PROBING TECHNIQUE FOR ENERGY DISTRIBUTION OF POSITIVE CHARGES IN GATE DIELECTRICS AND ITS APPLICATION TO LIFETIME PREDICTION

SHARIFAH FATMADIANA WAN MUHAMAD HATTA

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF LIVERPOOL JOHN MOORES UNIVERSITY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

December 2013

ABSTRACT

The continuous reduction of the dimensions of CMOS devices has increased the negative bias temperature instability (NBTI) of pMOSFETs to such a level that it is limiting their lifetime. This increase of NBTI is caused mainly by three factors: an increase of nitrogen concentration in gate dielectric, a higher operation electrical field, and a higher temperature. Despite of many years' research work, there are questions on the correctness of the NBTI lifetime predicted through voltage acceleration and extrapolation. The conventional lifetime prediction technique measures the degradation slowly and it typically takes 10 ms or longer to record one threshold voltage shift. It has been reported that NBTI can recover substantially in this time and the degradation is underestimated. To minimize the recovery, ultra-fast technique has been developed and the measurement time has been reduced to the order of microseconds. Once the recovery is suppressed, however, the degradation no longer follows a power law and there is no industry-wide accepted method for lifetime prediction. The objective of this project is to overcome this challenge and to develop a reliable NBTI lifetime prediction technique after freezing the recovery. To achieve this objective, it is essential to have an in-depth knowledge on the defects responsible for the recovery.

It has been generally accepted that the NBTI recovery is dominated by the discharge of trapped holes. For the thin dielectric (e.g. < 3 nm) used by current industry, all hole traps are within direct tunnelling distance from the substrate and their discharging is mainly controlled by their energy levels against the Fermi level at the substrate interface. As a result, it is crucial to have the energy distribution of positive charges (PC) in the gate dielectric, but there is no technique available for probing this energy profile. A major achievement of this project is to develop a new technique that

can probe the energy distribution of PCs both within and beyond the silicon energy gap. After charging up the hole traps, they are allowed to discharge progressively by changing the gate bias, Vg, in the positive direction in steps. This allows the Fermi level at the interface to be swept from a level below the valence band edge to a level above the conduction band edge, giving the required energy profile. Results show that PCs can vary by one order of magnitude with energy level. The PCs in different energy regions clearly originate from different defects. The PCs below the valence band edge are as-grown hole traps which are insensitive to stress time and temperature, and substantially higher in thermal SiON. The PCs above the valence band edge are from the created defects. The PCs within bandgap saturate for either longer stress time or higher stress temperature. In contrast, the PCs above conduction band edge, namely the anti-neutralization positive charges, do not saturate and their generation is clearly thermally accelerated. This energy profile technique is applicable to both SiON and high-k/SiON stack. It is found that both of them have a high level of as-grown hole traps below the valence band edge and their main difference is that there is a clear peak in the energy density near to the conduction band edge for the High-k/SiON stack, but not for the SiON.

Based on this newly developed energy profile technique and the improved understanding, a new lifetime prediction technique has been proposed. The principle used is that a defect must be chargeable at an operation voltage, if it is to be included in the lifetime prediction. At the stress voltage, some as-grown hole traps further below Ev are charged, but they are neutral under an operation bias and must be excluded in the lifetime prediction. The new technique allows quantitative determination of the correct level of as-grown hole trapping to be included in the lifetime prediction. A main advantage of the proposed technique is that the contribution of as-grown hole traps is

experimentally measured, avoiding the use of trap-filling models and the associated fitting parameters. The successful separation of as-grown hole trapping from the total degradation allows the extraction of generated defects and restores the power-law kinetics. Based on this new lifetime prediction technique, it is concluded that the maximum operation voltage for a 10 years lifetime is substantially overestimated by the conventional prediction technique. This new lifetime prediction technique has been accepted for presentation at the 2013 International Electron Devices Meeting (IEDM).

ACKNOWLEDGEMENTS

Alhamdulillah. I thank god for the strength and ease He has given me in completing this long-journey of PhD.

Firstly, I owe my deepest gratitude to my principle supervisors Professor Jian F. Zhang and Associate Professor Dr. Norhayati Soin. Without their guidance and unconditional support over the last 3 years, this thesis would not have been made possible. I also would never have made it this far by the unwavering supervision and support given by Dr. Zhigang Ji. I am forever indebted to you. Words could not express my gratitude to him.

Secondly, I would like to send my thank you to my second supervisor, Dr. Weidong Zhang who has been so kindly to support me and help me in solving the many questions I had.

I must also thank my dear colleagues in LJMU as well as in UM. Dr. Brahim Benbakhti, Dr. Meng Duan, Dr. Baojun Tang, Mr. Jigang Ma, Dr. Daniel Lin and Ms. Hanim Hussin, thank you all.

Last but not least, I send my deepest thanks to my loving husband, Mr Shahrin Abd. Samat. Thank you for being there always all throughout my journey. And, also thank you to my family for their support and love.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	*Signification*
CP	Charge Pumping
CV	Capacitance-Voltage
EOT	Equivalent Oxide Thickness
HCI	Hot Carrier Injection
HfO_2	Hafnium Oxide
HfSiON	Hafnium Silicate Oxide Nitride
IL	Interfacial Layer
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
NBTI	Negative Bias Temperature instability
SILC	Stress Induced Leakage Current
SiON	Silicon Oxide Nitride
TaN	Tantalum Nitride
TDDB	Time Dependent Dielectric Breakdown
TiN	Titanium Nitride

LIST OF SYMBOLS

Symbol	*Description*	*Unit*
μ_{eff}	Effective mobility	$cm^{-2}/V\text{-}s$
ε_{SiO2}	Dielectric constant of SiO_2	
ε_{Si}	Dielectric constant of Si	
ε_{0}	Electric permittivity of vacuum	F/cm
ε_{IL}	Dielectric constant of the interfacial layer	
C_{ox}	Oxide capacitance	F
E_f	Fermi level	eV
E_{eff}	Effective surface field in the Si substrate	MV/cm
f	Frequency	Hz
g_m, G_m	Transconductance	S
I_d	Drain current	A
L	Mask channel length	µm
L_D	Debye length	cm
N_A	Substrate doping density	cm^{-3}
N_{it}	Interface trap density	cm^{-2}
n_i	Intrinsic carrier concentration in Si substrate	cm^{-3}
q	One electron charge	C
R	Feedback resistance	Ω
T	Temperature	^oC
V_d	Drain voltage	V
V_{fb}	Flat band voltage	V
V_g	Voltage applied on the gate	V
V_t	Threshold voltage	V
ΔV_t	Threshold voltage shift	V
W	Mask channel width	µm

1 Introduction

1.1 History of the MOSFETs' Reliability

The integrated circuits (ICs) were invented in 1958 by J.Kilby at Texas Instruments. Soon after, Frank Wanlass from Fairchild had described the first CMOS logic gate (NMOS and PMOS) in 1963. The first ICs produced in the early seventies had only a few hundreds of transistors integrated based on bipolar technology. The Metal Oxide Semiconductor (MOS) had only arrived several years later due to its instability caused by the presence of minute amounts of alkali elements in the gate dielectric, which subsequently leads to a shift of the threshold voltage of the transistor during operation.

Due to the high power consumption in bipolar circuits, MOS technology had gradually made its way into the scene. The problem relating to the high power consumption can be dealt with as the dimensions of the MOS devices can be scaled down easily compared to other transistor types. Though MOS circuits have lower power consumption, the MOSFETs are relative slower than their bipolar counterparts. Initially the NMOS had obtained a wider acceptance but due to the increase in integration

density, its high power consumption became unacceptable once more. Eventually in the eighties, the CMOS processes were widely adopted. However even before the CMOS was widely accepted, the defects presented in the CMOS technologies had been investigated ever since its first generation in 1960s [1,2].

The rapid miniaturization of circuits by the downscaling trend of the transistor sizes has been seen to evolve ever since. Through transistor scaling, one can obtain a better performance-to-cost ratio of products which induces the exponential growth of the semiconductor market. The Moores' Law [3-7] is popular in the semiconductor industry which predicts the number of transistor used per chip to double every 18 months or so. The industry is currently faced with the increasing importance of new trends which are the "More Moore" and "More than Moore" (MtM) where added value to device is provided by incorporating functionalities that do not necessarily scale according to the Moore's Law. These trends are made effective in microelectronic products where there are non-digital functionalities incorporated such as an assembly of various components, for instance passive components on a printed circuit board (PCB). The combined need for digital and non-digital functionalities in an integrated system is distinctly reflected as a dual trend in the recent International Technology Roadmap for Semiconductors (ITRS). The "More Moore" trend is to define the rapid miniaturization of digital functions while the MtM defines functional diversification of semiconductor-based devices. The non-digital functionalities is claimed to be contributing to the rapid miniaturization of electronic systems too, though not to the same extent as that of the digital functionality. The relevance of this new trend in CMOS technologies has been extensively reported in the ITRS, 2011 edition [8].

The rapid scaling for enhanced performance had also pushed existing CMOS materials closer to their intrinsic reliability limits and hence reliability issues are observed to be of significance and tend to change as the transistor dimensions are downscaled. One of the early reliability issue emerged in the 1970s was the contamination, such as mobile ions which can induce instability [9]. This was eventually overcome by the realization of handling the devices in a clean-room environment. In the 1980s, even as the downscaling of the transistor dimensions continued, the operation voltage was maintained at a constant 5 V. This results in higher electrical field in the device and hot carriers were limiting the lifetime of nMOSFETs [10,11]. Gate leakage further became a big threat under the operating bias as the gate oxide was thinner than 3 nm. In the 1990s, the time-dependent-dielectric breakdown (TDDB) [12] was the main reliability concern. Later as the year steps into the new millennium, the lifetime of MOSFETs were threatened by negative bias temperature instability (NBTI), which will be elaborated in the following section. In the future, the degradation–induced time dependent device variability is predicted to be of a major issue [13]. However this variability issue is out of the scope conducted in the work of this thesis, and hence will not be further elaborated.

1.2 Negative Bias Temperature Instability

The negative bias temperature instability (NBTI) is one of the dominant reliability concerns in analog and digital CMOS technologies. This instability specifically occurs in pMOSFETs employed by current technology generations. The common observation resulting to this instability is that the threshold voltage of the transistor may increase

over time, followed by a reduction in the drive current. This ultimately degraded the circuit performance and decreases its lifetime.

The NBTI is one of the earliest reported instabilities occurring in the MOSFETs. In 1967, Deal et al [2] had studied NBTI in integrated circuits and in 1977, Jeppson and Svennsson [14] developed a theory based on the generalized reaction-diffusion (RD) model to explain the kinetics of NBTI. They observed that the defect generation followed a power law dependence against time, with a power factor ranging from 0.2 to 0.3. The role of relaxation and bulk traps were discussed.

NBTI received relatively less attention in the early 1980s however, due to the emergence of the NMOS technology and the buried channel PMOS [15]. In the late 1980s and early 1990s, the NBTI issue begins to become a concern due to the application of dual poly gate, though the hot carrier injection (HCI) dominates the device reliability issue. Not until the CMOS technology starts to scale below the 130 nm technology node [16-19] that the issue of NBTI is revisited. The rapid scaling of the CMOS technology results the following dominating trends: the first is that the oxide electric field increases significantly as a result the reduction in oxide thickness, without a corresponding reduction in the supply voltage. The voltage scaling reduces the HCI and TDDB effects, but the increase in field and temperature enhances the NBTI effects for both analog and digital circuits. Secondly, in order to reduce the gate leakage and boron penetration effects, oxynitride are adapted into the CMOS technology extensively. Since then, there were reports [20-22] to demonstrate that nitridation enhances NBTI and positive change formation.

The study of the NBTI involves the large negative bias stressing (which ranges from 5 – 12 MV/cm) at temperatures ranging from the room temperature to elevated temperatures (100° C - 200° C). After a preset stress time, the sample is cooled down to room temperature and its device characteristics are measured. The effects of the negative bias temperature stress is observed from the degradation suffered by the critical device parameters, specifically the threshold voltage, subthreshold slope, the transconductance, mobility and the drain current. This stress-measure-stress cycle is repeated and the time of stress extends to typically 10^3 to 10^5 s.

1.3 Models of Negative Bias Temperature Instability

There are various reports published in trying to explain the physical mechanism behind NBTI. The most prevalent of these reports are that of Alam [23] which suggests of the reaction-diffusion (RD) theory and Huard et al [24] which suggests the existence of two independent components (Recoverable, Dr, and Permanent, Dp) where each has its own different mechanism. In addition, Grasser et al [25] used a defect spectroscopy technique which entails the distribution map of defect properties. Another model which explains the kinetics of NBTI measured under the worst case condition is proposed by Z.Ji et al [26] which combines the effect of as-grown defects with that of the generated defects (The AG model).

1.3.1 Reaction-Diffusion (RD) model

The RD model was first proposed by Jeppson and Svennsson [14] and the model was further adapted by Alam [23] and has since been used extensively to explain the mechanism of the NBTI [27-35]. The physical process behind this RD model is that the degradation is initiated by the electrochemical reaction at the SiO_2/Si interface, which converts the precursors into interface states and releases a hydrogenous species. The hydrogenous species then diffuse away from the interface and this diffusion process limits the generation of interface states. The kinetic equation [36] that describes the interface reaction is as below

$$\frac{\delta N_{it}}{\delta t} = \underbrace{k_f (N_0 - N_{it})}_{Generation} - \underbrace{k_r N_{it} H_{it}^{1/a}}_{Annealing} \qquad (1)$$

where k_f is the oxide field dependent forward dissociation rate constant and k_r the annealing rate constant. N_0 denotes the initial number of electrically inactive Si-H bonds and H_{it} is the interfacial hydrogen concentration whereas a is the order of the reaction (a=1 for the atomic hydrogen, H_0 and a=2 for the molecular hydrogen, H_2) [37]. The forward and backward reaction is controlled by the hydrogen density at the interface. Therefore the transport mechanism of the hydrogen species away from the interface characterizes the degradation mechanism and thus controlling the device's parameters for instance the threshold voltage shift. After sufficient trap generation, the rate of the generation of traps is limited by the diffusion of hydrogen.

Fig. 1.1 illustrates the schematic description of the reaction-diffusion model which interprets interface trap generation. Broken Si-H bonds at the Si-SiO_2 interface create interface traps, Si+ and hydrogen species, H. Some H converts to hydrogen molecules, H_2. Diffusion of hydrogen away from the Si/SiO_2 interface controls the interface trap generation at the Si/SiO_2 interface, thus resulting to NBTI. Thinner oxides have brought the polysilicon gate closer to the Si/SiO_2 interface and therefore increasing NBTI susceptibility since hydrogen diffuse much faster in polysilicon than in oxide [38]. From the illustration presented in Fig. 1.1, it is obvious that as the gate oxide is further scaled to meet the current demand for scaled CMOS technologies, the interface trap concentration would significantly increase and thus, enhancing NBTI.

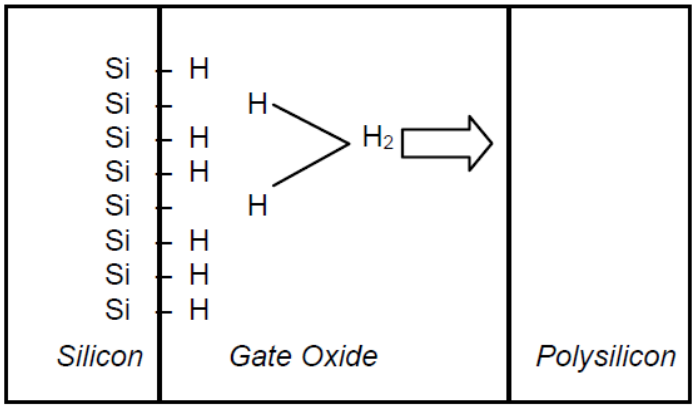


Fig. 1.1 Schematic description of the reaction-diffusion model to interpret the BTI-induced interface trap generation phenomena.

1.3.2 Hole Trapping and RD model

There are reports [39] which claim the pitfalls of the reaction-diffusion model. It is claimed that the RD model is unable to reproduce recovery characteristics [40-43]. In contrast to the experimental observation of the different relaxation magnitudes at

different recovery biases, this model predicts a bias-independent recovery. Hence, some reports concluded that the RD model is not the answer in explaining the NBTI mechanism and that the NBTI is not controlled by diffusion. The RD model theory was widely accepted until around the year 2006 when the hydrogen diffusion was suggested to be dispersive instead [44-50]. The pioneering authors of the RD model had then revised the model [51] by claiming that in addition to ΔN_{it}, stresses may also cause hole trapping in pre-existing bulk oxide traps (ΔN_{hole}) and at relatively higher stress bias, additional hole trapping in newly generated bulk oxide traps (ΔN_{ot}). The authors had conducted a lifetime prediction for differently processed SiON pMOSFETs by utilising their H-H_2 R-D framework for ΔN_{it} and link this framework with an analytical expression for ΔN_{hole} (= $A*(1-\exp(-t/\tau)^{\beta})$) in order to represent the mechanism for fast, quickly saturating (<1s) hole trapping in pre-existing bulk oxide defects.

In addition to the theory based in the diffusion and dispersion of hydrogen, the concept of hole trapping is widely acknowledged, particularly in explaining the NBTI, by the concept of recoverable hole trapping. Elastic tunnelling is used to model the process of hole trapping in which the holes tunnel into pre-existing traps at various distances away from the interface. Elastic hole trapping is defined to be first-order temperature independent and linearly dependent on the stress field [52].

There are also hole trapping models which are based on a dispersion of activation energies which results in a $1/f$ -noise behaviour for a homogeneous distribution. It was assumed that the holes can be captured through a thermally activated multiphonon emission (MPE) process into deep near-interfacial states/border traps and into oxygen vacancies called E′ centres. [53,54]. The difference between the MPE from the

conventionally invoked elastic tunnelling is that the MPE is temperature activated and its mechanism is assumed be derived under negligible electric fields. Hence, due to this non-field dependence assumption, it is difficult to use MPE to explain the mechanism of NBTI. An extension of the MPE is the multiphonon-field-assisted tunnelling (MPFAT) [55,56] in which it is an extension of the MPE to the large electric field, developed for the emission of particles from deep traps. Fig. 1.2 presents the energy band diagram of the MPFAT process for a clearer representation of this hole trapping mechanism.

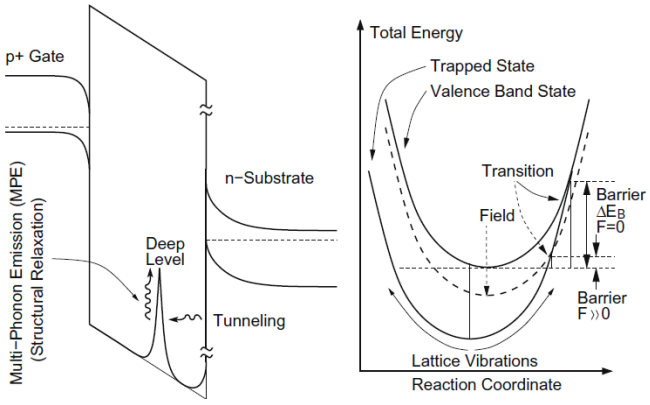


Fig. 1.2 The multiphonon-field-assisted tunneling (MPFAT) process used to explain the experimental data: elastic tunneling into deep states is only allowed when the excess energy of holes can be released via a multiphonon emission process during structural relaxation. The probability for a thermionic transition over the barrier ΔE_B has been estimated as $\exp(-\beta\Delta E_B)$ using 1D reaction-coordinate calculations, with $\beta = 1/kBT$. Application of an electric field shifts the total energy of the valence band state(dashed line), increasing the transition probability by $\exp(F^2=F^2_c)$. [25]

On the hole trapping kinetics and capture cross sections, after carefully eliminating the simultaneous neutralization and generation process by using the substrate hole injection technique, it has been shown that hole trapping follows the first order reaction model well with two well separated capture cross sections, σ:

$$Ne = \sum_{i=1}^{2} N_i[1-\exp(-\sigma_i Qh)] \qquad (2)$$

where *Ne* is the effective density of trapped holes by assuming all traps are at the oxide/Si interface [57-60]. N_i is the saturation level of traps with a capture cross section of σ_i. *Qh* is the number of holes injected into the oxide that can fill traps. The two extracted capture cross sections are in the order of $10^{-13} \sim 10^{-14}$ cm^2 and 10^{-15} cm^2, respectively [57-60]. It is proposed that oxygen vacancies are hole traps [58,60] and the smaller trap is hydrogen-related [60].

However, the RD populist would suggest that hole trapping remains to give a non-significant impact on the NBTI and had reflected that the hole trapping is merely a small threat in which its effect can be easily eradicated by removing the initial experimental data up to 1 sec so as to unveil the RD degradation mechanism [51].

1.3.3 Two-components Model (Huard)

A break from the RD tradition was brought about by the work of Huard [61] from STMicroelecronics where the author strongly suggested that hole trapping made important contribution to the degradation. He had explained that the NBTI is made up of two independent components, which are the recoverable component and the permanent component and these two distinct components demonstrate different voltage, temperature and process dependences. Fig. 1.3 describes the general scheme of the model presented by Huard et al [61]. The permanent component is suggested to be consisting of equal proportion of the interface states and positive fixed charges (1:1 relationship). It was also shown that this component will gradually reach total recovery after long thermal anneals [62].

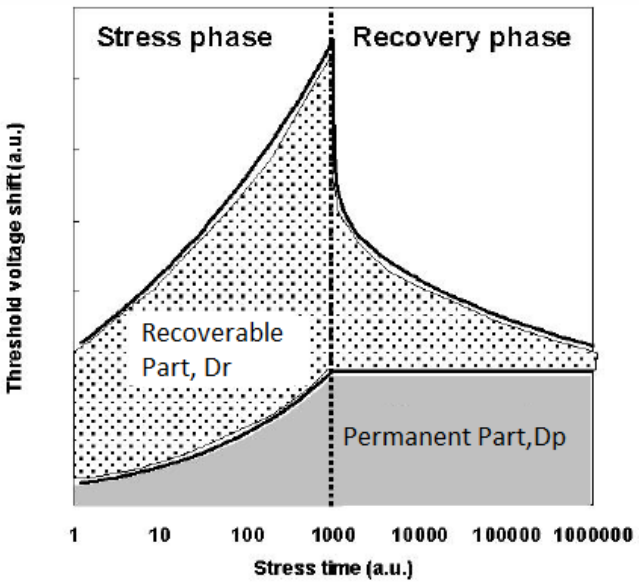


Fig. 1.3 General scheme of NBTI degradation consisting of permanent and recoverable parts [61].

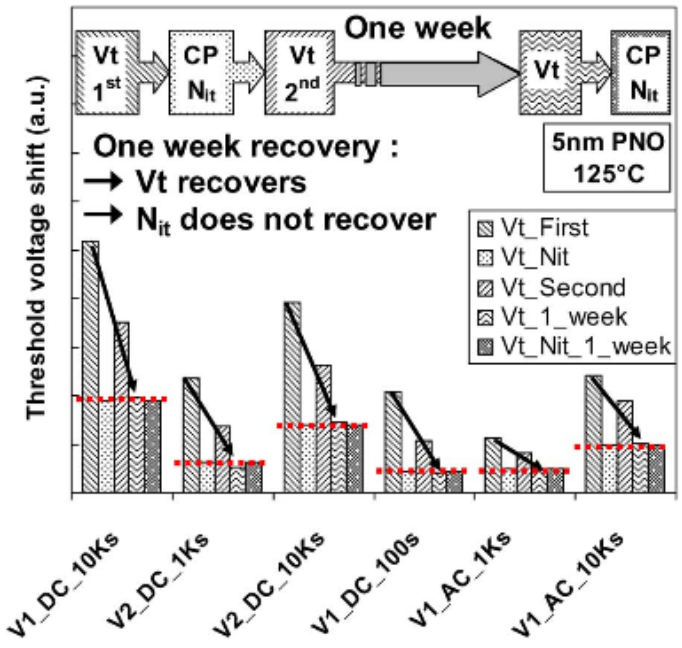


Fig.1.4 ΔV_t Charge Pumping-induced recovery observed after one week of anneal. The V_t shift continue to recover to reach the same level as Nit. [61]

Fig. 1.4 is the experimental observation which demonstrates the recovery of the threshold voltage after one week of annealing, eventually leading to the recovery of the positive charge to the same level as the Nit. Hence leading to the justification that of the 1:1 correlation between the interface states and the positive fixed charges which makes up the permanent component. The recoverable component on the other hand comprises of the hole trapping/detrapping processes to pre-existing defects in the dielectrics.

1.3.4 CET Mapping Model

Recently a physics-based analytic model was proposed by [63] to capture the BTI degradation mechanism during DC, AC and duty-factor dependent stress as well as BTI mechanism during recovery. The model is essentially based on the earlier proposed capture/emission time (CET) maps [64,65] which are used to extract accurate NBTI-relevant defect parameters. The extraction of capture and emission time constants, τ_c and τ_e , which corresponds to stress and recovery respectively, were reported in earlier literatures [66,67]. The CET map is formed by also including the information extracted from the temperature- and field-dependencies (τ_cs and τ_es) in order to gain insight in the physical processes. A determination of averaged time constants is obtained by extracting the averaged values for τ_c and τ_e from repetitive measurements. The typical number of measurements that needed for the formation of this model can resort to hundred or more and the main parameters are the gate voltage, the length of excitation pulse and the temperature.

1.3.5 As-grown Generation (AG) Model

It has been reported in many literatures that lifetime prediction using quasi-dc measurement may result in a substantial error due to the significant recovery which may occur during the measurement. This resorts in the use of ultra-fast pulse measurements, where time was reduced to the order of microseconds, when carrying out the Vg acceleration tests. Ji et al [26] however, has reported of the substantial error of the lifetime prediction at a worst case scenario when the recovery is suppressed and the Vg-acceleration method was used. The Vg acceleration models commonly used by various literatures in lifetime prediction are the power law model, $|Vg\_st|^{-\alpha}$ and the exponential model exp(-|Vg_st|). The detailed [26] investigation reports on how these models failed to predict the correct lifetime under operational voltage, Vg_op. This is since the NBTI kinetics no longer follow a simple power law.

Ji et al [26] proposed a new model to predict the NBTI lifetime at the operational bias, Vgop, and the ultra-fast pulse measurements were used to validate this model. They had observed a distinct 'shoulder' in the NBTI kinetics, which is insensitive to temperature. Fig.1.5 presents the kinetic feature of the 'shoulder' sensed by the ultra-fast pulse measurement. The presence of this 'shoulder' is claimed to be the effect of the saturated charging of as-grown defects which dominated during the initial period of stress. Subsequent to this initial period, a rise in the ΔV_{th} is observed suggesting the generation of new defects.

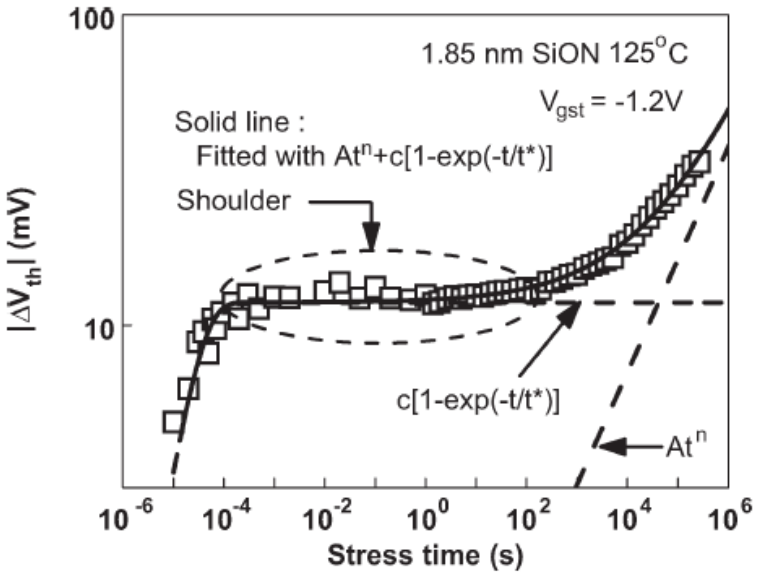


Fig. 1.5 Kinetic feature of a 'shoulder' measured by the ultra-fast pulse measurement ΔV_{th} sensed at $|V_g|$ =1.2V. [26]

The charging kinetics of the AHT is reported to generally follow the first-reaction model [57,58], whereas the generation of new defects follows a power law [68-70]. Hence, by combining the first-reaction model and the power law, the authors proposed the following expression to reflect the kinetics of the NBTI under the worst case scenario:

$$\Delta V_{th} = At^n + c\left(1 - e^{-t/t^*}\right) \qquad (3)$$

where for a given stress temperature and bias, A, n, c and t* are constants and were obtained by fitting experimental data with the least-square errors.

All the models mentioned above have been shown to be able to fit the test data, but the real acid-test for a model is whether it can predict NBTI when test data are not available. Unfortunately, little information and efforts have been made to test the

prediction ability of the proposed models. The AG model is the only model that the authors [26] demonstrated that it can not only fit the test data over ten orders of magnitude in time, but also predict the ΔV_{th} two orders of magnitude ahead where no test data were used to fit the model. By combining the first-order model for the AHTs and the power law for generating new defects, ΔV_{th} can be modelled over ten orders of stress time. This kinetic model is then used to predict the NBTI lifetime. This prediction ability makes the AG model stand out from the rest.

1.4 Positive Charges Formation in Gate Dielectric

1.4.1 Confusions in Positive Charges

Positive charges (PC) in gate dielectrics play an important role in NBTI. There are confusions relating to their understanding of their generation mechanisms and dependencies. Extensive works have been carried out in an attempt to characterize them. Two types of positive charges have been reported. One of them is the trapped hole that once fully recovered to its precursor state, will need a further hole injection to be recharged. Another type of positive charge is known to be recoverable and easily be positively recharged under a negative bias without switching on the hole injection [71,72]. This type of positive charge is referred to by various names, including anomalous positive charges [73,74], slow states [75], border traps [76], and switching oxide traps [77]. The formation of the anomalous positive charge (APC) remains to be a confusion among the published literatures and the understanding of the characteristics of this positive charge type is still poor. There might be more than one type of APC and there have been a lack of evidence to clarify the relation between the APC and the trapped hole.

This confusion has been brought to light by Zhang et al [78] in which they showed that the anomalous behaviour originates from the simultaneous presence of different types of positive charges in SiO_2. This is further elaborated in the next section where the framework of the positive charge is given.

1.4.2 Framework

It has been reported [78-81] that there are three different types of hole traps in the SiO_2 : the cyclic positive charge (CPC), the anti-neutralization positive charge (ANPC), and the as-grown hole trapping (AHT). It was shown that the CPC and ANPC are generated traps and that the AHT, regardless of its distance from the SiO2 interface, does not contribute to the generation of the new traps. The CPC can be repeatedly charged and discharged under E_{ox} = ±5 MV/cm. Another interesting property of the CPC is that it is insensitive to temperature. The ANPC can be easily positively charged without hole injection and but is difficult to neutralize. The ANPC, which is a thermally activated defect, can eventually be neutralised by high electron injection. However, the higher the energy level of the ANPC, the less number of electrons are able to reach the defect in order to neutralize it. In contrast to the ANPC, subsequent to neutralization, AHT cannot be recharged without hole injection. The initial charge state of the as-grown hole traps is neutral since it is energetically located below Ev.

Fig.1.6 [80] describes the separation of the three types of positive charges. Fig 1.6 (a) illustrates the procedure of how these positive charges was separated. A pMOSFET was stressed by Substrate Hole Injection (SHI) to form the positive charges. Subsequent

neutralization was carried out by Fowler Nordheim Injection (FNI). After the neutralization, a positive and negative bias with the magnitude of oxide field of E_{ox}=±5 MV/cm, was alternately applied with all other terminals grounded.

Figs.1.6 (b) and (c) [80] illustrate the energy levels of each types of these positive charges during their neutralization and recharging respectively. It was reported that the neutralization and charging of the CPC only involves the electron tunnelling at the same energy level. The ANPC on the other hand has an energy level above the conduction band edge of Si hence making its neutralization difficult. AHT cannot be recharged under the same Eox=-5 MV/cm without switching on of SHI, since their energy level is well below Si Ev and can only be reached by hot holes.

It was also shown that this framework can be applied to the PCs formed under different stress conditions, such as NBTI stresses, and in high-k/SiON stacks. For example, Figs. 1.7 (a) and (b) [80] present the positive charges in the HfO_2 and Hf-Silicate respectively after NBTI stresses. It has been reported that the PCs in high-k/SiON stack are dominated by the interfacial SiON layer [81].

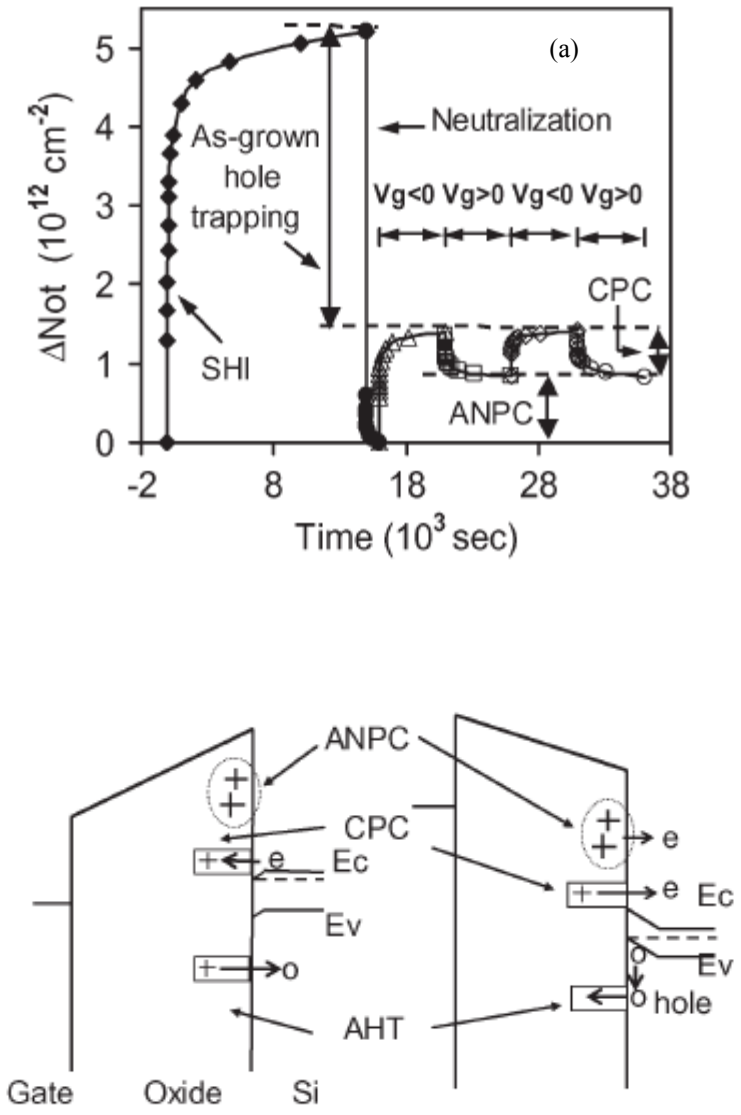


Fig.1.6 Separation of three types of positive charges: as-grown hole trapping (AHT), cyclic positive charges (CPC) and anti neutralization positive charges (ANPC). [80]

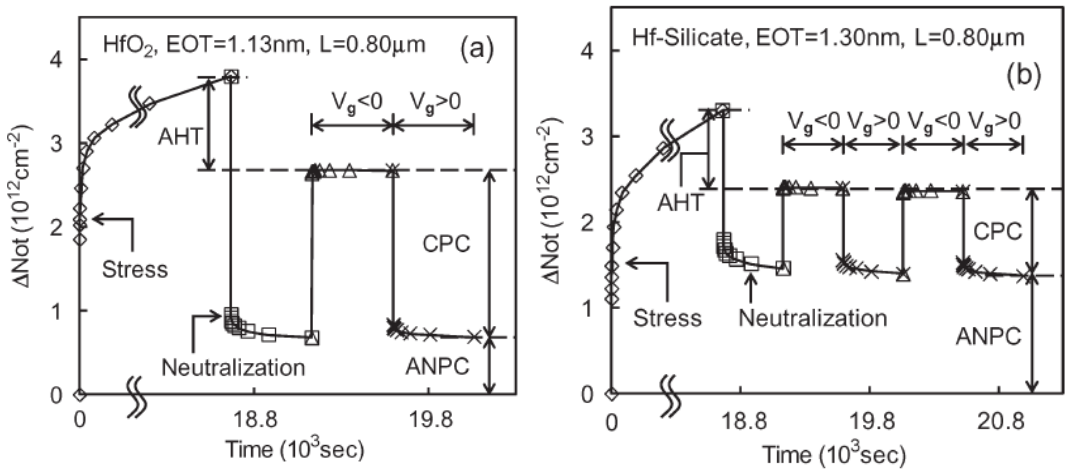


Fig.1.7 Separation of three types of positive charges: AHT, CPC and ANPC in Si-based Hf devices [80]

1.5 The rationale of the project work

Although a framework has been proposed for the PCs in gate dielectrics, it only gives a rule-of-the-thumb estimation of the energy levels of different types of PCs. Before this project, there is no probing technique that can extract the energy of PCs both within and beyond Si bandgap. Without this detailed energy distribution, not only our understanding of PCs is limited, but also there is a lack of confidence in the predicted NBTI lifetime for the threshold voltage. As a result, there is an urgent need for developing a new probing technique that can give detailed energy distribution of PCs in gate oxides both within and beyond Si bandgap. The **central objective of this project** is to develop such a technique and then use it to show the shortcomings of the lifetime prediction technique currently used by the industry.

Part of the PhD period (1 ½ years) was conducted in Kuala Lumpur, whereby the background research on NBTI was conducted together with initial simulation and experimental work. The work conducted in Kuala Lumpur had resulted in 3 journal

papers and 3 conference papers. The later part of the PhD period (2 years), consisting fully of experimental work, was conducted in Liverpool John Moores University (LJMU). The major output from the work conducted at the LJMU is the publication in the IEEE Transaction of Electron Devices and the invitation for a presentation at the 2013 International Electron Devices Meeting.

1.6 Organization of the thesis

This thesis is organized as follows:

Chapter 2 comprises a review of the characterization techniques used in investigating Negative Bias Temperature Instability (NBTI), which includes both conventional and fast techniques. The DC transfer characteristic, capacitance-voltage (C-V), On-The-Fly (OTF) and fast pulse I_d-V_g measurements are reviewed.

Chapter 3 comprises of the comprehensive demonstration of the probing technique in obtaining the energy distribution of positive charges (PCs) in the dielectric which is extracted within and beyond the Si bandgap. The study will show that the PCs have a broad energy distribution and the results strongly support the existence of different types of PCs.

Chapter 4 comprises of the impact of CMOS processes on the defects and their energy distributions. It will be demonstrated that the newly proposed energy distribution

technique is applicable not only to conventional single-layered SiON devices but also to various advanced high-k gate stacks. The energy distribution of different high-k devices will be investigated, including both HfO_2/SiON and the AlO-capped high-k/SiON stack. Attention will be paid to their differences from the pure SiON samples.

In Chapter 5, the application of the newly developed technique will be explored to predict the device lifetime. It will demonstrate how the current conventional technique used by the industry, as well as the advanced fast pulse measurement method suggested by other works is not suitable for lifetime prediction due to their incorrect inclusion of defects in the prediction. A new lifetime prediction method which can readily be implemented in the industry will be proposed, based on the energy distribution technique.

Chapter 6 summarizes the work presented in this thesis. Finally the direction for future work is suggested.

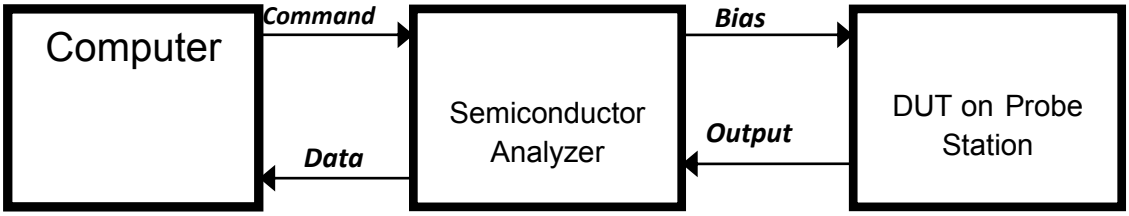
2 Devices and Test Facilities

2.1 Introduction

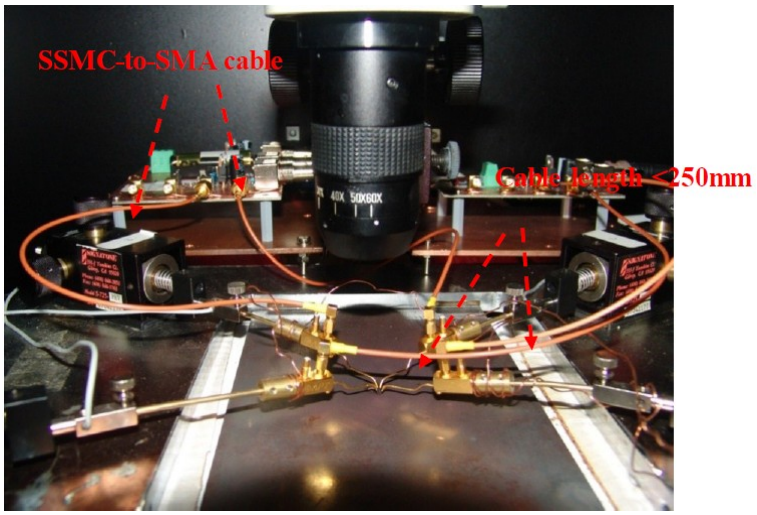
In this chapter, the test systems for the conventional measurement techniques, including the transfer characteristics (I-V), charge pumping (CP) and capacitance-voltage (C-V), are reviewed. The devices used in the experiments are also presented. Fast measurements, including the On-The-Fly (OTF), fast pulse I_d-V_g and single point measurements are further described, due to their significance to this project. The system accuracy checks of the pulse measurement are calibrated.

The essential equipment required for a standard wafer-level device characterization is illustrated in Fig. 2.1 (a). The device-under-test (DUT) is placed on a stage of a probe station and the probe station is housed in a black box to minimize interferences to the measurement from the outside environment. Four micro-positioners are used, of which each positioner is connected to one of the four terminals of a DUT. The micro-positioners are also connected to the semiconductor analyser and pulse generator controlled by a personal computer. Fig. 2.1 (b) presents a photo of the micropositioners located in the probe station. As depicted in the photograph, the SSMC-to-SMA cables

are needed to connect the amplifier circuit to the micropositioners. In order to avoid impedance mismatch, all system components must possess a 50Ω impedance and the length of the SSMC-to-SMA cables are minimized. The BNC cables between the circuit and the oscilloscope is required to be of the same lengths in order to synchronise the multiple output channels.



(a)



(b)

Fig. 2.1 (a) Schematic diagram of the conventional measurement system and (b) Photograph of the SSMC-to-SMA BNC cable to connect the testing device and circuit.

For the pulse measurement conducted in this work, the pulse generator used is the Arbitrary Waveform Generator Agilent HP 81150 A. An external circuit containing an op-amplifier is used to measure the drain current. The description of this circuit will be further elaborated in the following sections. The pulse system is able to provide a

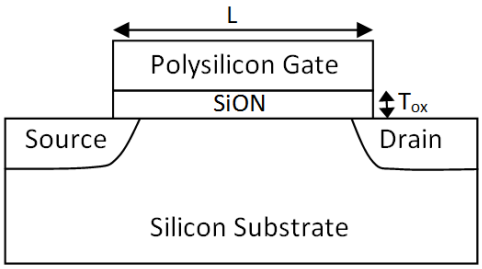
minimum measurement time of 100 ns with a noise margin of 5 mV. A program written in Visual Basic is used to control the pulse generator and automate the measurements conducted in this work. For the quasi-DC and the capacitance-voltage (CV) measurements, the CASCADE probe station and industrial standard parameter analysers, the Agilent E5270A and the Keithley 4200-SCS were used. It typically takes 20~150 ms for measuring one point and in order to obtain a transfer characteristic, tens of points are needed and the total measurement time will be in the order of seconds. Most of the experiments conducted in this work characterizes thin (< 3nm) gate oxides. Since both recovery and degradation can occur during the quasi-DC measurement, it becomes essential to increase the measurement speed by using the pulse measurements, which will be discussed later in this chapter. However, the DC measurement will still be carried out in order to compare the work conducted in this thesis to the standard slow measurement typically conducted in the industry.

