Characterization of dynamic NBTI by ultra-fast charge pumping and ultra-fast switching method

Teo, Andy Zhiqiang

2015

Teo, A. Z. (2015). Characterization of dynamic NBTI by ultra-fast charge pumping and ultra-fast switching method. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/65268>

<https://doi.org/10.32657/10356/65268>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**CHARACTERIZATION OF DYNAMIC
NBTI BY ULTRA-FAST CHARGE
PUMPING AND ULTRA-FAST
SWITCHING METHOD**

TEO ZHIQIANG, ANDY

**SCHOOL OF ELECTRICAL AND ELECTRONIC
ENGINEERING**

2015

CHARACTERIZATION OF DYNAMIC NBTI BY ULTRA-FAST CHARGE PUMPING AND ULTRA-FAST SWITCHING METHOD

TEO ZHIQIANG, ANDY

**SCHOOL OF ELECTRICAL AND ELECTRONIC
ENGINEERING**

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirements for the degree of
Doctor of Philosophy

2015

ABSTRACT

Till very recently, the most well-known model used to explain the NBTI phenomenon is the reaction-diffusion (R-D) model, which describes the evolution of Si/SiO₂ interface states (N_{it}) contributing to NBTI based on a hydrogen transport mechanism. However, there has been much on-going discussion regarding the commonly observed instantaneous threshold voltage (V_{th}) shift during stress and relaxation phase. Several believe that this instantaneous V_{th} shift is solely explained by the generation and passivation of N_{it} at the Si/SiO₂ interface based on the framework of the R-D model. On the other hand, some thinks that NBTI has two mechanisms, involving both hole trapping (N_{ot}) and N_{it} generation, which are used to explain the fast and slow V_{th} shift. In this work, an objective and systematic analysis was carried out to challenge the validity of the R-D model versus the two-mechanism NBTI model through dynamic NBTI stress/recovery testing. Further understanding of the proposed two-mechanism NBTI framework as well as its implication on NBTI will be discussed. Two fast measurement methodologies with the capability to characterize ΔV_{th} and ΔN_{it} within 100 ns measurement time were developed, namely the ultra-fast switching method (UFS) and its variant, ultra-fast switching charge pumping (UFS-CP). The hardware setup and working principle are presented as well.

Dynamic NBTI experimental results measured by UFS have revealed a specific fast responsive cyclical $|\Delta V_{th}|$ component which can be generated during stress and fully recovered upon relaxation, regardless of the stress and recovery intervals. This behavior, observed over many stress/recovery cycles, is inconsistent with the R-D

model, which stipulates that interface state relaxation should decrease progressively with stress/recovery cycling. Further insights through an Arrhenius experiment revealed two distinct degradation mechanisms, evident by a small and large activation energy (~ 45 meV and ~ 110 meV). The relatively small energy activation energy of 45 meV suggests a cyclical $|\Delta V_{th}|$ component arising from a hole trapping/detrapping mechanism (via tunneling). The stronger thermal activation of 0.11 eV was perceived as a more energy-intensive process involving the breaking of Si-H bonds for creation of N_{it} . Detailed testing revealed different time dependence for hole trapping and N_{it} generation, implying that these two mechanisms are not coupled.

The hole trapping/detrapping phenomenon, perceived by many to result in the rapid cyclical $|\Delta V_{th}|$ evolution in dynamic NBTI, may be ascribed to the well known oxygen vacancy defect (E' centers) and its variants (E'_δ centers) in the SiO_2 gate dielectric. As evident from the first principal calculation, the shallow energy state of E'_δ center is calculated to be 0.5 eV to 1 eV above the SiO_2 valance band, which favors the condition for a rapid charge exchange and short tunneling constant in a ultra-thin gate oxide. However, one must also consider the wide spread of the defect energy state and tunneling time constant in an amorphous nature gate oxide. Further evidence from a positive relaxation gate voltage revealed another important class of hole traps, classified as deep level hole trap (DLHT). A DLHT is ascribed to a defect located very close to the interface and as a consequence, their energy states are “pinned” by the Si/ SiO_2 band discontinuity. Due to the higher (deeper) energy state, a charged DLHT requires a much longer emission time (which could take up to several days) for it to

be discharged. However, the recovery of DHLT can be accelerated by a positive gate bias.

Through a combination of dynamic NBTI testing at high and low voltages, experimental results show that the number of hole traps that can be activated under a given stress condition is fixed, although this number can vary with the exact stress voltage applied. Even when the applied stress resulted in an increase of stress induced leakage current (SILC) or bulk oxide trap density, the number of hole traps activated under dynamic NBTI testing did not increase but exhibited a decrease instead. Analysis of $|\Delta V_{th}|$ measured immediately upon the termination of stress and after a substantial recovery, shows interestingly a faster increase of the latter. As the $|\Delta V_{th}|$ measured at the end of stress probes the total positive oxide trapped charge while that measured after relaxation only probes the unrecovered portion of the positive oxide trapped charge, the results show that the decrease in the number of hole traps activated during the dynamic NBTI test is a consequence of a portion of the hole traps being transformed into a more permanent form. This transformation is shown to always correspond to the increase of SILC, i.e. when the transformation is absent, so is the increase in SILC absent, implying that the transformation is responsible for the generation of more permanent bulk traps that are responsible for the observed SILC. The transformation is consistent with the irreversible structural change that may occur at an E' center upon the capture of a hole.

Further presumption that the R-D model can still be used to explain the N_{it} generation in the two-mechanism NBTI model was also ruled out. UFS-CP measurement of

dynamic NBTI shows that the amount of ΔI_{cp} induced in a given stress interval is completely removed in the following recovery interval, which disagrees with the notion of a hydrogen transport model that depicts the amount of recovery to decrease progressively with stress/recovery cycling. Instead, it was found that stress induced ΔN_{it} comprises of two distinct characteristics coupled together; (i) one is a consistent fast recoverable N_{it} component which is spontaneously recovered upon termination of stress (ii) while the other is the permanent N_{it} component. The fast recoverable N_{it} component, identified as a parasitic signal which originates from the N_{ot} was inevitably captured in the UFS-CP measurement. The other permanent component, identified as the generation of N_{it} , is shown to be an irreversible process, whereby the generation of N_{it} is solely dependent on stress period and the generated defects show no apparent recovery even after an extremely long recovery period (up to 50 days). A separation methodology is proposed to separate the parasitic oxide trap signal from as-measured signal in UFS-CP, yielding a power-law time exponent of ~ 0.27 for ΔN_{it} .

Based on the phenomenological observation that NBTI is determined by a cyclic hole trapping/detrapping mechanism and that interface-state generation is relatively permanent, a systematic separation approach is outlined to decouple both components to evaluate the impact on NBTI. In UFS methodology, the time dependence of hole trapping during stress is precisely determined through dynamic NBTI, and then subtracted from the overall degradation of the first cycle to yield the time dependence of N_{it} generation. N_{it} generation is shown to exhibit power-law time dependence with an initial exponent of ~ 0.5 , which subsequently decreases to a steady value of 0.25 after ~ 1000 s at the stress condition studied (oxide field ~ 10 MV/cm). This evolution

is shown to be consistent to that obtained via the charge-pumping method, confirming the underlying principle of the approach. Lastly, the impact of hole trapping at operating condition is examined for the first time. A device was pre-stressed for an extremely long time (up to 4.5 days), whereby further generation of N_{it} is insignificant. The impact of hole trapping is examined over a broad range of stress voltages and extrapolated to operating bias condition of 1.2V for the first time.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project supervisor, Dr. Ang Diing Shenp, for the continuous support and trust given to me during my Ph.D project. During this period, he has guided me well with his invaluable knowledge and experience. He is extremely knowledgeable in terms of electrical measurements and is also very capable in building up and adding on to his extensive knowledge through effective reading and delivering his idea clearly through professional writing. I am greatly impressed by his reading and writing abilities and am trying to learn from him as much as possible.

Next, I would like to express my sincere thanks to Dr. Lap Chan and Dr. Ng Chee Mang for their decisive support on my GlobalFoundries-NTU Graduate Research Scholarship application and continuous encouragement during my postgraduate studies as well as providing me with opportunities to increase my knowledge and analyze the current market industry, and the invaluable classes taught during my first year of Ph.D studies.

I also would like to extend my gratitude to my group members for their valuable contributions during our discussions and high degree of cooperation. They are Ms Wang Shuang (research on NBTI mechanism frequency dependence), Mr. Du Guoan (research on NBTI using fast measurement), Mr. Hu Youzhou (research on NBTI of high- k), Mr. Ho Jun Jie (research on NBTI channel length/ width dependence), Mr.

Lai Chung Sing (research on low frequency noise measurement) and Bao Ann Ann (research on mechanism of NBTI by ultrafast measurement).

Last but not least, I appreciate the help from, Mr. Lo Vui Lip, Mr. S.K. Cheong, Mr. Kenneth and Mr. Vincent Cher (Keithley Ins., providing expert advice on electrical measurement and equipment support)

SUMMARY OF ABBREVIATIONS

δV_{th}^S	Change in Threshold Voltage at the end of a particular stress phase
δV_{th}^R	Change in Threshold Voltage at the end of a particular recovery phase
δI_{cp}^R	Change in Charge Pumping Current the end of a particular recovery phase
Δg_m	Change in Transconductance
ΔI_{cp}	Change in Charge Pumping Current
ΔI_{cp}^{EoS}	Cumulative Change in Charge Pumping Current at the end of a particular relaxation cycle
ΔI_{cp}^{EoS}	Cumulative Change in Charge Pumping Current at the end of a particular stress cycle
ΔI_{cp}^{BoR}	Cumulative Change in Charge Pumping Current at the beginning of a particular relaxation cycle
ΔI_g	Change in Gate Leakage Current
ΔI_h	Change in Hole Recombination Current
ΔI_{it}	Change in Interface Trap Recombination Current
ΔI_{ot}	Change in Oxide Trap Recombination Current
ΔN_h	Change in Hole Trap Density
ΔN_{it}	Change in Interface Trap Density
ΔN_{ot}	Change in Oxide Trap Density
ΔV_{th}	Change in Threshold Voltage

ΔV_{th}^{EoS}	Cumulative Change in Threshold Voltage at the end of a particular stress cycle
ΔV_{th}^{EoR}	Cumulative Change in Threshold Voltage at the end of a particular relaxation cycle
ΔV_{th}^h	Change in Threshold Voltage by hole trap generation
ΔV_{th}^{it}	Change in Threshold Voltage caused by interface trap generation
ΔV_{th}^{ot}	Change in Threshold Voltage caused by oxide trap generation
AC	Alternative Current
CP	Charge Pumping
CVS	Constant Voltage Stress
CVS–DNBTI	Constant Voltage Stress–Dynamic Negative Bias Temperature
DC	Direct Current
DLHT	Deep Level Hole Trap
DLTS	Deep Level Transient Spectroscopy
DNBTI	Dynamic Negative Bias Temperature Instability
DUT	Device Under Test
ESR	Electron Spin Resonance
E' center	Oxygen Vacancies
E_{ox}	Gate Oxide Electric Field
EoS	End-of-Stress
EoR	End-of-Relaxation
FS	Fast Switching
f	Frequency

fA	Femto-Ampere
H_2	Molecular Hydrogen
H	Atomic Hydrogen
HfO_2	Hafnium Dioxide
pA	Pico-Ampere
I_{cp}	Charge Pumping Current
$I_{cp}^{prestress}$	Charge Pumping Current at time zero (before stress)
I_{cp}^{stress}	Charge Pumping Current at NBTI stress
I_{cp}^{relax}	Charge Pumping Current at NBTI relaxation
I_d	Drain current
I_d^{AC}	Drain current (Alternative Current component)
$I_{d,lin}$	Liner Drain current
I_g	Gate Leakage Current
NBTI	Negative Bias Temperature Instability
N_h	Hole Trapping Density
N_{it}	Interface Trap Density
N_{ot}	Oxide Trapping Density
N_0	Initial Density of unbroken Si-H bond
OTF	On- the- fly
OFIT	On- the- fly Interface Trap
p– MOSFET	p-Channel Metal-Oxide Silicon Field-Effect Transistor
Q_h	Hole Charge
Q_{it}	Interface Trap Charge
Q_{ot}	Oxide Trap Charge

R– D	Reaction-Diffusion
RBT	Remote Bias Tee
R	Recoverable Component of NBTI
Si	Silicon
SiO ₂	Silicon Dioxide
SiON	Silicon OxyNitride

TABLE OF CONTENTS

ABSTRACT.....	I
ACKNOWLEDGEMENT	VI
TABLE OF CONTENTS.....	X
TABLE OF FIGURES.....	X
1 INTRODUCTION	1
2 LITERATURE REVIEW.....	6
2.1 INTERFACE TRAPS AND OXIDE CHARGES IMPACT ON NBTI.....	10
2.1.1 <i>Interface Traps</i>	12
2.1.2 <i>Oxide Traps</i>	14
2.2 THE REACTION-DIFFUSION (R-D) MODEL	15
2.2.1 <i>Framework of R-D model</i>	16
2.2.2 <i>Formulation of R-D Model</i>	19
2.3 CHALLENGES FACED IN NBTI CHARACTERIZATION.....	23
2.3.1 <i>Recovery Effect during NBTI Measurement</i>	23
2.3.2 <i>Implications by conventional charge pumping</i>	24
3 ULTRA-FAST MEASUREMENT METHODOLOGY.....	27
3.1 BACKGROUND AND MOTIVATION	27
3.2 KEITHLEY PULSE-IV SETUP.....	30
3.2.1 <i>Remote Bias Tee (RBT)</i>	30
3.2.2 <i>3-Port Power Divider</i>	31
3.2.3 <i>Keithley Pulse-IV Hardware Setup</i>	31
3.3 ULTRA-FAST SWITCHING (UFS) MEASUREMENT METHODOLOGY	35
3.3.1 <i>Experimental Setup of UFS</i>	35
3.3.2 <i>Principal of UFS Measurement</i>	36
3.3.3 <i>Validation of UFS Measurement</i>	38
3.4 ULTRA-FAST SWITCHING-CHARGE PUMPING (UFS-CP) MEASUREMENT METHODOLOGY	41
3.4.1 <i>Experimental Setup of UFS-CP</i>	42
3.4.2 <i>Principal of UFS-CP Measurement</i>	43
3.4.3 <i>Validation of UFS-CP Measurement</i>	46
4 REASSESSING THE MECHANISM OF NBTI	50
4.1 CAN NBTI BE EXPLAINED BY REACTION-DIFFUSION MODEL.....	50
4.1.1 <i>Simulation of Dynamic NBTI using R-D Model</i>	51

4.1.2	<i>Comparison of UFS Dynamic NBTI Experimental Results with R-D Model</i>	56
4.2	REASSESSING NBTI MECHANISM : TWO MECHANISM FRAMEWORK	65
4.2.1	<i>Origin of Hole Trapping/Detrapping in NBTI</i>	66
4.2.2	<i>Observation of Hole Trapping/Detrapping in NBTI</i>	71
4.2.3	<i>Shallow-Level (Fast) Versus Deep-Level (Slow) Hole Traps</i>	73
4.2.4	<i>Correlation Between Oxide Trap Generation and Hole Trapping/Detrapping Mechanism</i>	83
4.2.5	<i>Observation of Interface Traps Mechanism in NBTI</i>	93
4.2.6	<i>Cyclic Behavior of ΔI_{cp}</i>	96
4.2.7	<i>Negligible Recovery of Interface Traps</i>	100
4.3	SUMMARY	107
5	SEPARATION OF OXIDE TRAP AND INTERFACE TRAP IN UFS AND UFS-CP MEASUREMENT METHODOLOGY	111
5.1	SEPARATION OF PARASITIC OXIDE TRAP SIGNAL IN UFS-CP	111
5.1.1	<i>Introduction</i>	111
5.1.2	<i>Experimental</i>	112
5.1.3	<i>Separation of Oxide Trap and Interface Traps in UFS-CP</i>	113
5.1.4	<i>Temperature-Dependent of Power-Law Exponent on Interface Trap</i>	120
5.1.5	<i>Summary</i>	123
5.2	SEPARATION OF HOLE TRAPPING AND INTERFACE-STATE GENERATION BY UFS	124
5.2.1	<i>Introduction</i>	124
5.2.2	<i>Experimental</i>	126
5.2.3	<i>Separation of Hole Trapping and Interface Trap</i>	126
5.2.4	<i>Summary</i>	133
5.3	IMPACT OF HOLE TRAPPING STUDIED ON A SEASONED DEVICE	134
5.3.1	<i>Introduction</i>	134
5.3.2	<i>Experimental</i>	135
5.3.3	<i>Preparation of a "Seasoned" Device</i>	136
5.3.4	<i>Impact of Hole Trapping/Detrapping (HTD)</i>	139
5.3.5	<i>Summary</i>	144
6	CONCLUSION AND RECOMMENDATION	145
6.1	CONCLUSION	145
6.2	RECOMMENDATION	150
	LIST OF PUBLICATIONS	152
	REFERENCES	156

