from the shift of quasi-dc transfer characteristics **[Zafar03]**. It will be shown that the significant detrapping during the quasi-dc measurement has a material effect on the extraction of capture cross section. To determine the capture cross section, efforts will be made to estimate the transient gate current and the real electron fluency. It is found that the trapping follows a kinetics

with two discrete capture cross sections, rather than a continuous distribution of capture cross section in section 3.4.

3.2 PROPERTIES AND DYNAMIC BEHAVIOUR OF ELECTRON TRAPS IN HfO_2/SiO_2 STACKS

3.2.1 Devices and Techniques

HfO_2 was prepared by atomic layer deposition. The gate dielectric consists of a ~1nm SiO_2 interfacial layer and a 4nm HfO_2, resulting in an equivalent oxide thickness of 1.8nm. The nMOSFET has an n^+ poly-si gate, a channel length of 1μm and a channel width of 10μm.

Two techniques were used for measuring electron traps. The first one is the traditional 'DC I_D~V_G' **[Zafar03]** and the second one is the 'pulsed I_D~V_G' **[Ker03]** introduce in Chapter 2 section 2.1.2 and 2.2.2 respectively. The main advantage for the pulsed I_D~V_G method is that the trapped charges can be measured within microseconds after filling, while it could take up to seconds for the 'DC I_D~V_G' technique. The impact of measurement techniques on the trapping will be addressed next.

3.2.2 Impacts of measurement techniques on trapping

The experimental results obtained by the pulsed I_D~V_G technique is shown in Fig. 3.1. The data shown in Fig 3.1 was taken using a voltage pulse with 200 μs and 100 μs rise and fall times. As can be seen the pulsed technique results in a higher drive current compared to the DC measurement (open symbols) this difference is explained in chapter 2 section 2.3.1. The V_T shifts obtained with the pulsed technique are strongly enhanced, as illustrated by the inset in Fig 3.1 (note delta V_T is in log scale). The

voltage shifts shown in the inset were extracted at a constant current level (Id max /2) of 150 μA for pulsed I_D~V_G and 125μA for DC.

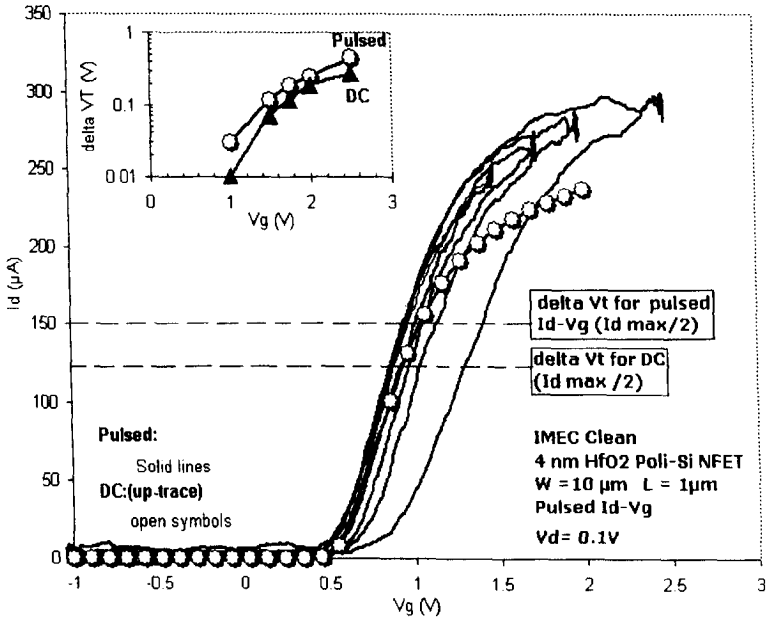


Fig. 3.1 I_D~V_G characteristics measured using the pulse I_D~V_G setup in chapter 2 section 2.2.2 Fig 2.7. A static I_D~V_G up trace is shown for comparison. V_T-shift as function of applied voltage V_G by both techniques is compared in the inset.

The enhanced voltage shift observed from the pulsed measurement technique indicates that in DC measurement a significant fraction of the trapped charges gets detrapped during the down trace. The evidence of fast detrapping is also confirmed by the stretch-out in the down trace of the pulsed I_D~V_G characteristics when compared with the up traces. When comparing the pulsed technique with DC measurement methods, it can be concluded that the conventional DC techniques severely underestimate the charging effects in SiO_2/HfO_2 dual layer gate dielectrics and if not carefully analyzed, they might even be overlooked. In the following, the difference between DC I_D~V_G and Pulse I_D~V_G technique will be examined further.

Fig.3.2a compares the trapping measured by the two different techniques against N_{im}. N_{im} is the electron fluency across the Si/SiO_2, rather than HfO_2/Gate, interface,

$$N_{im}=Jgm \times tw/q + \Delta V_G \times Cox/q, \quad (3.1)$$

where Jgm is the measured gate current density per unit area, q one electron charge ,tw the pulse width or stress time, C_{ox} the gate oxide capacitance per unit area and ΔV_G the shift caused by trapping. It should be pointed out that the N_{im} in Eq.(3.1) is the fluency of electrons injected into the gate dielectric, rather than exiting from the dielectric.

The inclusion of the second term in Eq.(3.1) is essential, since it makes substantial contribution at low injection level N_{im}. Without it, the trapping probability can be higher than unity **[Lerou04]**. The derivation of Eq.(3.1) is given in the chapter 2 section 2.3.3. Fig.3.2a shows that the trapping measured by the pulsed I_D~V_G is significantly higher than that by the DC I_D~V_G confirming early observation in Fig.3.1. For this figure, the DC I_D~V_G measurement took 1.9sec. When the pulsed I_D~V_G was used, trapping was measured from the falling edge of the pulse and the falling time was 30μs in Fig.3.2. Because of the longer time used for the DC I_D~V_G, considerable detrapping occurs during the measurement. To support the above explanation, the falling time of the pulsed I_D~V_G is gradually increased in Fig.3.2b. Although the trapping is insensitive to the falling time between 3 and 30μs, it is clear that further increase of measurement time leads to a substantial loss of trapped charges and the results are approaching those by the DC I_D~V_G in Fig.3.2a. It is concluded that the DC I_D~V_G underestimates the trapping level in HfO_2/SiO_2 stacks. A 30μs rising and falling time will be used for pulsed I_D~V_G, hereafter.

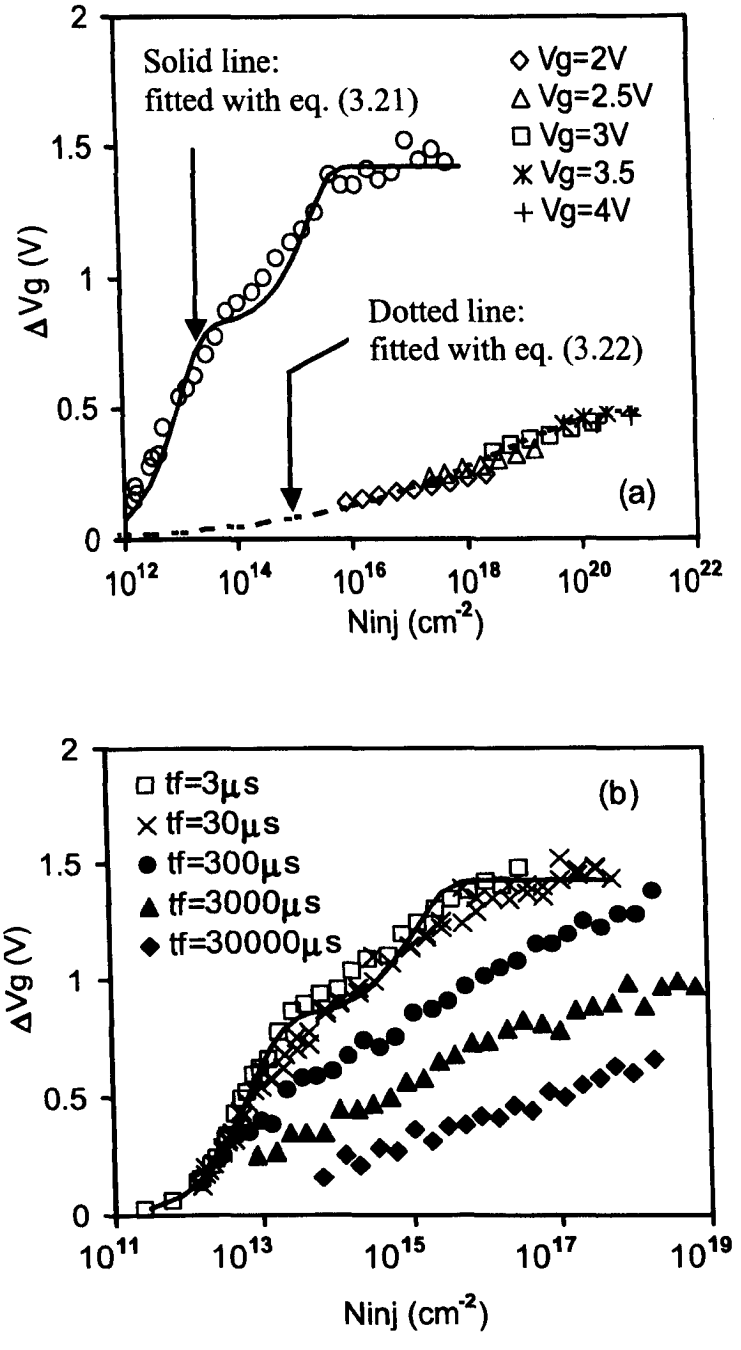
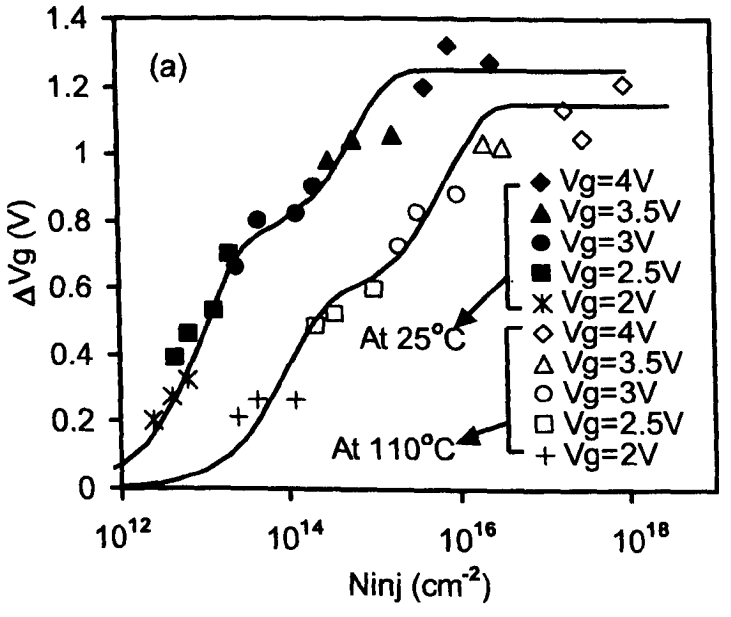


Fig. 3.2 Dynamic behaviour of electron trapping measured by different techniques. The symbol 'o' in (a) and data in (b) were measured by the pulsed I_D~V_G with V_G increasing in a step of 0.1V for each point. The pulse falling time, tf, is 30μs in (a) and different tf was used in (b). Other symbols in (a) were obtained by the traditional DC I_D~V_G, after electron injection under a V_G shown in (a). The solid and dotted lines were fitted with Eqs. (3.21)& (3.22), respectively.

3.2.3 Dependence on conduction mechanism

Since most of the high-k devices will operate at elevated temperature, it is of interest to study the effect of temperature on trapping. Fig.3.3a compares the trapping at 25°C with that at 110°C. The trapping is clearly less efficient at 110°C. This, however, does not mean that trapping will be less severe at 110°C. For a given time, higher current at 110°C (inset of Fig.3.3b) will result in higher electron fluency, N_{inj}, in Fig.3.3a. When N_{inj} at 110°C was calculated by using Jg(25°C), Fig.3.3b shows that there is little difference in trapping at these two temperatures.

Figs.3.3a&b suggest that trapping depends not only on how many electrons are passing through the oxide, but also on how they pass through. It appears that the tunneling electrons are more efficient in filling the traps than the electrons through trap-assisted Poole-Frenkel type of conduction. If this is true, one expects that trapping should also be insensitive to the stress-induced leakage current (SILC), since SILC is through trap-assisted conduction (TAC).



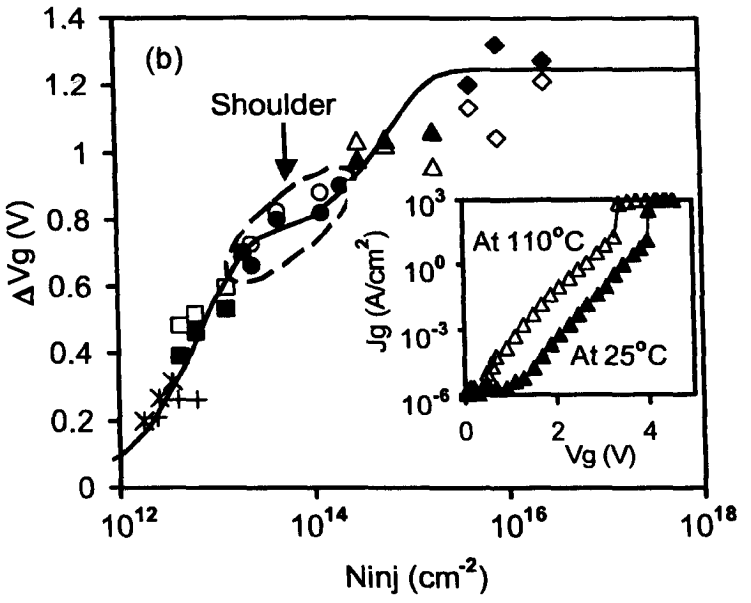
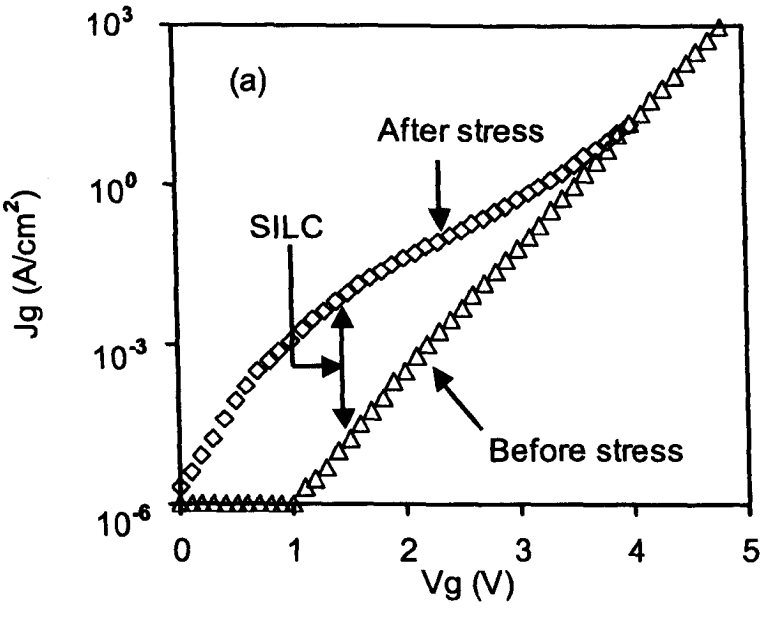
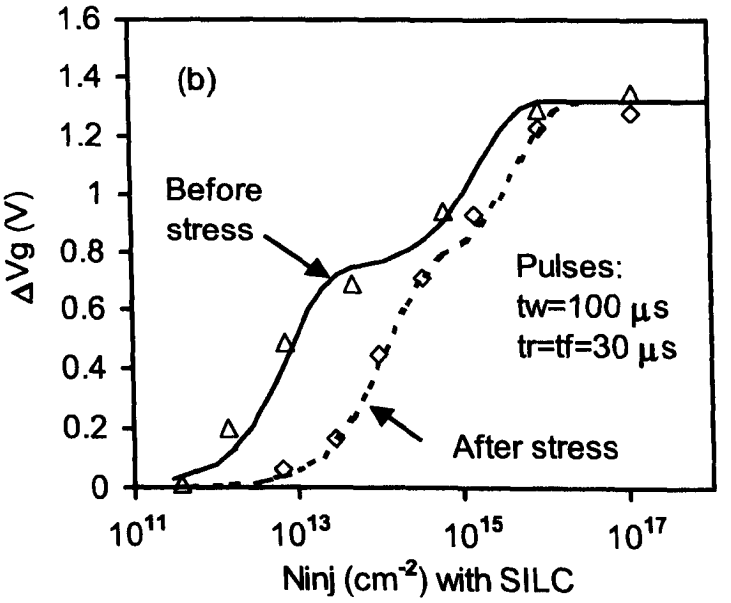


Fig. 3.3 Trapping at different temperatures. Data for a V_G given in (a) were measured by changing pulse width. In (a), Jg(110°C) and Jg(25°C) was used to calculate N_{inj} at 100°C and 25°C, respectively. In (b), Jg(25°C) was used to calculate N_{inj} at both 110°C and 25°C.

Fig.3.4a shows the gate current density before and after generating SILC. Figs.3.4b&c give the trapping with and without taking SILC into account when N_{inj} was calculated. It is clear that SILC contributes little to the trapping. It is not fully understood why trapping is insensitive to TAC, but some speculations can be given. There can be four possible explanations. First, traps participating in TAC will inevitably have a higher detrapping rate through hopping/tunneling. This leads to lower trapping. Secondly, it has been shown **[Degra04]** that the energy depth of generated traps is relatively deeper than that of as-grown traps


(a)
After stress
SILC
Before stress
Jg (A/cm²)
Vg (V)
10³
10⁰
10⁻³
10⁻⁶
0
1
2
3
4
5


(b)
Before stress
Pulses:
tw=100 µs
tr=tf=30 µs
After stress
ΔVg (V)
Ninj (cm⁻²) with SILC
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
10¹¹
10¹³
10¹⁵
10¹⁷

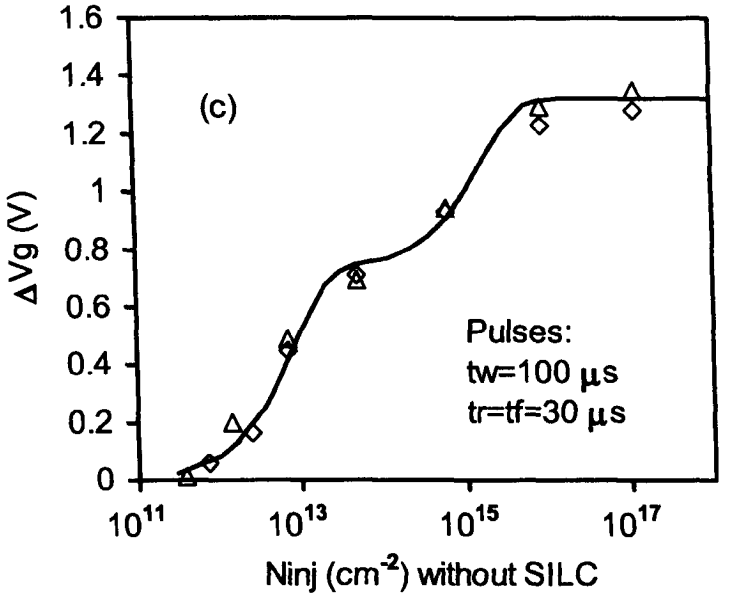


Fig. 3.4 Effects of SILC on trapping. (a) shows the Jg before and after creating SILC. (b) and (c) show the trapping with and without taking SILC into account when Ninj was calculated, respectively.

The electron hopping through these generated traps (i.e., SILC) will not fill the as-grown traps of a higher energy, as schematically shown in Fig.3.5a. Third, it is well known that oxide breakdown is a local phenomenon. As a precursor for the breakdown, SILC will not be uniform either at a microscopic scale. Although a large number of electrons can go through the localized path, traps outside of this path will not be filled by them, as illustrated in Fig.3.5b. Finally, generated traps can be more uniformly distributed in the direction between the gate and the substrate than the native traps. A small number of generated traps can give a significant SILC. For example, although there is no as-grown trap in the interfacial SiO_2 layer, traps can be created within it, which should enhance SILC, more detail work will be given in section 3.4.2.

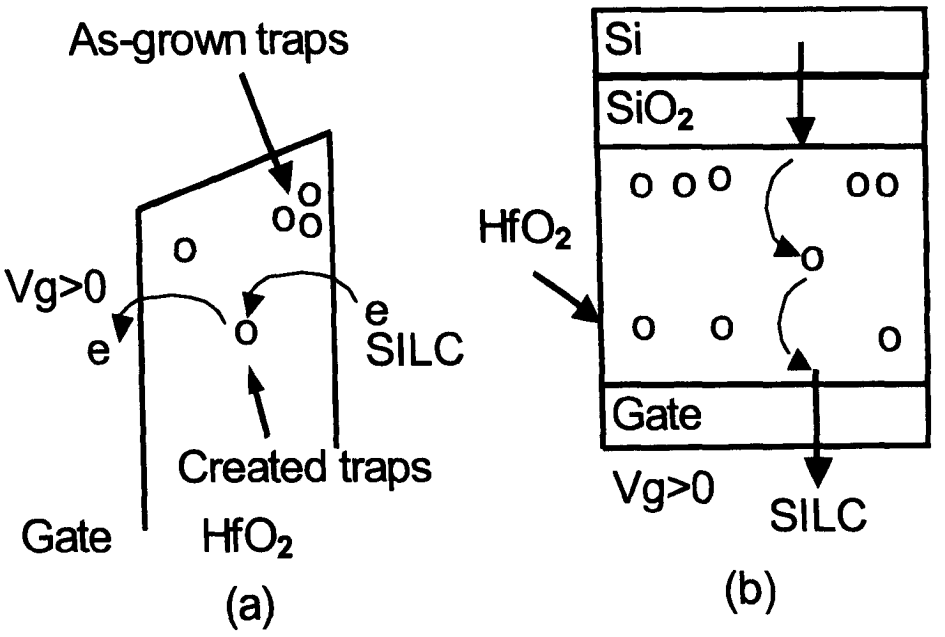


Fig.3.5 Schematic illustration on why SILC contributes little to filling traps. (a) shows that the SILC can occur at an energy level deeper than that of as-grown traps. (b) shows that traps can be outside of the localized conduction path.

3.3 LOCATION OF AS GROWN ELECTRON TRAPS IN HfO_2/HfSiO STACKS

3.3.1 Devices and Techniques

To assess the location of traps in HfO_2, five HfO_2 layers were prepared by ALCVD with a physical thickness of 1.8, 2.0, 2.5, 3.0 and 4.0nm, respectively. All other processing steps are the same for these five samples. Before HfO_2 deposition, there is a 0.4nm chemical oxide. To activate the source and drain, a 1000°C, 1sec anneal was used. This results in a 1.7nm HfSiO interfacial layer (IL), equivalent to a 0.9nm SiO_2 layer **[Schr05]**. The sample has a metal gate, consisting of 10nm TaN capped by a 70nm TiN layer. The size of nMOSFETs is 0.25μm in length and 10μm in width.

Several techniques can be used to measure traps in HfO_2. One of them is the frequency charge pumping **[Degra03]**, **[ZahM05]**. Here, the measured traps must be so

close to the substrate that tunneling happens within one pulse width of gate voltage, V_G. Since traps sufficiently away cannot be probed, we do not use this technique here. Traps can also be measured from the shift of gate voltage for a given drain current. This requires the filling process being periodically interrupted to record transfer characteristics.

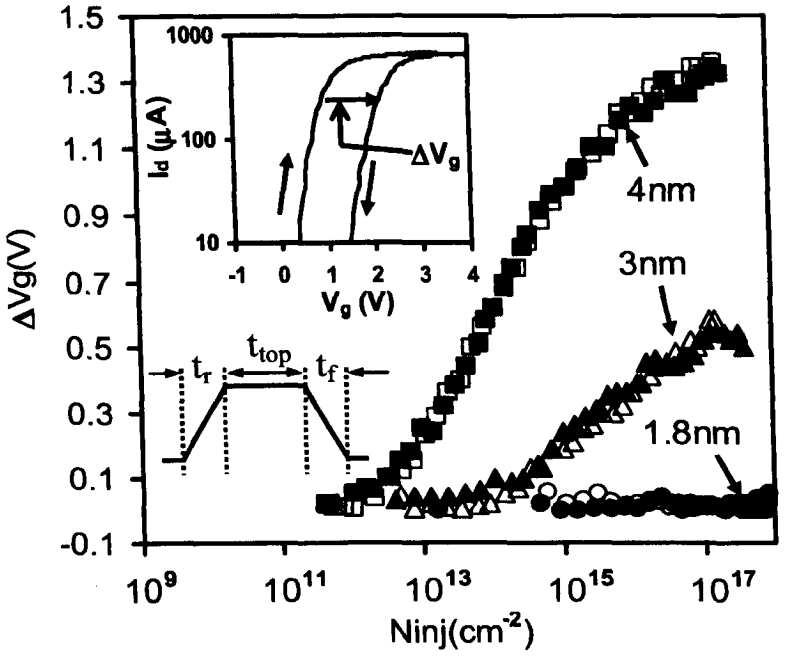


Fig. 3.6 Typical behavior of electron trapping. The inset schematically shows the gate pulse applied with t_r=t_f=10μs and t_{top}=23μs. The trapping induces a shift in the gate voltage, ΔV_G, between the two I_D–V_G recorded during the rising and falling edges of the gate pulse. Each open symbol was obtained by applying one gate pulse. N_{inj} is the number of electrons injected into the gate dielectric. The lowest N_{inj} for the open symbols was obtained by setting V_G=1.6V during t_{top}. The V_G at t_{top} was increased with a step of 0.1V and the highest N_{inj} was obtained with V_G=5V during t_{top}. The trapped electrons were then detrapped and the filling starts again by resetting V_G=1.6V during t_{top}. The trapping obtained during the second filling is represented by the filled symbols. The good agreement between the open and filled symbols indicates that trapping is dominated by as-grown traps. The thickness values given on the figure are for the HfO_2 layers.

Traditionally, transfer characteristics were measured under quasi-DC conditions. It has been shown in the section 3.2.2 that significant amount of trapped electrons are lost during the DC measurement and trapping level is underestimated. To overcome this

difficulty, the 'pulsed I_D~V_G 'technique introduced previously will be used again here. A gate pulse, illustrated by the inset of Fig.3.6, was applied and traps were filled during the peak period, t_{top}. This led to a 'hysterisis', ΔV_G, in the I_D~V_G recorded during the rising and falling edges of the gate pulse, as shown by the inset of Fig.3.6.

3.3.2 Test Condition

Since the objective of this work is to assess the location of as-grown electron traps in HfO_2, the contribution of generated traps to trapping should be negligible. This is confirmed by Fig.3.6. After stressing and filling traps, the trapped electrons were detrapped and then filled again (i.e. the second filling). If the contribution of generated traps to trapping is considerable, it should result in an enhancement in trapping during the second filling **[ZhangJ01]**. This is not observed in Fig.3.6 and, consequently, trapping is dominated by as-grown traps.

Ideally, one would like to fill and measure 100% of as-grown traps. Test conditions have to be selected carefully for this task. To start, V_G used for filling the traps must be sufficiently high. It has been reported that the as-grown electron traps have an energy level of 0.5-0.8eV below the bottom edge of the HfO_2 conduction band **[Xu02]**. If V_G is too low, tunneling electrons can be below the trap energy level, making the trap filling difficult, as illustrated by the inset of Fig.3.7

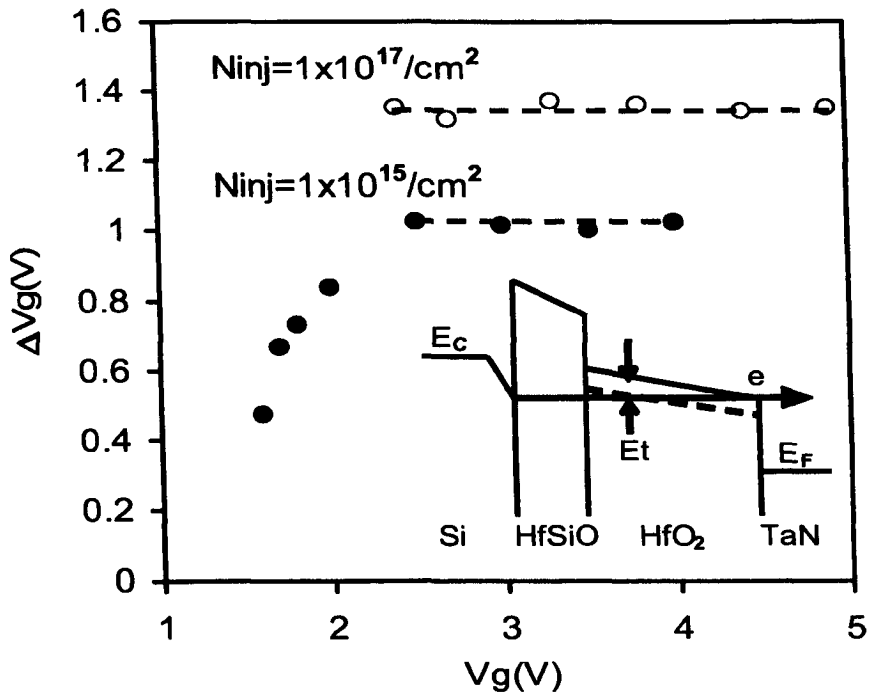


Fig. 3.7 Dependence of trapping on the gate voltage applied during filling, V_G. For an electron fluency of $N_{inj}=10^{17}cm^{-2}$, the filling is insensitive to V_G when $V_G \geq 2.4V$. $N_{inj}=10^{17}cm^{-2}$ could not be reached for $V_G < 2.4V$ due to the limited pulse duration provided by the pulse generator used here. An electron fluency of $N_{inj}=10^{15}cm^{-2}$ can be reached by lower voltages and the trapping reduces when V_G drops below 2.5V. The inset shows that electron energy can be below the trap energy level (dashed line) at low V_G. The sample has a 4nm HfO_2.

In this case, the filling across the HfO_2 is not uniform and more traps will be filled at higher V_G for a given electron fluency, which is not desirable. To simplify the experimental condition, V_G should be high enough, so that electron energy is above the trap energy level and the filling will not be sensitive to V_G. Fig.3.7 shows that this requires $V_G > 2.5V$ approximately. We will assess the trap location by filling at $V_G \geq 3V$. The insensitivity of trapping to V_G when $V_G > 2.5V$ indicates that trapping is not affected by the electron energy, once it is above the trap energy level. To fill as many traps as possible, we would like to have an electron fluency sufficiently high for the trapping to reach saturation. Fig.3.8 shows that saturation is obtained at an electron

fluency of $10^{17}cm^{-2}$. The location of traps will be assessed at this level of electron fluency.

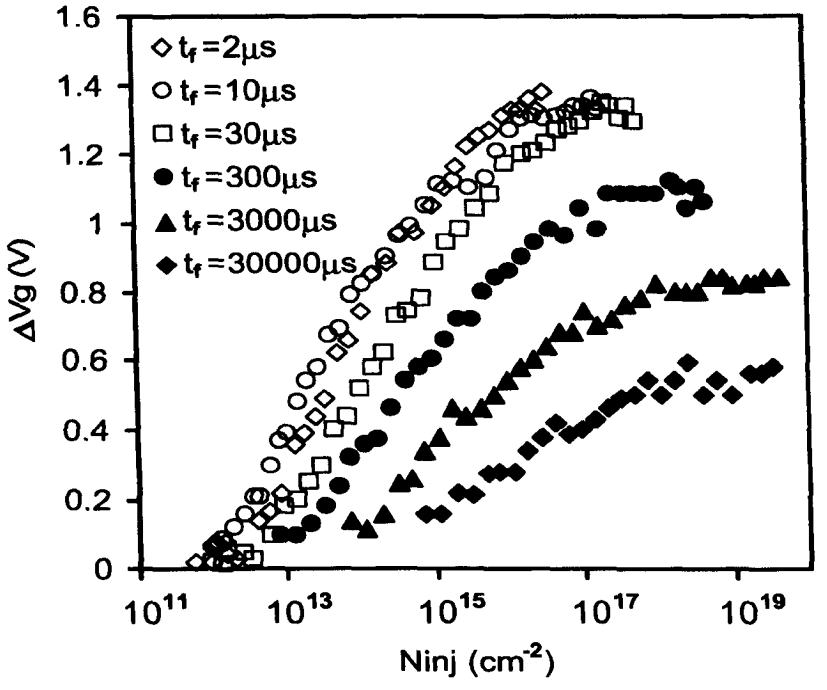


Fig.3.8 Effects of measurement time on trapping. t_f is the falling time of gate pulse, during which the trapping is measured. The rising and falling time is the same and the duration for the peak is set at $t_{top}=2.33t_f$ (see the inset of Fig.3.6). An increase of t_f beyond 10μs leads to a progressive reduction of trapping. t_f =10μs will be used for the rest of this work. The peak value of gate pulse is the same as those used in Fig.3.7 The sample has a 4nm HfO_2 here.

Care must be exercised in controlling the detrapping after filling. When the measurement time increases from 2μs to 10μs, Fig.3.8 shows that trapping changes little, indicating detrapping is negligible during measurement. When the measurement time increases beyond 10μs, however, trapping reduces progressively, because of detrapping. A measurement time of 10μs is chosen here for assessing trap location.

3.3.3 Location of as grown electron traps

After selecting the appropriate test conditions, it is the time for assessing the trap location. Figs.3.9a & b show the gate voltage shift, ΔV_G, recorded at a filling electron fluency of $10^{17}cm^{-2}$ as a function of HfO_2 thickness, X_{Hf}.

Trapping is insignificant for the thinnest HfO_2 (X_{Hf} =1.8nm), agreeing with early observation **[Bers04]**. An increase of X_{Hf} results in higher ΔV_G. If one assumes that the traps are located at the HfO_2/HfSiO interface, ΔV_G is related to X_{Hf} by **[Sze81]**,

$$\Delta V_G = -\frac{qX_{Hf}Q}{\varepsilon_o k} \tag{3.2}$$

where Q is the area density of trapped charges and k is the high-k dielectric constant.

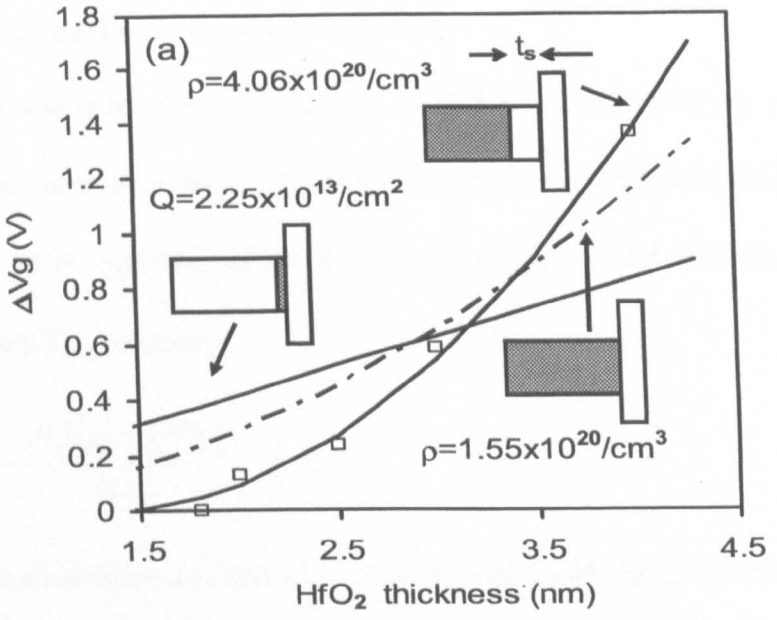


Fig. 3. 9a An assessment of as-grown electron trap location. The gray area of the insets show the assumed trap location for each case. The ΔV_G, symbols "ρ", was measured at an electron fluency of $10^{17}cm^{-2}$. The peak value of gate voltage used is 3.0, 3.6, 4.0, 4.4, and 4.7V for the 1.8, 2.0, 2.5, 3.0 and 4.0nm HfO_2 layers, respectively. The solid straight line was obtained by fitting the data with Eq. (3.2) and the extracted area density (Q) is given on the figure. The dashed-dotted and solid curves were obtained by fitting with Eqs. (3.3) and (3.4), respectively The extracted volume density, ρ, is given on the figure for each case.

Fig.3.9a shows that this linear relation between ΔV_G and X_{Hf} does not agree with the measured data. It is ruled out that the traps are concentrated near the HfO_2/HfSiO interface.

If one assumes that the traps are uniformly distributed throughout HfO_2 with a volume density of ρ, the relation between ΔV_G and X_{Hf} becomes,

$$\Delta V_G = -\frac{qX_{Hf}^2 \rho}{2\varepsilon_o k} \tag{3.3}$$

Although this is closer to the measured data than Eq. (3.2), the mismatch shown in Fig.3.9a is still too large to be acceptable. There are three possible reasons for this mismatch:

There can be a region of negligible trapping with a thickness of t_s from the HfSiO interface, as shown by the inset of Fig.3.9a.

One may speculate that HfO_2 can be modified in this region by interacting and mixing with the interfacial layer. If some silicon atoms enter this region, as-grown electron traps can be reduced, since they are negligible in Hf-silicates. The relation between ΔV_G and X_{Hf} becomes,

$$\Delta V_G = -\frac{q(X_{Hf} - t_s)^2 \rho}{2\varepsilon_o k} \tag{3.4}$$

Fig.3.9a shows that this assumption agrees well with the test data and the fitted t_s is 1.3nm.

A region of negligible trapping with a thickness of 't_g' can be located near the gate. The relation between ΔV_G and X_{Hf} becomes,

$$\Delta V_G = -\frac{q(X_{Hf}^2 - t_g^2)\rho}{2\varepsilon_o k} \tag{3.5}$$

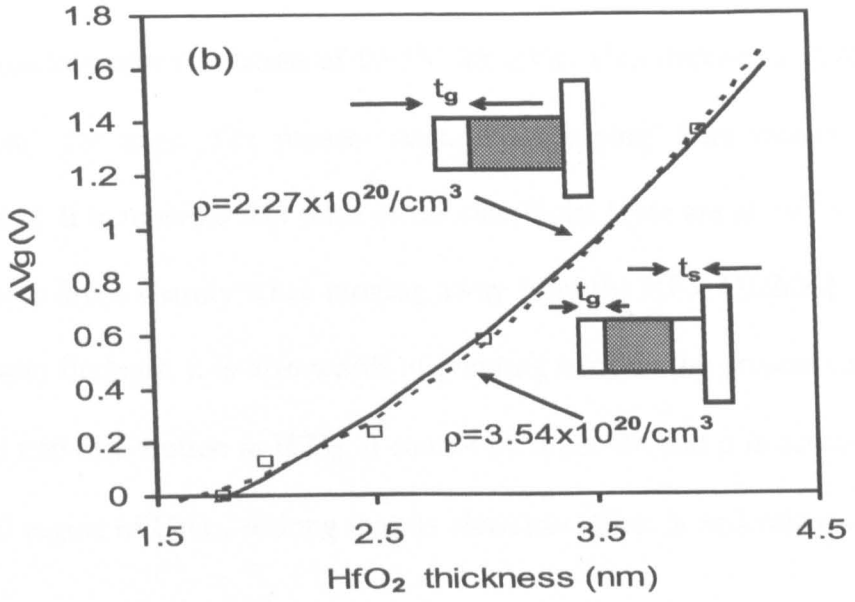


Fig. 3.9b An assessment of as-grown electron traps location. The gray area of the insets shows the assumed trap location for each case. The ΔV_G, symbols "ρ", was measured at an electron fluency of $10^{17}cm^{-2}$. The peak value of gate voltage used is 3.0, 3.6, 4.0, 4.4, and 4.7V for the 1.8, 2.0, 2.5, 3.0 and 4.0nm HfO_2 layers, respectively. The solid and dotted curves were obtained by fitting with Eqs. (3.5) and (3.6), respectively. The extracted volume density, ρ, is given on the figure for each case

Fig.3.9b shows that this assumption also agrees with test data with a fitted tg of 1.8nm. There are two possible explanations for the negligible trapping. One is the lack of traps, because HfO_2 near the gate can be modified by interacting with metal. The other is that traps sufficiently close to the metal cannot be filled steadily due to efficient tunneling to the gate under a positive V_G. A conclusion cannot be reached here.

Since the test data do not allow us to rule out anyone of the above two cases, it is possible that both interfacial regions with negligible trapping exist simultaneously. ΔV_G is now related to X_{Hf} by,

$$\Delta V_G = -\frac{q[(X_{Hf} - t_s)^2 - t_g^2]\rho}{2\varepsilon_o k} \tag{3.6}$$

The extracted t_s and t_g is 1nm and 0.6nm, respectively.

Finally, it should be pointed out that the pulsed I_D~V_G technique used in this work has a measurement resolution of 0.05V for ΔV_G. This imposes a detection limit around $10^{19}cm^{-3}$ for traps. The phrase 'negligible trapping' here means trapping is below this limit. It is reported that traps in the interfacial layer are at $10^{19}cm^{-3}$ or lower and trap density drops sharply when moving away from the HfO_2 **[Heh06]**. This agrees with the present findings. It is also worth of pointing out that the present results do not give detailed trap distribution in HfO_2. It cannot be ruled out that ρ is actually changing in the central region of HfO_2, so long that its electrical effect is equivalent to a constant ρ.

3.4 DETERMINATION OF CAPTURE CROSS SECTION FOR AS-GROWN ELECTRON TRAPS IN HfO_2/HfSiO STACKS

3.4.1 Devices

To assess the capture cross section of electron traps, it is desirable to select a test sample with a large amount of as-grown traps. In section 3.2, it is shown that the effective density of as-grown electron traps in a 4nm ALD HfO_2 is in the order of $10^{13}cm^{-2}$. The 4nm ALD HfO_2 was chosen as the test sample in this work, therefore. The physical vapour deposited (PVD) metal gate consists of a 10nm TaN layer capped by a 70nm TiN. Before HfO_2 deposition, there was a 0.4nm chemical oxide. To activate the dopant, a 1000°C and 1sec anneal was used. The TEM analysis shows that there is an interfacial layer (IL) of 1.7nm, which contains Hf-silicate **[Schr05]**. At the end of the processing, a forming gas anneal was carried out at 520°C for 20min.

The equivalent oxide thickness of the stack, EOT, is 1.75nm. The size of nMOSFETs used is 0.25μm in length and 10μm in width.

3.4.2 Test condition and measurement

In Chapter 2 section 2.3.2 the test condition was discussed .The pulsed I_D~V_G techniques was presented to be suitable for the capture cross section analysis but before carrying out a detailed analysis of experimental data, it is important to confirm their reproducibility. The experiment step described in chapter 2 section 2.3.2 was repeated on three different MOSFETs and the results are shown in Fig.3.10

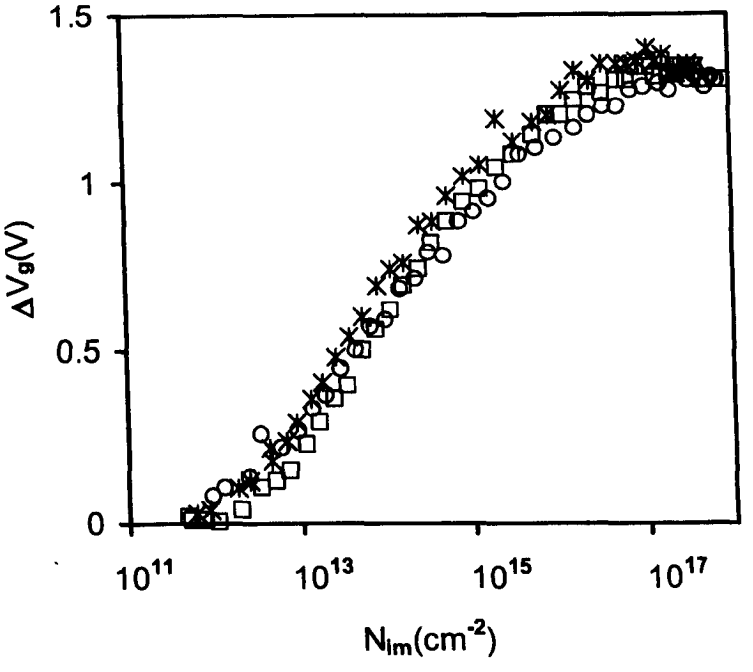


Fig.3.10 Typical results of as-grown electron trapping. The electron fluency, N_{im}, was obtained from Eq.(3.1), based on the gate current density given in Fig.3.11. Each point corresponds to one ΔV_g in Fig.2.8 in chapter 2 section 2.3.2 and was obtained by applying one pulse to the gate. The t_{top} and t_f was fixed at 23μs and 10μs, respectively. To increase N_{im}, the gate bias amplitude was increased progressively by a step of 0.1V, starting from V_G=1.6V. The same tests were carried out on three devices and the results are represented by the three sets of symbols. The sample-to-sample variations are insignificant.

Note also that to assess the impact of gate materials and related processes on electron trapping, limited tests were carried out on poly-si gated MOSFETs with the same HfO_2 thickness. The poly-si gated devices have an initial 1nm chemical oxide and the dopant activation was at 1000°C for 10sec, resulting in an EOT of 1.8nm. To further explore the effect of fabrication techniques on electron trapping, a 2nm HfO_2 was also prepared by physical vapour deposition (PVD). The oxidation was at 600°C for 5sec, followed by a post deposition anneal at 700°C in N_2 for 60sec.

The trapping is shown in Fig.3.10 and the gate current density J_{gm} is plotted in Fig.3.11. Fig.3.10 and Fig.3.11 show that the sample-to-sample variation is acceptable. As a result, the quality of the data is good enough for extracting electron capture cross sections. All measurements were at room temperature.

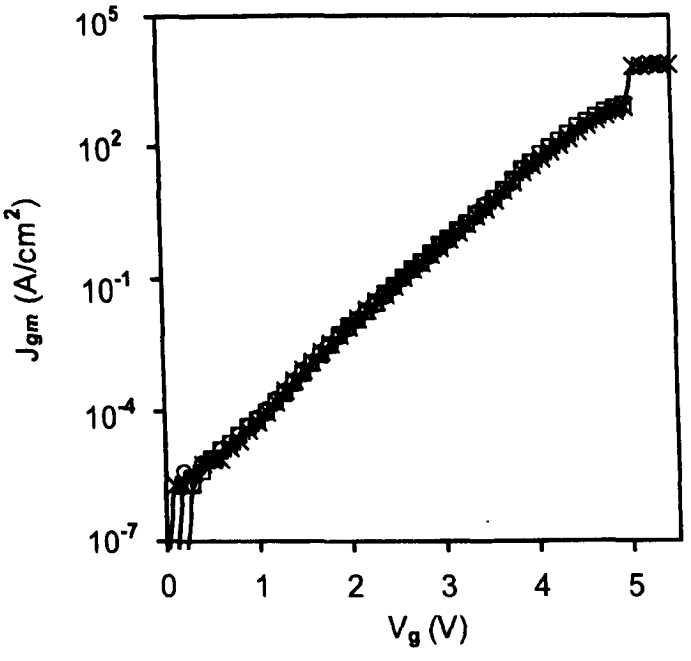


Fig.3.11 The measured gate current per unit area against gate biases. Here, the measurement of each point took approximately 0.15sec, which was much longer than the pulse period used in Fig.3.10. The same measurements were made on three devices and the sample-to-sample variations were small.

Fig.3.12 compares electron trapping in samples of different gates. Although trapping in metal-gated samples is lower, the impact of gate materials is modest. Since the industrial attention is focused on metal gates at present, attention will also be concentrated on metal-gated samples hereafter.

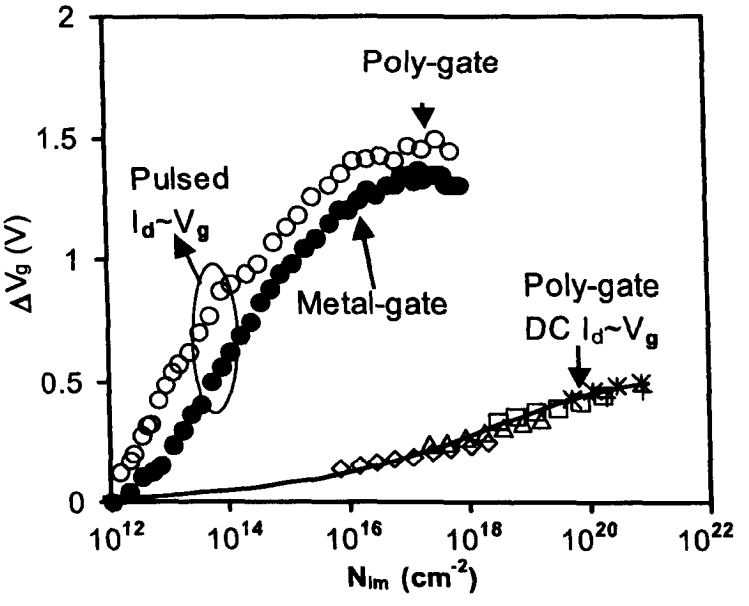


Fig.3.12 Effects of gate materials and measurement techniques on electron trapping. Changing gate from poly-si to metal leads to a modest reduction of trapping, but the overall features remain the same. The t_{top} and t_f for pulsed measurement is 23μs and 10μs, respectively. Trapping is significantly reduced through detrapping during the quasi-dc measurement. The solid line was obtained by fitting with Eq.(3.24). The gate voltage used for the quasi-dc measurement is: 'O'—2V; 'Δ'—2.5V; '□'—3V; '*'—3.5V; '+'—4V

As mentioned earlier, when the trapping in HfO_2 was assessed from the shift of quasi-dc transfer characteristics in the subthreshold region **[ZhangW02]**, the measurement can take seconds and Fig.3.12 shows that trapping is substantially underestimated through detrapping. If these quasi-dc data are used for extracting capture cross sections, a value in the order of $10^{-19}cm^2$ is obtained, which agrees well with that reported in early works **[Zafar03]**, **[Bers04]** and indicates that capture cross section is not sensitive to the difference in samples used by different groups. However, a

comparison with the pulsed data clearly shows that the detrapping during quasi-dc measurements leads to a severe underestimation of the capture cross section. The question is how fast the pulse has to be to suppress detrapping.