For experiment conducted in this work, the value of the Vd is chosen differently for the different devices used. These values were chosen according to the maximum current a particular device can output. The higher Vd will lead to a high maximum current. However, the maximum current needs to be limited not higher than 150 μA , due to the limitation of the measurement system. For the measurement conducted in this work, a feedback resistance of 10kΩ was used in the external circuit containing the op-amplifier. If the maximum current is higher than this set value, this might lead to the breakdown of the system. A tradeoff by having a lower Vd is that the current might be lower and hence will introduce noise interference to the measurement. Therefore, prior in commencing each experiment, the best Vd need to be chosen in order to avoid such noise, but at the same time, is able to maintain a well working measurement system.

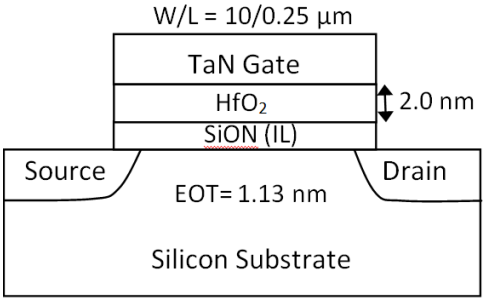
2.2 Devices

The gate dielectric of the conventional pMOSFETs used in this project includes the 1.4 nm, 1.85 nm, 2.0 nm and 2.3 nm plasma-nitrided (PNO) SiON and the 2.7 nm thermally-nitrided (TNO) SiON. High-k gate stacks used in this work were the 1.13 nm (equivalent oxide thickness) Ta/HfO_2/SiON stack, a 2.0 nm Al-capped HfO_2/SiO_2 stack, FUSI gated 1.52 nm Hf-Silicate, a 1.53 nm TiN/HfSiON/SiON stack and a 2.13 nm TaN/HfSiON/SiON. The 2.3 nm and 2.0 nm PNO SiON devices are standard devices used in the industry hence providing a good benchmark for the research presented in this thesis. Various types of high-k devices were used to demonstrate that the newly proposed energy probing technique is applicable not only to conventional devices, but also to advanced high-k processes. Figures 2.2 (a)-(f) depict the cross-section of the devices used in this work.

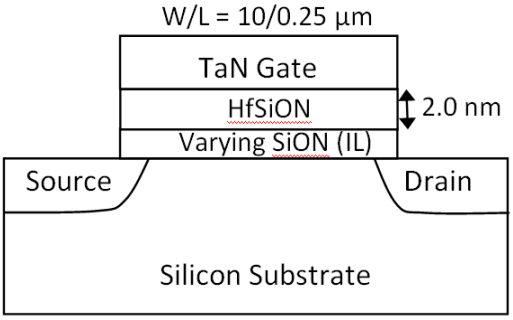
Nitridation	Gate oxide thickness, T_{ox}	W/L
Plasma Nitrided	1.4 nm	10/0.25 µm
	1.85 nm	10/0.13 µm
	2.0 nm	10/0.13 µm
	2.3 nm	1/0.35 µm
Thermally Nitrided	2.7 nm	10/0.13 µm



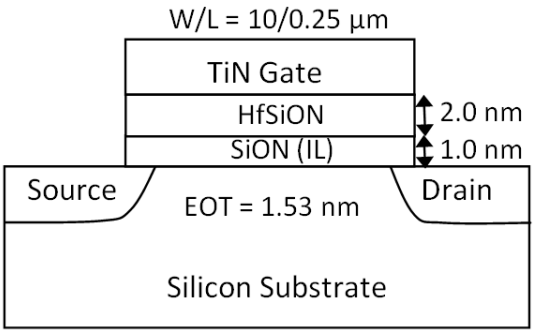
(a)



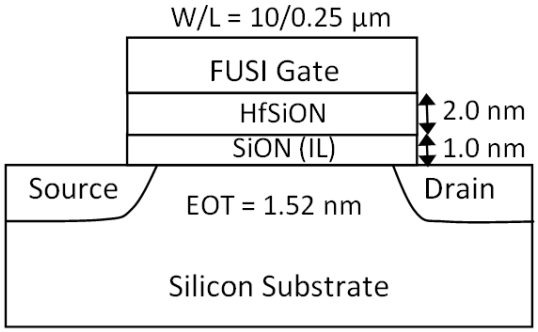
(b)



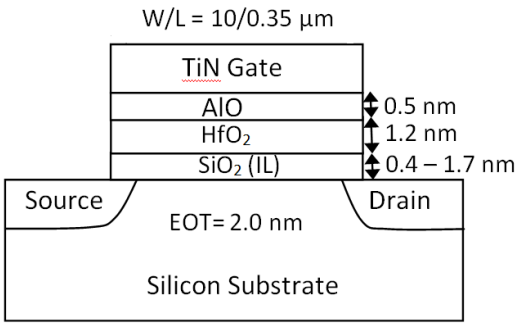
(c)



(d)



(e)



(f)

Fig. 2.2 Cross section of the devices used in this work. (a) Single Layer SiON device. The table shows the different W/L of the devices used. High-k gate stack of (b) 1.13 nm HfO2/SiON, (c) TaN/HfSiON with varying IL, (d) 1.53 nm TiN/HfSiON, (e) 1.52 nm FUSI/HfSiON and (f) 2.0 nm Al-capped HfO_2/SiO_2.

2.3 Conventional characterization and stress techniques

In this section, the conventional characterization and stress techniques for SiO_2-based MOSFETs will be presented. It includes system and equipment setup, samples used in the experiments, techniques for characterizing degradation, and techniques for stressing devices.

2.3.1 Conventional I_d-V_g technique

The conventional transfer characteristics (I_d-V_g) of a device are measured by applying DC voltages using a DC parametric analyser. This type of measurement is called the 'slow measurement' since the total time for one I_d-V_g measurement is typically in the order of 1-10 seconds. The threshold voltage of the device is extrapolated from the I_d-V_g curve by either applying the gm-max method [82] or the constant current method [83]. Fig. 2.3 (a) and (b) demonstrates how the threshold voltage, V_t is extracted by the gm-max extrapolation method and the constant current method respectively. The gm-max extrapolation method requires the transconductance which is calculated by differentiating the I_d-V_g curve. The threshold voltage is extracted from the gate voltage axis intercept of the linear extrapolation of the I_d-V_g curve at maximum transconductance. The experimental data presented in these figures were measured by the author, using the measurement system presented in section 2.1.

The conventional method of the negative bias temperature instability (NBTI) is carried out using the stress-measure-stress methodology. The test is started by characterising the properties of a fresh device, such as the measuring the threshold

voltage. The fresh value is then used as the reference for measuring the shift of parameters during the stress due to degradation. In order to produce a measurable degradation within a practical test time, the stress biases applied are typically considerably higher than that used in the real operation. The device under test is stressed at a certain temperature and the measurement can be either at the stress temperature or a value between the room and stress temperature. The measurement is interrupted at preset times to measure the I_d-V_g transfer characteristics. The stressed I_d-V_g transfer characteristics are expected to be shifted in the negative direction, which signifies the generation of positive charges. Figs. 2.4 (a) and (b) presents the typical shift of the I_d-V_g curve, compared with the fresh I_d-V_g and the typical threshold voltage shift under NBTI stress.

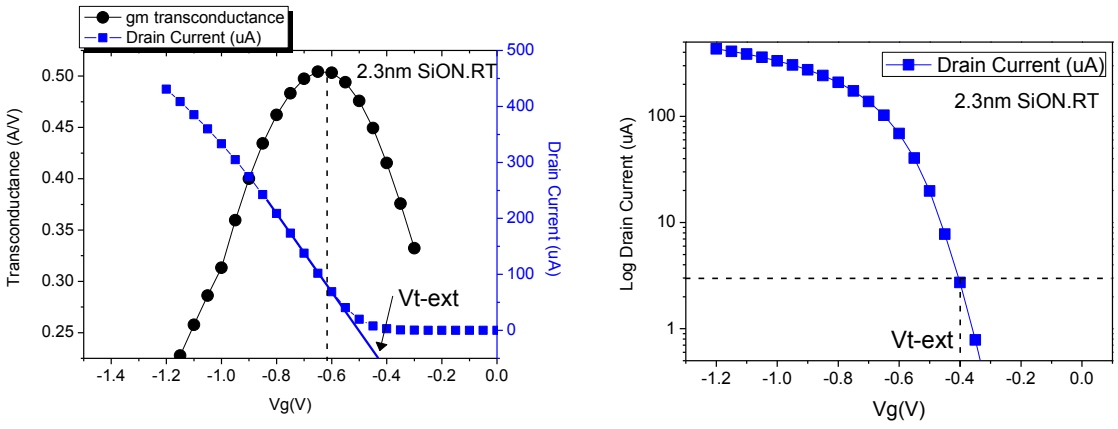
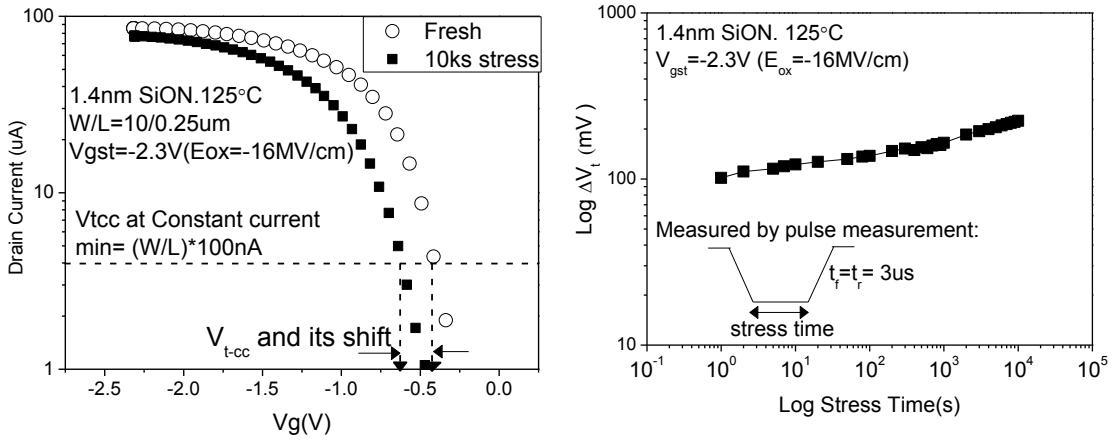


Fig.2.3 (a) demonstrates how the threshold voltage Vt is extracted by the gm-max extrapolation method and (b) demonstrates how the threshold voltage Vt is extracted by the constant current method.



(a) (b)

Fig. 2.4 (a) A typical result obtained by using the pulse measurement technique. I_d-V_g curves were measured after stressing the device under NBTI for a 10 ks. After stress, I_d-V_g curve is shifted towards higher |Vg|. Threshold voltage is extracted by using the constant current method. The standard industry practice is to use a constant current of W/L*100 nA. (b) Typical NBTI degradation against stress time.

2.3.2 Conventional Charge Pumping (CP) technique

The charge pumping is a technique used in this work for measuring the interface states in MOS devices. It was first introduced by Brugler and Jespers in 1969 [84] and this technique is widely accepted due to the excellent accuracy and applicability to small geometry MOS transistors, compared to the capacitance-voltage method which requires large area devices.

There are two standard types of the charge pumping technique namely the fixed amplitude charge pumping and the variable amplitude (fixed based) charge pumping. The basic experimental set up of the charge pumping measurement is presented in Fig. 2.5 (a). A pulse waveform is applied to the transistor gate using a pulse generator, and the corresponding charge pumping current I_{cp} is measured at the substrate. The source

and drain of the transistor are connected together to a certain reverse bias voltage with respect to the substrate or grounded. For the fixed amplitude charge pumping, the V_{base} is stepped, while for the variable amplitude charge pumping, the V_{top} is stepped. The waveform applied to the gate voltage is shown in Fig. 2.5 (b). Fig. 2.5 (c) demonstrates the schematic of the fixed amplitude CP and the variable amplitude CP.

Proper selection of charge pumping pulse amplitude, falling time and rising time is important in order to obtain reliable results. In this work, only the PMOS device was used and the pulse amplitude was fixed at 1.0 V and the fall time, t_f = rise time, t_r = 0.02 μs, the frequency, f = 1 MHz with duty cycle of 50. Using rise and fall time shorter than these criteria will induce parasitic effects such as geometric effect [85] which may cause significant errors to the results.

The basic charge pumping principle is described in the following. We start the gate voltage, V_g from V_{top} and for demonstration purposes, a pMOSFET is used as an example. When the magnitude of gate voltage is higher than the threshold voltage $|V_t|$, inversion occurs and all interface traps up to the Fermi Level, V_{fb} are filled with holes. These holes are drawn into the device from the source and drain. When the gate voltage drops below $|V_t|$, the Fermi level moves further above the valence band and the interface hole concentration is reduced exponentially with the decrease in Fermi level. When the V_g increases from V_{base} to V_t, the emission of holes from interface traps only reaches the energy level, $E_{em,h}$, by [86]

$$E_{em,h} = E_i + k \cdot T \cdot \ln\left(V_{th} \cdot \sigma_p \cdot n_i \cdot t_{em,h} + e^{\frac{E_{f,acc} - E_i}{kT}} \right) \qquad (2.1)$$

where E_i is the intrinsic energy level,

k is the Boltzmann's constant,

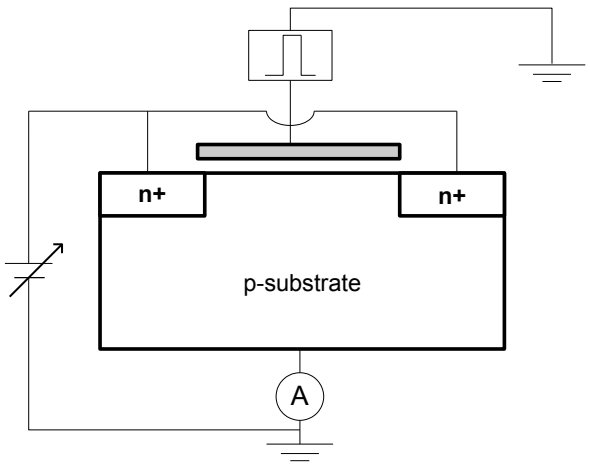
T is the absolute temperature,

V_{th} is the thermal velocity of carriers,

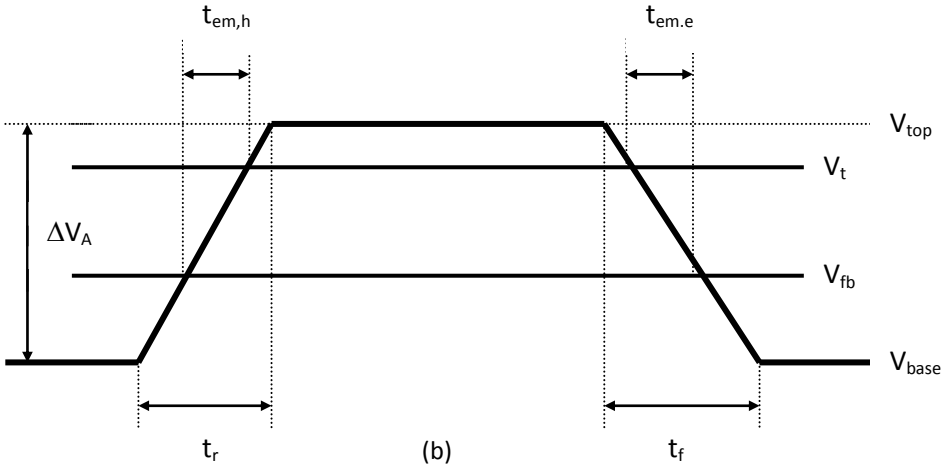
σ_p is the capture cross section for holes,

n_i is the intrinsic concentration,

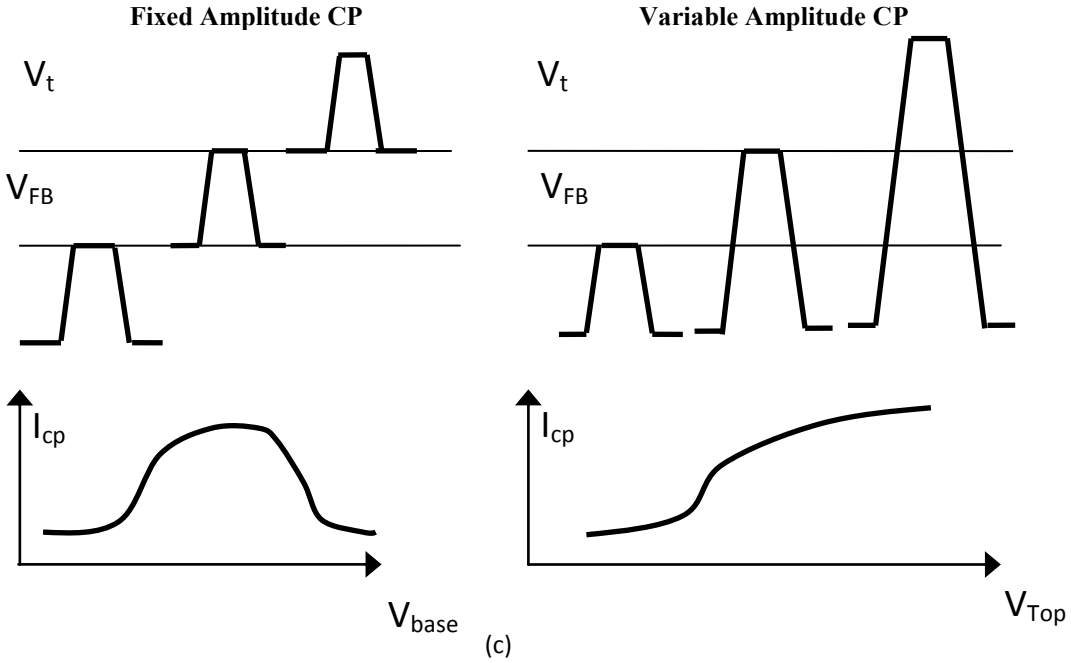
$t_{em,h}$ is the time available for the emission of holes from the interface traps during the fall time of the gate pulse, and $E_{f,acc}$ is the Fermi level in inversion.



(a)



(b)



(c)

Fig. 2.5 (a) Illustration of the experimental set up for charge pumping measurement and (b) The waveform applied on the gate during the charge pumping measurement. (c) Schematic of waveform for different charge pumping measurement techniques.

The holes still trapped in the interface states between $E_{em,h}$ and $E_{f,inv}$ will recombine with electrons from the source and drain. The repetitive occurrence of recombination processes gives rise to a net current, I_{cp}, which can be measured either at the substrate or at the source and drain. I_{cp} is determined by [87] the equation 2.2. Once all other parameters are known, D_{it} can be determined by measuring I_{cp}.

$$I_{cp} = 2 \cdot q \cdot f \cdot A_g \cdot k \cdot T \cdot D_{it} \cdot \ln\left(V_{th} \cdot n_i \cdot \sqrt{\sigma_n \cdot \sigma_p} \cdot \sqrt{t_{em,e} \cdot t_{em,h}}\right) \qquad (2.2)$$

where *f* is the gate pulse frequency,

A_g is the transistor area, and

D_{it} is the average interface state density between $E_{em,e}$ and $E_{em,h}$.

The $t_{em,e}$ and $t_{em,h}$ are determined by the rising time, t_r, and falling time, t_f, respectively.

$$t_{em,e} = \frac{|V_{fb} - V_{TH}|}{|\Delta V_G|} \cdot t_f \qquad (2.3)$$

$$t_{em,h} = \frac{|V_{fb} - V_{TH}|}{|\Delta V_G|} \cdot t_r \qquad (2.4)$$

Fig. 2.6 presents an example of the result obtained from the variable amplitude charge pumping measurement. The experimental data presented in these figures were measured by the author, using the measurement system presented in section 2.1.

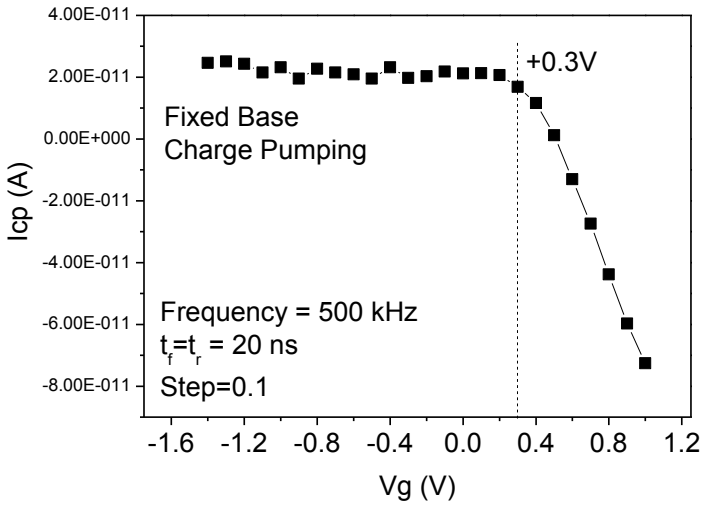


Fig. 2.6 Demonstration of charge pumping results obtained by fixed base charge pumping.

2.3.3 Conventional Capacitance-Voltage (C-V) technique

The capacitance-voltage measurements are a standard measurement used in studying the gate-oxide quality. Various MOS devices parameters such as the oxide thickness, flatband voltage, threshold voltage, bulk and interface charges information can be extracted from the C-V measurement. The capacitance-voltage behaviour of a MOS device can be described using the equivalent circuit presented in Fig. 2.7. Capacitance of a MOS capacitor is described by the change in the charge (Q_g) of a device, in which a simultaneous change in voltage (V_g) is also occurring:

$$C = \frac{dQ_g}{dV_g} \tag{2.5}$$

If to assume that there is no charge trapping in the dielectric, the concept of charge neutrality is upheld whereby $Q_g = -(Q_s + Q_{it})$. Here, Q_s is the substrate charge and Q_{it} is the trapped interface charge. The gate voltage on the other hand has a partial drop across

the dielectric and the semiconductor substrate: $V_g = V_{fb} + V_{ox} + \phi_S$, where V_{fb} is the flat-band voltage, V_{ox} is the voltage drop across the oxide and the ϕ_s is the Si surface potential. By taking in this assumption, the equation (2.5) is re-written as

$$C = -\frac{dQ_s + dQ_{it}}{dV_{ox} + d\phi_s} \tag{2.6}$$

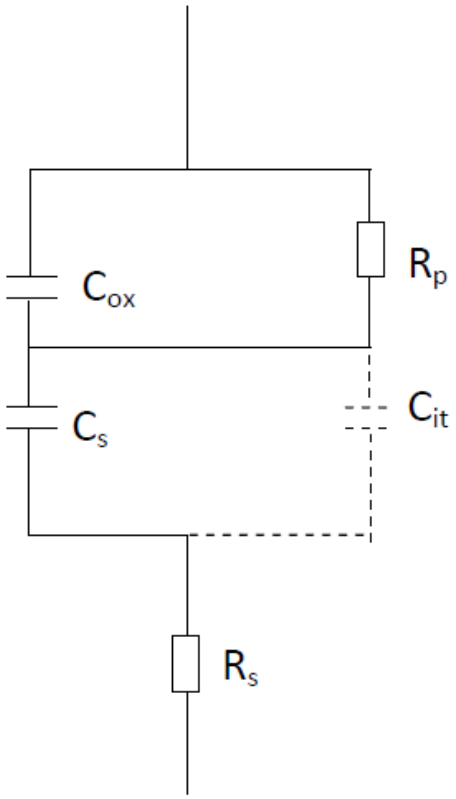


Fig. 2.7 Equivalent circuit of a MOS structure

The contribution of the majority, minority and the depletion charge associated with the substrate varies, depending on the Si surface potential. From the equivalent circuit, the total gate capacitance can also be written as:

$$C = \frac{1}{(1/C_{ox}) + (1/C_s + C_{it})} \tag{2.7}$$

The low-frequency substrate capacitance is given by [88] :

$$C_{s,lf} = \hat{U}_S \frac{\varepsilon_{Si}\varepsilon_o}{2L_D} \frac{e^{U_F}(1-e^{-U_S}) + e^{-U_F}(e^{U_S}-1)}{F(U_s, U_F)} \tag{2.8}$$

where the dimensionless surface electric field $F(U_S, U_F)$ is defined by:

$$F(U_S, U_F) = \sqrt{e^{U_F}(e^{-U_S} + U_S - 1) + e^{-U_F}(e^{U_S} - U_S - 1)} \tag{2.9}$$

U_S and U_F are normalized potentials, defined as $U_S = q\phi_s / kT$ and $U_F = q\phi_F / kT$. The Fermi potential is calculated by $\phi_F = (kT/q)ln(N_A/n_i)$ where N_A is the acceptor concentration and n_i the intrinsic carrier concentration in the Si substrate.

The symbol $\hat{U}_S$ stands for the sign of the surface potential and is given by

$$\hat{U}_S = \frac{|U_S|}{U_S} \tag{2.10}$$

Where $\hat{U}_S = 1$ for $U_S > 0$ and $\hat{U}_S = -1$ for $U_S < 0$. The extrinsic Debye length L_D is:

$$L_D = \sqrt{\frac{\varepsilon_{Si}\varepsilon_o kT}{2q^2 N_A}} \tag{2.11}$$

The basic setup of the C-V measurement conducted in this work is presented in Fig. 2.8 (a) and (b) where the gate-channel capacitance, C_{gc} and gate-bulk capacitance, C_{gb} is separately obtained through the split C-V technique [89]. The total gate capacitance is obtained by combining the C_{gc} and the C_{gb}. Fig. 2.9 (a) and (b) respectively presents the C_{gc} and the C_{gb}, and the combination of these two measurements. Parasitic capacitance

has been accounted for in the measurements. Parasitic capacitance will lead to an offset of the measurement, and this has been considered by nulling back to zero.

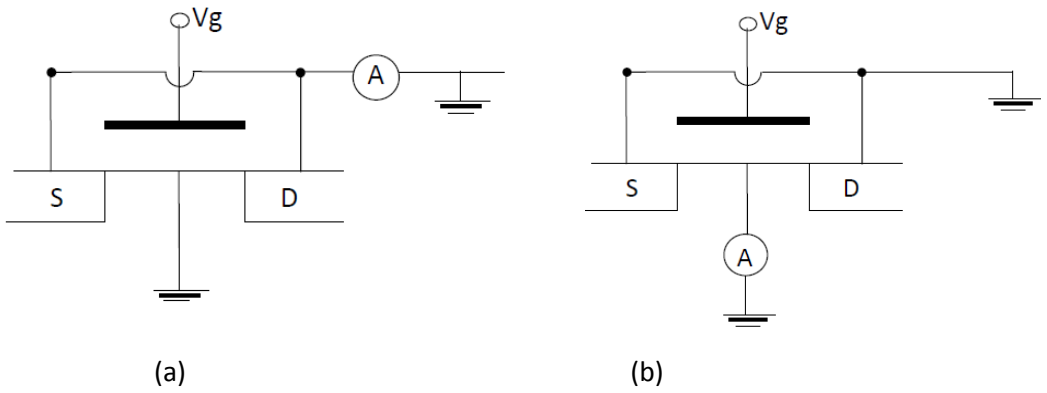


Fig. 2.8 Spilt C-V measurement technique configuration to obtain (a) $C_{gate\text{-}channel}$ against the Gate voltage and (b) $C_{gate\text{-}bulk}$ against the Gate voltage.

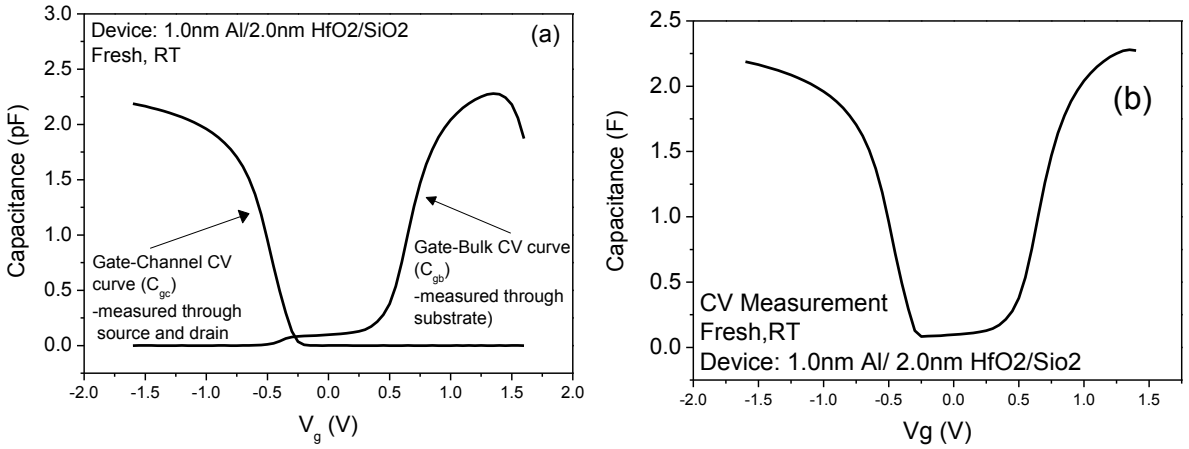


Fig. 2.9 Spilt CV measurement profiles (a) obtained from separate C-V measurement and (b) final profile obtained by the combination of the C_{gc} and C_{gb} measurement.

The total gate capacitance measurement can also be conducted using a single C-V measurement, as oppose to the split C-V measurement conducted in this work. Capacitance-Voltage measurement is conducted by superimposing a small oscillating AC voltage on a DC voltage, which is applied to the gate. The resulting AC current

through the source, drain or substrate is measured from which the capacitance, which is the change in charge in response to the AC voltage, is calculated.

The combined C-V can be divided into three regions: accumulation, depletion, and inversion, as described below for a p-type substrate.

Accumulation Region

When a negative voltage is applied to the gate of an n-type substrate MOS capacitor, the accumulation region of the C-V curve can be observed. The negative polarity will attract the majority carriers, which are the holes, towards the gate. These holes will accumulate at the oxide/substrate interface due to the oxide being a good insulator. The C-V measurement measures the oxide capacitance in the strong accumulation region at which the voltage is negative enough and the C-V curve is essentially flat. Hence, the oxide thickness can be extracted from the oxide capacitance.

Depletion Region

As the gate voltage moves toward the positive, the holes are repelled from the substrate oxide interface. Subsequently, a carrier- depleted area forms beneath the oxide. As the gate voltage becomes more positive, the depletion zone become deeper. The depletion capacitance thus becomes smaller and the total measured capacitance becomes smaller consequently.

Inversion Region

As the gate voltage increases, the mid-band energy level eventually falls below the Fermi-level at the interface, so that the interfacial region is inverted from p-type into n-type. The positive gate bias attracts, the minority carriers, which are the electrons, towards the gate. Due to the oxide being a good insulator, these minority carriers will

pile-up at the oxide/substrate interface and form an inversion layer. As a result, the positive charges on the gate are separated from the electrons in the oxide and the total capacitance returns to the oxide capacitance. The electrons in the inversion layer screen the positive charges on the gate from the substrate, so that the depletion depth will not increase further with Vg.

2.4 On-The-Fly (OTF) techniques

The primary motivation in the development of the OTF technique is to counter the occurrence of recovery during measurement. The conventional measurement techniques particularly the DC measurement can result in significant underestimation of degradation due to recovery during the measurement where the |Vg| was lowered from the stress level, as shown in Fig. 2.9 (a). Numerous literatures [90-96] detail the recovery effect that happens during the interruption of stress, leading to the conclusion that the relaxation of the NBTI degradation is dependent on the instrumentation and measurement technique applied. The common feature of the OTF is to ensure that the stress voltage is always applied to the gate, and the degradation of the drain current is measured at stress voltage. The OTF technique evaluates ΔVt at the stress gate bias so that the stress is always maintained during the measurement. The OTF technique monitors both I_d and the transconductance, g_m, at preset intervals under a low drain bias. To evaluate g_m, the stress V_g is perturbed by a small amount of ± DV and the corresponding current variation is recorded. The g_m at a time "n" is estimated from the equation 2.13.

The first OTF technique was developed by Rangan et al [97]. The author had initially measured I_d-V_g with the V_g ramped to the stress voltage. Subsequently the author recorded the drain current I_{d0} at $V_g = V_{gst}$ and the threshold voltage V_{t0}. The drain current is then continuously sampled at $Vg=V_{gst}$ in the consequent electrical stress. The threshold voltage shift is then calculated from the following equation

$$\Delta V_t = \frac{\Delta I_d}{I_{d0}} \bullet \left(V_g - V_{t0}\right) \qquad (2.12)$$

where the change in the drain current is $\Delta I_d = I_d - I_{d0}$. Further to the development of this technique by Rangan et al [16], other groups [98-102] further developed the OTF technique to mitigate the uncertainty such as the lack of the consideration of mobility variation with V_g in the eq. (2.12). The 2nd order OTF technique was proposed by Denais et al [86] as shown in Fig.2.10 (b). In order to take the mobility degradation into consideration, the transconductance, g_m is evaluated. To estimate the transconductance, $g_m(n)$, Vg was perturbed by a small ±DV, where D signifies perturbation.

$$g_m(n) = \frac{I_d(V_g + DV) - I_d(V_g - DV)}{2 \bullet DV} \qquad (2.13)$$

The degradation of drain current between two measurement points ‘n’ and ‘n-1’ is,

$$\Delta I_d(n) = I_d(n) - I_d(n-1) \qquad (2.14)$$

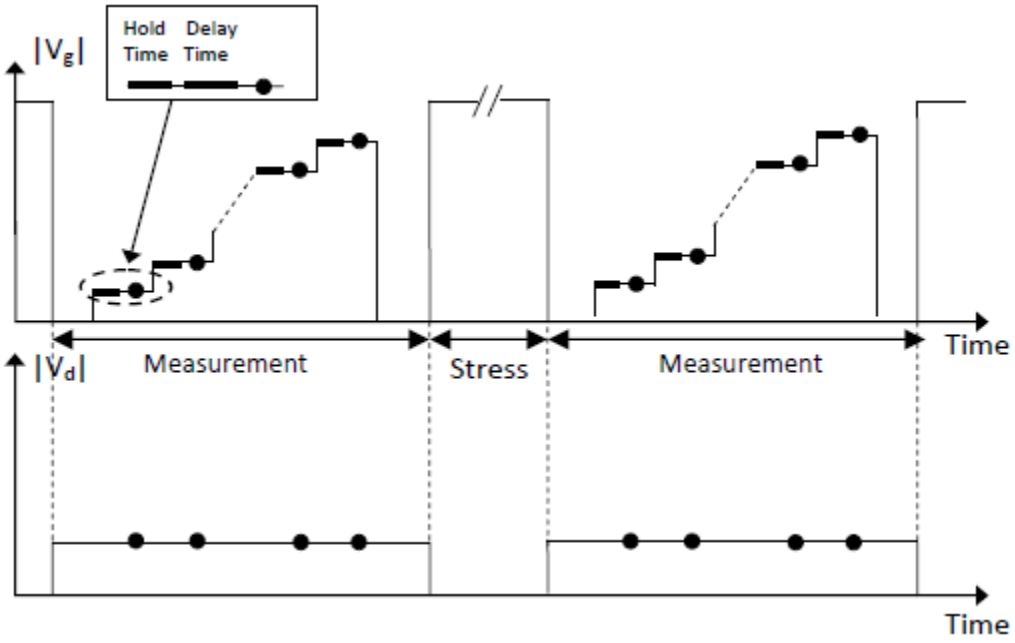
The shift of threshold voltage between these two points can be evaluated by,

$$\Delta V_t(n) = -\frac{\Delta I_d(n)}{g_m(n)} \qquad (2.15)$$

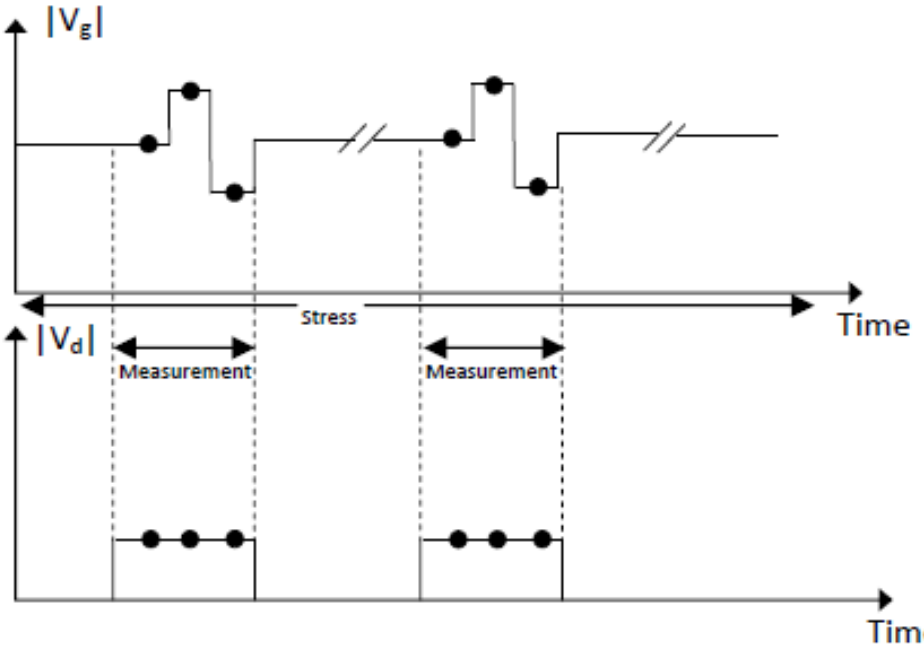
The accumulative shift of threshold voltage is,

$$\Delta V_t = -\sum_{n=1}^{M} \frac{I_d(n) - I_d(n-1)}{g_m(n)} \qquad (2.16)$$

where M is the number of I_d measurements and $g_m(n)$ is the mean value of the trans-conductance between the n^{th} and $n\text{-}1^{th}$ I_d measurements, as shown in Fig. 2.11. Hence, periodical three point I_d measurements are enough to monitor ΔI_d, g_m, ΔV_t during stress.



(a)



(b)

Fig 2.10 (a) Traditional NBTI test sequence (b) The 2nd order On-The-Fly measurement sequence

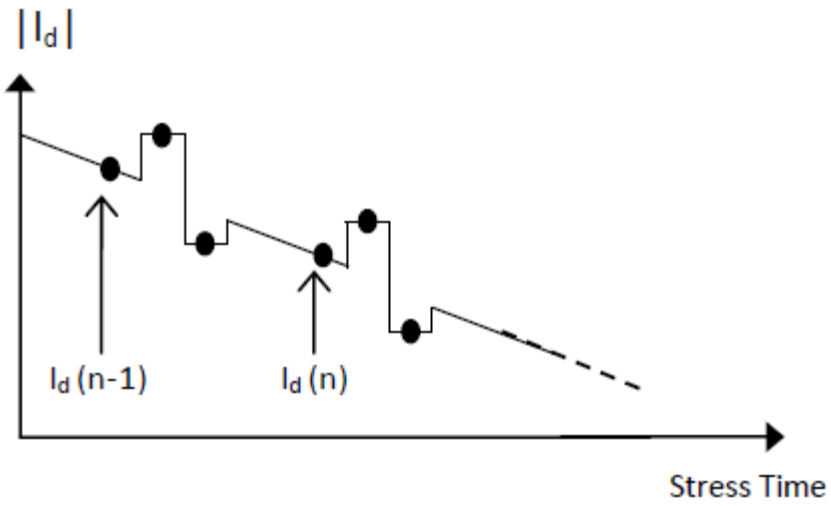


Fig. 2.11 The nth and n-1th Id measurements, together with the transconductance $g_m(n)$, can give the threshold voltage shift, ΔVt, between nth and n-1th measurement points.

2.5 Pulse I_d-V_g techniques

It is well known that conventional measurement techniques developed for SiO_2-based gate dielectrics are not capable of measuring the fast transient instabilities in high-*k* materials. Hence faster measurement techniques are required.

The pulse I-V technique is used to measure the transfer characteristics with a much faster speed than the conventional I_d~V_g measurement techniques. A pulse signal generated by the pulse generator is applied to the gate of the transistor. The drain current can be recorded by a digital oscilloscope during the pulse edges. The transfer characteristic I_d~V_g can be determined from the gate voltage and the corresponding drain current. The advantage of this technique is that the threshold voltage (V_t) can be determined after the application of a pulse by capturing the I_d~V_g during the falling

edge of the stress pulse. This can minimize the trapping/detrapping during measurement.

2.5.1 Experimental Setup

The first pulse I_d-V_g technique was developed by Kerber et al [87], in which he employed this technique to investigate on the large charge trapping occurring in high-k dielectric. The schematic measurement setup of the pulse measurement he had employed is presented in Fig. 2.12. The MOSFET is connected to an inverter circuit with the resistor load (R_L). A small constant DC bias (100mV) is applied to the resistor which subsequently works together with the channel resistor to form a voltage divider. The I_d~V_g characteristic is obtained by applying a trapezoidal (triangular) pulse to the gate and the drain voltage is subsequently recorded using a digital oscilloscope. From the measured voltage traces the I_d~V_g characteristic can be determined using

$$I_D = \frac{100mV}{V_D}\left(\frac{100mV - V_D}{R_L}\right) \tag{2.17}$$

where V_D is the measured drain voltage and R_L the resistive load of the inverter circuit.

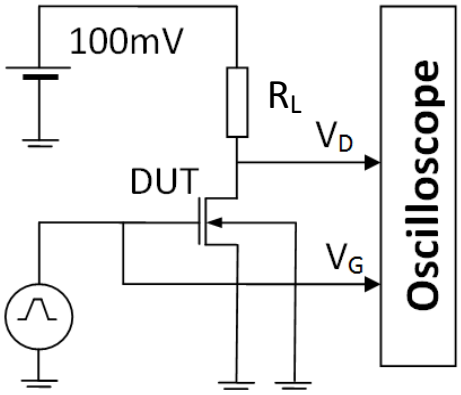


Fig 2.12 Schematic setup for the ultra fast pulse I_d-V_g technique

The use of a voltage divider can potentially cause the drain voltage to change during the measurement. This effect can be eliminated by normalizing the extracted drain current to a constant drain voltage, which is given by the term 100 mV/V_D in equation (2.17). It should be noted that this normalization is correct only when the MOSFET operates in the linear region, which limits the range for the DC bias applied to the resistor. To reduce the noise and further increase the accuracy of the measurement, this impedance along the signal path of this circuit needs to be matched and hence the resistive load should be around 50 Ω. However, this limits the gain of the circuit significantly.

An improved approach to increase the gain of the circuit while maintaining impedance matching along the signal path is to use an op-amplifier (op-amp), as presented in Fig. 2.13. The drain of the MOSFET is connected to the negative input of the op-amp. Since the voltages at the two input terminals are approximately equal when

negative feedback is present through R, the drain voltage of the MOSFET is fixed at V_d supplied by the voltage source.

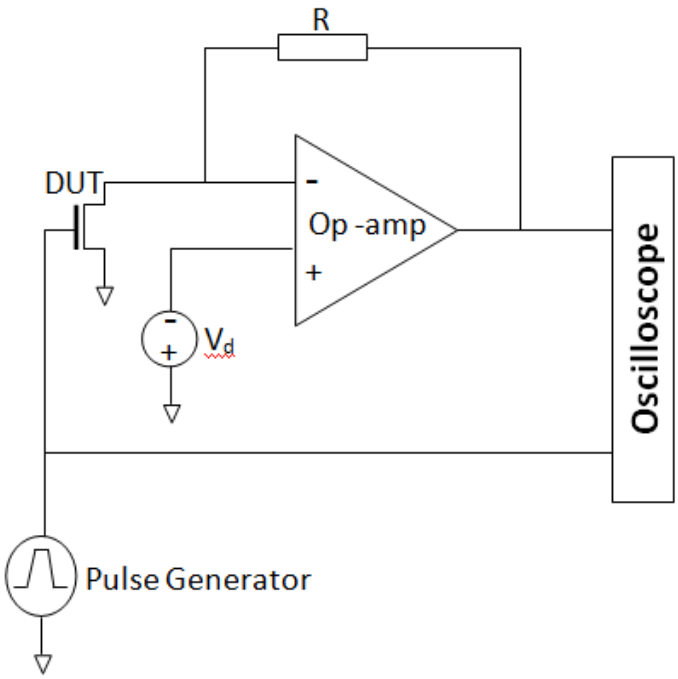


Fig.2.13 Schematic of our modified pulse I_d-V_g

Due to the fact that the input bias current of the op-amp is very low, the drain current flows almost entirely through the gain resistor, R. Resistors ranging from 1 to 10kΩ are used for different gain. The output voltage of the op-amp, in terms of the MOSFET drain current is given by the following equation:

$$V_{out} = I_d \bullet R + V_d \qquad (2.18)$$

2.5.2 Calibration of Pulse Measurement System

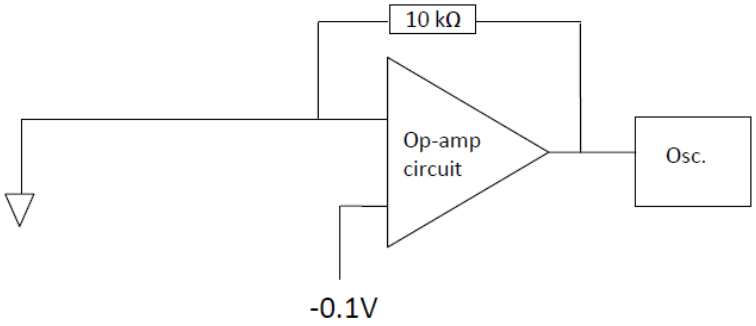
This section details the various check conducted on the pulse measurement system in ensuring that the system has an acceptable noise level and performance.