TABLE OF FIGURES

Figure 2.1: Structural model (100) Si surface.....	14
Figure 2.2: Model for hole-assisted field-enhanced thermal dissociation of Si–H bond resulting $k_f \propto E_c \exp(\gamma_T E_{ox}) \exp[-(E_F - aE_{ox}) / k_B T] \propto E_c \exp(\gamma E_{ox})$. Resultant field acceleration factor can be written as $\gamma = \gamma_T + a/k_B T$. [57].....	18
Figure 2.3: (top) Schematic description of the reaction-diffusion model used to interpret interface-trap generation during NBTI stress. Broken Si–H bonds at the Si–SiO ₂ interface create Si ⁺ (interface traps) and H. Initial interface-trap generation rate depends on Si–H bond dissociation (reaction), while the later rate depends on H removal (diffusion). Some H dimerized into H ₂ and since H is a slow diffuser, the H may diffuse (middle) or drift (bottom) away from the interface depending on the charge state of the diffusing species.....	18
Figure 2.4: (a) NBTI degradation over stress time for various delays between stress interruption and measurement (b) power law exponent versus delay time. [53].....	23
Figure 2.5: (a) The time evolution of ΔV_{th} measured by DC Id–V _g method and ΔN_{it} measured by CP method. Two characterization approaches: In the first, V_{th} was measured (curve 1) followed by N_{it} measurement (curve 3). In the second, N_{it} was measured (curve 2) followed by V_{th} measurement (curve 4). (b) Zoom-in times t_1 and t_2 [15].....	25
Figure 2.6: Evolution of the normalized threshold voltage shift $\Delta V_{th}/\Delta V_{th}^0$ of three identical p-MOSFETs during negative-bias temperature stressing and post-stress recovery. For all the three cases, only dc I_d – V_g measurement was made during the stress phase. In the recovery phase, different gate relaxation voltages (V_g^{relax}) were applied. In one case, a CP current measurement step was introduced after the dc I_d – V_g measurement. [62].....	26
Figure 3.1: Simplified schematic of Remote Bias Tee (RBT).....	30
Figure 3.2: Schematic of 3-Port Power Divider.....	31
Figure 3.3: Schematic of Keithley Pulse-IV measurement technique hardware setup.....	32
Figure 3.4: Waveform captured at selected points in the Pulse IV setup	34
Figure 3.5: Schematic diagram of the UFS measurement set-up.....	35
Figure 3.6: UFS test waveform consisting of 3 different phases: (i) pre-stress, (ii) stress and (iii) relax.....	36

Figure 3.7: Waveforms of gate voltage pulse, V_g^{AC} and drain current, I_d^{AC} . Gate voltage pulse has amplitude of $2.6V$, pulse width at $100ns$, rising/falling time at $10ns$. Data in between 60% and 80% of the pulse width is averaged and gives value of V_g^{AC} and I_d^{AC}	39
Figure 3.8: Drain current, $ I_d $ measured by both UFS and DC method by pulsing/sweeping gate voltage from 0 to $-1V$. Drain voltage is kept at $-0.1V$. (a) Linear regime, (b) Sub-threshold regime, (c) forward and reverse sweep of Pulse IV from $0V$ to $-1.2V$ in the liner regime (Fig. 1a of ref [65]), (d) forward and reverse sweep of Pulse IV from $0V$ to $-1.2V$ in the sub-threshold regime. (Fig. 1b of ref [65])	40
Figure 3.9: Comparison of $ \Delta V_{th} $ measured by UFS and conventional DC. (a) Plotted on linear scale to highlight the recovery effect during stress measurement but negligible during relaxation measurement (b) Plotted on log-log scale to highlight the power law exponent which is critical for long term extrapolation. Constant current method is used to extract $ \Delta V_{th} $ from I_d - V_g curve taken at various time intervals with drain voltage, V_d kept at $-0.1V$. Devices were stressed for time period of $50,000s$ and relax for $1000s$ at $V_g = -2.4V / 0V$ at $100^\circ C$	41
Figure 3.10: Schematic illustration of the ultra-fast switching-charge pumping (UFS-CP) measurement setup, adapted from ultra-fast switching method. The DUT is stressed or relaxed at a DC voltage, V_g^{DC} . CP measurement is performed at selected intervals by injecting a series of positive pulses of amplitude V_g^{AC} at the bias-tee, and the resultant I_{cp} is measured at the drain. The bias-tee allows very narrow pulse of duration $100ns$ or less to be used.	43
Figure 3.11: UFS-CP test waveform consisting of 3 different phases: (i) pre-stress, (ii) stress and (iii) relax. At respective phases (pre-stress/stress/relax), the V_g^{DC} is set to their respective value ($-1V \rightarrow V_g^{stress} \rightarrow -1V$). During UFS-CP measurement, V_g^{AC} is injected into the gate to bring the overall voltage from V_g^{DC} to accumulation ($+1V$) and back V_g^{DC} within $100ns$	44
Figure 3.12: Waveforms of gate voltage, V_g during UFS-CP measurement at different phases. During pre-stress and relax, the gate voltage is pulsed from $-1V$ to $+1V$. During stress, the gate voltage is pulsed from $-2.6V$ to $+1V$. The pulse width is defined at $100ns$ with rising/falling time at $10ns$	46
Figure 3.13: (a) Illustration showing constant base UFS-CP measurement (b) I_{cp} measured by UFS-CP at various V_{cp} for pre-stress/stress/relax phase.	

Measurements are done on a SiO ₂ gate oxide thickness of 28Å with $V_g^{\text{stress}} = -2.6V$. It can be observed for stress phase, there is a vertical shift as compared to pre-stress/relax phase. This is attributed to tunneling leakage current due to high V_g^{stress} and it is expected to increase for thinner gate oxide thickness which will affect the I_{cp} resolution.	48
Figure 3.14: Comparison of $ \Delta V_{\text{th}} $ measured by UFS-CP and conventional CP done on a SiO ₂ gate oxide thickness of 28Å. (a) Plotted on linear scale to highlight the recovery effect during stress measurements but negligible during relaxation measurements (b) Plotted on log-log scale to highlight the power law exponent which is critical for long term extrapolation. Devices were stressed for time period of 7,000s at $V_g^{\text{stress}} = -2.6V$ at 100°C.....	49
Figure 4.1: R-D model simulated results. Device undergoes dynamic NBTI stress and relaxation for 30 cycles at $V_g^{\text{stress}} = -2.6V$ and $V_g^{\text{relax}} = 0V$ for duration of 500s each. (a) $ \Delta N_{\text{it}} $ plotted with respect to stress time (b) Snapshots of hydrogen profile at the end of respective stress (left) and relaxation cycles (right). It can be clearly shown that the hydrogen will diffuse further away from the interface ($x > 0$) and the availability of hydrogen at interface ($x = 0$) will decrease with longer cycles.	51
Figure 4.2: Evolution in the hydrogen concentration profile and the corresponding interface state generation, ΔN_{it} under dynamic NBTI in the context of the reaction-diffusion model. (a) The stress period, t_s is fixed while the recovery period, t_r is varied. A longer t_r leads to greater recovery of ΔN_{it} due to more back diffusion of hydrogen. (b) t_s is varied while t_r is fixed. A longer t_s leads to lesser recovery of ΔN_{it} since hydrogen has been transported further away from the Si/SiO ₂ interface and thus more time is required for a given amount of hydrogen to diffuse back towards the interface. It should be pointed out that even during the recovery interval, some hydrogen continues to move away from the interface (see typical profiles at the end of the recovery period). This gives rise to self-limited recovery.....	54
Figure 4.3: (a) The amount of N_{it} recovered during relaxation is plotted with respect to dynamic cycles. It is observed that there is a ~50% decrease in the amount of N_{it} recovered for 1st and 30th cycle. (inset) The amount of N_{it} recovered can be calculated by taking the N_{it} difference between “End of Stress” (EoS) and “End of Relaxation”	

(EoR) for the respective dynamic cycles. (b) Profiles of hydrogen concentration as a function of distance x into the oxide at the end of selected stress cycles. As stress/recovery cycling progresses, hydrogen concentration at the Si/SiO ₂ interface decreases (arrow A) and the diffusion front advances further away from the interface (arrow B), indicating that more hydrogen has been transported away from the interface. These changes lead to the self-limiting recovery behavior depicted in Fig. 4.3(a), since the amount of hydrogen back-diffusion (and hence available for repassivation of interface states) in a given recovery interval would become gradually lesser as more hydrogen is transported further away from the interface with stress/recovery cycling.....	55
Figure 4.4: UFS experimental results showing $ \Delta V_{th} $ under dynamic NBTI stress and relaxation for 30 cycles at $V_g^{stress} = -2.6V$ and $V_{relax} = 0V$ for a similar time duration of 1000s at 100°C.....	57
Figure 4.5: Illustration showing definition of (i) $ \delta V_{th}^S $ ($ \delta V_{th}^R $) denotes the change in V_{th} shift at the end of a particular stress (recovery) phases; (ii) $ \Delta V_{th}^{EoS} $ ($ \Delta V_{th}^{EoR} $) denotes the cumulative $ V_{th} $ after a given number of stress (recovery) phases.....	58
Figure 4.6: (a) Evolution of $ \delta V_{th}^S $, $ \delta V_{th}^R $, $ \Delta V_{th}^{EoS} $ and $ \Delta V_{th}^{EoR} $ as a function of stress/recovery cycles for the device depicted in Fig. 4.3. The cumulative degradation ($ \Delta V_{th}^{EoS} $) is observed to increase progressively during the dynamic NBTI cycles. But the amount of $ V_{th} $ degradation and recovery per cycle, $ \delta V_{th}^S $ and $ \delta V_{th}^R $ is observed to be consistent throughout the dynamic NBTI cycles, (b) Normalized $ \delta N_{it}^R $ or $ \delta V_{th}^R $ recovered per cycle with respect to 1 st cycle based on R-D model simulation and experimental is compared.	59
Figure 4.7: Dynamic NBTI at various time ratio (t_{stress}/t_{relax}) performed over 20 cycles. The following parameters ($ \delta V_{th}^S $, $ \delta V_{th}^R $, $ \Delta V_{th}^{EoS} $ and $ \Delta V_{th}^{EoR} $) of the last cycle are plotted with respect to their time ratio. Given longer stress time (larger time ratio), the cumulative degradation ($ \Delta V_{th}^{EoS} $) is observed to suffer more degradation as compared to those with shorter stress time. But on the other hand, the amount of $ V_{th} $ degradation and recovery per cycle, $ \delta V_{th}^S $ and $ \delta V_{th}^R $ is still observed to be consistent despite the various time ratio.....	61
Figure 4.8: (a) Arrhenius plot of the cyclic component, i.e. $ \delta V_{th}^S $ (or $ \delta V_{th}^R $) and the permanent component $ \Delta V_{th}^{EoR} $; percentage of $ \delta V_{th}^S $ (or $ \delta V_{th}^R $) and $ \Delta V_{th}^{EoR} $ relative to the total $ V_{th} $ shift as a function of	

temperature; (c) dependence of $ \delta V_{th}^S $ or $ \delta V_{th}^R $ and $ \Delta V_{th}^{EoR} $ on the concentration of nitrogen in the gate oxide under a given stress condition ($V_g^{stress} = -2$ V; temperature = 100 °C).....	64
Figure 4.9: Schematic of neutral oxygen vacancy (left) and the positively charged E'_δ center (right), which is also called as shallow state in energy level point of view.	67
Figure 4.10: The electrical state energy distribution of shallow hole traps (E'_δ) when p-MOSFET is under negative gate voltage stress.	68
Figure 4.11: Schematic illustration of the $E_{\gamma 4}'$ center model for NBTI [79]. An important feature of this model is the significant asymmetrical structural relaxation which occurs upon hole capture (1 $\rightarrow$ 2), which results in one of the Si atom moving past the plane containing the oxygen atoms and getting back-bonded to a neighboring oxygen atom. This significant structural relaxation reduces the likelihood of the reformation of Si-Si bond (3 $\rightarrow$ 1). Transition from state 3 to 2 (during stressing) and from state 2 to 3 (during relaxation) are believed to explain the cyclic shifts observed in the threshold voltage $ \Delta V_{th} $	71
Figure 4.12: Typical evolution of threshold voltage shift, $ \Delta V_{th} $ during the stress (-2 V) and relaxation cycles of dynamic NBTI. (b) $ \Delta V_{th} $ at the end-of-stress (EoS), end-of-recovery (EoR), and the amount of $ \Delta V_{th} $ shift during each stress and relaxation cycle, $ \delta V_{th}^S $ and $ \delta V_{th}^R $, respectively as a function of the number of stress/relaxation (S/R) cycles, n . The constant $ \delta V_{th}^R $ in conjunction with $ \delta V_{th}^S \approx \delta V_{th}^R $, ($N > 10$) clearly depict a cyclic $ \Delta V_{th} $ shift behavior. (c) This is confirmed by the highly repetitive $ \Delta V_{th} $ shift over many stress and relaxation cycles (thin lines). Symbols denote the average.	73
Figure 4.13: Thin line: Evolution of the hole trapping component of threshold voltage shift, $ \Delta V_{th}^h $ during the stress and recovery cycles of dynamic NBTI (from the 10^{th} to 30^{th}). Symbols denote average characteristics. Only the average characteristics are shown for the case of 1.5 V gate recovery voltage. The gate stress voltage (-2.2 V) was the same for both cases.	76
Figure 4.14: Evolution of $ \Delta V_{th} $ of two devices undergoing NBTI stress (-2.6 V) for 8×10^3 s. One device is then relaxed at 0 V and monitored for 30 days. Another was relaxed at a positive voltage ($+1.5$ V) for 2000 s and relaxed back to 0 V for another 1000 s. Evidence showed that	

recovery of deep level hole traps is accelerated by a positive gate bias.....	77
Figure 4.15: (a) A comparison of the activation energy of the hole trapping component of threshold voltage shift, $ \Delta V_{th}^h $ and the relatively permanent component at the end of the recovery interval, $ \Delta V_{th}^{EoR} $ (due mainly to interface state generation). $ \Delta V_{th}^h $ was determined using a 1.5 V gate recovery voltage (cf. Fig. 4.13; filled triangle). (b) The corresponding power-law dependence of $ \Delta V_{th}^h $ on temperature T, giving an exponent of ~ 1.3 . Also shown is the result for $ \Delta V_{th}^h ^*$, an underestimated form of $ \Delta V_{th}^h $ assuming that the V_{th} estimated at 1 ms after the start of stress is the pre-stress V_{th} . Reducing the contribution of the fast hole trapping component clearly reveals a stronger dependence of slower hole trapping, which can be understood in terms of the multiphonon field-assisted tunneling model.	78
Figure 4.16: A comparison of the gate stress voltage dependence of fast and slow threshold voltage shift, $ \Delta V_{th} $. The fast $ \Delta V_{th} $ component corresponds to $ \Delta V_{th} $ taken 1 s after stress has commenced while the slow $ \Delta V_{th} $ component is the subsequent $ \Delta V_{th} $ shift for the remaining stress period (5×10^3 s).	79
Figure 4.17: Evolution of threshold voltage shift, $ \Delta V_{th} $ during NBTI. Starting from -1 V, the gate stress voltage was decreased, in step of -0.2 V, each time after the device had been stressed at a given voltage for 5×10^3 s.	80
Figure 4.18: A comparison of fast and slow threshold voltage shift, $ \Delta V_{th} $ resulting from the stress methodology shown in Fig. 4.17. The almost equal increase in fast $ \Delta V_{th} $ for a given increase in $ V_g^{stress} $ from one stress cycle to the next confirms the linear dependence of fast $ \Delta V_{th} $ on V_g^{stress} . The dashed line is an eye-guide only.	82
Figure 4.19: A possible explanation for the almost identical increment of $ \Delta V_{th} $ when the magnitude of the gate stress voltage was successively increased by a given step size (Fig. 4.18). Assuming a uniform energy distribution of hole trap precursors in the ultrathin gate oxide, a given negative change in the gate voltage, ΔV_g ($V_g \ll V_{th}$) will raise an equal amount of trap precursors above the Fermi energy, resulting in them becoming positively charge through spontaneous hole capture from the channel.	82