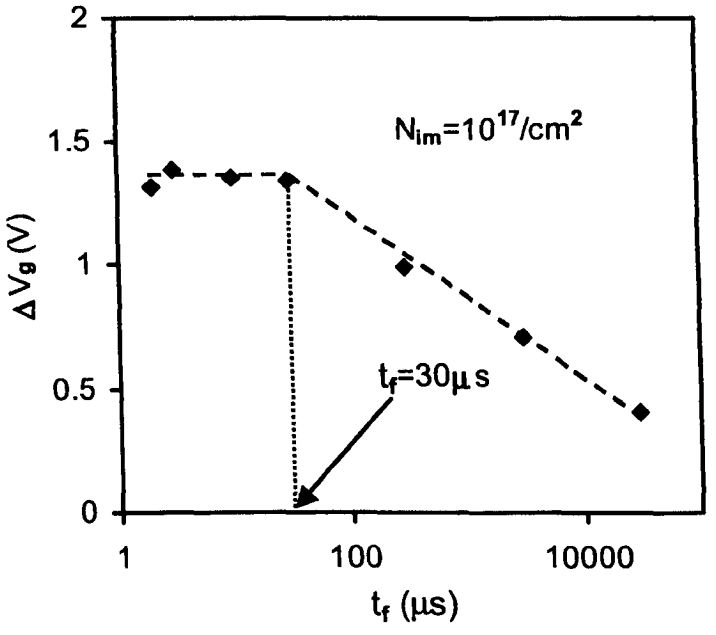


Fig.3.13 Dependence of trapping levels on the pulse falling time, t_f. The electron fluency was fixed at $10^{17}cm^{-2}$. When t_f is higher than 30μs, an increase of t_f enhances detrapping, which in turn reduces the trapping. For t_f<30μs, detrapping in not important and trapping is insensitive to t_f.

For the pulsed I_D~V_G technique, the delay between trap filling and measurement is controlled by the falling time, tf, of the pulse. Tests were carried out to assess the impact of tf on trapping. For a given N_{im}, Fig.3.13 shows that trapping indeed reduces substantially for long tf through detrapping. However, trapping is insensitive to tf for tf<30μs. This indicates that detrapping becomes insignificant when tf<30μs for the sample used in this work. tf=10μs is selected for the rest of this work.

3.4.3 Transient gate current estimation

To determine the capture cross section, the electron fluency must be known. One way for obtaining the fluency is measuring the transient gate current during the gate

pulse and integrating it against time. Unfortunately, in practice, this small and transient gate current could not be measured at present, since the parameter analyser is too slow. The J_{gm} measurement in Fig.3.11 took about 0.15sec, while the typical transient is in the order of microseconds. Electron trapping reduces the electrical field near the cathode and, consequently, leads to a reduction of the gate current **[ZhangJ92]**. Fig.3.10 shows that trapping can induce a gate voltage shift over 1V. An inspection of Fig.3.11 shows that, for this amount of reduction in V_G, J_{gm} can be reduced by two orders of magnitude. As a result, there will be a large change in the gate current when a pulse is applied to the gate. Initially, there is no trapping and gate current will be high. As trapping accumulates, gate current will approach the J_{gm} shown in Fig.3.11.

The N_{im} in Fig.3.10 was calculated from the measured gate current density, J_{gm}, by using Eq.(3.1). Here, the assumption is that the gate current is a constant during the pulse and equals to its value shown in Fig.3.11. This will lead to an underestimation of electron fluency, because the transient high current was not taken into account. In this section, efforts will be made to reduce the error in N_{im}. Since the transient gate current cannot be measured, a first order estimation will be made by numerical simulation.

3.4.3.1 Numerical modeling of gate current without trapping

Under a given gate bias, electrons can tunnel through the thin dielectric stack with or without the assistance of traps **[Houssa01]**. The current caused by trap-assisted-tunnelling (TAT) will not be considered here, since it does not contribute to trapping, as shown in section 3.2.3. Without TAT, tunnelling current per unit area from the substrate, J_g, can be evaluated by **[ZhaoY04]**, **[Govo04]**, **[Yang99]**,

$$J_g = q\sum_{ij} N_{ij} T_{ij} f_{ij} \tag{3.7}$$

The simulation can be broadly divided into four steps. First, one must know the density of charge carriers available for tunnelling, N_{ij}, from the inversion layer in the j^{th} energy sub-band of the i^{th} valley **[YZhao04], [Govo04], [Yang99]**. Second, the potential profile in the dielectric is determined, since tunnelling probability, T_{ij}, is sensitive to it. T_{ij} is then calculated based on the Wentzel-Kramers-Brillouin (WKB) approximation. Finally, the impact frequency at the dielectric/substrate interface, f_{ij}, is calculated and the J_g is obtained from Eq. (3.7). More information for each step is given below:

Calculation of N_{ij} and E_{ij}: To start the simulation, we chose the substrate surface potential, V_s, as the input parameter. The corresponding N_{ij} and the sub-band energy level, E_{ij}, were obtained by solving the following equations numerically **[ZhaoY04], [Govo04],**

$$E_{ij} = \left(\frac{\hbar^2}{2m_{zi}}\right)^{\frac{1}{3}} \left(A_j q F_{eff}\right)^{\frac{2}{3}}, \tag{3.8}$$

and

$$N_{ij} = \left(\frac{kT}{\pi\hbar^2}\right) g_i m_{di} \ln\left[1 + \exp\left(\frac{E_F - E_{ij}}{kT}\right)\right] \tag{3.9}$$

where $\hbar = h/(2\pi)$ and h is the Plank's constant, m_{zi} is the effective mass in the i^{th} valley, F_{eff} is the effective electrical field strength, and A_j originates from the j^{th} zero of the Airy function,

$$A_j \approx \frac{3}{2}\pi\left(j + \frac{3}{4}\right), \quad \text{with j=0,1,2....} \tag{3.10}$$

In Eq.(3.9), k is the Boltzmann's constant, T the temperature,m_{di} the density-of-states effective mass,g_i the degeneracy of the i^{th} valley, and E_F the Fermi level. The values of g_i, m_{zi}, and m_{di} used for the calculation are given in Table I .

	g_i[Govo04]	m_{ti}[Govo04] (m_0)	m_{di}[Govo04] (m_0)
i=0	2	0.916	0.190
i=1	4	0.190	0.417

Table 1 Parameters used for the i^{th} valley of Si conduction band. m_0 is the free electron mass. All other symbols are defined in the text. The data were taken from the references in the square bracket.

Determination of potential drop in the dielectric and gate voltage: Once the N_{ij} is known, the total carrier density in the inversion layer, N_{inv}, can be obtained by summing up N_{ij} in all subbands. The total potential drop over the oxide, V^0_{OX}, is given by **[ZhaoY04],**

$$V_{ox}^{0} = \frac{q(N_{depl} + N_{inv})EOT}{\varepsilon_{SiO2}\varepsilon_0}, \tag{3.11}$$

where N_{depl} is the charged dopant per unit area in the depletion layer and the superscript 'o' represents trapping-free. The equivalent oxide thickness, EOT, is evaluated by,

$$EOT = t_{IL}\varepsilon_{SiO2}/\varepsilon_{IL} + t_{HfO2}\varepsilon_{SiO2}/\varepsilon_{HfO2}, \tag{3.12}$$

where ε_{SiO2}, ε_{IL}, and ε_{HfO2} are the dielectric constant of SiO_2, the interfacial layer (Hf-silicates), and HfO_2, respectively. t_{HfO2} and t_{IL} is the thickness of HfO_2 and the interfacial layer, and ε_0 is the permittivity of free space. The voltage drop over each layer of the stack (see Fig.3.14) is determined from,

$$V_{IL}^{0} = \frac{V_{ox}^{0}}{EOT}\frac{t_{IL}\varepsilon_{SiO2}}{\varepsilon_{IL}} \tag{3.13}$$

and

$$V_{HfO2}^{0} = V_{ox}^{0} - V_{IL}^{0}. \tag{3.14}$$

The gate voltage is evaluated by,

$$V_g^0 = V_s + V_{ox}^0 + \psi_{ms}, \tag{3.15}$$

where ψ_{ms} is the work function difference.

Evaluation of T_{WKBij}: Once the potential drop over the dielectric is known, the tunnelling probability can be evaluated. As a first-order estimation, T_{ij} is evaluated based on the WKB approximation **[Yang99]**,

$$T_{ij} = T_{WKBij} T_{Rij}, \tag{3.16}$$

where T_{Rij} is a correction factor, taking into account of reflections at interfaces due to energy band discontinuities and its calculation can be found in **[Yang99]**. The WKB tunnelling probability, T_{WKBij}, can be expressed as **[Yang00]**,

$$T_{WKBij} = \exp\left[\frac{-2\sqrt{2m_{ox}}}{\hbar}\int_0^x \sqrt{\Phi(x) - E_{ij}}\,dx\right], \tag{3.17}$$

where m_{ox} is the effective mass of tunnelling electrons in the oxide and $\Phi(x)$ is the bottom edge of conduction band in the oxide. As shown in Figs.3.14a-d, there are four possible cases for electron tunnelling.

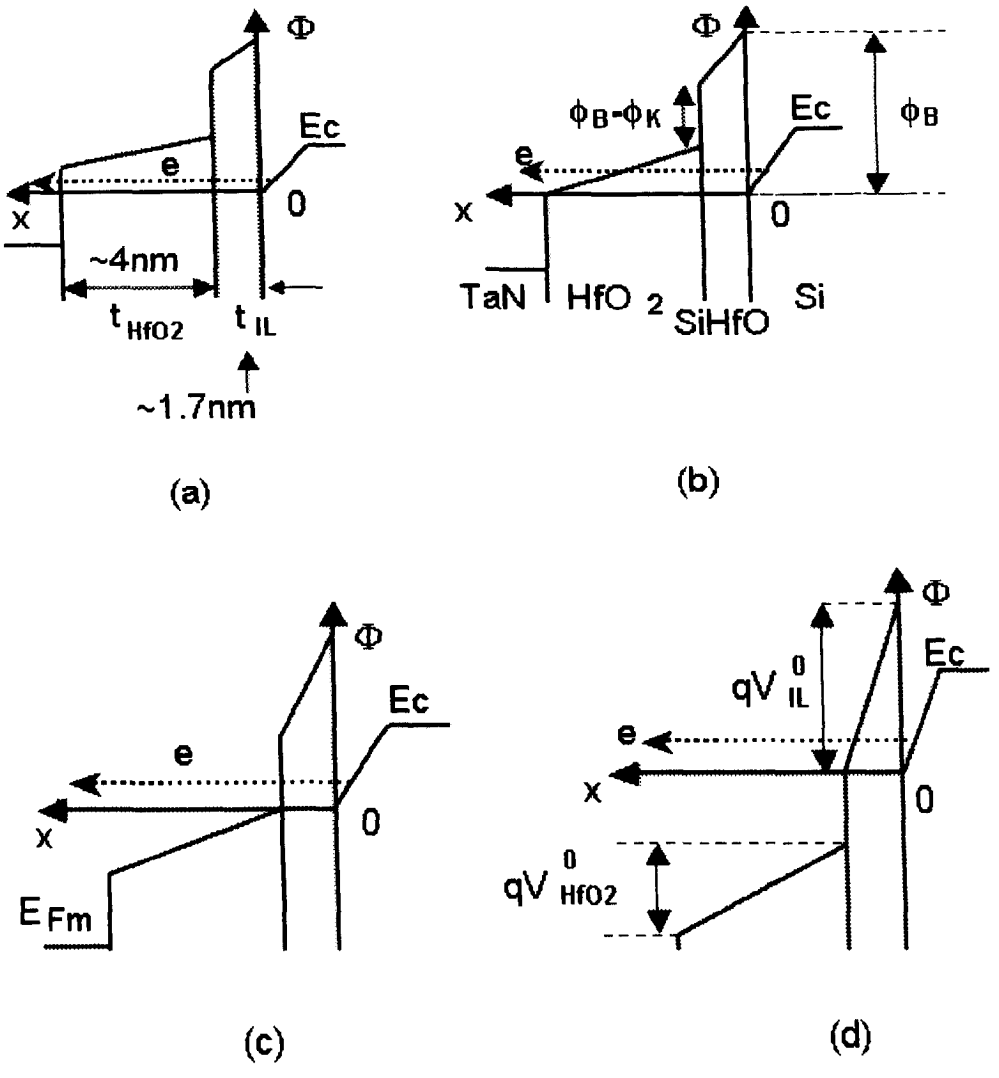


Fig.3.14 Schematic energy band diagrams for different tunnelling regimes. The gate bias increases from the case (a) to (d). (a) shows direct tunnelling through both HfO_2 and interfacial layers (IL). (b) is the case where electrons emerging from the conduction band of HfO_2. (c) illustrates that electrons only tunnel through IL directly. Finally, Fowler-Nordheim (FN) tunnelling occurs through the IL in (d). ϕ_B in (b) is the conduction band offset between HfO_2 and silicon.

Fig.3.14a shows direct tunnelling through both dielectric layers at low V_G. An increase of V_G leads to electron emerging from the conduction band of HfO_2 (Fig.3.14b). Fig.3.14c illustrates that electrons only tunnel through the interfacial layer directly at higher V_G. Finally, Fowler-Nordheim tunnelling through the IL occurs (Fig.3.14d).

Calculation of the J_g: The impact frequency at the dielectric/substrate interface in Eq.(3.7) is calculated from **[ZhaoY04], [Govo04],**

$$f_{ij} = \frac{E_{ij}}{2 A_j \hbar}. \tag{3.18}$$

Finally, J_g is obtained from Eq. (3.7). The gate current, I_g, is obtained from the product of J_g with gate area. We found that the I_g is dominated by the first sub-band in the two valleys. Higher sub-bands have lower N_{ij} and contribute little to I_g

The result is presented in Fig.3.15 below.

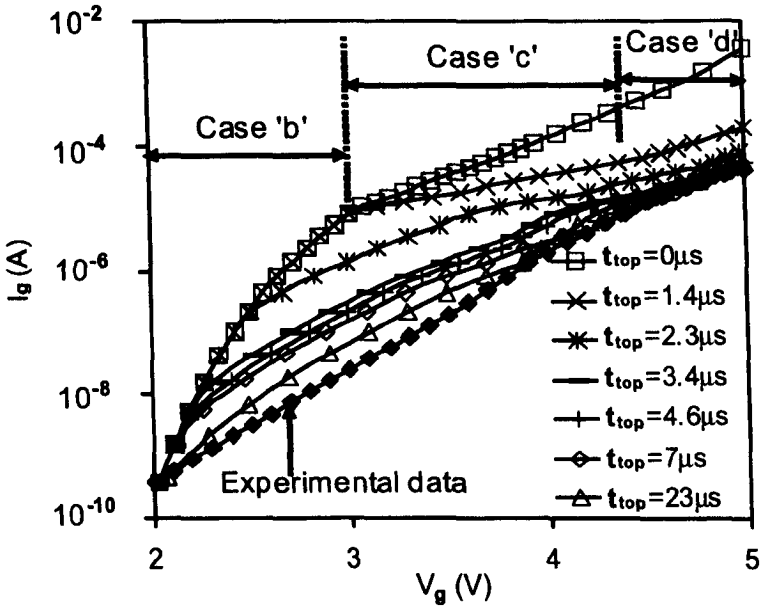


Fig.3.15 A comparison of the measured and calculated gate currents with various pulse time, t_{top}. The solid curves are guides-for-the-eye. The top curve (t_{top}=0) represents the trapping-free case. The range of V_G corresponding to each tunnelling regime of Figs.3.14b-d is marked out for t_{top}=0. The case corresponding to Fig.3.14a is not shown, since it occurs at V_G<2V. An increase of either V_G or t_{top} leads to a decrease of I_g and it is clear that trapping can reduce I_g substantially. At high V_G and t_{top}, trapping saturates and, consequently, the difference between the measured and calculated I_g disappears and I_g becomes insensitive to t_{top}

The top-curve is for the trapping free case, while the cases with trapping will be described in the next section. The parameters used for the simulation are given in Tables 1 and 2. It should be noted that we have a Hf-silicate interfacial layer **[Schr05]** and the energy barrier height at the IL/substrate interface, ϕ_B, and the effective electron mass, m_{IL}, were not known. ϕ_B was set at 2.35eV and m_{IL} at $0.45m_o$, so that the calculated I_g

agrees with the measured I_g, when trapping is insignificant at V_G=2V. m_o is the free electron mass. Since ϕ_B=3.15eV for SiO_2 and 1.5eV for HfO_2, ϕ_B=2.35eV is reasonable for Hf-silicates. $m_{IL}=0.45m_0$ is also acceptable for Hf-silicates, because of $m_{HfO2}=0.18m_0$ and $m_{SiO2}=0.5m_0$ **[Govo04]**.

[Schr05] ψ_{ms} (V)	ε_{IL}	[Schr05] ε_{HfO2}	[Schr05] t_{IL} (nm)	[Schr05] t_{HfO2} (nm)	[Schr05] EOT (nm)	ϕ_B (eV)	[ZhaoY04] ϕ_k (eV)	m_{IL} (m_e)	[HonY03] m_{HfO2} (m_e)	[Govo04,Yang99] N_A (cm^{-3})
-0.55	6.84	20	1.7	4	1.75	2.35	1.5	0.45	0.18	5×10^{17}

Table 2 Parameters used for calculating gate current through the HfO_2/HfSiO stack. ϕ_k is the conduction band offset between HfO_2 and silicon. N_A is the substrate doping density. In equation (3.17), the effective electron mass m_{ox} is replaced by m_{HfO2} and m_{IL} in the HfO_2 and interfacial layer, respectively. All other symbols are defined in the text. The data were taken from the references in the square bracket. ε_{IL}=6.84 was obtained from Eq.(3.12) to give an EOT of 1.75nm. ϕ_B and m_{IL} were selected so that the calculated I_g agrees with the measured value when trapping is insignificant at V_G=2V.

The voltage range corresponding to each tunnelling case illustrated in Figs.3.14b-d is marked out on Fig.3.15. The direct tunnelling through both layers (Fig.3.14a) cannot be seen from Fig.3.15, since it occurs at a voltage less than 2V. When electrons entering into the conduction band in HfO_2 (Fig.3.14b), the current rises rapidly with voltage. As V_G reaches 3V approximately, electrons only directly tunnel through the IL (Fig.3.14c). Since the tunnelling distance does not reduce for higher V_G in this case, I_g rises relatively slowly, which explains the smaller $\log(I_g)$~V_G slope in the case 'c' shown in Fig.3.15. Finally, Fowler-Nordheim (FN) tunnelling takes place through the IL when V_G> 4.4V.

Fig.3.15 clearly shows that the calculated I_g by assuming trapping-free can be two orders of magnitude higher than the measured I_g, where trapping took place. This confirms our expectation and we will attempt to estimate the I_g in the presence of trapping next.

3.4.3.2 Numerical modeling of gate current with trapping

Although many articles were published on the calculation of electron tunnelling through high-k/IL stacks **[ZhaoY04]**, **[Govo04]**, **[Muda00]**, **[FanY02]**, there is hardly any work that was carried out in the presence of transient trapping. When trapping occurs, the potential distribution profile in the dielectric changes with time. The main difficulty is how to estimate this transient distribution. It is overcome here by experimentally measuring the ΔV_G induced by trapping and some typical results are given in Fig.3.16. Once the ΔV_G at a given bias and time is known, a first order estimation of the potential distribution and the corresponding gate current can be made, as described below.

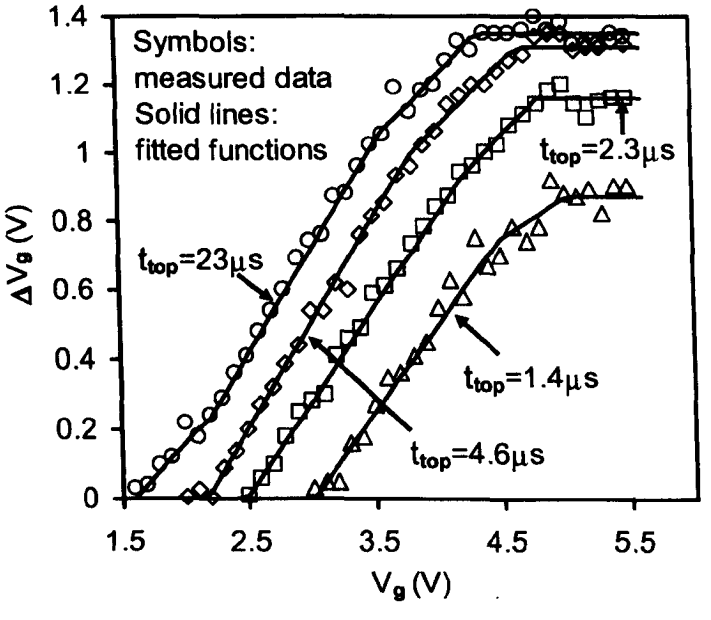


Fig.3.16 The electron trapping when different gate voltages were applied for a time of t_{top}. The solid lines are fitted functions. An increase of either voltage or t_{top} enhances trapping.

To determine the potential distribution in the dielectric for a given ΔV_G, the spatial distribution of trapped charges is required. Unfortunately, this information is not available at present. Since it is known that as-grown electron traps are negligible in SiO_2 **[Lerou04]**, **[ChangM06]** and Hf-silicate **[Shan03]**, the traps must locate either at the

HfO_2/IL interface or in the bulk of HfO_2. By examining how trapping depends on the HfO_2 thickness, we can rule out that traps are concentrated at the HfO_2/IL interface as demonstrated in part 3.3.3 **[Sim05], [ZhangJ06].** Early work **[Reim05]** also indicates that trapping occurs in the bulk of HfO_2. As a first order estimation, we assume that the centroid of trapped charges is at the middle of HfO_2 layer.

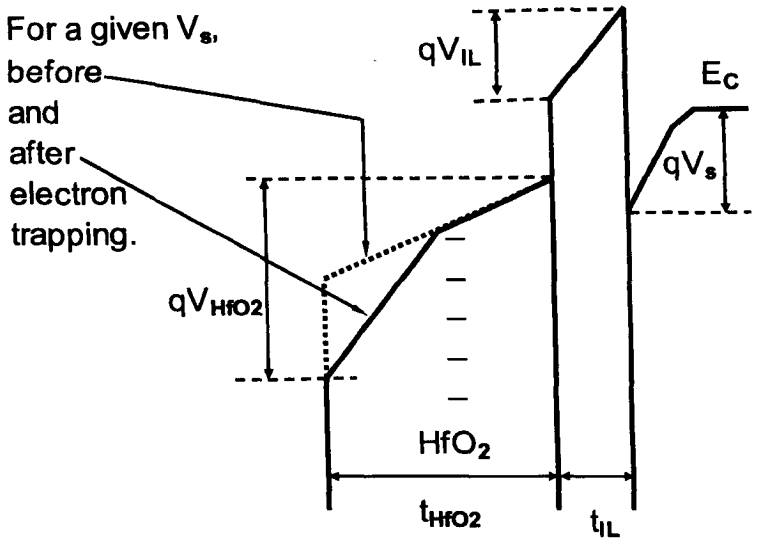


Fig.3.17 Schematic energy band diagram, showing the effect of electron trapping on potential distribution for a given substrate surface potential. The charge centroid is assumed being at the centre of HfO_2. The dotted and solid lines represent the cases without and with electron trapping, respectively. Trapping leads to an increase in the electrical field near the gate.

The calculation follows the same procedure as that without trapping, apart from the evaluation of potential drop in the dielectric. Here, the presence of trapped charges must be taken into account. As shown in Fig.3.17, for a given substrate surface potential, V_s, trapping in HfO_2 does not affect the potential drop over the IL, but increase the voltage drop over HfO_2.

The total potential drop over the stack is now,

$$V_G = V_G^0 + \Delta V_G. \tag{3.19}$$

When there is no trapping, V^0_G is fixed for a given V_s. Since trapping increases with time, ΔV_G and V_G will be a function of time for a given V_s now. As a result, to find a definitive ΔV_G and V_G, one has to fix both V_s and time.

As expected, Fig.3.16 shows that an increase of either gate voltage or time (t_{top}) results in higher trapping. Once V_s and t_{top} are fixed, V^0_G is known and we can set ΔV_G at a guessed initial value. This will give us two V_G: one from Eq.(3.19) and the other from the relevant curve in Fig.3.16. If these two V_G do not agree with each other, ΔV_G will be adjusted until their difference becomes negligible. Fig.3.15 shows the calculated I_g in the presence of trapping. For a given t_{top}, at relatively lower V_G, trapping is negligible and I_g agrees with the trapping-free value. An increase of V_G enhances trapping and brings down I_g. For a given V_G, an increase of t_{top} leads to higher trapping and lower I_g. It is important to note that when V_G and t_{top} are sufficiently high, the calculated I_g agrees with the measured value and becomes insensitive to t_{top}. This is because trapping reaches saturation now, as shown in Fig.3.10, so that I_g becomes insensitive to t_{top}. The reasonable agreement between the measured and calculated I_g at saturation makes us believe that the calculated I_g is acceptable as a first order approximation.

3.4.4 Capture cross section extraction

3.4.4.1 Electron fluency estimation

To determine the capture cross section of electron traps, electron fluency must be known. From Fig.3.15, the transient gate current for a given V_G can be obtained. One example is shown in Fig.3.18 for V_G=3.5V. As expected, Ig can drop by two orders of magnitude before reaching the measured value.

When electron fluency, N_{im}, was calculated from Eq.(3.1) based on the measured gate current, its 1st term is schematically represented by the grey region in Fig.3.18.

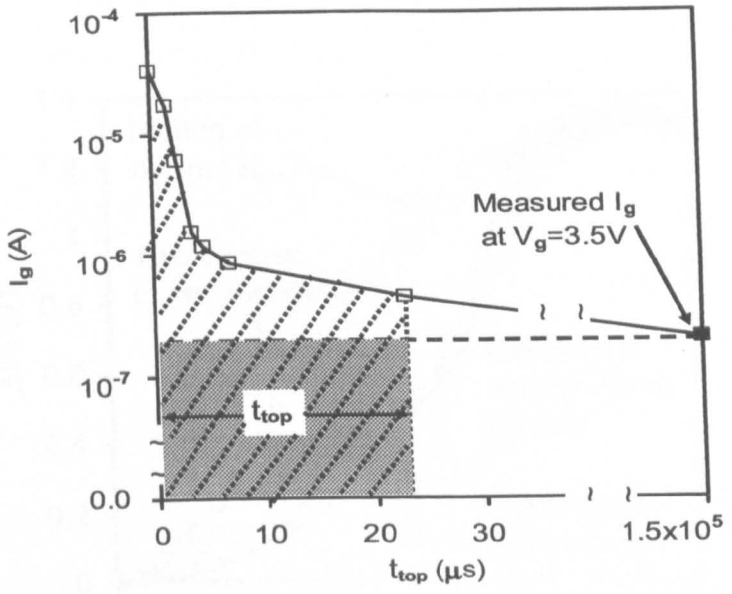


Fig.3.18 The calculated transient gate current when V_G=3.5V was applied to the gate. For comparison, the measured I_g is also shown. Trapping reduces I_g by two orders of magnitude and the substantial reduction occurs within microseconds. The grey area schematically represents the 1st term of Eq. (3.1), where the measured gate current was used to determine the electron fluency and the transient of I_g was not taken into account. The striped area schematically shows the 1st term of Eq.(3.20), where the transient gate current was used to calculate electron fluency. Note the logarithmic scale for the current.

By integrating the transient Ig against time, a more accurate electron fluency, N_{ic}, can be obtained by,

$$N_{ic} = \frac{\int_0^{t_{top}} I_g dt}{qLW} + \frac{\Delta V_g \varepsilon_{HfO2} \varepsilon_0}{t_{HfO2}/2} \quad , \tag{3.20}$$

where L and W is channel length and width, respectively. The 1st term of Eq.(3.20) is schematically represented by the striped region in Fig.3.18.

In Fig.3.19, the trapping-induced ΔVg is plotted against both N_{im} and N_{ic}. At low fluency, trapping is negligible and there is little difference in these two. As the fluency

and trapping increase, N_{ic} can be over one order of magnitude larger than N_{im}. In the saturation region, the relative difference between N_{im} and N_{ic} reduces, because the transient period of Ig becomes relatively less important

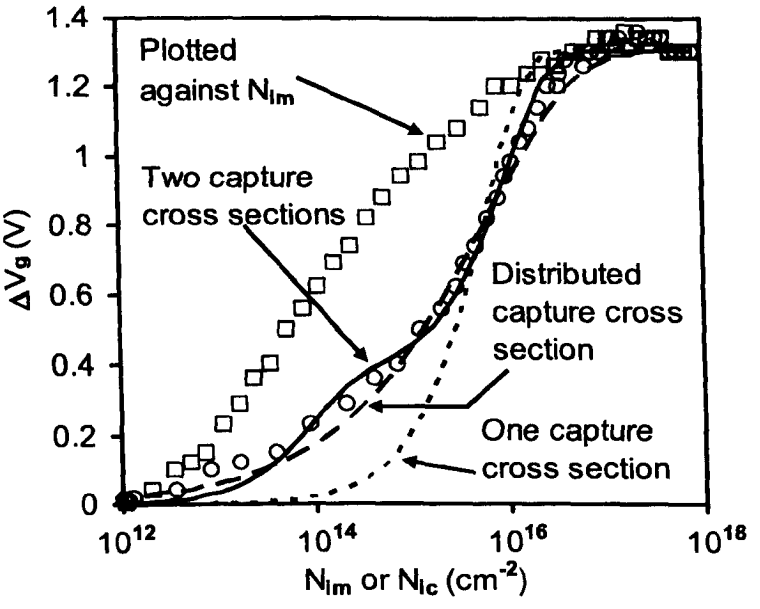


Fig.3.19 A comparison of electron fluency calculated with and without taking the transient of gate current into account. For the symbol '□' and 'o', the same data were used for the vertical axis, ΔV_G, but different data were used for the horizontal axis. The symbol '□' was plotted against N_{im}, calculated from Eq.(3.1) without taking the transient of gate current into account. The symbol 'o' was plotted against N_{ic}, calculated from Eq. (3.20) where the transient gate current is used. It shows that N_{ic} can be one order of magnitude higher than N_{im}. The dotted and solid lines were obtained by fitting the data with Eq.(3.23) when using one and two discrete capture cross sections, respectively. The dashed line is obtained by fitting the data with Eq.(3.24), assuming there is a distribution of capture cross sections.

3.4.4.2 Capture cross section extraction and trapping kinetics

We would like to assess the trapping kinetics measured by the pulsed I_D~V_G in Figs.3.2-3.4 and the range of capture cross sections from a direct observation of the data in Fig.3.19.

Figs.3.2-3.4 has several features. First, substantial trapping occurs before the electron fluency, N_{inj}, reaches $10^{14}cm^{-2}$. This indicates that there are traps with an

effective physical size in the order of $10^{-14}cm^2$ or larger. Second, the trapping clearly saturates for $N_{inj}>10^{16}cm^{-2}$, approximately. As a result, no traps smaller than $10^{-16}cm^2$ were observed. Third, there is a 'shoulder', as illustrated in Fig.3.3b. *This 'shoulder' does not originate from the use of different Vg, since Fig.3.3 shows that the trapping at a given N_{inj} is insensitive to Vg where the 'shoulder' appears.* After considering all models proposed in the past **[ZhangJ92]**, it is concluded that this 'shoulder' behavior is a signature of the 1st order trapping kinetics with two well separated capture cross sections, namely,

$$\Delta V_G=V_1[1-\exp(-\sigma_1 N_{inj})]+V_2[1-\exp(-\sigma_2 N_{inj})]. \quad (3.21)$$

The extracted capture cross sections, σ_1 and σ_2, are in the order of 10^{-14}~$10^{-13}cm^2$ and 10^{-16}~$10^{-15}cm^2$, respectively. The effective saturation density corresponding to σ_1 and σ_2 is estimated at $10^{13}cm^{-2}$ and $7\times10^{12}cm^{-2}$.

Early work **[Degra03]** estimated that the traps responsible for breakdown had a size of $0.8\times10^{-14}cm^2$. This is close to σ_1, since the accuracy of σ_1 will not be better than a factor of 4 **[ZhangJ92]**. The capture cross sections obtained here are much larger than those reported in **[Zafar03]**, based on,

$$\Delta V_G=V_o\times\{1-\exp[-(\sigma_o\times N_{inj})^{\beta}]\}. \quad (3.22)$$

By fitting the data measured by DC I_D~V_G, the extracted σ_o was in the order of $10^{-19}cm^2$ in **[Zafar03]**. When we fit our data measured by DC I_D~V_G in Fig.3.2a with Eq.(3.22), the extracted σ_o is also in the order of $10^{-19}cm^2$. As a result, this difference in capture cross sections is caused by different measurement techniques.

We would like to point out that there are some uncertainties, which could affect extracted capture cross sections. First, even for a fresh device at room temperature, Jg should have a component of TAC, which contributes little to the trapping. Second, Jg was assumed to be a constant during filling under a given V_G. In reality, Jg will drop as

trapping increases. At this stage, we can only give an 'order-of-magnitude' estimation for capture cross sections.

Further, as seen considerable trapping occurs before electron fluency reaches $10^{14}cm^{-2}$, indicating that there are traps with an effective capture cross section in the order of $10^{-14}cm^2$. There is little further trapping for $N_{ic}>10^{17}cm^{-2}$. This means that there are no as-grown traps of a capture cross section less than $10^{-17}cm^2$. To extract more definitive values for capture cross sections, a trapping model is needed.

As described in chapter 2 section 2.3.2, we have selected test conditions to ensure that detrapping is suppressed. In this case, there are two popular trapping models. One is the first order model with discrete capture cross sections **[ZhangJ92]**, **[ChangM06]**,

$$\Delta V_G = \sum_{k=1}^{K} \Delta V_{Gk} \left(1 - e^{-\sigma_k N_{ic}}\right), \tag{3.23}$$

where K is the number of discrete capture cross sections, σ_k. The other assumes that the value of capture cross section has a distribution **[Zafar03]**,

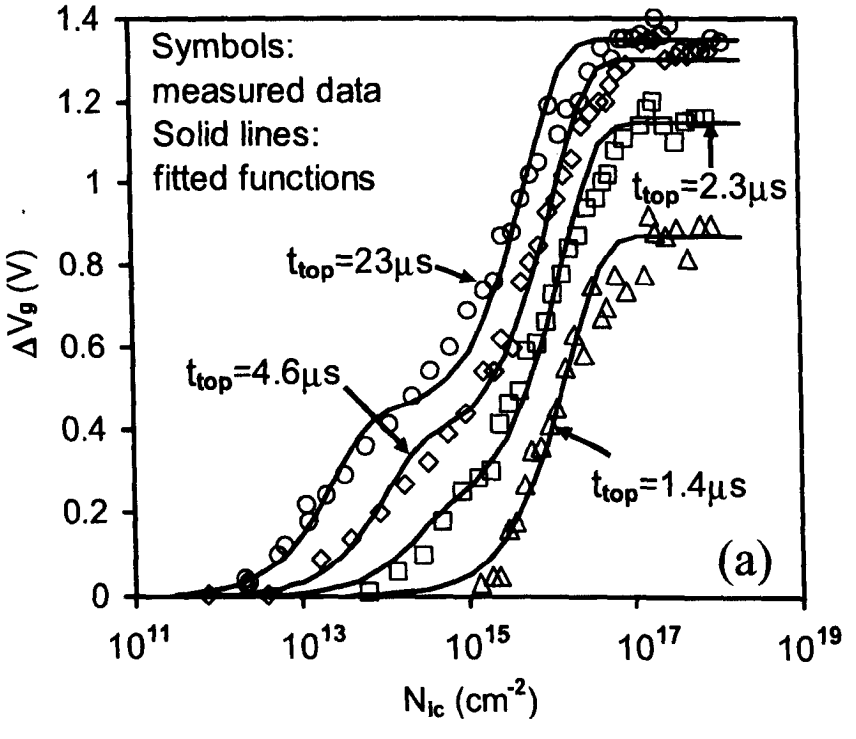
$$\Delta V_G = V_o[1 - e^{-(\sigma_o N_{ic})^{\beta}}] \tag{3.24}$$

If we use Eq.(3.23) with a single discrete capture cross section (K=1), the dotted line in Fig.3.19 shows that it does not agree with the data. The trapping occurs over a much larger range of N_{ic} than that covered by a single capture cross section.

Fig.3.19 shows that better agreement can be obtained by using either two discrete capture cross sections (K=2) in Eq.(3.23) or a distributed capture cross section

model, Eq.(3.24). In both cases, the number of adjustable parameters increased, when compared with a single discrete capture cross section.

The question is whether the improved agreement is a consequence of this increase of adjustable parameters. It is true that a good agreement with data can always be obtained if a large number of adjustable parameters are used and this is the reason why many researchers have their reservations with these models. In the following, we attempt to show that the improved agreement is genuine and to find if the capture cross section is distributed or discrete.



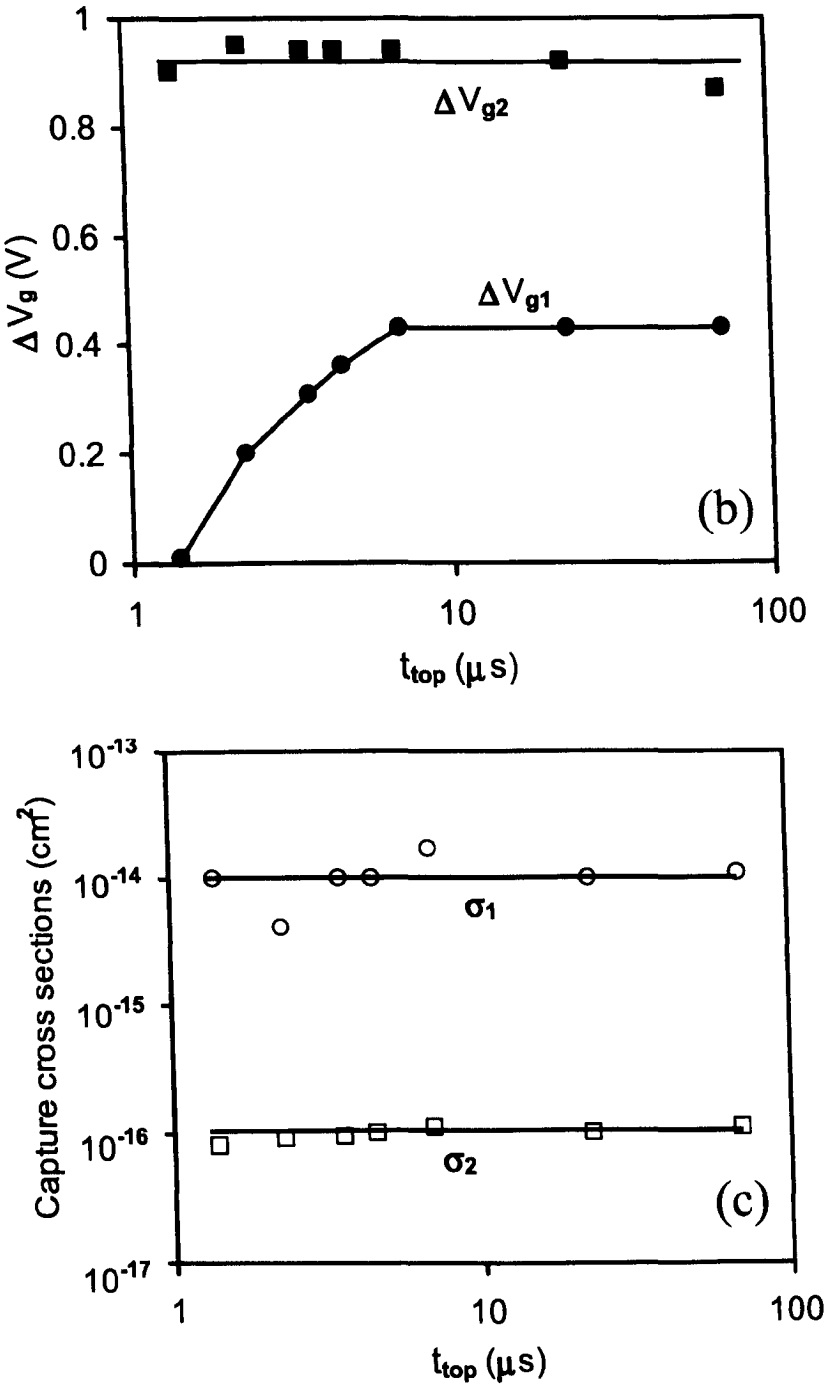


Fig.3.20 a,b&c Dependence of trapping on the filling time, t_{top}. (a) shows the trapping can reduce with the filling time. Solid curves were obtained by fitting data with two discrete capture cross sections. (b) and (c) show that the extracted saturation levels and capture cross sections, respectively. The extracted two capture cross sections are in the order of $10^{-14}cm^2$ and $10^{-16}cm^2$, respectively and they are insensitive to t_{top}. In (b), the trapping level by the smaller trap, ΔV_{g2}, which controls the further trapping for $N_{ic}>10^{15}cm^{-2}$ in (a), does not depend on t_{top}, either. In contrast, the trapping level by the larger trap, ΔV_{g1}, which corresponds to the region of $N_{ic}<10^{15}cm^{-2}$ in (a), clearly reduces for smaller t_{top}.

Fig.3.20a shows that a reduction of trap filling time, t_{top}, results in a progressive reduction of trapping in the relatively low range of N_{ic} (10^{12}~$10^{15}cm^{-2}$). However, further trapping in the relatively high range of N_{ic} ($>10^{15}cm^{-2}$) is hardly affected. Fig.3.20b plots the extracted two saturation trapping levels, ΔV_{g1} and ΔV_{g2}, corresponding to the capture cross section σ_1 and σ_2, given in Fig.3.20c, respectively. ΔV_{g1} decreases for $t_{top}<10\mu s$ and becomes negligible when $t_{top}=1.4\mu s$, but ΔV_{g2} changes little over this time range. As a result, the transition between ΔV_{g1} and ΔV_{g2} is 'discontinuous'. It is difficult to explain this discontinuous behaviour based on a continuous distribution of capture cross sections. The result strongly supports the existence of two types of traps with two discrete capture cross sections. For a 4nm ALD HfO_2 layer, the effective density is around $5.3\times10^{12}cm^{-2}$ and $1.1\times10^{13}cm^{-2}$ for the traps with a capture cross section in the order of $10^{-14}cm^2$ and $10^{-16}cm^2$, respectively.

3.4.5 Dependence on fabrication processes and techniques

After extracting the effective density and capture cross section for electron traps, we now explore the sensitivity of these trap properties to the fabrication processes and techniques. Before comparing results reported for different samples, it should be emphasized that, even for the same test sample, both the apparent effective density and capture cross section are highly sensitive to measurement conditions, such as the measurement time, gate bias, and trap filling time.

For example, Fig.3.12 shows that an increase of measurement time can reduce both the apparent density and capture cross section significantly. Thus, any comparison with results reported by other groups must be made under the same measurement conditions. It is also reported that trapping reduces for thinner HfO_2 layer **[Sim05]**, **[ZhangJ06]**. This thickness effect must be taken into account when making comparison.

It is found that the effective density obtained in our samples is within a factor of 2 of the values reported by earlier works for HfO_2 prepared by atomic layer deposition (ALD) **[Degra03], [ZhangW02]** and metal organic chemical vapor deposition (MOCVD) **[Bers04], [ShenC04]**. For other fabrication techniques, the differences in trap density can be larger.

For example, Fig.3.21a shows that the trapping in a 2nm HfO_2 prepared by the Physical Vapour Deposition (PVD) technique can be substantially higher than the trapping in a 2nm ALD HfO_2. At this stage, we consider the effective trap densities reported here are typical values only for ALD and MOCVD HfO_2 layers.

On the capture cross section, we can reproduce the values in the range of 10^{-12}~$10^{-19}cm^2$ reported earlier **[Zafar03], [Bers04]** on our samples, if the same measurement and formula are used. The capture cross section is similar for ALD **[Yang99]** and MOCVD **[Muda00]** samples. Furthermore, Fig.3.21b compares the trapping kinetics in PVD and ALD layers. It is obvious that there is little difference in these two. This suggests that, unlike the trap density, the capture cross section is insensitive to fabrication processes and techniques. The wide range of the apparent capture cross sections, 10^{-12}~$10^{-19}cm^2$, reported in early works **[Zafar03], [Bers04]** mainly originates from the difference in measurement techniques and conditions as shown in Fig.3.12, rather than the difference in fabrication processes and techniques.

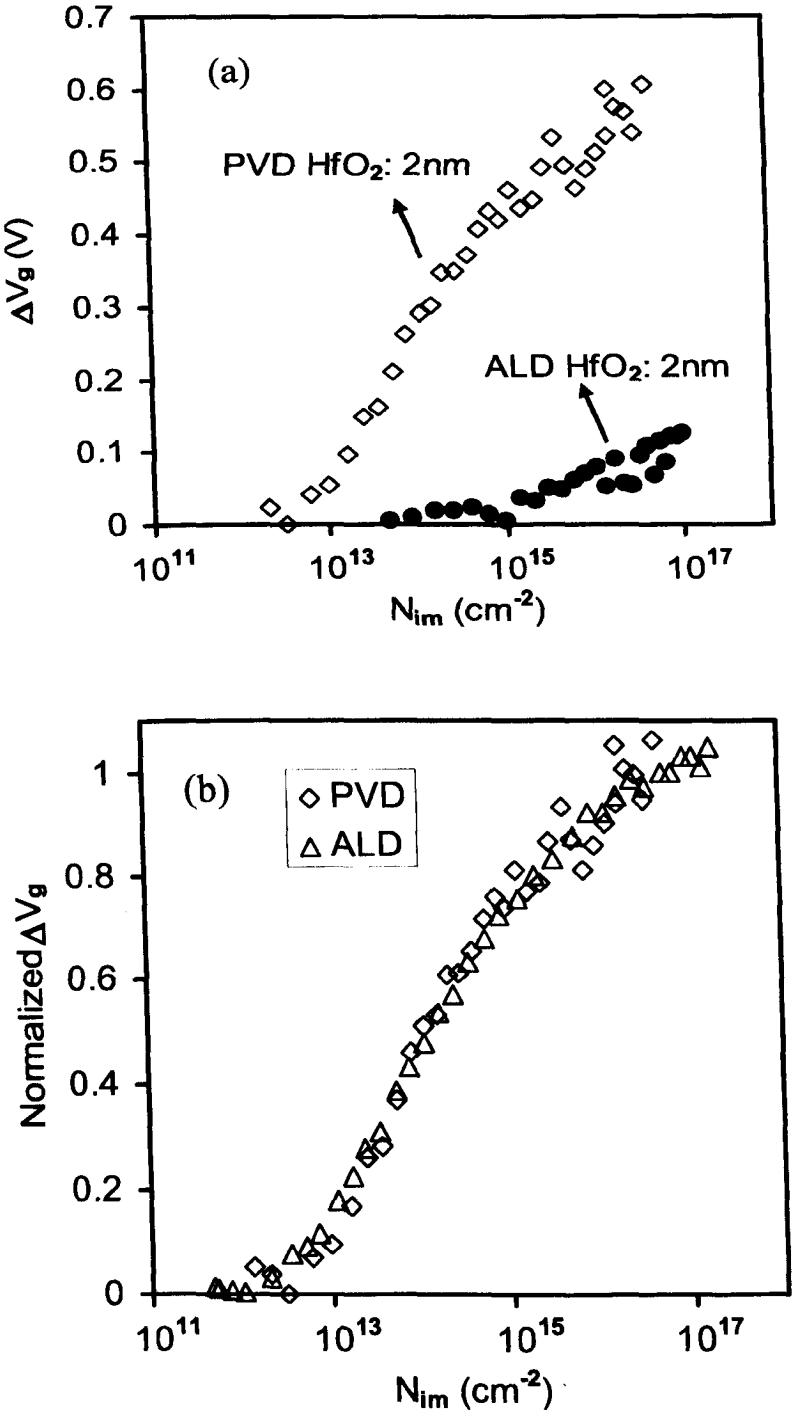


Fig.3.21a&b A comparison of electron trapping in HfO_2 prepared by ALD and PVD. (a) shows that trapping in a 2nm PVD HfO_2 is substantially higher than that in a 2nm ALD HfO_2. (b) shows that the trapping kinetics in a 2nm PVD HfO_2 is similar to that in a 4nm ALD HfO_2. The trapping for ALD sample is taken from a 4nm layer, since the trapping in a 2nm ALD sample is too low to give reliable trapping kinetics.