2.5.2.1 Calibration of the Op-amp Circuit

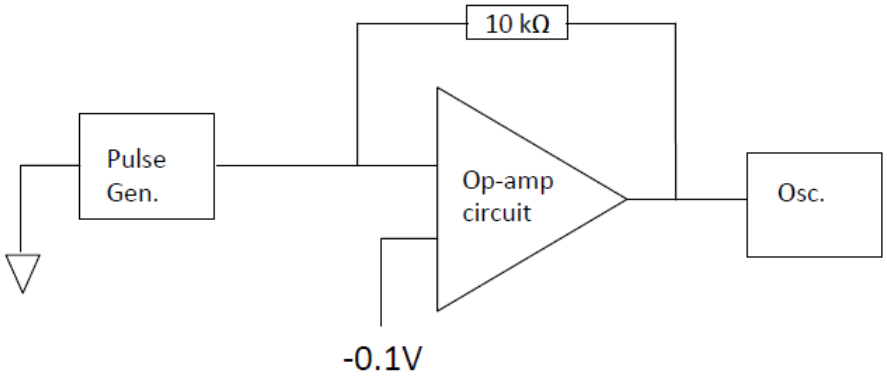
A calibration exercise was first carried out for the op-amp circuit without connecting it to the device under test. To start, the system noise without connecting the pulse generator is checked by grounding the 10 kΩ resistor in Fig. 2.14 (a). Fig. 2.14 (c) shows that the I_d can be measured with an accuracy of 0.09%, so that the noise is negligible. The response of the op-amp circuit to a pulse input is studied next. As shown in Fig. 2.14 (a), a pulse was applied to the input and the current at the pulse edge was measured. Fig. 2.14 (d) compares the measured current with different pulse edge times. The good agreement confirms that the op-amp circuit can respond in a time of a few micro-seconds.

2.5.2.2 Calibration with transistor connected

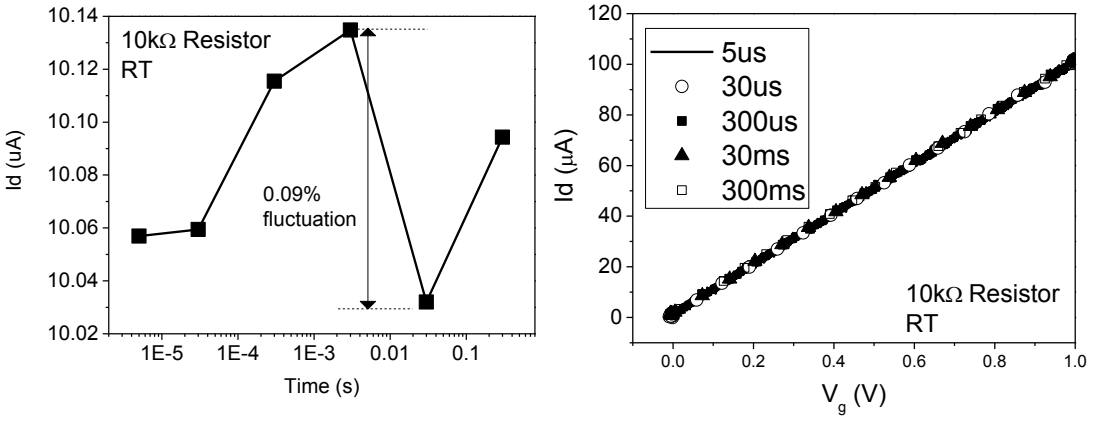
After connecting the device under test, the V_g waveform used for calibration is shown in the insets of Figs. 2.15 (a) to (c). In Fig. 2.15 (a) and (b), the I_d-V_g was measured from the pulse edges. The V_t was extracted from the V_g at a constant current in Fig. 2.15 (a) and an accuracy of 0.8 mV can be achieved when the same measurement was repeated many times. In Fig. 2.15 (b), V_t was extracted from the maximum G_m extrapolation and an accuracy of 2 mV was obtained. The V_g waveform in Fig. 2.15 (c) is different from that in Figs. 2.15 (a) and (b). The V_g was stepped to -0.41 V with an edge time of 100 ns and no measurement was made at the edge. Once V_g reached -0.41 V, I_d was measured within 1 μs and Fig. 2.15 (c) shows the accuracy is better than 0.032 μA. This accuracy is considered as acceptable.



(a)

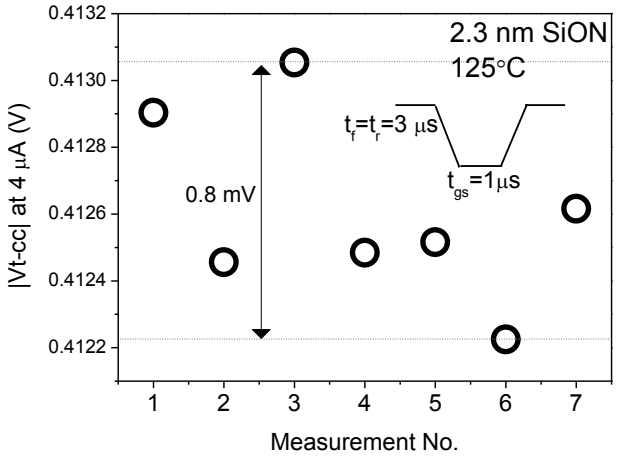


(b)

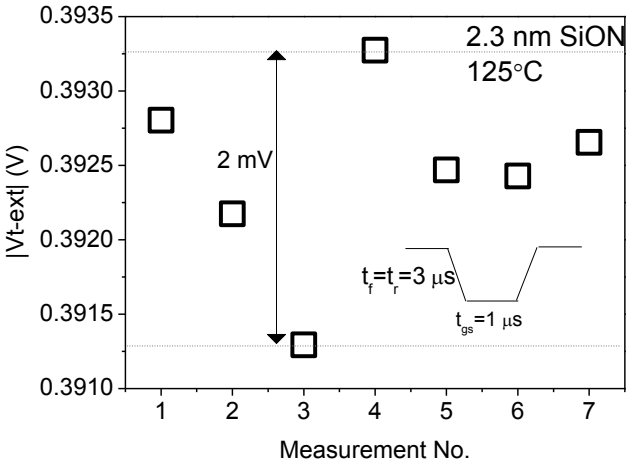


(c) (d)

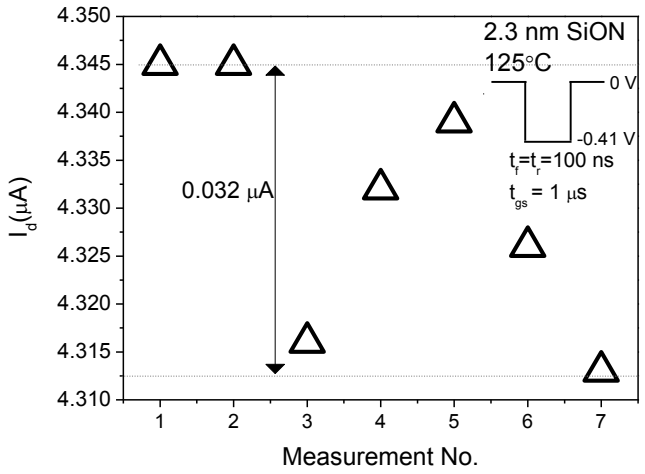
Fig. 2.14 Calibration of the op-amp circuit without connecting a MOSFET. (a) The 10 kΩ resistor is grounded. (b) The 10 kΩ resistor is connected to the pulse generator. (c) and (d) are the measured I_d obtained from configuration (a) and (b) respectively.



(a)



(b)



(c)

Fig. 2.15 Calibration of the pulse measurement with DUT connected by repeating the same measurement many times. The Vg waveforms applied are given by the insets. The threshold voltage was extracted at constant current in (a) and by gm-extrapolation method in (b). (c) shows the measurement variation at a constant voltage.

2.6 Data Smoothing Procedures

The measured TC obtained from the pulse measurement is smoothed by utilizing the smooth functions available in Curve Fitting Toolbox of MATLAB 7.0, prior to threshold voltage extraction in order to obtain better accuracy. There are several options for the smoothing of the measured data, namely implementing the moving average, or using the Savitzky-Golay filters. Alternatively there are also methods using local regression with and without weights and robustness (lowess, loess, robust-lowess and robust-loess).

Moving Average Filtering

A moving average filter smoothes the data by replacing each data point with the average of the neighbouring data points defined within the data span. This smoothing of the data is given by the difference equation given below:

$$y_s(i) = \frac{1}{2N+1}\left(y(i+N) + y(i+N-1) + ... + y(i-N)\right) \qquad (2.19)$$

Savitky-Golay Filtering

The Savitzky-Golay filter was developed in 1964 by Abraham Savitzky and Marcel J.E. Golay. The Savitky-Golay filtering can be claimed to be a generalized moving average. It is a filter that performs a local polynomial regression on a series of equally spaced data points and that the smoothed values or numerical derivatives are obtained from the coefficients of the polynomials. It can also be taken as a digital-smoothing polynomial filter or a least-squares smoothing filter due to that the filter coefficients are

derived by performing an unweighted linear least-squares fit using a polynomial of a given degree. The plots below present the smoothing conducted using the Savitky-Golay, with the span and the degree input varied accordingly.

Lowess and Loess Filtering

These smoothing methods uses locally weighted linear regression to smooth the data. Similar to the moving average method, each smoothed value is determined by neighbouring data points defined within a span. The process of smoothing the data is weighted because a regression weight function is defined for the data points contained within the span. The methods are then differentiated by the model used in the regression whereby the Lowess smoothing uses a linear polynomial while the Loess smoothing uses a quadratic polynomial.

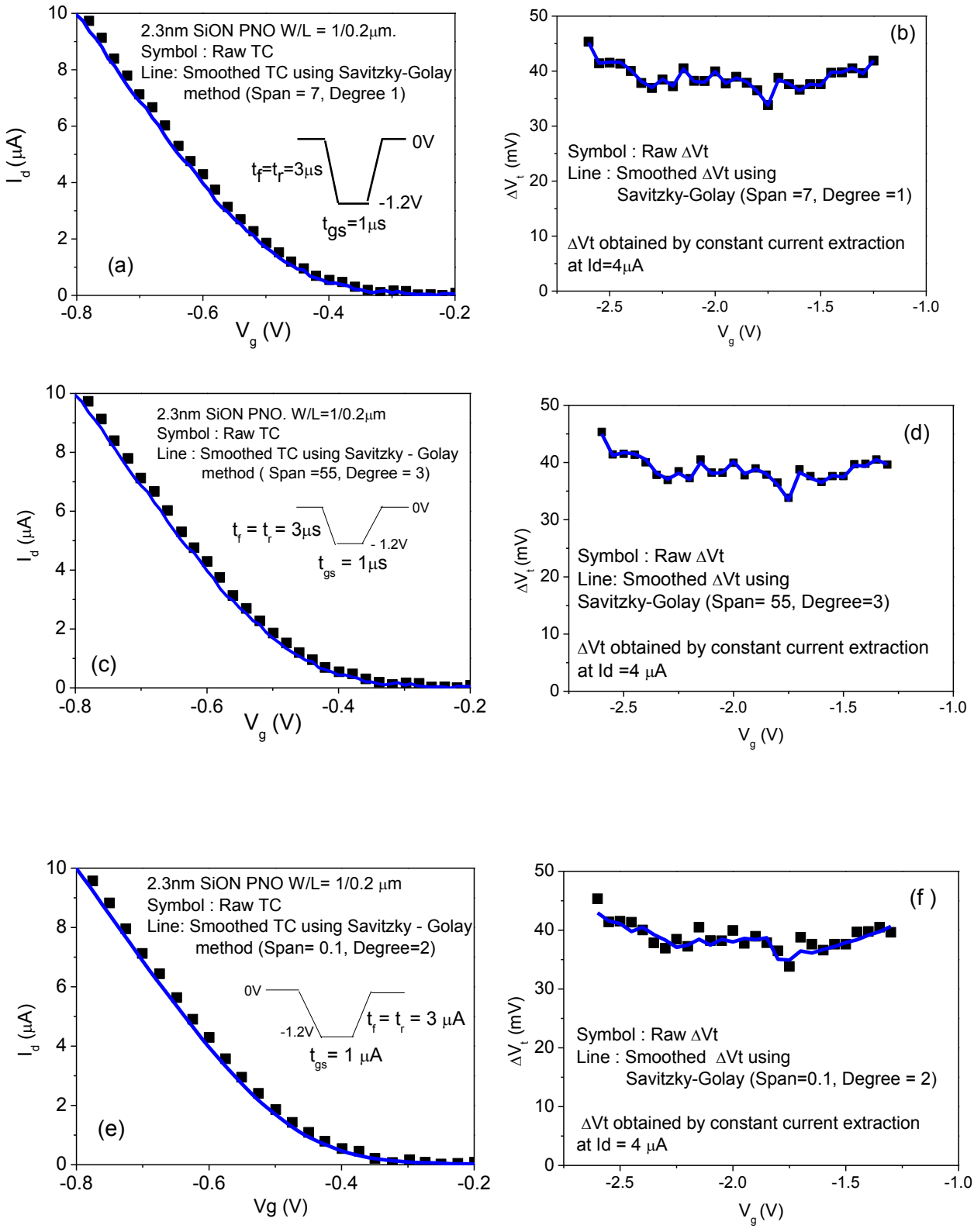


Fig. 2.16 Smoothing of data by using Savitzky- Golay method with varying smoothing variable.

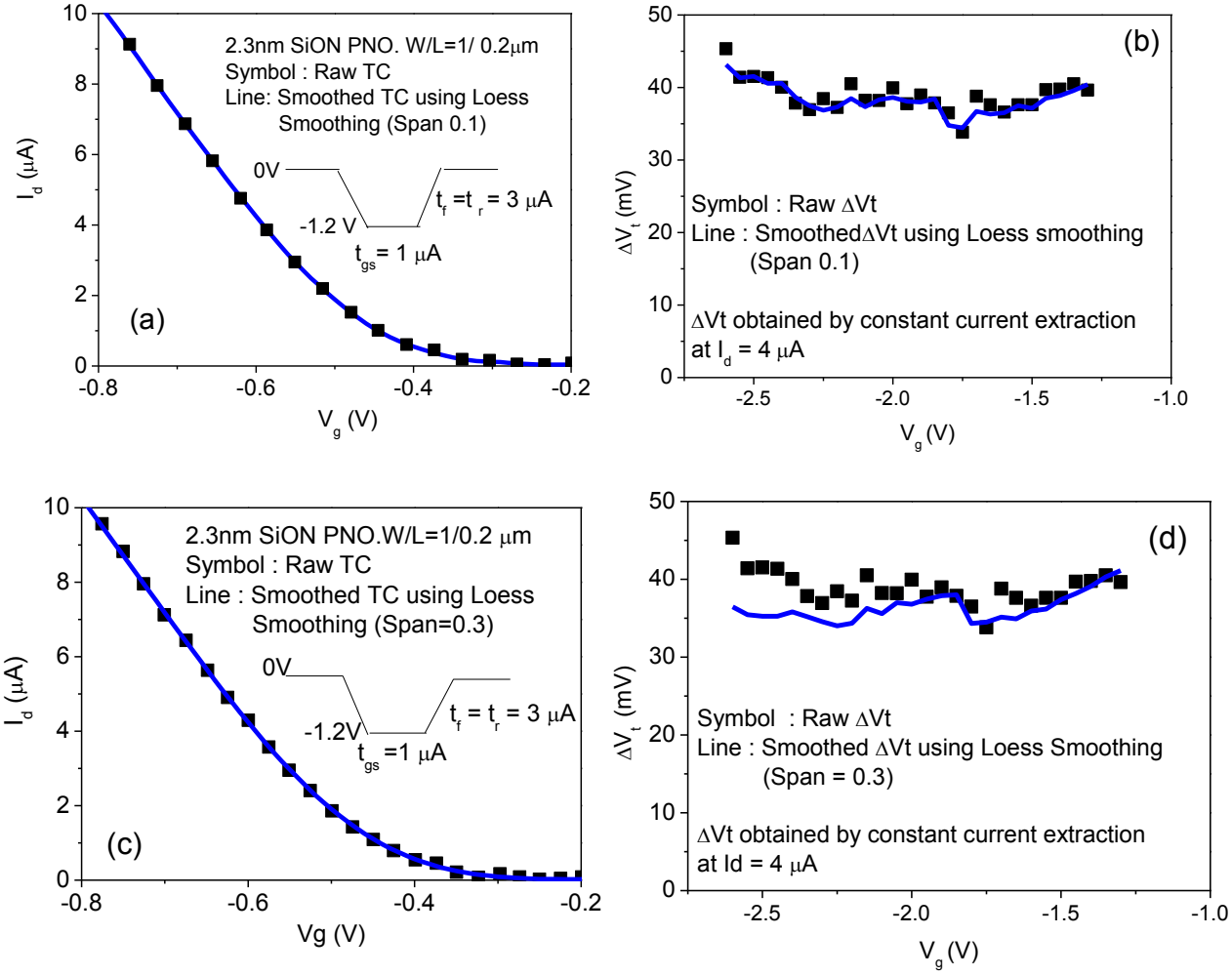


Fig. 2.17 Smoothing of data using Loess method applied with varying smoothing variable.

Selection of the smooth method

A comparison of the I_d in Figs. 2.16 (a), (c) and (e) with that in Figs. 2.17 (a) and (c) shows that the Loess Smoothing follows the raw data better. For the Loess smoothing given in Figs. 2.17 (b) and (d), it is shown that a span of 0.1 should be used. Loess smoothing applies local weighting methodology and outbounded data will not be considered in the smoothing process. Also, from the observation, smoothed result agrees very well with the measurement. As a result, the Loess smoothing with a span of 0.1 will be applied in this work.

2.7 Conclusions

In this chapter, the principles of various techniques for characterising the gate dielectrics in MOS devices are described. The different methods for extracting the threshold voltage have been demonstrated. The capacitance-voltage (C-V), charge pumping, the on-the-fly (OTF) technique and the pulse I-V measurement system have been presented and the mechanism behind the measurements has been discussed. The accuracy of the pulse I-V measurement is calibrated and the smoothing procedure is selected.

3 Energy Distribution of Positive Charges in Gate Dielectric : Probing Technique and Impacts of Different Defects

3.1 Introduction

Negative bias temperature instability (NBTI) causes time dependent device variability [103] and is a major threat to the reliability of CMOS circuits. It originates from the positive charge formation within gate dielectric and the generation of interface states [24, 104-109]. Positive charges in dielectric have been investigated since 1960s [110] and their complex behaviour has puzzled the international community ever since [76, 111-112]. Many terms were used to reflect their different properties, such as anomalous positive charges [111], slow states [112], and border traps [76]. To explain the complex dependence of PCs on biases, time, and temperature, it has been proposed that there are three different types of PCs: as-grown hole traps (AHT), cyclic positive charges (CPC), and anti-neutralization positive charges (ANPC) [80], [105], [113-115], as illustrated in Fig. 3.1. AHT has energy levels below the top edge of silicon valence band, i.e. E_v, making them easy to neutralize, but hard to charge. The as-grown hole trap (AHT) is a pre-existing defect in the device. Hence, initially, the defect is a precursor and is transformed into a positive charge when bias is applied. The defect can be

charged positively through carrier tunnelling. In contrast, ANPC has energy level above the bottom edge of silicon conduction band, i.e. E_c, making them hard to neutralize, but easy to charge. CPC is energetically located within the bandgap and can be repeatedly charged and discharged by alternating gate bias polarity. Although the above framework gives a broad picture of defect energy levels, a detailed energy distribution of PCs is still missing. This knowledge is useful for assessing the impact of PCs on circuits, since it gives PCs for each surface potential. Early works on NBTI generally divide the threshold voltage shift, ΔV_{th}, into two components: recoverable and permanent [24], [104], [116]. Recoverable component is fully charged only at the stress V_g, while permanent component is measured at either V_g=0 or a certain positive level after some time. This effectively gives PCs at only two V_g points: V_g(stress) before recovery and the V_g used for recovery. The PCs at other V_g levels are generally unknown.

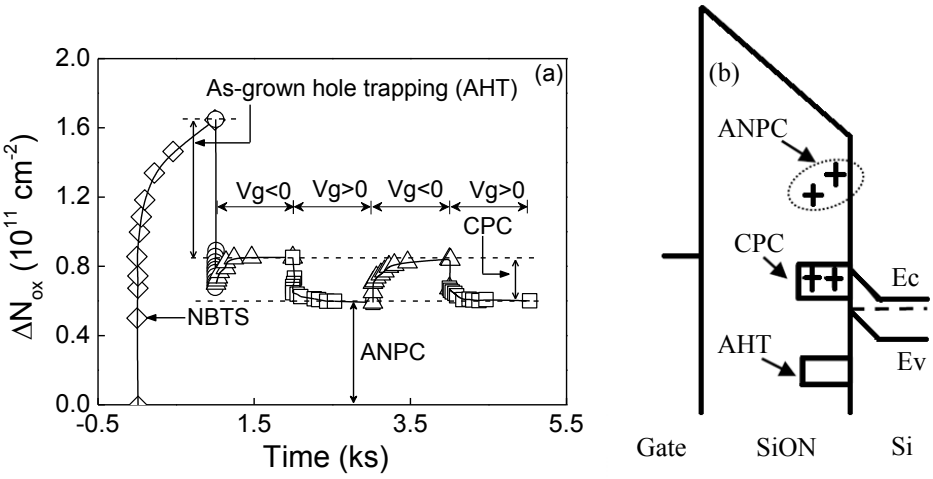


Fig. 3.1 An example of three different types of positive charges in gate dielectric. (a) shows the test sequence and (b) gives their energy location. Positive charges (PCs) built up during negative bias temperature stress (NBTS). Thereafter, when negative (V_g<0) and positive (V_g>0) gate biases with an oxide field of $|E_{ox}|$=5 MV/cm were applied alternately, some PCs can be repeatedly charged and discharged and are referred to as cyclic positive charges (CPC). Some PCs have energy level above E_c, cannot be neutralized, and are called as anti-neutralization positive charges (ANPC). Some PCs have energy level below E_v, cannot be recharged at E_{ox}= -5 MV/cm, and originates from as-grown hole traps (AHT) [115].

It has been recognized that, when $|V_g|$ reduces from the stress level, neutralization already occurs well before it reaches zero [26], [109], [117,118], indicating a continuous energy distribution of PCs, rather than an abrupt change at V_g=0. Efforts were made to extract the energy distribution of PCs [106,107], but they were based on the slow quasi-DC measurement, hence did not capture the defects that discharged rapidly. Moreover, they only provide distribution within the bandgap. As a result, two types of PCs, AHT and ANPC, were not covered by these early works [106, 107]. The central objective of this work is to develop a fast pulse technique that can evaluate the energy distribution of all types of PCs. It will be shown that PCs in different energy ranges originate from different types of defects.

3.2 Devices and Experiments

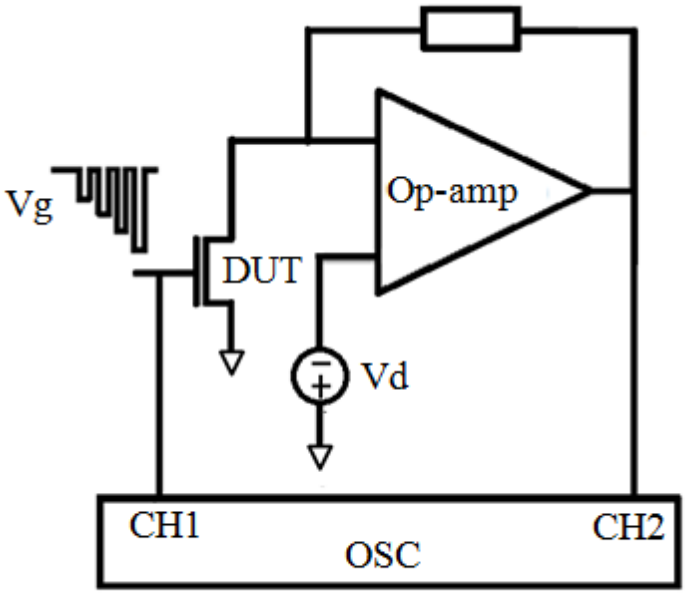
The gate dielectric of pMOSFETs used in this work includes a 2.3 nm plasma nitrided SiON and a thermally nitrided 2.7 nm SiON. For the case of the 2.3 nm plasma nitride SiON, two different devices of different nitiridation levels were investigated. The typical channel length is 0.25 μm and the width is 10 μm.

The typical V_g waveform and the circuit setup for the pulse measurement used in this work are given in Fig. 3.2. An op-amp is used to convert I_d into a voltage, which was recorded together with V_g by an oscilloscope. Care has been exercised to ensure impedance match and suppression of parasitic effects [26,119]. The 'Measurement time' is the time for a single pulse used to record one ΔV_{th} point. The 'Discharge time' is the time under a given $V_{discharge}$. While the discharge time can be hundred of seconds, the measurement time was fixed at 500 ns in this work, with the exception of Fig. 3.3.

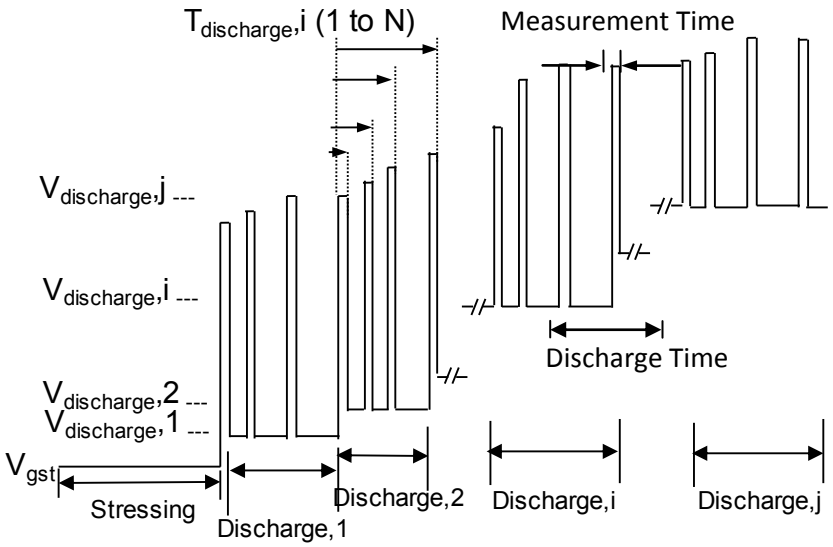
Prior to stress, a reference I_d-V_g curve was recorded by applying a V_g pulse under V_d = -25 mV. After stress at V_g=V_{gst} for a pre-specified time, Fig. 3.2 (b) shows that V_g was changed to $V_{discharge,}1$ to start the discharge and ΔV_{th} was monitored against the 'Discharge time' marked in Fig. 3.2 (b). Following the early work [120], ΔV_{th} was measured as the V_g shift at a constant sensing I_d= 100 nA×W/L. To reach this sensing I_d, a V_g pulse is used to suppress discharging during the measurement itself [104, 105]. Fig. 3.3 shows that the discharging is substantial in ms, so that quasi-DC measurement, that takes tens of ms for one point, cannot be used here [26], [105] [109],[117].

The discharging becomes negligible for a measurement time less than 10 μs and 500 ns will be used hereafter, to freeze discharging during the measurement time marked out in Fig. 3.2 (b).

As compared to the DC technique, the pulse probing technique is definitely important in this work. In principle, discharging can occur under a given $V_{discharge}$ and during the measurement itself. Hence, to measure the discharge under a given $V_{discharge}$, it is required to freeze the discharge during the measurement. If the DC measurement is used, one could not separate discharging under a given $V_{discharge}$ from that induced by the measurement itself.



(a)



(b)

Fig. 3.2 The measurement setup (a) and the V_g waveform (b). After stress, discharging occurs at $V_{discharge},1$ and PCs were periodically monitored by applying a pulse to the gate and recording the drain current at the pulse top through the op-amp, until the discharge completes. The gate bias was then changed to $V_{discharge},2$ for the next discharging phase and the same procedure is applied.

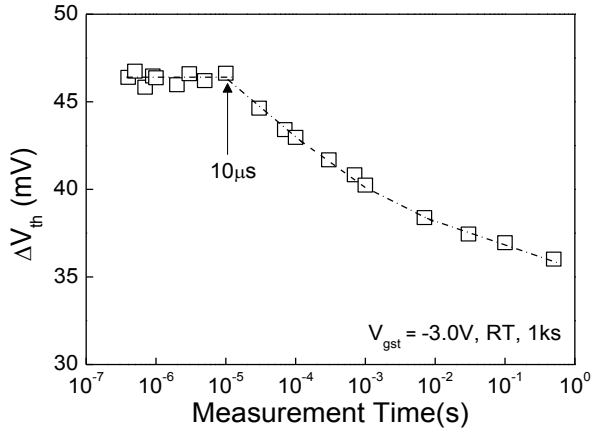


Fig. 3.3 Effect of measurement time on PCs. The measurement time is the pulse width, as marked in Fig. 2(b). When it is shorter than 10 μs, PCs are insensitive to measurement time, indicating discharge during measurement is negligible. A measurement time of 500 ns is used in this work.

ΔV_{th} was monitored under $V_{discharge,}1$ until its variation between two points becomes negligible (< 1 mV) [122, 123]. For thick oxides (e.g. > 7 nm), early work [123] shows that discharge can continue for days. Fig. 3.4 (a), however, shows that further discharge becomes insignificant beyond 500 sec, since the thin oxide (< 3 nm) used here allows efficient tunneling. After completing discharge at $|V_{discharge,}1|$, $|V_g|$ was further reduced to $|V_{discharge,}2|$ and the same procedure is followed. To capture a wide energy range, $V_{discharge}$ eventually becomes more positive than the V_g for the sensing current level and the direction of the applied pulse is changed. ΔV_{th} will not include interface states, Dit changing occupancy over discharging time, since it is widely accepted Nit will not recover. ΔV_{th} is always sensed at constant surface potential and therefore, the Vit will always the same inrespective of discharge time.

A typical discharge result under different $V_{discharge}$ is given in Fig. 3.4 (a). Fig. 3 4 (b) plots the V_g corresponding to the last point of each curve in Fig. 3.4(a), together with the reference I_d-V_g, measured on a fresh device.

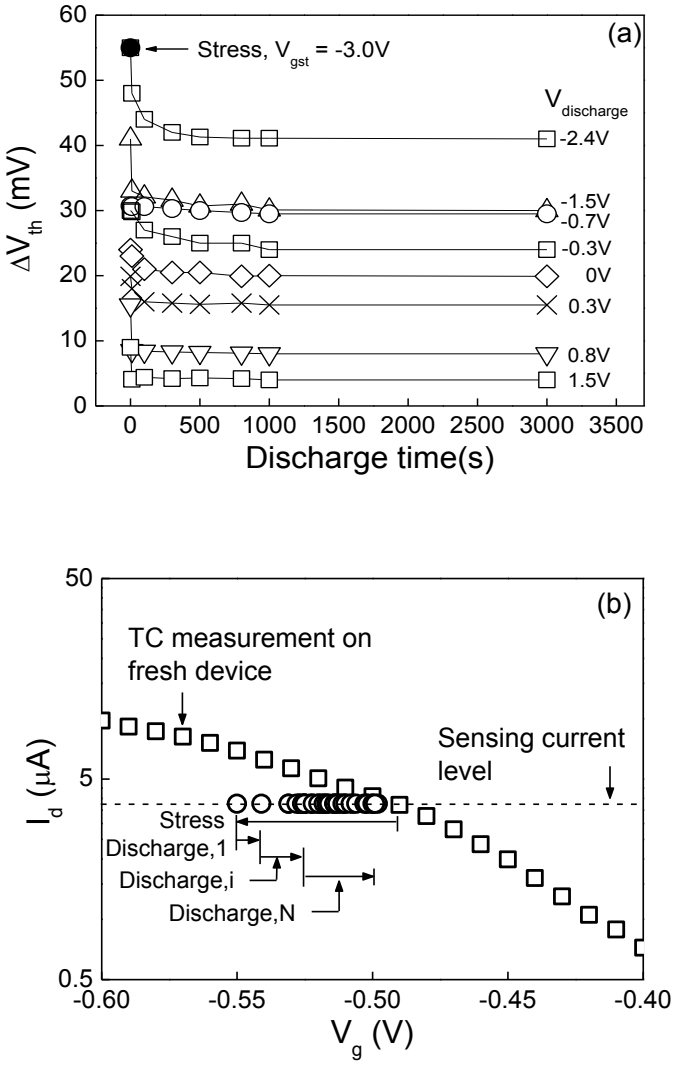


Fig. 3.4 Typical results for discharging under different $V_{discharge}$. The 'Discharge time is the time under a given $V_{discharge}$., as marked in Fig. 2(b). The device was stressed at V_{gst}=-3.0 V under room temperature for 10 ksec. The total threshold voltage shift before discharge is given by the symbol '•' in (a). The discharge under each $V_{discharge}$ was monitored against time. The V_g corresponding to the last point for each $V_{discharge}$ in (a) at the sensing I_d=4 μA was plotted as the symbol 'o' in (b), together with the reference I_d-V_g measured on a fresh device.

3.3 The Technique for Energy Distribution

Fig. 3.5 illustrates the principle of extracting the energy distribution of PCs. As a first order approximation, we follow early works [106, 107,124] by assuming that below the Fermi level, E_f, PCs will be neutralized throughout the thin oxide, if a sufficient discharge time is applied. As $V_{discharge}$ increases towards positive for each discharging step, the energy level of PCs is lowered against the substrate, thus accommodating a new shaded region below E_f for discharging, as shown in Fig. 3.5. By varying $V_{discharge}$ over a sufficiently large range, one can sweep E_f over a wide energy range at the interface, including the region beyond bandgap of silicon.

Although PCs within SiON is located close to the substrate [57], [125], we point out that they are not exactly at the Si/SiON interface and have a spatial distribution across the dielectric. As illustrated in Fig. 3.5, the vertical depth of the shaded area increases towards gate, namely, namely $\Delta\Phi_g > \Delta\Phi_s$. For device and circuit simulation, it is useful to know the change in PCs for a given Φ_s. Hence, the energy level at the Si/SiON interface is used for their energy distribution. The energy distribution defined in this way is considered as an "effective" energy distribution, as describe in details below.

On the spatial distribution of PCs, early works gave a rule-of-the-thumb estimation: the IBM group used photo-IV and found that the centroid of positive charges is within 5 nm from the interface [125]. There are also early works that reports that the positive charges are dominated by the interfacial layer [57],[126]. Hence it had been concluded that PCs pile up towards the substrate interface. A trap spectroscopy by charge injection and sensing (TSCIS) technique was developed to profile the energy and spatial

distribution of pre-existing electron traps in dielectric [127]. TSCIS requires a full discharge, which cannot be achieved for PCs under typical test conditions [116]. It also requires that there is no generation of new traps, but new defects are created during NBTI tests [26, 109]. As a result, TSCIS cannot be used to probe the spatial and energy distribution of PCs in dielectrics.

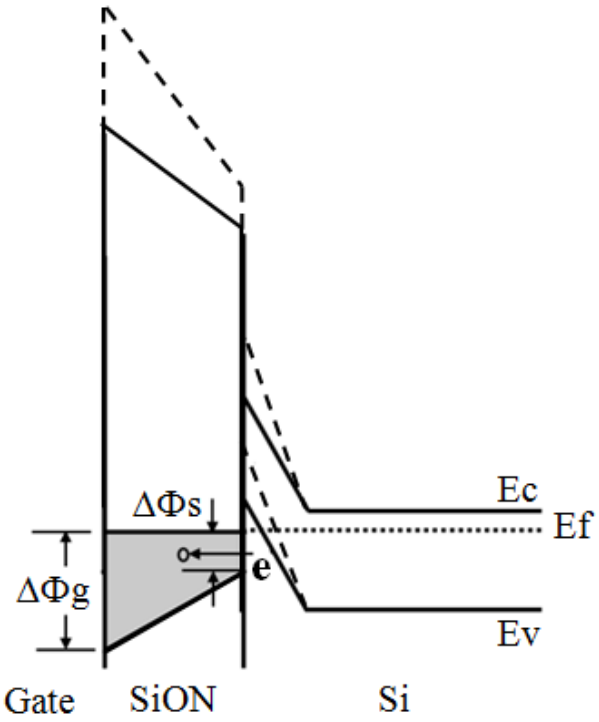


Fig. 3.5 An illustration of the principle for extracting the energy distribution of PCs. When $V_{discharge}$ was stepped towards positive direction each time, a shaded area with an energy depth of ΔΦs at the interface falls below E_f and the PCs within it start discharging. By sweeping $V_{discharge}$ from the negative stress level to positive, E_f can be driven from below E_v to above E_c.

When the spatial distribution of PCs is not known, the concept of 'effective charge density' often was used by early works [128,129]. If there is a sheet of positive charges with an area density of N at a distance of X from the gate interface, it will induce a threshold voltage shift of $\Delta V_{th} = - qNX/C_{ox}X_{ox}$, where C_{ox} and X_{ox} is oxide and

thickness capacitance. The "effective charge density", N_{eff} is the equivalent charge at the oxide/substrate interface that induces the same $\Delta V_{th,}$, as expressed below:

$$\Delta V_{th} = \frac{-qNX}{C_{ox}X_{ox}} = \frac{-qN_{eff} \times X_{ox}}{C_{ox}X_{ox}} = \frac{-qN_{eff}}{C_{ox}} \tag{3.1}$$

PCs were assumed to be at the substrate interface and the effective density is required to induce the same ΔV_{th} as that by the real spatially distributed PCs. Since the spatial distribution of the defects is unknown, an effective charge density is obtained by assuming the charge centroid is at the SiO_2/Si or SiON/Si interface, to comply with previous work in this area. This work also uses the effective density for PCs and assumes that they are neutral below E_f. It indeed will take time for a trapped charge away from the substrate to discharge. If the discharge time in the measurement is too short, it may not be able to discharge, thus invalidating the proposition that positive charges are neutral below E_f. Hence, to ensure that this proposition is valid, the discharge time under a given $V_{discharge}$ must be sufficiently long so that the PCs in dielectric below E_f can discharge, as shown in Fig. 3.4 (a) and the ΔV_{th} saturates. The energy distribution extracted in this way is referred to as "effective" energy distribution.

ΔV_{th} after completing discharge for 1000 sec under each $V_{discharge}$ is converted to the effective charge density, i.e. ΔN_{ox}, and plotted against $V_{discharge}$ in Fig. 3.6 (a). From the ΔV_{th} plotted against the discharging time in Fig. 3.4 (a), it is observed that the discharging has saturated at discharging time of 1000 sec. To obtain the energy density of PCs, a differentiation of these data will be needed. To improve the accuracy of differentiation, the data were smoothed by a robust-loess method which uses locally

weighted linear regression [130]. Care has been exercised for not altering the character of the data by the smoothing treatment, as shown in Fig. 3.6 (a).

The next task is to convert $V_{discharge}$ to the energy level of E_f with respect to E_v at the Si/SiON interface, i.e. (E_f-E_v). As shown by the inset of Fig. 3.6 (b), the difference between E_f and E_v at the interface is evaluated from, E_f-E_v = $E_g/2 + \Phi_B - \Phi_S$, where E_g is the Si bandgap. Fig. 3.6 (b) was obtained in three steps. First, following early works [122,131], the dashed curve was calculated from the CVC simulator [132] by inputting the equivalent oxide thickness, doping densities for both substrate and gate. Second, the V_{th} of stressed sample was evaluated from $V_{th}=V_{th0}+\Delta V_{th}$, where ΔV_{th} is the measured shift and V_{th0} is the threshold voltage of fresh sample and was obtained by the maximum transconductance extrapolation method [105, 109]. Third, to align the theoretical curve with the test sample, the solid curve was obtained by shifting the dashed curve to the left until E_f-E_v=$E_g/2$-Φ_B occurs at $V_g=V_{th}$.

Fig. 3.6 (c) plots ΔN_{ox} against E_f-E_v, converted from the $V_{discharge}$ in Fig. 3.6 (a). The measurement is taken after 1000 sec discharging and therefore the static ΔV_{th} is measured which is assumed to be corresponding to the energy level aligned to E_f. In principle, the Fermi Direct statistics may have a smearing effect. However, by changing the V_g step interval, the measured distribution is kept the same.

A differentiation of ΔN_{ox} against E_f-E_v gives the energy density, i.e. ΔD_{ox}, as shown in Fig. 3.6 (d). In the evaluation above, the impact of interface states has not been taken into account. The stress-induced interface states, i.e. ΔD_{it}, affect the measurement in

two ways: (i) they contribute to the ΔN_{ox} measured at a constant I_d and (ii) they affect the $V_g \sim (E_f\text{-}E_v)$ relation and, in turn, the evaluation of ΔD_{ox}.

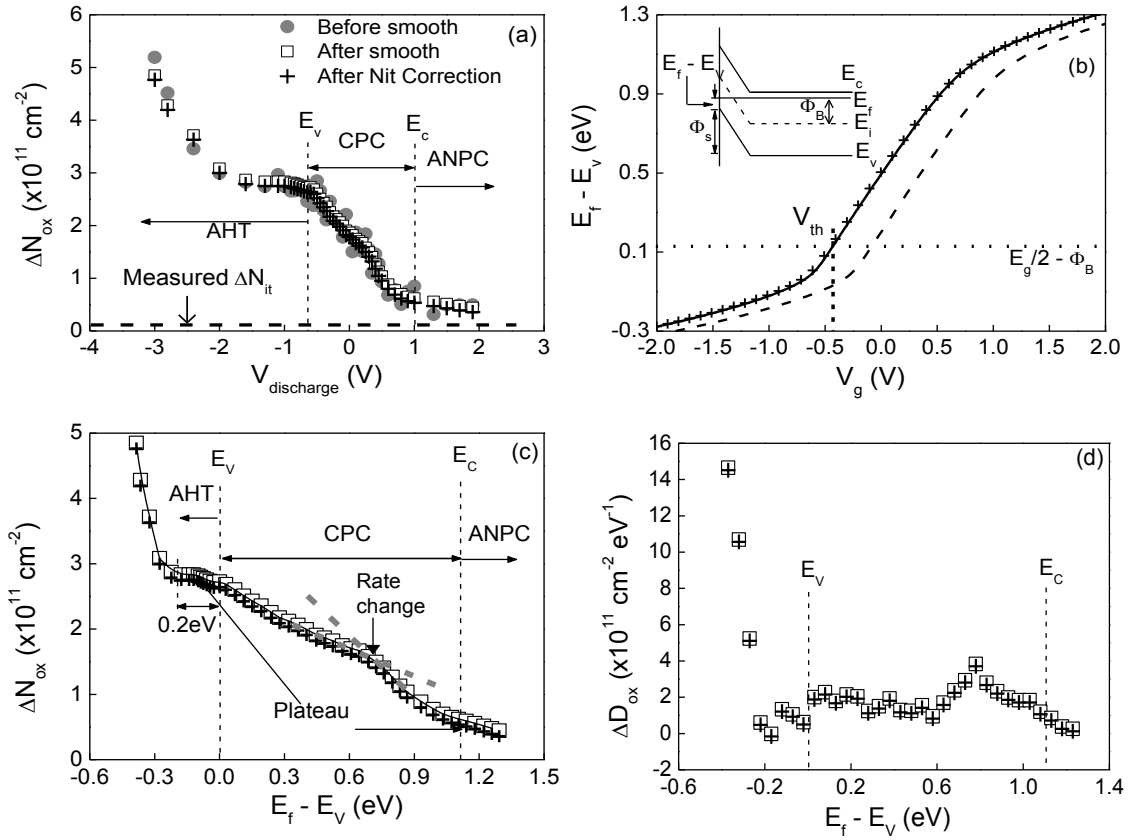


Fig. 3.6 Energy profile extraction for PCs. ΔN_{ox} in (a) is the PCs after competing discharge at each $V_{discharge}$. The device was stressed at V_g=-3.0 V under room temperature for 10 ksec. The '□' was obtained by smoothing raw data. In (b), the dashed curve is the theoretical value. The Φ_s was calculated from the CVC simulator [32] and then converted into $E_f\text{-}E_v=E_g/2+\Phi_B\text{-}\Phi_s$, as illustrated by the inset. The solid line is obtained by aligning $E_f\text{-}E_v=E_g/2\text{-}\Phi_B$ to the threshold voltage of stressed device. (c) is obtained after converting $V_{discharge}$ to $E_f\text{-}E_v$ by using (b). The ΔD_{ox} in (d) is the energy density of PCs, obtained by differentiating the data in (c). The symbol '+' denotes the data after ΔN_{it} contribution is corrected. The device used was a 2.3 nm plasma nitride SiON.