- Figure 4.20: For the range of negative V_g^{stress} commonly employed for NBTI testing, we found that $|\Delta V_{\text{th}}^{\text{R}}|$ at a given V_g^{stress} is unaffected by a preceding stress at a more negative V_g^{stress} . (a) A fresh DUT was first subjected to dynamic NBTI (DNBTI) stress at $V_g^{\text{stress}} = -1.9$ V. This was followed by a -2.3 V DNBTI stress. This sequence was repeated several times and evolution of $|\Delta V_{\text{th}}|$ captured. (b) Despite several more negative V_g^{stress} at -2.3 V, $|\delta V_{\text{th}}^{\text{R}}|$ at $V_g^{\text{stress}} = -1.9$ V remains the same as the very first time the DUT was stressed at -1.9 V.84
- Figure 4.21: (a) A fresh DUT was subjected to 4 stress (@ -1.15 V) and relaxation (@ 0 V) cycles. Each S/R cycle lasted 1×10^3 s. The test was then repeated on the same DUT with V_g^{stress} increased (negatively) by -0.25 V each time, until $V_g^{\text{stress}} = -2.65$ V. The V_g^{stress} was then decreased (positively) each time by the same voltage step and the test repeated. Evolution of $|\Delta V_{\text{th}}|$ for the entire test sequence is shown. (b) Despite having subjected the DUT to a higher stress oxide field, there is no change in the $|\delta V_{\text{th}}^{\text{R}}|$ at less negative gate stress voltages.86
- Figure 4.22: (a) $|\Delta V_{\text{th}}|$ recovery per cycle or $|\delta V_{\text{th}}^{\text{R}}|$ as a function of the number of DNBTs or CVS–DNBTs cycles. The inset shows the evolution of $|\Delta V_{\text{th}}|$ in a typical DNBTs cycle. (b) Corresponding gate leakage current after 20 DNBTs or 19 CVS–DNBTs cycles. CVS: $V_g^{\text{stress}} = 2.7$ V and time = 50 s.88
- Figure 4.23: (a) Fractional increase in the gate tunneling current $\Delta I_g/I_{\text{go}}$ (at $V_g = 1.5$ V) or stress-induced leakage current (SILC) is observed to “track” the decrease in $|\delta V_{\text{th}}^{\text{R}}|$ closely. (b) Excellent linear correlation between $\Delta I_g/I_{\text{go}}$ (extracted at various voltages) and change in $|\delta V_{\text{th}}^{\text{R}}|$92
- Figure 4.24: (a) With a larger V_g of CVS [cf. Fig. 4.22(a)], a more substantial decrease of $|\delta V_{\text{th}}^{\text{R}}|$ is observed. (b) Evolution of $|\Delta V_{\text{th}}|$ at the end of the stress cycle ($|\Delta V_{\text{th}}^{\text{EoS}}|$) and at the end of the relaxation cycle ($|\Delta V_{\text{th}}^{\text{EoR}}|$) as a function of the number of DNBTI or CVS–DNBTI cycles, for the same devices in (a).93
- Figure 4.25: (left) Schematic illustration of the ultra-fast switching-charge pumping (UFS-CP) measurement setup, adapted from ultra-fast switching method. (right) Schematic illustration of the gate waveforms during dynamic NBTI. For pre-stress and recovery, the DUT is biased at -1 V. The gate bias is switched to V_g^{stress} during stress. The top of the gate pulse for all cases is $+1$ V, except stated

otherwise. V_{inv} (-0.3 V) and V_{acc} ($+0.5$ V) refer to the threshold voltage and flat band voltage, respectively of charge pumping. The pulse period was fixed at 10 μs and the pulse width ranged from 100 - 900 ns; the corresponding duty cycle ranged from 1 - 9 %.	95
Figure 4.26: Increase in the charge pumping current (ΔI_{cp}), measured by an ultrafast 100 -ns pulse, as a function of stress time. Gate stress voltages are as indicated. Power-law dependence of ΔI_{cp} on stress time is evident and the exponent n ranges from 0.145 - 0.171 . This observation has been regarded as an experimental proof for a H_2 diffusion-limited N_{it} generation mechanism, which has a theoretical exponent n of 0.167 .	95
Figure 4.27: UFS-CP experimental results showing the evolution of ΔI_{cp} under dynamic NBTI stress and relaxation for 30 cycles at $V_g^{\text{stress}} = -2.6\text{V}$ and $V_{\text{relax}} = -1\text{V}$ for similar time duration of 1000s at 100°C . Illustration showing definition of (i) $\delta I_{\text{cp}}^{\text{R}}$ denotes the change in I_{cp} shift at the end of a particular recovery phases; (ii) $\Delta I_{\text{cp}}^{\text{EoS}}$ ($\Delta I_{\text{cp}}^{\text{EoR}}$) denotes the cumulative I_{cp} after a given number of stress (recovery) phases.	97
Figure 4.28: (a) Evolution of $\Delta I_{\text{cp}}^{\text{EoS}}$, $\Delta I_{\text{cp}}^{\text{EoR}}$ and $\delta I_{\text{cp}}^{\text{R}}$, as a function of stress/recovery cycles for the device depicted in Fig. 4.27. The amount of I_{cp} recovered per cycle, $\delta I_{\text{cp}}^{\text{R}}$ is observed to be consistent throughout the dynamic NBTI cycles, (b) Normalized $\delta N_{\text{it}}^{\text{R}}$ or $\delta I_{\text{cp}}^{\text{R}}$ recovered per cycle with respect to 1^{st} cycle based on R-D model simulation and experimental is compared.	98
Figure 4.29: $\delta I_{\text{cp}}^{\text{R}}$ (circle) is observed to be independent of $t_{\text{stress}}/t_{\text{relax}}$. This result does not corroborate a transport-limited mechanism of N_{it} generation and recovery as stipulated by the R-D model. It can also be observed that $\Delta I_{\text{cp}}^{\text{EoR}}$ (open triangle) is insensitive to t_{relax} and appears to be only determined by t_{stress} . This implies that a component of stress induced ΔI_{cp} is relatively permanent, i.e. irrecoverable.	99
Figure 4.30: (a) Identical $\Delta I_{\text{cp}}^{\text{EoR}}$ (open triangle) is observed at a given cumulative stress time (2×10^4 s), regardless of the cumulative recovery time which ranges from 400 to 2×10^5 s. This shows that a component of ΔI_{cp} which depends only on the stress time but not on the recovery time, i.e. irrecoverable. $V_g^{\text{stress}} = -2.6\text{V}$; temperature was 100°C . (b) $\Delta I_{\text{cp}}^{\text{EoR}}$ observed on a long term scale for up to 50 days.	102

- Figure 4.31: $|\delta V_{th}^S|$, $|\delta V_{th}^R|$, $|\Delta V_{th}^{EoS}|$, and $|\Delta V_{th}^{EoR}|$ as a function of the time ratio of stress to recovery interval, t_{stress}/t_{relax} . The triangle denotes $|\Delta V_{th}^{EoS}|$ and $|\Delta V_{th}^{EoR}|$ extracted after a given cumulative stress time (3×10^4 s). Stress intervals of 1000s undergo 30 dynamic cycles while stress intervals of 3000 s undergo 10 dynamic cycles in order to achieve the defined cumulative stress time. As a result, the cumulative recovery time is varied for all cases. 103
- Figure 4.32: Distinct activation energies of δI_{cp}^R and ΔI_{cp}^{EoR} imply that two different mechanisms are responsible for the measured ΔI_{cp} . This finding shows that ΔI_{cp} is made up of two different components..... 104
- Figure 4.33: Change in the charge pumping current, ΔI_{cp} at three different stages of dynamic NBTI as a function of the gate pulse width used during measurement. The dashed circle highlights the comparison among the three parameters at timing comparable to state-of-the-art integrated circuits..... 105
- Figure. 5.1: (a) Increase in the charge pumping current (ΔI_{cp}), measured by an ultrafast 100-ns pulse, as a function of stress time. Gate stress voltages are as indicated. Power-law dependence of ΔI_{cp} on stress time is evident and the exponent n ranges from 0.145-0.171. This observation has been regarded as an experimental proof for a H_2 diffusion-limited N_{it} generation mechanism, which has a theoretical exponent n of 0.167. (b) Evolution of ΔI_{cp} under dynamic NBTI (with equal stress and recovery intervals). ΔI_{cp}^{EoS} (ΔI_{cp}^{EoR}) denotes ΔI_{cp} at the end of the stress (recovery) cycle. The typical 2-stage ΔI_{cp} evolution, i.e. abrupt initial change followed by quasi-saturation is believed to correspond to the reaction and diffusion limited phases of the R-D model... 114
- Figure 5.2: Additional I_{cp} can be observed even by pulsing the gate for a longer duration between V_g^{stress} and $V_g^T < V_{th}$, i.e. always maintaining the device in strong inversion. In this regime, no electron-hole recombination can occur at interface states since their occupancy is not changed and the surface electron concentration is negligible. However, occupancy of hole traps in the oxide bulk would change in tandem with oxide band bending, modulated by the gate pulse (Fig. 5.3). Thus, the additional I_{cp} is due to “charge pumping” of hole traps. 115
- Figure 5.3: Schematic illustration of change in oxide band bending when the gate is pulsed from a negative stress voltage (V_g^{stress}) to a less negative level V_g^T ($< V_{th}$, threshold voltage), during constant base charge

pumping. (a) During stress ($V_g^T = V_g^S$), interface states and interfacial hole traps quasi Fermi energy (dotted line) are charged by holes from the inversion layer. (b) When V_g is pulsed to $V_g^T (< V_{th})$, interface trap occupancy does not change and surface electron concentration is negligible (surface potential is “pinned”). Hence, no net electron-hole recombination occurs at interface states. However, hole traps “falling” below the hole quasi Fermi energy are neutralized via electronic charge exchange with the gate and/or substrate valence band (dashed lines depicts oxide band bending at $V_g = V_g^{stress}$). A longer pulse width results in more hole traps being neutralized, and this is in turn registered as an increased supply of holes from the drain/source to recharge the traps when V_g is pulsed back to V_g^{stress} . When majority of hole traps are neutralized, further increasing the pulse width yields little ΔI_{cp} increase, giving an impression that ΔI_{cp} is independent on pulse timing [16]. This in turn leads to an incorrect inference of negligible hole trap contribution to the measured I_{cp} 115

Figure 5.4: (a) ΔI_{cp} measured during stress increases with CP pulse timing (or duty cycle). ΔI_{cp} under recovery is, however, is not affected by the change in duty cycle. The latter conforms to the CP theory [93], which stipulates that I_{cp} is independent of pulse duty cycle for a given N_{it} density. The former implies contribution of oxide traps at longer pulse timing. (b) Amount of ΔI_{cp} recovered per cycle, δI_{cp}^R is increased at a longer CP pulse timing, but it remains independent of stress/recovery cycling. The increase is due to more hole traps contributing to ΔI_{cp} measured during stress (as confirmed in Fig. 5.1 and explained in Fig. 5.2). The resultant more abrupt ΔI_{cp} decrease at the onset of recovery would potentially lead to a wrong conclusion of fast ΔN_{it} repassivation. 117

Figure 5.5: (a) Illustration showing combination of charged N_{ot} and N_{it} contributing to I_{cp} during stress. During charge pumping measurement, the gate pulse will allow charged N_{ot} near the interface to discharge, thus generating a parasitic component, I_{ot} which superimposes with I_{it} . Increasing the gate pulse from 100ns to 900ns will allow more time for charged N_{ot} located deeper into the gate oxide to discharge, which explains the increased I_{cp} observed during stress. (b) Illustration showing absence of charged N_{ot} during relaxation. Only contribution from N_{it} is recorded in I_{cp} measurement.

Increasing the gate pulse will have no impact on I_{it} as stipulated by CP theory [93].....	119
Figure 5.6: (a) Power-law exponent n for ΔN_{it} generation is shown to be ~ 0.28 , after subtracting a constant hole trap component (ΔI_{ot}) from the measured ΔI_{cp} . ΔI_{ot} is estimated from δI_{cp}^R obtained by dynamic NBTI (Fig. 5.4(b)). ΔI_{it} denotes the component of ΔI_{cp} due to ΔN_{it} . (b) Comparable exponents of ~ 0.27 are observed over a wide V_g^{stress} range. Thus, an exponent of ~ 0.16 cannot be regarded as an experimental proof of H_2 diffusion kinetics since the measured ΔI_{cp} is “contaminated” by a hole-trap signal. The relatively permanent nature of ΔN_{it} (Fig. 4.30(b)), however, does not suggest that ΔN_{it} arises from atomic hydrogen diffusion kinetics [7].....	120
Figure 5.7: (a) Exponent n of the as-measured ΔI_{cp} increases with temperature. This may be explained by the superposition of two mechanisms, i.e. ΔI_{ot} and ΔI_{it} , the latter having a stronger temperature dependence than the other (Fig. 5.8) (b) Power-law exponent n of ΔI_{it} is consistently $\sim 0.26 - 0.27$ over a wide temperature range, after subtracting the oxide-trap component ΔI_{ot} from the measured ΔI_{cp}	121
Figure 5.8: (a) Exponent n of the as-measured ΔI_{cp} increases with temperature before and after decoupling. This may be explained by the superposition of two mechanisms, i.e. ΔI_{ot} and ΔI_{it} , the latter having a stronger temperature dependence than the other (b) Distinct activation energies of δI_{cp}^R (ΔI_{ot}) and ΔI_{cp}^{EoR} (ΔI_{it}) imply that two different mechanisms are responsible for the measured ΔI_{cp} . This finding corroborates the results in Fig. 4.32, showing that ΔI_{cp} is made up of two different components.	122
Figure 5.9: Amount of threshold voltage shift during the stress period ($ \delta V_{th}^S $), the recovery period ($ \delta V_{th}^R $), and the threshold voltage shift at the end of the recovery period ($ \Delta V_{th}^{EoR} $) as a function of the number of stress/recovery (S/R) cycles (N). Inset shows evolution of $ \Delta V_{th} $ during stress and recovery.	127
Figure 5.10: (a) Evolution of $ \Delta V_{th}^{h_i} $ during stress and recovery, showing highly repetitive hole trapping and detrapping characteristic, respectively, over many stress/recovery cycles (thin lines; $N = 10$ to 30). Symbols denote average characteristics. For clarity, only average characteristics are shown for the 1.5 and 1.8 V recovery voltage. (b) Arrhenius relations for $ \delta V_{th}^S $ (or $ \delta V_{th}^R $) and $ \Delta V_{th}^{EoR} $, showing a	

much smaller activation energy of the former. Inset shows $\sim T^{1.3}$ temperature dependence of $ \delta V_{th}^S $.	128
Figure 5.11: (a) The average stress induced $ \Delta V_{th}^h $ characteristic of Fig. 5.10(a) is redrawn on a log-log plot. As-measured $ \Delta V_{th} $ of the first stress cycle is also shown. Difference between $ \Delta V_{th} $ and $ \Delta V_{th}^h $ yields $ \Delta V_{th}^{it} $ (the component of $ \Delta V_{th} $ due to interface state generation). Also shown is the evolution of $ \Delta V_{th}^{it} + \Delta V_{th}^{DLHT} $ for the case of 0 V recovery. (b) Evolution of $ \Delta V_{th}^{it} $ (circle), as extracted by the approach outlined in Fig. 5.11(a). Also shown is $ \Delta V_{th}^{it} $ based on ΔN_{it} measured by the CP method.	132
Figure 5.12: A device stressed at a moderately negative V_g^{stress} for an extended period beyond 1×10^5 s. Power law exponent of ΔI_{cp} evolves from ~ 0.266 in earlier stage of stress and ~ 0.084 in the latter stage of stress. Demonstrating the quasi-saturation of N_{it} . (inset) Generation rate of ΔI_{cp} decreases progressively with stress time, rendering further generation of N_{it} to be negligible	136
Figure 5.13: The “seasoned” device in this case was already stressed at a moderately negative V_g^{stress} for some time and deep level hole traps fully recovered using a positive gate voltage. The same “seasoned” device was then stressed again at a less negative V_g^{stress} to reactivate hole trapping for 8×10^3 s. A full recovery under a positive gate voltage indicates that only N_h is reactivated and there is no further generation of ΔN_{it} . This allows a precise examination of the time dependence of N_h at various less negative V_g^{stress} , independent of the influence of ΔN_{it} .	138
Figure 5.14: Evolution of hole trapping during stress, extracted from the as measured $ V_{th} $ in UFS on a “seasoned” device (a) For various gate stress voltages down to as low as -1.4 V; Logarithmic evolution, spanning many orders of time magnitude (i.e. no sign of saturation) is evident (b) Power law time exponent of ~ 0.05 across various gate stress shows a clear indication that classical assumption of saturation of hole trapping at ~ 1 s is inaccurate.	140
Figure 5.15: (a) Evolution of hole trapping across various temperature, extracted from the as measured $ V_{th} $ in UFS on a “seasoned” device (b) Arrhenius plot showing a relatively small activation energy of ~ 60 meV for HTD process.	141

Figure 5.16: (a) An illustration of how the contribution of ΔN_{it} towards the as-measured $ \Delta V_{th} $ may be readily extracted based on the time dependence data for hole trapping determined in Fig. 5.14. (b) The extracted $ \Delta V_{th} $ due to ΔN_{it} exhibits a power-law exponent of 0.25, in excellent agreement with that determined independently using an ultrafast charge pumping method, with appropriate oxide trap current correction as shown in Fig. 5.6.	142
Figure 5.17: (a) Gate stress voltage, V_g^{stress} dependence of HTD and ΔN_{it} , extracted after 8×10^4 s stressing at 100 °C, using the method outlined in Figs. 8-10). (b) A comparison of HTD extracted in this work to that obtained by a “1-s” assumption (i.e. assuming that hole trapping saturates in ~ 1 s). Significant underestimation of hole trapping using the 1-s assumption is apparent across the entire gate stress voltage range. This in turn overestimates the importance of interface trap generation.	143
Figure 5.18: For the devices examined, the importance of HTD, relative to ΔN_{it} , is observed to increase as V_g^{stress} scales towards the operating condition.	144

1 INTRODUCTION

With the recent surge in demand for high-tech gadgets and mobile applications, semiconductor companies have been in a race to scale down technology nodes to meet the demand for high performance consumer products. As the technology nodes scale down, challenges like device dimensions and thickness have reach their physical limits and new solutions like channel strain engineering and new gate dielectric (SiON/HfO₂) have been developed to meet the requirements. However, deriving good performance from the MOSFETs is not sufficient. As the cost of device fabrication increases exponentially with each technology node, the semiconductor industry also needs a reliable MOSFET that can keep costs of production low. Thus, the present day MOSFET is required to exhibit both strong performance and a longer device lifetime.