3.5 CONCLUSIONS

In this chapter, as grown electron traps in HfO_2/SiO_2 or HfO_2 / HfSiO stacks are investigated. The issues addressed include their properties and dynamic behaviour, their location, capture cross sections, and trapping kinetics. Following conclusions can be drawn:

The traditional DC I_D~V_G substantially underestimates the trapping, because of detrapping during the measurement. The trap-assisted conduction, which is responsible for SILC and the thermally enhanced current, contributes little to trapping.

On the location, it is ruled out that the traps are piled up at the HfO_2/HfSiO interface. A uniform distribution throughout HfO_2 layer does not agree with the test data either. The results support that trapping is negligible in the region around 1.3~1.8nm near to one or both ends of the HfO_2 layer, when compared with the trapping in the central region of HfO_2.

On the capture cross sections of as-grown electron traps in HfO_2 layer and trapping kinetics. The results support the trapping model with two well separated discrete capture cross sections, rather than a continuous distribution, indicating the presence of two different types of as-grown electron traps. The extracted values are in the order of $10^{-14}cm^2$ and $10^{-16}cm^2$, respectively. For a 4nm ALD HfO_2 layer, the effective trap density is around $5.3\times10^{12}cm^{-2}$ and $1.1\times10^{13}cm^{-2}$ for the larger and smaller traps, respectively. These densities are typical values for 4nm HfO_2 layers prepared by ALD and MOCVD techniques, but trap density can vary substantially for HfO_2 fabricated by other techniques, such as PVD. In contrast, it is found that the capture cross section is insensitive to the fabrication technique.

CHAPTER 4 GENERATED ELECTRON TRAPS AND BREAKDOWN

4.1 INTRODUCTION

Charge Pumping (CP) has shown to be a suitable technique to investigate defects in high-k based devices. In this chapter defects generation in SiO_2/HfO_2 stacks are investigated using CP and Time Dependent Dielectric Breakdown (TDDB).

In section 4.2.1 Base level and Amplitude sweep CP is used to investigate the charging and discharging effect on HfO_2 bulk with different gate electrodes at different frequency. Furthermore the influence of the channel length on the charging/ discharging of the bulk HfO_2 defect are studied.

Section 4.3.1 presents results based on charging pumping technique by variable frequency to investigate the bulk trap density on both fresh and degraded high-k dielectrics. Efforts were made to find out why the t_{BD} depends on the channel length of MOS transistors.

In section 4.3.2 we further develop the applicability of the variable frequency CP. It is shown that by independently controlling the pulse low timing, the "discharging time", and high level timing, the "charging time", we are able to separate the traps in the interfacial SiO_2 from the traps in the HfO_2 and observe the creation of new traps in both layers. This so called Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP) is used to investigate the creation of traps in the SiO_2 and HfO_2 separately in an ALD SiO_2/HfO_2 metal gate stack. The voltage acceleration of trap creation is also compared with the TDDB data.

Section 4.4 introduces new breakdown phenomena when metal gate electrode is used. For downscaled poly-Si gate MOSFET devices, reliability margin is gained by progressive wearout. However, when the poly-Si gate is replaced by metal gate, the slow wearout phase observed in ultra thin SiON and HfSiON dielectrics with poly-Si gate disappears, and with it, the reliability margin.

4.2 CONVENTIONAL CHARGE PUMPING (CP)

4.2.1 Base level and Amplitude sweeps charge pumping

In this section we will use the conventional charge pumping technique introduced in chapter 2 to study interface and bulk defect in SiO_2/HfO_2 with different gate electrodes. The conventional base level sweep with rather small amplitude is used to sense interfaces state, whereas amplitude sweep is applied to investigate in more details the interface and bulk defect in HfO_2. The details of the devices used are given below:

After an IMEC clean, nMOS transistors were prepared with an EOT of 1.9nm and Poly-Si gate electrodes. The dielectric stack consisted of 4nm (80cycles) ALCVD HfO_2 on top of 0.7nm interfacial oxide. The high–k deposition was followed by a 500 C anneal in oxygen and the poly-Si gate, source and drain were activated with a 1000 C spike. A final forming gas was performed at 520C for 20 min. *For simplicity, we will call this device PG (Poly-Si gate) device.*

nMOSFETs of TaN/TiN gate were also prepared. After an IMEC clean, the fabricated dielectric stack consisted of 4nm (80cycles) ALCVD HfO_2 with an EOT of 2.1nm. The high–k deposition was followed by a 500 C anneal in oxygen and the TaN/TiN gate, source and drain were activated with a 1000 C spike. A final forming gas

was performed at 520C for 20 min. *For simplicity, we will call this device MG (Metal Gate) device.*

4.2.1.1 Effects of charging and discharging of HfO_2 bulk.

To separate the contributions related to interface states from that associated with bulk defects the conventional base level charge pumping with small amplitude was applied to PG and MG device, Fig 4.1 and 4.2.

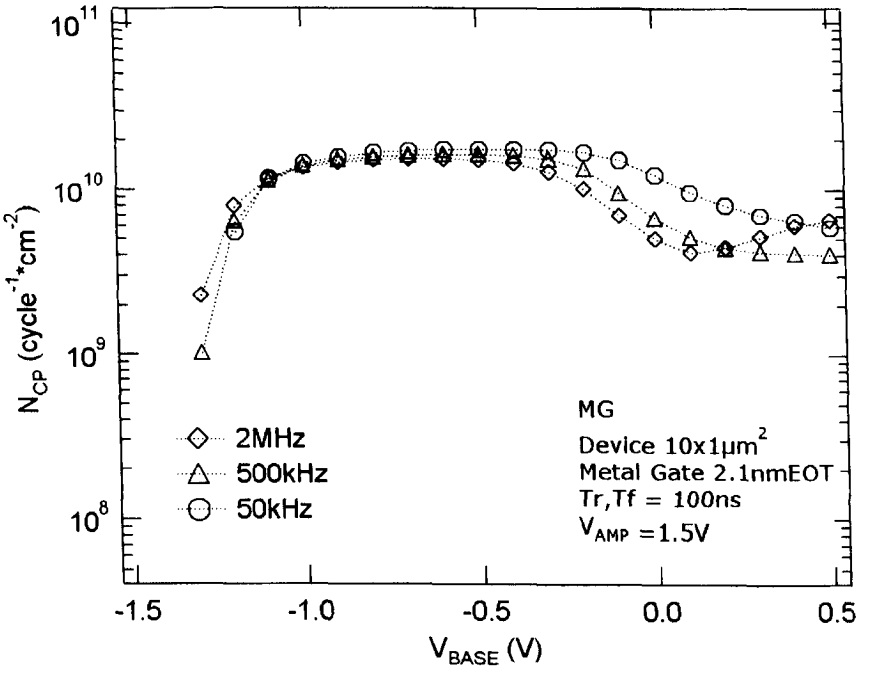


Fig.4.1 Charge per cycle measured by Base Level CP on SiO_2/HfO_2 MG using the parameters given in the figure.

When a voltage pulse with a small amplitude (V_{AMP}= 1.5V) is used to measure the charge pumping characteristic, the interface state densities (N_{CP}) are similar to that of the control oxide. Furthermore, the charge pumping current scales well with frequency, indicating no significant contribution from slow charging and discharging effects under these bias conditions.

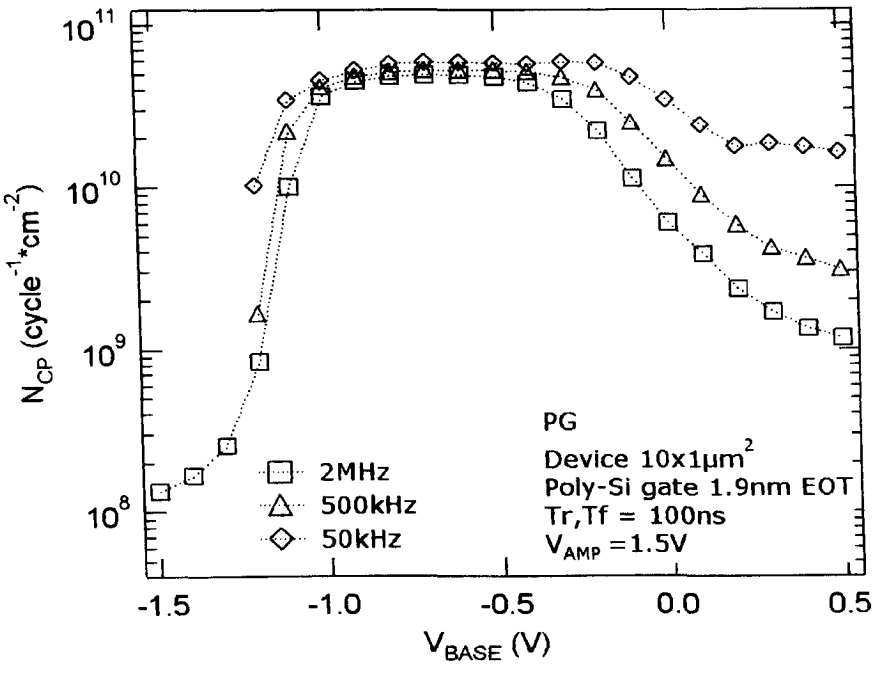


Fig.4.2 Charge per cycle measured by Base Level CP on SiO_2/HfO_2 PG using the parameters given in the figure.

Note that the charge per cycle doesn't scale with frequency when the Vbase is between -0.4V and 0.5V in figs 4.1 and 4.2. As the transistor have to switch on when the top level of the pulse applied at the gate reach the threshold voltage, the charge pumping current measured before this stage is drown by the leakage current which make the extraction of the charge per cycle difficult and as the frequency decrease (i.e 50kHz) the leakage increase.

When an amplitude sweep instead of the conventional base level sweep is applied, the contribution from charging and discharging of bulk defect states becomes clearly evident, as shown in figs 4.3 & 4.4 (lin/lin scale) figs 4.5 & 4.6 (log/lin scale). The charge per cycle strongly increases with increasing peak level and charging time, MG devices shows less interface traps, but higher bulk traps compare to PG devices.

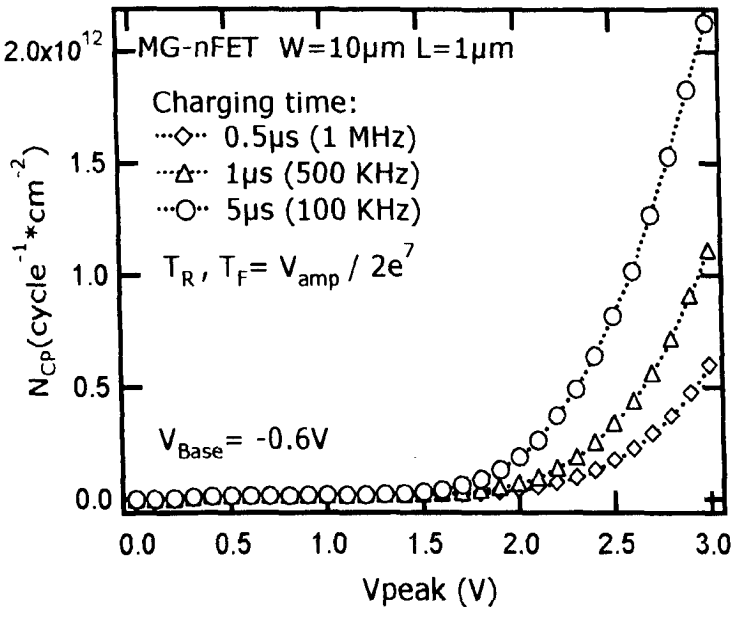


Fig.4.3 Charge per cycle measured by CP on SiO_2/HfO_2 metal gate using the amplitude sweep and parameter given in the figure. A strong increase in N_{CP} can be observed according to charging / discharging time on bulk defects.

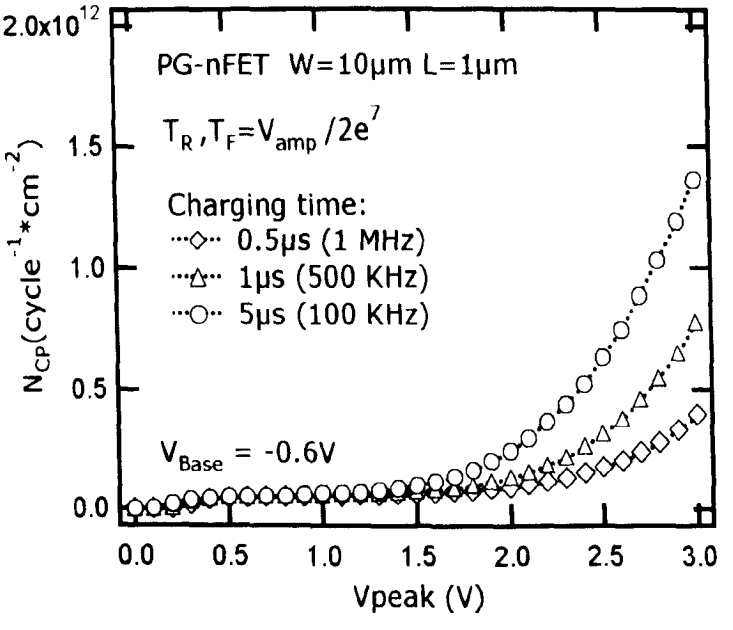


Fig.4.4 Charge per cycle measured by CP on SiO_2/HfO_2 Poly-Si gate using the amplitude sweep and parameter given in the figure. Less bulk defect can be observed when comparing with Metal Gate. (Fig 4.3)

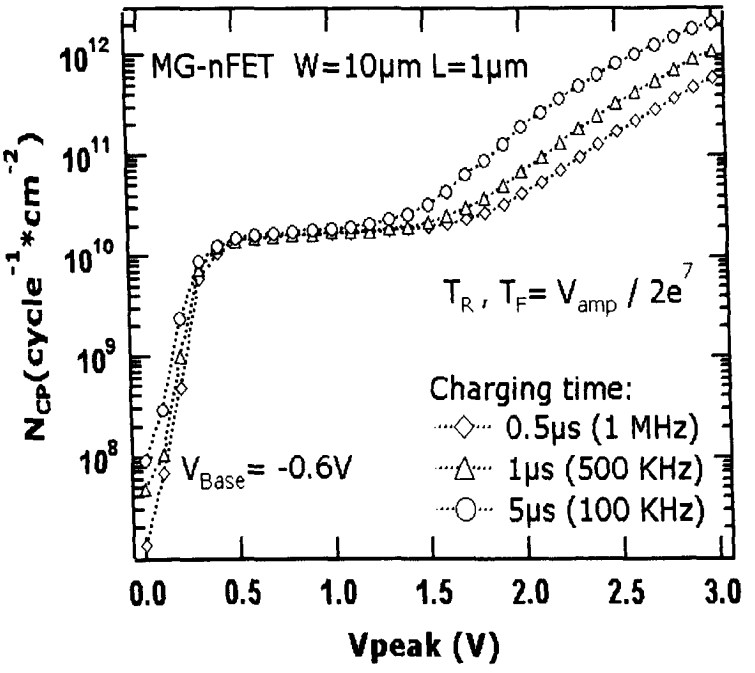


Fig.4.5 Charge per cycle measured by CP on SiO_2/HfO_2 metal gate using the amplitude sweep and parameter given in the figure. A similar interface state can be observed for different frequency.

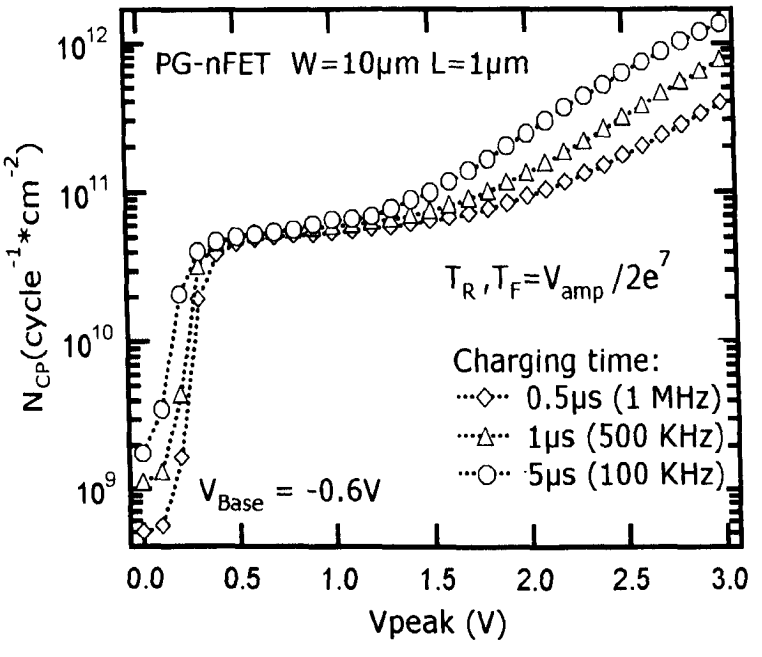


Fig.4.6 Charge per cycle measured by CP on SiO_2/HfO_2 Poly-Si gate using the amplitude sweep and parameter given in the figure. Higher interface state compare to Metal Gate.

4.2.1.2 Influence of the channel length on charging and discharging of HfO_2 bulk

To investigate the channel length dependence of the bulk trapping in SiO_2/HfO_2 dual layer stacks, the amplitude sweep is used with a sufficiently negative gate bias to have a complete detrapping throughout the entire sweep range. Electrons emitted from the HfO_2 layer to the substrate will recombine with the majority carriers and will contribute to the substrate current, so that the oxide charge will be measured. If pulses with identical duty cycle (=0.5) are used, the charging time is inversely proportional to the frequency. By varying the CP frequency we can therefore sense different fractions of the trap density. A high frequency corresponds to short charging time and only Si/SiO_2 interface traps and HfO_2 traps very close to the SiO_2/HfO_2 interface are pumped. A low frequency corresponds to a long charging time and traps deeper in the stack can also respond to the signal. Moreover, if the trap density sensed at high frequency is subtracted from the trap density measured at low frequency, a fraction of the bulk trap density only is obtained **[Degra03]**.

A high frequency of 1MHz is used to sense the Si/SiO_2 interface traps and HfO_2 traps very close to the SiO_2/HfO_2 interface. A relatively low frequency of 50 KHz is used to sense traps deeper in the stack. A long rise / fall time was used to minimize the geometrical component (V_{AMP} /$2e^7$) where V_{AMP} is the amplitude of the pulse; this factor is used to keep the same ramp rate at each amplitude.

From Figs. 4.7 to 4.10, the charge-pumping characteristic using the amplitude sweep for PG and MG at frequency 1 MHz and 50 KHz is shown for devices with different channel length. The charge carrier per cycle reduces strongly for short channel devices.

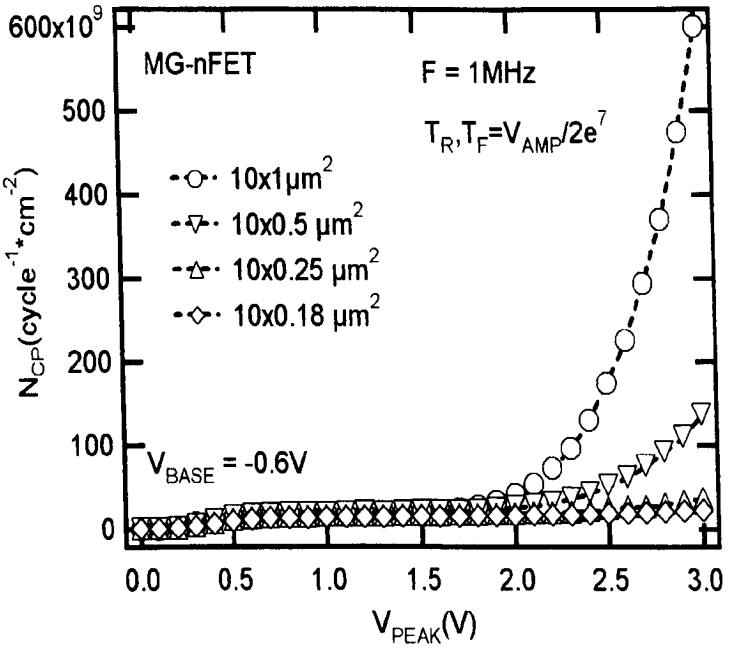


Fig.4.7 MG charge pumping characteristic versus V_{PEAK} for different channel length at 1 MHz using amplitude sweep technique. The current measured at the Si substrate strongly decreases with decreasing channel length.

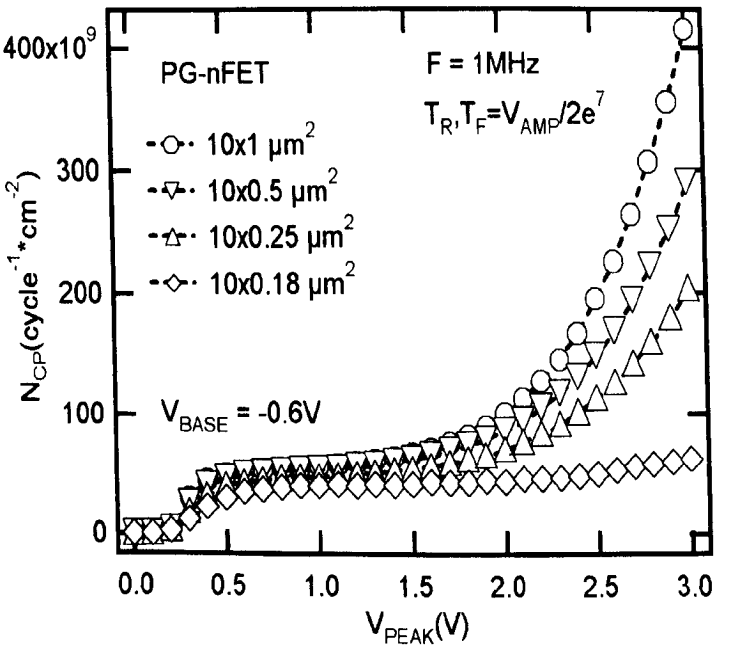


Fig.4.8 PG charge pumping characteristic versus V_{PEAK} for different channel length at 1 MHz using amplitude sweep technique. The current measured at the Si substrate strongly decreases with decreasing channel length.

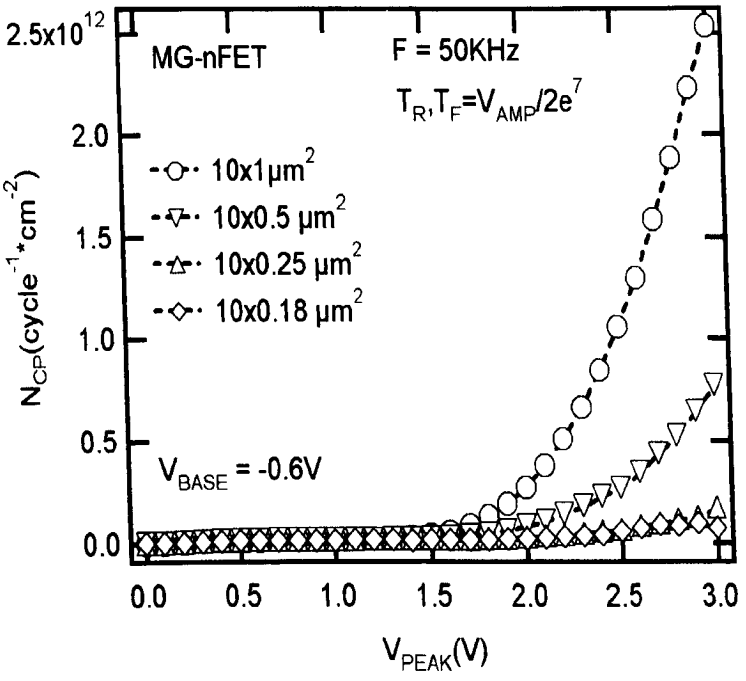


Fig.4.9 MG charge pumping characteristic versus V_{PEAK} for different channel length at 50 KHz using amplitude sweep technique. The current measured at the Si substrate strongly decreases with decreasing channel length.

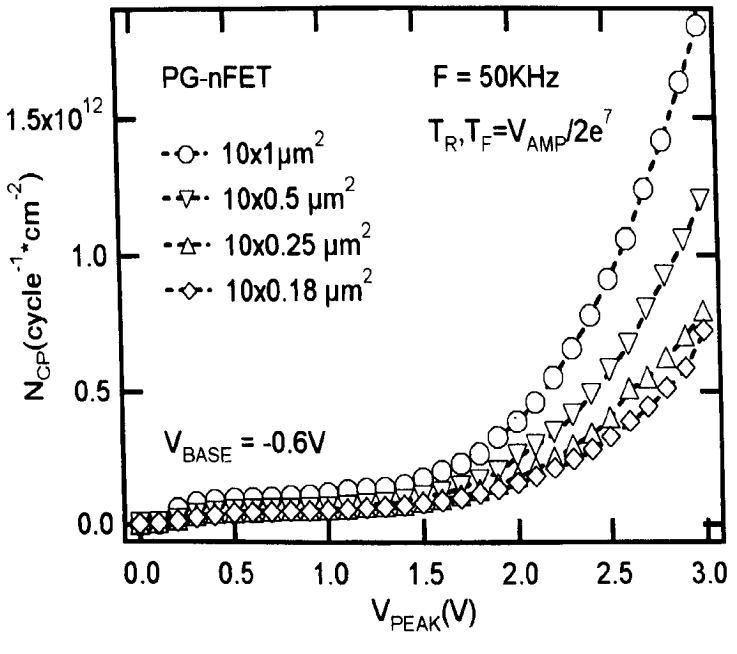


Fig.4.10 PG charge pumping characteristic versus V_{PEAK} for different channel length at 50 KHz using amplitude sweep technique. The current measured at the Si substrate strongly decreases with decreasing channel length.

When the charge per cycle is plotted versus channel length, Figs. 4.11 & 4.12 shows a clear dependence on the channel length independent of the gate electrode.

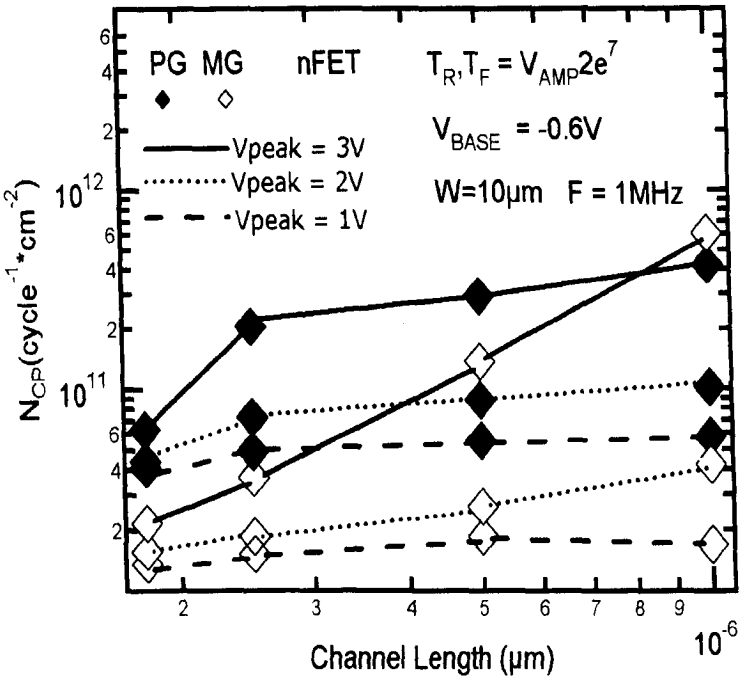


Fig.4.11 N_{CP} versus channel length obtained from data as shown in Fig 4.7 & 4.8 at frequency 1MHz. The decrease in N_{CP} with channel length is due to enhanced charge collection by the junctions.

The decrease of the charge per cycle for short channel is caused by charge loss to the source / drain junctions. When the trapped electrons are emitted from the HfO_2 layer back into the p-type Si substrate of the n-channel MOSFET they can diffuse into the substrate prior to recombination with majority carriers. The diffusion of electrons back to the source / drain junctions becomes more effective in devices with shorter channel length, which leads to a reduction of the recombination current measured at the Si substrate **[Ker04]**. For a correct assessment of the charge trapping using charge pumping technique, this effect has to be included accordingly.

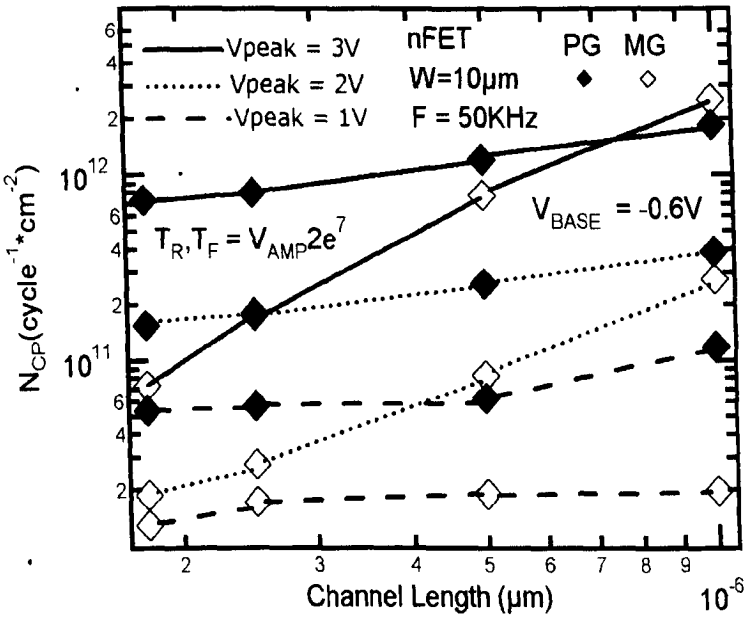


Fig.4.12 N_{CP} versus channel length obtained from data as shown in Fig 4.9 & 4.10 at frequency 50KHz. The decrease in N_{CP} with channel length is due to enhanced charge collection by the junctions.

In summary, charge pumping was used to study interface and bulk defect in SiO_2 / HfO_2 with different gate electrodes. The conventional base level sweep with rather small amplitude was used to sense interfaces states, whereas amplitude sweep was applied to investigate in more details the interface and bulk defect in HfO_2.

4.3 IMPROVED CHARGE PUMPING (CP)

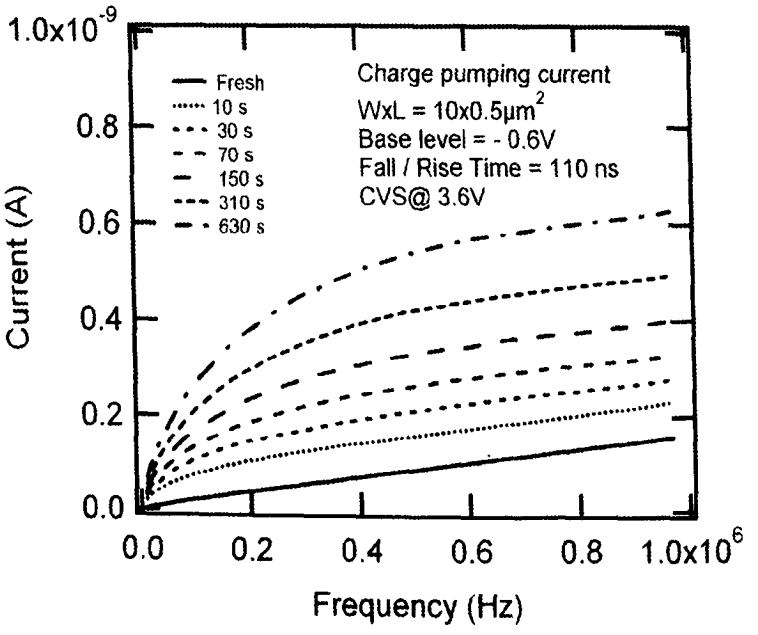
4.3.1 Variable Frequency CP

Variable frequency charge pumping has been shown to be a powerful analysis technique to investigate the bulk trap density on both fresh and degraded high-k dielectrics of MOS transistors.

Here, it will be used for the first time to investigate why the time to breakdown (t_{BD}) depends on the channel length in a SiO_2/HfO_2/metal gate MOS transistor. This channel length dependence is observed on several stacks with metal gates and it also appears, but less pronounced, for poly-Si gate devices.

The device taken in consideration is an nMOSFET transistors stressed in inversion, with either poly-Si (EOT 2.5nm) or metal gate (2.1nm EOT). The gate oxide dimensions WxL are $10x1\mu m^2$, $10x0.5\mu m^2$ and $10x0.25\mu m^2$. The fabrication process includes an IMEC clean, 80cycles with ALD HfO_2, a PDA @ 500°C, 1', O_2, activation of 1000°C, and a FGA at 520°C.

The CP current I_{CP} is measured at fixed base and top level, fixed rise and fall time, and variable frequency with a duty cycle of 50%. The typical results are shown in Fig 4.13



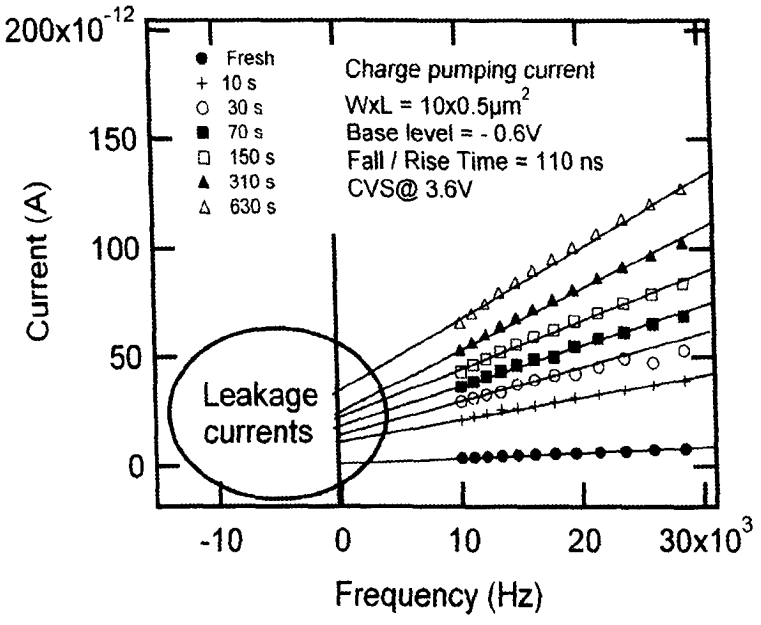


Fig.4.13 (Top) The figure on the top represents the I_{CP} current at low frequency. (Bottom). A linear extrapolation is used to determine the leakage current. The values in the inset are used to monitor the Icp current.

In its simplest form, the pumped charge per cycle Q_c shown in Fig 4.14 is expressed as follows:

$$Q_c = \frac{I_{CP}}{f} = qA[D_{it} + D_{ot}(f)] \qquad \text{with } D_{ot}(f) \propto D_{HFO}\log(f) \tag{4.1}$$

with q=elementary charge ($1.602\text{x}10^{-19}$ C), A=gate oxide area, f=CP frequency. D_{it} is the interface trap density at the substrate/SiO_2 interface.

This is the frequency independent CP component from the conventional theory **[Groes84]**. D_{ot} is the frequency dependent HfO_2 bulk trap density and its value also depends on the choice of the pulse amplitude and the base level voltage. At low frequency, the charging and discharging times during each pulse are long and traps deep in the HfO_2 bulk can respond to the applied signal. At high frequency, the charging and discharging times are shorter and only traps closer to the SiO_2/HfO_2 interface can respond to the signal. In first order, D_{ot} is proportional to log (f) with D_{HFO} the HfO_2 trap density per cm^2 that can be sensed per decade frequency change.

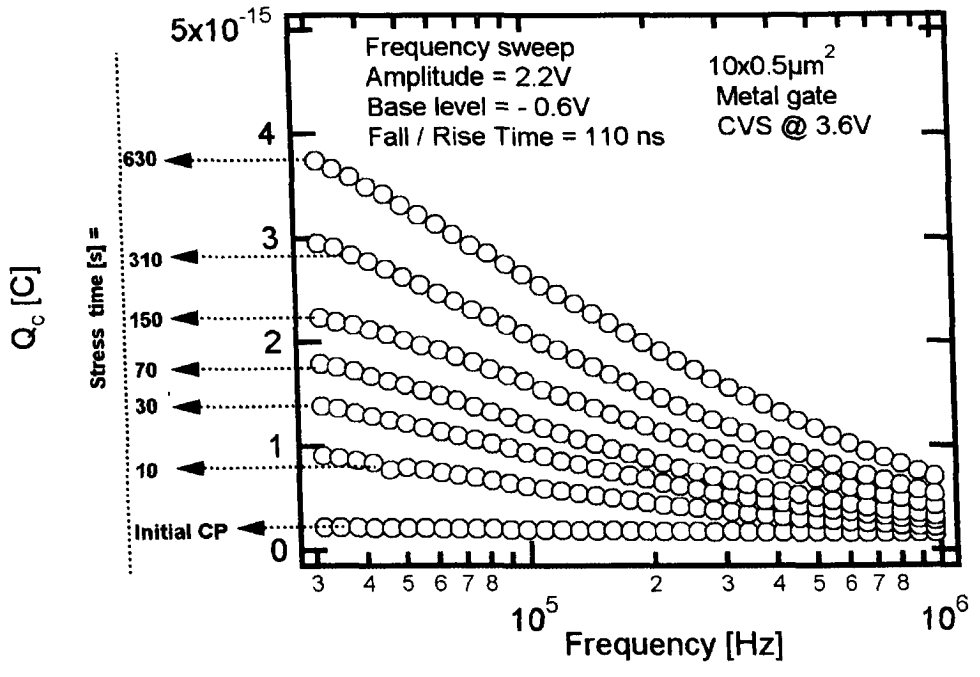


Fig.4.14 Pumped charge per cycle versus frequency. The device was stressed at CVS (3.6V) and the HfO_2 bulk trap density was calculated by taking the slope of each curve.

Reactions between the dielectric and the electrode, like in the case of HfO_2 and poly-Si, the formation of grains with highly defective grain boundaries, high temperature processing and mechanical stress lead to high defect densities in high-k gate stacks. In order to decrease the number of defects, various process steps like post-deposition anneals (PDA), forming gas anneals (FGA) or H_2O pulses after each high-k deposition cycle are included in the process flow. Unfortunately, the exact mechanisms of all these steps and their interactions are not very well understood yet.

Experiments however revealed a channel length dependence of the defect density: devices with shorter channel length have lower interface **[Chen03]** and bulk **[ZahM05]** defect densities than longer transistors located at the same die. (Fig.4.15)

In Fig. 4.15 this is demonstrated on HfO_2 dielectrics with both poly-Si and metal gate electrodes, where the extracted value of D_{HFO} on fresh devices for different channel length is presented. Two observations can be made: 1) poly-Si gate devices have higher

trap density compared with metal gate devices and 2) D_{HFO} increases with channel length.

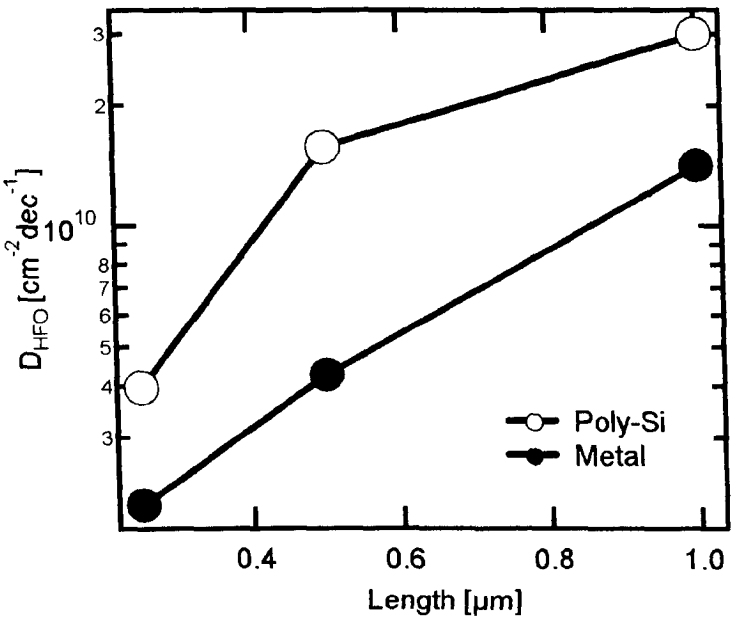


Fig.4.15 Extracted D_{HFO} values for fresh devices. Longest channels for poly-Si and metal gates show higher trap density.

When studying TDDB, differences in defect densities for the same gate stack can cause serious complications. According to the percolation theory a dielectric breaks down when a certain critical defect density is reached. Consequently, area scaling of time-to-breakdown distributions measured for various channel length under the same stress conditions and scaled to a reference area fails (Fig. 4.16).

Furthermore, Fig.4.17 shows the channel length dependence remains after stress.

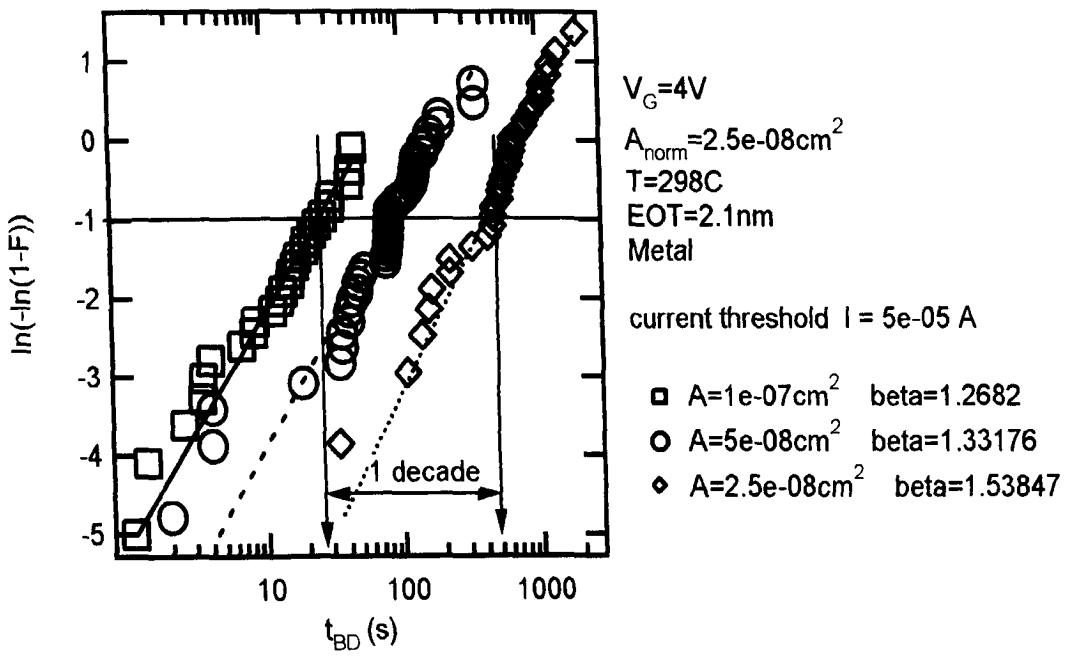


Fig.4.16 Time-to-hard breakdown t_{HBD} distributions normalized to the smallest area. A change of 1 decade is observed.

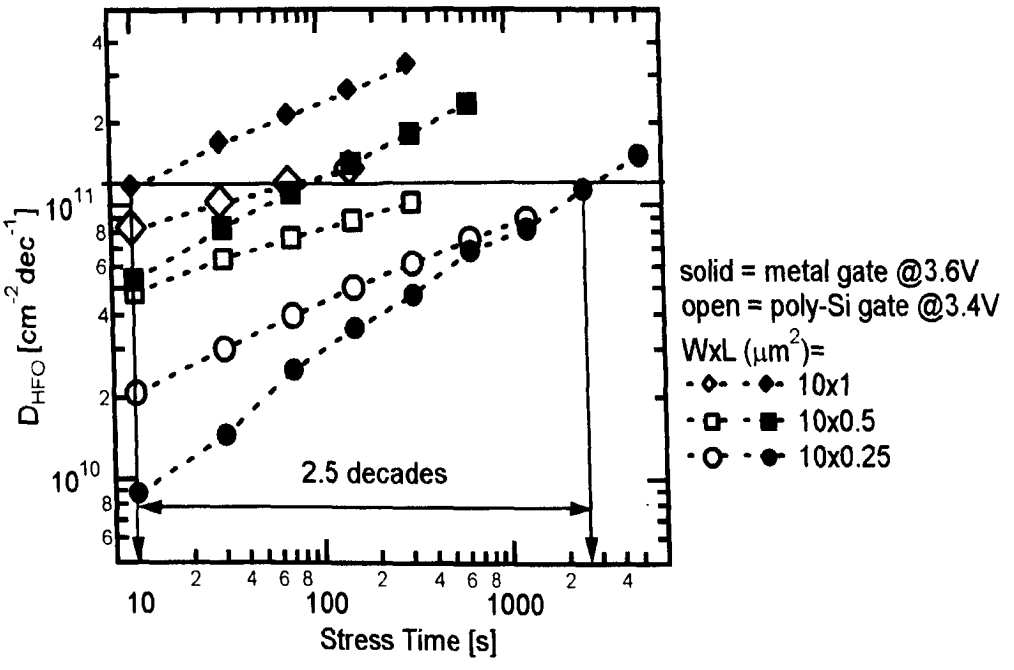


Fig.4.17 D_{HFO} versus stress time extracted from variable frequency charge pumping for different channel lengths. The creation time to reach the same level of D_{HFO} changes by 2.5 decades while the t_{BD} only varies by 1 decade (Fig.4.16)

This channel length dependence of the HfO_2 bulk trap density is presumable a consequence of non-uniform lateral passivation in the channel or caused by charge loss to the source and drain as described in section 4.2.1.2. It is not due to an oxide thickness variation, as I_G current scales well with area. Since the breakdown is expected to occur at a fixed critical trap density **[Sune90]**, **[Degra98]**, a difference in trap density will result in different time-to-breakdown.

For metal-gated devices changing, channel length from 0.25 to 1 μm causes a 2.5 decades change of trap creation, as shown in Fig.4.17. The t_{BD} change observed from Fig.4.16 is, however, only about 1 decade. This discrepancy is explained by examining Fig.4.18.

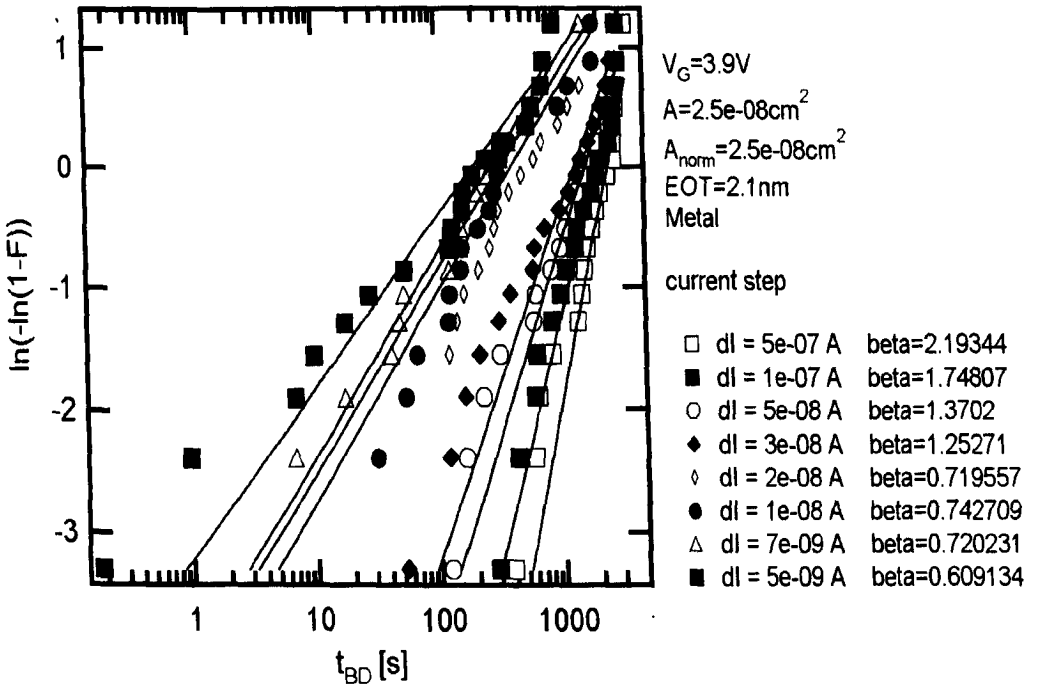


Fig.4.18 Time-to breakdown t_{BD} with different current step triggers for device area= 10x0.25μm^2. The use of a small current step results in a shallow distribution, signature of creation of traps. A higher current step results in a steeper distribution, signature of the wearout of breakdown spots [Kacz04].

When a small current step is applied to determine t_{BD} a shallow weibull distribution is obtained, while a large current step results in a steeper distribution. This is a typical signature **[Kacz04]** of two phases in the breakdown process: 1) creation of traps and 2) wearout of breakdown spots. Only the 1st phase correlates with the D_{HFO} measurement in Fig.4.17, as shown in Fig.4.19.