It has been reported that ΔN_{it} will not recover [24] and we also found that ΔN_{it} changes little after stress and during our measurements, as shown in Fig. 3.6 (a). At a given surface potential and I_d, an unchanging ΔN_{it} will give a fixed level of charging. In this technique, ΔV_{th} always was probed at a constant I_d, so that charges from ΔN_{it} can be treated as 'fixed charges', which contribute to the measured ΔN_{ox} by introducing an offset, as shown in Fig. 3.6 (a). The energy distribution of ΔN_{it} is not included in that of

ΔN_{ox}. To quantitatively estimate this offset, we measured $\Delta D_{it} = 2.1 \times 10^{10}$ cm $^{-2}$ eV^{-1} by using the charge pumping technique. The ΔV_{th} was sensed at the constant I_d corresponding to 0.4 eV below the midgap at the interface, i.e. E_i. The PCs from ΔD_{it} at this energy level can be estimated as $\Delta N_{it}=\Delta D_{it}*0.4 = 8.5 \times 10^9$ cm^{-2}, as shown in Fig. 3.6 (a). The contribution of ΔD_{it} to ΔN_{ox} will be removed hereafter. A demonstration in the later section of this chapter will purposely show that the measurement technique will not significantly be affected even at a substantially higher $\Delta N_{it.}$ Although the PCs from ΔD_{it} can be treated as 'fixed' when correcting the ΔN_{ox} measured at a constant I_d, they can no longer be treated as 'fixed' when evaluating ΔD_{ox}. To sweep the surface potential, different $V_{discharge}$ must be used and the PCs from ΔD_{it} change with $V_{discharge}$. The charge neutrality level of interface states is E_i and ΔD_{it} is acceptor-like above E_i and donor-like below E_i [134,135]. As a first-order approximation, we assume a uniform distribution of ΔD_{it} in the bandgap. The charges from interface states are calculated from $Q_{it}=q\Delta D_{it}$ x (E_i-E_f) and it shifts V_g by $\Delta V = -Q_{it}/C_{ox}$.

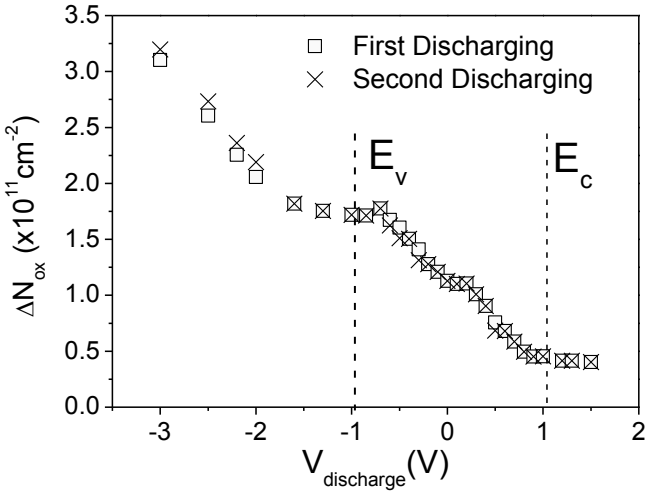


Fig 3.7. A fresh device was stressed at -3.0 V and room temperature for 100 s, followed by the first discharging measurement. The results are denoted by the symbol '□'. The test was then repeated on the same device and the data for the second discharge phase is represented by the '×'. Since the differences between the two sets of data are small, the additional degradation during discharging phase must be insignificant, when compared with the degradation during the stress phase.

The $(E_f\text{-}E_v)\sim V_g$ with and without introducing ΔD_{it} = 2.1 $\times 10^{10}$ $cm^{-2}eV^{-1}$ are compared in Fig. 3.6 (b). Based on each $(E_f\text{-}E_v)\sim V_g$, $\Delta N_{ox}\sim V_g$ can be converted to $\Delta N_{ox}\sim(E_f\text{-}E_v)$ in Fig. 3.6 (c), leading to the evaluation of ΔD_{ox} in Fig. 3.6 (d). The impact on ΔD_{ox} is insignificant and is taken into account in this work. To apply this technique, one would like to keep further degradation during the discharge phase negligible, when compared with that during the preceding stress by using sufficiently long stress time. This is the case for a stress time of 100 sec or longer, since the difference in results is found being insignificant when repeating the test twice in Fig. 3.7.

3.4 Results and Discussion

3.4.1 Different Types of PCs

The first impression of the energy distribution in Figs. 3.6 (a) and 3.6 (c) is that the PCs are sensitive to energy level and vary substantially over the energy range. This energy dependence should be taken into account when assessing the impact of PCs on devices and circuits.

A closer observation of Figs. 3.6 (a) and 3.6 (c) indicates that PCs behave differently in different regions of V_g and $E_f\text{-}E_v$. When E_f is below E_v, i.e. $E_f\text{-}E_v<0$, Fig. 3.6 (c) shows that ΔN_{ox} initially drops quickly, giving a high energy density, ΔD_{ox}, as shown in Fig. 3.6 (d) which is in the order of 10^{12} cm^{-2} eV^{-1}. The declining of ΔN_{ox}, however, slows down abruptly around $E_f\text{-}E_v$=-0.2 eV, creating a plateau before reaching E_v, where ΔD_{ox} becomes insignificant. Early works [105, 115] propose that the PCs

below E_v are as-grown hole traps (AHT) that are easiest to discharge, in agreement with the observed rapid discharging in $(E_f\text{-}E_v) < -0.2$ eV. Further support will be given in the next section to confirm that PCs below E_v are indeed as-grown hole traps.

Once above E_v, i.e. $E_f\text{-}E_v>0$, Fig. 3.6 (c) shows that the PCs start to decrease again, although at a rate substantially lower than that for $(E_f\text{-}E_v) < -0.2$ eV. Above midband, however, a rate increase can be observed, leading to a ΔD_{ox} peak around E_v+0.8 eV, in agreement with early work [124]. Although ΔD_{ox} within the bandgap is substantially lower than that in $(E_f\text{-}E_v) < -0.2$ eV, Fig. 3.6 (c) shows that the total PCs within the bandgap is comparable with the PCs below E_v When E_f reaches beyond E_c, Figs. 3.6 (a) and 3.6 (c) show that the decrease become insignificant, resulting in the second plateau. Early works [105,113,115] propose that these PCs are anti-neutralization positive charges (ANPC), whose energy levels are above E_c and thus making them beyond the reach of electrons for neutralization (see Fig. 3.1(b)). Further results will be given to confirm this.

3.4.2 Effects of stress time

The energy distributions of PCs after different stress times are compared in Figs. 3.8 (a) to (c). To separate the as-grown hole traps from those created during stress, one test was carried out by minimizing stress time. Unlike the typical procedure shown in Fig. 3.2 (b) in which the device was stressed at V_{gst} and then discharged by sweeping $V_{discharge}$ towards positive, a fresh device was used in this test and V_g was swept from positive towards negative with a time of only 1s at each point. It has been reported that generation is negligible within 1s and the filling of as-grown hole traps (AHT)

dominates during this time [26]. The result is shown as the symbol '▲' in Figs. 3.8 (a) to (c). PCs are negligible both within bandgap and above E_c in a fresh device. However, once below E_v, Fig. 3.8 (b) shows that ΔN_{ox} becomes substantial and its ΔD_{ox} in Fig. 3.8 (c) agrees with that after stress, supporting that the PCs below E_v originates from AHT. As stress time increases, both ANPC and the PCs within bandgap increase. Fig. 3.8 (c) shows that PCs within bandgap saturate after long stress time, in agreement with the saturation of creation of cyclic positive charges (CPC) [80, 105, 113].

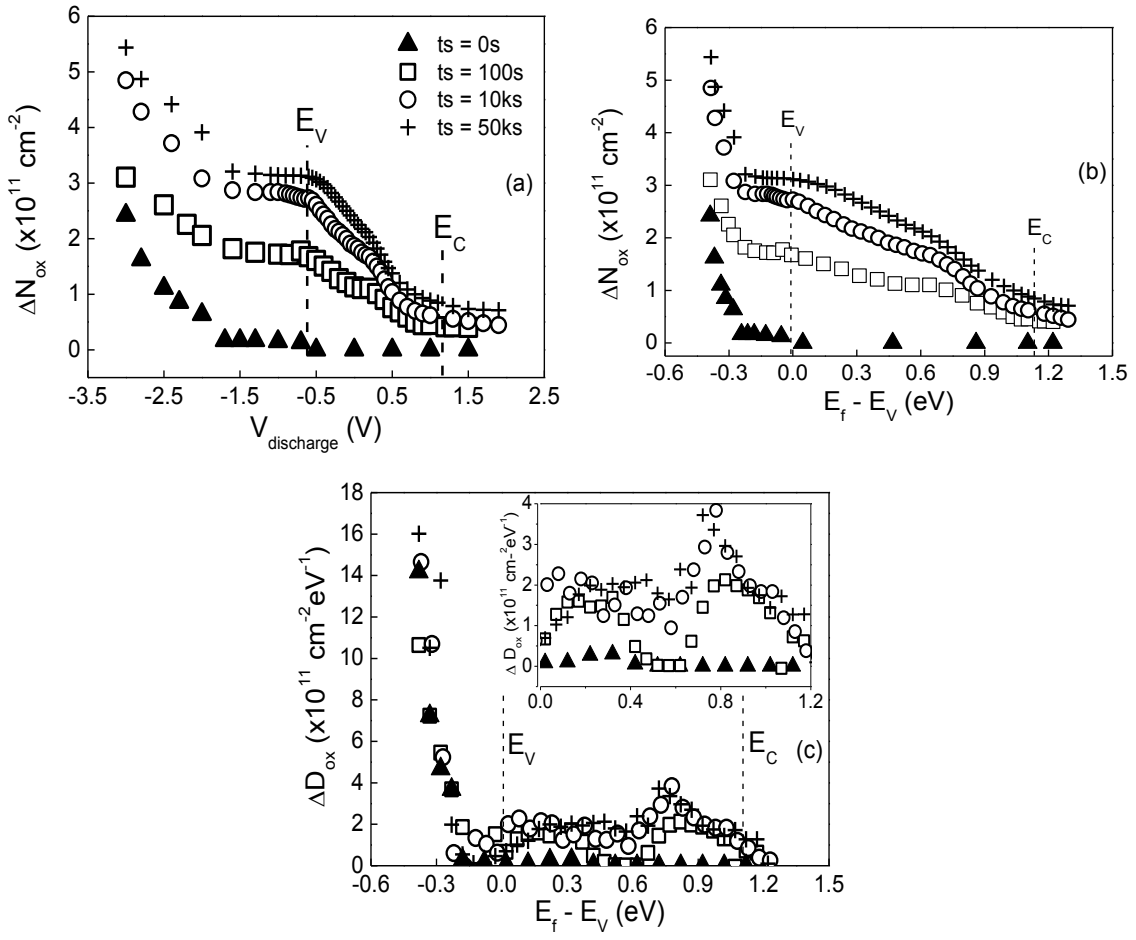


Fig. 3.8 Dependence of the energy distribution of PCs on stress time. Four devices were stressed at -3.0 V and room temperature for different time. The symbols '▲' were obtained on a fresh device and V_g was swept from +1.5 V towards negative and the time for each point is only 1 sec to minimize defect generation. All other symbols were measured by sweeping from stress V_g towards positive. The ΔN_{ox} was plotted against V_g in (a) and E_f-E_v in (b) and it increases with stress time. The ΔD_{ox} in (c) shows that the PCs below E_v actually do not increase with stress time, PCs within bandgap increase initially and then saturate.

3.4.3 Impacts of stress temperature

The results reported in Fig. 3.8 were obtained by stresses at room temperature. Figs. 3.9 (a) and (b) give the result after stressing at different temperatures.

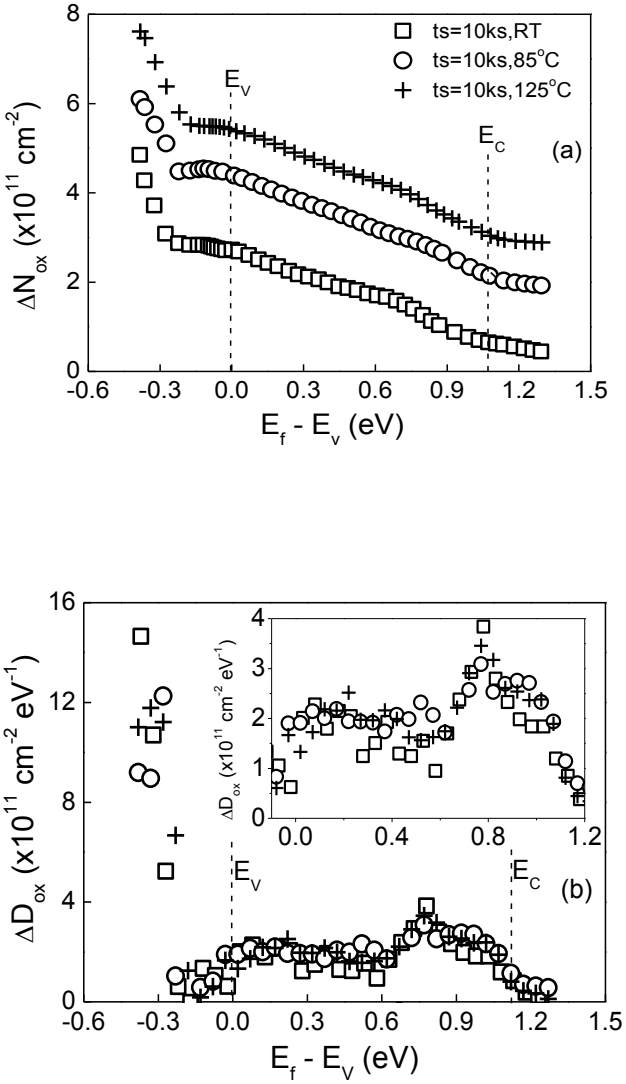


Fig. 3.9 Effects of stress temperature on the energy distribution of PCs. Devices were stressed at -3.0 V at different temperatures for 10 ksec. After cooling down to room temperature, V_g=-3.0 V was applied for another 10 ksec to recharge the neutralized defects during cooling. All data were measured at room temperature. (a) shows that higher stress temperature mainly lifts ΔN_{ox} up by increasing ANPC above E_c. (b) confirms that ΔD_{ox} is insensitive to stress temperature both within bandgap and below E_v for a stress time of 10 ksec.

As expected, AHT below E_v are the same for different stress temperature, since they are pre-existing defects. The PCs within bandgap are also insensitive to stress temperature, because the applied stress time of 10 ksec is long enough for them to reach saturation even at room temperature. Their number of precursors must be fixed and does not increase with stress temperature.

In contrast, the ANPC above the measurable energy level of $(E_f\text{-}E_v) \sim 1.3$ eV is substantially more for higher temperature and no saturation was observed [80, 105, 113]. This different dependence on temperature strongly supports that ANPC is a different type of defects. Since its generation does not saturate, its relative importance will increase for longer stress time. To avoid dielectric breakdown, tests were stopped at $V_{discharge}$= +2 V where $(E_f\text{-}E_v) \sim 1.3$ eV and ANPC remains charged. As a result, the ΔD_{ox} of ANPC were not covered in Fig. 3.9 (b). The operation voltage for modern CMOS technologies is less than 2 V and the ANPC can be considered as 'fixed charges'. In this ecperiment, it is assumed that there is negligible impact of the electron trapping.

3.4.4 Effects of nitridation technique

It has been reported that the nitridation technique has a large impact on NBTI [24,108] and we compare the energy distribution of plasma and thermally nitrided SiON in Fig. 3.10. When compared with plasma nitrided SiON, the thermal nitridation has different impacts on different types of defects in their respective energy regions. In the region above E_v, thermal nitridation increases AHTs substantially and its ΔD_{ox} is over three times of that for the plasma SiON. AHTs are the least stable PCs and contribute to

the recoverable components strongly. As a result, the large AHTs in thermal SiON observed here is in agreement with the early reports that thermal SiON generally has higher recoverable components than plasma SiON [24,108]. Two plasma SiON with different nitridation levels in Fig. 3.10 were compared, along with the results of a thermal SiON. Unlike the thermal SiON, the PCs in these two plasma SiON do not cross over in Fig. 3.10 (a) and a reduction of nitrogen level leads to smaller PCs over the whole energy range.

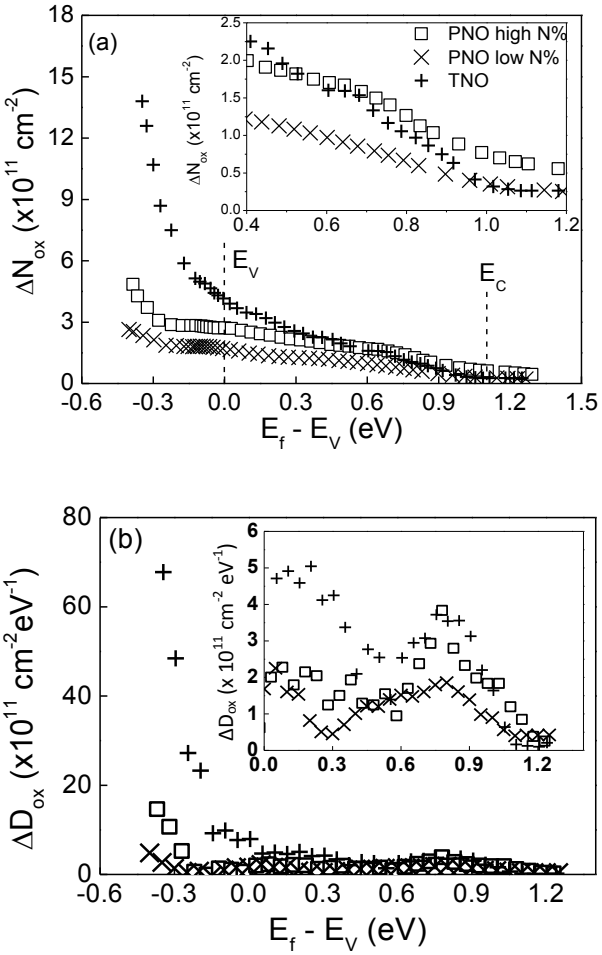


Fig. 3.10 Impacts of nitridation on the energy distribution of PCs. The samples were stressed at the same oxide field of 11 MV/cm at room temperature for 10 ksec. The thermally nitrided SiON has significantly higher as-grown hole traps (AHT) below E_v. Its ΔD_{ox} in the lower half of bandgap is also higher, but can be similar to that of plasma SiON in the upper half. The ANPC above E_c is not higher than the plasma SiON used here. For plasma nitrided SiON, an increase of nitridation leads to an increase of PCs over the whole energy range.

Despite the substantially higher PCs below E_v, Fig. 3.10 (a) show that PCs reduce at a faster rate for thermal SiON, eventually crossing over the plasma SiON of high nitridation and resulting in a lower ANPC above E_c. Fig. 3.10 (b) shows that ΔD_{ox} of thermal SiON is higher in the lower half of the bandgap, but similar to that of plasma SiON in the upper half. The permanent component of thermal SiON is not always higher than that of plasma SiON, therefore.

3.4.5 Effects of interface states

Fig.3.6 in section 3.3 demonstrates the technique applied to a case when the level of generated interface states is low. This section presents the results when the technique is applied to a case when the ΔD_{it} is more significant.

A device was stressed heavily to generate a $\Delta D_{it} = 2.8 \times 10^{11}$ $cm^{-2}eV^{-1}$, as shown in Fig.3.11, which is one order of magnitude higher that that presented in section 3.3 of this chapter. Under this stress condition, Fig. 3.12 shows that $\Delta V_{th} > 88$ mV in the whole energy spectrum. Since device lifetime is often defined as the time for ΔV_{th} to reach 50 mV, one can consider that this ΔD_{it} as significant. By applying the correction method demonstrated in the section 3.3, Fig. 3.13 (a) shows the impact of ΔD_{it} on ΔN_{ox}. Figs. 3.13 (b) and (c) show that the effects of ΔD_{it} on E_f -E_v and ΔD_{ox} are insignificant. Since the contribution of ΔD_{it} to ΔN_{ox} and ΔD_{ox} can be corrected and its effect on ΔD_{ox} is small anyway, the proposed technique is generally applicable to NBTI stresses, rather than limited to the stress condition of $\Delta D_{it} \sim 10^{10}$ $cm^{-2}eV^{-1}$.

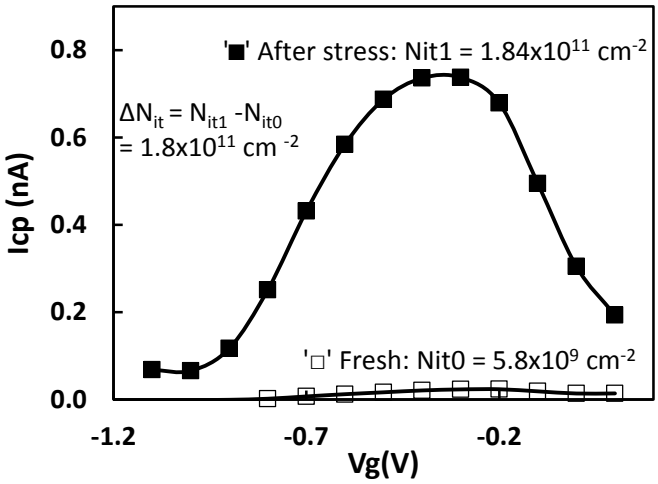


Fig. 3.11 Charge pumping measurement which measures the interface states of a heavily stressed device. The generated interface states is ΔD_{it} =2.8 x10^{11} $cm^{-2}eV^{-1}$

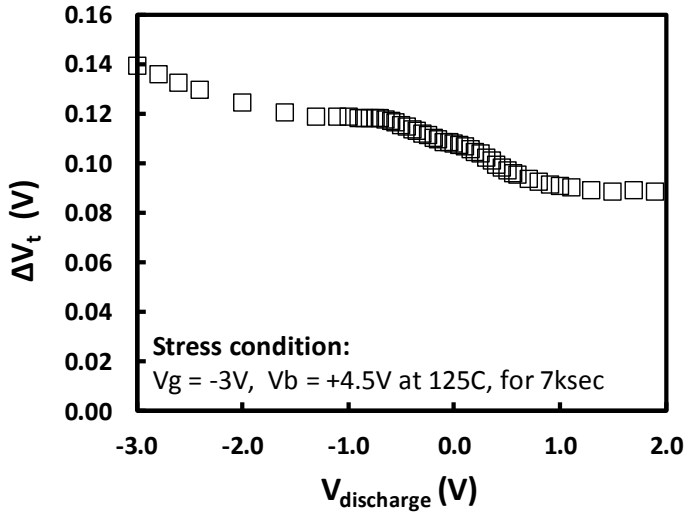


Fig. 3.12 Threshold voltage shift after heavy stress which consequently generated interface states of ΔD_{it} =2.8 x10^{11} $cm^{-2}eV^{-1}$

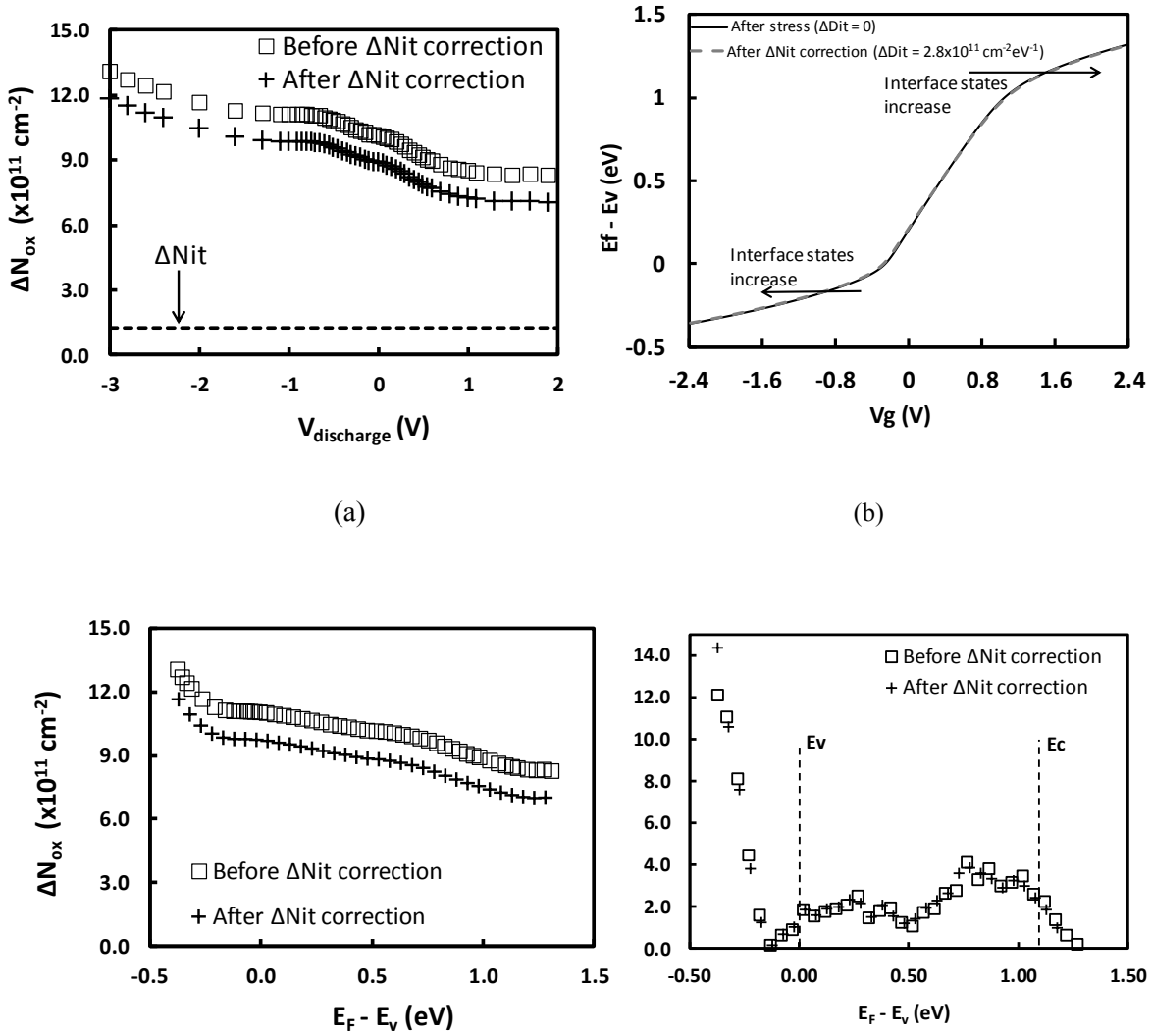


Fig. 3.13 (a) The impact of the generated interface states on the ΔN_{ox} measured at a constant I_d, respectively. (b) shows the impact on the V_g versus (E_f-E_v) relation. (c) shows the impact of the interface states on ΔN_{ox} versus (E_f-E_v). (d) shows the impact on ΔD_{ox}.

3.5 Conclusion

A fast pulse technique has been developed and, for the first time, the energy distribution of positive charges (PCs) in the dielectric is extracted both within and beyond Si bandgap. It is found that these PCs have a broad energy distribution and should be taken into account when assessing their impacts on devices and circuits. The results strongly support the existence of different types of PCs and each of them dominates different energy regions. The PCs below E_v originate from as-grown hole traps (AHT), whose energy density increases rapidly once (E_f-E_v) is below -0.2 eV and becomes significantly higher than the density within the bandgap. The AHT does not increase with either stress time or temperature and its energy density in a thermal SiON can be three times of that in a plasma SiON. The PCs distributed within the bandgap have a clear peak around Ev+0.8 eV. They are created by stresses and can reach saturation after long stress time. Moreover, the saturation level is independent of stress temperature. In contrast, the anti-neutralization positive charge (ANPC) above E_c is the only type of PCs, whose creation does not saturate with stress time and is substantially enhanced for higher stress temperature. This non-saturation characteristic increases its relative importance as stress continues.

4 Application of Energy Probing Technique on High-k devices

4.1 Introduction

Deep-sub-micron device scaling is rapidly evolving and requires stringent control of short-channel effects (SCE) and sub-threshold behaviour. With this in mind, the gate dielectrics should be thinned to less than 1.0–1.5 nm equivalent oxide thickness (EOT) [136]. It has been reported that due to quantum mechanical tunnelling, the typical leakage current of SiO_2 at gate voltage, Vg of 1 V can increase from 10^{-12} A/cm^2 with EOT of 3.5 nm to 10 A/cm^2 with EOT of 1.5 nm [137]. To achieve the EOT target and to counter the issue of leakage currents, dielectric materials with higher permittivity as compared to SiO_2 ($k \approx 3.9$) are introduced. Compounds of hafnium (Hf), zirconium (Zr), and aluminium (Al) were proposed as potential high-k dielectric materials and hafnium-based dielectrics have emerged as the winner to replace the conventional SiO_2 due to its high dielectric constant (k up to 25), wide bandgap (5.7 eV) [138], acceptable band offset with respect to silicon (Ec offset = 1.5 eV) [138], and process conditions which are compatible with silicon process flow.

The knowledge of energy distribution of hole traps in these stacks is incomplete. The characterization of energy distribution of both as-grown and stress-induced hole traps will benefit the understanding of the stress induced degradation of CMOS devices. There are only a few of reports [139-142] that address the energy/voltage dependency of hole traps. The methods and conditions used in these works for extracting the energy distribution have several shortcomings. Most of the existing works uses slow dc measurements [124, 143-145] or very low measurement temperature [146] in order to characterize the energy distribution. Some hole traps will be neutralized during the measurement delay and consequently will not be detected. Another clear limitation of the early published works [147-149] is that they can only probe the energy profile within the bandgap.

This chapter is dedicated to demonstrate a method which gives the energy profiling of as-grown hole traps and generated hole traps that are located either beyond or within the bandgap. It demonstrates that the energy probing technique developed in the Chapter 3 for the conventional SiON can also be applied to high-k/SiON stacks. Attention will be paid to the difference in the energy differences in the energy distribution between SiON and high-k/SiON stacks.

4.2 Devices and Experiments

The gate dielectrics of PMOSFETs used in this chapter include a 1.13 nm HfO_2/SiON, and a 2.0 nm Al-capped HfO_2/SiO_2. To investigate the effect of different gate metals on the energy distribution an EOT = 1.52 nm FUSI-gated HfSiON/SiON and an EOT = 1.53 nm TiN/HfSiON/SiON were used. To study the impact of the

thickness of the interfacial SiON on the energy distribution, the interfacial SiON was slant-etched and the HfSiON thickness was kept the same in a TaN/HfSiON/SiON wafer. The discharging during the measurement itself were effectively suppressed by applying fast technique (t_m=500 ns). The stress and measurements were carried out at room temperature.

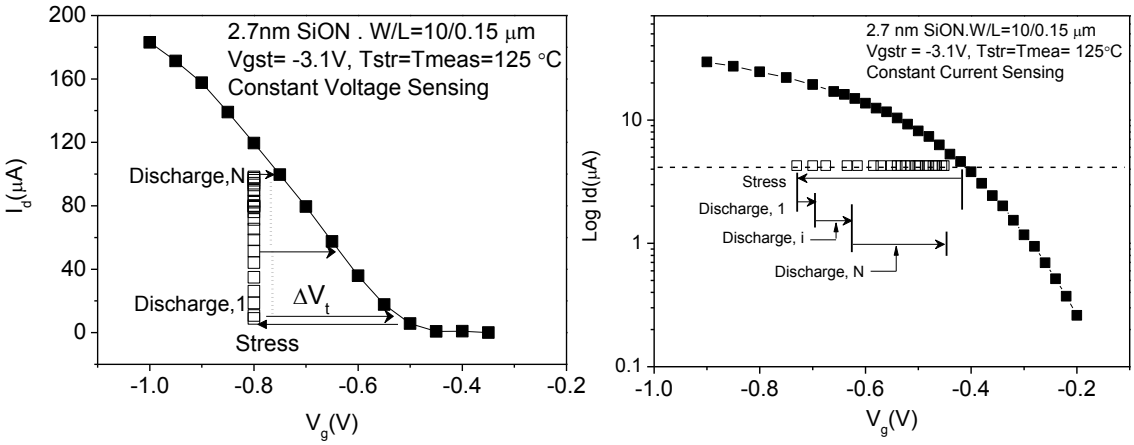
The energy probing technique was comprehensively demonstrated in Chapter 3. The technique uses fast pulses for measurement and the novelty in the technique is that the energy distribution of positive charges in the dielectric is extracted both within and beyond the Si bandgap. The results obtained from the technique demonstrate the existence of different types of positive charges (PCs) whereby each type of PCs will dominate different energy regions. Hence, the results of this technique support earlier works [80, 105,113- 115] which proposed that there are three different types of PCs in gate dielectrics: As-grown hole traps (AHTs), cyclic positive charges (CPCs) and anti-neutralization positive charges (ANPCs)

4.2.1 Comparison of Energy Probing at Constant current vs Constant voltage

The positive charges can be probed at either a constant voltage or a constant current. For comparison purposes, this section presents the results obtained at either a constant voltage or a constant current. The drawback of the constant voltage sensing will be pointed out and the advantages of probing at constant current will be discussed.

Fig. 4.1 (a) and (b) present how the constant voltage sensing and constant current sensing is obtained from the fresh I_d-V_g transfer characteristics. The experiment was

conducted on the 2.7 nm thermal nitrided SiON, at stress and measurement temperature of 125 °C. The stress voltage applied is -3.1 V (≈ 10 MV/cm), and the device was stressed for 10 ksec.



(a) (b)

Fig. 4.1 Energy probing technique by (a) constant voltage sensing and (b) constant current sensing. The filled symbol in both figures denotes the fresh I_d-V_g measurement and the empty symbols denotes the discharging results, after 1 ks discharging time at each discharging level.

Figs. 4.2 (a) - (c) present the results obtained on the 2.7 nm thermal nitrided SiON. The energy probing technique was conducted either by constant current sensing or by constant voltage sensing. The temperature for the stress and measurement is the same in Figs. 4.2 (a) and (c), but different in Fig. 4.2 (b).When the stress and measurement temperature were the same, the discharging procedure was carried out immediately after the stress had been conducted for the pre-specified time. However, for the case of when the stress and measurement temperature are different, as shown in Fig. 4.2 (b), the device was initially stressed at 125 °C for 10 ks. After the stress, the device was cooled down to room temperature with the stress being turned off before it was recharged at Vg = - 3.1 V for around 5 ks and followed by the discharging procedure at room temperature.

Figs. 4.2 (a)-(c) show that there are differences between the constant current and constant voltage measurements and a decision has to be made on which one should be used. To probe the energy distribution of the positive charges (PCs), PCs should be sensed at a fixed surface potential. A change of surface potential for the measurement itself introduces uncertainties in extracting the energy distribution and must be avoided. On one hand, under a constant voltage, a change of positive charges in the gate dielectric will change the substrate surface potential, so that the PCs are actually measured at different surface potential as charging or discharging progresses. As a result, the constant voltage sensing is not the best method for charging sensing. On the other hand, for a given current, the surface potential is essentially fixed, so that it minimizes the uncertainty and will be used in this project.

The surface potential-based compact models of MOS transistor, based on the surface potential equation [143] is $(V_{gb} - V_{fb} - \Psi_s)^2 = \gamma^2 \varphi_t H(u)$ where $H(u) = (\varepsilon_s E_s^2)/(2q\varphi_t P_b)$, represents the normalized square of the surface electric field E_s. V_{fb} is the flat-band voltage, $\varphi_t = K_b T/q$ is the thermal potential where q is the absolute value of the electron's charge, ε_s is the absolute dielectric constant of Si, P_b is the bulk concentration of holes and Ψ_s denotes the surface potential.

To show the advantages of constant current sensing over the constant voltage sensing, Fig. 4.3 (a) gives the energy probing results conducted by constant voltage sensing at different stress and measurement temperature conditions, while Fig.4.3 (b) presents that of the constant current sensing. It is observed in Fig. 4.3 (a) that there is little dependence in the Region 4, though in principle [113] the ANPC should be smaller for the device stressed and consequently measured at 125°C, denoted by the empty

symbol ‘o’. The energy level of ANPC is located above the bottom edge of the silicon conduction band.

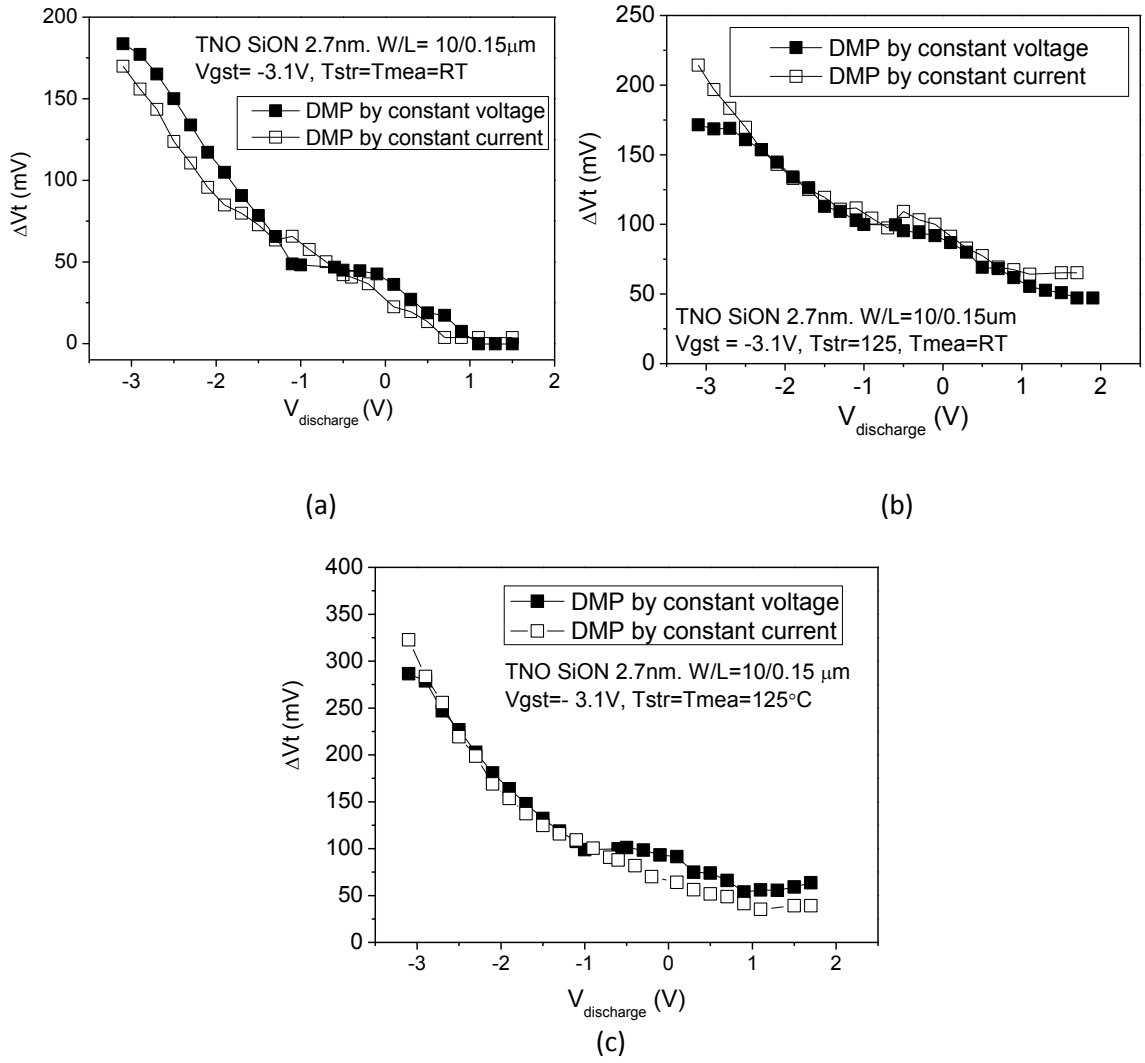
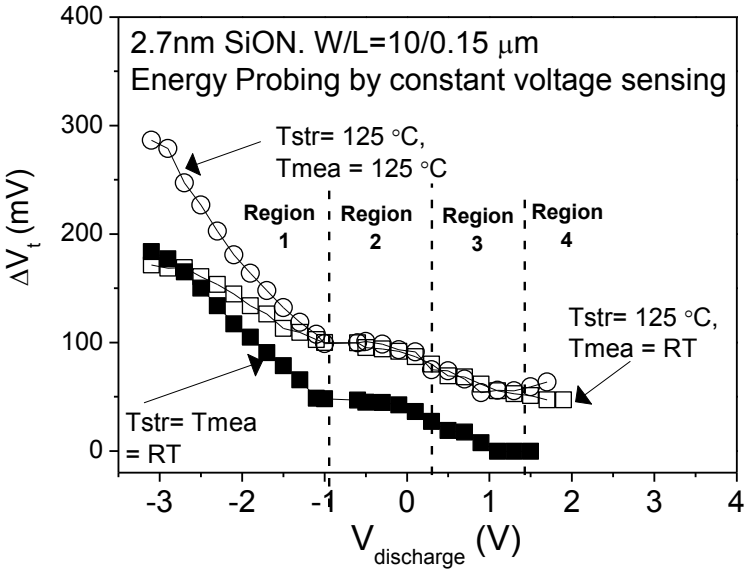


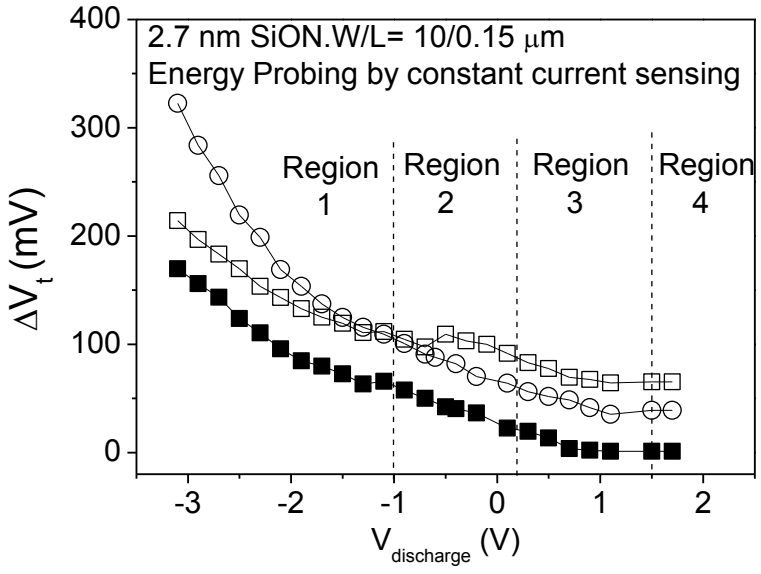
Fig. 4.2 Comparison of constant current sensing to constant voltage sensing for different stress and measurement temperature conditions (a) Stress and measurement temperature is at room temperature (RT), (b) Stress and measurement temperature is at 125 °C and RT respectively, (c) stress and measurement temperature are both at 125 °C.

In principle, for the curves denoted by ‘o’ and ‘□’, the ΔV_t in the region 4 ($E_f - E_v \geq 1.1$ eV,; where ANPC should be observed) obtained from high temperature measurement should be lower than the ΔV_t obtained from the room temperature measurement. The increase in temperature enhances the number of electrons that can reach the ANPC and, consequently neutralize them. The results obtained from the constant voltage sensing do not agree with the results reported in early works [113], therefore. Fig. 4.3 (b), on the other hand, exhibits a clear difference in the energy

profiling for the three different cases. The effect of the stress and measurement temperature conditions can be observed clearly in the region 4. It agrees with earlier works [148] that ANPC increases for lower measurement temperature.