During the process of scaling, there has been increased influence of factors like increased gate electric fields, higher operating temperature and increasing nitrogen concentration in SiON gate dielectric that will potentially affect the device reliability. Negative-bias temperature instability (NBTI) has been well-known as one of the novel failures in the device reliability field which limits the circuit performance since 1966 [1]. It has been ascribed to the progressive shift of device parameters, such as threshold voltage, transconductance, drain current, *etc.* of the p-channel metal-oxide-semiconductor field-effect transistor (p-MOSFET). It is only during the last couple of years, that it has become a serious reliability issue for silicon integrated circuits [2-5].

Since the beginning of this decade, NBTI has attracted widespread attention from both industry and researchers. A well-known model based solely on interface traps generation process, known as Reaction-Diffusion (R-D) model was first introduced by Jeppson and Svensson [6] to explain the phenomenon of NBTI. This model has been widely adopted to explain several observed experimental results [7-9]. This includes instantaneous threshold voltage shift during the onset of NBTI stress, power law time exponent ($n \sim 1/6$) of NBTI evolution and temperature activated process with activation energy of 0.1~0.2eV and many more.

However, it was later reported that NBTI characterization done by conventional DC methodology underestimates NBTI as it is corrupted by the spontaneous recovery effect of threshold voltage ($|V_{th}|$) upon termination of stress [10]. As a result, this has enticed the development of fast measurement methodology, like on-the-fly $I_{d,lin}$ [11, 12], pulse I_d-V_g [10, 13] and fast V_{th} [7] characterization technique. This forthcoming development in fast measurement techniques have surfaced many new experimental observations which was previously not observed in conventional DC measurements. As a result, these novel observations have triggered a series of questionable doubts regarding the soundness of R-D model.

Till date, there have been many debates within the NBTI community regarding the commonly observed transient $|V_{th}|$ shift during stress and relaxation phase. Several believe that this transient $|V_{th}|$ shift is solely explained by the generation and passivation of interface traps at the Si/SiON interface based on the framework of R-D theory [7, 12]. On the other hand, some think that NBTI involves an additional

mechanism, suggested by several literatures as hole trapping caused by positive charging of preexisting oxide traps to explain the fast $|V_{th}|$ shift [14, 15].

With the recent advancement in fast measurement techniques, many methods had been proposed to extract the $|V_{th}|$ shift. But with the understanding that the experimental data is extremely sensitive towards measurement delay. It is practically challenging to compare the results obtained from each fast measurement methodology [16]. This is mainly attributed to the measurement delays inherent to their respective fast measurement methodologies.

As a result, this has rendered observations and conclusions made from different measurement methodologies to vary. To clarify this confusion, it is necessary to develop a new ultra-fast measurement technique that can deliver an accurate extraction of $|V_{th}|$ shift with minimum measurement delay to act as a benchmark for future NBTI testing. In addition, it is also necessary to clarify the controversial debate regarding the validity of R-D model which will have a major impact on NBTI reliability modeling on p-MOSFETs.

In this thesis report, the author starts by a literature study on the impact of interface and oxide traps which are mainly responsible for the threshold voltage shift in NBTI. This is followed by a detailed study on the well-known R-D model and its variants, which have been widely used to explain NBTI. In addition, recent challenges like recovery effect during NBTI measurement and implication of charge pumping are presented.

In Chapter 3, the author demonstrates two recently developed ultra-fast measurement methodology known as “Ultra-Fast Switching” (UFS) [17] and “Ultra-Fast Switching-Charge Pumping” (UFS-CP) [18] measurement methodology to address the need for an accurate characterization of NBTI threshold voltage shift. These two measurement methodologies have a measurement pulse delay of 100 ns, which is much shorter than those achievable by current fast measurement technique [7, 10-13]. UFS measurement methodology is built upon the Keithley Pulse-IV solution to sample I_d - V_g with a measurement delay of 100ns. First, this was demonstrated by the hardware capability to deliver a 100ns pulse first. This pulse was then improvised onto gate stress during NBTI testing, allowing a rapid transition between gate stress to measurement conditions within a 100ns. Thus, allowing a minimum interruption for stress during sampling. A similar concept was also used to develop UFS-CP. Unlike conventional charge pumping, UFS-CP has demonstrated the capability to switch the gate waveform from negative strong inversion during stress to positive accumulation within a 100ns. This will minimize the stay time in the positive accumulation which can potentially distort the experimental results. Both methodologies are also verified in this work for an accurate characterization of NBTI threshold voltage shift and interface traps generation in comparison with conventional DC measurements.

In Chapter 4, the kinetics of R-D model is first investigated through dynamic NBTI to resolve the controversial debate on the NBTI. However, the new insights discussed based on experimental results obtained from UFS, have revealed several fundamental flaws inherent to the R-D model. Hence, a reassessment of NBTI degradation mechanism is necessary and an in-depth analysis through dynamic NBTI revealed two distinct degradation mechanisms in NBTI. These two distinct degradation

mechanisms were identified as interface trap generation and hole trapping/detrapping, occurring in parallel upon NBTI stress. These new observations would also clarify controversial observation of a single NBTI mechanism made in the literature.

In Chapter 5, the author proposes independent separation methods demonstrated on UFS and UFS-CP measurement data to decouple the two degradation components, based on their distinct characteristics exhibited. These observations were also used to reconcile the earlier experiment observation in Chapter 4. A novel method using a “seasoned” device was also demonstrated to understand the hole trapping/detrapping mechanism in the absence of interface trap generation. Lastly, the author will summarize the essential findings in Chapter 6.

2 LITERATURE REVIEW

NBTI was observed as early as 1967 by Deal *at al.* [19], but an apparently consistent explanation was only made much later in 1977 by Jeppson and Svensson [6], based on the Reaction-Diffusion (R–D) model. In the model, the interface trap generation (ΔN_{it}) was ascribed to the dissociation of the Si–H bonds (by holes from the inversion layer) and the subsequent diffusion of atomic hydrogen away from the Si/SiO₂ interface (as the hydrogen remaining at the interface would repassivate the dangling Si bonds and reduce ΔN_{it}). This model was very successful then because it was able to match the evolution of the p-MOSFET parameter drift, e.g., threshold voltage shift $|\Delta V_{th}|$, transconductance degradation $|\Delta g_m|$, etc., which typically exhibited power-law time dependence with a ubiquitous exponent n of 0.25.

Moreover, the R–D model self-consistently explained the recovery that took place when the stress was stopped, in terms of the hydrogen back-diffusion and repassivation of the interface traps. The same 0.25 exponent was also observed in the early part of this decade when the interest in the NBTI was reignited [20–22]. This strengthened the notion that the underlying mechanism remained unchanged despite the fact that significant advancement in the gate oxide technology had taken place since the model was first proposed [6]. Expectedly, the R–D model has since been the theoretical basis for the interpretation and analysis of the NBTI in advanced CMOS technologies [7].

However, NBTI recovery studies [23, 24], forced researchers to abandon the traditional stress-measure-stress (SMS) approach for novel fast measurement methods

as discrepancies in experimental observation begin to surface. A major observation lies in the fact that novel fast measurement methods usually yield a much smaller exponent n of ~ 0.1 [10, 25]. This smaller exponent indicates the underestimation of the actual degradation due to recovery, which is particularly significant in the early stage of stress and would therefore increase the apparent power-law time exponent of the drift curve on a log-log plot. Consequently, a longer measurement delay would artificially yield a larger exponent [24]. Hence, these findings cast serious doubt over the previously observed exponent of 0.25, which was regarded as the experimental basis of the R–D model [6].

Guided by the simulation study of the R–D model, the actual exponent for ΔN_{it} was later proposed to be 0.167 [14, 26], driven by hydrogen diffusion in the form of a molecular hydrogen (H_2) instead of the atomic hydrogen (H). This proposition appeared reasonable as H was deemed inherently unstable and would dimerize to form H_2 molecules. Improvement to the R–D model was made to include dimerization as an important stage in the initial evolution of the exponent n , from a relatively large value to a subsequently steady value of ~ 0.17 [27]. The exponents of the drift curves smaller than 0.167 were ascribed to trapped holes (ΔN_h), sensed at the same time as ΔN_{it} by fast measurement methods [10, 25]. It was debated that hole trapping occurred at preexisting oxide defect sites and should be viewed as an extrinsic mechanism, whose contribution to the as-measured $|\Delta V_{th}|$ must first be excluded before the underlying mechanism for ΔN_{it} could be revealed. This reconciliation became widely accepted following the seemingly successful separation of the individual contributions of ΔN_{ot} and ΔN_{it} towards the as-measured $|\Delta V_{th}|$ by Lee *et al.* and Mahapantra *et al.*

[14, 28]. The separation was done based on the assumption that hole trapping in an ultrathin gate oxide rapidly saturates in approximately or less than 1 second and that it predominated the $|\Delta V_{th}|$ shift only in the beginning stage. By subtracting the component of $|\Delta V_{th}|$ due to the hole trapping, the resultant evolution of $|\Delta V_{th}|$ (presumed to be caused only by ΔN_{it}), exhibited an exponent n of ~ 0.16 at the later stage of stress evolution.

Apart from the deviation of experimental stress data from the expected exponent of 0.167, deviation of the experimental recovery data from the prediction of the R–D model has been pointed out. Reisinger *et al.* [13] were among the first to report recovery characteristics that spanned logarithmically for more than ten orders of time magnitude (from as short as $1\ \mu\text{s}$ to longer than $1 \times 10^4\ \text{s}$). On the other hand, the R–D model predicted a slower initial recovery with most of the subsequent recovery taking place over about three decades of time which differed from experimental results. (e.g., see [13], Fig. 9 (a)), i.e., while short-term recovery occurred more slowly, the long-term recovery proceeded much faster than that observed experimentally.

Grasser *et al.* [29] pointed out that the main difficulty in matching the experimental data to the simulation results lies in the fact that the equation for the recovery regime does not depend on any model parameters (including those of the stress regime). This precluded any possibilities of optimizing the model to fit the recovery data. Grasser *et al.* [29] went on to show that the discrepancy could, however, be reconciled by a combined dispersive-hole and hydrogen-transport model, where the hole trap density was determined by the trapped hydrogen (assumed to create hole traps).

A dispersive hydrogen model had also been adopted by some authors [30, 31] to explain the variation of exponent n with temperature, i.e., the non-Arrhenius behavior of the NBTI. However, other authors [14, 32, 33] showed that $|\Delta V_{th}|$ due to ΔN_{it} displayed an Arrhenius behavior, after accounting for the hole trapping component of $|\Delta V_{th}|$. Thus, the validity of the dispersive hydrogen transport framework remains questionable.

It was later shown by Kumar *et al.* [34] that the discrepancy between the experimental recovery characteristics and the model simulation may be resolved by considering different hydrogen transport properties in the SiO₂ and the polysilicon gate. The need for the latter consideration was particularly crucial in the present landscape of the ultrathin gate oxide (< 2nm) technology. On the basis that hydrogen diffuses more slowly in the polysilicon than in the SiO₂ [35], Kumar *et al.* demonstrated a good match with the experimental recovery characteristics [34]. This enhanced version of R–D model was able to explain the faster and greater recovery rate in thicker SiO₂ as compared to the thinner SiO₂. Attributing to the higher concentration of hydrogen present in the thicker SiO₂ to repassivate the interface traps upon termination of stress but some disagreed with the notion [36]. In addition, several fundamental issues underlying the enhanced version of R–D model were highlighted by Grassler *et al.* [29]. Imposing different hydrogen diffusion coefficients in the SiO₂ and polysilicon would result in a “non-universal” behavior, with a “hump” appearing in the recovery curve at a time depending on the relative difference in the hydrogen diffusion coefficients of the two materials [35]. But this feature was not observed

experimentally. Moreover, being essentially a transport-based model, the enhanced version of R-D model suffers from an intrinsic limitation just like the classical R-D model. (as elaborated later in section 4.1)

With the recent development, Huard *et al.* used a two independent component, hole trapping and interface trap generation to model NBTI. Unlike the understanding of the classical R-D model, Huard *et al.* argued that the interface trap generation is a permanent component and does not recover [37]. On the contrary, Mahapantra *et al.* [38] explained the possibility of interface trap recovery with a reduced back diffusion coefficient in the form of molecular hydrogen to reconcile the experimental results.

Despite the intense research effort made and models proposed in recent years to understand NBTI, the widely accepted model to account for NBTI, is still the R–D model with hydrogen diffusion in the form of molecular H₂, displaying a power-law time exponent of 0.167. However, it is emphasized that most of the studies reported to date have focused either only on the stress data or on a limited amount of stress and recovery cycle (typically, 2-3 stress/relax cycles) to conclude their findings, which may indirectly confine the understanding as the data in longer dynamic cycles are not well discussed.

2.1 Interface traps and oxide charges impact on NBTI

The principle operation of p-MOSFET is based on the charge-controlled channel whereby the gate voltage is applied across the gate dielectric to induce inversion

charges in the channel. However, any presence of positive or negative charges within the gate dielectric would affect the inversion charge and hence changes the gate voltage at which the transistor is turn on/off, i.e. the threshold voltage.

The p-MOSFET threshold voltage, V_{TH} is denoted by Eqn. (2.1):

$$V_{TH} = V_{FB} - 2\phi_F - \frac{|Q_B|}{C_{OX}} \quad (\text{Equation 2.1})$$

where, $\phi_F = (kT/q) \ln(N_D/n_i)$, $|Q_B| = (4qK_s\epsilon_0\phi_F N_D)^{1/2}$ and C_{OX} is the oxide capacitance per unit area. The other symbols have their usual meaning. The flatband voltage, V_{FB} is expressed by the Eqn. (2.2):

$$V_{FB} = \phi_{MS} - \frac{Q_f}{C_{OX}} - \frac{Q_{it}(\phi_s) + Q_{ot}}{C_{OX}} \quad (\text{Equation 2.2})$$

where Q_f is the fixed charge density, Q_{it} is the interface trap density and Q_{ot} is the oxide charge density.

As seen from this equation 2.2, the three parameters that can determine the threshold voltage shifts are the fixed charge density Q_f , interface trapped charge density Q_{it} and oxide charge density Q_{ot} .

$$\Delta V_{th}(t) = - \frac{\Delta Q_{it}(t) + \Delta Q_{ot}(t)}{C_{OX}} \quad (\text{Equation 2.3})$$

The density of fixed charge Q_f is normally associated to the process conditions of the gate dielectric growth. With the long standing history of SiO₂ gate dielectric manufacturing, the amount of fixed charges present in the gate dielectric is considered negligible in current MOSFET technology. Therefore, the only two possibilities that

can cause the threshold voltage of p-MOSFET to shift during operations are the interface trap density Q_{it} and the oxide charge density Q_{ot} . It is highlighted that interface trap density Q_{it} depends on the surface potential Φ_s , because the occupancy of the interface trapped charge is surface potential dependent. Any changes in either or both of these charge densities during NBTI stress will result in threshold voltage shift as denoted in Eqn. (2.3).

2.1.1 Interface Traps

For quite some time, interface traps generated at Si/SiO₂ interface during NBTI stress have widely postulated as one of the charges contributing to threshold voltage shift. In the silicon substrate, silicon is tetrahedrally bonded with each Si atom bonded to four Si atoms in the wafer bulk. When this silicon substrate is oxidized, the bonding configuration at the surface is as shown in Fig.2.1 for the (100) orientation with most Si atoms bonded to oxygen atoms while some of the remaining Si bonds are left dangling. This unsatisfied bonding configuration results in an interface defect state (often called interface trap) which is an interface trivalent Si atom with an unsaturated (unpaired) valence electron at the Si/ SiO₂ interface, denoted by $[\text{Si}_3 \equiv \text{Si}\cdot]$. The “ $\equiv$ ” represents three complete bonds to other Si atoms (the Si₃) and the “ $\cdot$ ” represents the fourth, unpaired electron in a dangling orbital (dangling bond).