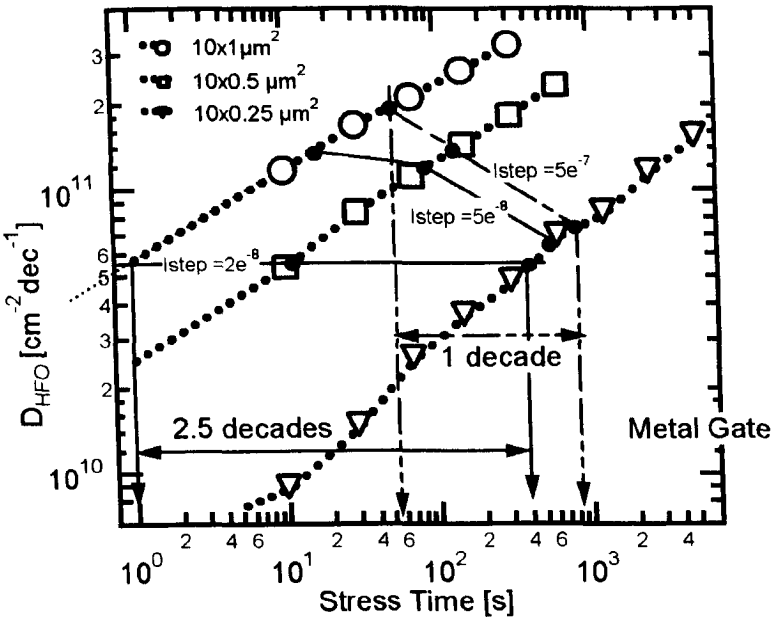


Fig.4.19 D_{HFO} versus stress time extracted from variable frequency charge pumping for different channel lengths. When triggering on a small current step a change of 2.5 decade is observed consistent with Fig.4.17 referring to the creation of traps and when triggering on a large current step a change of 1 decade is observe consistent with Fig.4.16 referring to the wearout phase.

When a small current step is used, the *creation* of traps determines t_{BD} and therefore a fixed trap density at breakdown is indeed observed. Changing the channel length from 0.25 to 1 μm indeed changes the t_{BD} by 2.5 decades. When a large current step is used (In Fig.4.16, the step = $5x10^{-5}$A), the t_{BD} is determined by the *wear out* of conductive paths and does not correlate with the trap density. The change of t_{BD} with channel length is in this case reduced to only 1 decade consistent with the result in Fig.4.16.

4.3.2 Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP)

The Variable T_{charge}-$T_{discharge}$ Charge Pumping (VT^2CP) described in chapter 2 section 2.4.4 is used to investigate and separate the creation of traps in the SiO_2 from the traps in the HfO_2 in an ALD SiO_2/HfO_2 metal gate stack and observe the creation of new traps in both constituent layers by independently controlling the pulse low timing "discharging time" and high level timing "charging time".

The device considered is a thin (EOT = 1.29 nm) Atomic Layer Deposition (ALD) SiO_2 / HfO_2 metal gate stacks. The stack was formed by an O_3-based clean of the substrate surface, which resulted in a chemically grown oxide of about 1nm physical ('IMEC clean'), followed by ALD of 2 nm HfO_2 with 10s of water pulse (H_2O). The devices used consist of n+/ pwell meander-type gate capacitors. The junction and well contacts along the poly-Si stripes are shorted by metal. Because of the junction, a source of carriers is present for inversion mode. The device has 2 poly stripes of 10μm wide and a total area of 10700 μm^2. With these large area structures, the charge pumping current at low frequency can still be resolved.

A gate pulse with fixed t_{charge} (=time at top pulse level) and $t_{discharge}$ (=time at base pulse level) is applied for a base level sweeping from -1.3 to 0.2V. Then $t_{discharge}$ is increased while t_{charge} remains unchanged and a new base level sweep is taken.

The pulse amplitude V_{AMP} is 1.3V and fall and rise times were chosen sufficiently long (tr = tf = 300 ns) to avoid geometric component. As an example, the base level sweeps was done with a fixed t_{charge} = 3μs and $t_{discharge}$ varying from 0.4μs to 5000μs (corresponding to 295 KHz to 200Hz). The results are shown in Fig.4.20. The discharge time was limited to 5000 μs since for longer values the charge pumping current was drowned by the gate leakage current and reliable extraction of D_T was no longer possible.

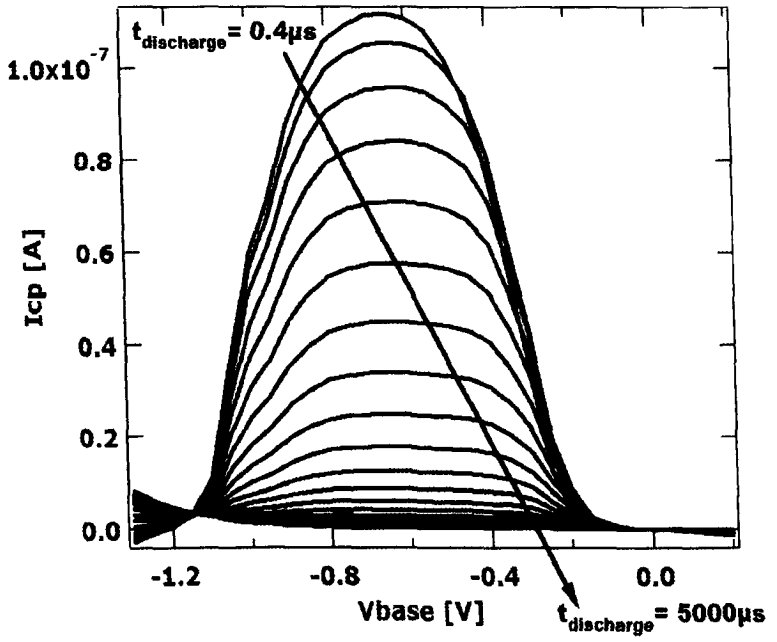


Fig.4.20 Charge Pumping current versus Vbase using VT²CP .Vamp is 1.3V and fall and rise times were chosen sufficiently long (tr = tf = 300 ns) to avoid geometric component. Each base level sweep was done with a fix t_{charge} = 3µs and $t_{discharge}$ varying from 0.4µs to 5000µs.

The trap density was extracted using Eq.2.26 introduced in chapter 2 section 2.4.2 with the Icp current monitored at fixed base level, Vbase = -0.4V.We selected a value as close to zero as possible (but still in charge pumping regime), to avoid leakage at low frequency.

The measurement in Fig.4.20 was repeated for different t_{charge} values and the extracted trap densities are summarized in Fig.4.21. A schematic drawing of the data in Fig.4.21 is shown in Fig.4.22

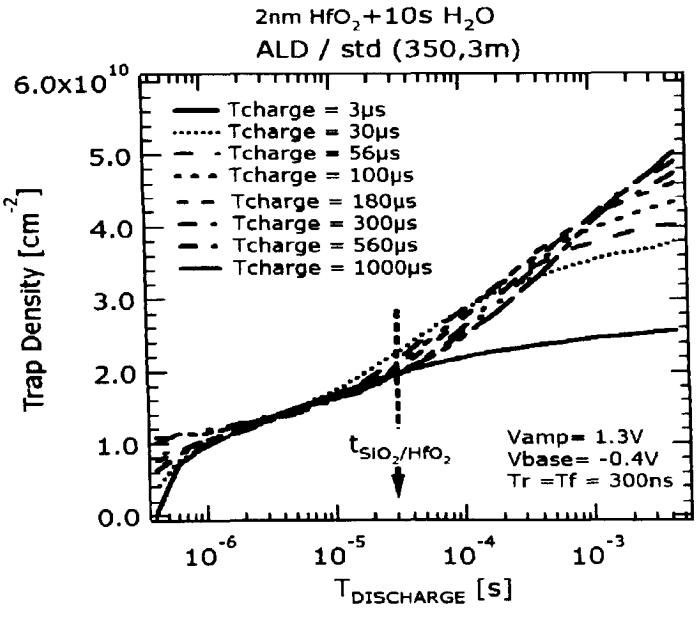


Fig.4.21 Trap density vs. $t_{discharge}$ at different t_{charge} for 2nm HfO_2 with 10s water pulse, using VT^2CP. The vertical dotted arrow indicates the $t_{discharge}$ time ($t_{SiO2/HfO2}$) when the SiO_2/ HfO_2 interface is reached.

In Fig.4.22 four regimes are indicated that can be interpreted as follows:

I. At the chosen base level (-0.4V in Fig.4.20), the CP-curve does not reach its full maximum, causing an apparent reduction of the D_T. The drop at short t_{charge} and $t_{discharge}$ is not fully understood at present. We will focus on the result taken at a t_{charge} sufficiently long that this drop is absent.

II. As $t_{discharge}$ increases more traps in the SiO_2 interface layer can respond to the gate pulse. The slope change in the curve at $t_{SiO2/HfO2}$ (=~30-40 µs in Fig.4.21) indicates the scanning depth has reached the SiO_2/HfO_2 interface.

III. The bulk states in the HfO_2 are scanned until a saturation level sets in at t_{sat}

IV. No additional traps are measured at longer $t_{discharge}$ time because no charged bulk states are left to be discharged. Note that in Fig.4.21 t_{sat} shifts towards longer $t_{discharge}$ as t_{charge} increases.

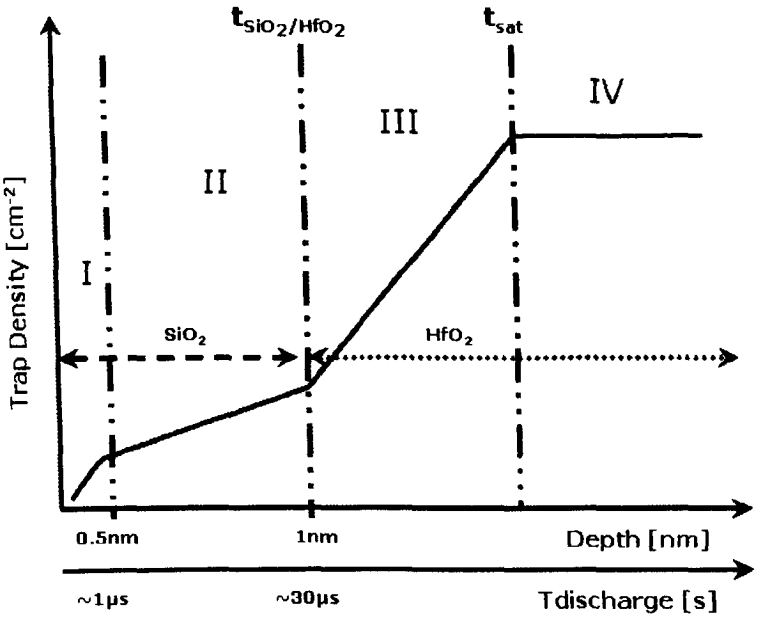


Fig.4.22 Using VT²CP 4 regimes can be interpreted: I) An apparent reduction of D_T; II) traps in the SiO_2 interface layer are scanned; III) Bulk states in the HfO_2 are scanned; IV) Saturation is reached as no charged bulk states are left to be discharged.

There exists disagreement in literature on how deep one can probe in the stack. We interpret the transition between region II and III as the transition between SiO_2 and HfO_2 ($t_{SiO2/HfO2}$ in Fig.4.22), but some groups **[YoungC06]** claim that region III is still in the interface layer. We have performed an experiment that correlates the SiO_2 thickness with $t_{SiO2/HfO2}$. This correlation provides *experimental* evidence for the interpretation that the traps measured in region III are indeed in the HfO_2, the details of this experiment is given below.

We used small wafer level variations to find correlations between parameters using TDDB and CP. CP is used to carry out the experiment shown in Fig 4.21 at t_{charge} = 560us only across the entire wafer to calculate the variation of the intersection $t_{SiO2/HfO2}$ between low and high frequency as shown in Fig 4.22. Fig 4.23 shows the intersection time $t_{SiO2/HfO2}$ across the wafer.

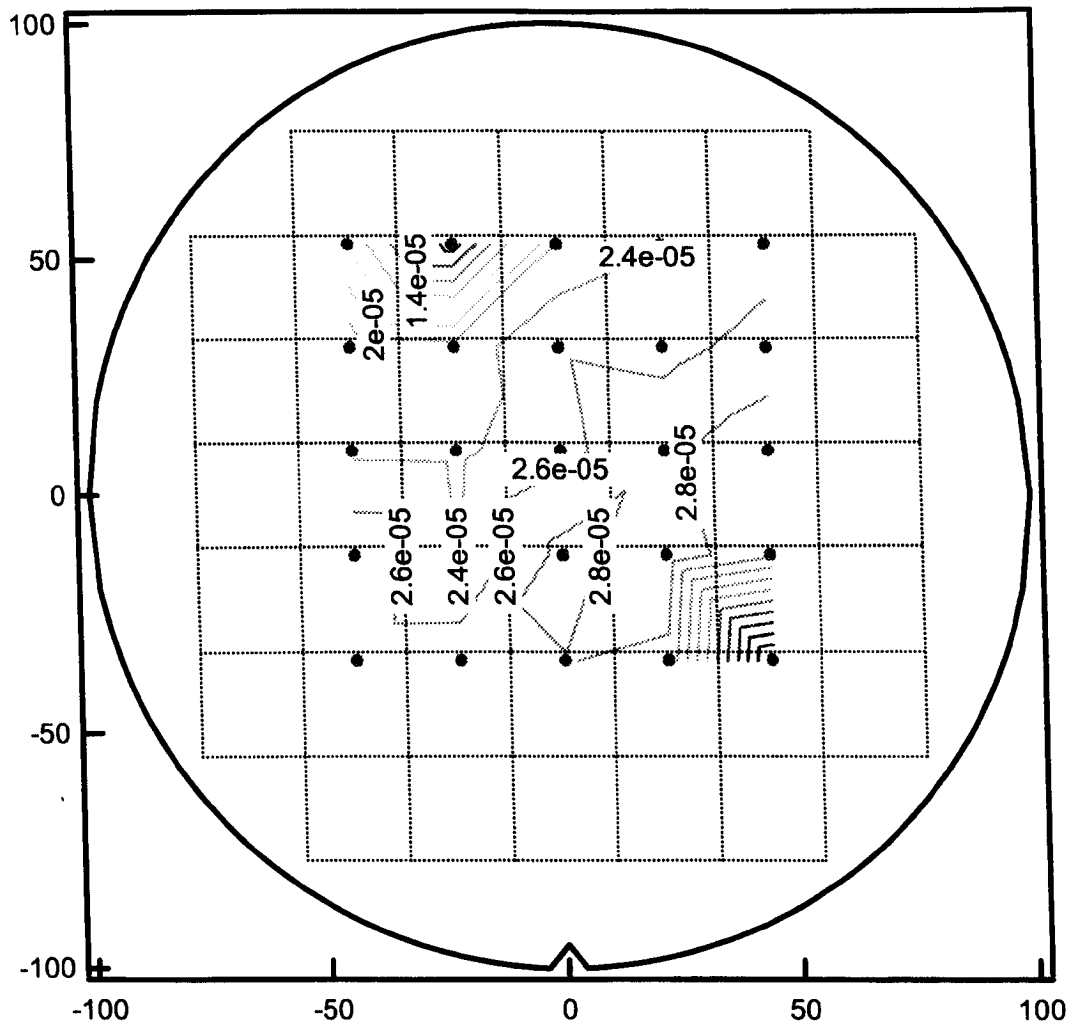


Fig.4.23 Wafer level variation showing the $t_{SiO2/HfO2}$ time at the intersection between low and high frequency measured by VT^2CP

It is known that the thickness of the interface layer can be measured by the gate leakage current at high positive voltage. TDDB is used here to measure the small variation of the gate leakage current on the entire wafer at high positive voltage (3.3 V). The result in given in Fig 4.24

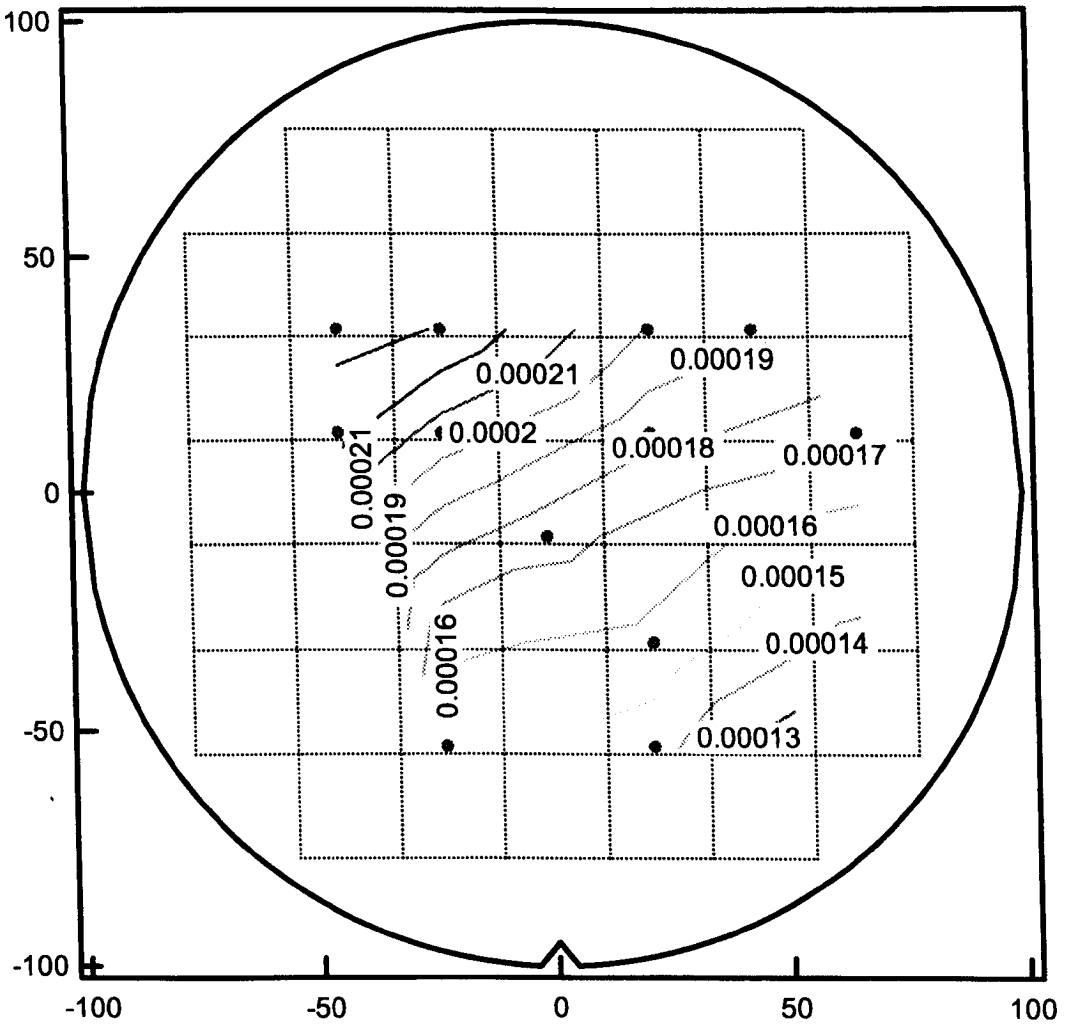


Fig.4.24 Wafer level variation showing the leakage current at 3.3V measured by TDDB

Fig. 4.25 shows that there is a good correlation between the time $t_{SiO2/HfO2}$ from Fig 4.23 and the gate leakage at high voltage (3.3V) from Fig 4.24.

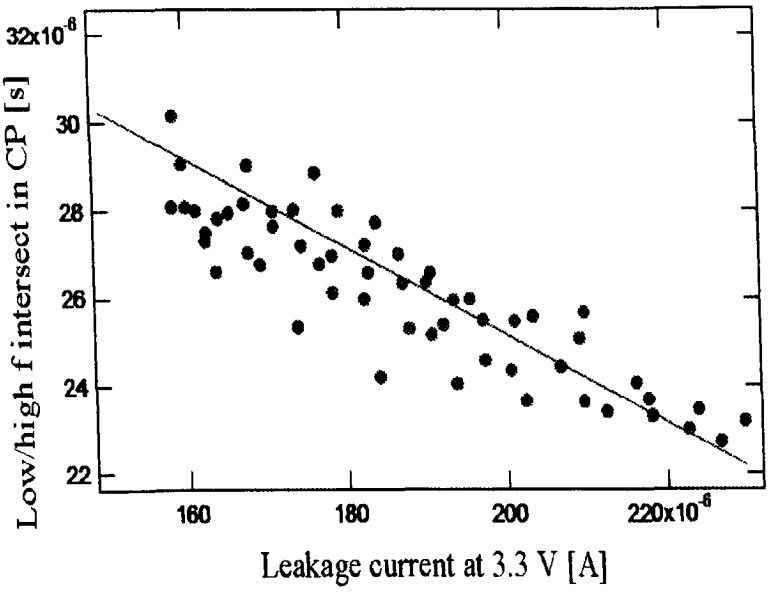


Fig.4.25 $t_{SiO2/HfO2}$ intersection time versus leakage current . A good correlation between both data can be well fitted.

Furthermore, if we take the SiO_2 layer to be ~1 nm (from electrical measurement, TEM, XPS), we can show that at $t_{discharge}=10^{-6}$ s, the scanning depth is ~0.5 nm (Fig.4.22).

Fig.4.26 shows the complement of Fig.4.21: the trap density is plotted versus the charge time for different discharge time. When the discharge time is very short (<30μs) the trap density remain flat, indicating that we are indeed scanning only the SiO_2 interfacial layer. When the discharge time is higher than 30μs, an increase in the trap density is seen, indicating that the traps in the HfO_2 are pumped. For larger charge time a reduction of trap density is observed. Note that the measured trap density reaches a maximum as a function of t_{charge}. This is possibly due to charge redistributed deep in HfO_2 that can no longer be discharged in the applied discharge time.

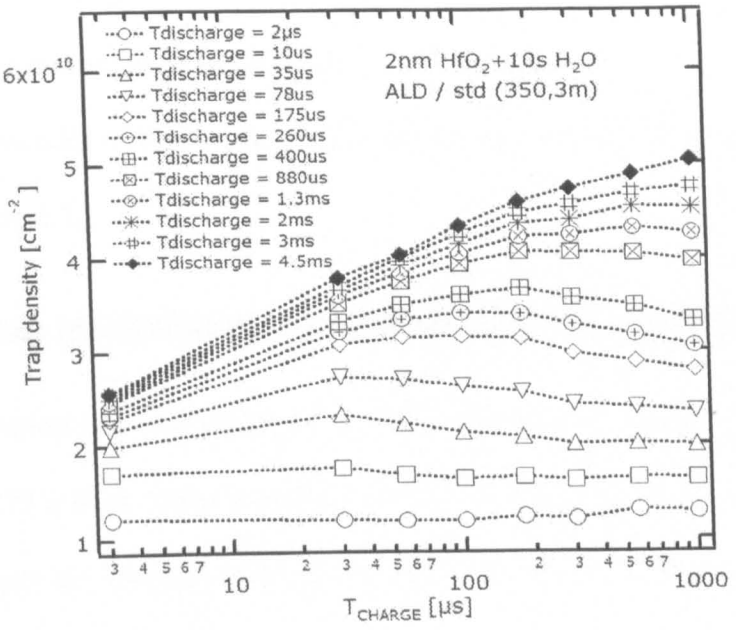


Fig.4.26 Trap density vs. t_{charge} at different $t_{discharge}$ for 2nm HfO_2 with 10s water pulse, using VT^2CP.

Note also for ~30 µs *charge* time, we can fill traps inside the HfO_2 as shown in the schematic drawing in Fig. 4.27a. A 30 µs *dis*charge time ($t_{SiO2/HfO2}$ in Fig.4.22) is, however, insufficient to remove all of this charge, since only the SiO_2 layer is discharged (Fig 4.27b)

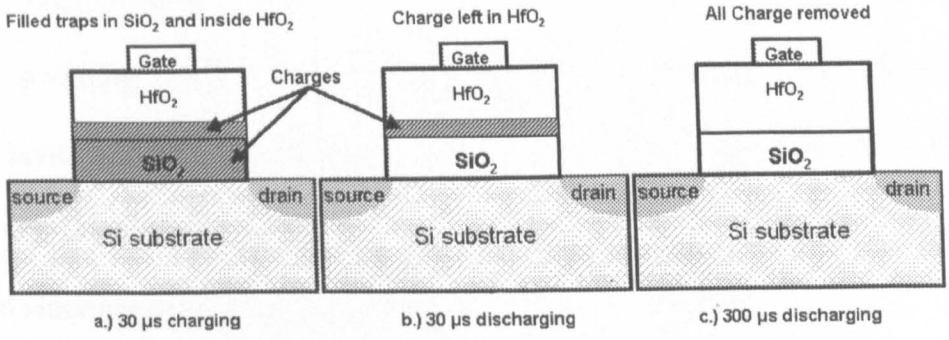


Fig.4.27 schematic drawing on how traps can be filled inside the HfO_2. a) 30 µs charge time, can fill traps inside the HfO_2; b) 30 µs discharge time is insufficient to remove all of this charge, only the SiO_2 layer is discharged; c) 300 µs discharge time, all trapped HfO_2 charge is also removed.

Only if the discharge time is increased up to 300µs (t_{sat} in Fig. 4.22), all trapped HfO_2 charge is also removed (Fig. 4.27c). In other words, the charging mechanism is

faster then the discharging mechanism. This can be due to a reduced energy level for a charged state as compared to an uncharged.

In the remainder of this work, we fix the charge time at 560 μs in order to avoid region IV as in Fig.4.22.

4.3.3 Trap generation in SiO_2/HfO_2 stacks measured by VT^2CP

In this section, we will apply the VT^2CP technique to analyze the generation of traps in a SiO_2/HfO_2 stack under positive Constant Voltage Stress (CVS). In particular, we will investigate the creation of traps in the SiO_2 and HfO_2 separately. The device used is the same as in section 4.3.2.

The trap density in the interfacial SiO_2, D_{SiO2}, and the HfO_2 trap density, D_{HfO2}, can be determined as shown in Fig.4.28.

Two linear fits were taken, one in region II (interval t1 in Fig.4.28) and one in region III (interval t2 in Fig.4.28). The slope in region II (III) gives D_{SiO2} (D_{HfO2}) expressed per area unit and sensed per decade of $t_{discharge}$. Typical values on a fresh device are $D_{SiO2}=1.5x10^{10}$ $cm^{-2}dec^{-1}$ and $D_{HfO2}=1.7$ $x10^{10}cm^{-2}dec^{-1}$. Note that with the estimated scanning depth as indicated in Fig.4.22, the volume trap density in the interface layer is $\sim 10^{17}cm^{-3}$, in agreement with published data for thicker SiO_2 and SiON layers **[Degra04A],[Ielmi02]**. The volume trap density in the HfO_2 cannot be calculated since the exact scanning depth in the HfO_2 is not known.

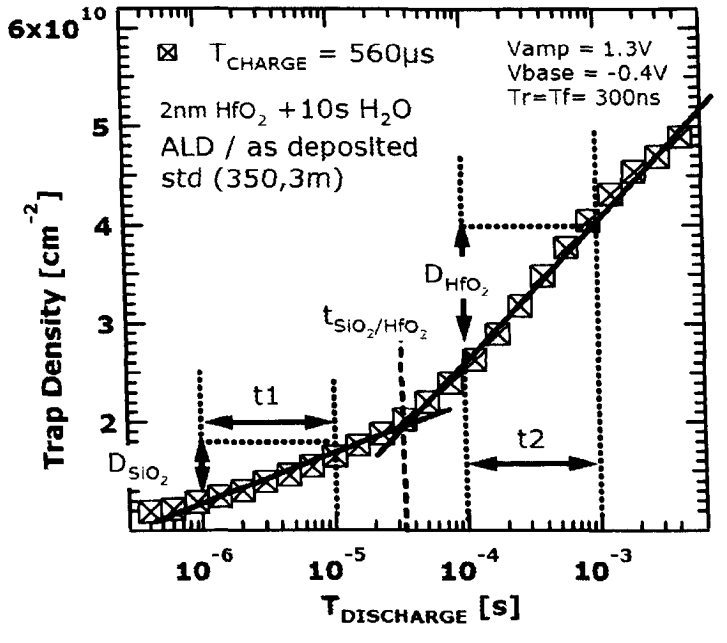


Fig.4.28 The trap density in the interfacial SiO_2, D_{SiO2}, and the HfO_2 trap density, D_{HfO2} determination.

In order to measure the trap generation during electrical stress, a Constant Voltage Stress (CVS) is interrupted at regular time intervals and VT^2CP is applied. Note that only one device was used for each CVS combined with VT^2CP measurement.

Figs.4.29, 4.30 and 4.31 show the time evolution of the trap density versus discharge time at different stress voltages. An increase of the total trap density in both region II and III is observed as summarized in Fig.4.32

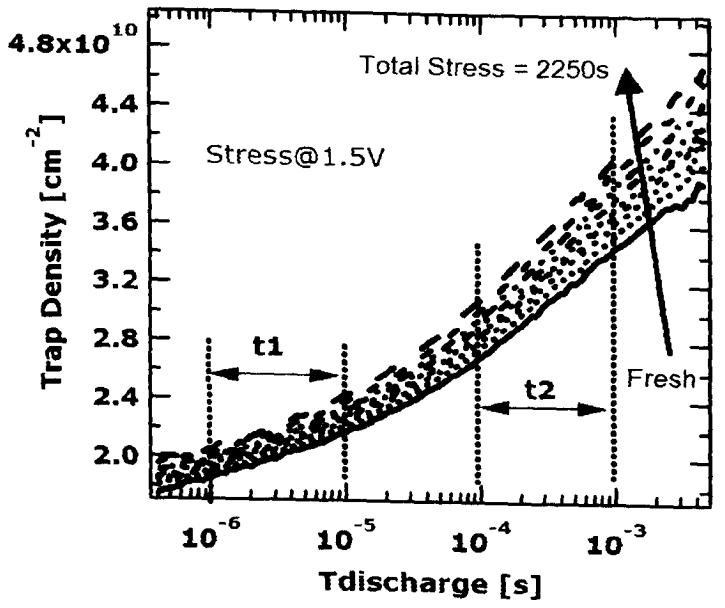


Fig.4.29 Time evolution of the trap density versus discharge time at constant voltage stress = 1.5V, t1 represent the interval fitted in region II for D_{SiO2} and t2 represent the interval fitted in region III for D_{HfO2}.

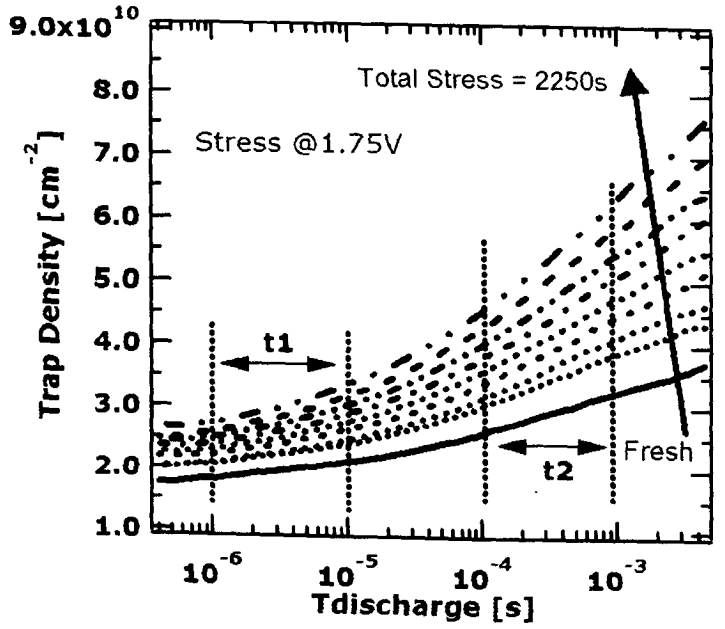


Fig.4.30 Time evolution of the trap density versus discharge time at constant voltage stress = 1.75V, t1 represent the interval fitted in region II for D_{SiO2} and t2 represent the interval fitted in region III for D_{HfO2}.

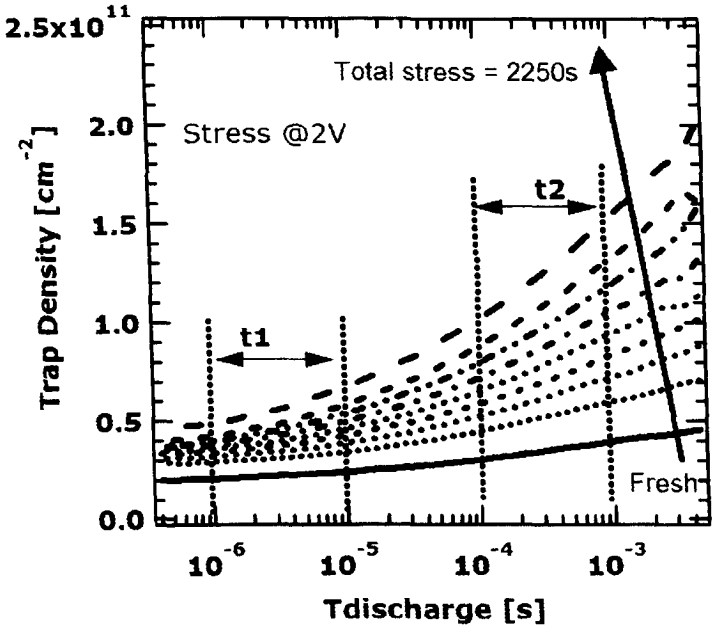


Fig.4.31 Time evolution of the trap density versus discharge time at constant voltage stress = 2 V, t1 represent the interval fitted in region II for D_{SiO2} and t2 represent the interval fitted in region III for D_{HfO2}.

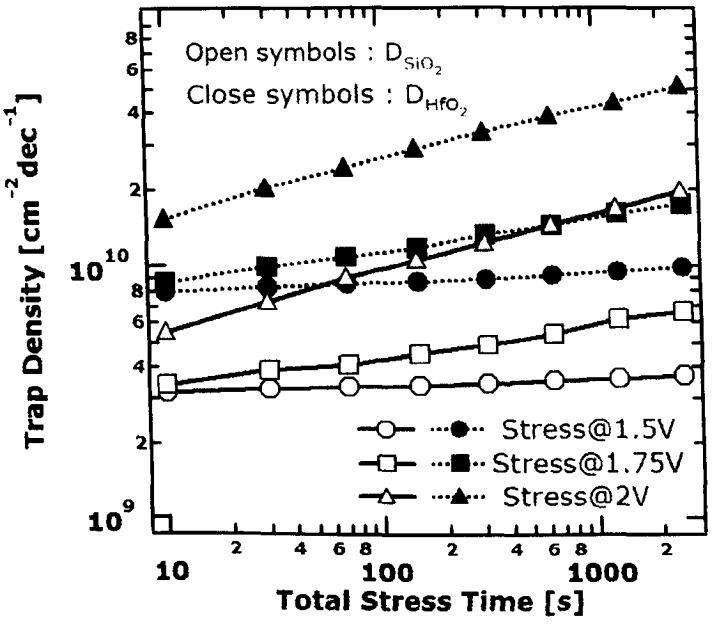


Fig.4.32 Total trap density sensed per decade frequency during stress in both region II (D_{SiO2}) and III (D_{HfO2}) at 1.5V, 1.75Vand 2V.

After subtracting the trap density of the fresh device, the data of both ΔD_{SiO2} and ΔD_{HfO2} can be well fitted by a power law with an exponent of ~0.32 and ~0.34 respectively and independently of stress voltage as shown in Fig.4.33.

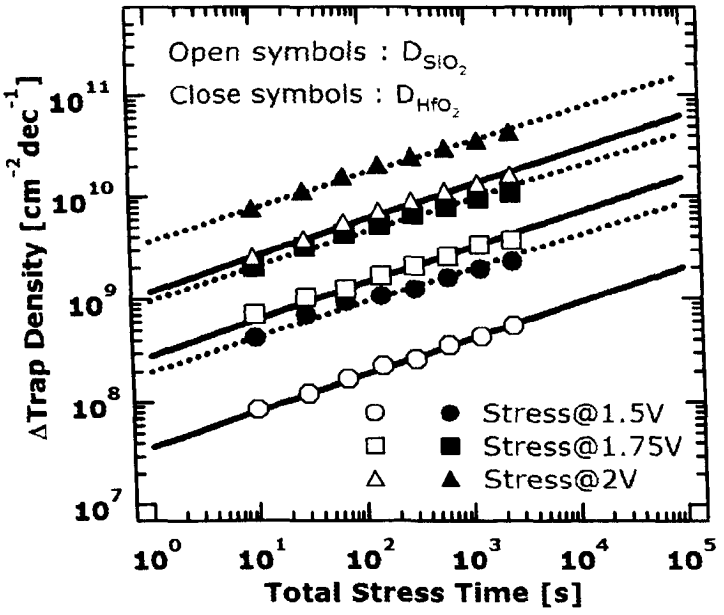


Fig.4.33 ΔT_D vs. Total stress time sensed per decade frequency during stress in both region II (D_{SiO2}) and III (D_{HfO2}) at 1.5V, 1.75V, and 2 V. data can be well fitted by a power law with an exponent of ~0.34(D_{HfO2}) and ~0.32(D_{HfO2}), independently of stress voltage.

Fig.4.34 shows the voltage acceleration of the trap generation process for ΔD_{SiO2} and ΔD_{HfO2} (derived from Fig.4.33).

A power law voltage acceleration is used and we find voltage acceleration exponent of -34 and -30 respectively. TDDB-measurements, also plotted in Fig.4.34, show a voltage acceleration exponent of -27, nearly identical to the D_{HfO2} acceleration exponent. This is consistent with the model that dielectric breakdown occurs when the trap density in the HfO_2 reaches a critical value **[Degra03]**, **[Degra05]**.

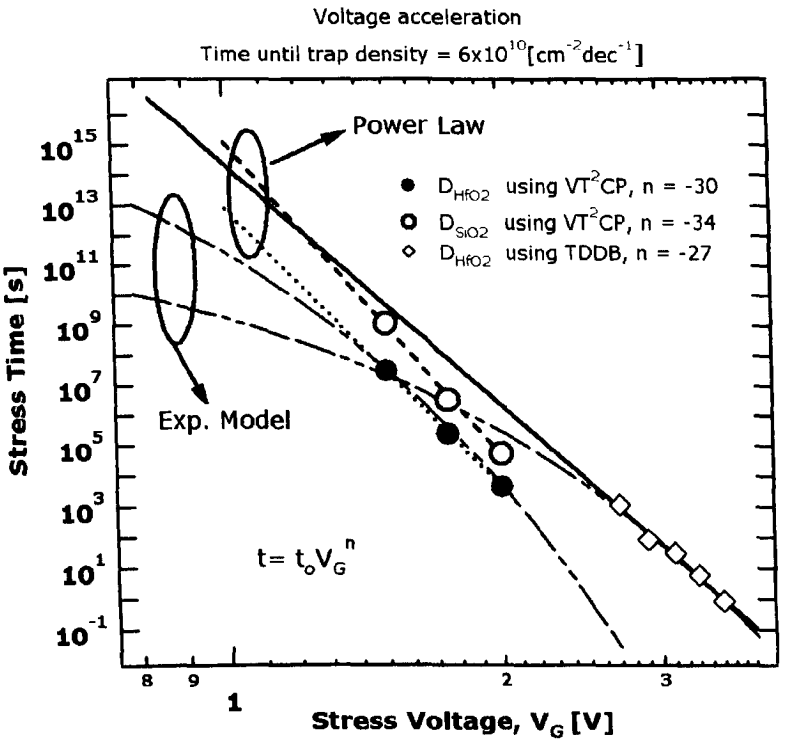


Fig.4.34 Voltage acceleration of the trap generation process at a fix trap density for D_{SiO2} and D_{HfO2} using VT^2CP and TDDB. A power law voltage acceleration of -34 for D_{SiO2} and -30 for D_{HfO2} is found when VT^2CP is used, which is nearly identical to the D_{HfO2} acceleration exponent found with TDDB-measurements.

Note that the VT^2CP can accurately detect degradation down to a much lower voltage than the dielectric breakdown measurement range. Furthermore, only one stress experiment combined with VT^2CP is sufficient to determine the degradation at a given voltage, while a TDDB test requires many measurements in order to construct an accurate distribution of failure times.

Note also that the agreement between high voltage TDDB and low voltage VT^2CP data is only obtained if a *power law* voltage acceleration is assumed and not if an exponential model is taken (Fig.4.34). The data therefore confirm that power law **[Degra05]**, **[YoungC06]** describes the voltage acceleration of dielectric degradation more consistently than exponential models.

4.3.4 Trap generation in different thickness of HfO_2 and different thickness and composition of HfSiO characterized by using VT^2CP

VT^2CP introduced in section 4.3.2 is used on different thickness of HfO_2 with several post deposition anneal (PDA) and different composition of Hf-Silicate combined with different PDA to investigate trap generation. TDDB is used to evaluate lifetime, leakage current generation and breakdown mechanism and the impact of processing on wafer level deposition uniformity.

The split table of the wafers studied is represented in Fig 4.35.

Samples	Interface	High-k	PDA	MG Degas	EOT (nm)	
1	IMEC	1.25 nm HfO_2	none	std (350, 3m)	1.15	As Deposited
2	IMEC	2 nm HfO_2	none	std (350, 3m)	1.24	
3	IMEC	3 nm HfO_2	none	std (350, 3m)	1.45	
4	IMEC	2 nm HfO_2 + 10s H_2O	none	std (350, 3m)	1.29	
5	IMEC	2 nm HfO_2 + 10s H_2O	none	none	1.67	
6	IMEC	2 nm HfO_2	800C N2 in polygon	std (350, 3m)	1.34	RTA
7	IMEC	2 nm 80% HfSiO	AB23-DPN	std (350, 3m)	1.51	Plasma Nitridation
8	IMEC	3 nm 80% HfSiO	AB23-DPN	std (350, 3m)	1.46	
9	IMEC	4 nm 80% HfSiO	AB23-DPN	std (350, 3m)	1.57	
10	IMEC	3 nm 50% HfSiO	AB23-DPN	std (350, 3m)	1.71	
11	IMEC	4 nm 50% HfSiO	AB23-DPN	std (350, 3m)	1.94	
12	IMEC	3 nm 50% HfSiO	800C NH3 in polygon	std (350, 3m)	1.56	Thermal Nitridation
13	IMEC	4 nm 50% HfSiO	800C NH3 in polygon	std (350, 3m)	1.79	

Fig.4.35 Details split of the sample used for the trap generation using VT^2CPcharacterization

The size and structure of devices used is the same as that in section 4.3.2. As mentioned in section 4.3.2, to avoid the saturation in region IV (Fig.4.22), a long T_{charge} is used (i.e 560μs) to investigate the trap generation. Figs. 4.36 to 4.38 represent the trap density obtained when VT^2CP is used for the samples listed in Fig 4.35.

Note that the data in Figs 4.36, 4.37, 4.38 represent the trap density after subtracting the initial trap density, to have a consistent comparison.

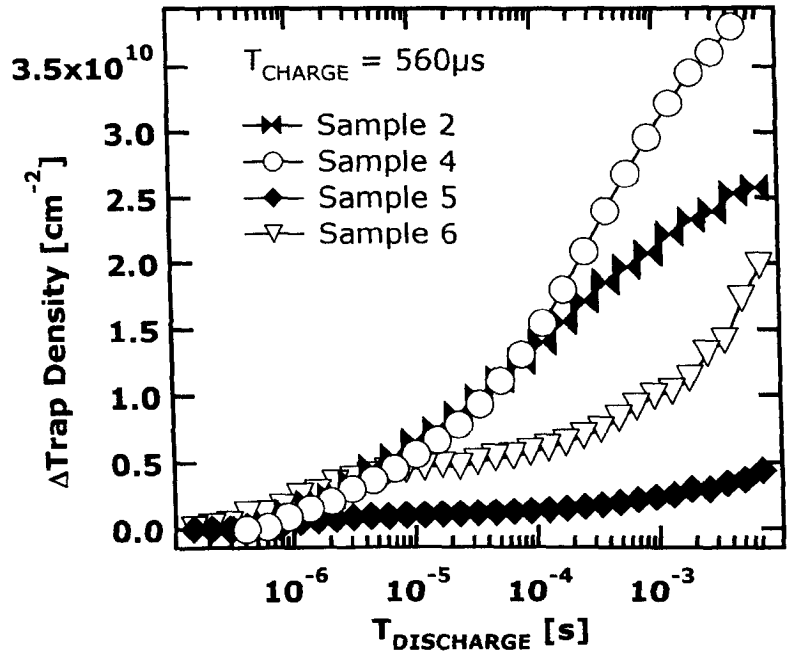


Fig.4.36 Comparison of the trap density vs. discharge time for 2 nm HfO_2, as deposited with/without 10s water pulse or Rapid Thermal Anneal (RTA) using VT^2CP

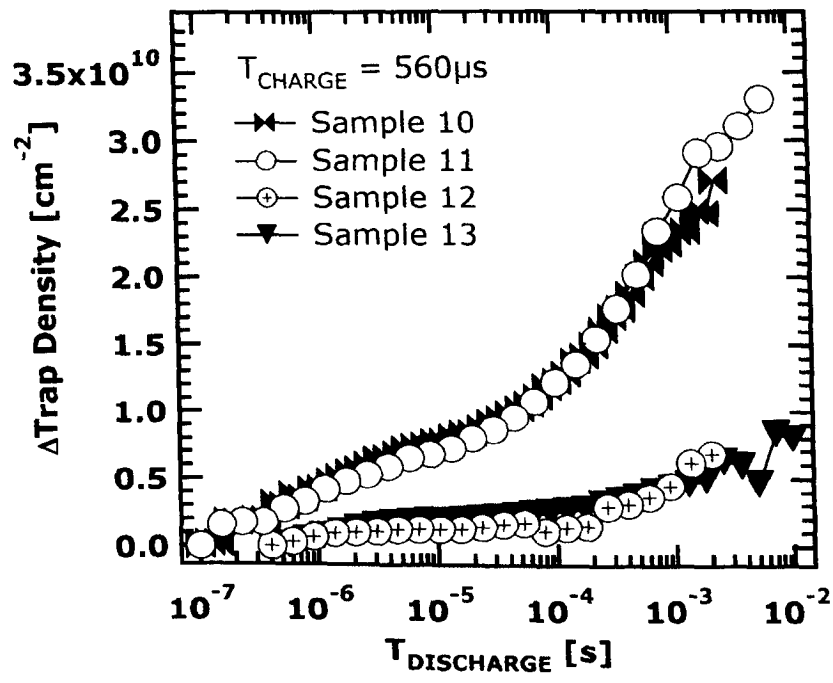


Fig.4.37 Comparison of trap density vs. discharge time for 3,4 nm 50% HfSiO with thermal nitridation and plasma nitridation using VT^2CP.

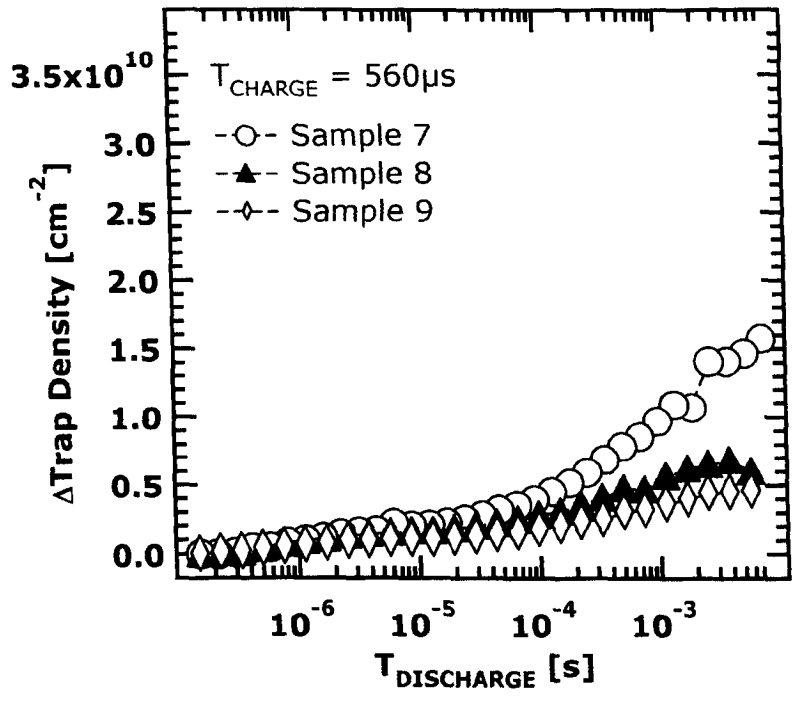


Fig.4.38 Comparison of the trap density vs. discharge time for 2, 3, 4 nm 80% HfSiO with plasma nitridation using VT²CP.

The trap generation at the interface (D_{it}) is taken at $T_{discharge}$ = 1μs and the trap generation in SiO_2 (D_{SiO2}) and Hf (D_{Hf}) is determined as shown previously in Fig 4.28 of section 3.2, and summarized in Figs 4.39 and 4.40.