(a)



(b)

Fig. 4.3 Energy probing technique by constant voltage sensing.. The symbol '▪' denotes the discharging result when both stress and measurement was at room temperature. The symbol '□' denotes the discharging result when the stress temperature was 125 °C and the measurement temperature was at room temperature. The symbol 'o' denotes the discharging result when both stress and measurement at 125 °C.

4.3 Comparison of the energy distribution in Hf-based devices of different gate stacks

The ever evolving advanced of sub-micron technology results in the complex and varying structures of the high-k gate stacks. By using the newly developed energy probing technique, the energy density of the hole traps in these devices can be investigated. This section presents the comparison of the energy distribution of the different types of positive charges existing in the Hf-based devices with different gate stacks. Through comparative analysis, the correlation of the energy density is observed for different high-k stacks with varying gate material, high-k bulk material and interfacial layer (IL) thicknesses. The amount of the AHT and the CPC will be compared, and the correlation in respect to the location of the peaks within and beyond the bandgap will be reviewed. The different location of the peaks observed from the energy density profiling may suggest different types of PCs exist in the different devices.

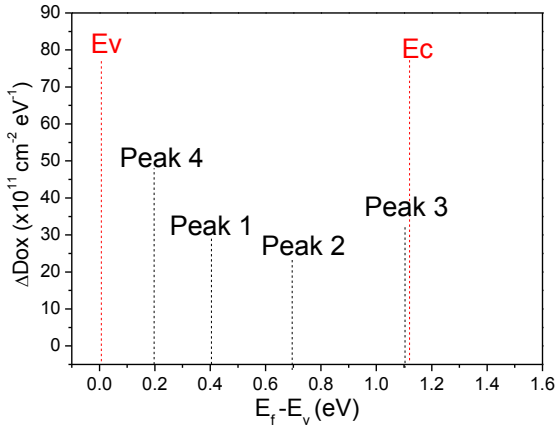


Fig. 4.4 Overall positions of the peaks observed within bandgap for the devices discussed.

Fig 4.4 presents the overall location of the peaks that can be observed from the investigated Hf-based stacks. The defects are perceived to only be hole traps since if there are contributions of electron traps, they can be activated once positive voltage is

applied. However, this has not been observed experimentally since only negative stress is applied. Furthermore, only $|\Delta V_{th}|$ reduction was observed during the discharging

4.3.1 Energy Distribution of positive charges in HfO_2/SiON

Fig. 4.5 shows the schematic of the HfO_2/SiON gate stack device presented in this section. This stack has an EOT of 1.13 nm. This device was prepared by atomic layer deposition. Before depositing the HfO_2, the chemical SiO_2 was nitrided in NH_3 at 900 °C for 60 s. A 2 nm HfO_2 was then prepared, resulting in an equivalent oxide thickness (EOT) of 1.13 nm. A 1000 °C spike anneal was used to activate the source and drain dopant, and the gate is TaN. The channel length and width of the pMOSFET is 0.25 μm and 10 μm respectively.

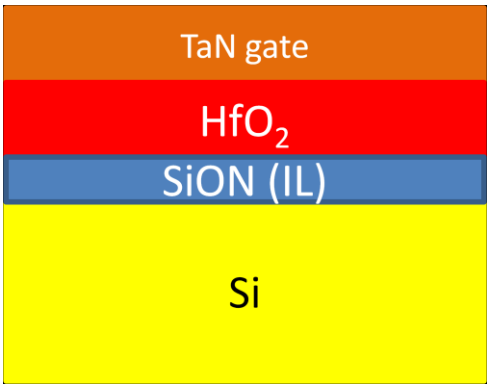
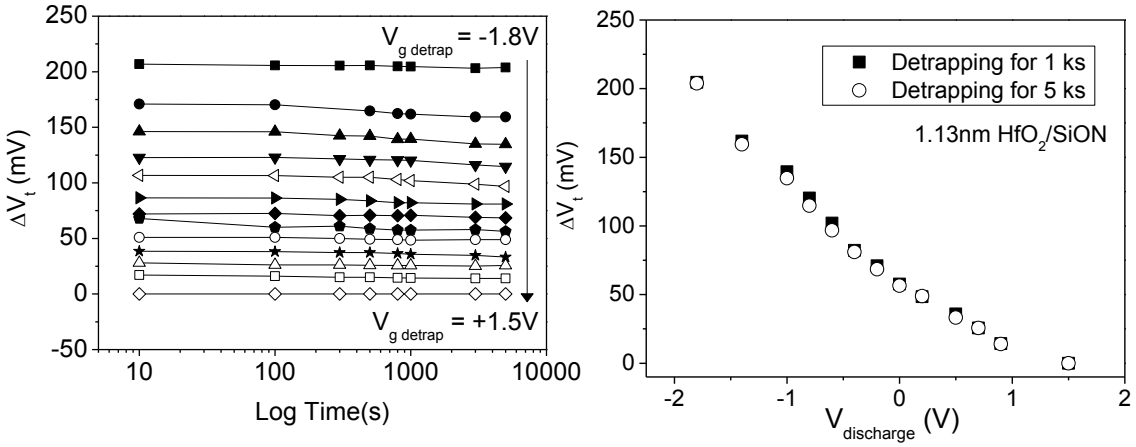


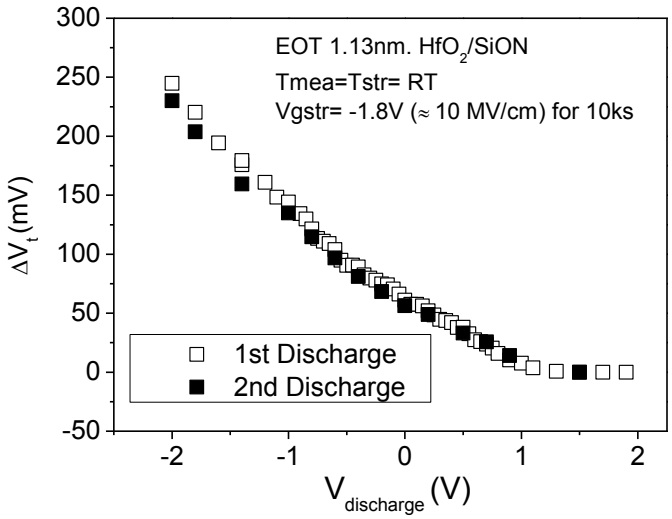
Fig. 4.5 Schematic of the HfO_2/SiON gate stack

The energy probing measurement was conducted on this device in order to observe its energy profiling. The device was stressed at -1.8 V (≈10 MV/cm) for 10 ks and Fig. 4.6 (a) presents the discharging result under different $V_{discharge}$. It is observed that the ΔV_t has almost no change after discharging for 1 ksec. This indicates that all hole traps within the dielectric which are energetically located below $V_{discharge}$ can be completely discharge. To further support this statement, Fig. 4.6 (b) presents the

threshold voltage shift after discharging for 1 ks and 5 ks, against the discharging voltage. The good agreement allows the selection of 1 ks of discharging in time hereafter for each discharging level.



(a) (b)



(c)

Fig. 4.6 (a) Results for discharging under different $V_{discharge}$. (b) Discharging time of 1 ks compared that of 5 ks. It is shown that the discharging time of 1 ks is sufficient. (c) Threshold voltage shift against the discharging voltage of the first discharging compared to the second discharging. It is shown that there is no significant traps generation during the first discharge.

Since discharging takes some time to complete, it is necessary to ensure that there is no further generation of defects during the discharging period. This is checked by recharging the same device for another 10 ks at the same stress bias level and then followed by the 2nd-discharging. As it is shown in Fig. 4.6 (c), there is no significant difference of energy profile from the 1st and 2nd discharge.

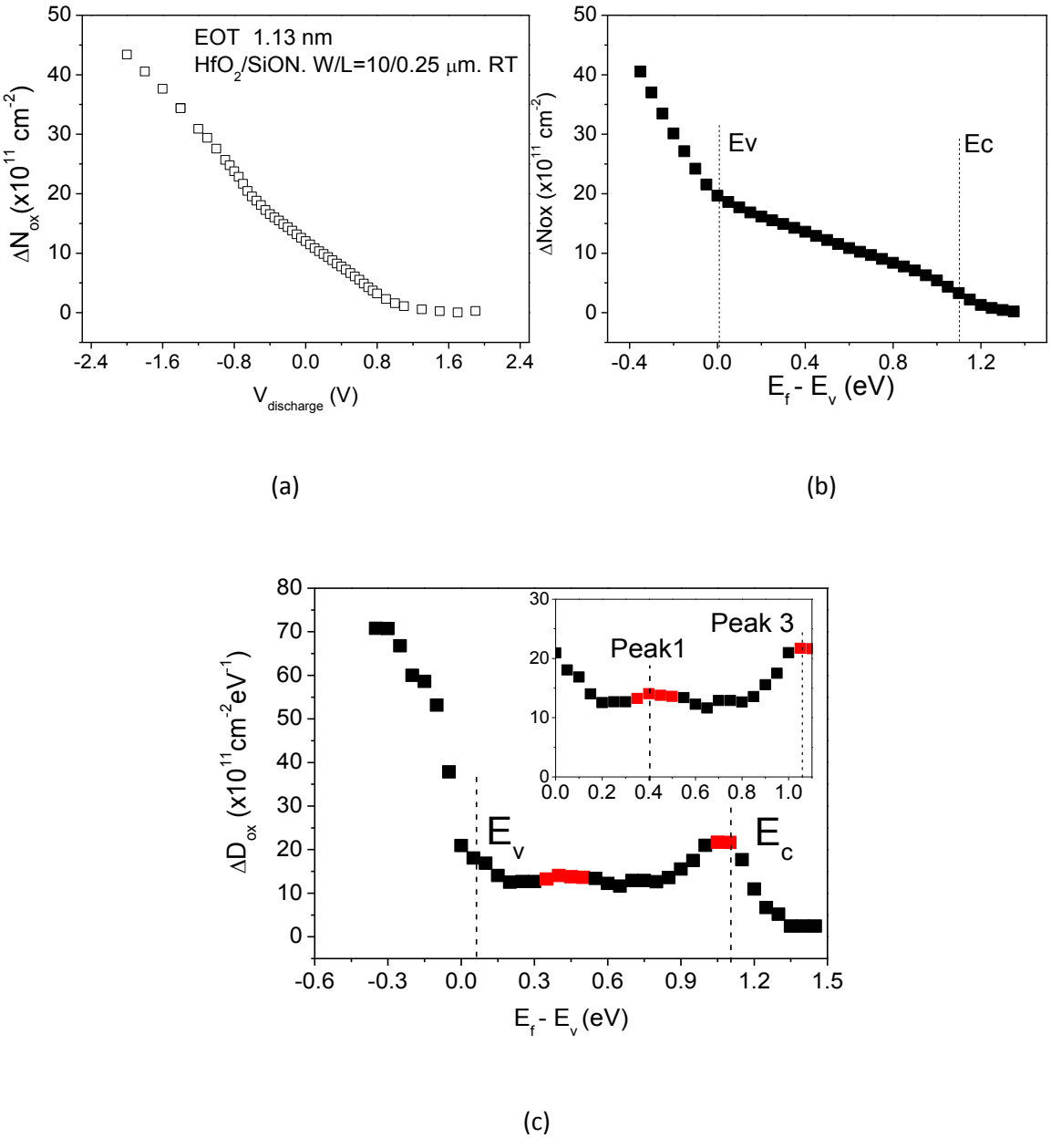
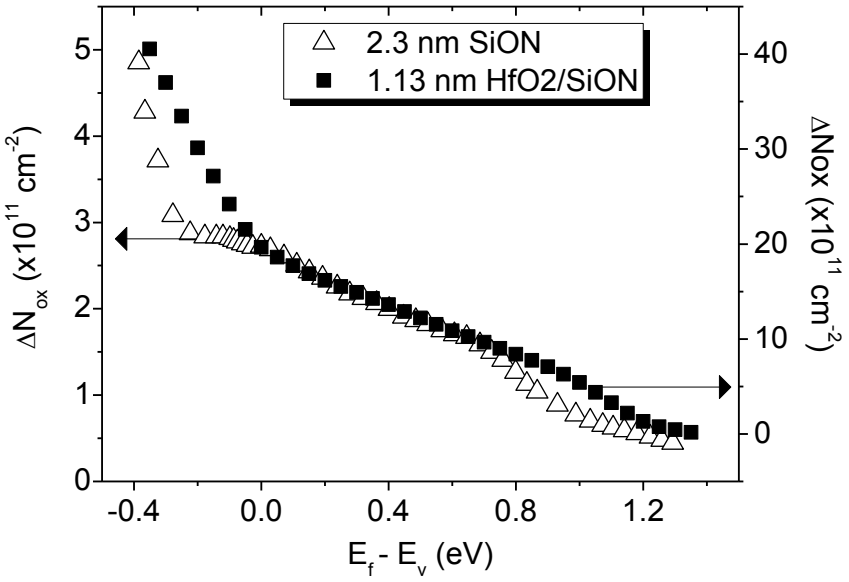


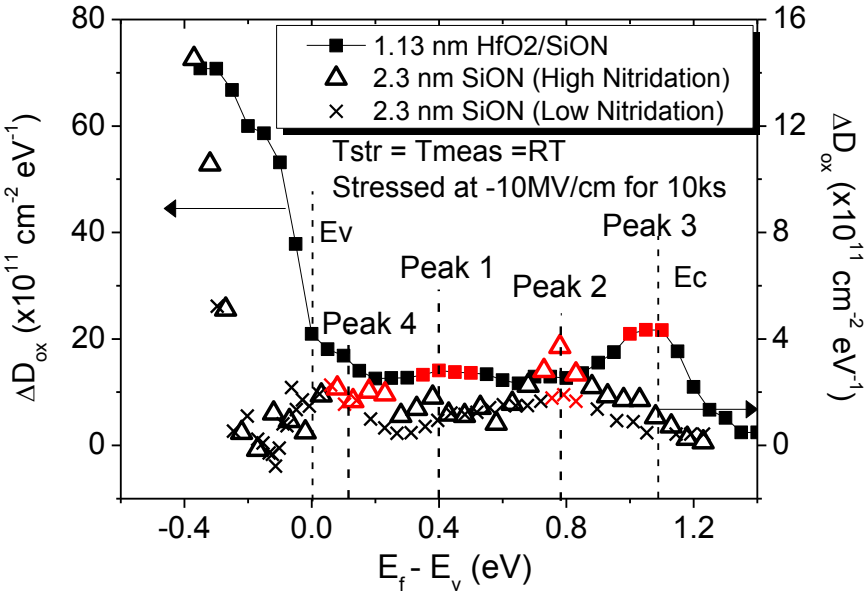
Fig. 4.7 (a) Trap density of the PCs against the discharging voltages. (b) is obtained by converting the $V_{discharge}$ to E_f - E_v . (c) Energy density of the PCs obtained by differentiating (b). Inset shows the magnified view of the energy density within bandgap.

Fig. 4.7 presents the energy density obtained from the energy probing technique for the HfO_2/SiON gate stack device. From Fig. 4.7 (b), it is observed that the PCs beyond the Ec are insignificant under this stress condition suggesting low level of ANPC above Ec. The significant energy density below Ev in Fig. 4.7 (c) suggests substantial as-grown trap (AHT) exist in this device. Early work conducted [113] suggests that the cyclic positive charges (CPC) is energetically located within the bandgap giving rise to the significant peak near to the Ec observed in Fig. 4.7 (c). Additionally, there is also a small peak at the lower half of the bandgap, located around $E_f - E_v = 0.4$ eV.

In order to investigate how the energy distribution of the positive charges in the high-k device varies from the conventional single-layered SiON, Fig. 4.8 (a) and (b) compares the energy profiling of the high-k stack with that of the single-layered 2.3 nm plasma-nitrided SiON devices (highly nitrided and low nitrided devices), which was used in Chapter 3 for the demonstration of the energy probing technique. It is observed that both the high-k and the conventional processes comprise of substantial AHT indicated from the significant peak below the Ev. It is worth to note that the amount of the AHT is nearly tripled in the high-k HfO_2/SiON gate stack, as shown in Fig. 4.8 (b). As far as the PCs within the bandgap is concern, the location of the peaks observed in the conventional SiON is at Ef- Ev = 0.2 eV and 0.8 eV, while in the HfO_2/SiON gate stack one low peak can be observed at Ef- Ev = 0.4 eV and another significant peak near the Ec . Another noteworthy observation is that the CPC can be substantially small in the conventional SiON as compared to the HfO_2/SiON gate stack.



(a)



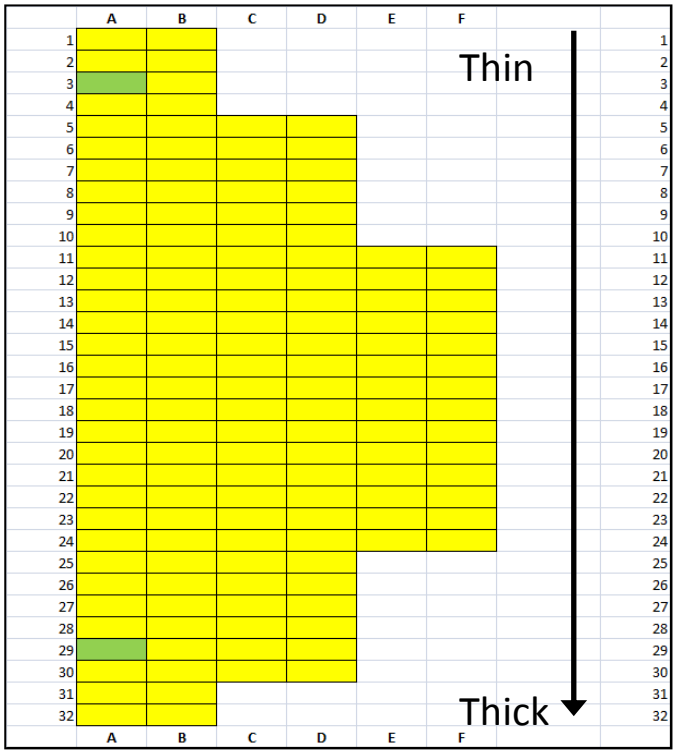
(b)

Fig. 4.8 Comparison of the energy distribution of the HfO_2/SiON gate stack with that of the conventional 2.3 nm plasma nitrided SiON (High and Low Nitridation). The results of the 2.3 nm SiON devices plotted here had been demonstrated in Chapter 3. (a) The trap density, ΔN_{ox} against the surface potential. (b) The energy density, ΔE_{ox} against the surface potential. The red symbols denote the positions of the peak observed.

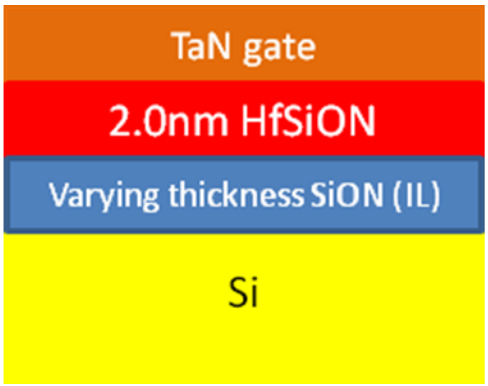
4.3.2 Energy Distribution of positive charges in TaN/HfSiON with varying IL SiON

The wafer investigated in this sub-section is a slant-etched SiON IL layer of varying thickness. The IL thickness of this 8-inch wafer gradually thins from one side of the wafer to the other. Fig. 4.9 (a) and (b) present the wafer map of the slanted wafer and the schematic illustration of the device respectively. The device has a 10 nm TaN – gate and the high-k layer is a 2.0 nm hafnium silicate.

The EOT of each block on the wafer is different due to the slanted feature of the IL. Hence, in order to investigate further on this wafer, the EOT of each block needs to be known. A capacitance-voltage (CV) measurement was conducted using the Keithley 4200-SCS. The measurement was on an NMOS capacitor, of W/L= 70 μm/ 70 μm, located at each block across the wafer. The voltage was swept from -3.0 V to 3.0 V, with a step of 0.1 V and at a frequency of 200 kHz. The result of the CV measurement was then inputted into the CVC simulator to extract the EOT. The simulated CV and the measured CV can be compared in order to ensure that the error margin of the simulated result is small. Fig. 4.10 (a) and (b) exhibits the comparison carried out between the simulated and the measured CV, which in this demonstration, is block A17 of the slanted wafer. Fig. 4.10 (b) is a magnified view of Fig. 4.10 (a) indicating that there is negligible difference between the simulated and the measured CV, beyond the Ec. Fig. 4.10 (c) shows the calculated EOT of the each of the block across the wafer. The solid line in this figure is a guide for the eyes to observe how the thickness varies across the wafer.

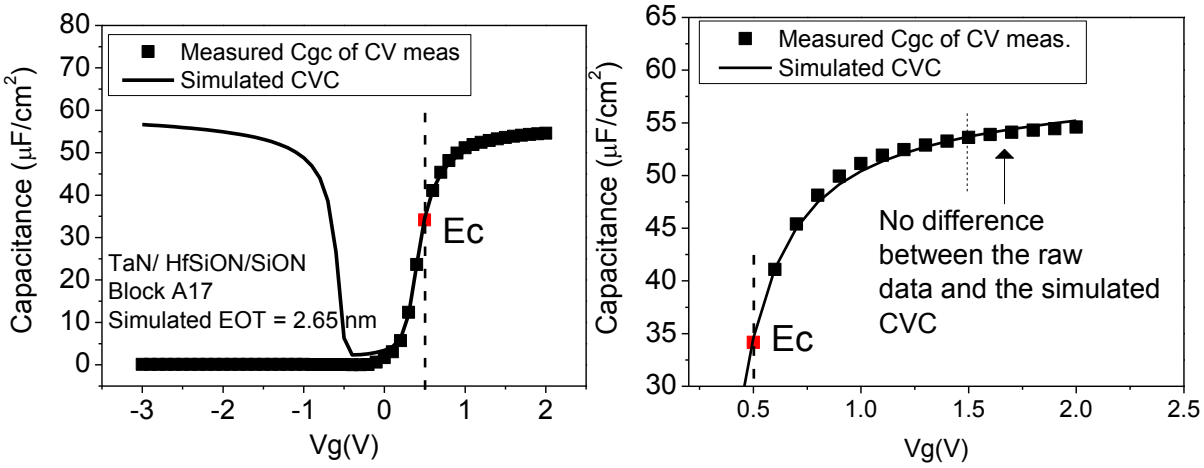


(a)

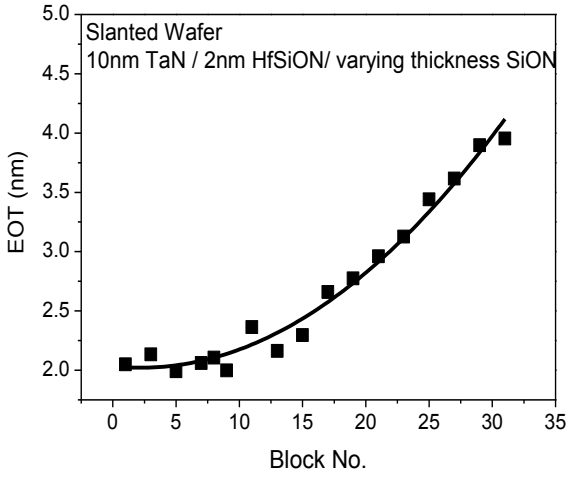


(b)

Fig. 4.9 (a) Wafer map of the slanted wafer TaN/HfSiON/SiON. The arrow indicates how the IL thickness is slanted. The green block (block A3 and A29) were used in this work to demonstrate the energy profiling of the slanted wafer. (b) Schematic illustration of the slanted TaN/HfSiON/SiON gate stack.



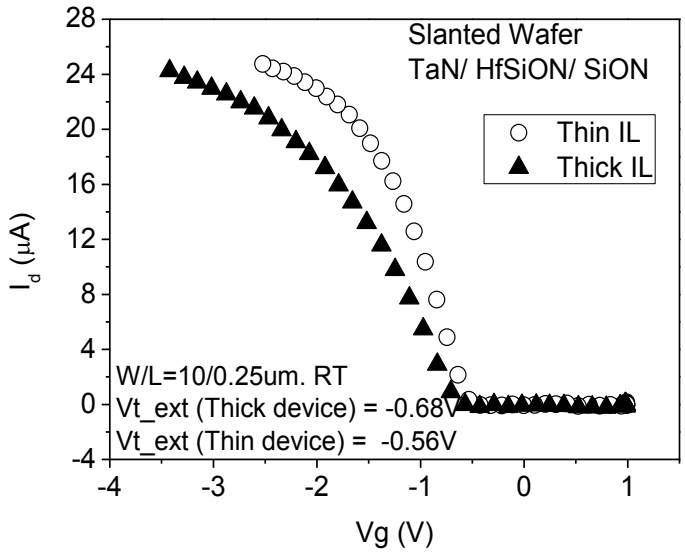
(a) (b)



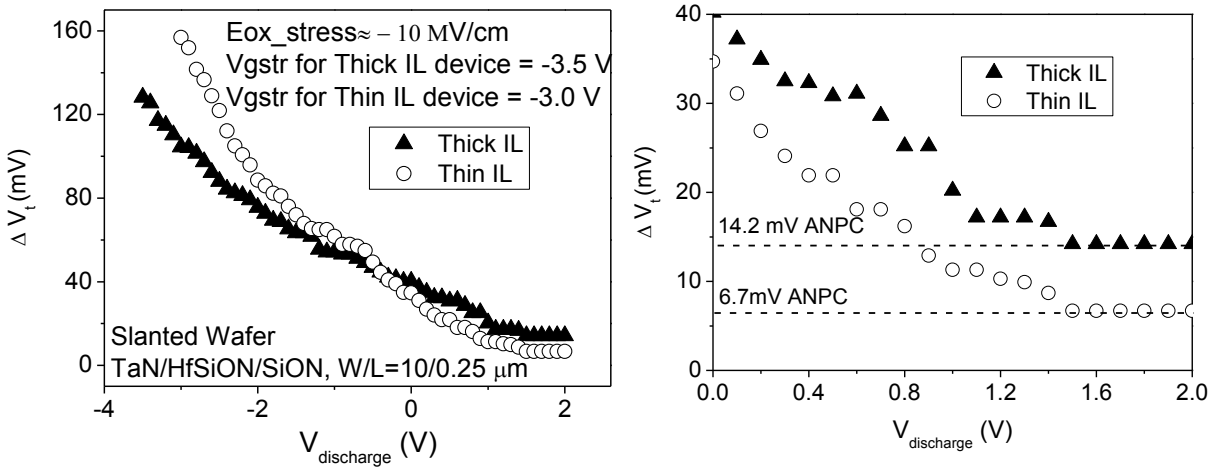
(c)

Fig. 4.10 (a) CV measurement conducted on an NMOS capacitor of block 17, compared to the simulated CVC measurement. (b) A magnified view of (a) above the Ec energy band. The symbol in red in (a) and (b) marks the Ec band. (c) EOT measurement of the device thickness according to the blocks on the wafer.

For the energy probing exercise conducted on this wafer, the thinnest IL and the thickest IL PMOS devices (of block A3 and A29, as depicted in green in Fig.4.9 (a)) were chosen, thus to observe how the varying IL thicknesses can affect the energy distribution. The thin IL and thick IL devices have an EOT of 2.13 nm and 3.89 nm respectively. Both devices have a width and length of 10 μm and 0.25 μm respectively. Fig. 4.11 (a) presents the transfer characteristics, along with the extracted threshold voltage, V_t of both devices. The V_d applied for the measurements is of -50 mV. The stress and measurement temperature is set at room temperature. The measurement was conducted by constant current sensing, whereby the oxide field for the constant current of each experiment was maintained to be constant in order for an effective comparison to be made. Hence, both measurements were conducted at a constant current with the oxide field, E_{ox}_CCS of -0.33 MV/cm (E_{ox}_CCS = (Vg-Vt)/ EOT). The discharging time at each discharging level is 1 ks. Subsequent to the measurement of the reference I_d-V_g on a fresh device, the devices were stressed at -10 MV/cm (-3.5 V and -3.0 V for the thick and thin device respectively), for a stress time of 10 ks at room temperature.

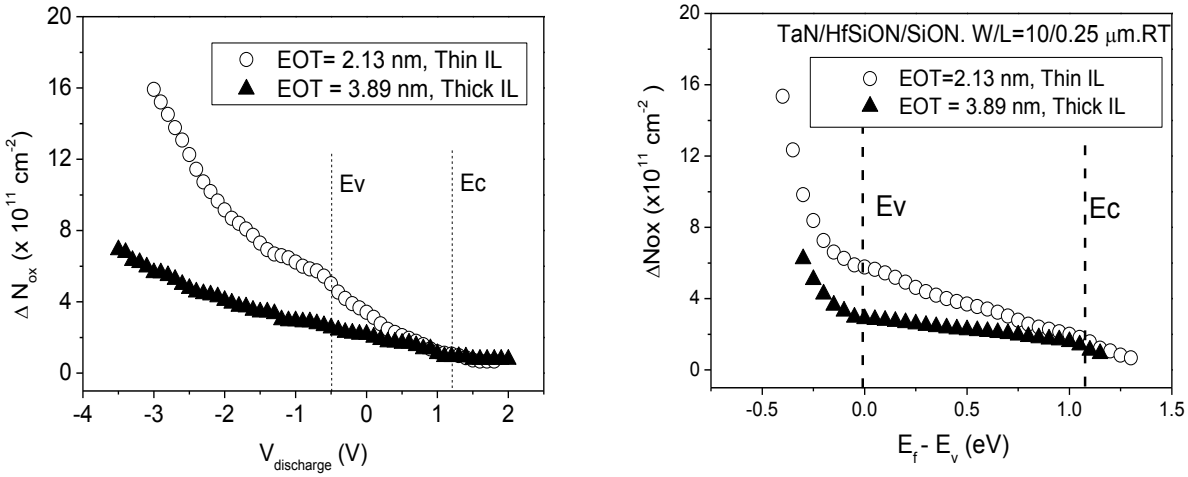


(a)

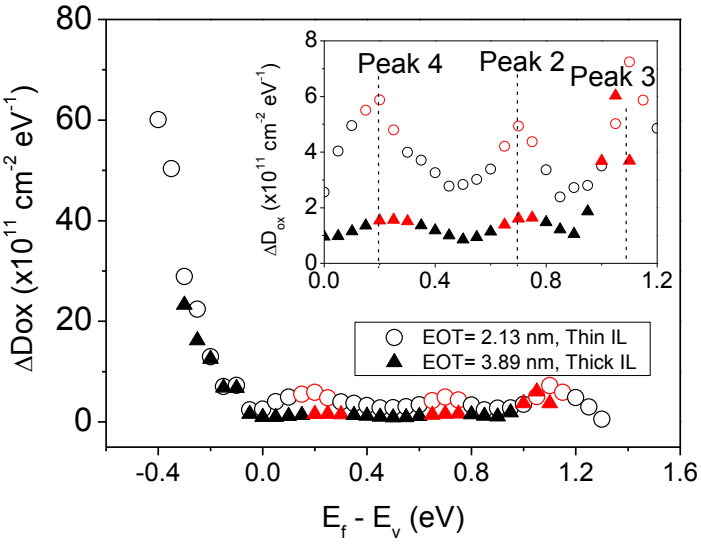


(b) (c)

Fig. 4.11 (a) Transfer characteristics (TC) of a thin IL PMOS device (EOT = 2.13 nm, block A3) and thick IL PMOS device (EOT=3.89 nm, block A29). The threshold voltage of each TC was extracted. (b) The threshold voltage shifts resulted from the discharging under different $V_{discharge}$. The devices were stressed at 10 MV/cm (-3.5V and -3.0V, for the thick IL and thin IL respectively). (c) A magnified view of (b) which shows the threshold voltage shift above Vg=0 V, and the indication of the ANPC generation.



(a) (b)



(c)

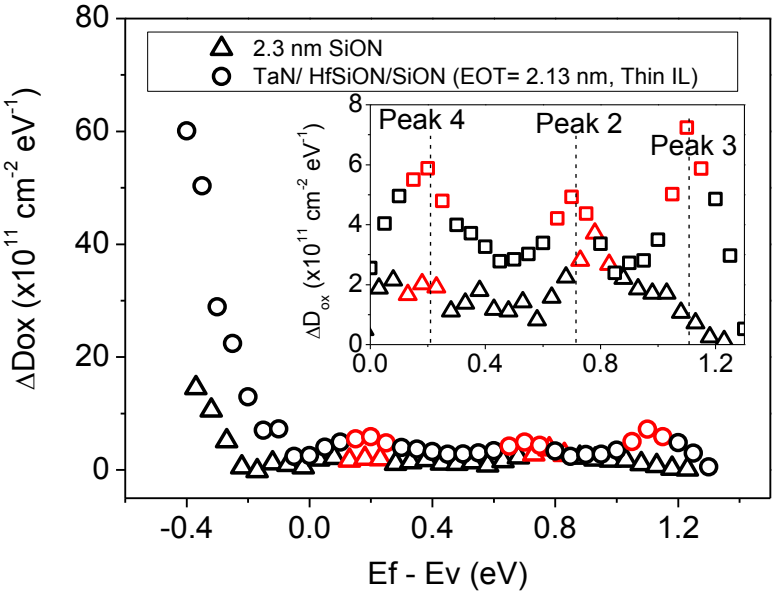
Fig. 4.12 (a) and (b) are the profiles of the trap density against the discharging voltages and surface potential respectively, for different IL thicknesses of the slanted wafer TaN/HfSiON/SiON. (c) The energy density as a result of the energy probing technique, for the different IL thickness of the slanted wafer.

An inspection of Fig. 4.12 (a) and (b) indicates that the PCs behave qualitatively similar for the thin and thick devices. Below the Ev, the ΔN_{ox} will initially drop for both of the device as E_f - E_v increases. Although ΔN_{ox} in Fig. 4.12 (b) is clearly higher for thinner IL, this does not mean that there are more hole traps for the thin IL below Ev. Fig. 4.11 (c) shows that the ΔD_{ox} is insensitive to the IL thickness below Ev. As a result, the as-grown hole traps (AHT) must pile up towards the SiON/Si interface, since an increase of the bulk volume contributes little to AHTs. The real interface traps, ΔN_{it}, measured experimentally by charge pumping, is always quite low in such devices and thus it it perceived that the profiling is not from the contribution of D_{it}.

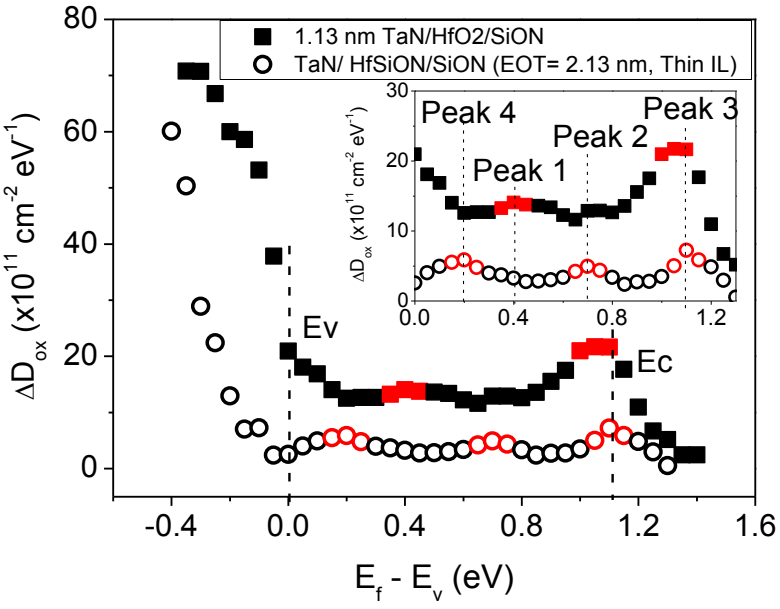
Within the bandgap, three peaks are observed: peaks 4 and 2 are in the lower and upper half of bandgap, while the peak 3 is close to Ec. The position of peak 4 and 2 is similar in a single layer SiON as shown in Fig 4.13 (a). The thinner IL clearly has higher peaks compared to the thicker IL, although the energy positions of these peaks appear insensitive to the IL thickness. This thickness effect is not fully understood at present. One may speculate that an increase of IL thickness allows the strained bond near the interface to relax more and slows down the generation of CPC. The sharp contrast in the IL thickness dependence for defects below Ev and within bandgap strongly supports that they are different types of defects. The energy density below the Ev of the thinner device is considerably significant compared to that of the thicker device, hence to suggest that the AHT is higher as the IL is reduced.

To downscale the EOT for future generation of CMOS technologies, the IL thickness must be reduced. Fig. 4.12 (b) clearly shows that the positive charging will be higher for thinner IL even if the stress was under the same field. Since the operation voltage cannot be reduced proportionally with the oxide thickness, it is expected that the

oxide field will rise in the future. As a result, NBTI will become increasingly important for future generation of CMOS technologies.



(a)



(b)

Fig. 4.13 Energy distribution of PCs in different Hf-based devices. (a) Comparing the TaN/HfSiON/SiON gate stack with the single layer SiON. (b) To observe the effect on energy profiliing for different high-k bulk : Comparison of TaN/HfO_2/SiON with the TaN/HfSiON/SiON.

Fig. 4.13 (a) presents the energy profile of the single layer SiON compared with the TaN/HfSiON/SiON high-k gate stack. It can be seen that while the locations of the peaks within the bandgap is similar, the peak 3 near the Ec is absent in the single layered SiON device. Fig. 4.13 (b) shows that the energy distribution for the two devices of different Hf-dielectric layer has similar structure. It can be seen that the peak 3 near the Ec is present for both devices. However, there are some differences in the locations of the peaks within the bandgap.

4.3.3 TiN / HfSiON/ SiON

The high-k process investigated in this subsection is the TiN/HfSiON/SiON PMOS transistor of an EOT of 1.53 nm. The device has a width and length of 10 μm and 0.25 μm respectively. Fig. 4.14 depicts the schematic illustration of this high-k gate stack.

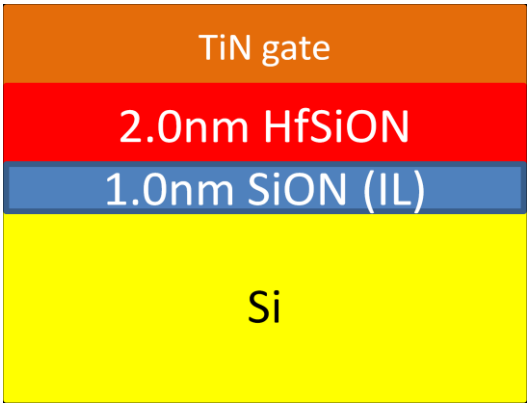


Fig. 4.14 Schematic illustration of the 2.0nm TiN/ HfO_2/ SiO_2 gate stack

The V_d applied in the energy probing measurement is -25 mV and the oxide field for the constant current sensing, Eox_CCS is of -0.13 MV/cm. The discharging time at each discharge level is 1 ks. The stress gate voltage is -10 MV/cm (-2.4 V) on the gate with other terminals grounded. The device was stressed for 10 ks at room temperature. The discharging measurement by constant current sensing was conducted at room temperature. Fig. 4.15 (a) – (d) presents the results obtained from the energy probing technique. The traps were completely discharged at Vg ≈ +1.5 V, where no ANPC can be observed for the stress condition applied.

Fig. 4.15 (c) shows an initial steep drop of the ΔN_{ox} resulting in the significant energy density, $\Delta D_{ox.}$ Once above Ev, the PCs reduce at a slower rate until the traps had completely discharged. The energy profile presented in Fig. 4.15 (d) exhibits three major peaks similar to those in the Fig. 4.12 (c), reproduced here in Fig 4.16 (b) where the energy profiling of the 1.53 nm TiN/HfSiON/SiON is compared to that of the 2.13 nm TaN/HfSiON/SiON. The highest peak again can be seen close to Ec. This suggests that the feature of the energy distribution of positive charges in high-k stack is insensitive to the fabrication condition.

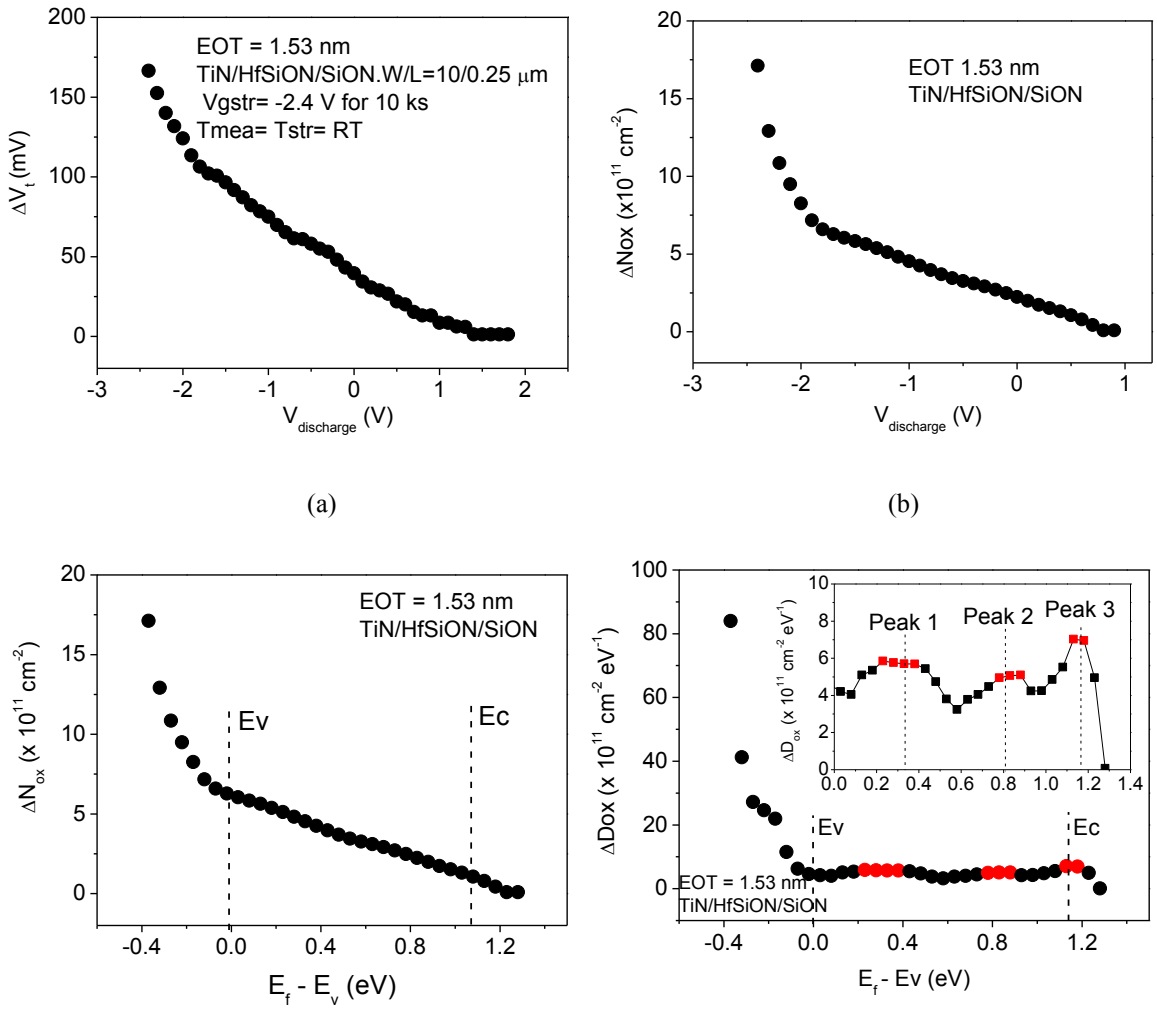
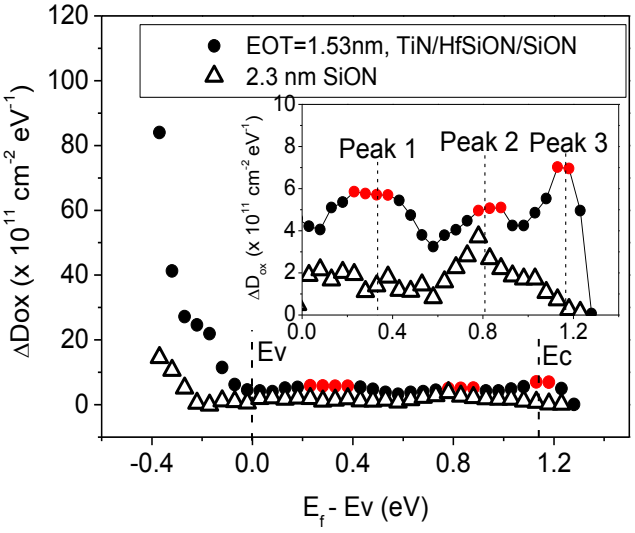
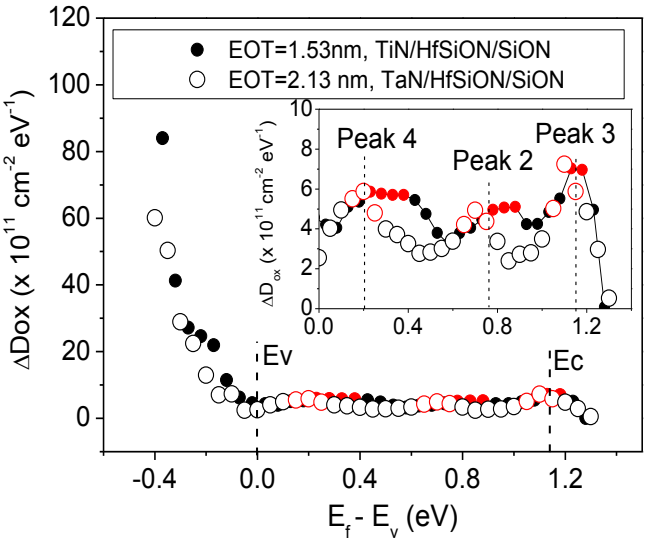


Fig. 4.15 Results of the energy probing technique on the 1.53 nm TiN/HfSiON/SiON. (a) Threshold voltage shifts against the discharging voltages, subsequent to -2.4 V stress for 10 ks at room temperature. (b) and (c) are the trap density against the discharging voltage and surface potential respectively. (d) Energy density obtained through the differentiation of (c). Inset is the magnified view of the energy density within the bandgap. Measurements were also conducted at room temperature.