As the presence of these interface traps formed during oxidation are electrically active and is undesirable due to threshold voltage control. Hydrogen forming gas annealing was performed to passivate the dangling bonds as denoted by $[\text{Si}_3 \equiv \text{Si-H}]$. However, this bonding energy formed between silicon and hydrogen, calculated by first

principal calculations is found to be relatively weak ($\sim 0.1\text{eV}$) [39], resulting in the Si–H bond to be easily broken again, when put under NBTI stressing. This interface trap generation mechanism and its impact on NBTI threshold voltage shift will be discussed further in reaction-diffusion (R-D) model in Chapter 2.2.

Detailed studies on the structure of this interface traps Si/SiO₂ interface have revealed a family of defects known as P_b centers dominating interface traps. This P_b center [40–43] is a [$\text{Si}_3 \equiv \text{Si}\cdot$] center with its unbounded central-atom orbital perpendicular to the interface and directed into a vacancy in the oxide immediately above it, as shown in Fig.2.1. And two of P_b center variants, known as P_{b0} and P_{b1} centers (detected by electronic spin resonance (ESR) measurement), have been identified to play an important role in NBTI [44]. However, the exact nature and structure of P_{b0} and P_{b1} centers still remains debatable. P_{b1} center was originally thought to be a Si atom backbonded to two substrate Si atoms, with the third saturated bond attached to an oxygen atom, designated as [$\text{Si}_2\text{O} \equiv \text{Si}\cdot$]. This identification was found to be incorrect, as the calculated energy levels for this defect do not agree with experiment [45]. A recent calculation suggested the P_{b1} center to be an asymmetrically oxidized dimer, with no first neighbor oxygen atoms [43]. By 2000, it was unambiguously established that both P_{b0} and P_{b1} center were chemically identical to the P_b center [39]. However, there is a charge state difference between these two centers indicating P_{b0} is electrically active, while some authors believed the P_{b1} to be electrically inactive. The two different effects are the result of strain relief in (100) silicon which results from the naturally occurring mismatch induced stress in the Si/ SiO₂ layer during oxide growth.

Figure 2.1: Structural model (100) Si surface

2.1.2 Oxide Traps

In the path of understanding the mechanism behind NBTI, some researchers have suggested that there is another additional degradation mechanism other than interface trap generation. They termed it as “hole trapping”, which is postulated to be caused by the presence of oxide traps in the current SiON gate dielectric. Ang *et al.* [33] provided experiment results that unambiguously showed the generation of positive trapped charge component, in conjunction with the generation of interface traps during NBTI stressing of p-MOSFET employing ultrathin silicon nitride gate dielectric. This observation was later further affirmed by several research groups [14, 46] which utilized fast electrical measurement and showed that generation of hole traps is a temperature insensitive process with a low activation energy ($\sim 0.06\text{eV}$).

However, the nature of these oxide defects and its contribution towards NBTI degradation mechanism is still not well understood at this point in time. Some researchers believe that trapping of H^+ , released via the dissociation of Si–H bonds, give rises to a positive oxide trapped charge which results in additional threshold voltage shift. But the transport of H^+ in the gate oxide is not consistent with the time dependence of parametric shift. Others believe that trapping of holes in the gate oxide gives rise to a positive oxide trapped charge. In 1995, P.M. Lenahan *et al.* revealed

that E'_{γ} center (a hole trapped at an oxygen vacancy) in SiO_2 gate dielectric can behave as a switching oxide trap through electron spin resonance (ESR) measurement [47, 48]. In addition, more paramagnetic defects were also highlighted in SiON gate dielectric, commonly known as K_n centers, which may also be a possible candidate in oxide traps as well [49]. This finding were also supported by Deep Level Transient Spectroscopy (DLTS) measurement. Poon *et al.* [50] observed an increase in the amount of hole trap in SiON gate dielectric as compared to thermal oxide (SiO_2), attributing to the oxygen deficiency in SiON gate electric as a result of oxygen removal during the nitridation.

However, with the continuous scaling down of the gate dielectric thickness to several angstrom, it is generally very challenging to spatially distinguish the oxide traps from interface traps for the range of practical measurement time.

2.2 The reaction-diffusion (R-D) model

In 1977, Jeppson and Svensson proposed a physical model, known as Reaction-Diffusion (R-D) model to provide a systematic explanation for NBTI [6]. This R-D model is primarily used to describe the mechanism of interface state generation at the Si/ SiO_2 interface, which translates to threshold voltage shift (ΔV_{th}) in contemporary pMOSFETs. Till date, this model has been widely adopted and fine-tuned by several researchers such as Alam *et.al*, Manahapatra *et.al* [8, 9] and Ogawa *et.al* [51] to explain many key experimental features as summarized below.

The R-D model can explain the key experimental features of NBTI summarized as follows:

1. The degradation is field-driven and is related to interface traps at the Si/SiO₂ interface [52].
2. Threshold voltage degradation due to NBTI is given by :

$$\Delta V_{th} \sim A \exp(-nE_D/kT) * t^n$$

where A is the stress factor, dependant on stress oxide field, n is the power law time exponent evolution of the degradation, ($n \sim 0.25$) and E_D is the activation energy, ($E_D \sim 0.5 \text{ eV}$). [52, 53].

3. As NBTI became a long term reliability concern, many research groups looked for and found long-term reduction of NBTI time-exponent n from 0.25 to 0.13-0.16. [54, 55].
4. Temperature dependent on NBTI reveals a characteristic of dispersive transport [12, 30]
5. Once NBTI stress is removed, a fraction of interface traps will self-anneal [7, 56].

2.2.1 Framework of R-D model

In general, the overall framework of R-D model comprises of a degradation process and a recovery process which explains the NBTI mechanisms during stress and relaxation respectively.

During the degradation process whereby negative gate stress voltage is applied to the gate of a p-MOSFET, inversion holes are generated in the channel. Some of this energetic inversion holes will be captured (via field assisted tunneling) by the passivated Si–H covalent bonds at the interface. Hence, weakening the existing Si–H covalent bond which is calculated to be easily broken at higher temperature based on first principle calculation as shown in Fig. 2.2 [57]. Also, this newly released hydrogen from the broken Si–H covalent bonds will diffuse away from the interface, leaving behind a positively charged interface state (Si^+) that is responsible for a higher threshold voltage and a lower transconductance of the p-MOSFETs. The former hole capture event and interface state generation is termed as the “reaction phase” while the latter hydrogen diffusion away from the interface is termed as the “diffusion phase”. This process is schematically illustrated in Fig. 2.3.

In the recovery process when the gate stress is removed, the dissociation of the Si–H bond that forced forward diffusion of hydrogen away from the interface no longer exists. Hence, the available hydrogen which is at/near the interface will re-passivate the broken Si–H covalent bonds almost instantaneously, restoring them to their passive Si–H state. Those remaining hydrogen atoms which are located further away from the interface will diffuse back and recombine with broken Si–H covalent bonds at a slower rate.

Figure 2.2: Model for hole-assisted field-enhanced thermal dissociation of Si–H bond resulting $k_f \propto E_c \exp(\gamma_T E_{ox}) \exp[-(E_F - a E_{ox}) / k_B T] \propto E_c \exp(\gamma E_{ox})$. Resultant field acceleration factor can be written as $\gamma = \gamma_T + a/k_B T$. [57]

Figure 2.3: (top) Schematic description of the reaction-diffusion model used to interpret interface-trap generation during NBTI stress. Broken Si–H bonds at the Si–SiO₂ interface create Si⁺ (interface traps) and H. Initial interface-trap generation rate depends on Si–H bond dissociation (reaction), while the later rate depends on H removal (diffusion). Some H dimerized into H₂ and since H is a slow diffuser, the H may diffuse (middle) or drift (bottom) away from the interface depending on the charge state of the diffusing species.

2.2.2 Formulation of R-D Model

The R-D model may be described by the following equations [9]:

$$\frac{dN_{it}}{dt} = k_F(N_0 - N_{it}) - k_R N_H(0) N_{it} \quad (x=0; \text{ Si/SiO}_2 \text{ interface}) \quad (\text{Equation 2.4})$$

$$\frac{dN_{it}}{dt} = D_H \frac{d^2 N_H}{dx^2} + N_H \mu_H E_{ox} + \frac{\delta}{2} \frac{dN_H(0)}{dt} \quad (0 < x < \delta) \quad (\text{Equation 2.5})$$

$$D_H (d^2 N_H / dx^2) = dN_H / dt \quad (\delta < x < T_{PHY}) \quad (\text{Equation 2.6})$$

$$D_H (dN_H / dx) = k_p N_H \quad (x > T_{PHY}) \quad (\text{Equation 2.7})$$

Where $x = 0$ denotes that Si/SiO₂ interface and $x > 0$ is (in the oxide) towards the gate, N_{it} is the density of interface traps at any given instant, N_0 is the initial density of unbroken Si–H bonds, N_H is the hydrogen concentration, k_F is the oxide field dependent forward dissociation rate constant, k_R is the annealing rate constant, D_H is the hydrogen diffusion coefficient, δ is the interface thickness, T_{PHY} is the oxide thickness and k_p is the surface recombination velocity at the oxide/polysilicon interface. The two-electron Si–H covalent bond is weakened once a hole is captured and this weakened bond (assisted by the electric field) is easily broken at relatively moderate temperature. The broken Si bonds act as donor traps and contribute to the shift in threshold voltage and reduction in transconductance. The H atoms released in the process can anneal the broken bonds, as described by the second term of the right hand side of Eqn. (2.4) ($N_H(0)$ is the H concentration at the interface, $x = 0$), or the H atoms may diffuse (or drift) away from the interface as Eqn. (2.5) shows. The last

term in Eqn. (2.5) is negligible. The H atoms may diffuse with diffusion constant (D_H) if the atoms are neutral, or drift with mobility μ_H if they are charged.

Since the rate of trap generation in Eqn. (2.4) is usually small compared to the dissociation and annealing rates and since $N_{IT} \ll N_0 \sim 5 \times 10^{12} \text{ cm}^{-2}$, therefore

$$\left(\frac{k_F N_0}{k_R} \right) \approx N_H(0) N_{IT} \quad (\text{Equation 2.8})$$

Also, Eqn. (2.5) may be expressed as a conservation equation which requires that the number of broken Si-H bonds equal that of total H concentration in the gate stack, *i.e.*,

$$N_{IT}(t) = \int_{x=0}^{x(t)=f(D_H, \mu_H, t)} N_H(x, t) dx \quad (\text{Equation 2.9})$$

Here $x(t)$ defines the tip of the diffusion or drift-front (Fig. 2.3, middle and bottom figures).

Consider diffusion of neutral atomic H as shown in Fig. 2.3 (middle figure). The diffusion distance at a given time t is $x(t) \sim (D_H t)^{1/2}$, therefore the right hand side of Eqn. (2.9) can be interpreted as the area under the triangle (sometimes referred to as ‘Triangle Method’) so that

$$N_{IT}(t) = \int_0^{\sqrt{D_H t}} N_H(x, t) dx = \frac{1}{2} N_H(0) \sqrt{D_H t} \quad (\text{Equation 2.10})$$

Inserting this expression for N_{IT} into Eqn. (2.8), we find

$$N_{IT}(t) = \sqrt{\frac{k_F N_0}{2k_R}} (D_H t)^{1/4} \quad (\text{Equation 2.11})$$

This reproduces the classic exponent of $n = 1/4$ which has been the signature of NBTI degradation based on a wide variety of experimental results. The success of the R-D model in interpreting the NBTI exponent led to prevailing view that NBTI degradation is rate-limited by the diffusion of atomic H across the gate dielectrics.

Based on this analysis, it is easy to see why some NBTI literature generally did not support interpretation based on diffusion of molecular H_2 or drift of H^+ (proton). For example, if the H-species are released as atomic H from Si–H bonds and then convert to and diffuse as molecular H_2 , then the ‘Triangle Method’ would predict

$$N_{IT}(t) = \int_0^{\sqrt{D_{H_2}t}} N_H(x,t) dx \sim \frac{1}{2} N_{H_2}(0) \sqrt{D_{H_2}t} \quad . \quad (\text{Equation 2.12})$$

In addition, the conversion between H and H_2 would be given by the Law of Mass Action, *i.e.*, $N_H(0)^2 / N_{H_2}(0) = \text{const.}$ Together with Eqn. (2.8), diffusion of neutral H_2 requires

$$N_{IT}(t) \propto \left[\frac{k_f N_0}{2k_R} \right]^{2/3} (D_{H_2}t)^{1/6} \quad . \quad (\text{Equation 2.13})$$

The exponent $n = 1/6$ reflects the bottleneck of H - H_2 conversion which results in a higher H concentration at the Si/SiO₂ interface [58]. This allows faster annealing of broken Si–H bonds and hence a reduced trap generation rate. Since there had been no experimental evidence in pre-2003 literature of NBTI degradation characterized by $n = 1/6$ exponent, a general consensus was that NBTI degradation by diffusion of molecular H_2 is unlikely.

Similarly, if H drifted away from the interface as a proton, we must retain the drift term in Eqn. (2.6), and then rewrite the conservation Eqn. (2.9) (in analogy to the Triangle Method) as

$$N_{IT}(t) = \int_0^{\mu_H E_{ox} t} N_H(x, t) dx = N_H(0) \mu_H E_{ox} t \quad (\text{Equation 2.14})$$

Here $\mu_H E_{ox}$ is the velocity with which the H^+ drift-front moves away from the interface as shown in Fig. 2.3 (bottom figure). Together with Eqn. (2.8), we find

$$N_{IT}(t) = \sqrt{\frac{k_F N_0}{k_R}} (\mu_H E_{ox} t)^{1/2} \quad (\text{Equation 2.15})$$

The $n = 1/2$ regime has also never been seen in NBTI measurements, leading one to conclude that NBTI degradation through proton transport is unlikely.

The above analysis shows that the time exponent n is dictated by and is a sensitive measure of the diffusing species, as also shown by numerical solution of R-D model [59]. Among the various exponents ($n = 1/2$ for proton, $n = 1/6$ for molecular H_2 , $n = 1/4$ for atomic H), only diffusion of atomic H appeared to be consistent with the pre-2003 literature.

2.3 Challenges faced in NBTI characterization

2.3.1 Recovery Effect during NBTI Measurement

In NBTI characterization, one of the crucial concerns was the corruption, impacted by the spontaneous recovery effect of threshold voltage ($|\Delta V_{th}|$) upon termination of stress [10]. In conventional current-gate (I - V) measurement, gate stress voltage was constantly interrupted to measure I_d - V_g curve. As a result, the measurement delay which occurred when gate stress voltage was removed, could range from a few seconds to tens of seconds, during which large amount (probably the majority) of $|\Delta V_{th}|$ have recovered. Rendering the $|\Delta V_{th}|$ extracted to be inaccurate (underestimate) as seen in Fig. 2.4.[53]. The power-law exponent for the time dependence increased while the delay of the interruption for measurement increased. Therefore, this fast recovery poses a very important challenge to the classical way by which device reliability is studied using the relatively slow conventional I_d - V_g measurement method.

Figure 2.4: (a) NBTI degradation over stress time for various delays between stress interruption and measurement (b) power law exponent versus delay time. [53]

Since then, this observation has motivated the development of fast measurement methodology, like on-the-fly $I_{d,lin}$ [11, 12], pulse I_d-V_g [10, 13] and fast V_{th} [7] characterization technique which has surfaced many new observations to explain NBTI. The advantage of the fast measurement is the minimization of the measurement delay, which has brought the stress interruption from seconds to the range of microseconds. This has greatly reduced the possibility of spontaneous recovery from taking place.

However, the other crucial concern is the lack of a practical measurement standard that would allow consistent benchmarking of NBTI in the face of fast and significant recovery. As the spontaneous recovery is extremely sensitive to measurement delay, the various measurement delays from different fast measurement methods are making the comparison difficult. (i.e. pulse I_d-V_g method has a delay of 100 μ s and fast V_{th} method has a delay of 0.5 μ s.). Hence, this triggers off a series of discussions over the infrequencies by the various fast measurement and experimental data [10, 60].

2.3.2 Implications by conventional charge pumping

Most of the time, conventional Charge Pumping (CP) and I_d-V_g measurement are coupled together to provide a comprehensive measurements of the $|\Delta V_{th}|$ and interface trap density $|\Delta N_{it}|$ during NBTI stress. However, similar to I_d-V_g measurement, the implication of measurement delay was also an issue for conventional CP measurement. T. Yang *et al.* [61] had illustrated the underlying problems associated both

conventional DC (V_{th}) measurements and CP (N_{it}) measurements together. As shown in Fig. 2.5(a), T. Yang *et al.* had illustrated two different measurement sequences whereby N_{it} measurement was done first, followed by V_{th} measurement and vice versa at measurement intervals. It was observed that $|\Delta V_{th}|$ for “ N_{it} first measurement” differed significantly from “ V_{th} first measurement”. The results clearly demonstrated the impact and importance of the V_{th} and N_{it} measurements order. This was explained by the implication of measurement delays whereby N_{it} measurement suffered from a longer delay than V_{th} measurement as shown in Fig. 2.5(b). Therefore, the recovery was greater during “ N_{it} first measurement” than during the “ V_{th} first measurement”, because of the longer measurement time employed.