Fig 4.39 shows that D_{it} generation is similar for the following samples: 3, 4 nm with 80% Hf (Plasma Nitridation), 50% Hf (Thermal Nitridation), and the 2 nm HfO_2 with Rapid Thermal Anneal (RTA). Relatively high D_{it} was observed for the samples of 3,4 nm 50% Hf and 2 nm 80% Hf (Plasma Nitridation)and the 3 nm HfO_2. Both thicker (3nm) and thinner (1.25nm) HfO_2 show higher Dit than the 2nm HfO_2.

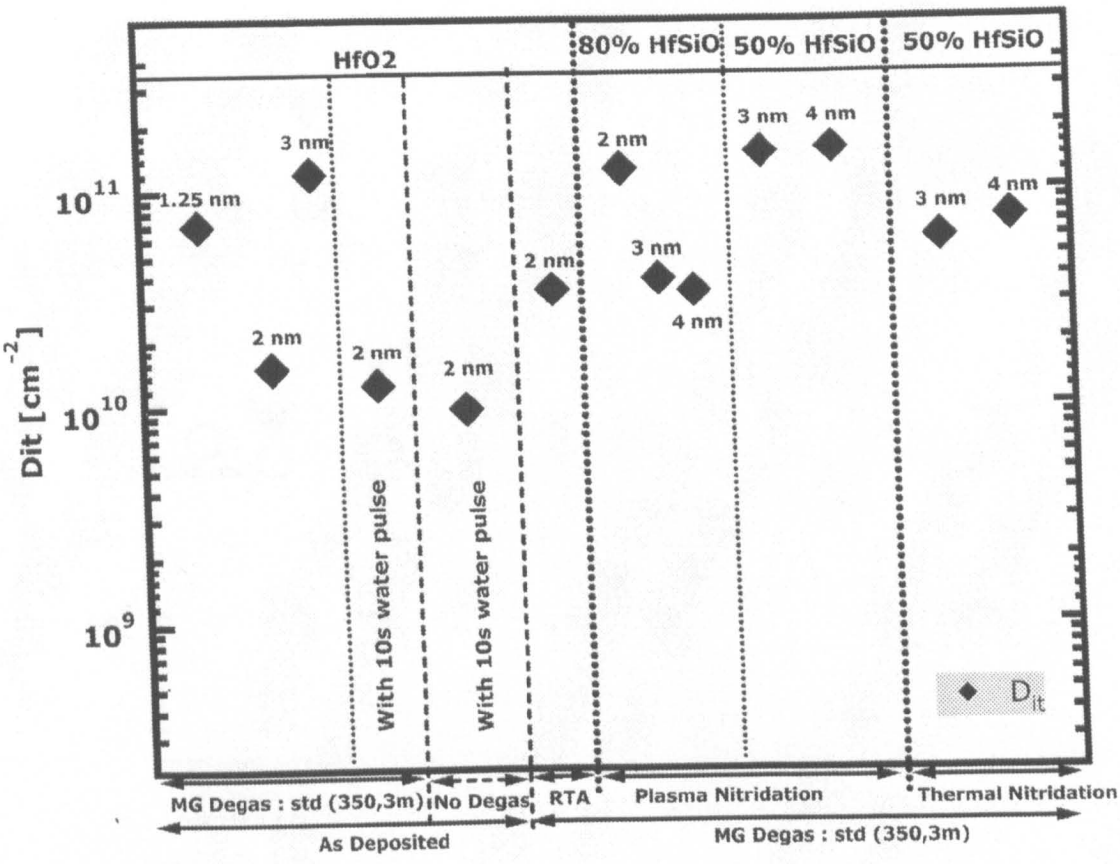


Fig.4.39 Interface trap density (D_{it}) taken at $t_{discharge}$ = 1μs using VT^2CP.

Fig.4.40 shows the traps generation in SiO_2 (D_{SiO2}) and in Hf (D_{Hf}). Similar D_{SiO2} and D_{Hf} was observed for 3,4 nm 80% HfSiO (plasma nitridation) and 50% HfSiO (thermal nitridation). Low D_{SiO2} and D_{Hf} was achieved for 2 nm HfO_2 with 10s water pulse and without MG Degas. The D_{SiO2} and D_{Hf} for 2 nm HfO_2 (RTA) is close to that of the 2 nm 80% Hf (Plasma nitridation). The data show that when using plasma nitridation the % Hf content has to be increased to have low D_{SiO2} and D_{Hf}.

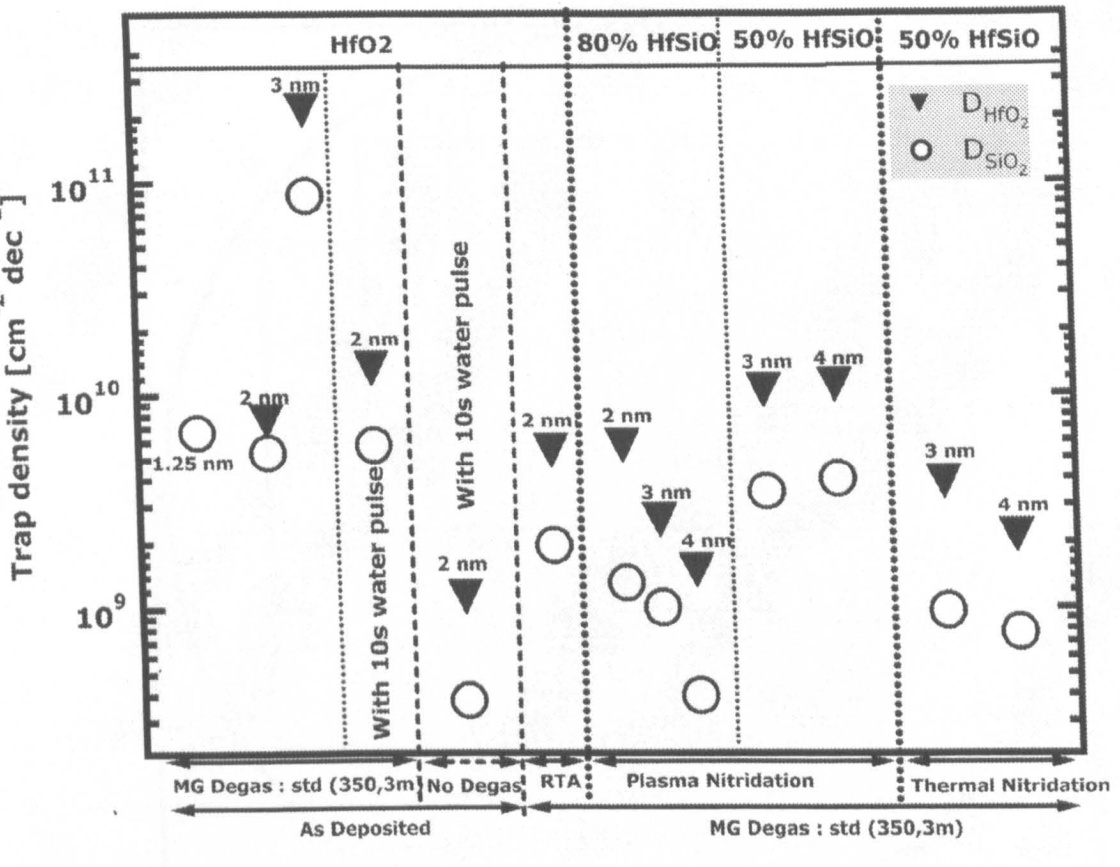


Fig.4.40 Trap density sensed per decade frequency on fresh device both in region II (D_{SiO2}) "open circle" and III (D_{HfO2}) "close triangle" calculated as shown in Fig.4.28 for samples presented in Fig.4.35.

4.3.5 Trap Generation evaluated by TDDB and waver level uniformity

4.3.5.1 Evaluation of 2nm HfO_2

In this section we will evaluate the trap generation using TDDB and wafer level uniformity. The device used consists of n^+/pwell and p^+/nwell square gate capacitors not overlapping on field. The capacitor consists of a 7 x 7 μm^2 poly square on active

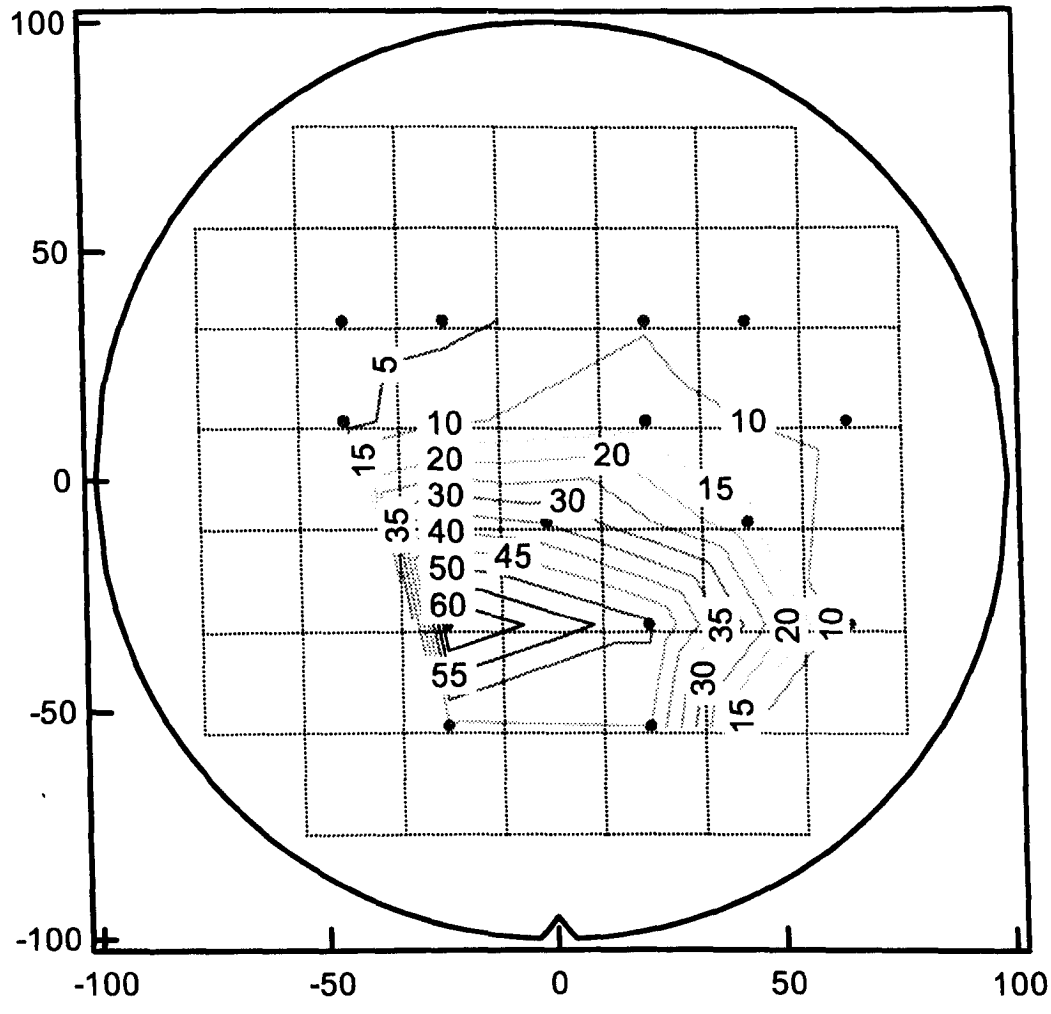


Fig.4.41 Wafer level uniformity for sample 2 in Fig.4.35(40 cy HfO_2 with 0.3 s H_2O pulse). The figure shows the time to breakdown in seconds across the wafer by applying a Constant Voltage stress.

Figs 4.41 & 4.42 shows the time to breakdown obtain by applying CVS on the entire wafer for sample 2 (short water pulse) and 4 (long water pulse).

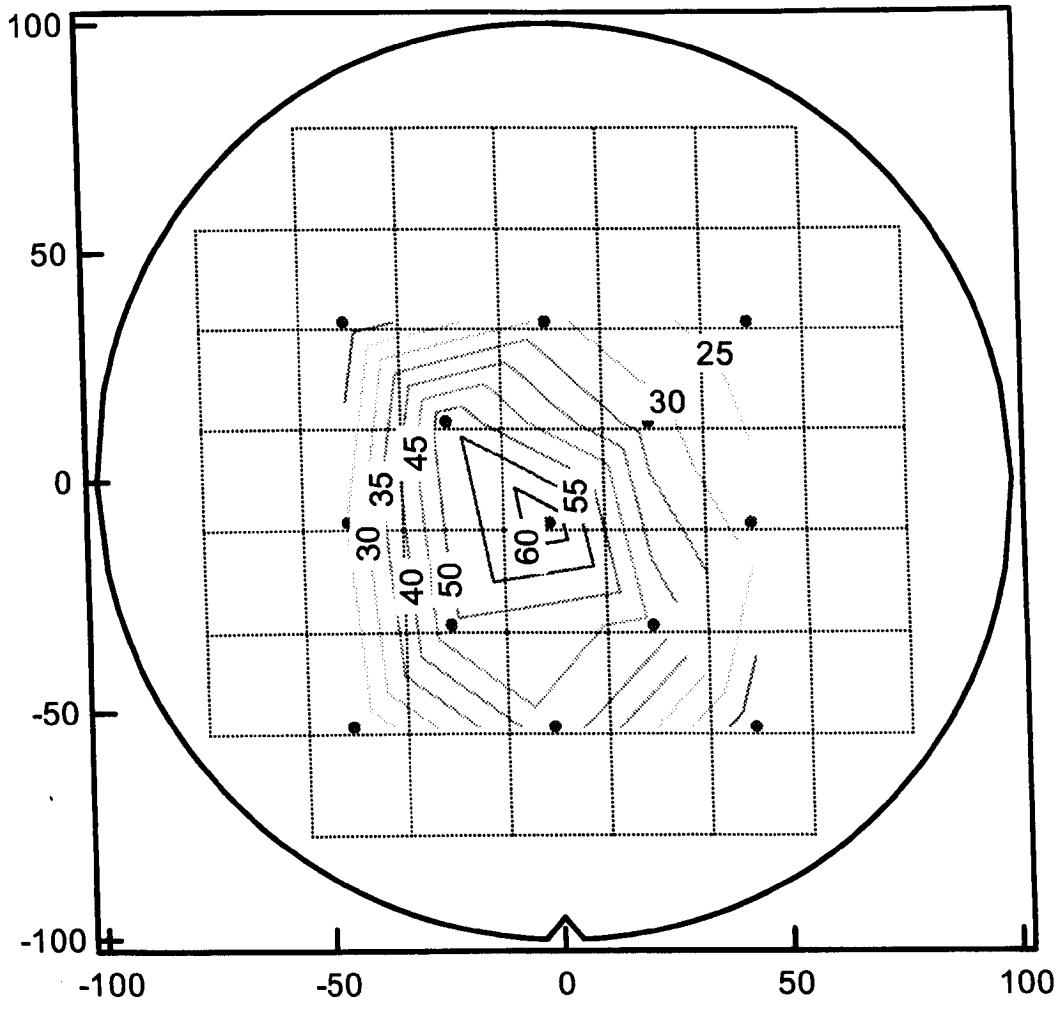


Fig.4.42 Wafer level uniformity for sample 4 in Fig.4.35(40 cy HfO_2 with 10 s H_2O pulse).The figure shows the time to breakdown across the wafer by applying a Constant Voltage stress.

It is shown that the time to breakdown is more uniform across the wafer with the effect of long water pulse by having a more uniform deposition over the wafer (Fig.4.42), which induces a higher trap density (Fig.4.43).

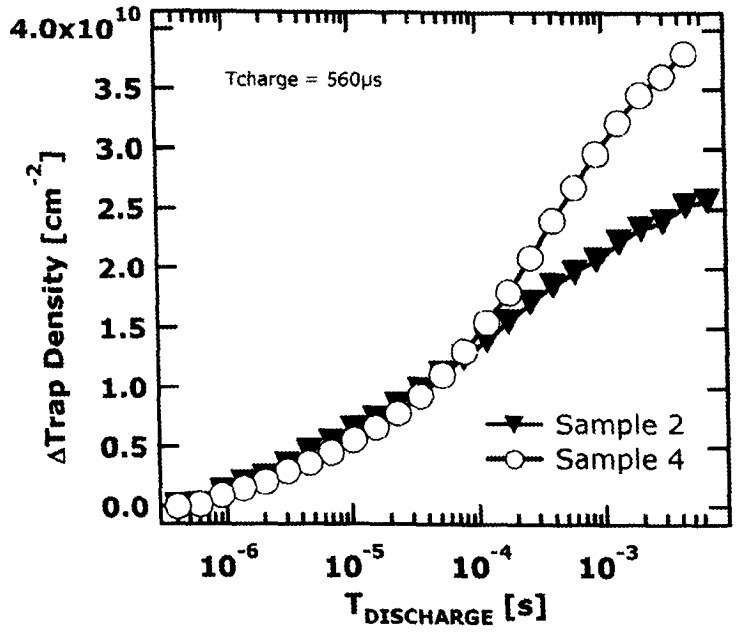


Fig.4.43 Comparison of the trap density vs. discharge time for sample 2 (40 cy HfO_2 with 0.3 s H_2O pulse) and sample 4 (40 cy HfO_2 with 10 s H_2O pulse) using VT^2CP.

Fig.4.44 shows that the lifetime is similar for these two wafers.

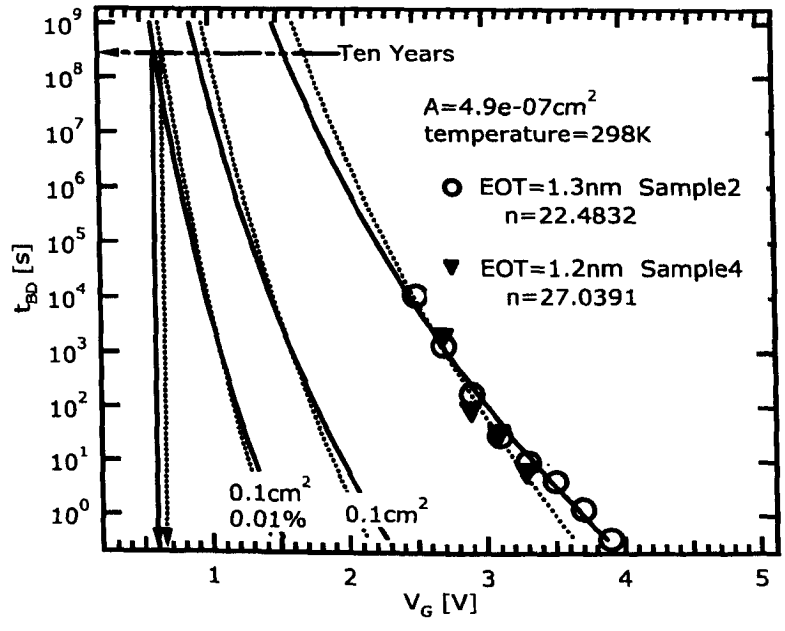


Fig.4.44 Low voltage extrapolation of t_{BD} for wafer 2 and 4 results in sufficient lifetime. No much effect of long water pulse on the lifetime. A power law extrapolation was used with exponent ~22 (sampler 2) and ~27 (sample 4). Maximum operating voltage at 10 years is ~0.6V for both wafers.

4.3.5.2 Evaluation of 2nm 50% HfSiON with Plasma Nitridation

The device considered is a thin (EOT = 1.3 nm) Atomic Layer Deposition (ALD) HfSiO metal gate stacks. The stack was formed by an O_3-based clean of the substrate surface, which resulted in a chemically grown oxide of about 1nm ('IMEC clean'), followed by ALD of 2nm 50% HfSiO and plasma nitridation. The device used consists of n^+/pwell and p^+/nwell square gate capacitors not overlapping on field. The contacts are to poly on active. The capacitor consists of a 7 x 7 μm^2 poly square on active

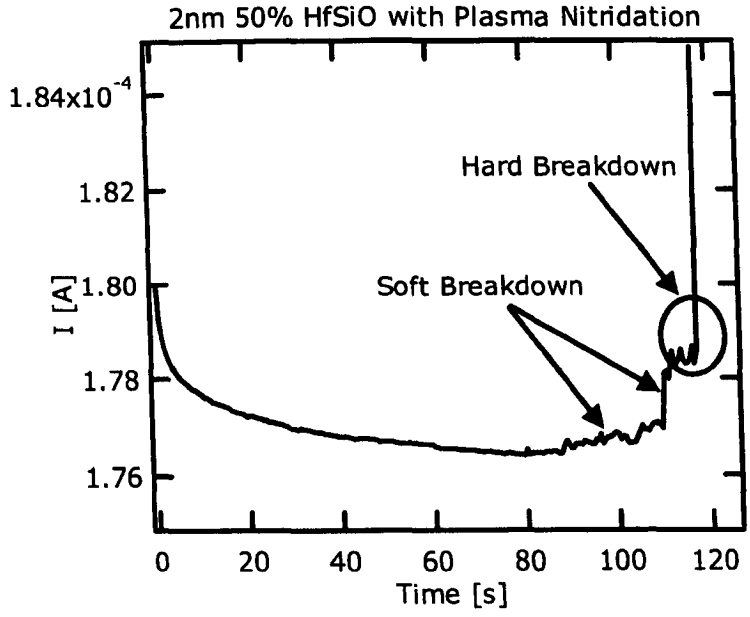


Fig.4.45 Gate current versus stress time for 2nm 50% HfSiO with plasma nitridation showing a soft breakdown followed by hard breakdown.

Fig 4.45 shows the gate current versus stress time until a soft breakdown occurs followed by hard breakdown. There is no significant current increase due to 2-trap paths **[Degra05]** between soft and hard breakdown, the slope indicate that 3-trap path triggers soft and hard breakdown (Fig.4.46).

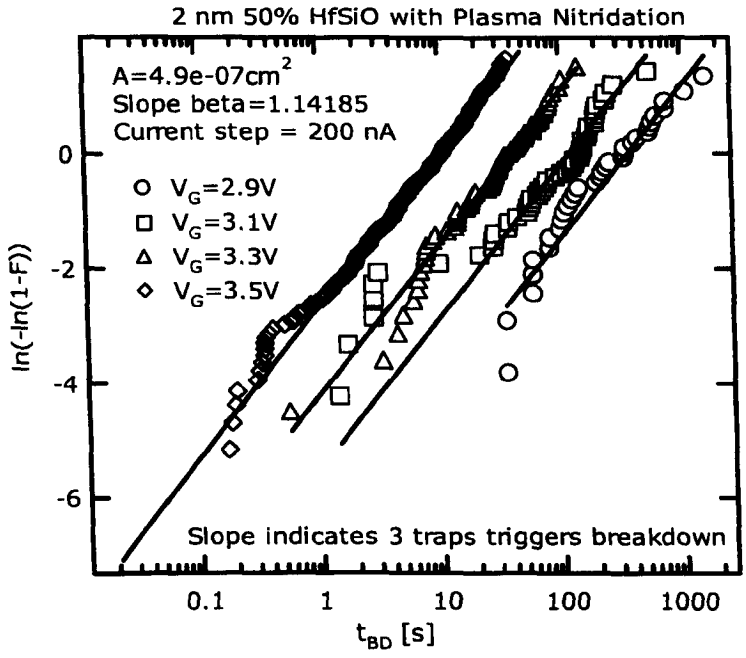


Fig.4.46 Time to breakdown t_{BD} distributions measured on 2nm 50% HfSiO with plasma nitridation for different stress voltage. Exponent beta β~1.15 indicate that 3trap triggers soft and hard breakdown.

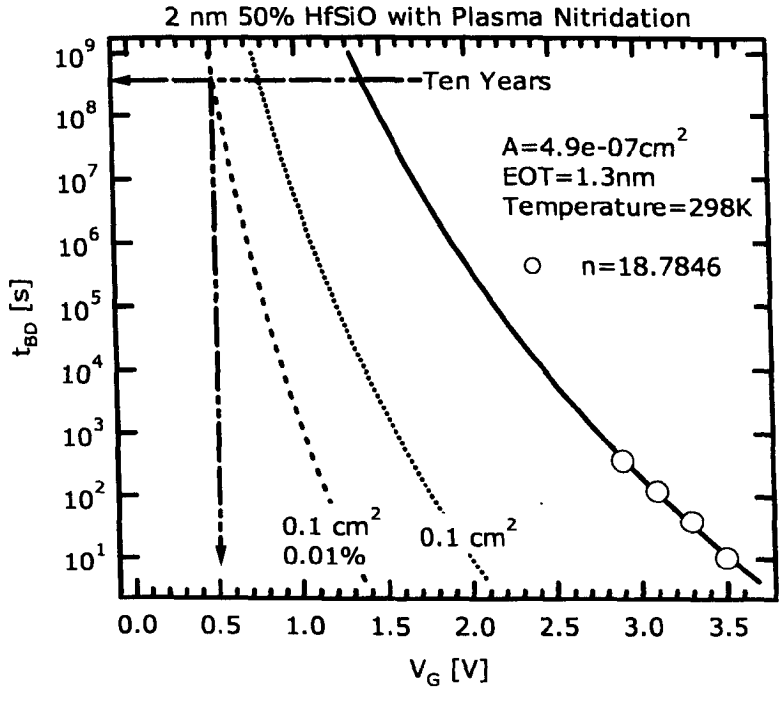


Fig.4.47 Low voltage extrapolation of t_{BD} for 2nm 50% HfSiO with plasma nitridation results in sufficient lifetime. A power law extrapolation was used with exponent ~18. Maximum operating voltage at 10 years is ~0.5V

Fig.4.47 shows the lifetime extrapolation for 2nm 50% HfSiO, maximum operating voltage at 10 years is 0.7V

4.3.5.3 Evaluation of 3nm 80% HfSiON with Plasma Nitridation

The HfSiON considered here is 3nm 80% HfSiON with an EOT of 1.46nm and a plasma nitridation. Other process conditions are the same as those in section 4.3.5.2.

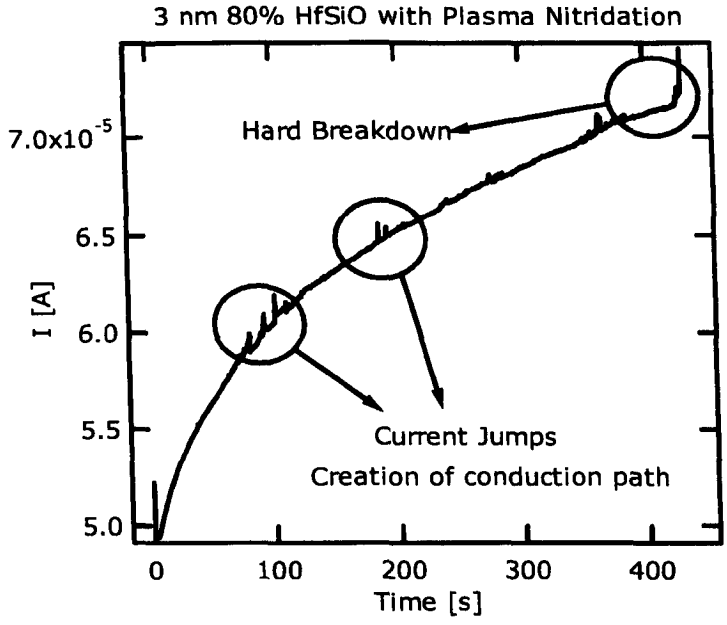


Fig.4.48 Gate current versus stress time for 3nm 80% HfSiO with plasma nitridation showing the creation of conduction paths followed by hard breakdown.

Fig.4.48 shows the gate current versus stress time with some current jumps followed by hard breakdown. There is a significant current increase due to 2 trap paths contrary to the data shown in section 4.3.5.2 for 2nm 50% HfSiO. This result behave similar to HfO_2 and 3-trap path triggers soft and hard breakdown (Fig 4.49)

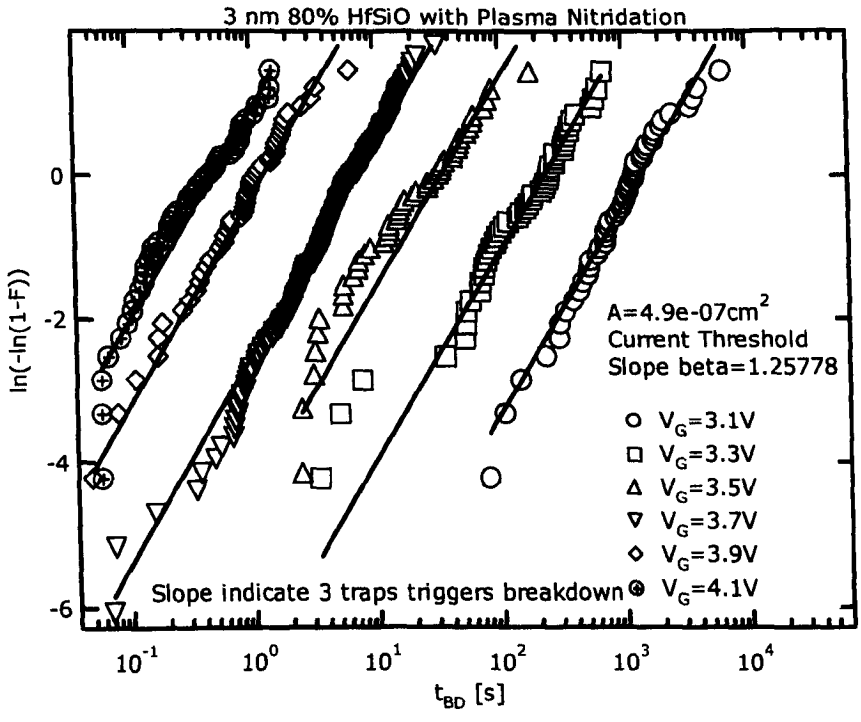


Fig.4.49 Time to breakdown t_{BD} distributions measured on 3nm 80% HfSiO with plasma nitridation for different stress voltage. Exponent beta β~1.25 indicate that 3trap triggers soft and hard breakdown.

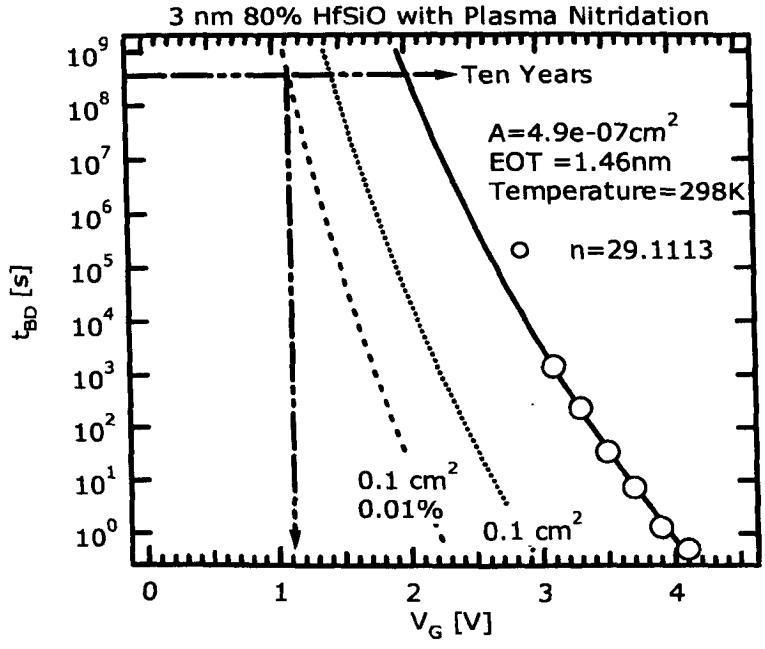


Fig.4.50 Low voltage extrapolation of t_{BD} for 3nm 80% HfSiO with plasma nitridation. A power law extrapolation was used with exponent ~29. Maximum operating voltage at 10 years is ~1.1V

Fig.4.50 shows that the lifetime extrapolation for 3nm 80% HfSiO, maximum operating voltage at 10 years is 1.1V

4.3.5.4 Evaluation of 3nm 50% HfSiON with Thermal Nitridation

The HfSiON considered here is 3nm 50% HfSiON with an EOT of 1.56nm and a thermal nitridation at 800°C. Other process conditions are the same as those in section 4.3.5.2.

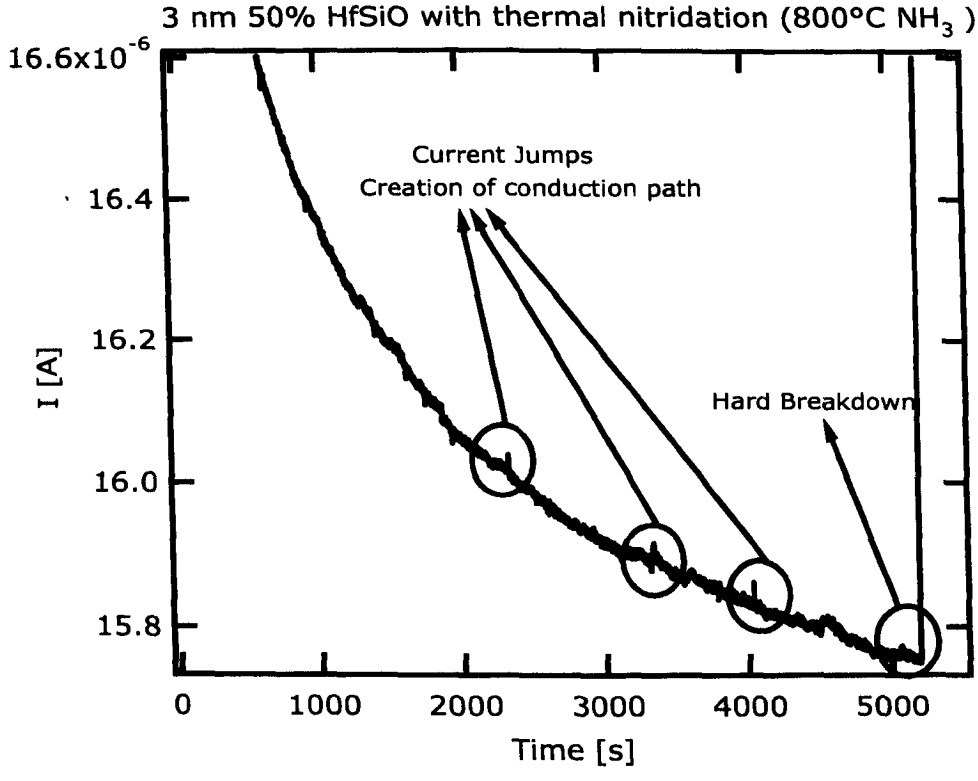


Fig.4.51 Gate current versus stress time for 3nm 50% HfSiO with thermal Nitridation showing multiple soft breakdown followed by hard breakdown

Fig.4.51 shows the gate current versus stress time until soft breakdown occurs followed by hard breakdown, There is no significant current increase due to 2 trap paths contrary to the data shown in section 4.3.5.3 for 3nm 80% HfSiO. 3-trap path triggers soft and hard breakdown (Fig 4.52).

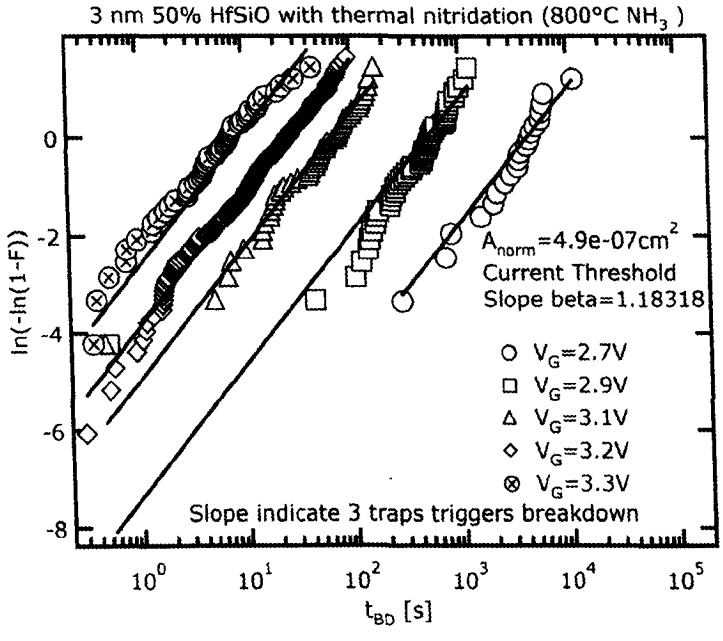


Fig.4.52 Time to hard breakdown t_{HBD} distributions measured on for different stress voltage. Exponent beta β~1.18 indicate that 3 traps triggers soft and hard breakdown

Fig.4.53 shows that the maximum operation voltage at 10 years is 1.1V.

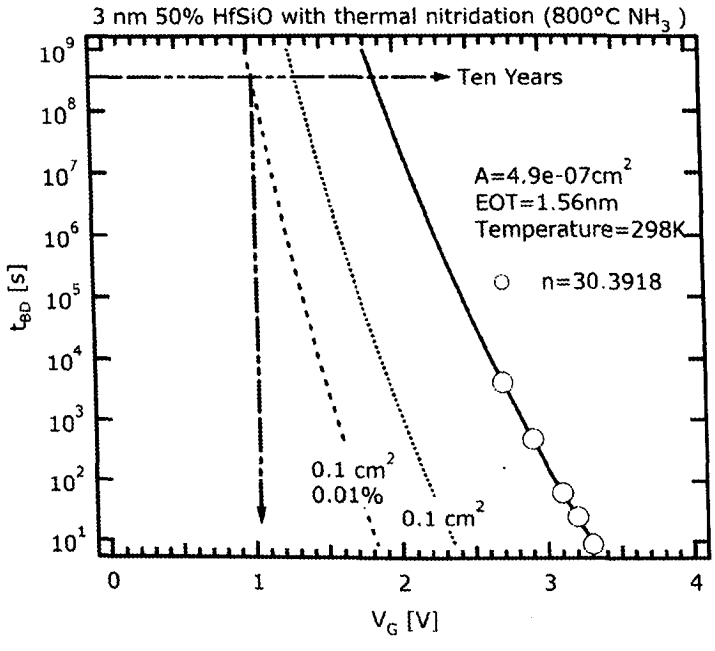


Fig.4.53 Low voltage extrapolation of t_{BD} for 3nm 50% HfSiO with thermal Nitridation A power law extrapolation was used with exponent ~30. Maximum operating voltage at 10 years is ~1.1V

4.4 TIME DEPENDENT-DIELECTRIC BREAKDOWN

The most common methodology for reliability studies and lifetime predictions of MOS devices is to apply a constant voltage until the breakdown (BD) of the dielectric. We discussed in chapter 2 that dielectric breakdown in insulating layers under electrical stress occurs when stress-induced traps form a percolation path between anode and cathode **[Degra98]**. The critical defect density necessary to form such a percolation path depends on the physical thickness of the dielectric, the capture cross-section of the traps and the device area **[Degra98]**. Several trap generation mechanisms, as anode hole injection **[Chen86]**, neutral trap generation **[Sune90]**, interface trap generation **[DiMa93]** and oxide charge trapping **[Nigam99]** were identified and subject of multiple studies **[Stath97]**, **[DiMa97]**, **[Degra99]**.

The local degradation under CVS of ultra-thin SiON with poly-Si gate dielectrics can be classified into 3 main stages (Fig.4.54) **[Kacz04]**: in the beginning (stage I) the dielectric at this spot is defect free, until at the time t_c the generation of a defect leads to the creation of a conductive path. It follows the phase of progressive wearout t_{pw} (stage II). Due to the locally enhanced current, more defects are accumulated along the conduction path and as a result, the current further increases. At certain current levels the whole process will saturate (stage III), because series resistance limits the current. For lifetime estimations, however, stage III can be ignored. The description and modeling of the voltage dependence of t_c and t_{pw} is the actual task.

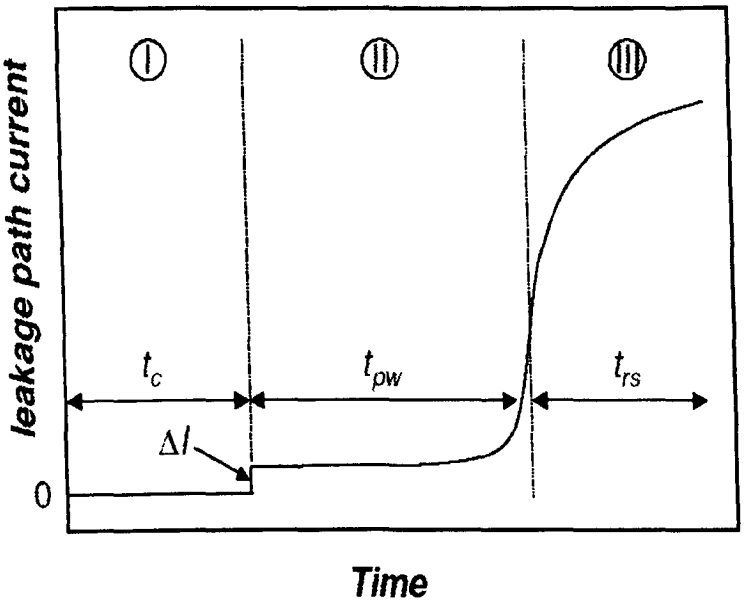


Fig.4.54 In ultra-thin dielectrics with poly-Si gates the oxide degradation is a 3 step process: I) the time to creation of a localized conduction path with leakage current ΔI; II) the progressive wearout of this path and eventually III) a saturation due to series resistance.

When SiON is replaced with high-k dielectrics like HfSiON or SiON/HfO_2 stacks, the features of the degradation process remain and at least for HfSiON with poly-Si gates, reliability can still be guaranteed **[Kau05]**.

Further reduction of the EOT requires the introduction of metal or Fully Silicided (FUSI) gates. With metal gates, we observe that the slow wearout process of breakdown spots has disappeared and is replaced by a systematic large current jump without any measurable transient. Different material combinations and thicknesses are investigated and show that large ΔI is a generally observed phenomenon with metal gates. These observations are consistent with earlier work on SiO_2 with tungsten gate by F. Palumbo et al. **[Paulm04]**.

Several dielectric/metal gate stacks with process conditions have been studied: SiON/TaN **[Hens04]**, SiON/FUSI **[Velos04]**, HfSiON/TaN, HfSiON/FUSI **[Velos04]**, ZrO2/TiN **[Kau02]** and HfO_2/TaN **[Schr05]**. The thickness of the dielectric ranged

from 0.8 nm EOT (HfO_2/TaN) to 5.9 nm EOT (ZrO2/TiN) and the Time-Dependent Dielectric Breakdown (TDDB) behaviour on various areas has been measured using Constant Voltage Stress (CVS) on nMOSFETs in inversion mode.

4.4.1 Hard breakdown in dielectric/metal gate stack

On all stacks with EOT < 3 nm it is consistently observed an abrupt current increase in the order of 0.1-10 mA (Fig.4.55) without any measurable transient or progressive wearout. Only on devices with an EOT > 3 nm some features of trap creation and soft breakdown (SBD) are found before a final hard breakdown (HBD) is triggered.

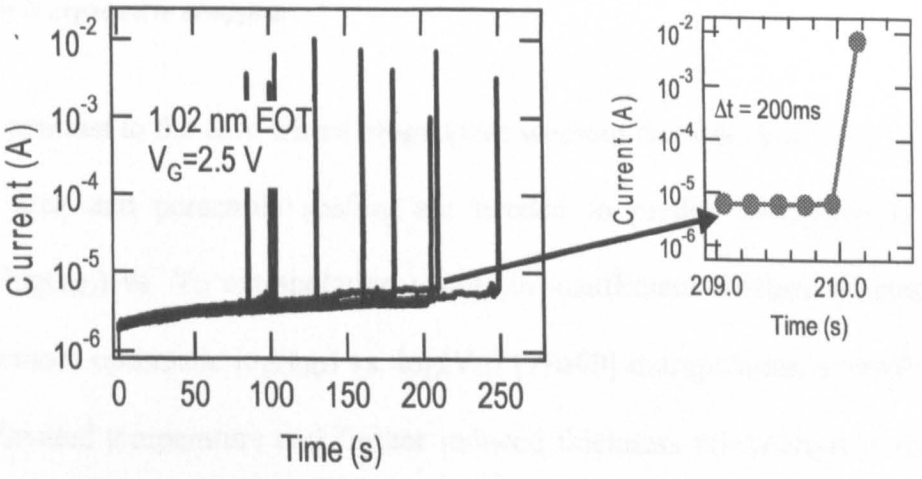


Fig.4.55 A typical set of I-t traces measured on HfO_2/TaN transistors. No SBD or progressive BD is observed and after the creation of a conducting path the current immediately jumps to several mA.

The findings indicate that the large ΔI is an inherent property of the breakdown process in dielectrics when combined with a metal gate. In particular, for some application the large ΔI will result in sudden device failure and if they persist at low voltage, this will seriously jeopardize the reliability of devices with metal gates as shown in Fig.4.56.

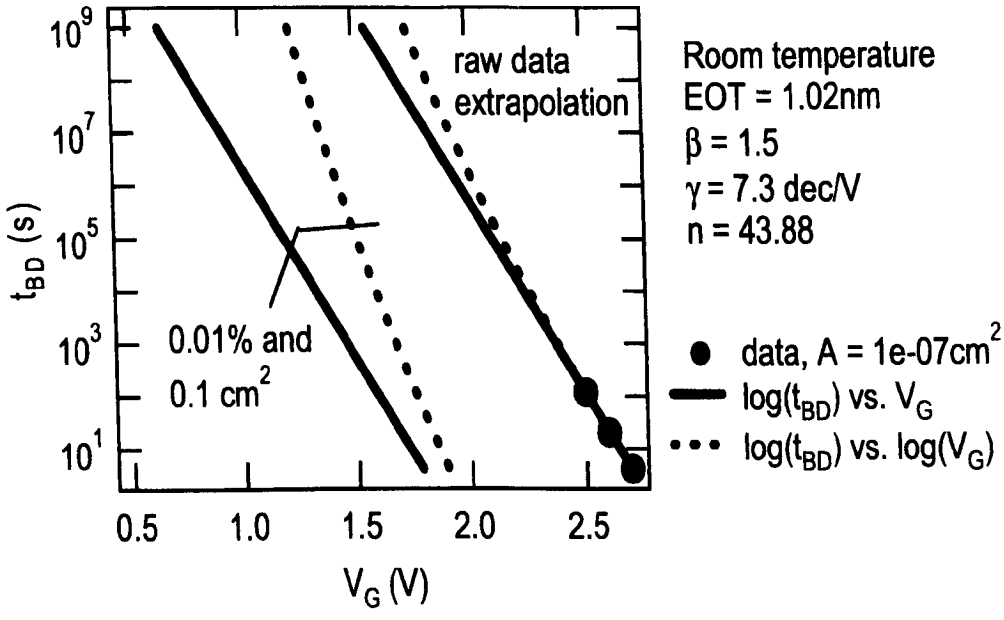


Fig.4.56 TDDB for one example stack (HfO_x/TaN). Using the most optimistic log(t_{BD}) vs. log(V_G) [Wu00] extrapolation a small margin is left. At elevated temperature and further reduced thickness this margin is expected to disappear.

In contrast to the case where progressive wearout dominates reliability **[Kacz04]**, **[Kau05]**, area and percentile scaling are needed to predict reliability at low V_G. Classical log(t_{BD}) vs. V_G extrapolation results in insufficient lifetime at operating V_G. Using the more optimistic log(t_{BD}) vs. log(V_G) **[Wu00]** extrapolation a small margin is left. At elevated temperature and further reduced thickness this margin is expected to disappear.

4.5 CONCLUSION

In this chapter we applied the techniques described in chapter 2 to investigate the trap generation in high-k metal gate dielectrics by using conventional and improved charge pumping. TDDB was used for the wafer level uniformity and for a reliable assessment of different wafers. The main conclusions of this chapter are: **In section 4.2.1** it is shown that by using amplitude sweep charge pumping the contribution from charging and discharging of bulk defect states becomes clearly evident and the charge per cycle strongly increases with increasing peak level and charging time. Furthermore a comparison for different gate electrode is studied with different channel length, the results shows that short channel shows low trap density independent of gate electrodes.

In section 4.3.1 it is demonstrated that the HfO_2 bulk trap density shows channel length dependence for fresh and degraded devices. Also, the time-to-hard-breakdown is channel length dependent. The numeric discrepancy between the channel length dependence of trap density and breakdown is consistently explained when trap generation and subsequent wear out are taken into account. This work also illustrates how extrapolating large area data to small area will lead to an underestimation of the reliability.

Section 4.3.2 investigated traps by using the VT^2CP, where the pulse low timing ("discharging time") and high level timing ("charging time") is independently controlled. This allows separating the traps in the interfacial SiO_2 from the traps in the HfO_2. The creation of new traps in both constituent layers are observed and this is proven by using the wafer level variation technique.

In section 4.3.3 trap generation during stress was investigated, the results show that during degradation the increase of traps, both in the SiO_2 as well as in the HfO_2,

follows a power law behavior as a function of time with an exponent ~0.32 and ~0.34 respectively independent of stress voltage. The voltage acceleration of creation of HfO_2 traps (-30) is nearly identical of the TDDB (-27), confirming the earlier published model that TDDB occurs when the density of traps in the HfO_2 reaches a critical value. VT^2CP can accurately detect degradation down to a much lower voltage than the dielectric breakdown measurement range. Furthermore, only one stress experiment combined with VT^2CP is sufficient to determine the degradation at a given voltage, while a TDDB test requires many measurements in order to construct an accurate distribution of failure times. The agreement between high voltage TDDB and low voltage VT^2CP data is only obtained if a *power law* voltage acceleration is assumed and not if an exponential model is taken. The data therefore confirm that power law [12] describes the voltage acceleration of dielectric degradation more consistently than exponential models.