(a)



(b)

Fig. 4.16 Energy distribution of PCs in different Hf-based devices. (a) Comparing the TaN/HfSiON/SiON gate stack with the single layer SiON. (b) To observe the effect on energy profiliing for different gate material: Comparison of TiN/HfSiON/SiON with the TaN/HfSiON/SiON.

4.3.4 FUSI/ HfSiON / SiON

The high-k process investigated in this subsection is the FUSI/ HfSiON/ SiON PMOS transistor with an EOT of 1.52 nm. The device has a width and length of 10 μm and 0.25 μm respectively. Fig. 4.17 depicts the schematic illustration of this high-k gate stack.

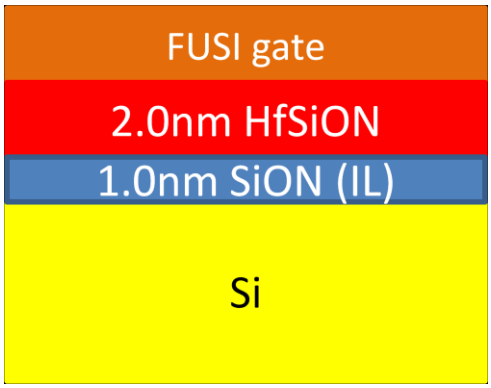


Fig. 4.17 Schematic illustration of the 1.52 nm FUSI/ HfSiON/ SiON gate stack

The V_d applied in the energy probing measurement is -50 mV and the oxide field of the constant current sensing is -0.13 MV/cm. The discharging time at each discharge level is 1 ks. The stress gate voltage applied is -2.5 V and the device was stressed for 10 ks at room temperature. The measurement was also conducted at room temperature. Figs. 4.18 (a) – (d) present the results obtained from the energy probing technique.

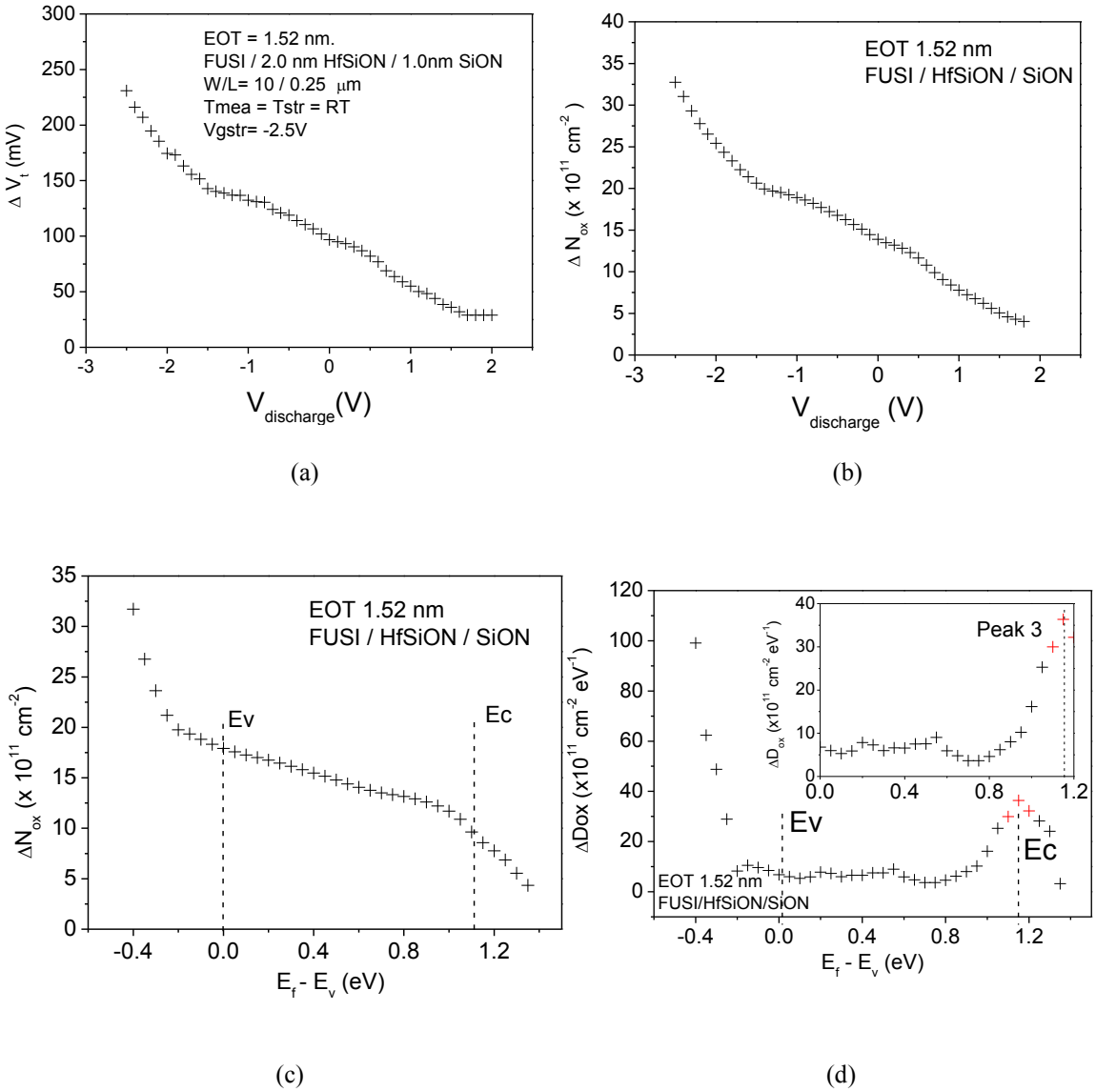


Fig. 4.18 Results of the energy probing technique on the 1.52 nm FUSI/HfSiON/SiON. (a) Threshold voltage shifts against the discharging voltages, subsequent to -2.5 V stress for 10 ks at room temperature. (b) and (c) are the trap density against the discharging voltage and surface potential respectively. (d) Energy density obtained through the differentiation of (c). Inset is the magnified view of the energy density within the bandgap. Measurements were also conducted at room temperature.

Fig. 4.18 (c) shows that the ΔN_{ox} initially drops rapidly which in turn results in the high energy density presented in Fig.4.18 (d). As the discharging level reaches Ev and beyond, the drop in ΔN_{ox} persists although less rapidly. Reaching beyond Ec, Fig. 4.18 (a) shows the ΔV_t saturates at 30 mV, indicating the presence of generated ANPC.

A clear observation that can be deduced from Fig. 4.18 (d) is of the large peak near the Ec band caused by cyclic positive charges (CPC).

When compared with the energy distribution of other high-k samples presented earlier, the peak near Ec in Fig. 4.14 (d) is larger, indicating higher CPC in this sample. Early work [81,105] shows that the precursors of CPC are fixed by the fabrication processes and CPC generation is sensitive to hydrogen exposure during the fabrication. It is possible that the sample with FUSI gate contains a higher density of hydrogenous species.

4.4 Energy Distribution of the Al-capped HfO_2

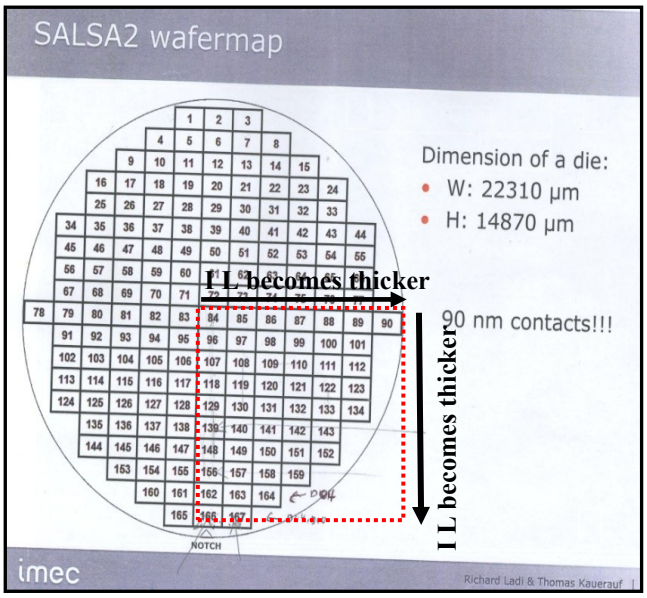
Devices with a high-k/metal gate (HKMG) stack result in low gate leakage and scaled EOT for advanced technology. However, threshold voltage control in these HKMG devices remains challenging [150-152] and requires gate workfunction tuning to control threshold voltage using new materials and new integration schemes [153-155] . To realize higher V_t controllability, V_t-control cap materials such as Al_2O_3 and La_2O_3 are incorporated into the MG/High-k stacks of PMOS and NMOS respectively [156-158]. The AlO capping layer incorporated in metal gated high-k stack devices have been reported to effectively shift the work function of the gate stack with respect to the referenced gate stack [159-161]. This leads to the ease of V_t tuning of the device. The introduction of AlO cap layer also acts as a buffer layer [162] between the MG and the dielectric, hence making the device thermally stable.

This section of the chapter investigates the characteristics of the Al/HfO_2/SiO_2 gate stack and the energy probing technique is applied in order to study the energy distribution of the PCs.

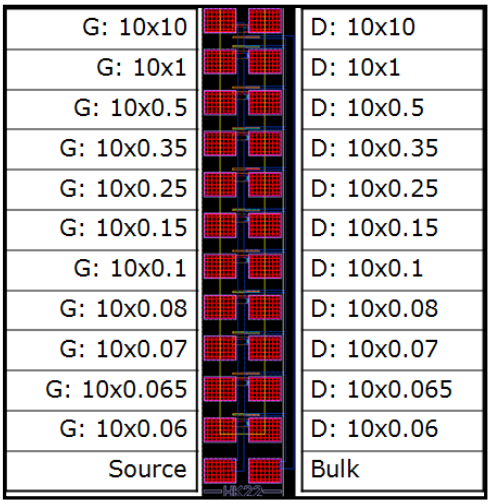
4.4.1 Device characteristics

The investigated Al/HfO_2/SiO_2 gate stack has an EOT of around 2.0 nm. Fig. 4.19 (a) and (c) respectively depicts the wafer map of the device and the schematic illustration of this particular high-k gate stack, while Fig. 4.19 (b) shows the schematic of the terminals on the device. A 0.5 nm thin Al-capping layer is located in between the TiN gate and a 1.2 nm HfO_2 bulk dielectric. The structure of the wafer under test is IL-slanted at which the thickness of the interfacial layer at the middle of the wafer gradually thickens as it widens towards the edge. This work had only used the PMOS transistor in investigating the device's characteristics and its energy distribution.

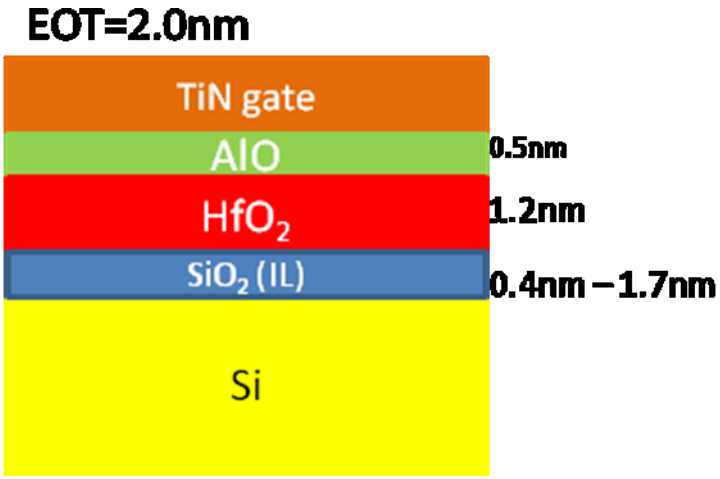
The Keithley 4200-SCS was used to conduct a CV measurement on a W/L= 10 µm / 10 µm PMOS transistor located at each block of the wafer under test. The measurement, which was conducted at room temperature, had been carried out in order to obtain the IL thickness of each block across the wafer. The result of the CV measurement is further inputted into the CVC simulator in order to obtain the specific EOT of a particular block. Fig. 4.20 presents the CV measurement results. Figs. 4.20 (a) and (b) are the screen captures of the quasi-DC split CV measurement, whereby (a) exhibits the C_{gb} plot while (b) gives the C_{gc} plot. Figs. 4.20 (c) and (d) each present the calculated IL thickness for the blocks across the wafer under test.



(a)

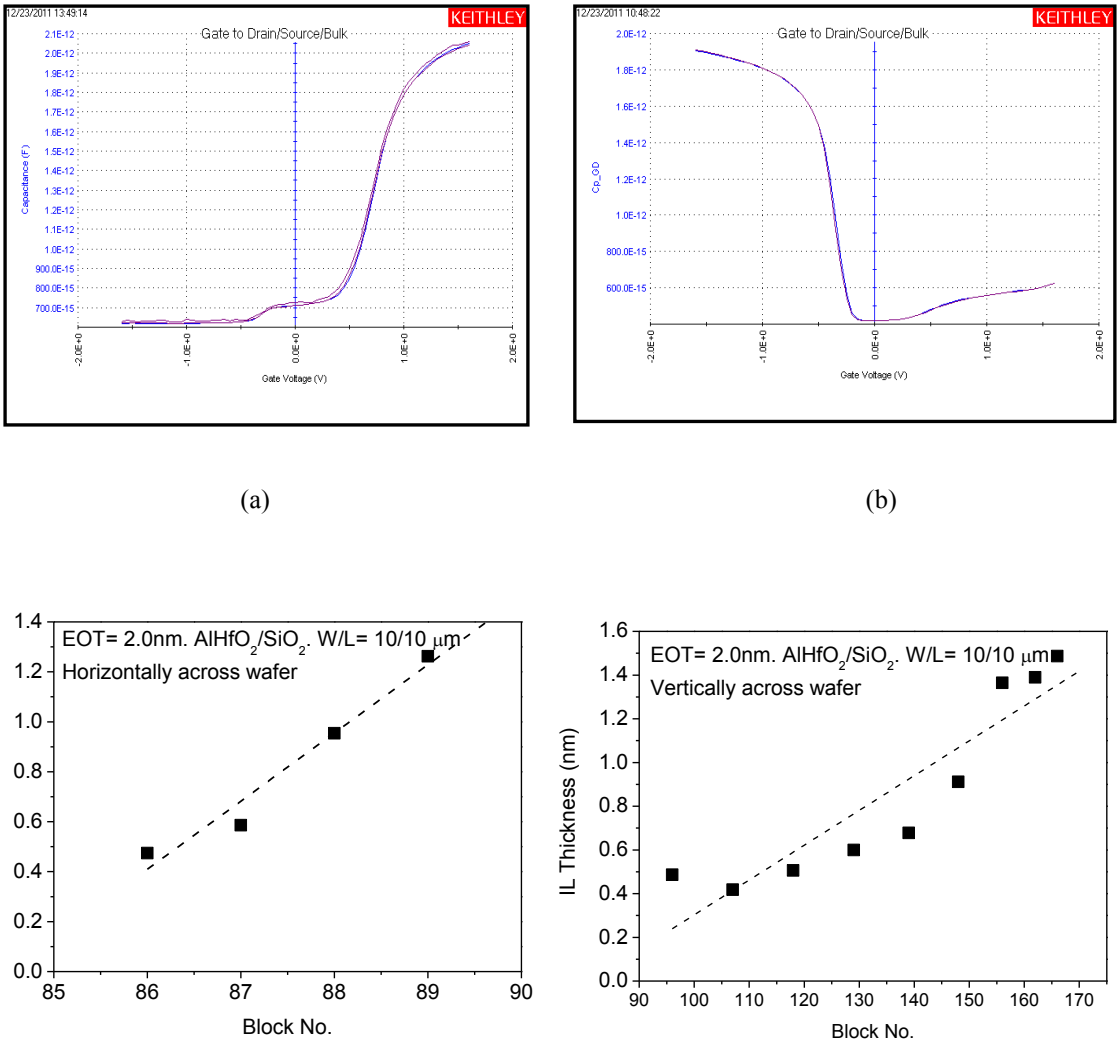


(b)



(c)

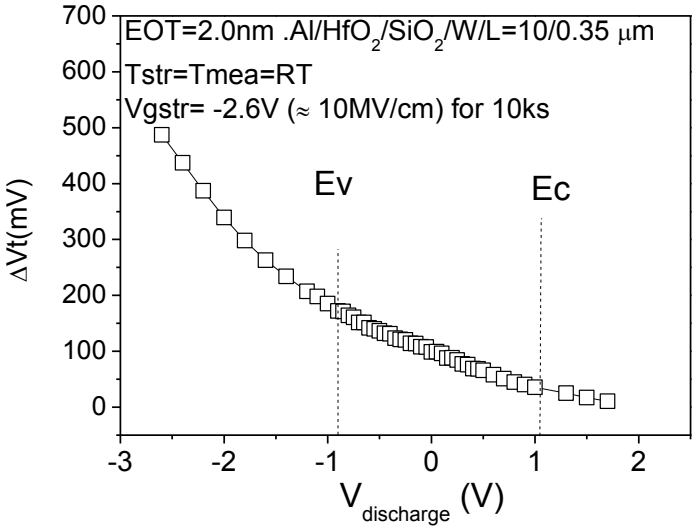
Fig. 4.19 Details on the investigated Al/HfO_2/SiO_2 gate stack. Figs (a) and (b) are the scanned image from IMEC documentation on the SALSA2 300mm mask set (last updated on 13 November 2009). (a) presents the wafer map of the wafer under test, where the dotted red lines indicate the part of the wafer being investigated. The interfacial layer (IL) gradually thickens from the middle of the wafer to edge. (b) gives the schematic of the probing layout of the PMOS transistor terminals. (c) Schematic illustration of the 2.0 nm Al/HfO_2/SiO_2 gate stack



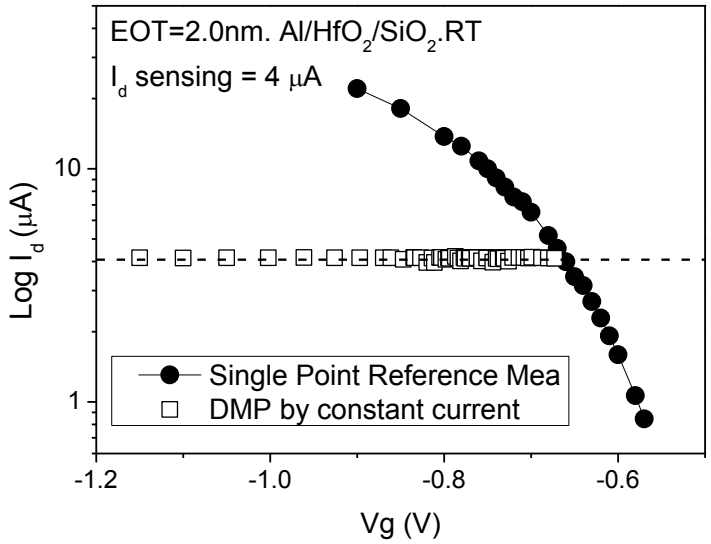
(a) (b)

(c) (d)

Fig 4.20 Capacitance-Voltage (CV) measurement conducted on the wafer under test to obtain the IL thickness of each block across the wafer. (a) and (b) are screen captures of the Cgb and Cgc measurement. (c) and (d) are the calculated IL thicknesses for the different blocks across the wafer. The dotted line is a guide for the eyes to indicate how the thickness varies across the wafer. Measurement is conducted at room temperature.

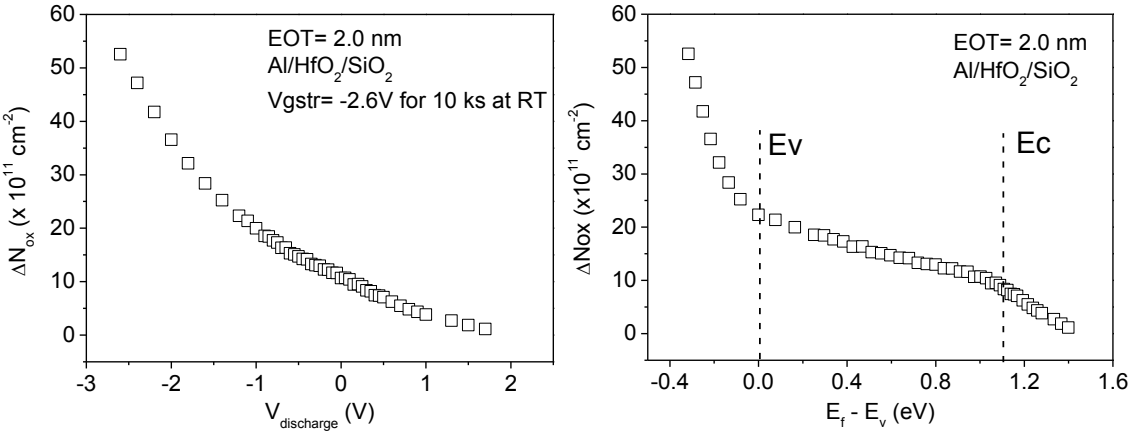


(a)

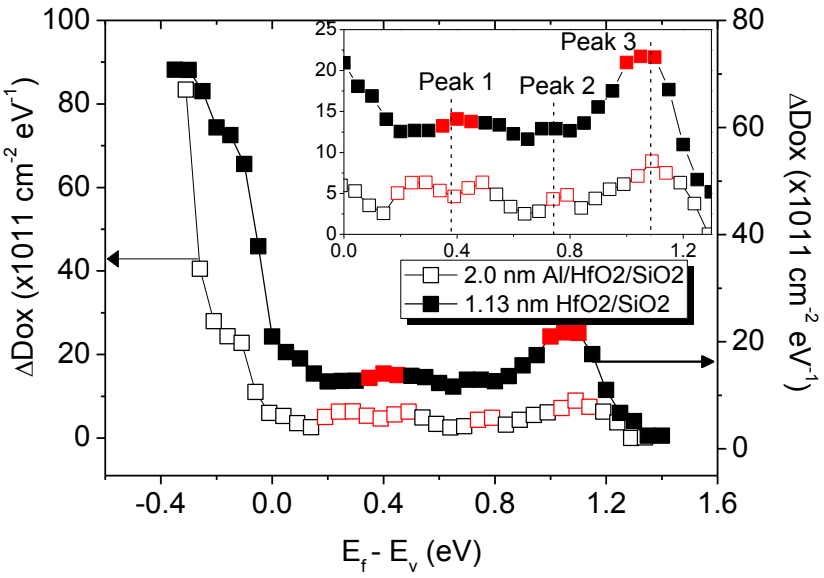


(b)

Fig. 4.21 Energy distribution measurement. (a) Threshold voltage shift, ΔV_t against the discharging voltage after stressing at -2.6 V for 10 ks, in room temperature. (b) presents the constant current sensing at $I_d = 4$ μA denoted by '□'.



(a) (b)



(c)

Fig. 4.22 Energy profiling of the Al-capped HfO_2/SiO_2. (a) and (b) Trap density of the $Al/HfO_2/SiO_2$ after stress, against the discharging voltage and surface potential respectively. (c) Energy density of the $Al/HfO_2/SiO_2$ gate stack compared to the $HfO_2/SiON$. The result of the $HfO_2/SiON$ was previously presented in Fig. 4.7 (b). The red symbols indicate peaks within the bandgap formed subsequent to the stress applied. The inset is the magnified view of the energy density within the bandgap.

The energy distribution measurement technique was applied to the $Al/HfO_2/SiO_2$ gate stack and the threshold voltage shifts observed and the current sensing applied are presented in Figs. 4.21 (a) and (b) respectively. The device under test is taken from block 162 (IL=1.35 nm) and the device's width and length are 10 μm and 0.35 μm respectively. Both stress and measurement were conducted at room temperature. After the measurement of the reference I_d-V_g on a fresh device, the device was stressed at -2.6 V for 10 ks. After stress, the discharging was conducted until the $V_{discharge}$ reaches 1.7 V. The discharging time at each V_g level was 1 ks.

Fig. 4.22 presents the results obtained from the energy profiling technique applied on this gate stack. The steep initial decline of the trap density ΔN_{ox}, below the Ev, become less so as the $E_f - E_v$ reaches 0 eV. Fig. 4.22 (c) presents the energy density of the gate stack obtained through the differentiation of Fig. 4.22 (b). The energy density of the 1.13 nm HfO_2/SiON was plotted together with that of the $Al/HfO_2/SIO_2$ gate stack for comparison. The peak 3 is apparent in the Al-capped device, as observed in the Hf-based devices presented previously, suggesting that it is Hf-related. Peak 1 is slightly wider in this device which might be due to the incorporation of the Al-cap. A closer inspection shows that the peak 2 is also present in this device.

4.5 Conclusion

The energy distribution technique developed in Chapter 3 has been applied to high-k gate stack devices. This chapter comprehensively demonstrate the applicability of the technique, not only to conventional single-layered device but also to various advanced high-k gate stacks. Energy profiling of different Hf-based devices were investigated and the distribution of PCs generated subsequent to stress were reviewed. The energy profiling of the high-k gate stacks were compared to that of the single layer SiON device.

Similar to SiON, a high level of as-grown hole traps were observed below Ev for high-k dielectric stacks. Within the bandgap, although peaks were observed for both SiON and the stack, their energy positions are not always the same. All Hf-dielectric stacks have a clear peak near to Ec, which was not observed for SiON. The defect which results in this peak may be induced by the incorporation of Hf since it only exhibits itself in the hf-based devices and not in any of the conventional single layered SiON devices. The SiON has a clear peak in the upper half of the bandgap, but there is no clear peak in the lower half. In contrast, some high-k stacks have clear peaks in both upper and lower half and the peak in the lower half can be higher than the one in the upper half. For the stress conditions used here, the positive charges in high-k stack are dominated by AHTs below Ev and CPCs within the bandgap, although modest ANPC was observed in some samples above Ec.

Reducing the IL SiON has little effect on AHT below Ev, but increases CPC within the bandgap, indicating higher NBTI for future CMOS technologies. The metal gate can impact NBTI substantially. The FUSI gate results in a high peak near Ec. Process optimization is essential for minimizing NBTI of high-k stack.

The Al-capped HfO_2 stack was also investigated in this chapter in order to observe how the Al-cap can affect the energy profiling of the HfO_2/SiO_2. It is observed that the presence of the capping do not affect the presence of the significant peak near the Ec (peak 3) which can be seen in most hf-based devices, and the capping apparently results in the widening of the peak 1, located around 0.4 eV, near the Ev energy band.

5 Negative Bias Temperature Instability Lifetime Prediction: Problems and Solutions

5.1 Introduction

After developing the technique for probing the energy distribution of positive charges in dielectric, this chapter will explore how to use it to improve the accuracy of NBTI lifetime prediction. The current CMOS manufacturers are using the NBTI lifetime, i.e. τ, as one of the criteria for determining the maximum operating voltage [163] and the assessment of lifetime prediction is used as a figure-of-merit for process screening [164]. In order to reduce the test time and reach a measurable degradation level, the classical lifetime prediction typically accelerate the tests by stressing devices at a bias level higher than that used in the real operation of a technology level . Tests at high stress levels are carried out on individual devices and the resultant lifetimes will be extrapolated towards 10 year to obtain the maximum operating voltage, Vgop_max [165].

Degradation is conventionally evaluated by threshold voltage, V_t extracted from slow DC I_d-V_g measurement which takes several seconds to complete. It has been widely accepted that the negative bias temperature instability (NBTI) degradation

recovers substantially upon the removal of stress, particularly when the conventional stress-measure-stress (SMS) techniques are used. Due to the significant recovery during the slow measurement, the monitored degradations can be highly under-estimated and thus resulting in the overestimation of device lifetime.

Various fast techniques have been proposed to suppress the recovery in the measurement [99,166], but these proposed techniques do not give an industry-wide accepted method for τ prediction, because of the problems which will be detailed in the following section. Since a reliable prediction technique is not available without recovery, some industrial researchers [167] purposely inserted a delay between stress and measurement to give a level of recovery. It is not known, however, what is the correct standardised level of recovery that should be used.

The objective of this chapter is to show that neither the DC nor the pulse measurement is suitable for lifetime prediction due to either over-estimation or under-estimation of the maximum operational voltage, Vgop_max. After analysing the sources of errors by using the newly developed energy probing technique this chapter will propose a new technique that can correct these errors. This new lifetime prediction technique can readily be implemented in industry.

5.2 Pitfalls of the Lifetime Prediction based on the slow DC and fast pulse measurements

After suppressing recovery, one would expect that the larger degradation shortens the τ. It was reported, however, that the extracted τ can either be similar, as depicted in

Fig.5.1 (a) [168] or even longer, as presented in Fig. 5.1 (b) [51]. Such reported results obviously do not make sense since larger degradation should result in shorter lifetime. Lifetime prediction requires extrapolation from high stress bias, Vg_st to low operational Vg_op. After suppressing recovery, log(τ) versus log(|Vg_st|) does not always follow a straight line, as presented in Figs.5.1 (c) and (d) [169]. There is an apparent change in the slope as the gate voltage stress, Vg_st reduces towards the operational voltage, Vg_op, as indicated in the figures. This invalidates the extrapolation.

The occurrence in the change of the slope as the stress voltage is reduced towards the Vg_op has also been examined by early published reports [169-171] in which they had remarked of such 'turn-around points' occurring in the lifetime measurement. This occurs to both conventional single layered devices as well as high-k devices. Hence, these existing experimental methods should be perceived as unreliable due to the defective extrapolation towards the 10 years criteria.

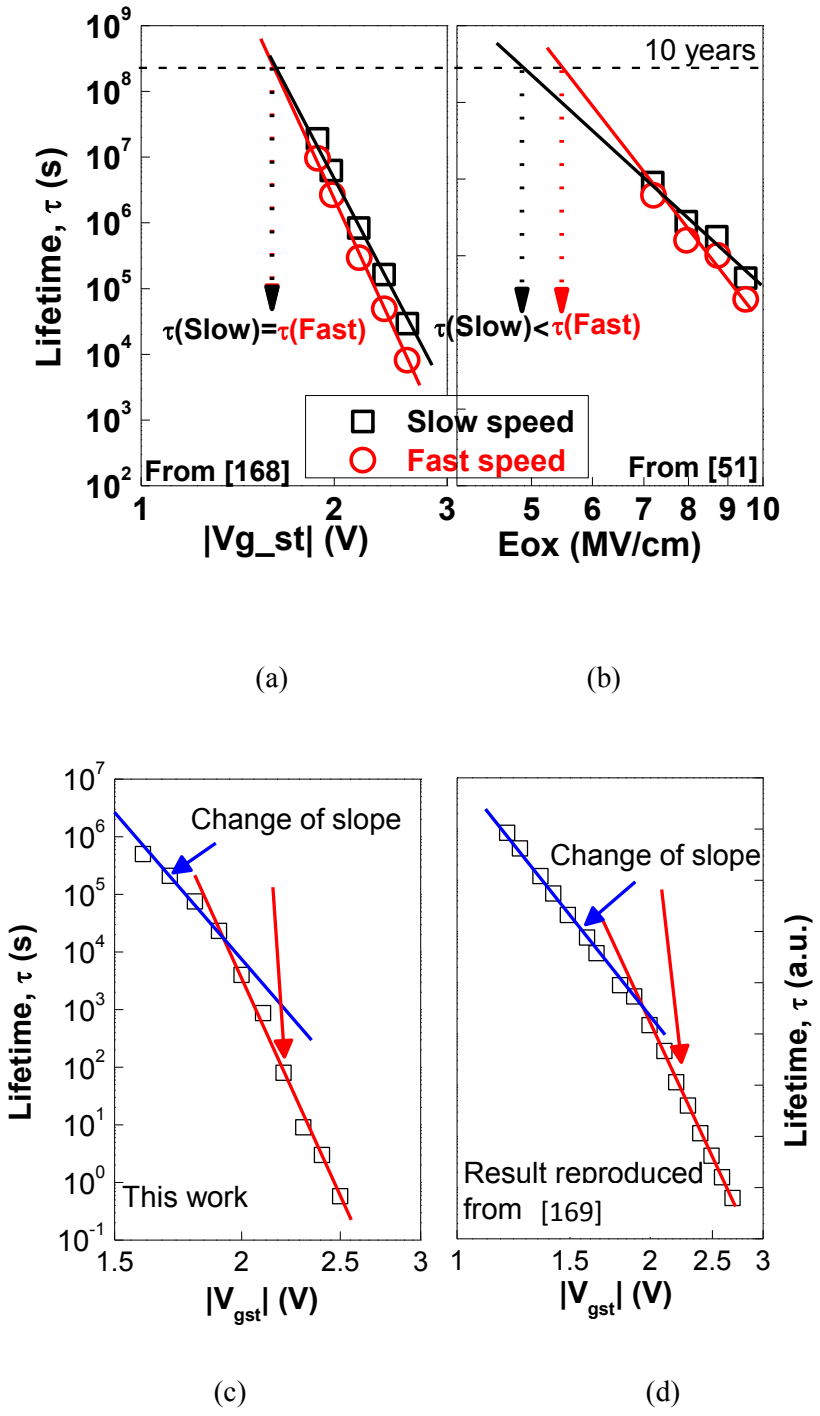


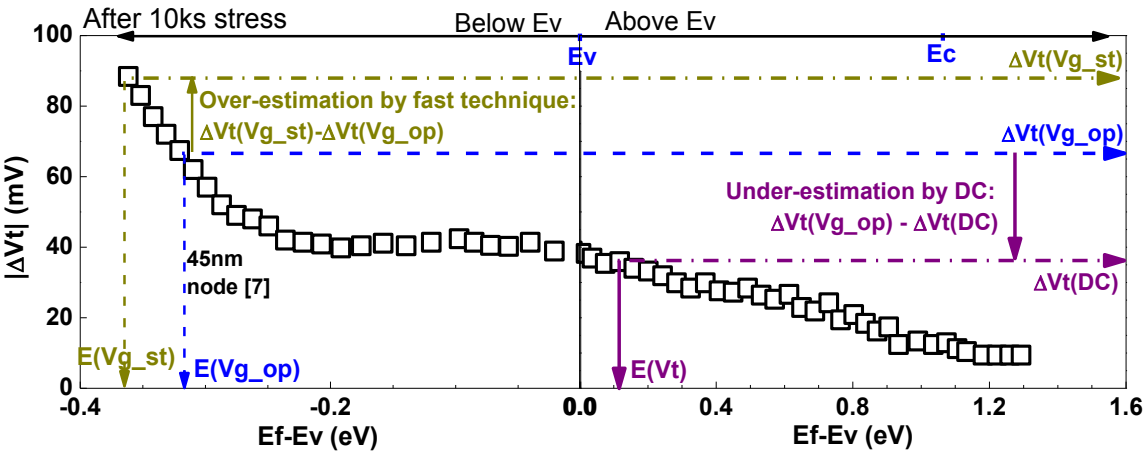
Fig. 5.1 Problems with using fast techniques: I. Lifetime, τ, from fast techniques can be similar (a) or even longer (b) than that from slow measurements. II. Lifetime cannot be reliably extracted by Vg-acceleration method because of the non-linearity between $\log(\tau)$ and $\log(|V_{gstr}|)$. Data in (c) was reproduced from this work. Data in (d) shows a similar trend observed by the work reproduced from Chen et al [169].

5.2.1 The Sources of Overestimation and Underestimation of the Lifetime

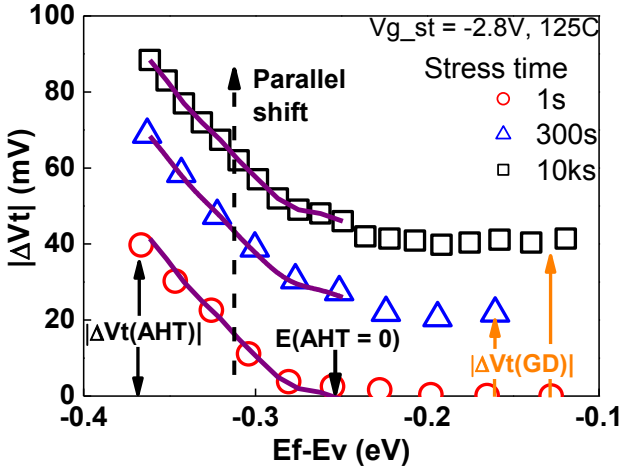
Fig. 5.2 (a) shows the energy profile of positive charges (PCs) after 10 ks stress obtained from the energy distribution method developed in Chapter 3. In principle, a defect must be chargeable at the operation field E(Vg_op), if it is to be included in τ prediction. However, the existing lifetime measurements applied by most academic works [91, 172-174] as well as the semiconductor industry [175-178] is to initially stress at substantially high gate voltages and consequently carry out the lifetime extrapolation. During stress at high |Vg_st|, the defects below E(Vg_op) are charged, but they are not chargeable at |Vg_op| and, consequently should be excluded in τ prediction.

On the other hand, the fast techniques published in early works [164,165] freeze the defects after stress, leading to the marked 'Over-estimation' of $|\Delta Vt|$ in Fig. 5.2 (a), which induces an extra lowering of τ for higher |Vg_st| and causes the problems in Figs. 5.1 (a) – (d). On the other hand, the defects that are charged at E(Vg_op) can be partially neutralized at E(Vt) during a slow DC measurement, resulting in the marked 'Under-estimation' of $|\Delta Vt|$. Fig. 5.2 (b) presents the threshold voltage shifts, ΔVt, obtained from the energy profiling measurement, for stress times of 1 s, 300 s and 10 ks. It can be seen that the three curves below E_v are parallel-shifted. Above E_v, the $|\Delta Vt|$ increases with stress time. This observation will be discussed in terms of the defects in the following section.

In future, the oxide field during operation will further increase, driving E(Vg_op) lower [51] and, in turn, exasperating the over-estimation of ΔVt . The challenge is how to avoid both the over- and the under- estimation in τ prediction.



(a)



(b)

Fig. 5.2 (a) An illustration of the origins of problems: ΔVt depends on Ef-Ev and, in turn, Vg. After stress under Vg_st, the fast technique freezes recovery and charges. Under Vg_op, 'ΔVt(Vg_st)' was used wrongly, leading to the 'Over-estimation'. For DC method, ΔVt was measured at Vg= Vt, resulting in the 'Under-estimation'. The correct ΔVt for τ prediction under Vg_op should be 'ΔVt(Vg_op)'. Note the different scales for Ef below and above Ev. (b) ΔVt at different stress time of 1 s, 300 s and 10 ks; each stress time are denoted by the symbols 'o', 'Δ' and '□' respectively. 'GD' and 'AHT' marks the generated defects and as-grown hole traps respectively.

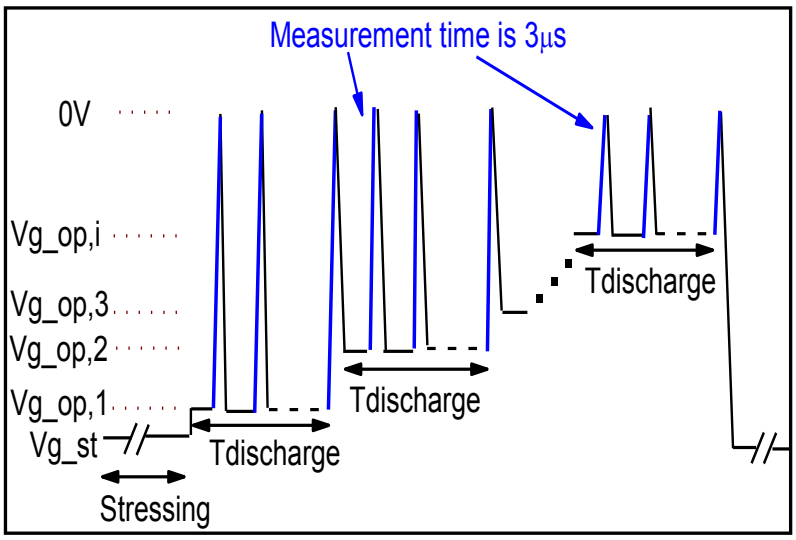
5.3 A new method for Lifetime Prediction

The waveform of the gate bias for the new experimental method to predict the lifetime is presented in Fig. 5.3 (a). A fresh device is initially stressed under Vg_st for a pre-specified time, and subsequently the Vg is stepped towards positive, to Vg_op1, in order to discharge. After completing discharge at Vg_op1, Vg is consequently stepped to Vg_op2 and the same procedure is followed until Vg reaches the threshold voltage, Vt. The Vg_st then is re-applied for the next pre-specified time.

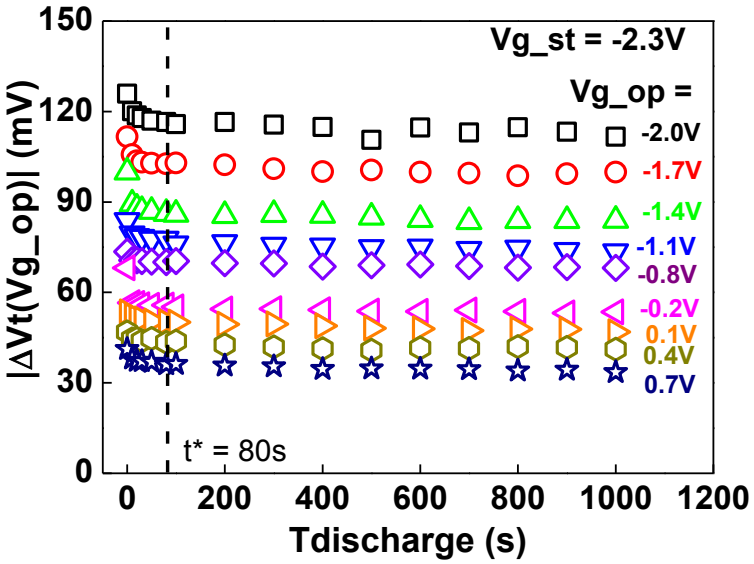
For the purpose of demonstrating this new technique, the measurement procedure was applied on a fresh 2.0 nm plama-nitrided SiON. Fig. 5.3 (b) presents the threshold voltage shift, ΔVt obtained against the discharging time of 1 ks. The ΔVt is monitored at different discharging voltages, of which the discharging voltage increases towards the positive, in steps of 0.3 V from a Vg_op of -2.0 V to a Vg_op of +0.7 V. The measurement was taken during the pulse edges of 3 μs at constant current of Id=100nA*W/L.