Figure 2.5: (a) The time evolution of ΔV_{th} measured by DC Id-Vg method and ΔN_{it} measured by CP method. Two characterization approaches: In the first, V_{th} was measured (curve 1) followed by N_{it} measurement (curve 3). In the second, N_{it} was measured (curve 2) followed by V_{th} measurement (curve 4). (b) Zoom-in times t_1 and t_2 [15]

Another implication of conventional CP was the presence of positive bias during measurement. This problem was highlighted by Ang *et al.* [62], whereby he

demonstrated that by applying a positive gate bias of 1.5V would enhance $|\Delta V_{th}|$ recovery as shown in Fig. 2.6. As conventional CP is required to bias the gate voltage between accumulation and inversion, the application of the positive gate bias was simply unavoidable. As a result, any I_d - V_g measurement after conventional CP measurement would render the $|\Delta V_{th}|$ measured to be inaccurate.

Figure 2.6: Evolution of the normalized threshold voltage shift $\Delta V_{th}/\Delta V_{th}^0$ of three identical p-MOSFETs during negative-bias temperature stressing and post-stress recovery. For all the three cases, only dc I_d - V_g measurement was made during the stress phase. In the recovery phase, different gate relaxation voltages (V_g^{relax}) were applied. In one case, a CP current measurement step was introduced after the dc I_d - V_g measurement. [62]

3 ULTRA-FAST MEASUREMENT

METHODOLOGY

3.1 Background and Motivation

In NBTI reliability studies, an accurate measurement of threshold voltage shift, (ΔV_{th}) is essential for understanding the degradation mechanism as well as the extrapolation of device lifetime. However, as discussed in Chapter 2, NBTI measurement results were noticed to be highly sensitive to the measurement delay [60, 61] and positive gate bias [62], due to the fast recovery during the measurement. In conventional DC NBTI measurement, gate stress voltage is interrupted periodically according to user-defined time array, to measure I_d - V_g curve by sweeping gate voltage from 0 to the operating voltage with a small drain bias voltage. The measuring sequence of stress – measure – stress – measure ..., (termed “SMS” measurement sequence in short). The interruption of stress, termed as measurement delay, could range from a few seconds to tens of seconds, during which large amount (probably the majority) of ΔV_{th} might have recovered. And in some measurements where charge-pumping (CP) method was used to monitor generation of interface traps density, (ΔN_{it}), the voltage at the gate would be switched between accumulation (positive bias) and inversion (negative bias) at a defined frequency. As a result, this periodic positive bias would enhance the ΔV_{th} recovery effect. Therefore, the inconsistency in measurement delay in conventional DC measurement and effect of the positive gate pulse during CP measurement were identified as the major factors which have led to an inaccurate analysis of NBTI.

Currently, there are several fast measurement methods which have been developed to eliminate the recovery effects. As the main focus of this work is not on fast measurement methods, a detailed comparison of the various fast measurement methodologies were discussed in a separate work [63].

- **On-the-fly (OTF)** measurement applies only small variations on gate stress voltage during measurement [11] and captures the I_d degradation and real time G_m . $|\Delta V_{th}|$ is calculated from $|\Delta I_d|$ while considering G_m variation. However, OTF measurement has a fatal limitation, which is that it can never measure the threshold voltage shift before first measurement point. The $|\Delta V_{th}|$ calculated at stress voltage differs from $|\Delta V_{th}|$ extracted in sub-threshold region.
- **On-the-fly Interface Trap (OFIT)** measurement [16] sweeps from the stress voltage to the accumulation and back to stress voltage within a short time (1 μ s). The charge pumping current is sampled either from substrate or drain and converted to interface state.
- **Fast switching (FS)** measurement raises gate voltage and measures drain current [13, 60] degradation in linear / saturate operating region. If needed, $|\Delta V_{th}|$ is calculated based on assumption that $|\Delta V_{th}|$ is linearly proportional to $|\Delta I_d|$.

- **Ultra-fast V_{th} (UF V_{th})** measurement [13] uses a negative feedback loop to transit gate voltage to V_{th} . UF V_{th} can take measurement within 1 μ s.
- **Ultra-fast I_dV_g (UF I_dV_g)** measurement [60] sweeps gate voltage from stress voltage to 0 V and back to stress voltage in extremely short time (1 μ s) by pulse generator. The drain current is converted to voltage by low noise pre-amplifier and subsequently captured by oscilloscope together with gate voltage. $|\Delta V_{th}|$ is extracted from the I_d - V_g curves which are generated by reconstruction of oscilloscope data.

In the effort to eliminate recovery effects caused by measurement delay, Ultra-Fast Switching (UFS) characterization method [64] was developed with the capability to minimizing measurement delay down to 40ns. The measurement hardware setup is based on Keithley PulseIV setup, which has the capability of achieving gate pulsing down to 40ns. Drain current can also be sampled at measurement voltage with similar measurement delay of 40ns.

In this chapter, two ultra-fast measurement methodologies will be discussed in detail. Namely, these are the Ultra-Fast Switching (UFS) and Ultra-Fast Switching Charge Pumping (UFS-CP) methods. Both were developed based on the Keithley 4200 Pulse IV setup.

3.2 Keithley Pulse-IV Setup

3.2.1 Remote Bias Tee (RBT)

RBT (Remote Bias Tee) adapter is a coupler for DC bias and pulse output from pulse generator. The output of the RBT provides pulse output riding on the DC bias. Fig. 3.1. shows the simplified schematic of the RBT. The capacitor of RBT functions as a low-impedance component for high speed pulses, and as a high-impedance element for DC. This allows the high-speed pulses from pulse generator to pass through to the output, while blocking DC from the SMU. The inductors function as low-impedance components for DC, and as high impedance components for high speed pulses. This allows the DC bias from the SMU to pass through to the output, while blocking the high speed pulses from the pulse generator.

Figure 3.1: Simplified schematic of Remote Bias Tee (RBT)

3.2.2 3-Port Power Divider

A 3-Port Power Divider is included to divide the electrical power equally among its three connectors using a 16.67Ω resistor in each port. The Power Divider is used on the gate of a FET to provide an impedance matched signal path.

Figure 3.2: Schematic of 3-Port Power Divider

3.2.3 Keithley Pulse-IV Hardware Setup

The schematic setup of Keithley Pulse IV test system is shown in Fig. 3.3. In this setup, the following equipments are used:

- Pulse Generator (4200-SCP2)
- Source Measurement Unit (4200-SMU)
- Oscilloscope (4200-SCP2)
- Remote Bias Tee (4200-RBT) / 3-port power divider

Figure 3.3: Schematic of Keithley Pulse-IV measurement technique hardware setup

In general, the Pulse IV setup applies a pulse to the gate of the device under test (DUT), while DC biasing the drain. The source and body connections are connected to ground/shield. The dual-channel scope card measures the gate voltage and drain voltage. The drain current is calculated by the voltage drop across the 50Ω termination of the scope. Fig. 3.4 shows the corresponding waveforms for the test system. They are from select points to provide snapshots of the signals and to show the effects of various aspects of the setup. Although many waveforms are displayed in Fig. 3.4, there are only two waveforms that are actually measured by the scope card (waveforms B and E). Showing all waveforms together permits comparison of the DC offset and relative amplitudes in various parts of the schematic.

At the start of the test, the drain of the device under test (DUT) is biased with SMU2 and is waiting for a gate pulse to turn on the transistor and cause drain current, I_d , to flow.

- A. **Gate Voltage Pulse:** This waveform is the output from Channel 1 of the PG2 pulse generator card. Its amplitude is about the same as that seen by the DUT gate (C), except for some small cable losses. Note that there is no DC offset on this waveform, as any DC voltage would cause excessive power to be dissipated in the power divider.
- B. **Gate voltage pulse measured by the oscilloscope (oscilloscope channel 1):** This is the portion of the pulse that is transmitted through the power divider. Because of the divider and the 50W input impedance of the scope, the amplitude of the pulse is reduced 33%, plus any additional cable losses.
- SMU1:** This is the voltage output by SMU1 during Pulse IV testing. In this illustration, the DC voltage is fixed at 0 V.
- C. **Voltage pulse at DUT gate:** This waveform shows what the DUT gate “sees.” This is the pulse that turns on the transistor and causes the drain current to flow.
- SMU2:** DC Voltage bias applied to the drain, through the RBT.
- D. **Drain response:** This waveform shows the DC bias and the resulting response to the gate pulse. Note that the drain is DC biased and the pulse in this waveform is the result of the gate pulse. There is no pulse applied to the drain, rather the pulse shown is the response of the drain current flowing.
- E. **Drain voltage response measured by oscilloscope (4200-SCP2 Channel 2):**
- The response of the drain is due to the gate pulse turning on the transistor and allowing drain current to flow. This drain current (I_d) flows through the 50 W input impedance of the scope, which acts as a current sense resistor (R_{SENSE}), resulting in the voltage drop in the waveform (V_{SENSE}). This voltage response is negative due to the direction of current flow through the 50 W resistor and the

fact that the scope is ground-referenced. Note that this waveform is AC shifted to 0 V, due to the coupling capacitor in the RBT.

Figure 3.4: Waveform captured at selected points in the Pulse IV setup

3.3 Ultra-Fast Switching (UFS) Measurement Methodology

3.3.1 Experimental Setup of UFS

The physical setup of UFS as shown in Fig. 3.5 is exactly the same as the Keithley Pulse-IV hardware setup. The gate terminal of DUT is connected via a gate bias tee which acts as a coupler for both DC and AC voltage. The DC input is sourced from a SMU while the AC signal comes from a pulse generator which has the capability of achieving pulse width as short as 40ns. The drain terminal of DUT is also connected to a drain bias tee with DC connected to a SMU. Unlike gate bias tee, the AC output of drain bias tee is connected to an oscilloscope whereby any abrupt change in the drain current will be captured on the oscilloscope.

Figure 3.5: Schematic diagram of the UFS measurement set-up.

3.3.2 Principal of UFS Measurement

UFS measurement was originally developed from the concept of Keithley Pulse I_d - V_g by pulsing the gate terminal to the measurement voltage within a short period of time. The abrupt change in the drain current was sampled within this duration by a high speed oscilloscope. In order to prevent recovery effect in the midst of measurement, a specially designed test waveform for NBTI characterization is shown in Fig. 3.6. Generally, it can be broken down into 3 phases: (i) pre-stress, (ii) stress and (iii) relax.

Figure 3.6: UFS test waveform consisting of 3 different phases: (i) pre-stress, (ii) stress and (iii) relax

In the initial pre-stress phase, the fresh I_d - V_g curve was extracted to serve as a reference for the subsequent phases. Here, the DC voltage, V_g^{DC} was kept constant at $0V$. The measurement is taken when a short pulse of 100ns with the amplitude of the measurement voltage from the pulse generator is injected into the gate signal. For example, in the case if drain current is desired to be sampled at $V_g = -0.5V$. A negative gate pulse with amplitude of $V_g^{AC} = -0.5V$ is superimposed onto $V_g^{DC} = 0V$. As a result, the resultant gate voltage will be $V_g = -0.5V$. During the gate pulse transition,

the drain current denoted as I_d^{AC} , will be measured by a high-speed oscilloscope whereby the current-to-voltage conversion takes place automatically via the oscilloscope 50Ω internal resistance, R_{in} . I_d^{AC} can be back-calculated by the formula ($I_d^{AC} = V_g^{AC} / R_{in}$). As each gate pulse is pulsed to a defined measurement voltage, only one drain measurement can be measured. Therefore, multiple gate pulses are required to measure the whole I_d - V_g sweep.

In the subsequent stress phase, the DUT (p-MOSFET) is continuously stressed with a constant negative DC gate voltage denoted as V_g^{stress} . This will correspond to a constant DC drain current, I_d^{DC} which is sampled by the SMU at drain bias tee since DUT is in strong inversion. At selected time intervals, measurement was achieved by injecting a positive voltage pulse, V_g^{AC} into the gate bias-tee, which brings the overall gate voltage at the output of gate bias tee down to a measurement voltage magnitude ($V_{meas} = V_g^{DC} + V_g^{AC}$). The AC component of the drain current, I_d^{AC} arising from V_g^{AC} is measured using a high-speed oscilloscope whereby the current-to-voltage conversion takes place automatically via the oscilloscope 50Ω internal resistance. The drain current at the specific measurement voltage can be calculated by taking $I_d = I_d^{DC} + I_d^{AC}$.

Lastly, in the relax phase, the DC gate voltage is kept at relaxation voltage, V_g^{relax} . Depending on the aim of the experiment, it could be set to 0 V or +1/+2 V to investigate the effect of enhanced fast $|\Delta V_{th}|$ recovery by positive gate voltage. The gate pulse used for UFS measurement is set at pulse width of 100ns and is periodically generated at an interval of $100\mu s$, which represent 0.01% of the overall

stress cycle during measurement. Hence, minimum stress interruption is ensured in UFS measurement methodology to achieve an accurate ΔV_{th} measurement.

3.3.3 Validation of UFS Measurement

In order to validate UFS measurement, it was necessary to ensure that the hardware used in UFS would be able to perform up to expectation. It is mentioned that UFS is able to achieve a pulse width of 40ns. However, a pulse width of 100ns was chosen to ensure better voltage and current stability. Fig. 3.7 shows waveforms of gate voltage V_g^{AC} and drain current I_d^{AC} , which was captured by oscilloscope during pulsing. It can be observed that the V_g^{AC} ramped from 0V to 2.6V within 40ns. However, the oscillation of V_g^{AC} may result in a noisy I_d^{AC} versus time. However, this problem can be avoided by carefully choosing pulse width and measurement window. Data in between 60% and 80% of the pulse width is averaged and gives value of V_g^{AC} and I_d^{AC} . Due to the damping effect, the best choice of pulse width is in between 70ns and 100ns. More number of pulses can be used and averaged to increase accuracy of measurement. And in this work, the average number of pulses used for a single point measurement is 5 pulses and the time taken is approximately 800ms. For a full transfer 21 points I_d - V_g sweep from 0 to $-1V$ in steps of $0.05V$, it will take up to approximately 15s.

Figure 3.7: Waveforms of gate voltage pulse, V_g^{AC} and drain current, I_d^{AC} . Gate voltage pulse has amplitude of $2.6V$, pulse width at $100ns$, rising/ falling time at $10ns$. Data in between 60% and 80% of the pulse width is averaged and gives value of V_g^{AC} and I_d^{AC} .

Nevertheless, the validity of the UFS measurement still remains to be validated. I_d - V_g measurements were conducted on a similar unstressed device by both UFS and conventional DC measurements. Fig. 3.8 shows the results obtained by both UFS and conventional DC in linear and sub-threshold regime. It is clearly shown that the I_d - V_g measurement obtained by UFS is in well in agreement with the DC in the linear regime. However, in the sub-threshold regime, the agreement lasts up to $\sim 5\mu A$. Unlike high resolution SMU which has a resolution down to $\sim pA$ range, the physical connection of the drain terminal is connected via drain bias tee which has caused the resolution limit to be degraded to $\sim 5\mu A$. A comparison between the forward and reverse I_d - V_g sweep using UFS measurement between 0V to $-1.2V$ and vice versa was also made in Fig. 3.8(c) and (d) [65]). However, this does not pose a problem as the sub-threshold current in advanced technology like 65nm is expected to increase

due to the scaling of gate oxide and channel length. Therefore, this method can also be utilized in NBTI characterization for advanced technological nodes.

Figure 3.8: Drain current, $|I_d|$ measured by both UFS and DC method by pulsing/sweeping gate voltage from 0 to -1V. Drain voltage is kept at -0.1V. (a) Linear regime, (b) Sub-threshold regime, (c) forward and reverse sweep of Pulse IV from 0V to -1.2V in the liner regime (Fig. 1a of ref [65]), (d) forward and reverse sweep of Pulse IV from 0V to -1.2V in the sub-threshold regime. (Fig. 1b of ref [65])

The final step of validation involved NBTI characterization by UFS method in comparison with conventional DC method. Two identical devices were stressed at similar conditions to highlight the difference caused by measurement methodologies.

It was clearly observed from Fig 3.9, that the magnitude of $|\Delta V_{th}|$ captured by DC measurement is lower as compared to UFS measurement, for the given similar stress

period. This was explained by the effect of recovery due to SMS measurement sequence in DC measurement as discussed earlier in Chapter 2.3.1. Therefore, by using UFS measurement, the impact of recovery can be shown to be greatly minimized, resulting in a smaller power law exponent of ~ 0.11 as compared to DC measurement with a power law exponent of ~ 0.19 .