In section 4.3.4 we used the VT^2CP to investigate trap generation in different layer composition and thickness. The result shows:

Low D_{SiO2} and D_{HfO2} for HfSiO is obtained for 3, 4 nm 80% Hf with plasma Nitridation and 50% Hf with thermal nitridation.

When using plasma nitridation the % Hf content have to be increased to have low D_{SiO2} and D_{HfO2} .

High D_{IT} for 3,4 nm 50% Hf and 2 nm 80% Hf with plasma nitridation, similar to the 3nm HfO_2.

High D_{SiO2} and D_{HfO2} for thicker HfO_2

High D_{IT} for thicker HfO_2 and Low D_{IT} for all 2nm HfO_2

Low D_{SiO2} and D_{HfO2} is obtained for 2 nm with 10s water pulse without MG degas.

Section 4.3.5 evaluates the trap generation using TDDB and wafer level uniformity. It is shown that water pulse length in ALD has no effect on t_{BD} value, but the deposition uniformity improves with 10 s water pulse length. HfSiO with 50 % Hf shows no significant leakage current increase before breakdown independent of nitridation technique and HfSiO with 80 % Hf shows leakage current increase before breakdown just as in pure HfO_2.

Section 4.4 is focused on hard breakdown. The findings indicate that the large ΔI is an inherent property of the breakdown process in dielectrics when the poly-Si gate is replaced with a metal gate and TDDB lifetime is consistently limited by an abrupt large current increase of several hundred μA

CHAPTER 5 POSITIVE CHARGES AND V_{TH} INSTABILITY

5.1 INTRODUCTION

As the gate leakage current through SiON approaches an intolerable level, intensive efforts have been made to find an alternative gate dielectric for future CMOS technologies. The leading candidate is HfO_2/SiO_2 or Hf-silicates, which have higher dielectric constant (high-k) than SiON. For an electrically equivalent oxide thickness (EOT), high-k layers are physically thicker, resulting in smaller gate leakage current **[Groes04]**. When poly-si gate is used, the gate depletion is highly undesirable, since it contributes to EOT considerably **[Groes04]**. The interaction between poly-si and HfO_2 is also problematic **[Groes04], [Ma05]**. This is the main reason why metal gates have attracted a lot of attention recently **[Groes04], [Rag03], [WangX05]**. Apart from the elimination of gate depletion, it has been reported that metal gates improve the electron mobility **[ChauR04]** and reduce the electron trapping, induced threshold voltage instability **[Ma05], [WangX05].** The question is how metal gates and the processing condition affect other device instabilities, such as positive charge formation in gate dielectric and the resultant threshold voltage, V_{TH}, instability which is one major issue for future MOSFETs with Hf-based gate dielectrics. V_{TH} instability has been studied intensively, but the attention was focused on nMOSFETs and there is little information for pMOSFETs. This chapter focuses on two effects, positive charge formation using different gate materials in nMOSFETs and an investigation of the V_{TH} instability in pMOSFETs.

5.2 IMPACT OF THE GATE MATERIAL ON POSITIVE CHARGE

To assess the impact of gate materials on positive charge formation, samples were prepared with either metal or n^+-poly-si gates. The metal gate was prepared by PVD at 150°C and consists of a 10nm TaN layer capped by a 70nm TiN [6]. For convenience, this TaN/TiN gate will be referred to as 'metal gate' hereafter. HfO_2 layer was prepared by ALCVD to a physical thickness of 4.0nm. Before HfO_2 deposition, there is a 0.4nm chemical oxide. To activate the dopant, a 1000°C anneal was used. The equivalent interfacial SiO_2 layer is around 0.9nm and the total EOT is 1.75 and 1.8nm for metal and poly-gated samples, respectively. The size of nMOSFETs used is 0.8μm in length and 10μm in width. The substrate is p-type silicon.

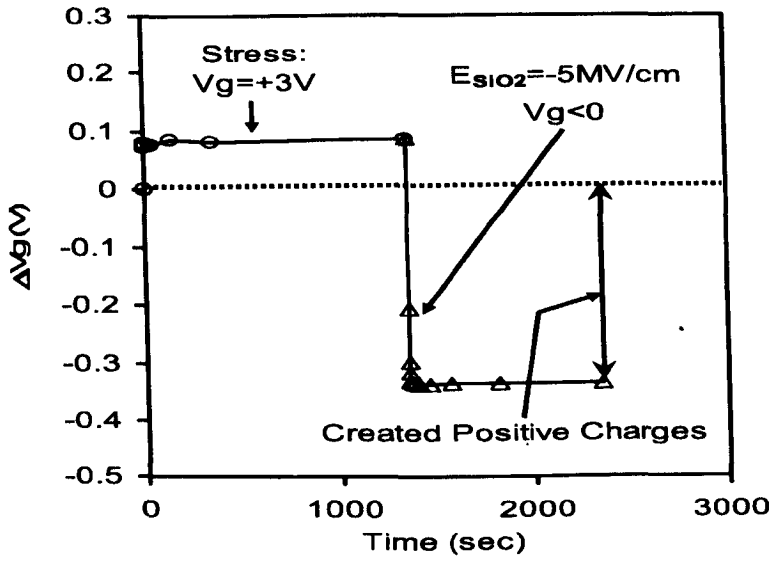


Fig.5.1 Typical test procedure. Devices were first stressed under high electrical field. V_G=3V gives an electrical field across the interfacial SiO_2 layer of $Esio_2$=11.5MV/cm. The electron trapping during the stress leads to a positive gate voltage shift, ΔV_G. After the stress, a relatively low negative field of $Esio_2$=-5MV/cm was applied to detrap electrons from the dielectric and allow the positive charges being measured from the negative ΔV_G. The ΔV_G was measured from the gate voltage shift in the subthreshold region of the transfer characteristics.

To compare the positive charge formation in samples of different gates, it is desirable to stress them by the same gate bias and current. A gate bias of 3V was chosen for stressing devices, since it produces approximately the same gate current for both sets of samples. The corresponding electrical field across the interfacial SiO_2 layers, $Esio_2$, is 11.5MV/cm. The test procedure is shown in Fig.5.1

The trapped charges were monitored from the gate voltage shift ΔV_G at a constant drain current of 10^{-7} A with a drain bias of 0.1V. During the stress at V_G=3V, there is a positive gate voltage shift, ΔVg, because of the well-known electron trapping in HfO_2 **[Groes04], [Ma03], [ZhaoC05]**. To measure the positive charge after the stress, a relatively low negative electrical field of $Esio_2$= –5MV/cm was applied for 1000sec, which is sufficiently long to detrap electrons from HfO_2 and to reach a steady negative ΔV_G. The ΔV_G was measured from the shift of transfer characteristics in the subthreshold region. The measurement time is in the order of one second and we would like to point out that some positive charges close to the interface can be neutralized during this period. These lost charges will not be investigated here and we focus on how the gate material affects the positive charges that are not neutralized within one second. Our measurement shows that the effective area density of these surviving positive charges can be in the order of 10^{12} cm^{-2}, which is considerable.

We are now ready to compare the positive charge formation in samples of different gates. Fig.5.2 shows that the positive charge formation for metal gate is significantly higher. We speculate that the enhanced positive charge formation originates from a higher hydrogen concentration for the sample with metal gate. Fig.5.3a shows that hydrogen can be located near the gate/dielectric interface. Under a positive gate bias, tunnelling electrons lose their energy near the gate interface and release these hydrogenous species. As hydrogen travels through the dielectric, positive

charges were formed, probably by trapping the proton on trivalent oxygen centers **[Houssa05]**

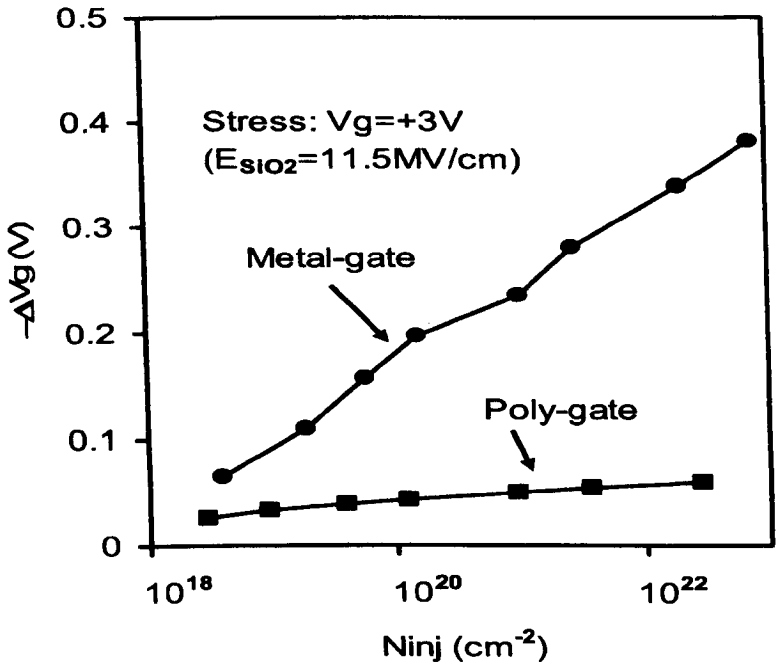


Fig.5.2 Impact of gate materials on positive charge formation under positive gate bias stresses. The Ninj is the electron fluency during the stress. The positive charge formation with the metal gate is significantly higher than that with poly-si gate.

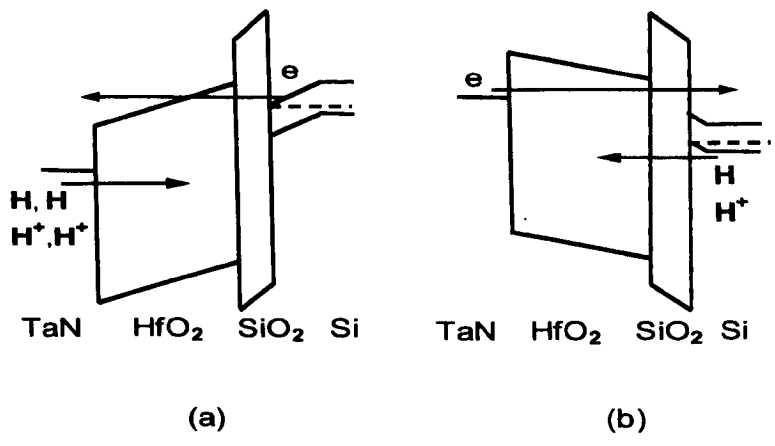


Fig.5.3 Schematic energy band diagrams for metal-gated samples under positive (a) and negative (b) gate bias stresses. It is assumed that hydrogen released from the anode dominates the positive charge formation and the hydrogen density near the TaN gate interface is higher than that near the substrate interface.

The support for this explanation is given below:

Based on the above assumption, we can make two predictions. One is that positive charge formation will be insensitive to the gate material under a negative gate bias. This is because, as illustrated by Fig.5.3b, hydrogen is released from the substrate interface under negative gate bias and the substrate interface should be similar for both types of gates. The other prediction is that there will be a gate polarity dependence for the positive charge formation with the metal gate, since the hydrogen concentration at the metal/dielectric interface is different from that at the dielectric/substrate interface, as shown in Figs.5.3a&b. Fig.5.4 confirms both predictions and supports our explanation strongly.

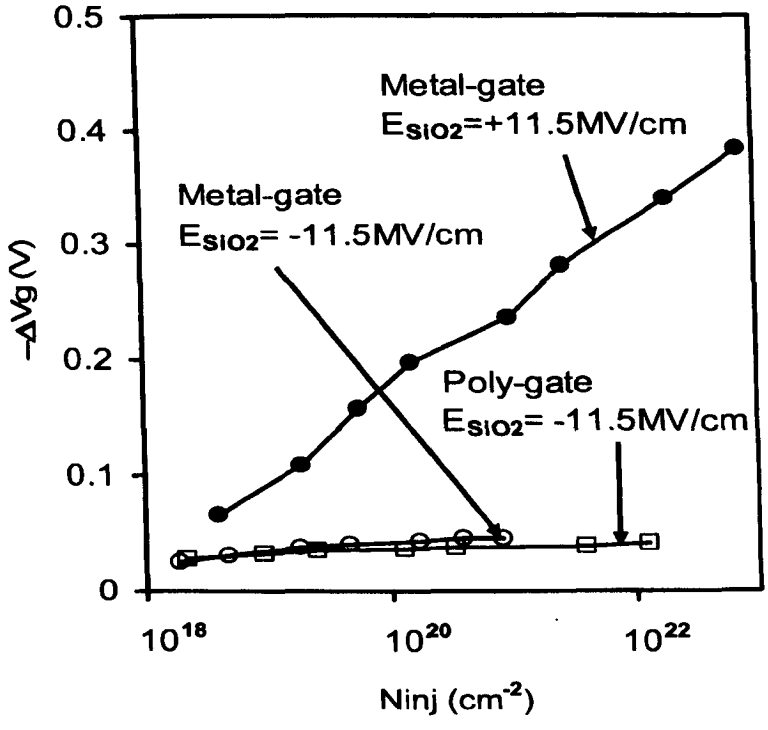


Fig.5.4 Dependence of positive charge formation on gate bias polarity and gate materials. When stressed under negative gate bias of $Esio_2$=-11.5MV/cm, positive charge formation in metal gate is similar to that in poly-si gate. For the metal gate, positive charge formation under $Esio_2$=-11.5MV/cm is significantly less than that under $Esio_2$=+11.5MV/cm.

We do not know why hydrogen concentration is higher near the metal/dielectric interface. It is possible that this is caused by the material difference between metal and poly-si and there are more hydrogen traps near the metal/dielectric interface. It is also possible that metal was deposited at 150°C, but poly-si was prepared at 600°C. More hydrogenous species could be trapped in the device at lower temperature. This explanation is supported by two early observations. One is that Si-H will be dissociated and hydrogen will be released from the device when the temperature rises above 550°C, approximately **[Stes00]**. This implies 600°C is not in favor of trapping hydrogen near Si/dielectric interface. The other observation is when metal gate (i.e. aluminum) was used in the early generation of MOSFETs, it is known that there were a large amount of hydrogenous species trapped in the device, resulting in the formation of positive charges **[RohY93], [Tromb91], [YoungD79],.[ZhangJ92]** When aluminum was replaced by poly-si, hydrogen and positive charge formation reduce significantly **[RohY93]**.

To further explore the connection between the positive charges observed for metal gate here with the positive charge reported for SiO_2, we compare their behaviour. In SiO_2, there can be two types of hydrogenous species: one is reactive and the other is not **[ZhangJ01]**. Under a positive gate bias, the positive charge induced by the non-reactive hydrogen is driven towards the oxide/substrate interface, leading to ΔV_G moving in the negative direction **[ZhangJ01], [Mac00], [Vanh97]**. In the opposite, the positive charge induced by reactive hydrogenous species, commonly known as anomalous positive charges (APCs), can be neutralized under a positive gate bias, resulting in ΔV_G moving in the positive direction **[RohY93], [Tromb91], [YoungD79],.[ZhangJ92]**. The APC is recharged when gate bias polarity is switched to negative **[RohY93], [Tromb91], [YoungD79], [ZhangJ92]**.

It is interesting to study which type the positive charge formed in HfO_2/SiO_2 belongs to. **[ZhangJ04], [ZhaoC04].** In Fig.5.5, we first check the effect of alternating gate bias polarity on fresh devices. The $Esio_2$ chosen here is ±5MV/cm. The ΔV_G is negligible under $Esio_2$ = -5MV/cm, but a small positive ΔV_G can be observed under $Esio_2$ = +5MV/cm, which is caused by electron trapping **[ZhaoC05]**

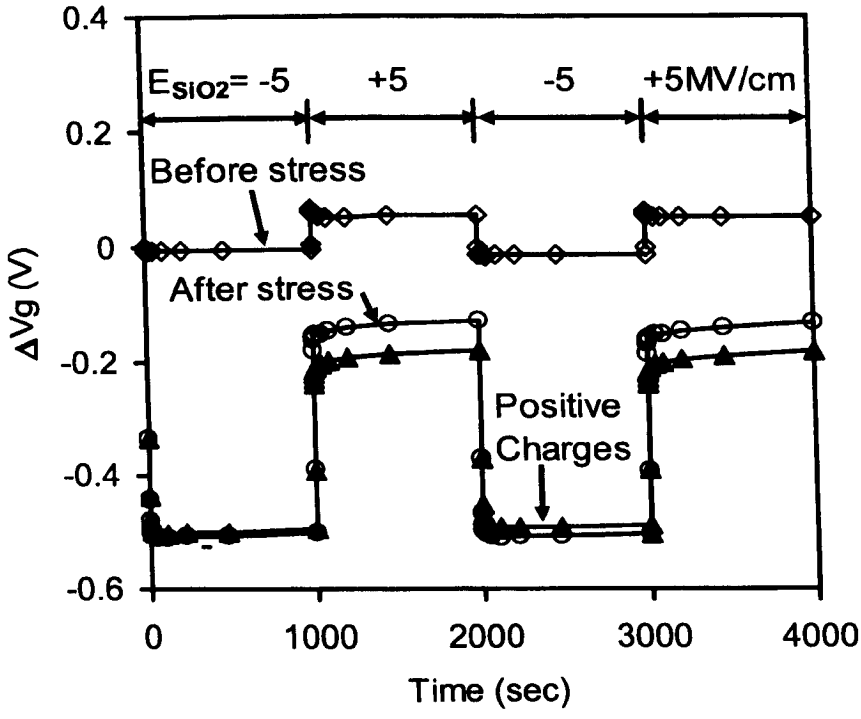


Fig.5.5 Behavior of positive charges formed in metal-gated HfO_2/SiO_2 stack under alternating gate bias polarities ($Esio_2$=±5MV/cm). When $Esio_2$=±5MV/cm was applied to a fresh device, the symbol '◊' shows that charging is negligible under $Esio_2$=-5MV/cm, but some electron trapping occurs under $Esio_2$=+5MV/cm, leading to a positive ΔV_G. After a stress under V_G=3.5V for an electron fluency of 8.8×10^{22} cm^{-2}, a large part of the formed positive charges can be repeatedly discharged under $Esio_2$=+5MV/cm and recharged under $Esio_2$=-5MV/cm. The symbol 'o' is the directly measured ΔV_G. The symbol ' ' represents positive charges, after correcting the electron trapping effect.

After stresses, the cyclic ΔV_G is significantly enhanced. After correcting the electron trapping under $Esio_2$=+5MV/cm, Fig.5.5 shows that a large part of positive charges can be neutralized under positive gate bias and recharged under Vg<0.

This clearly shows that the positive charge formed in HfO_2/SiO_2 behaves similarly to the APC in SiO_2

5.3 THRESHOLD VOLTAGE INSTABILITY OF PMOSFETS WITH HF-BASED DIELECTRICS

The commercial debut of the Hf-based high-k layers as gate dielectric has been held back by a number of issues, including low carrier mobility, interaction between gate and high-k layer, thermal instability, dielectric breakdown, and threshold voltage, V_{TH}, instability **[Groes04], [Ma05]**. For the V_{TH} instability, attentions were focused on nMOSFETs **[Ker03B], [Shan03], [Bers04], [Sim05], [ZhangJ06], [ZhaoC06]. [Groes04], [Ma05].** It is caused by electron trapping in the HfO_2 and the dominant trap is 'as-grown', rather than generated **[ZhangJ06], [ZhaoC06]**. The V_{TH} instability of nMOSFETs can be suppressed in two ways: using Hf-silicates **[Shan03]** and reducing the thickness of HfO_2 layers **[Bers04], [Sim05]**.

There is little information on the V_{TH} instability of pMOSFETs and the implicit assumption is that it is not important, since the electrons trapped in HfO_2 can be efficiently detrapped under a negative gate bias **[Sim05], [ZhangJ06]**. In this project it is shown that this is not true and positive charge formation can lead to considerable V_{TH} instability for pMOSFETs. For the first time, it is found that the characters of V_{TH} instability of pMOSFETs are different from those of nMOSFETs. In the following, a systematic investigation on the V_{TH} instability in pMOSFETs is carried out.

5.3.1 Comparisons of V_{TH} instability between nMOS and pMOS

The pMOSFETs used has a 2nm HfO_2 layer, prepared by the atomic layer deposition (ALD) with an interfacial layer nitrited in NH_3 at 900°C for 60 sec.

The equivalent oxide thickness (EOT) is 1.13nm. The gate metal is TaN. A 1nm Hf-silicate (30%SiO_2) was also prepared and nitrided in NH_3 at 800°C for 60sec, which gave an EOT of 1.3nm. To compare the features of V_{TH} instability of pMOSFETs and nMOSFETs, nMOSFETs were prepared with a 4nm ALD HfO_2 layer (EOT=1.75nm) or a 1.5nm Hf-silicate (77%SiO_2, EOT=1.53nm) also nitrided at 800°C for 60sec in NH_3. The channel length and width for both pMOSFETs and nMOSFETs is 0.25μm and 10 μm, respectively.

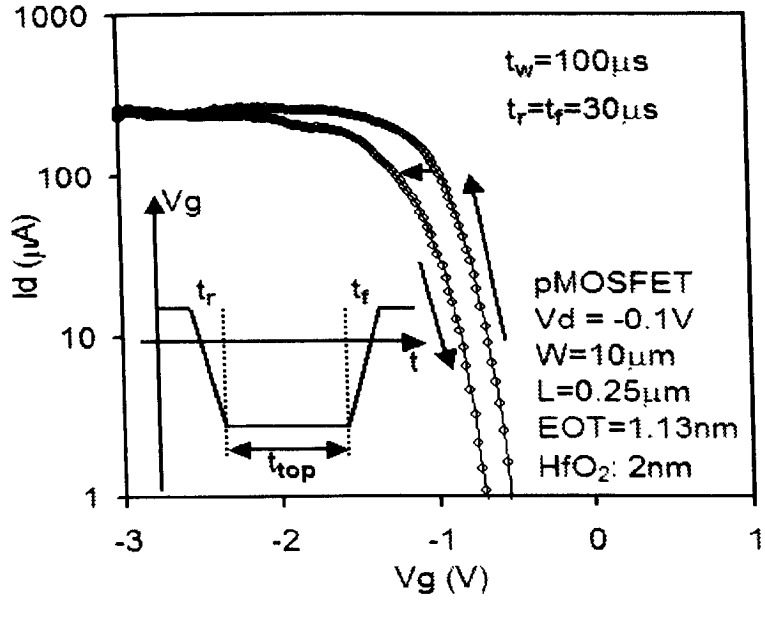


Fig.5.6 The instability of pMOSFETs. The inset shows the gate pulse applied. The transfer characteristics (TC) during the two edges of the pulse were recorded. The negative shift of the TC results from the positive charge formation in the gate dielectric.

For nMOSFETs, it is known that the V_{TH} instability is highly dynamic and can be substantially underestimated, if measured from the shift of the traditional quasi-dc transfer characteristics (I_D~V_G) **[Ker03B], [Shan03], [ZhangJ06], [ZhaoC06]**. To suppress the recovery during the measurement, the pulsed I_D~V_G technique was developed, which can reduce the delay between the stress and the measurement to the

order of microseconds **[Ker03B]**. We will use this pulsed I_D~V_G technique here to assess the V_{TH} instability of pMOSFETs.

The pulse applied to the gate is shown in the inset of Fig.5.6. Fig.5.6 also gives the I_D~V_G during the two edges of the pulse recorded under a drain bias of –0.1V. Like nMOSFETs, a 'loop' can be seen for the two I_D~V_G, indicating positive charging in the dielectric stack during the period of t_{top}. However, the similarity to nMOSFETs ends here and the important differences are addressed next

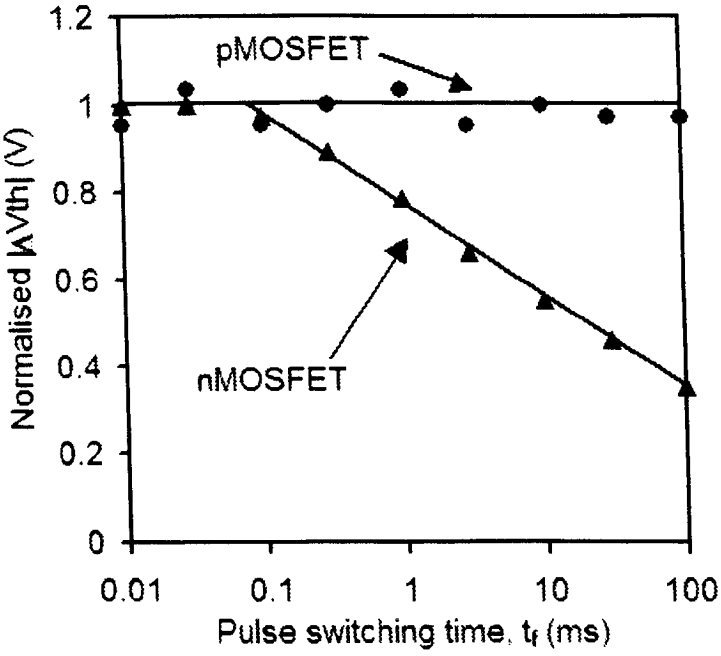


Fig.5.7 Dependence of the measured threshold voltage instability, ΔV_{TH}, on the measurement time. The measurement time is controlled by the pulse switching time, t_f, shown in the inset of Fig.5.6 Unlike nMOSFETs, the Vth instability of pMOSFETs is insensitive to the measurement time. The solid lines are guides for the eye.

<u>*Dependence on the measurement time:*</u> For the pulsed I_D~V_G technique, the delay between the stress and measurement is controlled by the second edge of the pulse, t_f, in the inset of Fig.5.6. Fig.5.7 shows that the V_{TH} instability of nMOSFETs decreases progressively as the measurement time increases. However, the V_{TH} instability of

pMOSFETs is insensitive to the measurement time. This implies that the positive charges in pMOSFETs are more stable than the trapped electrons in nMOSFETs.

Impact of using Hf-silicates: It has been reported that the V_{TH} instability of nMOSFETs can be suppressed by using Hf-silicates **[Shan03]**. This is also observed in our nMOSFETs in Fig.5.8a. However, Fig.5.8b shows that considerable V_{TH} instability is present for the pMOSFETs with a sub-2nm nitrided Hf-silicate and the instability in pMOSFETs can be higher than that in nMOSFETs. To explore this difference, it is worth of examining the electron trapping and positive charging in pure SiO_2 or SiON. It is well known that as-grown electron traps are negligible in a modern MOSFET with SiO_2 or SiON **[ZhangW02]**, **[ChangM06]**. However, substantial amount of as-grown hole traps exist in SiO_2 or SiON **[Stah93]**, **[ZhangJ01A]**. As a result, it is not surprising that adding Si to high-k layer is less effective for suppressing positive charging.

Dependence on the stress gate voltage, V_{GS}: V_{GS} is the gate voltage during the period of t_{top} in Fig.5.6. Fig.5.8a shows that the V_{TH} instability of nMOSFETs saturates at higher V_{GS}. This is because, once all as-grown electron traps were filled, the V_{TH} will not shift further for higher stress levels. Such a saturation behavior cannot be observed for pMOSFETs in Fig.5.8b.

Dependence on the carrier fluency: Since saturation cannot be observed for the V_{TH} instability of pMOSFETs in Fig.5.8b, it is possible that the defect responsible for the positive charging is created, rather than as-grown. If this is true, the positive charging should increase continuously with the fluency of carriers injected into the dielectric during the stress **[ZhaoC05A]**. To test this assumption, the V_{TH} instability is plotted against the carrier fluency in Fig.5.9. It is obvious that the V_{TH} instability of pMOSFETs does not increase with the carrier fluency. This is against the assumption that the instability of pMOSFETs originates from new defects generated during stress.

In the following, some speculations will be given to explain why the instability of pMOSFETs does not saturate with increasing $|V_{GS}|$ and does not increase with carrier fluency.

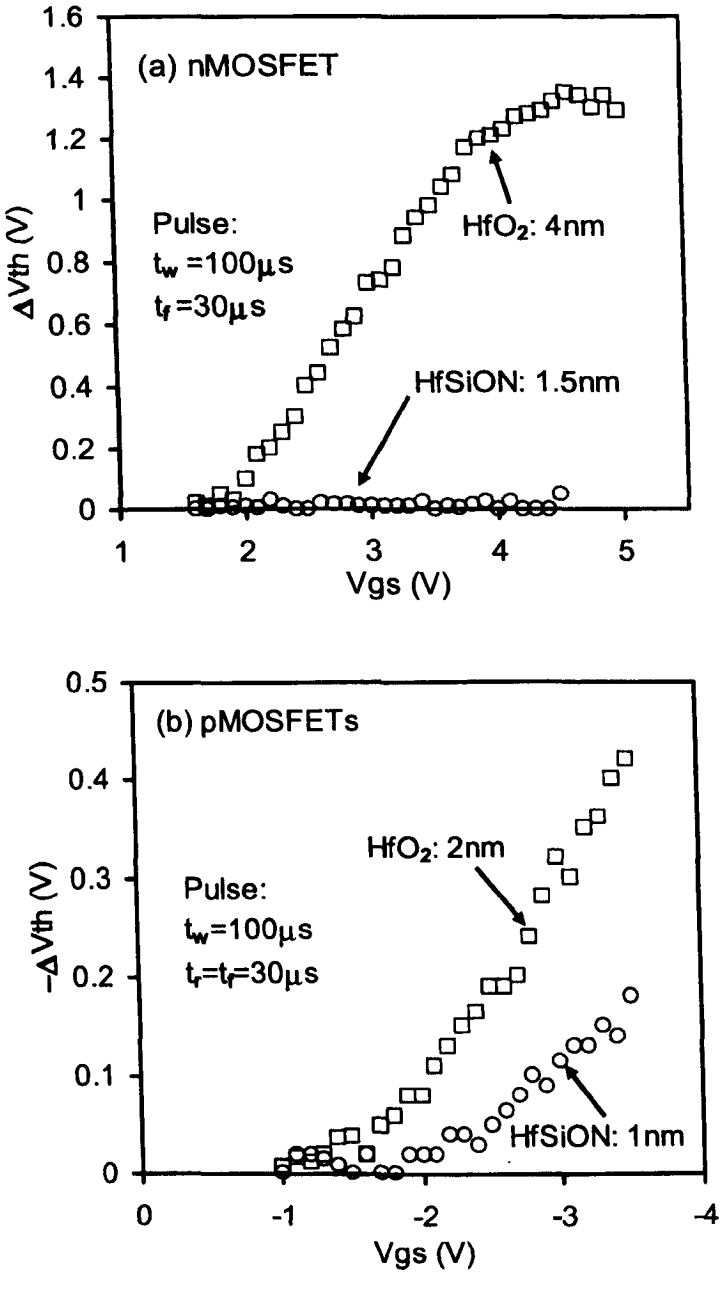


Fig.5.8 Dependence of ΔV_{TH} on the stress gate voltage and the impact of using Hf-silicates. V_{GS} is the bias level during the period of t_{top}, as shown in Fig.5.6. (a) nMOSFETs: ΔV_{TH} saturates as V_{GS} increases and the instability is negligible in the Hf-silicate. (b) pMOSFETs: ΔV_{TH} does not saturate with V_{GS} and considerable instability occurs in a sub-2nm Hf-silicate.

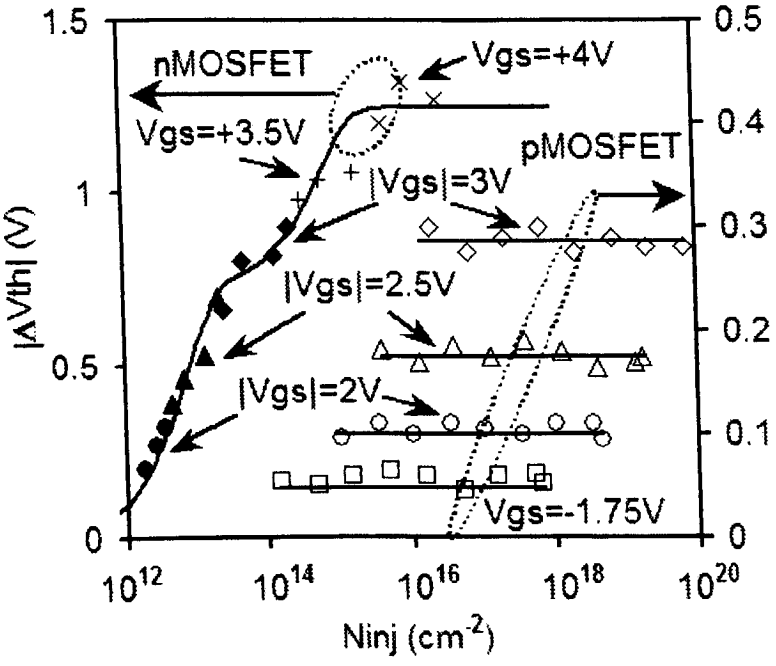


Fig.5.9 Dependence of ΔV_{TH} on the carrier fluency injected into the gate dielectric, Ninj. Under a constant V_{GS}, the ΔV_{TH} of pMOSFETs is insensitive to N_{inj}, but ΔV_{TH} of nMOSFETs increases with N_{inj}. Each symbol represents results obtained at a given V_{GS}. The open symbols represent pMOSFETs. The filled symbols and '+' and 'x' represent nMOSFETs. V_{GS} and ΔV_{TH} are negative for pMOSFETs and positive for nMOSFETs. The different N_{inj} at a given V_{GS} was obtained by varying the time t_{top} shown in Fig.5.6. The lines are guides for the eye.

For the V_{TH} instability of nMOSEFTs, the acceptor-like electron traps are initially neutral, since the flatband voltage is found to be insensitive to the dielectric thickness **[Cart04]**. A trap has to capture an injected electron to be negatively charged. The higher carrier fluency causes the higher trapping, as is shown in Fig.5.9. Here, once the energy of tunneling electrons is higher than the trap energy level, it is the carrier fluency that controls the V_{TH} instability **[ZhangJ06]**.

For the V_{TH} instability of pMOSFETs, its insensitivity to carrier fluency suggests that the charging does not involve capturing injected carriers. We speculate that there are as-grown donor-like defects in the dielectric stack, as shown in Fig.5.10.

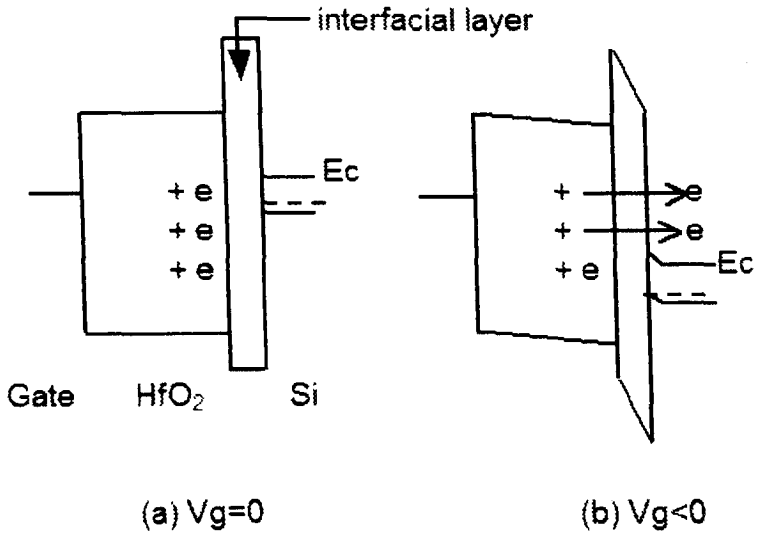


Fig.5.10 A schematic illustration of the defect and physical process responsible for the ΔV_{TH} of pMOSFETs. Under V_G=0, the donor-like defects are neutral (a). Under V_G<0, electrons tunnel away, leaving positive charges in the dielectric.

For a fresh device, these defects are neutral. Under a negative gate bias, their energy level rises above the bottom edge of silicon conduction band, Ec. The defect gives away its electron and becomes positively charged. The injected carriers are not needed here. The higher $|V_{GS}|$, the more defects move above the Ec and consequently, the more positive charging. This explains the non-saturation of V_{TH} instability of pMOSFETs as $|V_{GS}|$ increases.

5.4 SEPARATION OF AS-GROWN TRAPS FROM GENERATED TRAPS

In the previous section, it is shown that, for sub-2nm Hf-dielectrics, pMOSFETs can have more V_{TH} instability than nMOSFETs.

In this section, the pulse I_D-V_G and Negative Bias Temperature Instability (NBTI) is used to consistently investigate and separate positive charging caused by as-grown traps from those by generated traps in pMOSFETs.

Different samples are used and summarized in Fig.5.11. Both Hf-silicates and HfO_2 were prepared by ALD for pMOSFETs with TaN gate. NH_3 anneal at 900°C for the interface was used for the HfO_2 and a post deposition anneal in NH_3 at 800°C for 60s was carried out for Hf-silicate. The channel length and width is 0.25~1μm and 10 μm, respectively.

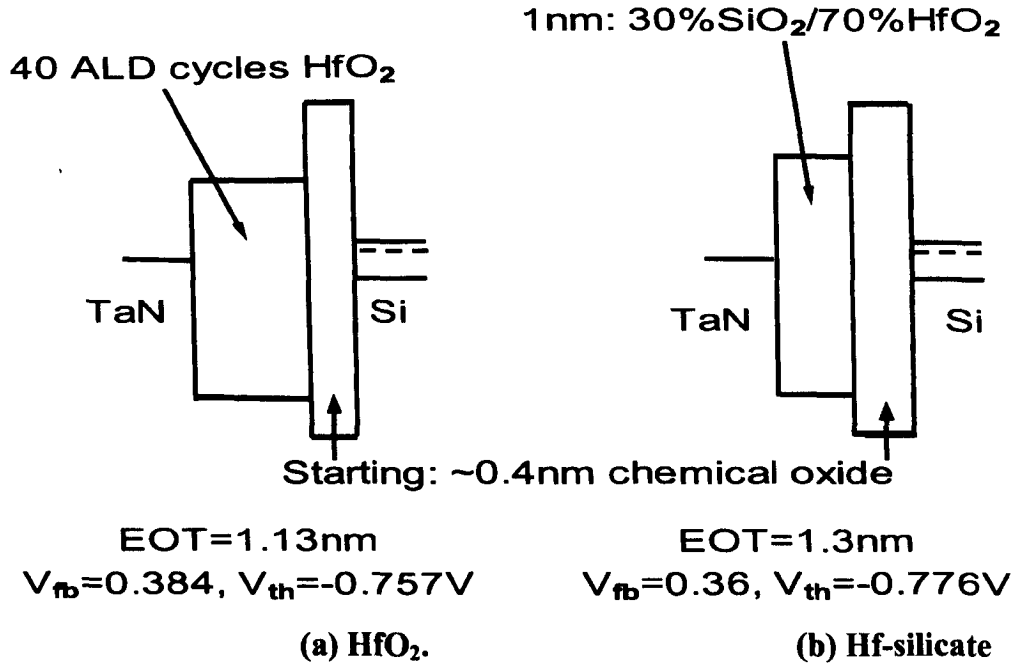


Fig.5.11 Schematic energy band diagrams with key data for the dielectric stacks.

5.4.1 Generated V_{TH} instability

The pMOSFETs V_{TH} instability discuss in the previous section is investigated in more details. The assumption made so far are that the as-grown traps cause the V_{TH} instability of pMOSFETs.

In this section a systematic experiments are carry out to unambiguously understand if the V_{TH} instability observed is only caused by as-grown traps. In Fig.5.12, a gate pulse as shown in the inset is first applied to a fresh device and the obtained I_D-V_G curves during up and down ramp are represented by the solid curves. After the device was stressed at Vg=-2.5V for 10^4sec, the same gate pulse was applied again and the resultant I_D-V_G during up and down ramp are represented by the dotted curves. It is clear that two different effects can be seen: 1) the long stress leads to a shift of the loop and 2) the loop size remain the same and is not affected by stress. The different impact of the two types of instabilities on ΔV_{TH} is illustrated in Fig 5.13.

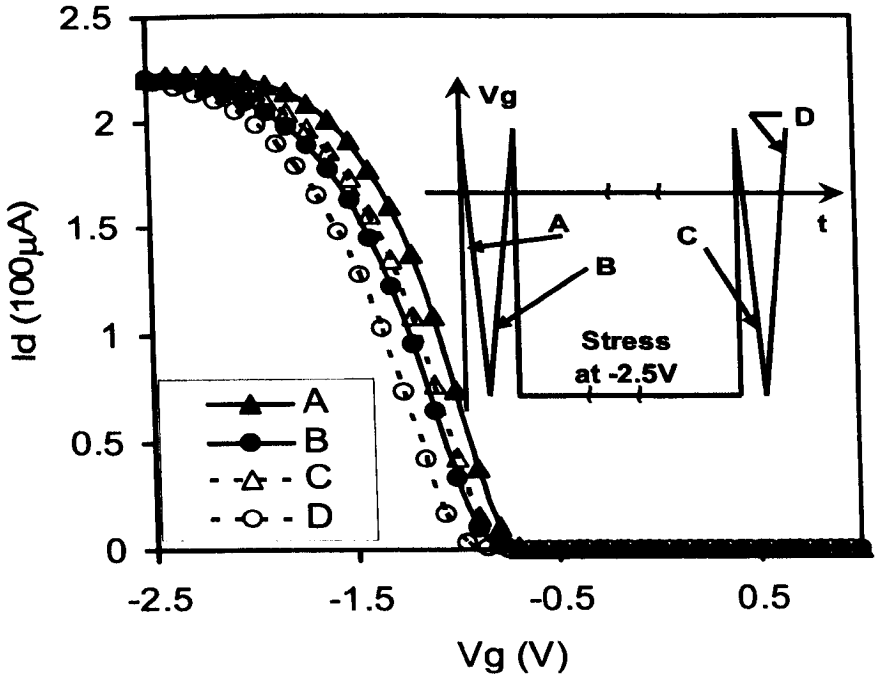


Fig.5.12 Separation of as-grown defects from the generated defects. As-grown VTH instability appears as a 'loop' (two solid curves) and the NBTI shifts the loop to the two dotted curves.

Two types of V_{TH} instability clearly exist:1) as-grown defects for the loop and 2) generated defects for the shift of the loop. The negative shift of the whole loop after a long stress in Fig.5.12 indicates that positive charges were generated during the stress. Unlike the as-grown donor-like defects which are neutralized under $V_G \geq 0$, the created

positive charge, curve 'C' in Fig.5.12, responsible for NBTI can survive the neutralization when $V_G \geq 0$ was applied. As a result the created positive charges can have higher energy levels (see Fig.5.10), making them anti-neutralizing **[ZhangJ04]**, **[ZhaoC05]**.

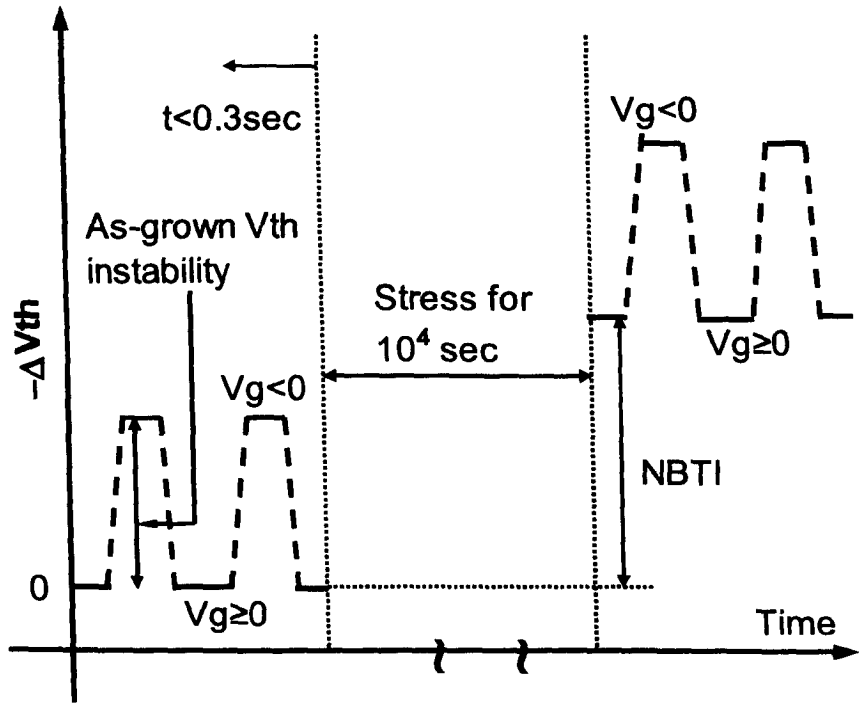


Fig.5.13 A schematic illustration of the different impact of the two types of instabilities on ΔV_{TH}

To provide further evidence on the presence of two different types of defects, the same loop size was measured at different voltage and pulse timing. Fig.5.14 shows that the size of the loop is not affected by the stress (V_G = -3V) and Fig.5.15 shows that the created defect leads to a shift of the loop.

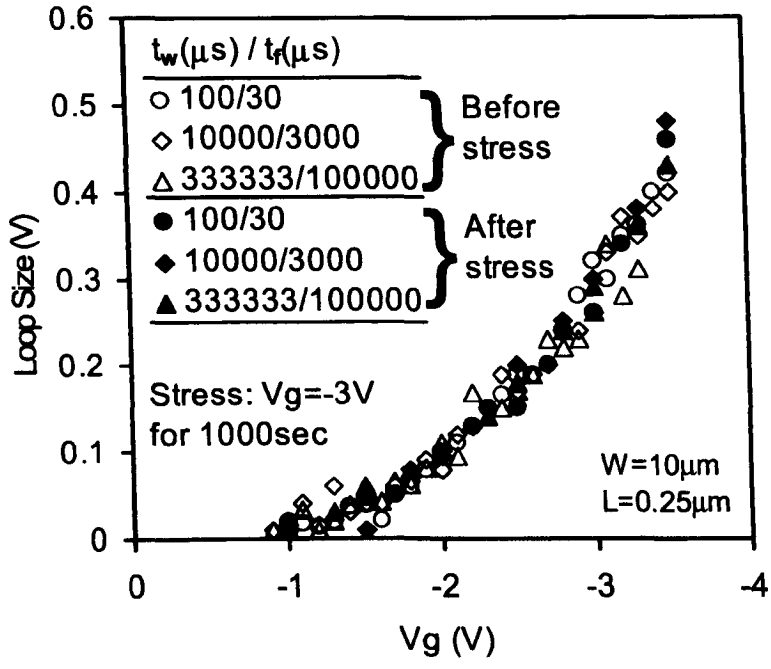


Fig.5.14 Separation of as-grown V_{TH} instability from NBTI: Shows the loop size is not affected by the stress

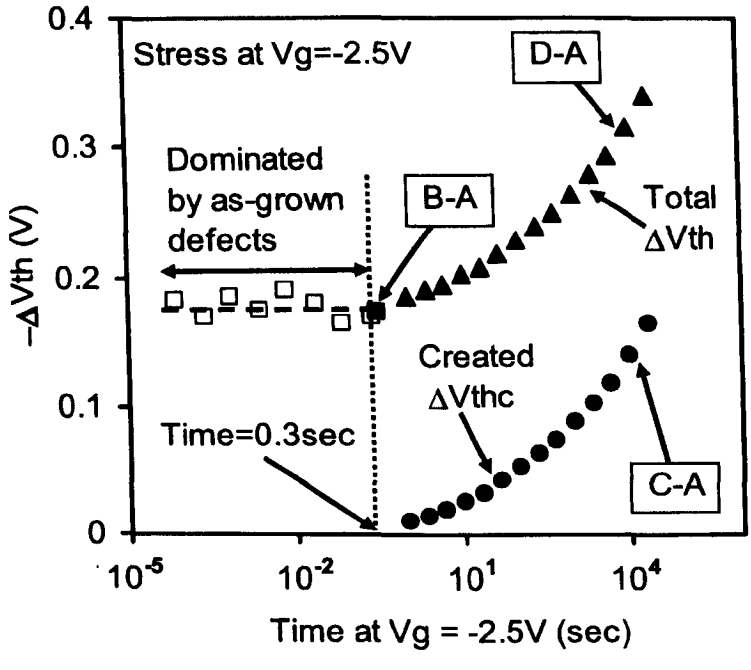


Fig.5.15 Separation of as-grown defects from the generated defects: As-grown defects dominate for time <0.3sec. The letter A,B,C,D refers to the inset in Fig.5.12.

Fig.5.16 shows the ΔV_{TH} against NBTS time. Instead of following the same power law, a constant ΔV_{TH} was observed for t<0.3sec. This behavior supports that there are two types of instabilities. The insensitivity of ΔV_{TH} for t<0.3sec suggests that it originates from as-grown defects and will be referred to as 'as-grown V_{TH} instability'. This instability is similar to the so-called V_{TH}-instability in nMOSFETs **[Ker03B]**.