Fig. 5.4 (a) presents the ΔVt after full discharge at each Vg_op. Before modeling ΔVt versus time, the defects are analysed first. The Fig. 5.2 (b) shows the energy profiles after different stress times. From this figure, two groups of defects can be identified : i) The defects below E(AHT=0) are fully filled after only 1sec and do not increase further with stress time, as confirmed by the parallel shift of the three curves, supporting that they are “as-grown hole traps” and depicted as “AHT” in the figure

[179] ; ii) Defects above E(AHT=0) are negligible at 1sec but increase with stress time, indicating they are generated defects, which is depicted as “GD” in the figure.



(a)

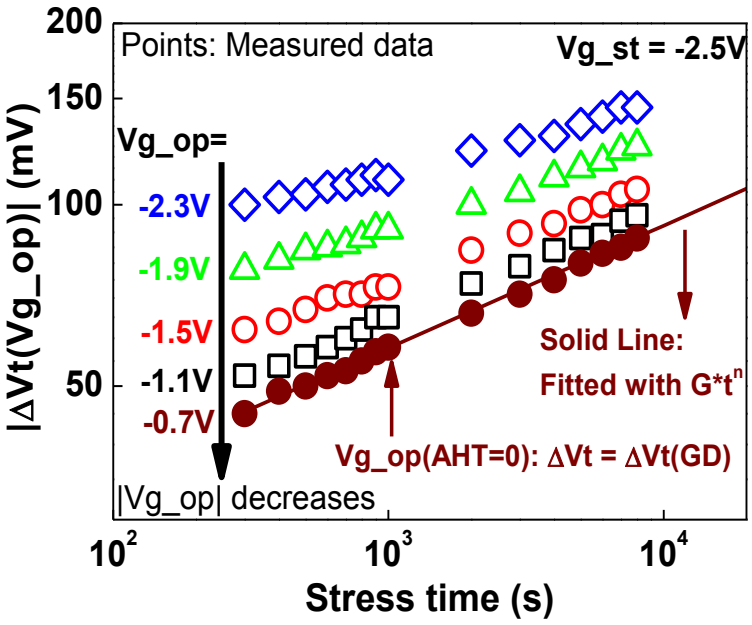


(b)

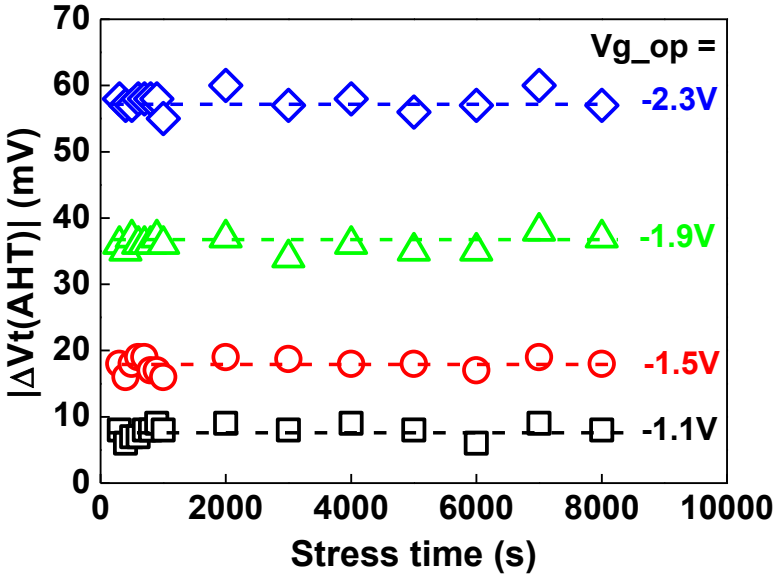
Fig. 5.3 (a) The test Vg waveform. After stressing for a pre-specified time, |Vg| was lowered in steps. Under each Vg_op, ΔVt is monitored against discharge time at the pulse edge of 3μs. (b) Typical results for discharging under different Vg_op. The device was stressed at Vg_st = -2.3 V under 125 ^{o}C for 10 ksec. The sample is a 2.0 nm plasma-nitrided SiON.

To separate the threshold voltage shifts due to the AHT, ΔVt(AHT), from the threshold voltage shifts resulting from the generated defects, ΔVt(GD), a test was first carried out to find the Vg_op for AHT=0. Fig. 5.4 (a) and (b) demonstrates the experimental procedure for separating the generated defects from the as-grown defects. Fig. 5.4 (a) presents the ΔVt for different Vg_op, plotted against stress time. The stress voltage, Vg_st is – 2.5 V and the total ΔVt was monitored at pre-specified times, from 300 s up to 10 ks. After each pre-specified stress time, the device is discharged in steps, starting from Vg_op of -2.3 V until Vg_op of -0.7 V, in steps of 0.3 V. Each data point plotted in Fig. 5.4 (a) is taken after discharging time, Tdischarge of 80 sec, where the discharge essentially has completed as shown in Fig. 5.3 (b). Under a given |Vg_st|, ΔVt at Vg_op(AHT=0) (denoted as '●' in the Fig. 5.4 (a)) can originate only from generated defects, since the as-grown traps are not charges at Vg_op(AHT=0). Fig. 5.4 (a) shows that the ΔVt(GD) follows a power law [180]. Fig. 5.4 (b) presents ΔVt(AHT) against stress time. The ΔVt(AHT) for each Vg_op is obtained from ΔVt(Vg_op) - ΔVt(GD). It can be seen that the higher the |Vg_op| is, hence the higher |ΔVt(AHT)| will be. Fig. 5.4 (b) also shows that the |ΔVt(AHT)| is a constant against stress time for a given Vg_op since AHT-filling saturates around 1s, as shown from the saturation that can be observed in the Fig. 5.5. Exclusion of these AHTs by subtracting ΔV_t at 1 s [181] overestimates the lifetime.

The characteristic of the AHT is that the traps are pre-existing and hence should not be affected by the stress. Fig. 5.6 shows that the ΔVt(AHT) is independent of the Vg_st, therefore confirming their 'as-grown' nature. It is worth to note that the ΔVt(AHT) is directly determined from the measured data, without the application of any trap filling model.



(a)



(b)

Fig. 5.4 Separation of generated defects (GD) from as-grown hole traps (AHT). (a) |ΔVt(Vg_op)| against stress time. The ΔVt for each different Vg_op and at each stress time was taken after Tdischarge = 80 s. The solid line is fitted with a power law. (b) ΔVt(AHT) against stress time . ΔVt(AHT) under each Vg_op is determined from ΔVt(Vg_op) - ΔVt(GD).

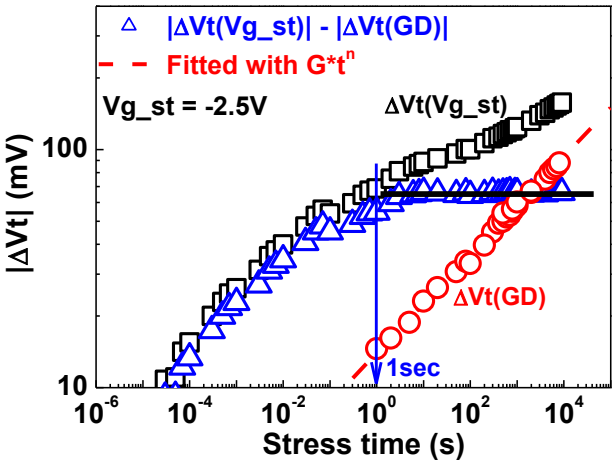
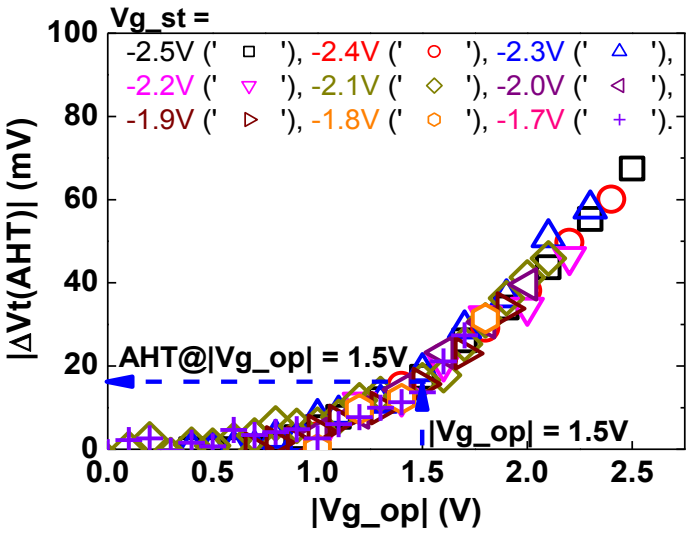


Fig. 5.5 A comparison of the total degradation |ΔVt(Vg_st)| ('□') and the generated |ΔVt(GD)| ('○'). ΔVt(AHT) ('Δ') is evaluated from ΔVt(Vg_st) - ΔVt(GD) and saturates around 1 sec.

Fig. 5.6 presents the impact of the different stress voltages on AHT and the generated defects. It is clearly observed that the ΔVt(AHT) is not at all affected by the different stress biases, justifying the characteristics of the as-grown hole traps. For $\Delta Vt(GD)=Gt^n$, the prefactor, G follows a power law against the stress voltage, Vg_st as shown in the Fig. 5.6 (b). Fig. 5.6 (c) shows a constant power factor, n against the Vg_st. Hence, it warrants the Vg-extrapolation.



(a)

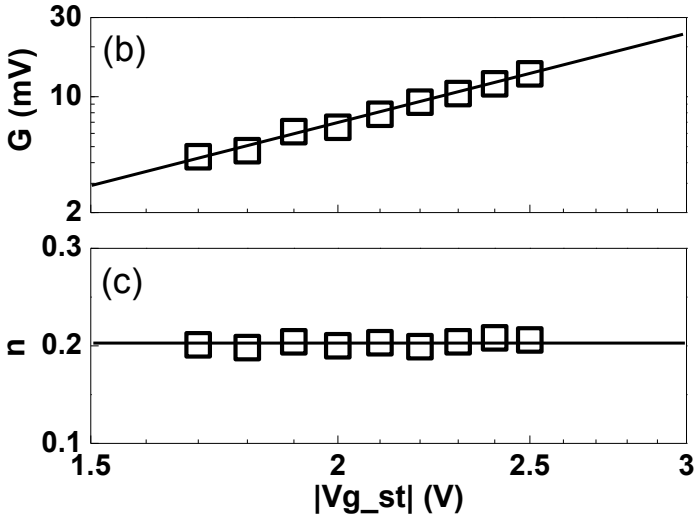
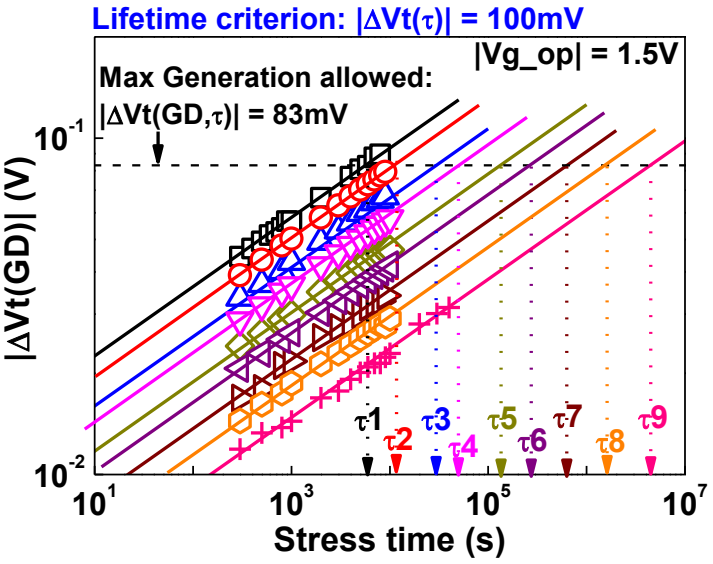


Fig. 5.6 Impact of Vg_st on AHT and GD: (a) The independence of ΔVt(AHT) on the stress Vg_st confirms they originating from 'as-grown hole traps'. At |Vg_op|=1.5 V, |ΔVt(AHT) |=17 mV. (b) The prefactor 'G' in $\Delta Vt(GD)=Gt^n$ follows a power law against Vg_st. (c) The time power exponent 'n' is independent of Vg_st.

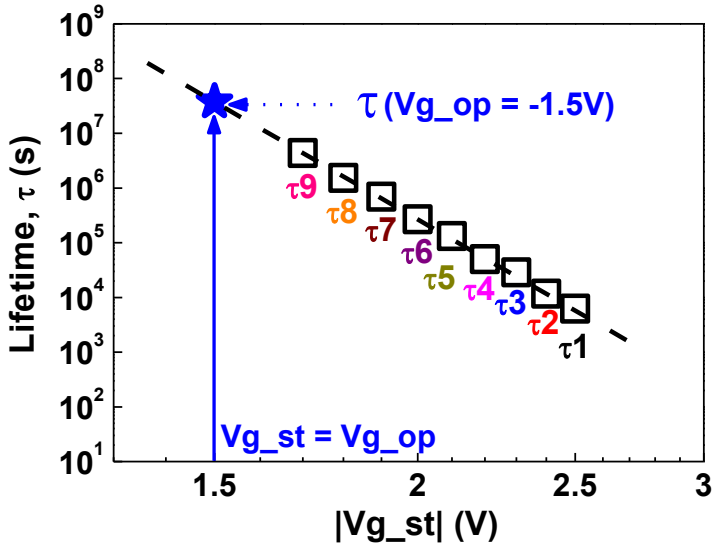
5.3.1 A step-by-step guide for lifetime prediction

This subsection gives a step-by-step guide for the proposed new lifetime prediction method. The device used in this demonstration is the 2.0 nm plasma-nitrided SiON. The device was stressed typically for 10 ks. For this demonstration, the operational voltage,Vg_op = -1.5 V is chosen. The lifetime criterion, $|\Delta Vt(\tau)|$ is selected to be of 100 mV.

Fig. 5.7 illustrates the step-by-step procedure for the lifetime prediction. The threshold voltage shifts due to the generated defects, |ΔVt(GD)| for different stress voltages, Vg_st from – 2.5 V to -1.7 V , in steps of -0.1 V as shown in Fig. 5.7 (a). The Vg_st of each measured dataset is given in Fig. 5.6 (a). From Fig. 5.6 (a), for a |Vg_op| = 1.5 V, the |ΔVt(AHT)| is 17 mV. Hence, with the lifetime criterion of 100 mV, the allowable $|\Delta Vt(GD, \tau)| = \Delta V(\tau) - \Delta V(AHT) = 100 - 17 = 83$ mV. This allowable maximum generation is indicated by the black dotted line in Fig. 5.7 (a). The lifetimes, τ, predicted under different Vg_st are extrapolated from the allowable maximum generation and are indicated by the coloured dotted lines denoted by $\tau 1$ to $\tau 9$. Each lifetime is then plotted against the stress voltages as presented in Fig. 5.7 (b). The lifetime at Vg_op is consequently obtained by extrapolating |Vg_st| to the predefined |Vg_op|, which in this demonstration is 1.5 V.



(a)



(b)

Fig. 5.7 An example of the lifetime prediction under a Vg_op= -1.5 V and a criterion of a permitted |ΔVt(τ)| =100 mV, giving an allowed ΔV(GD,τ)= ΔV(τ)-ΔV(AHT)=100-17=83 mV. (a) τ prediction under different Vg_st. The Vg_st for each dataset is given in Fig. 5.6 (a). (b) Lifetime at Vg_op is obtained by extrapolating |Vg_st| to |Vg_op|=1.5 V. τ1 to τ9 in (b) were taken from (a).

The step-by-step guide for the lifetime prediction is summarised is as below. For a given Vg_op (e.g. -1.5 V) and a permitted ΔVt(τ) (e.g. 100 mV):

i. Find ΔVt(AHT)= 17 mV at this Vg_op as shown in Fig. 5.6 (a) ;

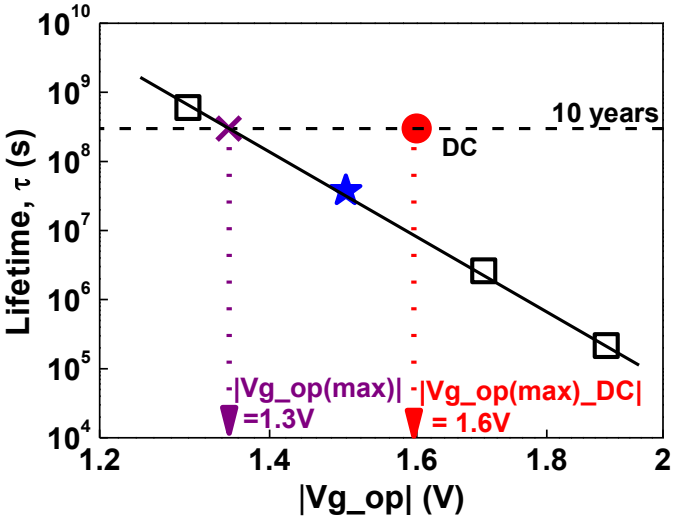
ii. Work out ΔVt(GD,τ)=ΔVt(τ)-ΔVt(AHT)=83mV;

iii. Extract τ from ΔVt(GD) vs stress time, see Fig. 5.7 (a);

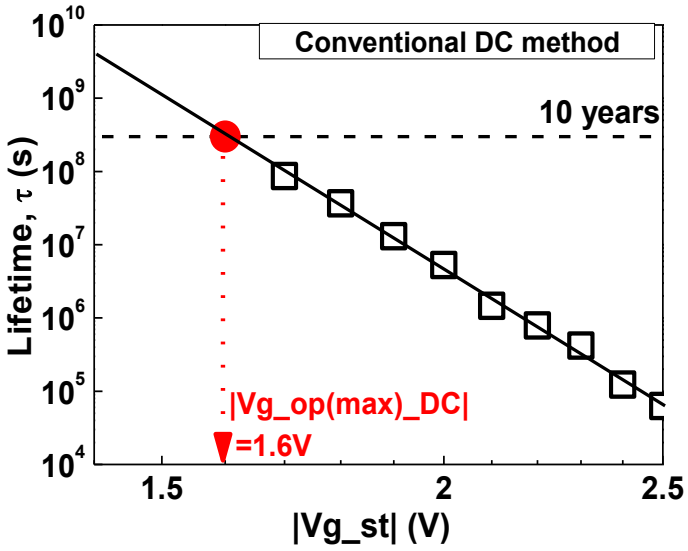
iv. Estimate τ(Vg_op=-1.5 V) by extrapolating τ against Vg_st as presented in Fig. 5.7 (b).

5.3.2 Estimating Vg_op (max) for a τ of 10 years

A standard practice in the industry is to find the maximum operational voltage, Vg_op(max) for a lifetime of 10 years. The Vg_op is now varied and for each Vg_op, the lifetime, τ is estimated by following the steps detailed in the previous subsection. This procedure is conducted until the calculated τ covers the lifetime above and below the 10 years criterion. The maximum Vg_op of τ =10 years, Vg_op(max), is then determined by interpolation from the two neighbouring points as shown in Fig. 5.8 (a). In this case, the |Vg_op(max)| extrapolated is of 1.3 V. Fig. 5.8 (a) shows that this value is substantially lower than the extrapolated value of 1.6 V where the extrapolation was conducted from the dataset obtained using the conventional DC method (see Fig. 5.8 (b).



(a)



(b)

Fig. 5.8 (a) τ at different Vg_op. '★' was taken from Fig. 5.7 (b) for |Vg_op|=1.5 V. Each τ for other Vg_op ('□') was obtained by following the procedure in Fig. 5.7. The maximum allowable Vg_op, Vg_op(max) ('X'), for τ=10 years was obtained by interpolating between two neighbouring points. '●' is the Vg_op(max) extracted by the conventional DC method, as shown in (b).

Fig. 5.9 plots the estimated Vg_op(max) under different lifetime criterion obtained from the proposed technique, as compared to that of the conventional DC method. It is evident that the difference in Vg_op(max) by the new and DC methods is insensitive to the lifetime criteria. This may be due to that both the |ΔVt(AHT)| and |ΔVt(GD)| reduce for smaller |ΔVt(τ)|, and smaller |ΔVt(τ)| always leads to lower |Vg_op(max)|. It can also be seen that the relative difference ('▲') is higher for lower |ΔVt(τ)|. From the figure, it is apparent that the DC method substantially overestimates Vg_op(max).

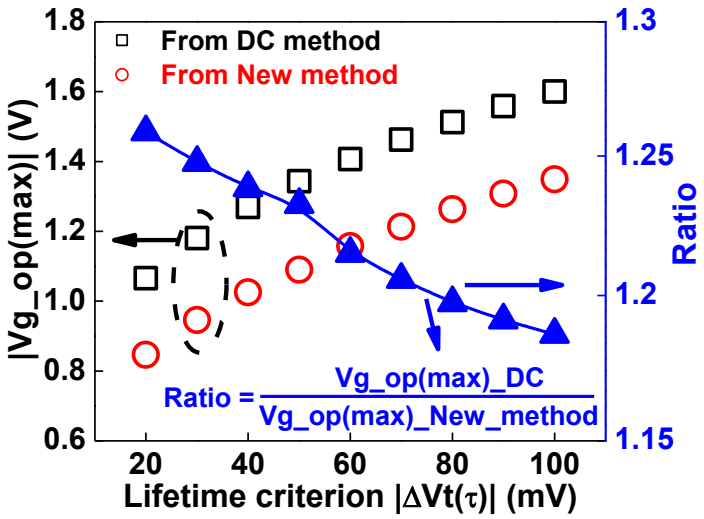


Fig. 5.9 Vg_op(max) estimated under different lifetime criteria. Smaller |ΔVt(τ)| leads to lower |Vg_op(max)|, but DC method always overestimates Vg_op(max) and the ratio ('▲') increases for smaller |ΔVt(τ)|.

5.4 Application of the proposed lifetime prediction technique to devices from different fabrication processes

This section applies the newly proposed lifetime prediction technique to devices fabricated under different process conditions. Two different processes were selected: a 2.3 nm plasma-nitided SiON with low ΔVt(AHT) and another is a 2 nm/1 nm HfSiON/SiON stack with high ΔVt(AHT). The ΔVt(AHT) against the Vg_op extracted from the two processes are presented in Fig. 5.10.

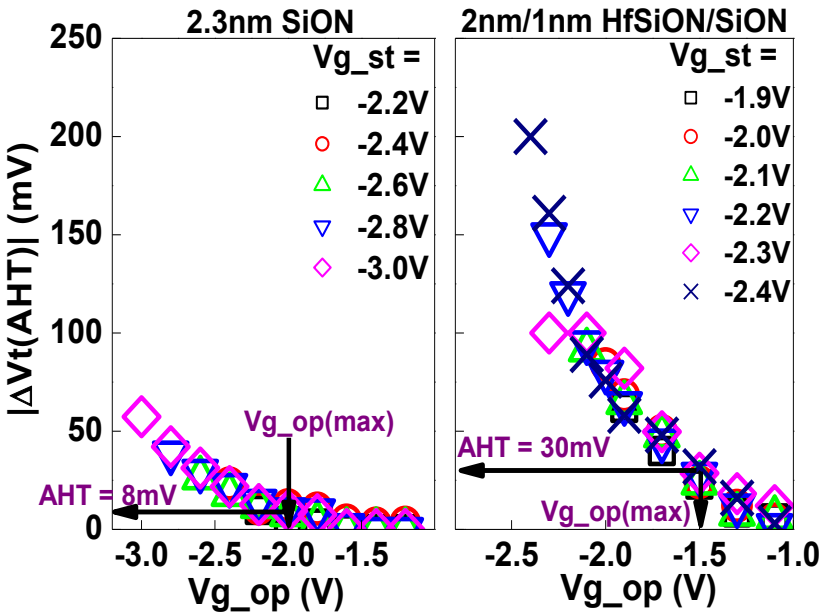


Fig. 5.10 ΔVt(AHT) vs Vg_op extracted from two processes. (a) A plasma 2.3 nm SiON with relatively low AHTs. (b) A 2nm/1nm HfSiON/SiON stack with relatively high AHTs.

The difference in the extracted Vg_op(max), as shown in Fig. 5.11 for the 2.3 nm SiON with relatively low AHT from the new and DC methods is small, which is expected. Fig. 5.12 presents that of the HfSiON/SiON high-k stack, which is large.

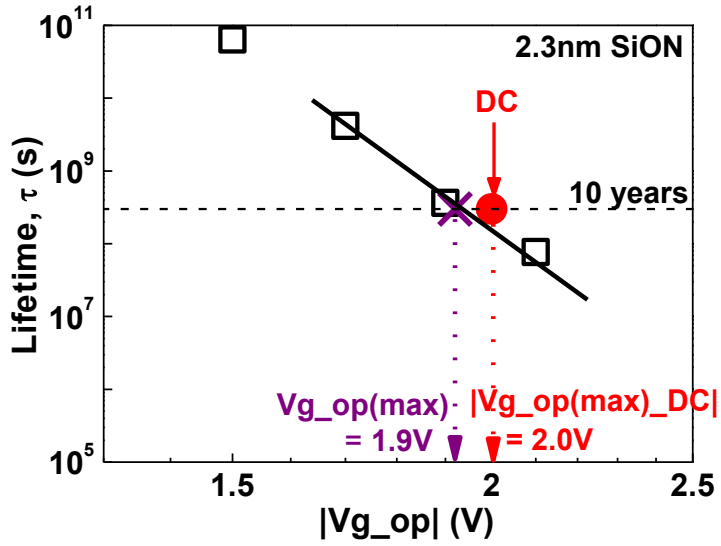


Fig. 5.11 Vg_op(max) evaluation for the 2.3 nm SiON. The relatively small difference between the new and the DC methods is because of the small ΔVt(AHT)~8 mV at Vg_op(max), as shown in Fig. 5.10(a).

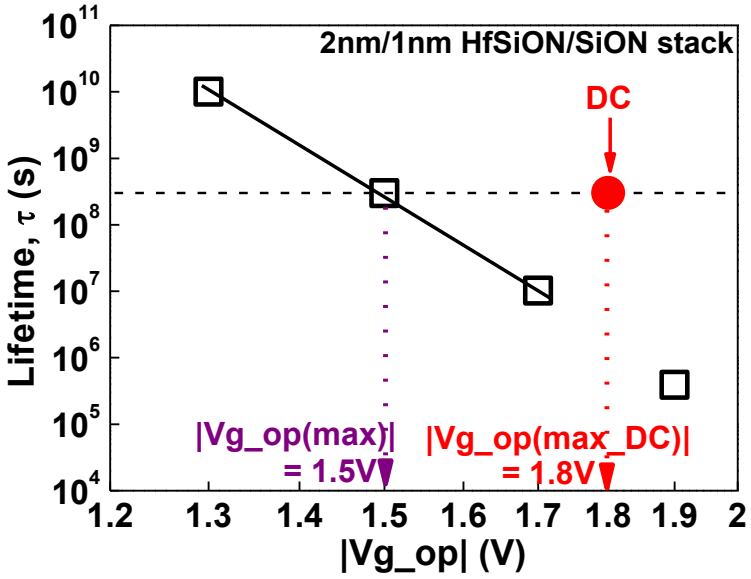


Fig. 5.12 Vg_op(max) evaluation for the 2 nm/1 nm HfSiON/SiON. The large difference between the new and the DC methods is because of the large ΔVt(AHT) ~30 mV at Vg_op(max) in Fig. 5.10 (b).

5.5 Conclusion

The DC method used by the current industry substantially overestimates Vg_op(max) due to NBTI recovery during measurements. The fast techniques suppress recovery, but wrongly include defects that are not chargeable at Vg_op for lifetime prediction. A new technique is proposed to include the correct amount of recoverable components (AHTs) in τ prediction. It does not need a trap-filling model and, consequently minimizes the number of fitting parameters and uncertainties. A major feature of this new technique is that the contribution of as-grown hole traps to the threshold voltage shift is experimentally determined. At a given temperature, it only uses two fitting parameters: G and n, the same as the DC technique, making it readily implementable in industry. The parameter 'G' covers Vg acceleration effect and the time exponent 'n' specifies the degradation kinetics.

The extracted Vg_op(max) is above typical Vg_op used in industry for the oxides studied here. In the future, Vg_op will reduce at a lower rate than EOT and the operation electrical field across the oxide will increase. A higher oxide field will drive E_f further below E_v and charge up more AHTs. A larger ΔVt(AHT) will shorten the device lifetime and NBTI can limit Vg_op(max) in the future.

6 Conclusion and Future Work

6.1 Conclusions

The work presented in this thesis has focused on the negative bias temperature instability (NBTI) and the positive charges responsible for this key reliability issue. **Chapter 1** reviewed the existing models which including the RD model, hole trapping model, two components model, the CET mapping model and the as-grown generation model, whereby each model has its own way in explaining the physical and electrical dynamics of the NBTI. The framework of the positive charges in dielectric was also been reviewed, and the chapter concluded with the rationale in the new work undertaken. **Chapter 2** had described the test facilities available and the characterization techniques applied in this work. The main research works undertaken are covered in the next three chapters. **Chapter 3** describes the newly proposed energy profiling technique which extracts the energy distribution of positive charges (PCs) in the dielectric, within and beyond the Si bandgap. The results obtained showed that the PCs vary significantly with energy level and that the PCs in different energy regions clearly originate from different defects. **Chapter 4** demonstrates that the newly proposed energy distribution technique is applicable not only to conventional single-layered SiON device but also to various advanced high-k gate stacks. The energy profiling of different Hf-based devices were investigated and the distribution of the PCs

generated subsequent to stress were reviewed. **Chapter 5** describes a newly proposed technique in the device lifetime prediction which adapts the energy profiling method and most importantly this technique can readily be implemented in industry. It is demonstrated that neither the conventional DC nor the pulse measurement is suitable for lifetime prediction due to their incorrect inclusion of PCs in the prediction.

A more detailed conclusion on the main three chapters of this thesis is given below:

6.2 Conclusions on the Energy distribution of Positive Charges in Gate Dielectric: Probing Technique and Impacts of Different Defects

The focus of this chapter was to develop a fast pulse technique that can evaluate the energy distribution of all types of PCs. Previously, a general framework which described a broad picture of the defect energy levels in the dielectric was proposed, but a detailed energy distribution of the PCs is still missing. The existing methods of NBTI measurements typically only give two levels of degradation: one before the recovery and one after the recovery at either Vg = 0 V or a certain positive level. This effectively offers information of the PCs at only two Vg points which are the stress bias before recovery and the bias used for recovery. The PCs at other bias level remain unknown. Hence, this work filled in the knowledge gap by providing a detailed energy profile in which to evaluate the PCs in the gate dielectric, within and beyond the Si bandgap. The information on the energy distribution of PCs is beneficial for assessing the impact of PCs on circuits since it gives the amount of PCs for each surface potential. Step-by-step procedure for extracting the energy distribution is summarised as below:

i. A reference I_d-V_g curve is recorded on a fresh device

ii. Device is stressed for a pre-specified time

iii. Immediately after stress, the bias is changed from V_g=Vg_st to $V_{discharge,1}$. ΔV_{th} is measured under $V_{discharge,1}$ until its variation between two points become negligible.

iv. After completing discharge at | $V_{discharge,1}$|, $|V_g|$ is reduced to | $V_{discharge,2}$| and the step iii was repeated, until eventually $V_{discharge}$ reaches the highest positive value allowable.

v. ΔV_{th} is converted to effective charge density, $\Delta N_{ox,}$ and plotted against $V_{discharge}$

vi. Convert $V_{discharge}$ to the energy level of E_f with respect to E_v at the Si/SiON interface (ie. E_f-E_v).

vii. Differentiate ΔN_{ox} to obtain the energy density of the PCs, $\Delta D_{ox.}$

The results obtained from the energy distribution had shown that the PCs are sensitive to energy level and vary substantially over the energy range. Further investigation was carried out to observe the effects of the stress time, stress temperature and nitridation technique on the energy distribution. The results strongly support the existence of different types of PCs and each of them dominates different energy regions. The PCs below Ev originate from as-grown hole traps (AHT). The AHT does not increase with either stress time or temperature and its energy density in a thermal SiON can be three times of that in a plasma SiON. The PCs distributed within the bandgap have a clear peak around the upper half of the bandgap for SiON. In contrast, the anti-neutralization positive charge (ANPC) above Ec is the only type of PCs, whose creation does not saturate with stress time and is substantially enhanced for higher stress temperature.

6.3 Conclusions for Application of Energy Probing Technique on High-k devices

The focus of this chapter is to demonstrate that the energy probing technique developed in Chapter 3 is applicable not only to the single layered devices, but also to the advanced high-k devices. Attentions had been paid to the differences in the energy distributions between the single layered SiON and the high-k/SiON stacks. The gate dielectrics of PMOSFETs that was used to demonstrate the technique include a 1.13 nm HfO_2/SiON, a 2.0 nm Al-capped HfO_2/SiO_2, a 1.52 nm FUSI-gated HfSiON/SiON, a 1.53 nm TiN/ HfSiON/SiON, and a slant-etched TaN/ HfSiON/SiON. Through comparative analysis, the correlation of the energy density is observed for different high-k stacks with varying gate material, high-k bulk material and interfacial layer (IL) thicknesses. The amount of the AHT and the CPC had been compared, and the correlation in respect to the location of the peaks has been reviewed. Similar to SiON, a high level of as-grown hole traps were observed below Ev for high-k dielectric stacks. An obvious peak had been observed near to Ec for all of the high-k stacks but this peak is missing in the single layered SiON. The defect which is responsible for this peak may be induced by the incorporation of hafnium. It was also observed that the SiON has a clear peak in the upper half of the bandgap, but there is no clear peak in the lower half. In contrast, some high-k stacks have clear peaks in both upper and lower half and the peak in the lower half can be higher than the one in the upper half. The effect of the reduction in the IL SiON of the high-k gate stack has little effect on AHT below Ev, but increases CPC within the bandgap, indicating higher NBTI for future CMOS technologies. It was concluded that process optimization is essential for minimizing NBTI of high-k stack.

6.4 Conclusions for Negative Bias Temperature Instability Lifetime Prediction: Problems and Solutions

The focus of this chapter was to overcome the limitations in the lifetime prediction by the existing techniques, including both the conventional DC and the fast measurements. Device degradation is conventionally evaluated by threshold voltage shift, ΔV_t extracted from a slow DC I_d-V_g measurement which can take from 10 ms up to several seconds long to complete. Due to the significant recovery during the slow measurement, the monitored degradations can be considerably under-estimated and thus result in the overestimation of device lifetime.

Fast techniques using pulse measurements have been proposed by many to suppress the recovery in the measurement but these proposed techniques do not give an industry-wide accepted method for lifetime prediction. The fast techniques tend to freeze the defects after stress, leading to an overestimation of the $|\Delta V_t|$ which induces an extra lowering of lifetime particularly at high |Vg_st|. This chapter proposed a new technique which was based from the energy profiling of the PCs. The method in general is based on the principle that a defect must be chargeable at the operation E(Vg_op), if it is to be included in lifetime prediction. A step-by-step guide for the lifetime prediction is summarised as below. The AHT and GD denotes the as-grown hole traps and generated defects respectively. For a given operational voltage, Vg_op and a permitted lifetime criterion, $\Delta V_t(\tau)$:

i. Find $\Delta Vt(AHT)$ at this Vg_op ;

ii. Work out $\Delta Vt(GD,\tau)=\Delta Vt(\tau)-\Delta Vt(AHT)$;

iii. Extract τ from $\Delta Vt(GD)$ vs stress time

iv. Estimate τ(Vg_op) by extrapolating τ against Vg_st

To estimate the maximum operation al voltage, Vg_op (max) for a device lifetime of 10 years, Vg_op is now varied and for each Vg_op, the lifetime,τ is estimated by following the steps i to iv. This procedure is conducted until the calculated τ covers the lifetime above and below the 10 years criterion. The Vg_op(max), is then determined by interpolation at the intersection of the 10 year criterion and the measured Vg_op points.

It is observed that the |Vg_op(max)| obtained from this technique was substantially lower from the |Vg_op(max)| obtained from the conventional DC measurement. It was also found that the difference in Vg_op(max) by the new and DC methods is insensitive to the lifetime criteria. Demonstration of the applicability of this new lifetime prediction technique to different device fabrication processes was also conducted.

6.5 Future Work

Despite the progress made in this project, there are many problems remaining to be solved, including but not limited to, the following:

Device lifetime prediction under AC stresses

The stress in this work is DC stress, representing the worst degradation scenario. The pMOSFETs in some circuits, such as SRAM where a memory bitcell does not flip, are subjected to the DC stress. Many circuits, however, will operate under AC conditions, where recovery of NBTI will occur when the pMOSFETs were switched off. A typical practice to take this recovery into account is to use a duty cycle. The proposed new technique for lifetime prediction does not include this duty cycle and cannot be used for predicting the lifetime under the AC stress. Further work should be carried out to extend this technique to cover the AC stress.

Device lifetime prediction for nano-size pMOSFETs

The work in this project was carried out on relatively large devices, where the device-to-device variation is negligible. For the nano-meter-size MOSFETs, there are substabtial device-to-device variations. After fabrication, the variation originates from the random dopant fluctuation, line edge roughness, gate work function fluctuation, and oxide thickness variation. For NBTI, the discreteness of positive charges will introduce a time-dependent device-to-device variation. This time-dependent variation has not been taken into account by the newly proposed lifetime prediction technique and further work is needed to address this issue.

Energy Profiling for multi-gate MOSFETs and nano-wire MOSFETs:

Multi-gate MOSFETs and nano-wire MOSFETs have been developed to increase the gate control on the substrate and reduce the drain-induced barrier lowering leakage current. However, there is limited information on the defects properties and their lifetime. Hence, the applicability of the proposed energy profile and lifetime techniques to these types of devices should be explored in the future.

Energy Profiling for other dielectric/semiconductor structures:

A lot of attentions have been paid to further improve transistor speed as the downscaling of silicon based MOSFETs reaches to its limit. The NBTI properties of both Germanium and III-V semiconductor MISFETs have been investigated. Due to the fact that the dielectric stack used in these transistors do not involve the well –known SiON, there are still significant gaps on the knowledge of the defect properties and the lifetime information on these new devices. It is worth to study the energy profiling of these devices in order to understand the devices characteristics and limitations.

References

[1] Terman, L. M. (1962). An investigation of surface states at a silicon/silicon oxide interface employing metal-oxide-silicon diodes. *Solid-State Electronics*, *5*(5), 285-299.

[2] Deal, B. E., Sklar, M., Grove, A. S., & Snow, E. H. (1967). Characteristics of the Surface-State Charge (Qss) of Thermally Oxidized Silicon. *Journal of The Electrochemical Society*, *114*(3), 266-274.

[3] Schaller, R. R. (1997). Moore's law: past, present and future. *Spectrum, IEEE*, *34*(6), 52-59.

[4] Moore, G. E. (1976). Microprocessors and integrated electronic technology. *Proceedings of the IEEE*, *64*(6), 837-841.

[5] Moore, G. E. (1998). The role of Fairchild in silicon technology in the early days of "Silicon Valley". *Proceedings of the IEEE*, *86*(1), 53-62.

[6] Moore, G. E. (1965). Cramming more components onto integrated circuits. *Proceedings of the IEEE,* vol. 86, pp. 82-85, 1998

[7] Goodall, R., Fandel, D., Allan, A., Landler, P., & Huff, H. R. (2002, May). Long term productivity mechanisms of the semiconductor industry. In *Ninth International Symposium on Silicon Materials Science and Technology*.

[8] International Technology Roadmap for Semiconductors, 2011 edition, 'Emerging Research Devices'.

[9] Pantelides, S. T. (1978). *The Physics of SiO2 and Its Interfaces, Proceedings of the International Topical Conference, Yorktown Heights, New York, March 22-24, 1978*. IBM THOMAS J WATSON RESEARCH CENTER YORKTOWN HEIGHTS NY.

[10] Weber, W., & Thewes, R. (1995). Hot-carrier-related device reliability for digital and analogue CMOS circuits. *Semiconductor science and technology*, *10*(11), 1432.

[11] Zhang, J. F., & Eccleston, W. (1995). Effects of high field injection on the hot carrier induced degradation of submicrometer pMOSFET's. *Electron Devices, IEEE Transactions on*, *42*(7), 1269-1276.

[12] Degraeve, R., Groeseneken, G., Bellens, R., Ogier, J. L., Depas, M., Roussel, P. J., & Maes, H. E. (1998). New insights in the relation between electron trap generation and the statistical properties of oxide breakdown. *Electron Devices, IEEE Transactions on*, *45*(4), 904-911.

[13] B. J. Cheng, A. R. Brown, and A. Asenov, IEEE Elec. Dev. Lett. 32, 740 (2011).

[14] Jeppson, K. O., & Svensson, C. M. (1977). Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices. *Journal of Applied Physics*, *48*(5), 2004-2014.

[15] W Schemmert, W., Gabler, L., & Hoefflinger, B. (1976). Conductance of ion-implanted buried-channel MOS transistors. *Electron Devices, IEEE Transactions on*, *23*(12), 1313-1319.

[16] Alam, M. A., Kufluoglu, H., Varghese, D., & Mahapatra, S. (2007). A comprehensive model for PMOS NBTI degradation: Recent progress. *Microelectronics Reliability*, *47*(6), 853-862.

[17] Kaczer, B., Arkbipov, V., Degraeve, R., Collaert, N., Groeseneken, G., & Goodwin, M. (2005, April). Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification. In *Reliability Physics Symposium, 2005. Proceedings. 43rd Annual. 2005 IEEE International* (pp. 381-387). IEEE.

[18] Huard, V., Denais, M., Perrier, F., Revil, N., Parthasarathy, C., Bravaix, A., & Vincent, E. (2005). A thorough investigation of MOSFETs NBTI degradation. *Microelectronics Reliability*, *45*(1), 83-98.

[19] Veloso, A., Ragnarsson, L. Å., Schram, T., Chew, S. A., Boccardi, G., Thean, A., & Horiguchi, N. (2013). Integration Challenges and Options of Replacement High-κ/Metal Gate Technology for (Sub-) 22nm Technology Nodes. *ECS Transactions*, *52*(1), 385-390.

[20] Tan, S. S., Chen, T. P., Soon, J. M., Loh, K. P., Ang, C. H., & Chan, L. (2003). Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations. *Applied physics letters*, *82*(12), 1881-1883

[21] Kimizuka, N., Yamamoto, T., Mogami, T., Yamaguchi, K., Imai, K., & Horiuchi, T. (1999). The impact of bias temperature instability for direct-tunneling ultra-thin gate

oxide on MOSFET scaling. In *VLSI Technology, 1999. Digest of Technical Papers. 1999 Symposium on* (pp. 73-74). IEEE.

[22] Kimizuka, N., Yamaguchi, K., Imai, K., Iizuka, T., Liu, C. T., Keller, R. C., & Horiuchi, T. (2000). NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10-μm gate CMOS generation. In *VLSI Technology, 2000. Digest of Technical Papers. 2000 Symposium on* (pp. 92-93). IEEE.

[23] Alam, M. A. (2003, December). A critical examination of the mechanics of dynamic NBTI for PMOSFETs. In *Electron Devices Meeting, 2003. IEDM'03 Technical Digest. IEEE International* (pp. 14-4). IEEE.

[24] Huard, V. (2010, May). Two independent components modeling for negative bias temperature instability. In *Reliability Physics Symposium (IRPS), 2010 IEEE International* (pp. 33-42). IEEE.

[25] Grasser, T., Kaczer, B., Goes, W., Aichinger, T., Hehenberger, P., & Nelhiebel, M. (2009, April). A two-stage model for negative bias temperature instability. In *Reliability Physics Symposium, 2009 IEEE International* (pp. 33-44). IEEE.

[26] Ji, Z., Lin, L., Zhang, J. F., Kaczer, B., & Groeseneken, G. (2010). NBTI lifetime prediction and kinetics at operation bias based on ultrafast pulse measurement. *Electron Devices, IEEE Transactions on*, *57*(1), 228-237.

[27] Ogawa, S., & Shiono, N. (1995). Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO_ {2} interface. *Physical Review B*, *51*(7), 4218.

[28] Alam, M., Weir, B., & Silverman, P. (2001, November). The prospect of using thin oxides for silicon nanotransistors. In *Gate Insulator, 2001. IWGI 2001. Extended Abstracts of International Workshop on* (pp. 30-34). IEEE.

[29] M. A. Alam. (2003) A critical examination of the mechanism of dynamic NBTI for PMOSFET. In *International Electron Devices Meeting (IEDM) Digest*, pp. 346-350. IEEE

[30] Kufluoglu, H., & Ashraful Alam, M. (2004, December). A geometrical unification of the theories of NBTI and HCI time-exponents and its implications for ultra-scaled

planar and surround-gate MOSFETs. In *Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International* (pp. 113-116). IEEE.