Figure 3.9: Comparison of $|\Delta V_{th}|$ measured by UFS and conventional DC. (a) Plotted on linear scale to highlight the recovery effect during stress measurement but negligible during relaxation measurement (b) Plotted on log-log scale to highlight the power law exponent which is critical for long term extrapolation. Constant current method is used to extract $|\Delta V_{th}|$ from I_d - V_g curve taken at various time intervals with drain voltage, V_d kept at $-0.1 V$. Devices were stressed for time period of 50,000s and relax for 1000s at $V_g = -2.4 V / 0 V$ at $100^\circ C$

3.4 Ultra-Fast Switching-Charge Pumping (UFS-CP) Measurement Methodology

The fundamental of UFS-CP operates similarly to conventional CP measurement which was used to monitor the generation of interface traps density, ΔN_{it} . Similar to DC measurement, conventional CP follows the SMS measurement sequence whereby

the stress is being interrupted and CP waveform is applied to measure charge pumping current, I_{cp} . It is also understood that during CP measurement, whereby the gate voltage will be swept between accumulation (positive gate bias) and inversion (negative gate bias) repeatedly, the presence of the positive gate bias will also enhance or forced V_{th} recovery. As a result, both the interruption of stress and the presence of positive gate bias in the CP waveform will affect the measurement of ΔV_{th} in NBTI. In UFS-CP method, these problems are resolved as the device is constantly kept at stress condition. At selected interval, CP measurements are done by injecting a gate pulse to change the overall gate voltage from strong inversion to accumulation and back within $\sim 100\text{ns}$. Unlike conventional CP, this positive gate bias duration is extremely short and will have minimum impact on ΔV_{th} recovery.

3.4.1 Experimental Setup of UFS-CP

The physical setup of Ultra-Fast Switch-Charge Pumping (UFS-CP) shown in Fig. 3.10. has some minor changes as compared with UFS. In UFS, the gate and drain terminal is connected via a remote bias tee to couple both DC and AC signal together. But in the UFS-CP setup, the gate connection remains unchanged as compared to UFS. The only difference is at the drain terminal where the drain bias tee is removed from the setup and replaced by a direct connection directly to a high-resolution SMU. This is necessary as it is understood that the I_{cp} is small ($\sim pA$). The presence of the drain bias tee will affect the resolution limit detected by the SMU which is around the range of a few $\sim \mu A$, making I_{cp} impossible to be detected. By removing the drain bias tee and connecting the drain directly to the high-resolution SMU, the resolution limit

can be increased up to $\sim fA$ which will provide extremely high sensitivity measurement for any changes in I_{cp} .

Figure 3.10: Schematic illustration of the ultra-fast switching-charge pumping (UFS-CP) measurement setup, adapted from ultra-fast switching method. The DUT is stressed or relaxed at a DC voltage, V_g^{DC} . CP measurement is performed at selected intervals by injecting a series of positive pulses of amplitude V_g^{AC} at the bias-tee, and the resultant I_{cp} is measured at the drain. The bias-tee allows very narrow pulse of duration 100 ns or less to be used.

3.4.2 Principal of UFS-CP Measurement

The fundamentals of UFS-CP measurement operates in the same way as conventional charge pumping. Using the ultra-fast pulsing capability of Keithley 4200 Pulse-IV hardware setup, a voltage pulse of 100ns is injected into the gate terminal periodically which will causes the overall gate voltage to oscillate between the inversion (below threshold voltage) and accumulation (above flatband voltage) regime of the DUT (p-MOSFET). This will trigger a recombination event occurring at the interface traps

which can be detected from drain (substrate) terminal where the net in-flow and out-flow of holes (electrons) during the gate pulsing will be integrated over a fixed period of time to obtain an excess current known as charge pumping current, I_{cp} .

The magnitude of I_{cp} can be converted into interface trap density, N_{it} using the following formula:

$$N_{it} = I_{cp} / (q.A.f)$$

where $q = 1.602 \times 10^{-19}C$, A is the area of the device and f is the frequency of the charge pumping

To further elaborate on how the measurement is done at different stages, a simple illustration of the test waveform is shown in Fig. 3.11. In general, the test waveform is used to characterize $|\Delta N_{it}|$ during NBTI stress and can be categorized into 3 different phases: (i) pre-stress, (ii) stress and (iii) relax.

Figure 3.11: UFS-CP test waveform consisting of 3 different phases: (i) pre-stress, (ii) stress and (iii) relax. At respective phases (pre-stress/stress/relax), the V_g^{DC} is set to their respective value ($-1V \rightarrow V_g^{stress} \rightarrow -1V$). During UFS-CP measurement, V_g^{AC} is injected into the gate to bring the overall voltage from V_g^{DC} to accumulation ($+1V$) and back V_g^{DC} within 100 ns.

As explained previously, the gate voltage swing must encompass the inversion regime (below threshold voltage) to the accumulation regime (above threshold voltage) in order to trigger the recombination event to measure I_{cp} . Therefore, in the initial pre-stress phase, the V_g^{DC} is set to $-1V$ instead of $0V$. The UFS-CP measurement is done by injecting a series of 100ns pulse of amplitude $+2V$ into gate bias tee, which will bring the overall gate voltage to $+1V$. During the gate pulsing, the SMU at the drain terminal samples the I_d over a fixed time period and integrates it to measure I_{cp} . This I_{cp} measured will be termed as $I_{cp}^{prestress}$ which will serve as a reference for subsequent measurements. As it transits into the stress phase, the gate is stressed continuously with a constant DC voltage ($V_g^{DC} = V_g^{stress}$). UFS-CP measurements are done in a similar manner as the pre-stress phase by pulsing the overall gate voltage to $+1V$ and I_{cp} is measured by drain SMU. The magnitude of V_g^{AC} in stress phase can be calculated by ($V_g^{AC} = |V_g^{stress}| + 1V$). Upon relaxation, the V_g^{DC} is set back to $-1V$ and the operation mechanism is exactly similar to pre-stress phase.

In all the UFS-CP measurements, the gate pulse used is set at pulse width of 100ns and is periodically generated at an interval of $10\mu s$, which represents 0.1% of the overall stress cycle during measurement. Therefore, both stress interruption time and positive pulse duration is kept to the minimum in UFS-CP measurement methodology to ensure an accurate $|\Delta I_{cp}|$ measurement.

3.4.3 Validation of UFS-CP Measurement

The first step in validating UFS-CP measurement is to ensure the performance of the gate pulsing waveform within $\sim 100\text{ns}$. Fig. 3.12 shows the actual waveform captured at the output of gate bias tee by the oscilloscope during pre-stress/stress/relax phases of the UFS-CP test waveform. It can be observed that the waveform of the stress phase is over pulsed to $+1.2\text{V}$ instead of the defined $+1\text{V}$. This is attributed to voltage overshoot during the gate pulsing whereby the pulse generator is expected to pulse a large magnitude of $+3.6\text{V}$ from -2.6V to $+1\text{V}$. However, this overshoot voltage of $+1.2\text{V}$ is not expected to affect the recombination event at the interface which corresponds to the I_{cp} measurement

Figure 3.12: Waveforms of gate voltage, V_g during UFS-CP measurement at different phases. During pre-stress and relax, the gate voltage is pulsed from -1V to $+1\text{V}$. During stress, the gate voltage is pulsed from -2.6V to $+1\text{V}$. The pulse width is defined at 100ns with rising/ falling time at 10 ns .

In order to monitor the I_{cp} measured at pre-stress/stress/relax phases, a constant base UFS-CP measurement (Fig. 3.13) is performed. The sample used in this experiment were done on a SiO_2 gate oxide thickness of $28\text{\AA}$. This is done by pulsing the top gate voltage to $+1V$ in steps $[-1V, -0.8V, \dots, +0.8V, +1V]$ with a constant bottom voltage at V_g^{DC} . Fig 3.13 shows the resultant I_{cp} samples at various charge pumping voltages, V_{cp} . The I_{cp} due to interface traps during pre-stress and relax phases, $I_{cp}^{\text{prestress}}$ and I_{cp}^{relax} can be obtained by taking the current difference between V_{cp} at $-1V$ and $+1V$. For UFS-CP measurements done at stress phase, the magnitude of I_{cp} measured is much higher as compared to pre-stress and relax phase. This is expected as high V_g^{stress} will induce a higher oxide field, inducing higher gate leakage current which will subsequently be measured by drain SMU during UFS-CP measurements. The I_{cp} due to interface traps during stress, I_{cp}^{stress} can be obtained by taking the current difference between V_{cp} at $-2.6V$ and $+1V$. It is highlighted that UFS-CP measurement cannot be done on devices with ultra-thin gate oxide ($< 28\text{\AA}$) as the tunneling leakage current will increase exponentially at high V_g^{stress} . As a result, the resolution and the sensitivity of I_{cp} (in pA range) measured due to interface traps will be greatly affected by the large tunneling current (μA range).

Figure 3.13: (a) Illustration showing constant base UFS-CP measurement (b) I_{cp} measured by UFS-CP at various V_{cp} for pre-stress/stress/relax phase. Measurements are done on a SiO_2 gate oxide thickness of $28\text{\AA}$ with $V_g^{\text{stress}} = -2.6\text{V}$. It can be observed for stress phase, there is a vertical shift as compared to pre-stress/relax phase. This is attributed to tunneling leakage current due to high V_g^{stress} and it is expected to increase for thinner gate oxide thickness which will affect the I_{cp} resolution.

The final step of validation involves two identical devices which were stressed at $V_g^{\text{stress}} = -2.6\text{V}$ for a period of 7000s. Interface traps characterization is measured by UFS-CP and conventional CP methods respectively and the results are plotted in Fig. 3.14. It can be clearly shown that the amount of ΔN_{it} measured by UFS-CP is higher as compared to conventional CP measurements. This difference is expected as conventional CP suffers from the effect of recovery due to SMS measurement sequence as discussed earlier. Therefore, by using UFS-CP measurements, it is able to suppress the impact of recovery, resulting in a smaller power law exponent of ~ 0.14 as compared to conventional CP of ~ 0.35 .

Figure 3.14: Comparison of $|\Delta V_{th}|$ measured by UFS-CP and conventional CP done on a SiO_2 gate oxide thickness of $28\text{\AA}$. (a) Plotted on linear scale to highlight the recovery effect during stress measurements but negligible during relaxation measurements (b) Plotted on log-log scale to highlight the power law exponent which is critical for long term extrapolation. Devices were stressed for time period of 7,000s at $V_g^{\text{stress}} = -2.6\text{V}$ at 100°C .

4 REASSESSING THE MECHANISM OF NBTI

4.1 Can NBTI be explained by Reaction-Diffusion Model

At present, generation of Si/SiO₂ interface traps, rate-limited by hydrogen diffusion (in the context of the reaction-diffusion (R-D) model [7-9]), is believed to be responsible for NBTI. This framework of the R-D model explained the fast “degradation transient” by a hole capture via the interface state, which resulted in the lowering of energy barrier, causing interfacial Si-H bonds to be broken at higher temperature. It also explained the initial fast recovery by interface traps passivation by interfacial hydrogen and the subsequent slow recovery by back-diffusing hydrogen. However, there is another school of thought whereby hole trapping, reported in numerous studies [14, 15], is regarded as another NBTI mechanism which resulted in a fast “degradation transient” arising from the positive charging of preexisting oxide traps. This hypothesis is also deemed consistent with (i) the initial fast recovery, explained by spontaneous hole detrapping and interface traps passivation by interfacial hydrogen (the latter remains a theoretical proposition); (ii) the subsequent slow passivation of interface traps, rate-limited by back-diffusing hydrogen [8].

Till date, R-D model still remains as a controversial model. In order to truly understand these mechanisms of NBTI, a systematic study was carried out to evaluate the validity of R-D model [7-9] by analyzing the transient characteristics of threshold voltage shift, $|\Delta V_{th}|$ through dynamic stress-and-relaxation sequences, also termed as dynamic NBTI. The test devices used in this work consists of p+ polysilicon gate p-MOSFETs with SiO₂ or SiON gate dielectrics with physical oxide thickness of 2.8nm and 1.7nm respectively. The nitrogen concentration for SiON gate dielectrics ranges

from 6-13 atomic percent. Using UFS and UFS-CP fast measurement, the mechanisms involved in NBTI are systematically studied and presented.

4.1.1 Simulation of Dynamic NBTI using R-D Model

The mechanism of interface traps generation and passivation (annealing) during the stress-relaxation cycles in dynamic NBTI (DNBTI) was examined using a modeling tool built based on the numerical and analytical solutions of the Reaction-Diffusion (R-D) [66]. The R-D model parameters used are $k_f = 6 \times 10^{-3} E_c \exp(0.65E_{ox}) \text{ s}^{-1}$ (where E_c is the oxide electric field due to mobile carriers and E_{ox} is the total oxide electric field), $k_r = 3 \times 10^{-9} \text{ cm}^3\text{s}^{-1}$, $N_0 = 5 \times 10^{12} \text{ cm}^{-2}$, $D_H = 3 \times 10^{-13} \text{ cm}^2\text{s}^{-1}$ and $k_H = 1.4 \times 10^{-3} \text{ cm}^3\text{s}^{-1}$ as reported by Islam *et al.*[57]. Fig. 4.1(a) shows the simulated result of a p-MOSFET undergoing dynamic NBTI stress. The gate voltage at stress (relaxation) was set at -2.6V (0V) with similar time duration of 500s at 100°C . In the right panel of Fig. 4.1(b), three simulated snap-shots of the hydrogen diffusion fronts during the 1st, 10th and 30th stress (S) and relaxation (R) cycles are captured.

Figure 4.1: R-D model simulated results. Device undergoes dynamic NBTI stress and relaxation for 30 cycles at $V_g^{\text{stress}} = -2.6\text{V}$ and $V_g^{\text{relax}} = 0\text{V}$ for duration of 500s each. (a) $|\Delta N_{it}|$ plotted with respect to stress time (b) Snapshots of hydrogen profile at the end

of respective stress (left) and relaxation cycles (right). It can be clearly shown that the hydrogen will diffuse further away from the interface ($x > 0$) and the availability of hydrogen at interface ($x = 0$) will decrease with longer cycles.

As discussed in Chapter 2, the framework of R-D model depicted the raise in $|\Delta V_{th}|$ during the onset of NBTI stress; the rapid transient threshold voltage shift $|\Delta V_{th}|$, was explained by the spontaneous generation of interface traps (N_{it}), caused by the interaction of hole capture which weakens and dissociate Si-H bonds at elevated temperature. By-products of dissociated hydrogen would later diffuse away from Si/SiO₂ interface locality with the highest hydrogen concentration, still remaining at interface, limiting the subsequent generation of N_{it} by a generation-repassivation phenomenon, which explained the latter slowly down of $|\Delta V_{th}|$ generation.

During the onset of recovery when NBTI stress is removed, the rapid decrease in $|\Delta V_{th}|$, is explained by an spontaneous annealing process whereby the previously generated N_{it} repassivated with the hydrogen “rich” interface present at/near the interface. However, this repassivation process would deplete the hydrogen concentration at the interface with time, which caused a hydrogen concentration segregation, thus, triggering a back diffusion of hydrogen back towards the interface which explained the latter’s slow recovery of $|\Delta V_{th}|$ in recovery phase.

If interface state generation or annealing were indeed determined by hydrogen transport away from or towards the Si/SiO₂ interface, then one would expect, for a given number of stress/relaxation (S/R) cycles under DNBTI, the net change in ΔN_{it} to vary according to the individual stress (t_s) and recovery (t_r) period (e.g. Fig. 8 of

ref. [7]). Another important attribute of the hydrogen transport model is that the relaxation characteristic is “universal”, i.e. individual time-dependent relaxation curves when normalized to the last stress value and plotted against the ratio t_r/t_s , would all converge to a common curve (e.g. Fig. 2 of ref. [29]). Nevertheless, once t_r/t_s is varied, the net ΔN_{it} would still be changed correspondingly.

Dependence of ΔN_{it} on t_s and t_r may readily be explained based on the evolution of the hydrogen concentration profile. On the basis of mass conservation, the area underneath the hydrogen profile should be equal to ΔN_{it} . Fig. 4.2 illustrates evolution of the hydrogen profile over a single S/R cycle for two typical scenarios. For a given t_s (Fig. 4.2(a)), ΔN_{it} at the end of the relaxation cycle clearly shows the dependence on t_r ; a long t_r would enable back-diffusion of more hydrogen species to the Si/SiO₂ interface, thus decreasing the net ΔN_{it} . On the other hand, fixing t_r and increasing t_s would lead to a higher net ΔN_{it} . An increased t_s allows more N_{it} to be generated as well as more hydrogen to be transported further away from the Si/SiO₂ interface, thereby limiting the amount of relaxation for a given t_r (Fig. 4.2(b)).