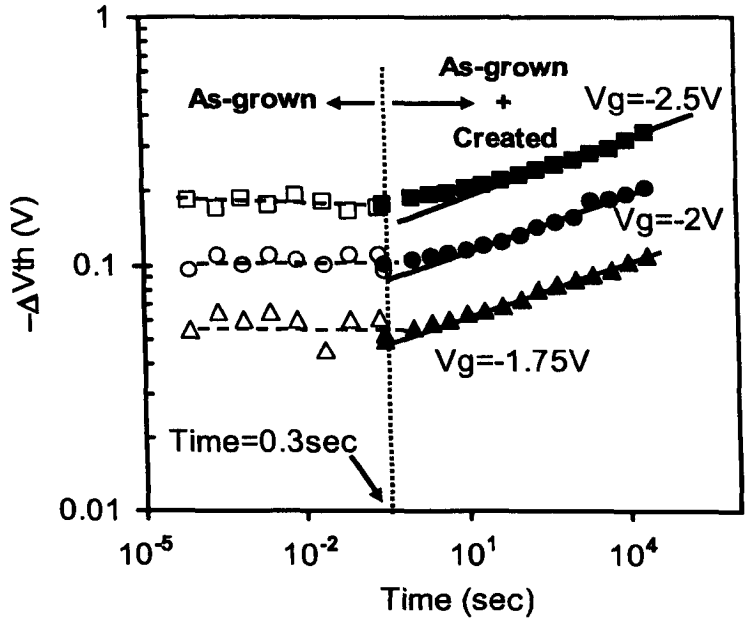


Fig.5.16 Different dependence on stress time in different ranges of time for HfO_2 at RT.

When t>0.3sec, the continuing build-up of ΔV_{TH} in Fig.5.16 indicates that new defects are created and the increased part of ΔV_{TH}, can be evaluated by

$$\Delta V_{THC}=\Delta V_{TH}-\Delta V_{TH}\,(\text{at } t=0.3\text{sec}), \tag{5.1}$$

which is the traditional NBTI.

5.5 IMPACT ON OPERATION VOLTAGE

In the previous session we have seen how we can separated the as-grown traps from the generated traps using Eq.5.1. The same sample as described in Fig.5.11 was used to evaluate the impact of V_{TH} instability on operation voltage

5.5.1 NBTI generation in HfO_2

ΔV_{TH} under different V_G is plotted in Figs.5. 17a, b & c for room temperature (RT), 100°C and 150°C, respectively.

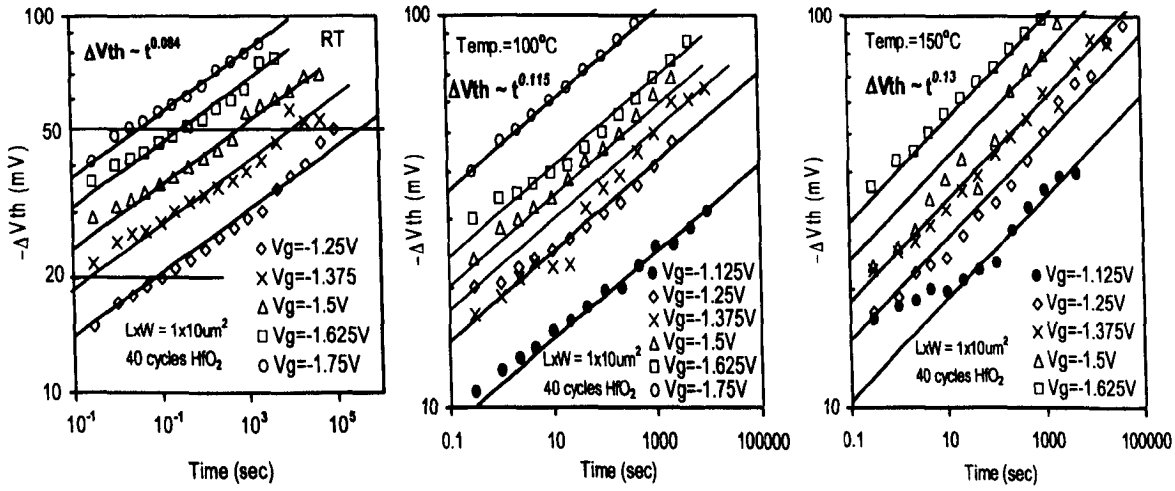


Fig.5.17 V_{TH} instability under different V_G for HfO_2. (a) Room temperature (RT). (b) 100°C. (c) 150°C

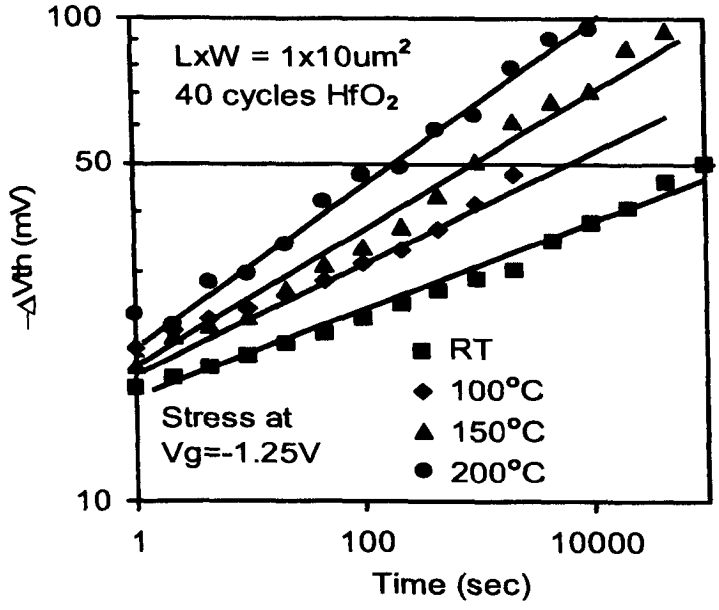


Fig.5.18 A comparison of V_{TH} instability at different temperatures for HfO_2.

Fig.5.18 compares ΔV_{TH} for different temperature at V_G = -1.25V for HfO_2 sample. If ΔV_{TH} is to be limited at 50mV, the 'lifetime' is plotted in Fig.5.19.

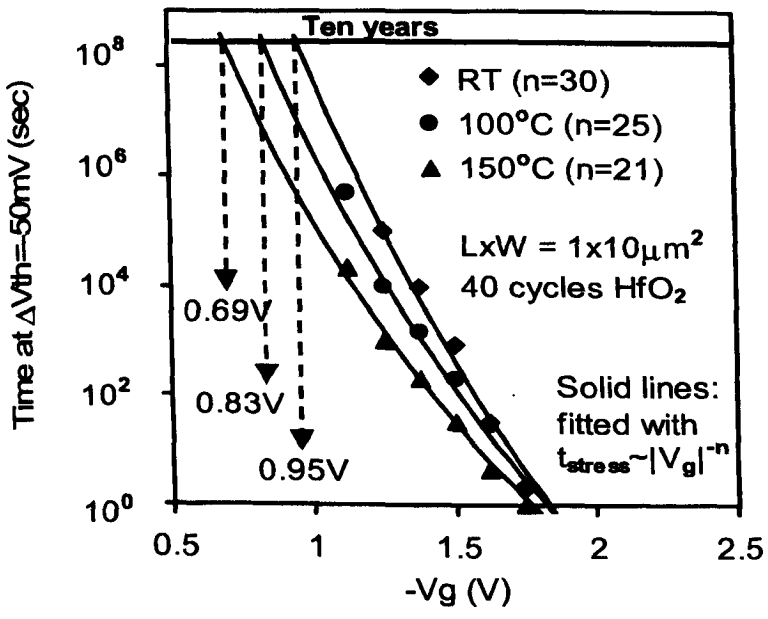


Fig.5.19 Estimation of the maximum operation voltage for $\Delta V_{TH} \leq 50mV$.

Based on the power law voltage acceleration **[Wu00]**, the maximum operation voltage for a 10 years' lifetime, V_{Gmax}, is estimated and the result is shown in Fig.5.20. As a comparison, the V_{Gmax} estimated for a 0.9nm EOT HfO_2 caused by breakdown reported at the IEDM-2005 **[Degra05]** is also given. The ΔV_{TH} clearly imposes a lower V_{Gmax} than the breakdown induced V_{Gmax} of 1.6V.

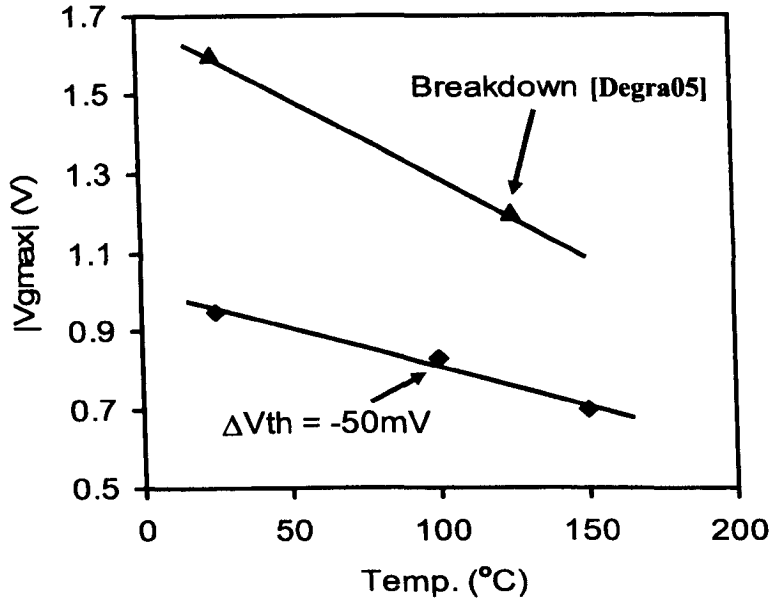


Fig.5.20 A comparison of the maximum operation voltage, V_{Gmax}, imposed by ΔV_{TH}=-50mV and breakdown [Degra05]

Since the V_{TH} instability of pMOSFETs can limit V_{Gmax}, it is important to improve our understanding of it. The ΔV_{TH} in Figs.5.17a-c appears following a power law with a power factor independent of V_G. This power factor is only 0.084 at RT and 0.13 at 150°C, which is much lower that the typical 0.2~0.3 reported for NBTI. The ΔV_{TH} includes contributions from both as-grown and generated defects. The instability induced by created defects alone as shown in previous section can be estimated by using Eq.5.1. The results are plotted in Figs 5.21a,b&c for RT, 100°C and 150°C, respectively.

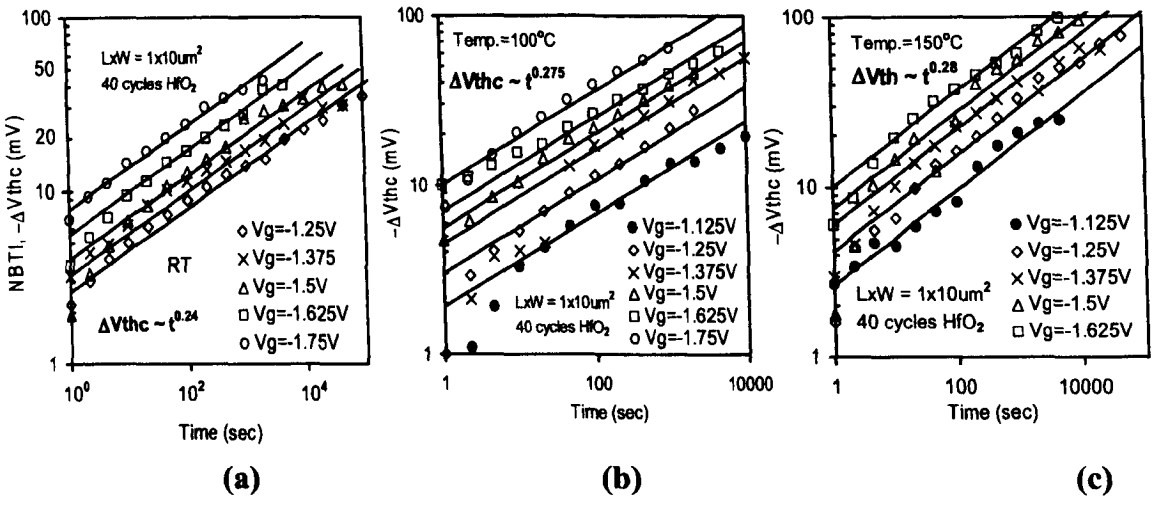


Fig.5.21 The created V_{TH} instability under different V_G for HfO_2 at (a) Room temperature (RT). (b) 100°C. (c) 150°C

After separating the total ΔV_{TH} into as-grown V_{TH} instability and NBTI, we can remove the effect of as-grown defects on NBTI kinetics. Figs.5.21a, b & c show that generation kinetics of NBTI (ΔV_{THC}) alone for HfO_2.The power factor is increased to 0.24~0.28 for HfO_2 agreeing with the expected value for NBTI.

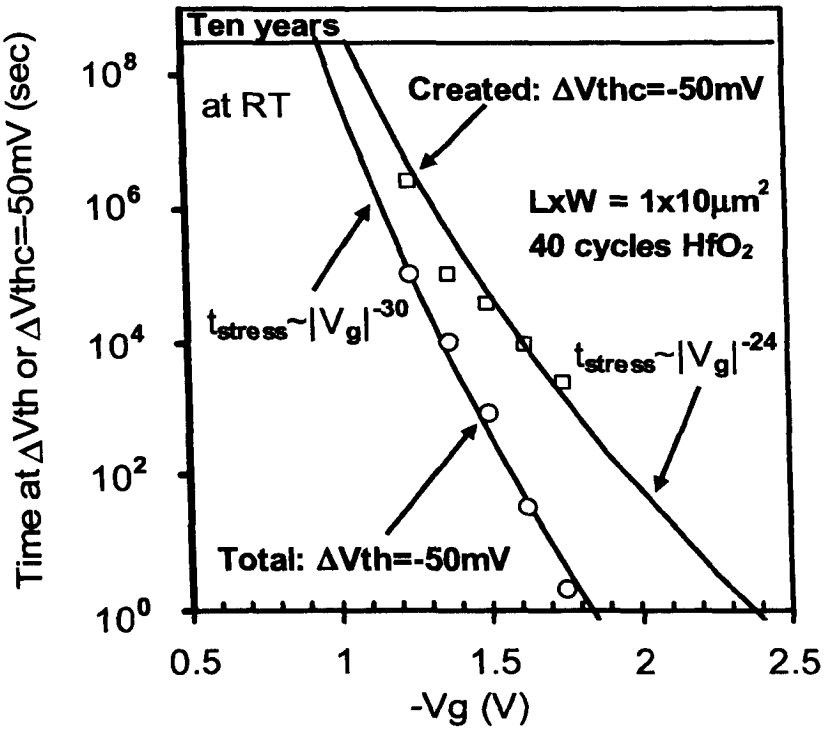


Fig.5.22 Estimation of the max. V_G for 10 years based on power law V_G acceleration.

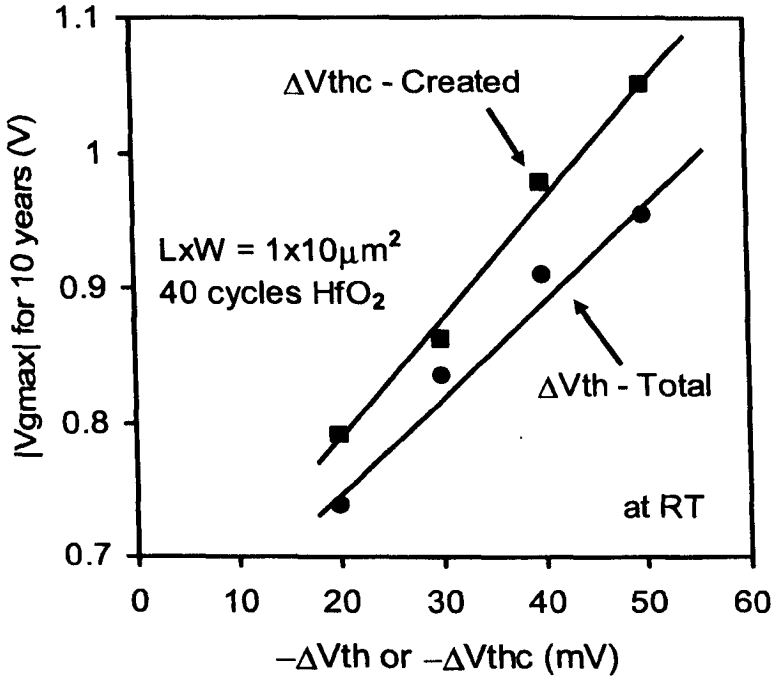


Fig.5.23 Dependence of V_{Gmax} on the allowed V_{TH} instability.

It is interesting to estimate the V_{Gmax} induced by the created defects alone. Fig.5.22 compares the lifetime for ΔV_{THC} = 50mV and ΔV_{TH} = 50mV. Fig.5.23 shows the V_{Gmax} dependence on the allowed ΔV_{TH} or ΔV_{THC}. If $\Delta V_{TH} \leq$ 20mV is required, $|V_{Gmax}|$ is only 0.75V, which is about the same as the threshold voltage

5.5.2 NBTI generation in Hf-silicate

As described in the previous section for HfO_2 on the separation of as-gown from the generated defects, the same experiment is carried out on Hf-silicate samples. ΔV_{TH} under different Vg is plotted in Fig.5.24 for room temperature (RT) as an example.

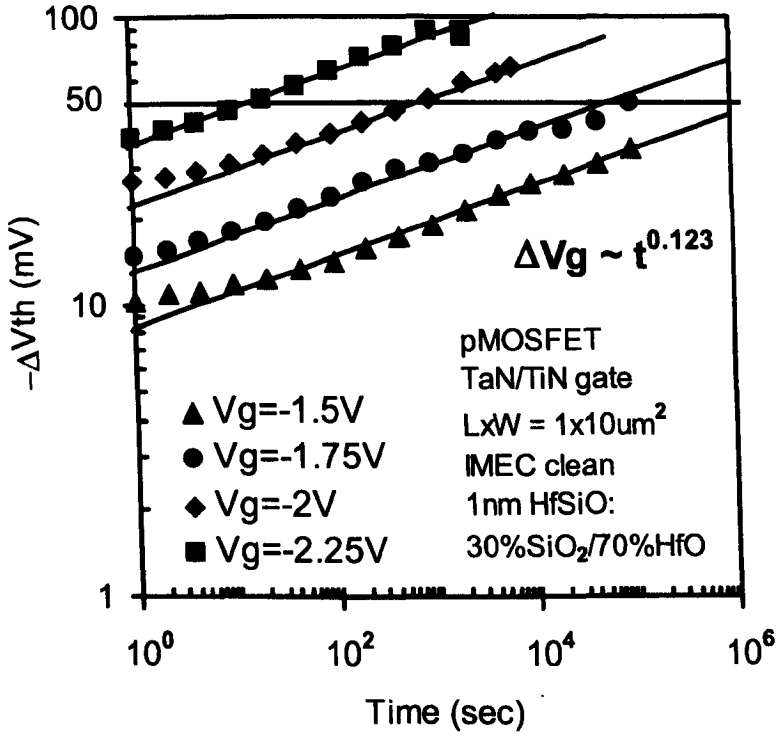


Fig.5.24 V_{TH} instability under different V_G for the Hf-silicate at RT.

Fig.5.25 compares the V_{Gmax} estimation, based on power law and exponential V_G acceleration for Hf-silicate. The results of HfO_2 are also given for comparison.

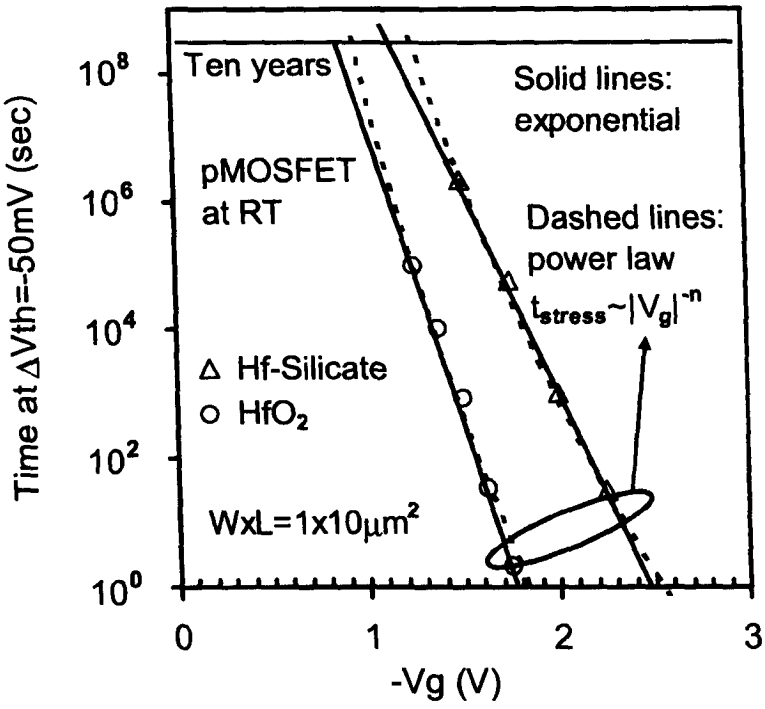


Fig.5.25 A comparison of V_{Gmax} estimation, based on power law and exponential V_G acceleration

It is shown that if an exponential V_G acceleration is used, the estimated V_{Gmax} is lower than that based on the power law V_G acceleration. For the Hf-Silicate, V_{Gmax} at RT is 1.25V, still lower than the breakdown induced V_{Gmax} shown in Fig.5.20. The power factor is only 0.12 at RT, far from the NBTI value expected.

After using Eq 5.1 the instability induced by created traps only is shown in Fig.5.26 for the Hf-silicate. The power factor is increased to 0.28 agreeing with expected value for NBTI. If ΔV_{TH} or ΔV_{THC} is to be limited at 50mV, the 'lifetime' is plotted in Fig.5.27 for the Hf-silicate samples. Based on the power law voltage acceleration **[Wu00]** the maximum V_G for the silicate sample for a 10 years lifetime, V_{Gmax}, is approximately 1.25V for $\Delta V_{TH} \leq 50$mV and 1.26V for $\Delta V_{THC} \leq 50$mV.

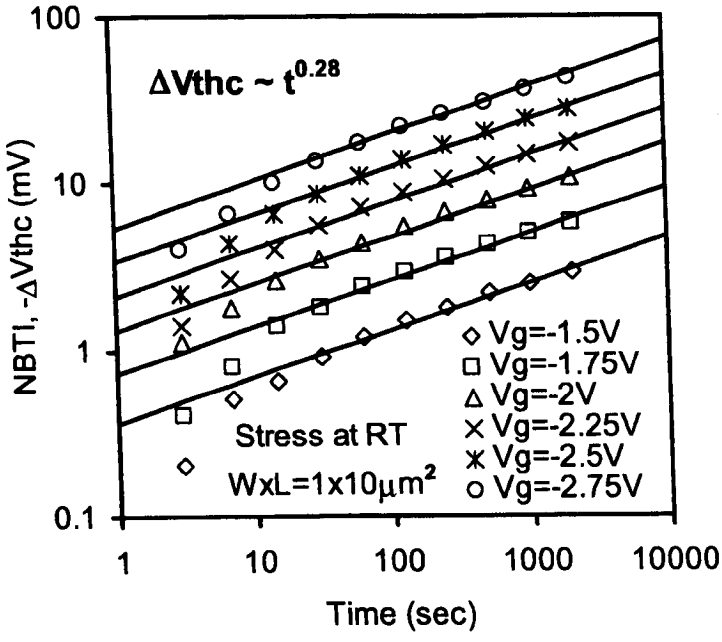


Fig.5.26 Generation kinetics of NBTI (ΔV_{THC}) under different V_G for 2nm Hf-silicates annealed in NH_3

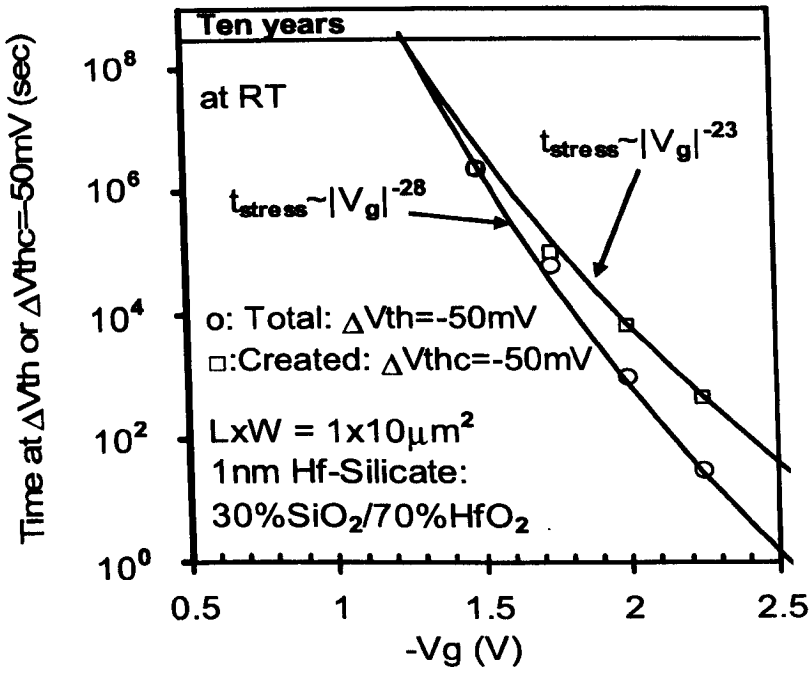


Fig.5.27 Estimation of the max operation voltage V_{Gmax} for 10 years for Hf-silicate stacks, based on power law V_G acceleration.

5.6 IMPACT OF NITRIDATION ON V_{TH} INSTABILITY: COMPARISON BETWEEN ALD VS MOCVD

Different samples are used and summarized in Fig.5.28. Hf-silicates were prepared by both MOCVD and ALD for pMOSFETs with TaN gate. Two nitridation techniques were used: DPN and NH_3 anneal. Reference samples without any anneal or with anneal in N_2 were also investigated. The channel length and width is 0.25~1μm and 10 μm, respectively.

Interf.	Thick. (nm)	%SiO_2 / %HfO_2	PDA	CET [A]	V_{th} [V]
IMEC	2	MOCVD (45/55)	N_2, 800 °C, 60 s	23.2	-0.59
IMEC	2	MOCVD (45/55)	O2, 800 °C, 15 s	23.5	-0.63
IMEC	2	MOCVD (45/55)	DPN (52.5 kJ) + 800 °C N_2	20.7	-0.6
IMEC	2	MOCVD (30/70)	NH_3, 800 °C,60s	21.7	-0.88
IMEC	2	ALD (45/55)	N_2, 800 °C, 60 s	24.8	-0.56
IMEC	2	ALD (45/55)	NH_3, 800 °C, 60 s	22.7	-0.52
IMEC	2	ALD (45/55)	DPN (52.5 kJ) + 800 °C, O_2	23	-0.56
IMEC	2	ALD (45/55)	No Nitridation	27.2	-0.76

Fig.5.28 Key data of ALD and MOCVD Hf-silicate stacks used in this work.

To assess the impact of nitridation on V_{TH} instability, the pulse I_D-V_G introduced in chapter 2 is used. Figs 5.29 & 5.30 clearly show that both DPN and NH_3 anneal substantially enhance the instability of pMOSFETs with Hf-silicates, prepared by either MOCVD or ALD. The instability here is measured as the gate voltage difference art fixed drain current between ramp-up and ramp down I-V curve. The samples with anneal in N_2 at the same temperature and the same anneal time do not show this high instability, which rules out that the instability is caused by a simple thermal effect.

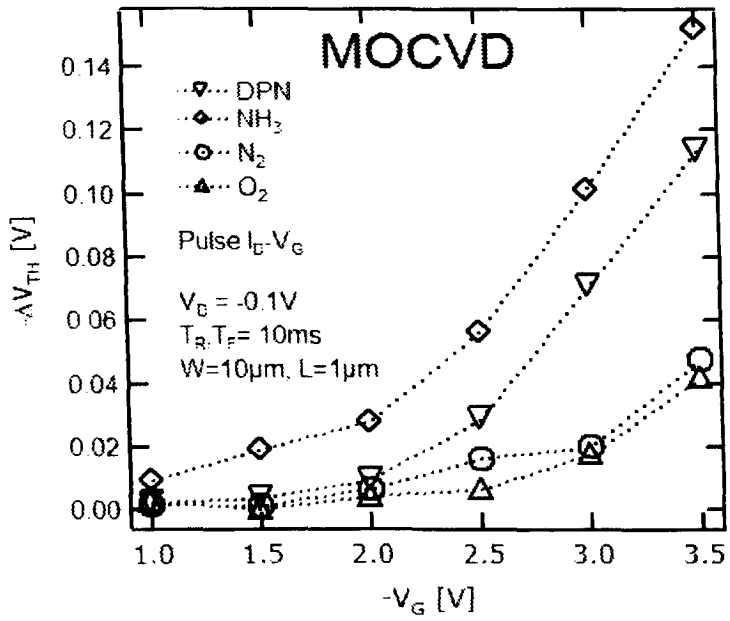


Fig.5.29 Impact of nitridation on the instability of pMOSFETs with 2nm MOCVD Hf-silicates with different processing condition.

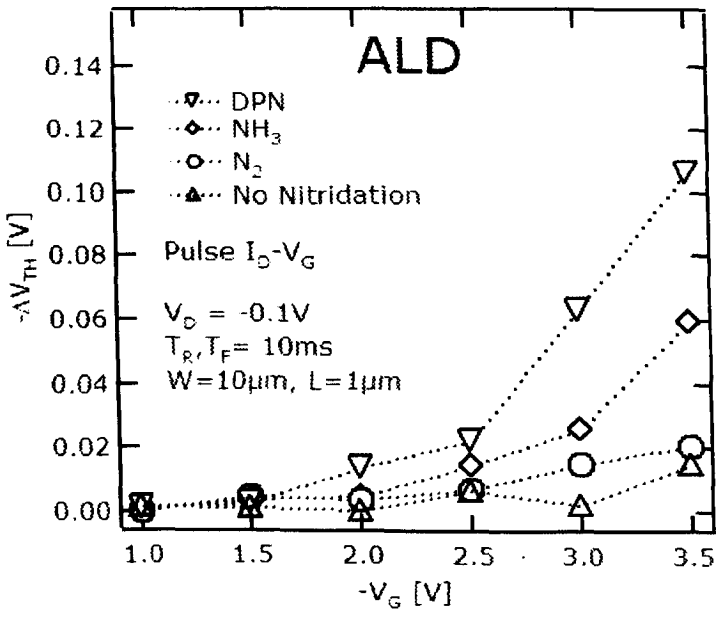


Fig.5.30 Impact of nitridation on the instability of pMOSFETs with 2nm ALD Hf-silicates with different processing conditions

The long term degradation of the threshold voltage, ΔV_{TH}, is shown in Fig.5.31 under different gate biases. It follows a power law well with a power factor insensitive to the gate bias. If $|\Delta V_{TH}|$ is to be limited at 50mV, the 'lifetime' is plotted in Fig.5.32.

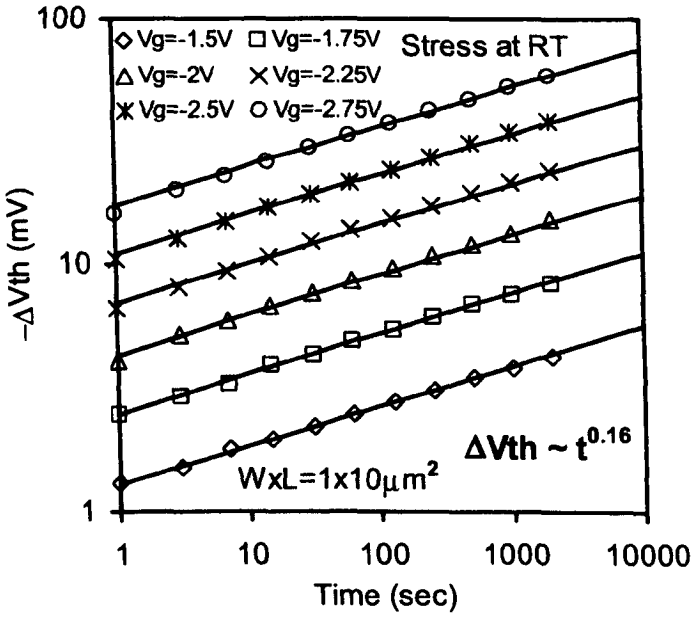


Fig.5.31 V_{TH} instability under different V_G for a 2nm MOCVD Hf-silicate, after annealing in NH_3 at 800°C for 60s.

Based on an exponential V_G acceleration, the maximum operation voltage for a 10 years' lifetime is estimated. The nitridation of Hf-silicates can reduce the operation voltage by 1V, as shown in Fig.5.32.

Figs.5.33 and 5.34 shows a comparison of the V_{TH} instability measured with pulsed I_D-V_G and NBTI for ALD and MOCVD samples respectively. For the case of NBTI the V_{TH} instability shown in the figures was measured after 1s stress at each voltages and the parameter used was for the pulsed I_D-V_G are the same as described in Figs.5.29 and 5.30.

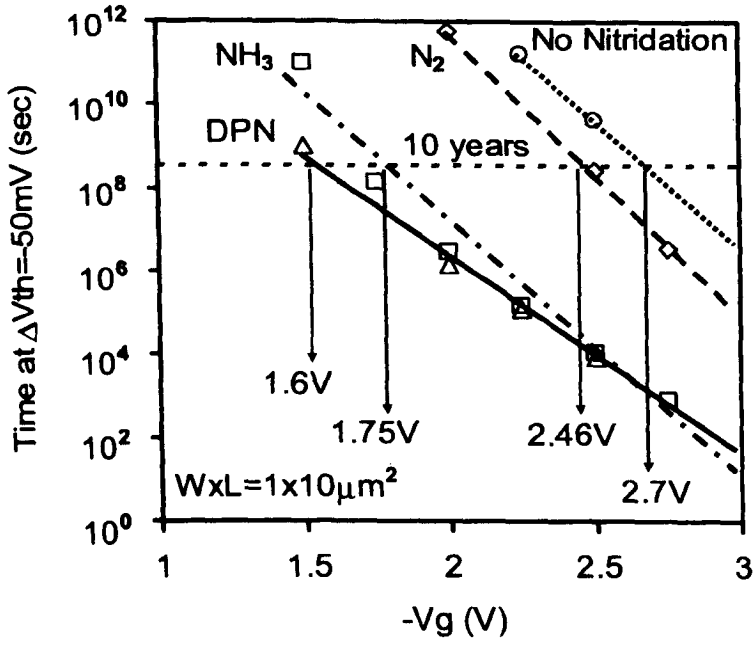


Fig.5.32 Impact of nitridation on the operation voltage for 10 years' lifetime with $|\Delta V_{TH}| \leq 50mV$ for 2nm Hf-silicates

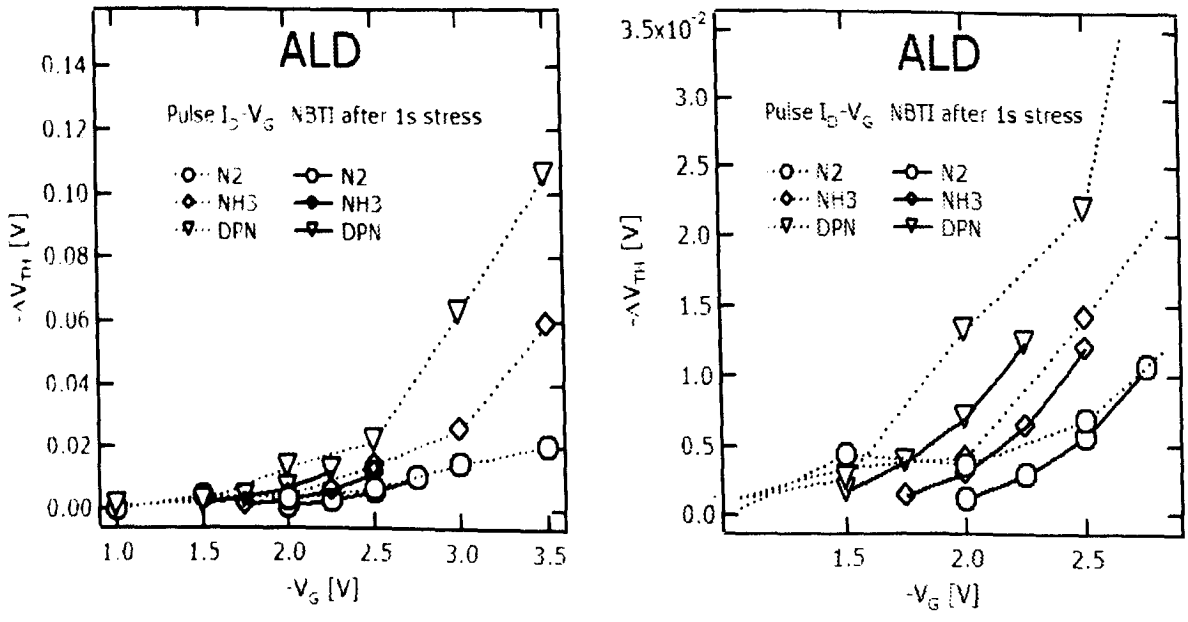


Fig.5.33 A comparison of the V_{TH} instability for ALD Hf-silicate stacks with different processing condition using pulsed I_D-V_G and NBTI generation after 1s of stress

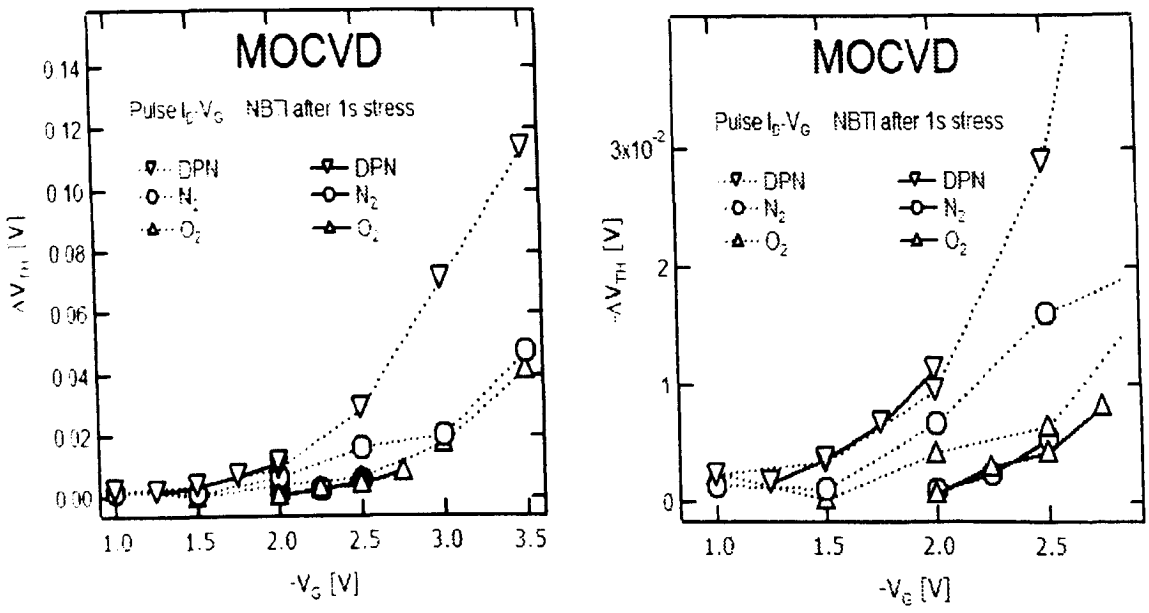


Fig.5.34 A comparison of the V_{TH} instability for MOCVD Hf-silicate stacks with different processing condition using pulsed I_D-V_G and NBTI generation after 1s of stress

Figs.5.35 and 5.36 shows a comparison of the V_{TH} instability measured with pulsed I_D-V_G and NBTI for ALD and MOCVD samples respectively. For the case of NBTI the V_{TH} instability shown in the figures was measured after 4095s stress at each voltages and the parameter used was for the pulsed I_D-V_G are the same as described in Figs.5.29 and 5.30.

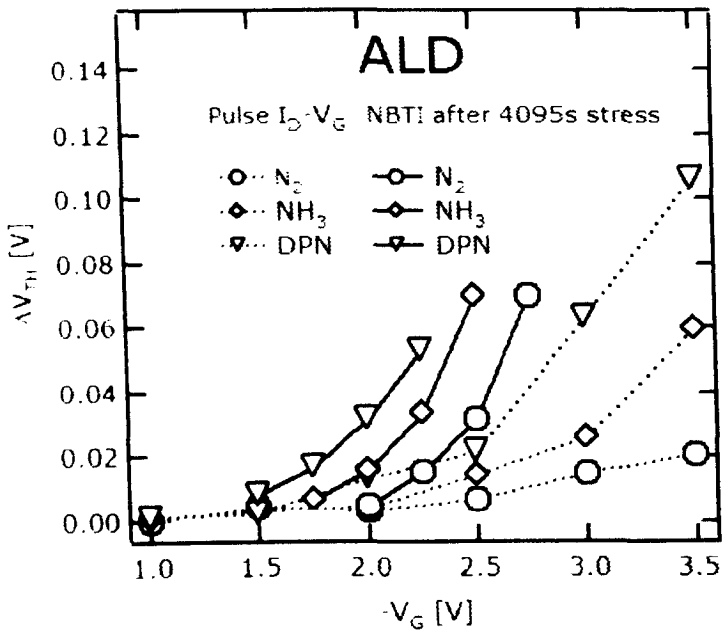


Fig.5.35 A comparison of the V_{TH} instability for ALD Hf-silicate stacks with different processing condition using pulsed I_D-V_G and NBTI generation after 4095s of stress.

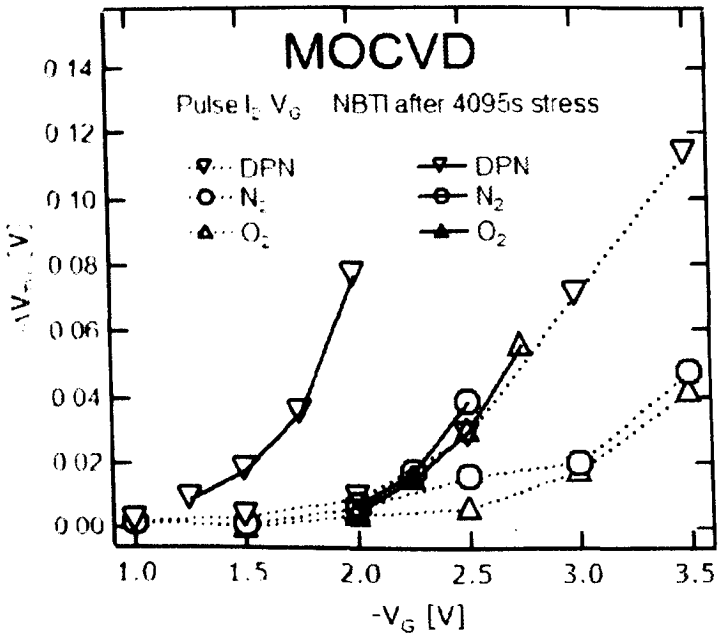


Fig.5.36 A comparison of the V_{TH} instability for MOCVD Hf-silicate stacks with different processing condition using pulsed I_D-V_G and NBTI generation after 4095s of stress

The ΔV_{TH} measured by the pulse I_D~V_G is compared with the traditional NBTI measurement in Figs. 5.33 & 5.35 for ALD samples and in Figs. 5.34 & 5.36 for MOCVD samples. For both sets of samples, it indeed shows that the V_{TH} measured by

NBTI after 1s stress matches the pulsed I_D-V_G (Figs 5.33 & 5.34). A generation of defects after 4095sec stress enhances the ΔV_{TH} for both ALD and MOCVD (Figs.5.35 & 5.36). This confirms that pulse I_D-V_G indeed scan the pre-existing traps rather that the generated. After a long time stress, the traps are created, confirming our early observation in section 5.4 and 5.5.

5.6.1 Impact of process conditions on V_{TH} instability

The V_{TH} instability in ALD pMOSFETs and nMOSFETs are elaborated by using different thickness variation of HfO_2 and Hf-silicate composition, combined with different process conditions. The device used is 10μm width and 0.25μm length. The experiment steps are as shown in Fig 5.37. Fig.5.38 shows the V_{TH} instability for both pMOSFETs and nMOSFETs.:

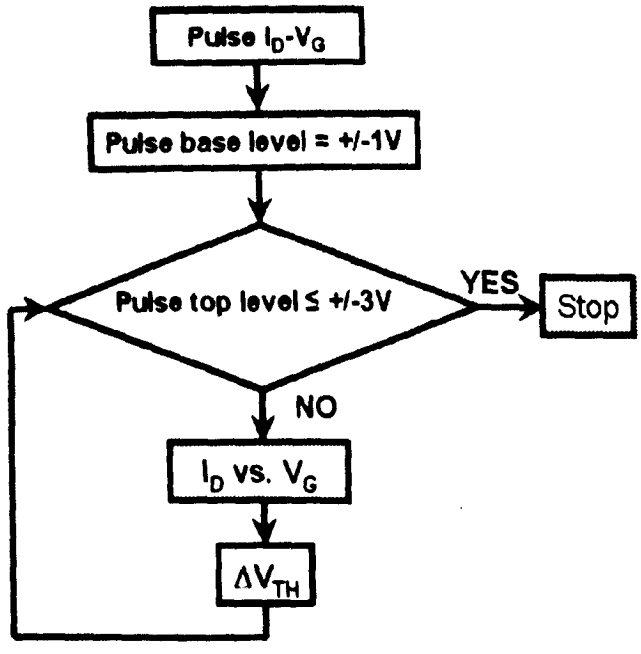


Fig 5.37 Schematic flow diagram for Pulse I_D-V_G to monitor the V_{TH} instability for nMOSFETs and pMOSFETs

A pulse is applied at the gate by keeping the base level fixed at 1V and increasing the top level from -0.1V to -3V. The ΔV_{TH} shown in Fig.5.38 is taken for both pMOSFETs and nMOSFETs when the top level of the pulse is at -2.5V. As a reference the V_{TH} instability of a SiON is shown.

From the comparisons in Fig.5.38, the main conclusions are:

1. NH_3 nitridation for pMOSFETs show high V_{TH} instability for 2nm HfO_2 than nMOSFET.
2. High percentage of Hf content (80%) combine with NH_3 nitridation rise to high V_{TH} instability for pMOSFETs compare to nMOSFETs.
3. Only low Hf content (50%) and plasma nitridation shows less or similar V_{TH} instability for pMOSFETs compared with nMOSFET.

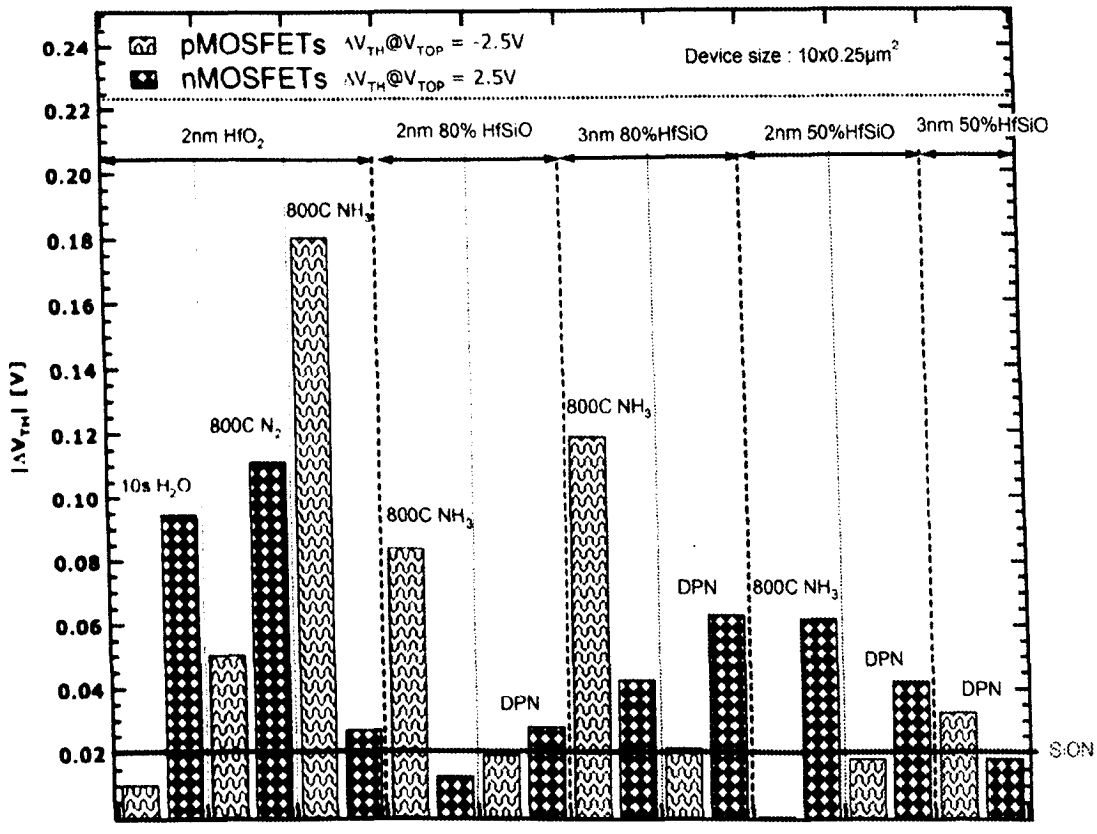


Fig.5.38 An investigation of the V_{TH} instability for different process and thickness of HfO2 and Hf content for nMOSFETs and pMOSFETs using pulse I_D-V_G. The V_{TH} shift is calculated when the top level of the pulse is 2.5V (nMOSFETs case) or -2.5V (pMOSFETs case)

5.7 CONCLUSIONS

In this chapter, we systematically investigated the V_{TH} instability and positive charge generation in pMOSFETs and compared it with that of nMOSFETs. This research shows that using metal gate material can induce more positive charge in nMOSFETs. Depending on the processing condition and the composition of the high-k layer, more V_{TH} instability can be found for pMOSFETs than for nMOSFETs. The main points are summarized below:

Section 5.2 investigates the impact of the gate material on the positive charge formation in HfO_2/SiO_2 nMOSFETs stacks. It is found that using metal gate, the positive charge generation after stressing under positive gate bias are significantly high than in poly-si gate. However, the positive charging is similar for the two types of gates, when stressed under negative gate bias. A strong gate bias polarity dependence was observed for positive charge formation in metal-gated samples, but not for poly-gated devices. These phenomena are explained by assuming that hydrogen release from the anode during stress dominates the positive charge formation and there are more hydrogenous species at the metal/dielectric interface. A large portion of the positive charges in HfO_2/SiO_2 stacks can be repeatedly neutralized and then recharged by alternating gate bias polarity, similar to the anomalous positive charges created in a single SiO_2 layer.