[31] Krishnan, A. T., Reddy, V., Chakravarthi, S., Rodriguez, J., John, S., & Krishnan, S. (2003, December). NBTI impact on transistor and circuit: models, mechanisms and scaling effects [MOSFETs]. In *Electron Devices Meeting, 2003. IEDM'03 Technical Digest. IEEE International* (pp. 14-5). IEEE.

[32] Blat, C. E., Nicollian, E. H., & Poindexter, E. H. (1991). Mechanism of negative-bias-temperature instability. *Journal of Applied Physics*, *69*(3), 1712-1720.

[33] Kakalios, J., Street, R. A., & Jackson, W. B. (1987). Stretched-exponential relaxation arising from dispersive diffusion of hydrogen in amorphous silicon. *Physical review letters*, *59*(9), 1037.

[34] Zafar, S., Lee, B. H., Stathis, J., Callegari, A., & Ning, T. (2004, June). A model for negative bias temperature instability (NBTI) in oxide and high κ pFETs In *VLSI Technology, 2004. Digest of Technical Papers. 2004 Symposium on* (pp. 208-209). IEEE.

[35] K. Hess, A. Haggag, W. McMahon, B. Fischer, K. Cheng, J. Lee, and J.Lyding, "Simulation of Si-SiO2 defect generation in CMOS chips: from atomistic structure to chip failure rates," in *International Electron DevicesMeeting (IEDM) Digest*, pp. 93-96, 2000.

[36] Grasser, T., Entner, R., Triebl, O., Enichlmair, H., & Minixhofer, R. (2006, September). TCAD modeling of negative bias temperature instability. In *Simulation of Semiconductor Processes and Devices, 2006 International Conference on* (pp. 330-333). IEEE.

[37] R.Etner, T.Grasser, O.Triebl, H.Enichlmair, R. Minixhofer, "Negative Bias Temperature Instability Modeling For High Voltage Oxides At Different Stress Temperatures", Microelectronic Reliability 47, pg 697-699, 2007.

[38] Mahapatra. S, Bharat Kumar P, Alam MA, A new obervation of enhances bias temperature instability in thin gate oxide p-MOSFETs, Proc. Int Electronic Device Meet, 2003, p.337-41

[39] Grasser, T., Goes, W., & Kaczer, B. (2009). Critical modeling issues in negative bias temperature instability. *ECS Transactions*, *19*(2), 265-287

[40] Huard, V., Denais, M., & Parthasarathy, C. (2006). NBTI degradation: From physical mechanisms to modelling. *Microelectronics Reliability*, *46*(1), 1-23

[41] Reisinger, H., Blank, O., Heinrigs, W., Muhlhoff, A., Gustin, W., & Schlunder, C. (2006, March). Analysis of NBTI degradation-and recovery-behavior based on ultra fast VT-measurements. In *Reliability Physics Symposium Proceedings, 2006. 44th Annual., IEEE International* (pp. 448-453). IEEE.

[42] Grasser, T., Gos, W., Sverdlov, V., & Kaczer, B. (2007, April). The universality of NBTI relaxation and its implications for modeling and characterization. In *Reliability physics symposium, 2007. proceedings. 45th annual. ieee international* (pp. 268-280). IEEE.

[43] Kaczer, B., Grasser, T., Roussel, P. J., Martin-Martinez, J., O'Connor, R., O'Sullivan, B. J., & Groeseneken, G. (2008, April). Ubiquitous relaxation in BTI stressing—New evaluation and insights. In *Reliability Physics Symposium, 2008. IRPS 2008. IEEE International* (pp. 20-27). IEEE.

[44] Kaczer, B., Arkbipov, V., Degraeve, R., Collaert, N., Groeseneken, G., & Goodwin, M. (2005, April). Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification. In *Reliability Physics Symposium, 2005. Proceedings. 43rd Annual. 2005 IEEE International* (pp. 381-387). IEEE.

[45] Houssa, M., Aoulaiche, M., De Gendt, S., Groeseneken, G., Heyns, M. M., & Stesmans, A. (2005). Reaction-dispersive proton transport model for negative bias temperature instabilities. *Applied Physics Letters*, *86*(9), 093506-093506.

[46] Zafar, S. (2005). Statistical mechanics based model for negative bias temperature instability induced degradation. *Journal of applied physics*, *97*(10), 103709-103709.

[47] Varghese, D., Saha, D., Mahapatra, S., Ahmed, K., Nouri, F., & Alam, M. (2005, December). On the dispersive versus Arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: Measurements, theory, and implications. In *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International* (pp. 684-687). IEEE.

[48] Kaczer, B., Arkhipov, V., Degraeve, R., Collaert, N., Groeseneken, G., & Goodwin, M. (2005). Temperature dependence of the negative bias temperature instability in the framework of dispersive transport. *Applied Physics Letters*, *86*(14), 143506-143506.

[49] Islam, A. E., Kufluoglu, H., Varghese, D., Mahapatra, S., & Alam, M. A. (2007). Recent issues in negative-bias temperature instability: Initial degradation, field dependence of interface trap generation, hole trapping effects, and relaxation. *Electron Devices, IEEE Transactions on*, *54*(9), 2143-2154.

[50] Grasser, T., Gos, W., & Kaczer, B. (2008). Dispersive transport and negative bias temperature instability: Boundary conditions, initial conditions, and transport models. *Device and Materials Reliability, IEEE Transactions on*, *8*(1), 79-97.

[51] Deora, S., Maheta, V. D., & Mahapatra, S. (2010, May). NBTI lifetime prediction in SiON p-MOSFETs by H/H2 Reaction-Diffusion (RD) and Dispersive hole trapping model. In *Reliability Physics Symposium (IRPS), 2010 IEEE International* (pp. 1105-1114). IEEE.

[52] Tewksbury III, T. L. (1992). Relaxation Effects in MOS Devices due to Tunnel Exchange with Near-Interface Oxide Traps.

[53] Fleetwood, D. M., Xiong, H. D., Lu, Z. Y., Nicklaw, C. J., Felix, J. A., Schrimpf, R. D., & Pantelides, S. T. (2002). Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices. *Nuclear Science, IEEE Transactions on*, *49*(6), 2674-2683.

[54] Kirton, M. J., Uren, M. J., Collins, S., Schulz, M., Karmann, A., & Scheffer, K. (1989). Individual defects at the Si: SiO2 interface. *Semiconductor Science and Technology*, *4*(12), 1116.

[55] Makram-Ebeid, S., & Lannoo, M. (1982). Quantum model for phonon-assisted tunnel ionization of deep levels in a semiconductor. *Physical Review B*, *25*(10), 6406.

[56] Ganichev, S. D., Prettl, W., & Yassievich, I. N. (1997). Deep impurity-center ionization by far-infrared radiation. *Physics of the Solid State*, *39*(11), 1703-1726

[57] Zhang, J. F., Sii, H. K., Groeseneken, G., & Degraeve, R. (2001). Hole trapping and trap generation in the gate silicon dioxide. *Electron Devices, IEEE Transactions on*, *48*(6), 1127-1135.

[58] Zhang, J. F., Sii, H. K., Chen, A. H., Zhao, C. Z., Uren, M. J., Groeseneken, G., & Degraeve, R. (2004). Hole trap generation in gate dielectric during substrate hole injection. *Semiconductor science and technology*, *19*(1), L1.

[59] Zhang, J. F. (2009). Defects and instabilities in Hf-dielectric/SiON stacks. *Microelectronic Engineering*, *86*(7), 1883-1887.

[60] Zhang, J. F., Zhao, C. Z., Sii, H. K., Groeseneken, G., Degraeve, R., Ellis, J. N., & Beech, C. D. (2002). Relation between hole traps and hydrogenous species in silicon dioxides. *Solid-State Electronics*, *46*(11), 1839-1847.

[61] Huard, V., Parthasarathy, C., Rallet, N., Guerin, C., Mammase, M., Barge, D., & Ouvrard, C. (2007, December). New characterization and modeling approach for NBTI degradation from transistor to product level. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International* (pp. 797-800). IEEE.

[62] Huard, V. (2010, May). Two independent components modeling for negative bias temperature instability. In *Reliability Physics Symposium (IRPS), 2010 IEEE International* (pp. 33-42). IEEE.

[63] Grasser, T., Wagner, P., Reisinger, H., Aichinger, T., Pobegen, G., Nelhiebel, M., & Kaczer, B. (2011, December). Analytic modeling of the bias temperature instability using capture/emission time maps. In *Electron Devices Meeting (IEDM), 2011 IEEE International* (pp. 27-4). IEEE.

[64] H. Reisinger, T. Grasser, W. Gustin, and C. Schlunder, in *Proc Intl.Rel.Phys.Symp. (IRPS)* (2010), pp. 7–15.

[65] H. Reisinger, T. Grasser, K. Ermisch, H. Nielen, W. Gustin, and C.Schlunder, in *Proc. Intl.Rel.Phys.Symp. (IRPS)* (2011), pp. 597–604.

[66] Huard, V., Parthasarathy, C. R., & Denais, M. (2005, October). Single-hole detrapping events in pMOSFETs NBTI degradation. In *Integrated Reliability Workshop Final Report, 2005 IEEE International* (pp. 5-pp). IEEE.

[67] Ma, H. C., Chiu, J. P., Tang, C. J., Wang, T., & Chang, C. S. (2009, April). Investigation of post-NBT stress current instability modes in HfSiON gate dielectric pMOSFETs by measurement of individual trapped charge emissions. In *Reliability Physics Symposium, 2009 IEEE International* (pp. 51-54). IEEE.

[68] Kumar, E. N., Maheta, V. D., Purawat, S., Islam, A. E., Olsen, C., Ahmed, K., & Mahapatra, S. (2007, December). Material Dependence of NBTI Physical Mechanism in Silicon Oxynitride (SiON) p-MOSFETs: A Comprehensive Study by Ultra-Fast On-The-Fly (UF-OTF) IDLIN Technique. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International* (pp. 809-812). IEEE.

[69] Aoulaiche, M., Kaczer, B., De Jaeger, B., Houssa, M., Martens, K., Degraeve, R., & Heyns, M. M. (2008, April). Negative bias temperature instability on Si-passivated Ge-interface. In *Reliability Physics Symposium, 2008. IRPS 2008. IEEE International* (pp. 358-362). IEEE.

[70] Tan, S. S., Chen, T. P., Soon, J. M., Loh, K. P., Ang, C. H., & Chan, L. (2003). Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations. *Applied physics letters*, *82*(12), 1881-1883.

[71] Trombetta, L. P., Feigl, F. J., & Zeto, R. J. (1991). Positive charge generation in metal-oxide-semiconductor capacitors. *Journal of applied physics*, *69*(4), 2512-2521.

[72] Freitag, R. K., Brown, D. B., & Dozier, C. M. (1994). Evidence for two types of radiation-induced trapped positive charge. *Nuclear Science, IEEE Transactions on*, *41*(6), 1828-1834.

[73] Young, D. R., Irene, E. A., DiMaria, D. J., De Keersmaecker, R. F., & Massoud, H. Z. (1979). Electron trapping in SiO 2 at 295 and 77 K. *Journal of Applied Physics*, *50*(10), 6366-6372.

[74] Stahlbush, R. E., Cartier, E., & Buchanan, D. A. (1995). Anomalous positive charge formation by atomic hydrogen exposure. *Microelectronic Engineering*, *28*(1), 15-18.

[75] Lai, S. K., & Young, D. R. (1981). Effects of avalanche injection of electrons into silicon dioxide—Generation of fast and slow interface states. *Journal of Applied Physics*, *52*(10), 6231-6240.

[76] Fleetwood, D. M. (2002). Effects of hydrogen transport and reactions on microelectronics radiation response and reliability. *Microelectronics Reliability*, *42*(4), 523-541.

[77] Lelis, A. J., & Oldham, T. R. (1994). Time dependence of switching oxide traps. *Nuclear Science, IEEE Transactions on*, *41*(6), 1835-1843.

[78] Zhang, J. F., Zhao, C. Z., Chen, A. H., Groeseneken, G., & Degraeve, R. (2004). Hole traps in silicon dioxides. Part I. Properties. *Electron Devices, IEEE Transactions on*, *51*(8), 1267-1273.

[79] Zhao, C. Z., Zhang, J. F., Groeseneken, G., & Degraeve, R. (2004). Hole-traps in silicon dioxides. Part II. Generation mechanism. *Electron Devices, IEEE Transactions on*, *51*(8), 1274-1280.

[80] Zhao, C. Z., Zhang, J. F., Chang, M. H., Peaker, A. R., Hall, S., Groeseneken, G. & Heyns, M. (2008). Stress-induced positive charge in Hf-based gate dielectrics: impact on device performance and a framework for the defect. *Electron Devices, IEEE Transactions on*, *55*(7), 1647-1656.

[81] Zhao, C. Z., & Zhang, J. F. (2005). Effects of hydrogen on positive charges in gate oxides. *Journal of applied physics*, *97*(7), 073703-073703.

[82] Ortiz-Conde, A., Garcia Sánchez, F. J., Liou, J. J., Cerdeira, A., Estrada, M., & Yue, Y. (2002). A review of recent MOSFET threshold voltage extraction methods. Microelectronics Reliability, 42(4), 583-596

[83] Ang, D. S., & Wang, S. (2006). Recovery of the NBTI-stressed ultrathin gate p-MOSFET: The role of deep-level hole traps. Electron Device Letters, IEEE, 27(11), 914-916.

[84] Brugler, J. S., & Jespers, P. G. (1969). Charge pumping in MOS devices. Electron Devices, IEEE Transactions on, 16(3), 297-302.

[85] Van den Bosch, G., Groeseneken, G., & Maes, H. E. (1993). On the geometric component of charge-pumping current in MOSFETs. Electron Device Letters, IEEE, 14(3), 107-109.

[86] Denais, M., Parthasarathy, C., Ribes, G., Rey-Tauriac, Y., Revil, N., Bravaix, A., & Perrier, F. (2004, December). On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's. In Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International (pp. 109-112). IEEE.

[87] Kerber, A., Cartier, E., Pantisano, L., Rosmeulen, M., Degraeve, R., Kauerauf, T. & Schwalke, U. (2003). Characterization of the Vt-instability un SiO2 HFO2 gate dielectrics. status: published, 41-45.

[88] Semiconductor Material and Device Characterization, D. K. Schroder, John Wiley and Sons Inc., 2nd edition, 1998

[89] Keithley Application Note Series: Gate Dielectric Capacitance-Voltage Characterization Using the Model 4200 Semiconductor Characterization System, 2006.

[90] Chen, G., Li, M. F., Ang, C. H., Zheng, J. Z., & Kwong, D. L. (2002). Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling. Electron Device Letters, IEEE, 23(12), 734-736.

[91] Ershov, M., Lindley, R., Saxena, S., Shibkov, A., Minehane, S., Babcock, J. & Redford, M. (2003, April). Transient effects and characterization methodology of negative bias temperature instability in pMOS transistors. In Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International (pp. 606-607). IEEE.

[92] Ershov, M., Saxena, S., Karbasi, H., Winters, S., Minehane, S., Babcock, J. & Shibkov, A. (2003). Dynamic recovery of negative bias temperature instability in p-type metal–oxide–semiconductor field-effect transistors. Applied physics letters, 83(8), 1647-1649.

[93] Usui, H., Kanno, M., & Morikawa, T. (2003). Time and voltage dependence of degradation and recovery under pulsed negative bias temperature stress. In Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International (pp. 610-611). IEEE.

[94] Zhu, B., Suehle, J. S., Chen, Y., & Bernstein, J. B. (2002, October). Negative bias temperature instability of deep sub-micron p-MOSFETs under pulsed bias stress. In Integrated Reliability Workshop Final Report, 2002. IEEE International (pp. 125-129). IEEE.

[95] Ang, D. S., & Wang, S. (2006). Recovery of the NBTI-stressed ultrathin gate p-MOSFET: The role of deep-level hole traps. Electron Device Letters, IEEE, 27(11), 914-916.

[96] Denais, M., Bravaix, A., Huard, V., Parthasarathy, C., Guerin, C., Ribes, G. & Roy, D. (2006, March). Paradigm shift for NBTI characterization in ultra-scaled CMOS technologies. In Reliability Physics Symposium Proceedings, 2006. 44th Annual., IEEE International (pp. 735-736). IEEE.

[97] Rangan, S., Mielke, N., & Yeh, E. C. C. (2003, December). Universal recovery behavior of negative bias temperature instability [PMOSFETs]. In Electron Devices Meeting, 2003. IEDM'03 Technical Digest. IEEE International (pp. 14-3). IEEE.

[98] Reisinger, H., Blank, O., Heinrigs, W., Gustin, W., & Schlunder, C. (2007). A comparison of very fast to very slow components in degradation and recovery due to NBTI and bulk hole trapping to existing physical models. Device and Materials Reliability, IEEE Transactions on, 7(1), 119-129.

[99] Islam, A. E., Kumar, E. N., Das, H., Purawat, S., Maheta, V., Aono, H. & Alam, M. A. (2007, December). Theory and Practice of On-the-fly and Ultra-fast VT Measurements for NBTI Degradation: Challenges and Opportunities. In Electron Devices Meeting, 2007. IEDM 2007. IEEE International (pp. 805-808). IEEE.

[100] Hehenberger, P., Aichinger, T., Grasser, T., Gos, W., Triebl, O., Kaczer, B., & Nelhiebel, M. (2009, April). Do NBTI-induced interface states show fast recovery? A study using a corrected on-the-fly charge-pumping measurement technique. In Reliability Physics Symposium, 2009 IEEE International (pp. 1033-1038). IEEE.

[101] Kerber, A., Cartier, E., Pantisano, L., Rosmeulen, M., Degraeve, R., Kauerauf, T.. & Schwalke, U. (2003). Characterization of the Vt-instability un SiO2 HFO2 gate dielectrics. status: published, 41-45.

[102] Van den Bosch, G., Groeseneken, G., & Maes, H. E. (1993). On the geometric component of charge-pumping current in MOSFETs. Electron Device Letters, IEEE, 14(3), 107-109.

[103] B. J. Cheng, A. R. Brown, and A. Asenov, “Impact of NBTI/PBTI on SRAM Stability Degradation,” *IEEE Electron Device Lett.,* vol. 32, no. 6, pp. 740-742, Jun, 2011.

[104] M.-F. Li, D. Huang, W. J. Liu, Z. Y. Liu, and X. Y. Huang, "New Insights of BTI degradation in MOSFETs with SiON Gate Dielectrics," *Silicon Nitride, Silicon Dioxide, and Emerging Dielectrics 10*, ECS Transactions 2, pp. 301-318, 2009.

[105] J. F. Zhang, “Defects and instabilities in Hf-dielectric/SiON stacks (Invited Paper),” *Microelectron. Eng.,* vol. 86, no. 7-9, pp. 1883-1887, 2009.

[106] T. Aichinger, M. Nelhiebel, S. Einspieler, and T. Grasser, “Observing two stage recovery of gate oxide damage created under negative bias temperature stress,” *J. Appl. Phys.,* vol. 107, no. 2, Jan, 2010.

[107] X. Ji, Y. Liao, F. Yan, Y. Shi, G. Zhang, and Q. Guo, "The energy distribution of NBTI-induced hole traps in the Si band gap in PNO pMOSFETs," in *Int. Reliab. Phys. Symp.*, pp. XT.12.11-XT.12.15, 2012.

[108] S. Mahapatra, V. D. Maheta, S. Deora, E. N. Kumar, S. Purawat, C. Olsen, K. Ahmed, A. E. Islam, and M. A. Alam, "Material Dependence of Negative Bias Temperature Instability (NBTI) Stress and Recovery in SiON p-MOSFETs," *Silicon Nitride, Silicon Dioxide, and Emerging Dielectrics 10*, ECS Transactions 2, R. E. Sah, J. M. Deen, A. Toriumi *et al.*, eds., pp. 243-263, Pennington: Electrochemical Society Inc, 2009.

[109] D. S. Ang, Y. Z. Hu, Z. Q. Teo, G. A. Du, S. C. S. Lai, and T. J. J. Ho, "Negative-Bias Temperature Instability: Measurement and Degradation Mechanisms," *Silicon Nitride, Silicon Dioxide, and Emerging Dielectrics 10*, ECS Transactions 2, R. E. Sah, J. M. Deen, A. Toriumi *et al.*, eds., pp. 147-176, Pennington: Electrochemical Society Inc, 2009.

[109] Z. Ji, J. F. Zhang, M. H. Chang, B. Kaczer, and G. Groeseneken, “An Analysis of the NBTI-Induced Threshold Voltage Shift Evaluated by Different Techniques,” *IEEE Trans. Electron Devices,* vol. 56, no. 5, pp. 1086, 2009.

[110] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, “Characteristics of the surface-state charge (Qss) of thermally oxidized silicon,” *J. Electrochem. Soc.,* vol. 114, pp. 266-274, 1967.

[111] R. E. Stahlbush, E. Cartier, and D. A. Buchanan, “ANOMALOUS POSITIVE CHARGE FORMATION BY ATOMIC-HYDROGEN EXPOSURE,” *Microelectron. Eng.,* vol. 28, no. 1-4, pp. 15-18, Jun, 1995.

[112] S. K. Lai, and D. R. Young, “EFFECTS OF AVALANCHE INJECTION OF ELECTRONS INTO SILICON DIOXIDE - GENERATION OF FAST AND SLOW INTERFACE STATES,” *J. Appl. Phys.,* vol. 52, no. 10, pp. 6231-6240, 1981.

[113] J. F. Zhang, C. Z. Zhao, A. H. Chen, G. D. Groeseneken, and R. Degraeve, “Hole traps in silicon dioxides - Part I: Properties,” *IEEE Trans. Electron Devices,* vol. 51, no. 8, pp. 1267, 2004.

[80] C. Z. Zhao, J. F. Zhang, M. H. Chang, A. R. Peaker, S. Hall, G. Groeseneken, L. Pantisano, S. De Gendt, and M. Heyns, “Stress-induced positive charge in Hf-based gate dielectrics: Impact on device performance and a framework for the defect,” *IEEE Trans. Electron Devices,* vol. 55, no. 7, pp. 1647-1656, Jul, 2008.

[114] C. Z. Zhao, J. F. Zhang, G. Groeseneken, and R. Degraeve, “Hole-Traps in Silicon Dioxides - Part II: Generation Mechanism ” *IEEE Trans. Electron Devices,* vol. 51, no. 8, pp. 1274, 2004.

[115] M. H. Chang, and J. F. Zhang, “On positive charge formed under negative bias temperature stress,” *J. Appl. Phys.,* vol. 101, no. 2, 2007.

[116] M. Duan, J. F. Zhang, Z. Ji, W. Zhang, B. Kaczer, S. D. Gendt, and G. Groeseneken, “Defect loss: A new concept for reliability of MOSFETs,” *IEEE Electron Device Lett.,* vol. 33, no. 4, pp. 480-482, 2012.

[117] J. F. Zhang, Z. Ji, M. H. Chang, B. Kaczer, and G. Groeseneken, "Real Vth instability of pMOSFETs under practical operation conditions," in *IEDM Tech Dig.* , pp. 817, 2007.

[118] C. R. Parthasarathy, M. Denais, V. Huard, G. Ribes, E. Vincent, and A. Bravaix, "New insights into recovery characteristics post NBTI stress," in *Int. Reliab. Phys. Symp.*, pp. 471-477, 2006.

[119] Z. Ji, J. F. Zhang, and W. Zhang, "A New Mobility Extraction Technique Based on Simultaneous Ultrafast Id-Vg and Ccg-Vg Measurements in MOSFETs," *IEEE Trans. Electron Devices,* vol. 59, no. 7, pp. 1906-1914, Jul, 2012.

[120] H. Reisinger, O. Blank, W. Heinrigs, W. Gustin, and C. Schlunder, "Analysis of NBTI degradation- and recovery-behavior based on ultra fast Vt-measurements," in *Int. Reliab. Phys. Symp.*, pp. 448-453, 2006.

[121] W. D. Zhang, B. Govoreanu, X. F. Zheng, D. R. Aguado, M. Rosmeulen, P. Blomme, J. F. Zhang, and J. Van Houdt, "Two-pulse C-V: A new method for characterizing electron traps in the bulk of SiO2/high-k dielectric stacks," *IEEE Electron Device Lett.,* vol. 29, no. 9, pp. 1043-1046, Sep, 2008.

[122] X. F. Zheng, W. D. Zhang, B. Govoreanu, J. F. Zhang, J. van Houdt, and Ieee, *A discharge-based multi-pulse technique (DMP) for probing electron trap energy distribution in high-k materials for Flash memory application*, 2009.

[123] I. S. AlKofahi, J. F. Zhang, and G. Groeseneken, "Generation and annealing of hot hole induced interface states," *Microelectron. Eng.,* vol. 36, no. 1-4, pp. 227-230, Jun, 1997.

[124] T. Aichinger, M. Nelhiebel, S. Decker, and T. Grasser, "Energetic distribution of oxide traps created under negative bias temperature stress and their relation to hydrogen," *Appl. Phys. Lett.,* vol. 96, no. 13, 2010.

[125] D. J. Dimaria, Z. A. Weinberg, and J. M. Aitken, “Location of positive charge in SiO2-films on Si generated by VUV photons, X-rays, and high-field stressing,” *J. Appl. Phys.,* vol. 48, no. 3, pp. 898-906, 1977.

[57] [29] J. F. Zhang, H. K. Sii, G. Groeseneken, and R. Degraeve, “Hole trapping and trap generation in the gate silicon dioxide,” *IEEE Trans. Electron Devices,* vol. 48, no. 6, pp. 1127-1135, Jun, 2001.

[126] J. F. Zhang, M. H. Chang, Z. Ji, L. Lin, I. Ferain, G. Groeseneken, L. Pantisano, S. De Gendt, and M. M. Heyns, “Dominant Layer for Stress-Induced Positive Charges in Hf-Based Gate Stacks,” *IEEE Electron. Dev. Lett.,* vol. 29, no. 12, pp. 1360-1363, Dec, 2008.

[127] R. Degraeve, B. Govoreanu, B. Kaczer, M. B. Zahid, J. V. Houdt, M. Jurczak, and G. Groeseneken, “Trap Spectroscopy by Charge Injection and Sensing (TSCIS): A quantitative electrical technique for studying defects in dielectric stacks,” *in IEDM Tech. Dig.*, pp. 775-778, 2008.

[128] J. F. Zhang, I. S. Al-kofahi, and G. Groeseneken, “Behavior of hot hole stressed SiO2/Si interface at elevated temperature,” *J. Appl. Phys.,* vol. 83, no. 2, pp. 843-850, Jan 15, 1998.

[129] D. J. DiMaria, “The properties of electron and hole traps in thermal silicon dioxide layers grown on silicon,” *in The Physics of SiO2 and its Interfaces, S. T. Pantelides, Ed. New York: Pergamon*, pp. 160-178, 1978.

[130] W. S. Cleveland, “ Robust locally weighted regression and smoothing scatterplots,” *Journal of the American Statistical Association*, pp. 829-836, 1979.

[131] X. F. Zheng, W. D. Zhang, B. Govoreanu, D. R. Aguado, J. F. Zhang, and J. Van Houdt, “Energy and Spatial Distributions of Electron Traps Throughout SiO2/Al2O3 Stacks as the IPD in Flash Memory Application,” *IEEE Trans. Electron Devices,* vol. 57, no. 1, pp. 288, 2010.

[132] J. R. Hauser, and K. Ahmed, "Characterization of ultra-thin oxides using electrical CV and IV measurements," *Characterization and metrology for ULSI Technology*, pp. 235-239, 1998.

[133] M. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, J. Babcock, R. Lindley, P. Clifton, M. Redford, and A. Shibkov, "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field-effect transistors," *Applied Physics Letters,* vol. 83, no. 8, pp. 1647, 2003.

[134] J. F. Zhang, S. Taylor, and W. Eccleston, "ELECTRON TRAP GENERATION IN THERMALLY GROWN SIO2 UNDER FOWLER-NORDHEIM STRESS, "*J. Appl. Phys.,* vol. 71, no. 2, pp. 725-734, Jan 15, 1992.

[135] M. H. Chang, J. F. Zhang, and W. D. Zhang, "Assessment of capture cross sections and effective density of electron traps generated in silicon dioxides," *IEEE Trans. Electron. Devices,* vol. 53, no. 6, pp. 1347-1354, Jun, 2006.

---

[136] Kang JF, Yu HY, Ren C, Li M-F, Chan DSH, Liu XY, et al. Ultrathin HfO2 (EOT @0.75 nm) gate stack with TaN/HfN electrodes fabricated using a high temperature process. Electrochem Solid-State Lett 2005;8(11):G311–3.

[137] Yang Chia-Han, Kuo Yue, Lin Chen-Han, Wan Rui, Kuo Way. Relaxation behavior and breakdown mechanisms of nanocrystals embedded Zr-doped HfO2 high-k thin films for nonvolatile memories, MRS procs. Mater Sci Technol Nonvolatile Mem 2008;1071-F02-09.

[138] Wilk GD, Wallace RM, Anthony JM. High-j gate dielectrics: current status and materials properties considerations. J Appl Phys 2001;89(10):5243.

[139] Jo, M., Kim, S., Lee, J., Jung, S., Park, J. B., Jung, H. S. & Hwang, H. (2010). Characterization of fast charge trapping in bias temperature instability in metal-oxide-semiconductor field effect transistor with high dielectric constant. *Applied Physics Letters*, *96*(14), 142110-142110.

[140] Liu, Z., & Ma, T. P. (2010). Determination of energy and spatial distributions of traps in ultrathin dielectrics by use of inelastic electron tunneling spectroscopy. *Applied Physics Letters*, *97*(17), 172102-172102.

[141] Schmid, A., Bollmann, J., & Oestreich, C. (2011, September). Determination of the trap energy distribution in oxynitride charge trapping layers by temperature dependent retention measurement. In *Semiconductor Conference Dresden (SCD), 2011* (pp. 1-3). IEEE.

[142] Cho, H. J., Son, Y., Lee, S., Lee, J. H., Park, B. G., & Shin, H. (2011). Study on the Oxide Trap Distribution in a Thin Gate Oxide from Random Telegraph Noise in the Drain Current and the Gate Leakage Current. *Journal of the Korean Physical Society*, *58*(5), 1518-1521.

[143] Gildenblat, G., Zhu, Z., & McAndrew, C. C. (2009). Surface potential equation for bulk MOSFET. *Solid-State Electronics*, *53*(1), 11-13.

[144] X.Ji, Y.Liao, F.Yan, Y.Shi, "The energy distribution of NBTI-induced hole traps in the Si band gap in PNO in pMOSFETs"ICSiCT 2012.

[145] Heh, D., Young, C. D., Brown, G. A., Hung, P. Y., Diebold, A., Vogel, E. M., & Bersuker, G. (2007). Spatial Distributions of Trapping Centers in HfO2/SiO2 Gate Stack. *Electron Devices, IEEE Transactions on*, *54*(6), 1338-1345.

[146] T.Aichinger, M.Nelhiebel, S.Decker, T.Grasser, "Observing two gate stage recovery of gate oxide damage created under NBTS", JAP 2010.

[147] Liao, Y., Ji, X., Wu, F., Zhu, X., Yan, F., Shi, Y., & Guo, Q. (2010, November). Investigation on the role of hole traps under NBTI stress in PMOS device with plasma-nitrided dielectric oxide. In *Solid-State and Integrated Circuit Technology (ICSICT), 2010 10th IEEE International Conference on* (pp. 1695-1697). IEEE.

[148] Ang, D. S., Lai, S. C. S., Du, G. A., Teo, Z. Q., Ho, T. J. J., & Hu, Y. Z. (2009). Effect of hole-trap distribution on the power-law time exponent of NBTI. *Electron Device Letters, IEEE*, *30*(7), 751-753.

[149] Seo, Y. J., Kim, K. C., Kim, H. D., Kim, T. G., & An, H. M. (2008). Study of hole traps in the oxide-nitride-oxide structure of the SONOS flash memory. *Journal of Korean Physical Society*, *53*, 3302.

[150] Kesapragada, S.; Rongjun Wang; Liu, D.; Guojun Liu; Zhigang Xie; Zhenbin Ge; Haichun Yang; Yu Lei; Xinliang Lu; Xianmin Tang; Jianxin Lei; Allen, M.; Gandikota, S.; Moraes, K.; Hung, S.; Yoshida, N.; Chorng-Ping Chang, "High-k/metal gate stacks in gate first and replacement gate schemes," *Advanced Semiconductor Manufacturing Conference (ASMC), 2010 IEEE/SEMI* , vol., no., pp.256,259, 11-13 July 2010

[151] Tseng, H. H., Kirsch, P., Park, C. S., Bersuker, G., Majhi, P., Hussain, M., & Jammy, R. (2009). The progress and challenges of threshold voltage control of high k/metal-gated devices for advanced technologies. *Microelectronic Engineering*, *86*(7), 1722-1727.

[152] Cartier, E., Kerber, A., Krishnan, S., Linder, B., Ando, T., Frank, M. M. & Narayanan, V. (2011). (Invited) Voltage Ramp Stress Based Stress-And-Sense Test Method For Reliability Characterization of Hf-Base High-k/Metal Gate Stacks For CMOS Technologies. *ECS Transactions*, *41*(3), 337-348.

[153] Auth, C., Cappellani, A., Chun, J. S., Dalis, A., Davis, A., Ghani, T. & Wiegand, C. (2008, June). 45nm high-k+ metal gate strain-enhanced transistors. In *VLSI Technology, 2008 Symposium on* (pp. 128-129). IEEE.

[154] Henson, K., Bu, H., Na, M. H., Liang, Y., Kwon, U., Krishnan, S. & Khare, M. (2008, December). Gate length scaling and high drive currents enabled for high performance SOI technology using high-κ/metal gate. In *Electron Devices Meeting, 2008. IEDM 2008. IEEE International* (pp. 1-4). IEEE.

[155] Ando, T., Copel, M., Bruley, J., Frank, M. M., Watanabe, H., & Narayanan, V. (2010). Physical origins of mobility degradation in extremely scaled SiO/HfO gate stacks with La and Al induced dipoles. *Applied Physics Letters*, *96*, 132904

[156] Lee, T. H., Chen, S. M., Hsu, C. W., Fang, Y. K., Juang, F. R., Hsu, C. H. & Chen, Y. W. (2011, June). Capping layer induced degradations in nano MOSFETs with scaled IL. In *Nanoelectronics Conference (INEC), 2011 IEEE 4th International* (pp. 1-2). IEEE.

[157] Kang, C. Y., Kirsch, P. D., Lee, B. H., Tseng, H. H., & Jammy, R. (2009). Reliability of La-doped Hf-based dielectrics nMOSFETs. *Device and Materials Reliability, IEEE Transactions on*, *9*(2), 171-179.

[158] Kuhn, K. J. (2012). Considerations for ultimate CMOS scaling. *Electron Devices, IEEE Transactions on*, *59*(7), 1813-1828

[159] Chang, V. S., Ragnarsson, L. A., Yu, H. Y., Aoulaiche, M., Conard, T., Yin, K., & Biesemans, S. (2007). Effects of Al 2 O 3 Dielectric Cap and Nitridation on Device Performance, Scalability, and Reliability for Advanced High-κ/Metal Gate pMOSFET Applications. *Electron Devices, IEEE Transactions on*, *54*(10), 2738-2749.

[160] Aoulaiche, M., Kaczer, B., Cho, M., Houssa, M., Degraeve, R., Kauerauf, T., & Groeseneken, G. (2009, April). Positive and negative bias temperature instability in La2O3 and Al2O capped high-k MOSFETs. In *Reliability Physics Symposium, 2009 IEEE International* (pp. 1014-1018). IEEE.

[161] Lu, C. C., Chang-Liao, K. S., Tsao, C. H., & Wang, T. K. (2010). Comparison of positive and negative bias-temperature instability on MOSFETs with HfO2/LaO x and HfO2/AlOx dielectric stacks. *Solid-State Electronics*, *54*(11), 1474-1478.

[162] Sellner, S., Gerlach, A., Schreiber, F., Kelsch, M., Kasper, N., Dosch, H. & Ulbricht, G. (2006). Mechanisms for the enhancement of the thermal stability of organic thin films by aluminum oxide capping layers. *Journal of materials research*, *21*(02), 455-464.

[163] Stathis, J. H. (2002). Reliability limits for the gate insulator in CMOS technology. *IBM Journal of Research and Development*, *46*(2.3), 265-286.

[164] JEDEC-JEP122G, "Failure Mechanisms and Models for Semiconductor Devices", 2011.

[165] Groeseneken, G., Degraeve, R., Kaczer, B., & Roussel, P. (2005, April). Recent trends in reliability assessment of advanced CMOS technologies. In *Microelectronic Test Structures, 2005. ICMTS 2005. Proceedings of the 2005 International Conference on* (pp. 81-88). IEEE.

[166] Zhang, J. F., Ji, Z., Chang, M. H., Kaczer, B., & Groeseneken, G. (2007, December). Real Vth instability of pMOSFETs under practical operation conditions. In *Electron Devices Meeting, 2007. IEDM 2007. IEEE International* (pp. 817-820). IEEE.

[167] Hicks, J., Bergstrom, D., Hattendorf, M., Jopling, J., Maiz, J., Pae, S. & Wiedemer, J. (2008). 45nm transistor reliability. *Intel Technology Journal*, *12*(2), 131-144.

[168] Heinrigs, W., Reisinger, H., Gustin, W., & Schlunder, C. (2007, April). Consideration of recovery effects during NBTI measurements for accurate lifetime predictions of state-of-the-art pMOSFETs. In *Reliability physics symposium, 2007. proceedings. 45th annual. ieee international* (pp. 288-292). IEEE.

[169] Chen, C. L., Lin, Y. M., Wang, C. J., & Wu, K. (2005, April). A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2 nm ultra thin oxide. In *Reliability Physics Symposium, 2005. Proceedings. 43rd Annual. 2005 IEEE International* (pp. 704-705). IEEE.

[170] Aono, H., Murakami, E., Okuyama, K., Nishida, A., Minami, M., Ooji, Y., & Kubota, K. (2005). Modeling of NBTI saturation effect and its impact on electric field dependence of the lifetime. *Microelectronics Reliability*, *45*(7), 1109-1114.

[171] Chen, S. C., Chien, C. H., & Lou, J. C. (2007). Anomalous negative bias temperature instability behavior in p-channel metal-oxide-semiconductor field-effect transistors with HfSiON/ SiO2 gate stack. *Applied physics letters*, *90*(23), 233505-233505.

[172] Mahapatra, S., & Alam, M. A. (2002). A predictive reliability model for PMOS bias temperature degradation. In *Electron Devices Meeting, 2002. IEDM'02. International* (pp. 505-508). IEEE.

[173] Chen, G., Chuah, K. Y., Li, M. F., Chan, D. S., Ang, C. H., Zheng, J. Z. & Kwong, D. L. (2003, April). Dynamic NBTI of PMOS transistors and its impact on device lifetime. In *Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International* (pp. 196-202). IEEE.

[174] Danković, D., Manić, I., Djorić-Veljković, S., Davidović, V., Golubović, S., & Stojadinović, N. (2006). NBT stress-induced degradation and lifetime estimation in p-channel power VDMOSFETs. *Microelectronics Reliability*, *46*(9), 1828-1833.

[175] Cho, M., Aoulaiche, M., Degraeve, R., Kaczer, B., Franco, J., Kauerauf, T & Groeseneken, G. (2010, May). Positive and negative bias temperature instability on sub-nanometer EOT high-K MOSFETs. In *Reliability Physics Symposium (IRPS), 2010 IEEE International* (pp. 1095-1098). IEEE.

[176] Reisinger, H., Vollertsen, R. P., Wagner, P. J., Huttner, T., Martin, A., Aresu, S. & Schlunder, C. (2009). A study of NBTI and short-term threshold hysteresis of thin nitrided and thick non-nitrided oxides. *Device and Materials Reliability, IEEE Transactions on*, *9*(2), 106-114.

[177] Schlunder, C., Heinrigs, W., Gustin, W., & Reisinger, H. (2006, October). On the impact of the NBTI recovery phenomenon on lifetime prediction of modern p-MOSFETs. In *Integrated Reliability Workshop Final Report, 2006 IEEE International* (pp. 1-4). IEEE.

[178] Franco, J., Kaczer, B., Eneman, G., Mitard, J., Stesmans, A., Afanas' ev, V.& Groeseneken, G. (2010, December). 6Å EOT Si 0.45 Ge 0.55 pMOSFET with optimized reliability (V DD= 1V): Meeting the NBTI lifetime target at ultra-thin EOT. In *Electron Devices Meeting (IEDM), 2010 IEEE International* (pp. 4-1). IEEE.

[179] Yamada, R. I., & King, T. J. (2003). Variable stress-induced leakage current and analysis of anomalous charge loss for flash memory application. In *Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International* (pp. 491-496). IEEE.

[180] Wu, E. Y., & Suñé, J. (2005). Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, *45*(12), 1809-1834.

[181] Neugroschel, A., Bersuker, G., Choi, R., Cochrane, C., Lenahan, P., Heh, D. & Jammy, R. (2006, December). An accurate lifetime analysis methodology incorporating governing NBTI mechanisms in high-k/SiO2 gate stacks. In *Electron Devices Meeting, 2006. IEDM'06. International* (pp. 1-4). IEEE

APPENDIX

i. List of Publications

LIST OF PUBLICATIONS

Journals

[1] **Hatta, S. W. M**., Ji, Z., Zhang, J. F., Duan, M., Zhang, W. D., Soin, N., ... & Groeseneken, G. (2013). Energy Distribution of Positive Charges in Gate Dielectric: Probing Technique and Impacts of Different Defects. *Electron Devices, IEEE Transactions on*, *60*(5), 1745-1753. (*ISI-Cited* , Q1)

[2] **Hatta, S. F**., Soin, N., Hadi, D. A., & Zhang, J. F. (2010). NBTI degradation effect on advanced-process 45nm high-k PMOSFETs with geometric and process variations. *Microelectronics Reliability*, *50*(9), 1283-1289. (*ISI-Cited* , Q2)

[3] Hussin, H., Soin, N., Karim, N. M., & **Wan Muhamad Hatta, S. F**. (2012). On the effects of NBTI degradation in p-MOSFET devices. *Physica B: Condensed Matter*, *407*(15), 3031-3033. (*ISI-Cited* , Q3)

[4] **Hatta, W. M., Fatmadiana**, S., Abdul Hadi, D., & Soin, N. (2011). Laser Anneal-Induced Effects on the NBTI Degradation of Advanced-Process 45nm High-K PMOS. *Advanced Materials Research*, *189*, 1862-1866. (*ISI-Cited*)

[5] **S. F. W. M. Hatta** , Z. Ji, X. Zhang, J. F. Zhang, W. Zhang, M. Duan, N. Soin, B. Kaczer, S. De Gendt, and G. Groeseneken, "A critical assessment on the predictive capability of different NBTI models," To be published to IEEE Trans. Electron. Dev.

Conferences

[1] Z.Ji, **S.F.W.M. Hatta**, J.F.Zhang, J.G.Ma, W.Zhang, N.Soin, B.Kaczer, S.De Gendt, G.Groeseneken, " Negative Bias Temperature Instability Lifetime Prediction: Problems and Solutions" – accepted for publication in the 2013 International Electron Devices Meeting (IEDM). (*ISI-Cited* , Q1)

[2] **Hatta, S. W. M**., Soin, N., & Zhang, J. F. (2010, June). The effect of gate oxide thickness and drain bias on NBTI degradation in 45nm PMOS. In *Semiconductor Electronics (ICSE), 2010 IEEE International Conference on* (pp. 210-213). IEEE.(ISI-cited)

[3] Hadi, D. A., **Wan Muhamad Hatta, S. F**., & Soin, N. (2010, June). Effect of oxide thickness on 32nm Pmosfet reliability. In *Semiconductor Electronics (ICSE), 2010 IEEE International Conference on* (pp. 244-247). IEEE. .(ISI-cited)

[4] **Hatta, S. W. M**., Soin, N., & Zhang, J. F. (2010, June). The effect of process variation on NBTI degradation in 90nm PMOS. In *Semiconductor Electronics (ICSE), 2010 IEEE International Conference on* (pp. 206-209). IEEE. .(ISI-cited)

[5] Hadi, D. A., Soin, N., & **Hatta, S. W. M**. (2011, May). Reliability Study of the 90 nm CMOS Inverter. In *AIP Conference Proceedings* (Vol. 1341, p. 181).

Invited paper at international conferences

[1] J. F. Zhang, Z. Ji, **S. F. W. M. Hatta**, J. G. Ma, W. Zhang, and N. Soin "Impact of different types of defects on NBTI lifetime prediction"China Semiconductor Technology International Conference (CSTIC), March, 2014.