Section 5.3 investigates the threshold voltage, V_{TH}, instability of pMOSFETs with Hf-based dielectrics. It is found that substantial V_{TH} instability can occur in pMOSFETs. For a HfO_2/nitrided interfacial layer stack with an EOT of 1.13nm, a negative shift of V_{TH} up to 0.4V is observed. For the first time, it is found that the characters of V_{TH} instability of pMOSFETs are different from those of nMOSFETs in several aspects. The measured V_{TH} instability reduces significantly as the measurement

time increases for nMOSFETs, but not for pMOSFETs. As the magnitude of stress gate voltage increases, the instability saturates for nMOSFETs, while not for pMOSFETs. The carrier fluency can control the instability of nMOSFETs, but not that of pMOSFETs. It is speculated that the V_{TH} instability of pMOSFETs originates from donor-like as-grown defects, whose charging does not involve injected carriers, in contrast with the electron trapping in nMOSFETs. Using Hf-silicates is less effective in suppressing the instability of pMOSFETs and it can be higher than that of nMOSFETs for a sub-2nm nitrided layer. Process optimization will be needed to suppress the instability of pMOSFETs.

Section 5.4 separated as-grown traps from generated traps using Pulsed I_D-V_G and NBTI. It is found for the first time, that nitridation introduces two different types of defects, leading to two types of instabilities with distinctive characteristics for pMOSFETs, as-grown and generated instabilities. To match the power factor to the NBTI, separation between as-grown and generated traps is needed. Furthermore for an EOT around 1nm this work shows that V_{TH} instability of pMOSFETs is more severe than that of nMOSFETs.

Section 5.5 shows that the V_{TH} instability of pMOSFETs can become a potential limiting factor for the operation voltage. Also note that this work clearly shows that nitridation of Hf-based dielectrics significantly enhances the instability and substantially reduce the operation voltage of pMOSFETs.

Section 5.6 investigates the V_{TH} instability in ALD and MOCVD pMOSFETs. It is found that nitridation enhances the V_{TH} instability for both ALD and MOCVD. Further comparison between NBTI and as-grown instability shows that the pulsed I_D-V_G capture the pre-existing traps, which matches well with the NBTI data after 1s stress. New traps are created after long stress. An investigation on the V_{TH} instability with

different thickness and composition of ALD Hf-silicates and ALD HfO_2 with different post deposition anneals is also performed. The result shows that larger ΔV_{TH} is observed for 2nm HfO_2 and 2,3 nm 80% HfSiO for pMOSFETs with NH_3 nitridation, when compared with the ΔV_{TH} of nMOSFETs. Smaller or similar ΔV_{TH} is observed for 2,3 nm 50% HfSiO for pMOSFETs, when compared with the ΔV_{TH} of nMOSFETs.

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

6.1 CONCLUSION

This thesis addressed the pre-existing and generated defects in SiO_2/ high-k dielectrics. It was consequently investigate the characterization of as grown electrons traps, the generated electron traps, breakdown and finally the positive charge formation in nMOSFETs and pMOSFETs.

The main conclusions and contributions are summarized as follows:

6.1.1 Conclusions on the characterizations of as grown electrons traps

The issues addressed in chapter 3 were on the properties and dynamic behaviour, the location, capture cross sections, and trapping kinetics of as grown electron traps in HfO_2/SiO_2 and HfO_2 / HfSiO. The contributions are:

The traditional DC I_D~V_G substantially underestimates the trapping, because of detrapping during the measurement. The trap-assisted conduction, which is responsible for SILC and the thermally enhanced current, contributes little to trapping.

On the location, it is ruled out that the traps are piled up at the HfO_2/HfSiO interface. A uniform distribution throughout HfO_2 layer does not agree with the test data, either. The results presented in this project support that trapping is negligible in the region around 1.3~1.8nm near to one or both ends of the HfO_2 layer, when compared with the trapping in the central region of HfO_2.

On the capture cross sections of as-grown electron traps in HfO_2 layer and trapping kinetics. The results presented support the trapping model with two well separated discrete capture cross sections, rather than a continuous distribution, indicating the presence of two different types of as-grown electron traps. The extracted values are in the order of $10^{-14}cm^2$ and $10^{-16}cm^2$, respectively. For a 4nm ALD HfO_2 layer, the effective trap density is around $5.3\times10^{12}cm^{-2}$ and $1.1\times10^{13}cm^{-2}$ for the larger and smaller traps, respectively. These densities are typical values for 4nm HfO_2 layers prepared by ALD and MOCVD techniques, but trap density can vary substantially for HfO_2 fabricated by other techniques, such as PVD. In contrast, it is found that the capture cross section is insensitive to the fabrication technique.

6.1.2 Conclusions on the generated electron traps and breakdown

The issues addressed in chapter 4 were on the generation of traps in fresh and degraded device and a reliability characterization of different process. The contributions are:

Using amplitude sweep charge pumping the contribution from charging and discharging of bulk defect states becomes clearly evident and the charge per cycle strongly increases with increasing peak level and charging time. Furthermore a comparison of different gate electrode is studied with different channel length, the results shows that short channel shows low trap density independent of gate electrodes.

Using frequency sweep charge pumping we demonstrated that the HfO_2 bulk trap density indeed shows channel length dependence for fresh and degraded devices and the time-to-hard-breakdown is channel length dependent. The numeric discrepancy between the channel length dependence of trap density and breakdown is consistently

explained when trap generation and subsequent wear out are taken into account. Further we develop the applicability of frequency sweep charge pumping , where the pulse low timing ("discharging time") and high level timing ("charging time") is independently controlled to investigate the trap generation in SiO_2 and HfO_2. Using this so called Variable T_{charge} $T_{discharge}$ charge pumping (VT^2CP) technique we are able to separate the traps in the interfacial SiO_2 from the traps in the HfO_2. The separation of the scanned traps between the SiO2 (D_{SiO2}) and HfO2 (D_{HfO2}) was confirmed by using the wafer level variation technique.

VT^2CP was used to investigate the trap generation during stress, the results show that during degradation the increase of traps, both in the SiO_2 as well as in the HfO_2, follows a power law behavior as a function of time with an exponent ~0.32 and ~0.34 respectively independent of stress voltage. The voltage acceleration of creation of HfO_2 traps (-30) is nearly identical of the TDDB (-27). This technique shows that we can accurately detect degradation down to a much lower voltage than the dielectric breakdown measurement range and only one stress experiment combined with VT^2CP is sufficient to determine the degradation at a given voltage, while a TDDB test requires many measurements in order to construct an accurate distribution of failure times.

Using the VT^2CP the trap generation in SiO_2 and HfO_2 and the interface traps D_{IT} was investigated with different percentage combination of Hf and thickness, the results shows in one hand that plasma nitridation combine with high % Hf content and 2 nm HfO_2 with 10s water pulse without MG degas gives low D_{SiO2} and D_{HfO2} and in the other hand the lowest D_{IT} is obtained with 2nm HfO_2.

TDDB has proven to be a good reliability tools for high-k devices, it is shown that water pulse length in ALD has no effect on t_{BD} value, but the deposition uniformity improves with 10 s water pulse length. HfSiO with 50 % Hf shows no significant

leakage current increase before breakdown independent of nitridation technique and HfSiO with 80 % Hf shows leakage current increase before breakdown just as in pure HfO_2.that

Using TDDB a new phenomena of breakdown is shown "hard breakdown" The findings indicate that the large current jump ΔI is an inherent property of the breakdown process in dielectrics when combined with a metal gate. Finally it is concludes that when the poly-Si gate is replaced with a metal gate, TDDB lifetime is consistently limited by an abrupt large current increase of several hundred μA. The occurrence of those large ΔI is a potential limitation for the reliability of metal gate devices.

6.1.3 Conclusion on positive charges and V_{TH} instability

The issues addressed in chapter 5 were on the positive charge generation in pMOSFETs compared to nMOSFETs and a systematically investigation on the V_{TH} instability. This research shows that using metal gate material can induce more positive charge in nMOSFETs. Depending on the processing condition and the composition of the high-k layer, more V_{TH} instability can be found for pMOSFETs than for nMOSFETs. The main contributions are summarized below:

It is found that using metal gate, the positive charge generation after stressing under positive gate bias are significantly high than in poly-si gate in HfO_2/SiO_2 nMOSFETs sta. However, the positive charging is similar for the two types of gates, when stressed under negative gate bias. A strong gate bias polarity dependence was observed for positive charge formation in metal-gated samples, but not for poly-gated devices. These phenomena are explained by assuming that hydrogen release from the anode during stress dominates the positive charge formation and there are more hydrogenous species at the metal/dielectric interface. A large portion of the positive

charges in HfO_2/SiO_2 stacks can be repeatedly neutralized and then recharged by alternating gate bias polarity, similar to the anomalous positive charges created in a single SiO_2 layer.

It is found that substantial V_{TH} instability can occur in pMOSFETs. For a HfO_2/nitrided interfacial layer stack with an EOT of 1.13nm, a negative shift of V_{TH} up to 0.4V is observed. For the first time, it is found that the characters of V_{TH} instability of pMOSFETs are different from those of nMOSFETs in several aspects. The measured V_{TH} instability reduces significantly as the measurement time increases for nMOSFETs, but not for pMOSFETs. As the magnitude of stress gate voltage increases, the instability saturates for nMOSFETs, while not for pMOSFETs. The carrier fluency can control the instability of nMOSFETs, but not that of pMOSFETs. It is speculated that the V_{TH} instability of pMOSFETs originates from donor-like as-grown defects, whose charging does not involve injected carriers, in contrast with the electron trapping in nMOSFETs. Using Hf-silicates is less effective in suppressing the instability of pMOSFETs and it can be higher than that of nMOSFETs for a sub-2nm nitrided layer. Process optimization will be needed to suppress the instability of pMOSFETs.

Pulsed I_D-V_G and Negative Bias Temperature Instability "NBTI" were used to separate as-grown traps from generated traps. It is found for the first time, that nitridation introduces two different types of defects, leading to two types of instabilities with distinctive characteristics for pMOSFETs, as-grown and generated instabilities. To match the power factor to the NBTI, separation between as-grown and generated traps is needed. Furthermore for an EOT around 1nm this work shows that V_{TH} instability of pMOSFETs is more severe than that of nMOSFETs and can become a potential limiting factor for the operation voltage. Also note that this work clearly shows that nitridation

of Hf-based dielectrics significantly enhances the instability and substantially reduce the operation voltage of pMOSFETs.

Finally is shown that nitridation enhances the V_{TH} instability for both ALD and MOCVD. Further comparison between NBTI and as-grown instability shows that the pulsed I_D-V_G capture the pre-existing traps, which matches well with the NBTI data after 1s stress. New traps are created after long stress. An investigation on the V_{TH} instability with different thickness and composition of ALD Hf-silicates and ALD HfO_2 with different post deposition anneals is also performed. The result shows that larger ΔV_{TH} is observed for 2nm HfO_2 and 2,3 nm 80% HfSiO for pMOSFETs with NH_3 nitridation, when compared with the ΔV_{TH} of nMOSFETs. Smaller or similar ΔV_{TH} is observed for 2,3 nm 50% HfSiO for pMOSFETs, when compared with the ΔV_{TH} of nMOSFETs.

6.2 FUTURE WORK

In order to successfully integrate high-k materials into future CMOS technologies, strong improvements in the device performance (carrier mobility) and the charge trapping (V_T-instability) behaviour are required. The improvements may come from varying the deposition process, the deposition temperature and possibly the post-deposition treatment. Nitrogen and silicon incorporation into the high-k dielectric may improve the stability and performance of CMOS devices.

Despite the efforts in this project and many previous works, the understanding of the defect is limited at present. A lot of work remains to be carried out to achieve a full understanding of the defects in Hf-dielectric. A full understanding will lay the foundation for controlling the device instability of future MOSFETs. Suggestions for the future work are given below:

ELECTRON TRAPPING AND TRAP GENERATION

On electron trapping and trap generation, we showed in chapter 3 that, under positive gate bias, a large amount of electron traps could be created in the bulk of Hf-dielectrics. Future work should investigate the generation in the stack under negative gate bias. The present study is based on the VT^2CP technique, but the scanning depth has not been quantitatively known yet. It is hoped that a modeling of the results in the future can determine the scanning depth in the high-k layer. The present understanding on the properties of electron traps is still poor. For example, there is little information on their microscopic structure. A systematic comparison with the electron trap reported for SiO_2 is also missing.

The VT^2CP technique has been proven to be a good reliability tool when the breakdown-induced lifetime is assessed, compared with the TDDB method. Further

investigation can be carried out for the impact on trap generation by different gate materials and deposition techniques (i.e Fully Silicide , ALD TiN, PVD TiN, ALD TaN, PVD TaN), different process conditions, and different thickness of the interfacial and the high-k layers.

POSITIVE CHARGING AND HOLE TRAPPING

On the V_{TH} instability of pMOSFETs, it is shown that nitridation can have a large impact on the instability, when compared with nMOSFETs. Future work can probe this defect by using the newly developed VT^2CP technique and estimate its location. The questions to be answered include if it is located at the interface or in the high-k layer, whether there is a difference in the location for pre-existing and generated defects, and how this instability depends on the N incorporation in the dielectric. On the impact of nitrogen, different nitridation techniques can be adopted to change the nitrogen profile in the film. For example, it is expected that the nitridation by NH_3 anneal leads to a pile-up of nitrogen at the interface, while plasma nitridation will not. Research should be carried out on the correlation between the nitrogen distribution and the positive charging and how to optimize the nitrogen profile in the gate stack.

Work should also be carried out on the nature and properties of defects responsible for the positive charging. The relation between the observed positive charging and hole trapping needs exploring. Unlike electron traps, there is little information on the energy level on the positive charging defect at present. Like electron traps, the microscopic structure of hole traps is not known. Investigation should also be carried out on the relation between the built-in hole traps and the bias temperature instability and their impact on the lifetime of pMOSFETs

WORK ON EOT<1NM AND SUB-50NM CHANNEL LENGTH

Hf-based dielectric layers with EOT<1nm are actively investigated for 22nm node and beyond. EOT scalability of these films is simultaneously achieved by reducing the high-k thickness as well as optimizing the N-profile in the thin film. The hole traps depend on the nitrogen profile and the metal gate chemistry and deposition. Furthermore, since Hf-based compounds and metal gate are very sensitive to oxygen sources, the scalability and chemical stability has to be guaranteed for short channel lengths used in future CMOS technologies.

When the channel length is below 100nm, leakage current leads to an unreliable extraction of trap density if charge pumping is used. A good contribution will be to find out a model for correcting this leakage current. This will allow using the VT^2CP to probe either electron traps or hole traps for short channel length and investigate how these defects affect device lifetime. Such a systematic investigation on device instability will assist process optimization.

REFERENCES

[Agar02] A.K.Agarwal, C.C.Chao, R.H.Vogel, and M.H.White. "On the transient and steady-state transport of electrons and holes in the MNOS and MONOS devices". In IEDM-Technical Digest of the International Electron Device Meeting, pp 400-403,2002.

[Amer02] A. Amerasekera, C. Duvvury, W. Anderson, "ESD in silicon integrated circuits", Wiley, 2002

[Arre05] A.Arreghini, F.Driussi, D.Esseni, L.Selmi, M.J. van Duurenand R. van Schijk," New Charge Pumping model for the analysis of the spatial trap distribution in the nitride layer of SONOS devices", Microelectronic Enginerring, Vol.80, p. 333-336, 2005

[Bauza94] D.Bauza and G. Ghibaudo," Analytical study of the contribution of fast and slow oxide traps to the charge pumping current in MOS structures", Solid State Electronics, vol. 39, p.563,1996.

[Berg92] Bergonzoni. C., and Libera. G. D. "Physical Characterization of Hot-Electron-Induced MOSFET Degradation Through an Improved Approach to the Charge-Pumping Technique". *IEEE Trans.Electron Devices 39*, 8 (1992), 1895-1901.

[Bers04] G. Bersuker, J. H. Sim, C. D. Young, R. Choi, P. M. Zeitzoff, G. A. Brown, B. H., Lee, R. W. Murto," Effect of pre-existing defects on reliability assessment of high-K gate dielectrics", Microelectronics Rel., Vol 44, no 9-11 pp. 1509-1512, 2004.

[Bosch93] G. Van den Bosch, G. Groeseneken, and H.E. Maes. "On the Geometric Component of Charge-Pumping Current in MOSFETs". IEEE Electron Device Letters, Vol. 14, No. 3, pp. 107-109, 1993

[Brug69] BRUGLER, J. S., AND JESPERS, P. G. A."Charge Pumping in MOS Devices", *IEEE Trans.Electron Devices ED-16*, 3 (1969), 297-302

[Cart04] E. Cartier, A. Kerber, L. Pantisano, 'Charge Trapping in SiO_2/HfO_2 Dual Layer Gate Stacks', RC23080 (W0401-119) January 23, pp.1-6,2004 Electrical Engineering.

[Cell02] G. Cellere, L. Larcher, M.G. Valentini, A. Paccagnella, "Micro breakdown in small-area ultrathin gate oxides", IEEE Trans. Electron Devices, Volume 49, Issue 8, pp. 1367-1374, 2002

[ChangM06] M. H. Chang, J. F. Zhang, and W. D. Zhang, IEEE Trans. Electron Dev., 53, 1347 (2006).

[ChauR04] R. Chau, S. Datta, M. Doczy, B. Doyle, J. Kavalieros, and M. Metz", High-/spl kappa//metal-gate stack and its MOSFET characteristics", IEEE Electron Device Letters, Vol 25, Issue 6, pp. 408-410, 2004.

[ChenW92] Chen, W., Balasinski, A., Zhang B., and Ma, T.-P. "Hot-Carrier Effects on Interface-Trap Capture Cross Sections in MOSFET's as Studied by Charge Pumping". *IEEE Electron Device Lett. 13*, 4 (1992), 201-202.

[Chen86] I. C. Chen, S. Holland, K. K. Young, C. Chang, C. Hu, "Substrate hole current and oxide breakdown", Applied physics letters 49 (11), pp. 669-671 (1986)

[ChungS02] S.S. Chung, S.J. Chen, C.K. Yang, et al. "A Novel and Direct Determination of the Interface Traps in Sub-100nm CMOS Devices with Direct Tunneling Regime (12 ˜ 16 A) Gate Oxide". In Digest of IEEE VLSI Technology Symposium, pages 74-75, Hawaii, 2002

[Crupi04] F. Crupi, R. Degraeve, A. Kerber, D. H. Kwak, G. Groeseneken," Correlation between stress-induced leakage current (SILC) and the HfO_2 bulk trap density in a SiO_2/HfO_2 stack" in: Proc. 42nd IRPS (2004) 181.

[Degra98] R. Degraeve, G. Groeseneken, R. Bellens, J.L. Ogier, M. Depas, Ph. Roussel, H.E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown", IEEE Trans. Elec. Dev., vol. 45, No. 4, pp. 904-911, 1998.

[Degra99] R. Degraeve, G. Groeseneken, "Degradation and breakdown in thin oxide layers: mechanisms, models and reliability prediction", Microelectronics Reliability, vol. 39, no. 10, pp. 1445-1460, 1999

[Degra01] R. Degraeve, B. Kaczer, F. Schuler, M. Lorenzini, D. Wellekens, P. Hendrickx, J. Van Houdt, L. Haspeslagh, G. Tempel, G. Groeseneken, "Statistical model for stress-induced leakage current and pre-breakdown current jumps in ultra-thin oxide layers", IEDM Technical Digest, pp. 621-624, 2001

[Degra01A] R. Degraeve, B. Kaczer, A. De Keersgieter, G. Groeseneken, "Relation between breakdown mode and breakdown location in short channel NMOSFETs and its impact on reliability specifications", Proceedings IRPS, pp. 360-366, 2001

[Degra03] R. Degraeve, A. Kerber, Ph.J. Roussel, E. Cartier, T. Kauerauf, L. Pantisano, G. Groeseneken, "Effect of bulk trap density on HfO_2 reliability and yield", IEDM Technical Digest, pp.935-938, 2003

[Degra04] R. Degraeve, F. Crupi, D. H. Kwak, G. Groeseneken," On the defect generation and low voltage extrapolation of Q/sub BD/ in SiO/sub 2//HfO/sub 2/ stacks", in: VLSI Tech. Symp., pp. 140-141, 2004

[Degra04A] R.Degraeve, F.Schuler, B. Kaczer, M. Lorenzini, D. Wellekens, P.Hendrickx, M van Duuren , G.J.M. Dormans, J.Van Houdt, L.

Haspeslagh, G. Groeseneken, G. Tempel, " Analytical percolation model for predicting anomalous charge loss in flash memories", IEEE Trans.Electron Dev.,vol.51, no.9, pp. 1392-1400,2004.

[Degra05] R. Degraeve, T. Kauerauf, M. Cho, M. Zahid, L-Å. Ragnarsson, D.P. Brunco, B. Kaczer, Ph. Roussel, S De Gendt, G. Groeseneken, "Degradation and breakdown of 0.9 nm EOT SiO2/ALD HfO2/metal gate stacks under positive Constant Voltage Stress", IEDM Technical Digest, pp. 419-422, 2005

[Degra05A] R. Degraeve, B. Govoreanu, B. Kaczer, J. Van Houdt, G. Groeseneken, "Measurement and statistical analysis of single trap current-voltage characteristics in ultrathin SiON", Proceedings IRPS, pp. 360-365, 2005

[DiMar93] D.J. DiMaria, E. Cartier, D. Arnold, "Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon", Journal of Applied Physics, Volume 73, Number 7, pp. 3367-3384, 1993

[DiMa97] D.J. DiMaria, J.H. Stathis, "Explanation for the oxide thickness dependence of breakdown characteristics of metal-oxide-semiconductor structures", Applied physics letters 70 (20), pp. 2708-2710 (1997

[FanY02] Y. Y. Fan, R. E. Nieh, J. C. lee, G. Lucovsky, G. A. Brown, L. F. Register, and S. K. Banerjee, "Voltage- and temperature-dependent gate capacitance and current model: application to ZrO/sub 2/ n-channel MOS capacitor", IEEE Trans. Electron Dev., Vol.49, Issue.11, pp 1969-1978 , 2002.

[Felnh05] D. Felnhofer, E. P. Gusev, P. Jamison, and D. A. Buchanan, "Charge trapping and detrapping in HfO2 high-k MOS capacitors using internal photoemission," Microelectronic Eng., vol.80, pp.58-61, 2005.

[Green01] M.L. Green, E.P. Gusev, R. Degraeve, E.L. Garfunkel, "Ultrathin (<4 nm) SiO2 and Si–O–N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits", Journal of Applied Physics, Volume 90, Number 5, pp. 2057-2121, 2001

[Goetz70] Goetzberger. A., and Nicollian. E. H. U.S. Patent 3,523,190. filed October 17, 1968; issued August 4, 1970.

[Groes84] G. Groeseneken, H.E. Maes, N. Beltran, and R.F. De Keersmaecker. "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors".IEEE Transaction on Electron Devices, Vol. 31, No. 1, pp. 42-53, 1984

[Groes04] G. Groeseneken, L. Pantisano, L-A. Ragnarsson, R. Degraeve, M. Houssa, T. Kauerauf, P. Roussel, S. De Gendt, and M. Heyns, "Achievements and challenges for the electrical performance of MOSFET's with high-k gate dielectrics," in Proc. 11th Int. Symp. Physical and Failure Analysis of IC, 2004, pp.147-155.

[Govo04] B. Govoreanu, P. Blomme, K. Henson, J. V. Houdt, and K. D. Meyer," An effective model for analysing tunneling gate leakage currents through ultrathin oxides and high-k gate stacks from Si inversion layers", Solid-State Electrons, Vol.48, Iss.4, pp 617-625, 2004.

[Gus04] E.P. Gusev ,C. D_Emic, S. Zafar, A. Kumar,"Charge trapping and detrapping in HfO2 high-k gate stacks', IBM Semiconductor Research and Development Center (SRDC), IBM Thomas J. Watson Research Center,Yorktown Heights, NY 10598, USA, Microelectronic Engineering 72 (2004) 273-277

[Heh06] D. Heh, C D. Young, G. A. Brown, P. Y. Hung, A. Diebold, G. Bersuker, E. M. Vogel, and J. B. Bernstein, "Spatial distributions of trapping centers in HfO2/SiO2 gate stacks", Appl. Phys. Lett., vol.88, 152907, 2006.

[Hens04] K. Henson, R.J.P. Lander, M. Demand, C.J.J. Dachs, B. Kaczer, W. Deweerd, T. Schram, Z. Tokei, J.C. Hooker, F.N. Cubaynes, S. Beckx, W. Boullart, B. Coenegrachts, J. Vertommen, O. Richard, H. Bender, W. Vandervorst, M. Kaiser J-l. Everaert. M. Jurczak and S. Biesemans, "45nm nMOSFET with metal gate on thin SiON driving 1150μA/μm and off-state of 10na/μm", IEDM Tech. Dig., 2004, pp. 851-854.

[Here88] Heremans, P., Bellens, R., Groesenken, G., And Maes, H. E. "Consistent Model for the Hot-Carrier Degradation in n-Channel and p-Channel MOSFET's". *IEEE Trans.Electron Devices ED-35*, 12 (1988), 2194-2209.

[Here89] Heremans, P., Witters, J., Groeseneken, G., And Maes, H. E. "Analysis of the Charge Pumping Technique and Its Application for the Evaluation of MOSFET Degradation". *IEEE Trans.Electron Devices 36*, 7 (1989), 1318-1335.

[Hick95] Ian Hickman,"Oscilloscope 4th Edition" Butterworth-Heinemann, pp 138-143 ,.1995

[HouY03] Y. T. Hou, M. F. Li, H. Y. Yu, and D. L. Kwong," Modeling of tunneling currents through HfO_2 and HfO_2/Al_2O_3 gate stacks", IEEE Electron Device Letters Vol. 24, no. 2, pp. 96-98. Feb. 2003

[Houd90] Houdt, J. V., Heremans, P., Witters, J. S., Groeseneken, G., And Maes, H. E."Study of the enhanced hot-electron injection in split-gate transistor structures". In Proc: *ESSDERC* (Nottingham (U.K.), 1990), Adam Hilger, pp. 261-264.

[Houssa01] M. Houssa, A. Stesmans, and M. M. Heyns, "Model for the trap-assisted tunnelling current through very thin SiO_2/ZrO_2 gate dielectric stacks ", Semiconductor Sciences and Technology, Vol.16, pp. 427-432, 2001.

[Houssa05] M. Houssa, G. Pourtois, M. M. Heyns, and A. Stesmans," Defect generation in high κ gate dielectric stacks under electrical stress: the impact of hydrogen", J. Phys.: Condens. Matter, **17**, S2075-S2088, 2005.

[Ielmi02] D. Ielmini, A.S. Spinelli, A.L. Lacaita, A. Modelli, " A statistical model for SILC in flash memories", IEEE Trans. Electron Devices, Volume 49, Issue 11, pp. 1955-1961, 2002

[KangA03] A. Y. Kang, P. M. Lenahan, J. F. Conley," Electron spin resonance observation of trapped electron centers in atomic-layer-deposited hafnium oxide on Si", Appl. Phys. Lett., Vol.83, Issue 16, pp. 3407-3409, 2003.

[Kacz00] B. Kaczer, R. Degraeve, N. Pangon, G. Groeseneken, "The influence of elevated temperature on degradation and lifetime prediction of thin silicon-dioxide films", *IEEE Trans. Elec. Dev.*, vol. 47, No. 7, pp. 1514-1521, 2000.

[Kacz04] B. Kaczer, R. Degraeve, R. O'Connor, P. Roussel, G. Groeseneken, "Implications of progressive wear-out for lifetime extrapolation of ultra-thin (EOT ~ 1 nm) SiON films", IEDM Technical Digest, pp. 713-716, 2004

[Kau02] T. Kauerauf, R. Degraeve, E. Cartier, B. Govoreanu, P. Blomme, B. Kaczer, L. Pantisano, A. Kerber, G. Groeseneken, "Towards understanding degradation and breakdown of SiO2 /high-k stacks", IEDM Technical Digest, pp. 521-524, 2002

[Kau05] T. Kauerauf, R. Degraeve, F. Crupi, B. Kaczer, G. Groeseneken, H. Maes, "Trap generation and progressive wearout in thin HfSiON", Proceedings IRPS, pp. 45-49, 2005

[Kaur05A] T. Kauerauf, R. Degraeve, M. B. Zahid, M. Cho, B. Kaczer, Ph. Roussel, G. Groeseneken, H. Maes, S. de Gendt, "Abrupt breakdown in dielectric/metal gate stacks: a potential reliability limitation?", IEEE Elec. Dev. Lett., vol. 26, No. 10, pp. 773-775, 2005

[Ker02] A. Kerber, E. Cartier, R. Degraeve, L. Pantisano, P. J. Roussel, and G. Groeseneken, "Strong correlation between dielectric reliability and charge trapping in SiO_2/Al_2O_3 gate stacks with TiN electrodes," in Symp. VLSI Tech. Dig., 2002, p. 76. VLSI Technology Symposium, pp.76-77,2002.

[Ker03] A.Kerber,E.Cartier,L.Pantisano et al.," Origin of the threshold voltage instability in Si02iHf02 dual layer gate dielectrics ",IEEE EDL, Vol.24,No.2,pp.87-89,2003.

[Ker03] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Characterization of the V_T-instability in SiO_2/HfO_2 gate dielectrics," in Proc. 41st Anuu. Int. Reliability Physics Symp., 2003, pp.41-45.

[Ker03A] Andreas Kerber, Methodology for Electrical Characterization of MOS Devices with Alternative Gate Dielectrics, October 2003, pp 50-55

[Ker03B] A. Kerber a, E. Cartier , L. Pantisano , M.Rosmeulen,R. Degraeve ,T.Kauerauf, G. Groeseneken ,H.E. Maes , U. Schwalke, " Characterization of the V_T-instability in SiO_2/HfO_2 gate dielectrics" IEEE 41st Annual International Reliability, Dallas, Texas, 2003

[Ker03B] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, in: *Proc. of Int. Reliability Phys. Symp.*, 2003, pp.41-45.

[Ker04] A. Kerber a, E. Cartier , L. Pantisano , R. Degraeve , G. Groeseneken ,H.E.Maes , U. Schwalke, " Charge trapping in SiO2/HfO2 gate dielectrics: Comparison between charge-pumping and pulsed ID-VG", Microelectronic Engineering 72 (2004) 267-272

[Lerou04] C. Leroux, J. Mitard, G. Ghibaudo, X. Garros, G. Reimbold, B. Guillaumot, F. Martin," Characterization and modeling of hysteresis phenomena in high K dielectrics ", in IEDM Tech Dig., 2004, pp.737-740.

[Ma05] T. P. Ma, H. M. Bu, X. W. Wang, L. Y. Song, W. He, M. Wang, H. -H. Tseng, P. J. Tobin, "Special reliability features for Hf-based high-k gate dielectrics," IEEE Trans. Device and Materials Reliability, vol.5, pp.36-44, 2005.

[Macfa00] P. J. Macfarlane and R. E. Stahlbush, Appl.Phys.Lett.," Near interface H^+ trapping and its influence on H^+ transport in hydrogenated Unibond buried oxides ",Vol.77, Issue 19, pp. 3081-3083, 2000.

[Mat95] R.E. Matick, "Transmission lines for digital and communication networks : an introduction to transmission lines, high-frequency and high-speed pulse characteristics and applications", 1995

[Maes82] Maes, H. E., And Groeseneken, G."Determination of spatial surface state density distribution in MOS and SIMOS transistors after channel hot electron injection". *Electron.Lett. 18*, 9 (1982), 372-374.

[Ma05] T. P. Ma, H. M. Bu, X. W. Wang, L. Y. Song, W. He, M. Wang, H. -H. Tseng, P. J. Tobin, "Special reliability features for Hf-based high-k gate dielectrics," *IEEE Trans. Device and Materials Reliability,* vol.5, pp.36-44, 2005.

[Mahn88] Mahnkopf, R., Przyrembel, G., And Wagemann, H. G."A new method for the determination of the spatial distribution of hot carrier damage". In Proc: *ESSDERC* (Montpellier (France), 1988), pp. 775-778.

[Muda00] S. Mudanai, Y. Y Fan, Q. Ouyang, A. F. Tasch, and S. K. Banerjee," Modeling of direct tunneling current through gate dielectric stacks", IEEE Trans. Electron Dev., Vol.47, Issue 10, 1851-1857, 2000.

[Ncona88] Ncona, M. G., Saks, N. S., And Mccarthy, D."Lateral Distribution of Hot-Carrier-Induced Interface Traps in MOSFET's". *IEEE Trans.Electron Devices ED-35*, 12 (1988), 2221-2228.

[Nico82] E.H. Nicollian and J. R. Brews."MOS (Metal Oxide Semiconductor) Physics and Technology". Wiley New York (N. Y.),1982

[Nig98] T. Nigam, R. Degraeve, G. Groeseneken, M.M. Heyns, H.E. Maes, "Constant current charge-to-breakdown: still a valid tool to study the reliability of MOS structures?", Proceedings *IRPS*, pp. 62-69, *1998*.

[Nigam99] T. Nigam, R. Degraeve, G. Groeseneken, M.M. Heyns, H.E. Maes, "A fast and simple methodology for lifetime prediction of ultra-thin oxides", Proceedings IRPS, pp. 381-388, 1999

[Paulm04] F. Palumbo, S. Lombardo, J.H. Stathis, V. Narayanan, F.R. McFeely, J.J. Yurkas, "Degradation of ultra-thin oxides with tungsten gates under high voltage: wearout and breakdown transient", Proc. IRPS 2004, pp. 122-125.

[Poor84] Poorter, T., And Zoestbergen, P."Hot carrier effects in MOS transistors". In Proc: *Int.Electron Devices Meeting* (1984), pp. 100-103.

[Oka94] K. Okada, S. Kawasaki, and Y. Hirofuji, "New experimental findings on stresss induced leakage current of ultra thin silicon dioxides", *Ext. Abst. of the 1994 SSDM*, p. 565, 1994

[Olivo88] P. Olivo, T.N. Nguyen, B. Ricco, "High-Field-Induced Degradation in Ultra-Thin SiO2 Films", IEEE Trans. Electron. Devices, vol. 35, pp. 2259-2267, 1988.

[Paul94] R.E. Paulsen and M.H. White," Theory and application of charge pumping for the characterizationof Si-SiO2 interface and near-interface oxide traps", IEEE Trans. on Elect. Devices, vol. 41, p. 1213-1216, 1994

[PeyK04] K.L. Pey, R. Ranjan, C.H. Tung, L.J. Tang, W.H. Lin, M.K. Radhakrishnan, "Gate dielectric degradation mechanism associated with DBIE evolution", Proceedings IRPS, pp. 117-121, 2004

[Rag03] L.-Å. Ragnarsson, L. Pantisano, V. Kaushik, S.-I. Saito, Y. Shimamoto, S. De Gendt, and M. Heyns," The impact of sub monolayers of HfO_2 on the device performance of high-κ based transistors", in *IEDM Tech. Dig.*, 2003, pp.87-90.

[Reim05] G. Reimbold, J. Mitard, M. Casse, X. Carros, C. Leroux, L. Thevenod, F. Martin, in Proc. Silicon Nitride,"Silicon Dioxide thin Insulating Films, and Other Emerging Dielectrics VIII", 2005, pp.437-455.

[RohY93]. Roh. Y, L. Trombetta, and J. Stathis," New model of a common origin for trapped holes and anomalous positive charge in MOS capacitors", Microelectronic Eng. **22**, pp.227-230 (1993).

[Ross94] Barry Ross, "Hands-On Guide to Oscilloscopes" McGraw-Hill Book Company, pp 167-168, 1994

[RoyM03] M. L. Roy, E. Lheurette, O. Vanbesien, and D. Lippens," Wave-mechanical calculations of leakage current through stacked dielectrics for nanotransistor metal-oxide-semiconductor design", J. Appl. Phys., Vol.93, Issue 5, pp. 2966-2971, 2003

[Sak98] T. Sakura, H. Utsunomiya, Y. Kamakura, K. Taniguchi, "A detailed study of soft- and pre-soft-breakdowns in small geometry MOS structures", *IEDM Tech. Dig.*, pp. 183-186, 1998.

[Schr05] T. Schram, L-A Ragnarsson, G. Lujan, W. Deweerd, J. Chen, W. Tsai, K. Henson, R. J. P. Lander, J. C. Hooker, J. Vertommen, K. De Meyer, S. De Gendt, and M. Heyns, "Performance improvement of self-aligned HfO_2/TaN and SiON/TaN nMOS transistors," Microelectronics Reliability, vol.45, pp.779-782, 2005.

[Shaw89] Shaw, J.-J., And Wu, K."Determination of spatial distribution of interface states on submicron, lightly doped drain transistors by charge pumping measurement". In Proc: *Int.Electron Devices Meeting* (1989), pp. 83-86.

[Shan03] A. Shanware, M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, J. McPherson, L. Colombo, G. A. Brown, C. H. Lee, Y. Kim, M. Gardner, and R. W. Murto," Characterization and comparison of the charge trapping in HfSiON and HfO_2 gate dielectrics" in IEDM Tech Dig., 2003, pp.939-942.

[ShenC04] C. Shen, M. F. Li, X. P. Wang, H. Y. Yu, Y. P. Feng, A. T. L. Lim, Y. C. Yeo, D. S. H. Chan, and D. L. Kwong,"Negative U traps in HfO gate dielectrics and frequency dependence of dynamic BTI in MOSFETs", in IEDM Tech Dig., 2004, pp.733-736.

[Sim05] J. H. Sim, S. C. Song, P. D. Kirsch, C. D. Young, R. Choi, D. L. Kwong, B. H. Lee, and G. Bersuker, "Effects of ALD HfO2 thickness on charge trapping and mobility," Microelectronic Eng., vol.80, pp.218-221, 2005.

[Stah93] R. E. Stahlbush, A. H. Edwards, D. L. Griscom, and B. J. Mrstik," Post-irradiation cracking of H_2 and formation of interface states in irradiated metal-oxide-semiconductor field-effect transistors", J. Appl. Phys., **73**, 658 (1993).

[Stath97] J.H. Stathis, "Quantitative model of the thickness dependence of breakdown in ultrathin oxides", Microelectronics Enigeering, vol. 36, no. 1-4, pp.325-328, 1997

[Stes00] A. Stesmans, "Dissociation kinetics of hydrogen-passivated Pb defects at the (111)Si/SiO2 interface", Phys. Rev. B, 61, 8393 (2000).

[Sue04] J.S. Suehle, B. Zhu, Y. Che, J.B. Bernstein, "Acceleration factors and mechanistic study of progressive breakdown in small area ultra-thin gate oxides", Proceedings IRPS, pp. 95-101, 2004

[Sue04] J.S. Suehle, B. Zhu, Y. Che, J.B. Bernstein, "Acceleration factors and mechanistic study of progressive breakdown in small area ultra-thin gate oxides", Proceedings IRPS, pp. 95-101, 2004

[Sune90] J. Suñé, I. Placencia, N. Barniol, E. Farrés, F. Martín, X. Aymerich, "On the breakdown statistics of very thin SiO2 films", Thin solid films, vol. 185, pp. 347-362, 1990

[Sze81] S. M. Sze, Physics of Semiconductor Devices, 2nd Edition, New York: John Wiley & Sons, Inc., 1981, pp.390-395.

[Trom91] L. P. Trombetta, F. J. Feigl, and R. J. Zeto," Positive charge generation in metal-oxide-semiconductor capacitors", J. Appl. Phys., Vol.**69**, Issues 4, pp.2512-2521, 1991.

[Vanh97] K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur, and Z. J. Lemnios," Non-volatile memory device based on mobile protons in SiO2 thin films ", Nature 386, 587 – 589 , April 1997;

[Velos04] A. Veloso, K.G. Anil, L. Witters, S. Brus, S. Kubicek, J-F. de Marneffe, B. Sijmus, K. Devriendt, A. Lauwers. T. Kauerauf, M. Jurczak and S. Biesemans, "Work function engineering by FUSI and its impact on the performance and reliability of oxynitride and Hf-silicate based MOSFETs", IEDM Tech. Dig., 2004, pp. 855-858.

[VogelE98] E. M. Vogel, K. Z. Ahmed, B. Hornung, W. K. Henson, P. K. McLarty, G. Lucovsky, J. R. Hauser, and J. J. Wortman," Modeled tunnel currents for high dielectric constant dielectrics", IEEE Trans. Electron Dev., 45, 1350 (1998).

[WangX05] X. Wang, J. Peterson, P. Majhi, M. I. Gardner, and D.-L. Kwong," Impacts of gate electrode materials on threshold voltage (V/sub th/) instability in nMOS HfO/sub 2/ gate stacks under DC and AC stressing ", IEEE Electron Device Letters, 26, 553 (2005).

[Wach86] Wschnik, R. A."The Use of Charge Pumping to Characterize Generation by Interface Traps". *IEEE Trans.Electron Devices ED-33*, 7 (1986), 1054-1061.

[Wein86] Z.A. Weinberg and M.V. Fischetti."SiO_2-induced substrate current and ist relation to positive charge in filed-effect transistors". Journal of Applied Physics, 59(3):824-832,1986.

[Wilk01] G.D. Wilk, R.M. Wallace, J.M. Anthony, "High- gate dielectrics: Current status and materials properties considerations", Journal of Applied Physics, Volume 89, Number 10, pp. 5243-5275, 2001

[Witt87] Witters, J. S., Groeseneken, G., And Maes, H. E."Programming mode dependent degradation of tunnel oxide floating gate devices". In Proc: *Int.Electron Devices Meeting* (1987), pp. 544-547.

[Witt87A] Witters, J. S., Groeseneken, G., And Maes, H. E."Degradation phenomena of tunnel oxide floating gate EEPROM devices". In Proc: *ESSDERC* (Bologna (Italy), 1987), pp. 167-170.

[Wu00] E.Y. Wu, J. Aitken, E. Nowak, A. Vayshenker, P. Varekamp, G. Hueckel, J. McKenna, D. Harmon, L.K. Han, C. Montrose, R. Dufresne, "Voltage-dependent voltage-acceleration of oxide breakdown for ultra-thin oxides", IEDM Technical Digest, pp. 541-544, 2000

[Wu02] E.Y. Wu, A. Vayshenker, E. Nowak, J. Sune, R.P. Vollertsen, W. Lai, D. Harmon, "Experimental evidence of TBD power-law for voltage dependence of oxide breakdown in ultrathin gate oxides", IEEE Trans. Electron Devices, Volume 49, Issue 12, pp. 2244-2253, 2002

[Xu02] Z. Xu, M. Houssa, S. De Gendt, and M. Heyns, "Polarity effect on the temperature dependence of leakage current through HfO2/SiO2 gate dielectric stacks," Appl. Phys. Lett., vol.80, pp.1975-1977, 2002.

[Yang99] N. Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, IEEE Trans. Electron Dev., 46, 1464 (1999).

[Yang00] K. N. Yang, H. T. Huang, M. C. Chang, C. M. Chu, Y. S. Chen, M. J. Chen, Y. M. Lin, M. C. Yu, S. M. Jang, D. C. H. Yu, and M. S. Liang, IEEE Trans. Electron Dev., 47, 2161 (2000).

[YoungC05] C. D. Young, G. Bersuker, Y. Zhao, J. J. Peterson, J. Barnett, G. A. Brown, J. H. Sim, R. Choi, B. H. Lee, P. Zeitzoff, Microelectronics Rel., 45, 806 (2005).

[YoungC05] Young, C.D.; Choi, R.; Sim, J.H.; Lee, B.H.; Zeitzoff, P.; Zhao, Y.; Matthews, K.; Brown, G.A.; Bersuker, G.; "Interfacial layer dependence of HfSixOy gate stacks on Vt-instability and charge trapping using ultra-short pulse in characterization", Reliability Physics Symposium, 2005, 75-79.

[YoungC06] C. D. Young et al.," Detection of electron trap generation due to constant voltage stress on high-k gate stacks", IRPS., pp.169-172,2006

[YoungD79] D. R. Young, E. A. Irene, D. J. DiMaria, R. F. De Keersmaecker, H. Z. Massoud," Electron trapping in SiO_2 at 295 and 77 °K", J. Appl. Phys., Vol.50, Issue 10, October 1979, pp.6366-6372.

[Zafar03] S. Zafar, A. Callegari, E. Gusev, and M. V. Fischetti, "Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks",J. Appl. Phys., Vol.93, Issue 11, pp. 9298-9303, 2003

[ZahM05] M. B. Zahid, R. Degraeve ,T. Kauerauf, G. Groeseneken, J.Zhang,"Investigation of channel length dependent time-to-breakdown with variable frequency charge pumping" 2005, IEEE, SISC

[ZhangJ92] J. F. Zhang, S. Taylor, and W. Eccleston, "Electron trap generation in thermally grown SiO_2 under Fowler–Nordheim stress", J. Appl. Phys., -- Vol.71, Issue 2, pp. 725-734, 1992

[ZhangJ01] J. F. Zhang, C. Z. Zhao, G. Groeseneken, R. Degraeve, J. N. Ellis, and C. D. Beech,"Hydrogen induced positive charge generation in gate oxides", J. Appl. Phys., Vol.90, Issue 4, pp. 1911-1919, 2001

[ZhangJ01A] J. F. Zhang, H. K. Sii, G. Groeseneken, and R. Degraeve, "Hole trapping and trap generation in the gate silicon dioxide," IEEE Trans. Electron Dev., vol.48, pp.1127-1135, 2001.

[ZhangJ04] J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken, and R. Degraeve," Hole traps in silicon dioxides. Part I. Properties", IEEE Trans. Electron Dev. Vol. 51, Issue 8, pp.1267-1273, 2004.

[ZhangJ06] J. F. Zhang, C. Z. Zhao, M. B. Zahid, G. Groseneken, R. Degraeve, and S. De Gendt," An assessment of the location of as-grown electron traps in HfO_2 /HfSiO stacks", IEEE Electron Dev. Lett., 27:1010, 817-820, 2006.

[ZhangW02] W. D. Zhang, J. F. Zhang, M. Lalor, D. Burton, G. Groeseneken, and R. Degraeve, "Two types of neutral electron traps generated in the gate silicon dioxide," IEEE Trans. Electron Dev., vol.49, pp.1868-1875, 2002.

[ZhaoY04] Y. Zhao and M. H. White,"Modeling of direct tunneling current through interfacial oxide and high-K gate stacks Solid-State Electrons", Vol.48, Issues 10-11 , October-November 2004, Pages 1801-1807

[ZhaoY] Yuegang Zhao, Keithley Instruments, Inc., Chadwin D. Young and George Brown,"Qualifying High κ Gate Materials With Charge-Trapping Measurements" International SEMATECH

[ZhaoC04] C. Z. Zhao, J. F. Zhang, G. Groeseneken, and R. Degraeve," Hole-traps in silicon dioxides. Part II. Generation mechanism", IEEE Trans. Electron Dev. 51, 1274-1280, 2004.

[ZhaoC05] C. Z. Zhao, M. B. Zahid, J. F. Zhang, G. Groeseneken, R. Degraeve, and S. De Gendt, "Properties and dynamic behavior of electron traps in HfO_2/SiO_2 stacks," Microelectronic Eng., vol.80, pp.366-369, 2005.

[ZhaoC05A] C. Z. Zhao and J. F. Zhang," Effects of hydrogen on positive charges in gate oxides", J. Appl. Phys. 97 art. no. 073703 (2005.)

[ZhaoC06] Z. Zhao, J. F. Zhang, M. B. Zahid, B. Govoreanu, G. Groeseneken, and S. De Gendt," Determination of capture cross sections for as-grown electron traps in HfO_2/HfSiO stacks", J. Appl. Phys. 100, 093716 (2006), accepted

[Zhu03] W. Zhu, J. P. Han, and T. P. Ma," Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics", IEEE Tran. Electron Devices Vol. 51, no. 1, pp. 98-105. 2004.