Figure 4.2: Evolution in the hydrogen concentration profile and the corresponding interface state generation, ΔN_{it} under dynamic NBTI in the context of the reaction-diffusion model. (a) The stress period, t_s is fixed while the recovery period, t_r is varied. A longer t_r leads to greater recovery of ΔN_{it} due to more back diffusion of hydrogen. (b) t_s is varied while t_r is fixed. A longer t_s leads to lesser recovery of ΔN_{it} since hydrogen has been transported further away from the Si/SiO_2 interface and thus more time is required for a given amount of hydrogen to diffuse back towards the interface. It should be pointed out that even during the recovery interval, some hydrogen continues to move away from the interface (see typical profiles at the end of the recovery period). This gives rise to self-limited recovery.

An important feature of the R-D model, not frequently discussed but evident in the hydrogen profiles in Fig. 4.3, relates to the fact that a fraction of the hydrogen species continues to diffuse away from the Si/SiO_2 interface, even during the recovery phase [7, 26] (dotted lines in Fig. 4.3(b)). This corresponds to the findings that, (i) ΔN_{it} recovery would not be complete in a given S/R cycle and (ii) the amount of ΔN_{it} recovery per S/R cycle should reduce progressively with increase in the number of S/R cycles. This is because in each cycle, more hydrogen species had diffused away from the Si/SiO_2 interface during the stress and recovery phases, thus further limiting

the amount of ΔN_{it} recovery in the ensuing cycles. When the number of S/R cycles is small, hydrogen is close to the Si/SiO₂ interface and thus ΔN_{it} recovery is relatively large. As the S/R cycle ensues, the hydrogen “diffusion front” will advance significantly further away from the Si/SiO₂ interface, and later outcomes are in agreement with the self-limiting recovery characteristic depicted in Fig. 4.3(a). These inferences are borne out by the R-D simulation [7, 26] results depicted in Fig. 4.3.

Figure 4.3: (a) The amount of N_{it} recovered during relaxation is plotted with respect to dynamic cycles. It is observed that there is a ~50% decrease in the amount of N_{it} recovered for 1st and 30th cycle. (inset) The amount of N_{it} recovered can be calculated by taking the N_{it} difference between “End of Stress” (EoS) and “End of Relaxation” (EoR) for the respective dynamic cycles. (b) Profiles of hydrogen concentration as a function of distance x into the oxide at the end of selected stress cycles. As stress/recovery cycling progresses, hydrogen concentration at the Si/SiO₂ interface decreases (arrow A) and the diffusion front advances further away from the interface (arrow B), indicating that more hydrogen has been transported away from the interface. These changes lead to the self-limiting recovery behavior depicted in Fig. 4.3(a), since the amount of hydrogen back-diffusion (and hence available for repassivation of interface states) in a given recovery interval would become gradually lesser as more hydrogen is transported further away from the interface with stress/recovery cycling.

Fig. 4.3(a) depicts typical evolution of ΔN_{it} recovered under DNBTI by taking the difference between End-of-Stress (EoS) and End-of-Relax (EoR). The amount of ΔN_{it} recovered per S/R cycle is observed to decrease progressively as the number of S/R cycles increases. At the end of the 30th S/R cycle, the ΔN_{it} recovered decreased by $\sim 50\%$ which corroborates with the substantial $\sim 50\%$ decrease in the hydrogen concentration at the interface from $2 \times 10^{17} \text{ cm}^{-3}$ (1st S/R cycle) to $1 \times 10^{17} \text{ cm}^{-3}$ (30th S/R cycle) as shown in Fig. 4.3(b). Therefore, from the simulation results based on R-D model, it is clearly shown that the decrease in N_{it} recovery which also corresponds to $|\Delta V_{th}|$ recovery for increasing DNBTI cycles, is readily explained by the hydrogen diffusion away from interface.

4.1.2 Comparison of UFS Dynamic NBTI Experimental Results with R-D Model

In order to validate the understanding of the R-D model with experimental results, similar dynamic NBTI experiments were carried out using UFS fast measurement methodology. In this section 4.1.2, SiON (4% nitrogen concentration) device sample with 1.8nm gate oxide thickness were stressed and device degradation was monitored in terms of $|\Delta V_{th}|$, extracted by the constant subthreshold drain current (20 μA) method. Channel length and width is 120 nm and 10 μm . Fig. 4.4 shows the $|\Delta V_{th}|$ whereby the device undergoes stress (relaxation) at -2.6V (0V) for 30 cycles with similar stress time (t_{stress}) and relaxation time (t_{relax}) of 1000 s. As observed from the result, there is a fast degradation and recovery transient which seems to conform to the framework of interface traps generation and passivation in the context of R-D

model. However, more detailed checks are mandatory and will be presented in this section.

Figure 4.4: UFS experimental results showing $|\Delta V_{\text{th}}|$ under dynamic NBTI stress and relaxation for 30 cycles at $V_g^{\text{stress}} = -2.6\text{V}$ and $V_{\text{relax}} = 0\text{V}$ for a similar time duration of 1000s at 100°C .

Figure 4.5: Illustration showing definition of (i) $|\delta V_{th}^S|$ ($|\delta V_{th}^R|$) denotes the change in V_{th} shift at the end of a particular stress (recovery) phases; (ii) $|\Delta V_{th}^{EoS}|$ ($|\Delta V_{th}^{EoR}|$) denotes the cumulative $|V_{th}|$ after a given number of stress (recovery) phases.

Based on the results obtained in Fig. 4.4, four parameters were defined to facilitate further explanation as shown (Fig. 4.5) : (i) $|\delta V_{th}^S|$ ($|\delta V_{th}^R|$) denotes the $|\Delta V_{th}|$ shift at the end of a particular stress (recovery) phase; (ii) $|\Delta V_{th}^{EoS}|$ ($|\Delta V_{th}^{EoR}|$) denotes the cumulative $|\Delta V_{th}|$ after a given number of stress (recovery) phases. These four parameters were plotted as a function to their respective DNBTI cycle number in Fig. 4.6(a). In the first stress/recovery cycle, it was noticed that the amount recovered, $|\delta V_{th}^R|$ is smaller than the initial stress, $|\delta V_{th}^S|$, which seems to conform to the understanding of the R-D model by the explanation of instantaneous re-passivation of hydrogen near the interface after stress is stopped immediately. And in the subsequent cycles, $|\Delta V_{th}^{EoS}|$ and $|\Delta V_{th}^{EoR}|$ increases at a slow pace of $\sim 1\text{mV}/\text{per cycle}$.

However, it was highlighted that the amount of $|V_{th}|$ generated and recovered per cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$ is observed to be fairly consistent for throughout 29 cycles of dynamic NBTI (ignoring 1st cycle). According to the understanding of the R-D model, the hydrogen present at the interface would have diffused further away after subsequent stress/recovery cycle [7]. That would have resulted in a decreasing amount of N_{it} recovery; causing the extent of recovery for a given interval to decrease progressively in subsequent dynamic cycles. However, the observed consistent component of $|\delta V_{th}^R|$ exhibited no sign of a progressively decreasing recovery. This result does not corroborate the notion which ascribed $|\Delta V_{th}|$ generation/recovery in the slow phase to hydrogen transport (Fig. 4.6(b)). If one ascribes the incomplete recovery to only a fraction of hydrogen returning to the Si/SiO₂ interface, it failed to explain the nearly constant $|\delta V_{th}^R|$ for all the DNBTI cycles (ignoring the 1st DNBTI cycle).

Figure 4.6: (a) Evolution of $|\delta V_{th}^S|$, $|\delta V_{th}^R|$, $|\Delta V_{th}^{EoS}|$ and $|\Delta V_{th}^{EoR}|$ as a function of stress/recovery cycles for the device depicted in Fig. 4.3. The cumulative degradation ($|\Delta V_{th}^{EoS}|$) is observed to increase progressively during the dynamic NBTI cycles. But the amount of $|V_{th}|$ degradation and recovery per cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$ is observed to be consistent throughout the dynamic NBTI cycles, (b) Normalized $|\delta N_{it}^R|$ or $|\delta V_{th}^R|$

recovered per cycle with respect to 1st cycle based on R-D model simulation and experimental is compared.

To further affirm on the dissimilarity in the two distinct characteristic trends between experimental and R-D model simulation as observed in Fig. 4.6(b), the time dependent of the hydrogen transport is reviewed in this section. Recall that the R-D model is governed by the time dependent kinetics based on hydrogen transport, which is sensitive to time ratio of stress (recovery) interval, ($t_{\text{stress}}/t_{\text{relax}}$), i.e. one would expect $|\delta V_{\text{th}}^{\text{S}}|$ and $|\delta V_{\text{th}}^{\text{R}}|$ to deviate from each other as the difference between t_{stress} and t_{relax} is enlarged. Further similar dynamic NBTI experiments were repeated for 20 S/R cycles with similar stress (relax) conditions at $-2.6V$ ($0V$), but with t_{stress} (t_{relax}) conducted at various time periods at 3000s (100s), 1000s (100s), 1000s (1000s) and 1000s (3000s) respectively. Similar parameters $|\Delta V_{\text{th}}^{\text{EoS}}|$, $|\Delta V_{\text{th}}^{\text{EoR}}|$, $|\delta V_{\text{th}}^{\text{S}}|$ and $|\delta V_{\text{th}}^{\text{R}}|$ of the last 20th DNBTI cycle were presented in Fig. 4.7, in order of the respective $t_{\text{stress}}/t_{\text{relax}}$, spanning from 0.33 – 30 (two orders of magnitude).

In the case ($t_{\text{stress}}/t_{\text{relax}} = 30$), where t_{stress} is prolonged to 3000s and t_{relax} shorten to 100s. One would expect N_{it} generation to increase with prolonged t_{stress} and in conjunction with shortened t_{relax} , back diffusion of hydrogen species back to Si/SiO₂ for N_{it} passivation would be limited. Based on the previous discussion on the understanding of R-D model in Fig. 4.3(b), the hydrogen “diffusion front” will inevitably be driven further away from the interface, displaying a salient characteristic of a decreasing N_{it} recovery for ensuing S/R cycles. However, the experimental results differ from the understanding of the R-D model. Instead, a consistent experimental observation of $|\delta V_{\text{th}}^{\text{S}}|$ and $|\delta V_{\text{th}}^{\text{R}}|$ components, extracted at the 20th S/R cycle showed

no sign of a decreasing recovery (Fig. 4.6). This observation again does not corroborate with the understanding of the hydrogen transport framework.

Similar conclusions were also derived from the other 3 cases, ($t_{\text{stress}}/t_{\text{relax}} = 10, 1$ and 0.33) where, t_{stress} was kept at 1000s but with t_{relax} varying from 100s to 3000s . With the relaxation period prolonged, one should have observed the evolution of $|\Delta V_{\text{th}}|$ recovery to be more pronounced based on the time dependent hydrogen transport kinetics. However, the experimental results did not correlate with this understanding. A consistent experimental observation of $|\delta V_{\text{th}}^{\text{S}}|$ and $|\delta V_{\text{th}}^{\text{R}}|$ components extracted at the 20th S/R cycle shows no sign of a increasing recovery across all 3 cases shows the ambiguity in the understanding of the R-D model.

Figure 4.7: Dynamic NBTI at various time ratio ($t_{\text{stress}}/t_{\text{relax}}$) performed over 20 cycles. The following parameters ($|\delta V_{\text{th}}^{\text{S}}|$, $|\delta V_{\text{th}}^{\text{R}}|$, $|\Delta V_{\text{th}}^{\text{EoS}}|$ and $|\Delta V_{\text{th}}^{\text{EoR}}|$) of the last cycle are plotted with respect to their time ratio. Given longer stress time (larger time ratio), the cumulative degradation ($|\Delta V_{\text{th}}^{\text{EoS}}|$) is observed to suffer more degradation as compared

to those with shorter stress time. But on the other hand, the amount of $|V_{th}|$ degradation and recovery per cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$ is still observed to be consistent despite the various time ratio.

One salient observation highlighted was the amount of $|\Delta V_{th}|$ generated and recovered per cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$ was consistent despite the fact that the time ratio was varied widely over two orders of magnitude. If the fundamental understanding of DNBTI was indeed governed by the hydrogen transport mechanism, the varied t_{stress} and t_{relax} would have caused a difference in the hydrogen profile after 20 S/R cycles and yielding a variation of net N_{it} recovery (observed by response in $|\delta V_{th}^R|$) for all four varied time ratio cases. However, the experimental observation of a consistent $|\delta V_{th}^R|$ across all varied time ratio cases in Fig. 4.7, clearly showed that DNBTI kinetics does not conformed to the time sensitive hydrogen based R-D model. This finding prompted the author to consider the possible phenomena behind NBTI.

There have been similar reported explanations in numerous studies which have considered hole trapping [14, 15] to be an additional mechanism in-situ with interface traps generation to explain NBTI. With the hindsight from the experimental results, it was speculated that the nature of this component appeared to exhibit characteristics of positive charged up state during stress intervals and recover to a neutral state during relaxation interval. Grasser *et al.* [67] arrived at a similar conclusion via a mathematical approach. It was mentioned that hole trapping (detrapping) is responsible for the rapid change in $|\Delta V_{th}|$ during NBTI stress (recovery) with the subsequent slow $|\Delta V_{th}|$ explained by interface traps generation (passivation).

If we revisit the earlier discussed DNBTI experimental characteristics in Fig. 4.7, the results can be re-interpreted as two distinct nature components coupled together; (i) one is a significant fast “cyclic” switching nature $|\Delta V_{th}|$ component ($\sim 60\text{eV}$), which appears to be able to charge up positively during stress intervals and recover fully during relaxation interval as observed by the consistency in (i.e. $|\delta V_{th}^S|$ or $|\delta V_{th}^R|$) over 20-30 dynamic cycles. (ii) the other is a $|\Delta V_{th}|$ component that degrades over stress time which was observed by gradual increase evolution in $|\Delta V_{th}^{EOR}|$ (Fig. 4.6(a) and Fig. 4.7). Theoretically, this proposed hypothesis of two degradation mechanisms is deemed consistent with the UFS experimental results.

To provide a more sound understanding about the physical mechanism behind this “cyclic” component, further supporting evidence for a two mechanism framework is provided in Fig. 4.8. An Arrhenius experiment was conducted at fixed oxide field ($E_{ox} = 11\text{MV/cm}$) with temperature ranging from $50^\circ\text{C} \sim 150^\circ\text{C}$ in Fig. 4.8(a). The rather small activation energy ($\sim 45\text{ meV}$) suggested that the cyclical $|\Delta V_{th}|$ component arises from a hole trapping/detrapping mechanism (via tunneling). *Ab initio* simulation has shown that H^+ near the Si interface could behave as a hole trap [14]. Hole trap generation in SiO_2 thus not necessarily only refers to Si–O bond breaking, which is generally perceived as more energy-intensive. On the other hand, stronger thermal activation of $|\Delta V_{th}^{EOR}|$ ($\sim 0.1\text{ eV}$) [39] shows that a larger energy barrier must be overcome to create this component of $|\Delta V_{th}|$. The larger energy barrier is probably needed to initiate some local change in the oxide structure, thus making the resultant defect immune to recovery. As expected from this stronger thermal activation, the fraction of these recovery-immune defect states is increased at higher temperature

[Fig. 4.8(b); open triangle]. However, negligible recovery over extended periods does not conform to the hydrogen-diffusion controlled recovery mechanism stipulated by the R-D model.

Figure 4.8: (a) Arrhenius plot of the cyclic component, i.e. $|\delta V_{th}^S|$ (or $|\delta V_{th}^R|$) and the permanent component $|\Delta V_{th}^{EoR}|$; percentage of $|\delta V_{th}^S|$ (or $|\delta V_{th}^R|$) and $|\Delta V_{th}^{EoR}|$ relative to the total $|V_{th}|$ shift as a function of temperature; (c) dependence of $|\delta V_{th}^S|$ or $|\delta V_{th}^R|$ and $|\Delta V_{th}^{EoR}|$ on the concentration of nitrogen in the gate oxide under a given stress condition ($V_g^{stress} = -2$ V; temperature = 100 °C).

Further insight revealed that the physical origin of “cyclic” component shares some commonality in atomic origin linked to nitrogen defect in the SiON dielectric (Fig. 4.8(c)). With the increase in density of nitrogen precursor within the gate dielectric, the amount of “cyclic” component was found to increase as well. With the hindsight that nitrogen precursor in SiON dielectric plays an active role for hole trapping/detrapping mechanism. For the current technologically nitrided oxide, another possible switching hole-trap candidate, namely, the k_N center ($N_3 \equiv Si^+ \cdot Si \equiv N_3$), had been identified earlier, in addition to the E' center, via an electron spin

resonance measurement [68]. Owing to their proximity to the Si substrate, the hybridized “dangling” sp^3 orbital on one of the Si atoms can readily exchange an electronic charge (via quantum mechanical tunneling) with the underlying Si, causing the E' or k_N defect centers to rapidly “cycle” back and forth a positively charged state and a neutral state [69]. This finding was also supported by Deep Level Transient Spectroscopy (DLTS), indicated a larger amount of hole traps observed in SiON [50]. Evidence that the evolution of a large part of $|\Delta V_{th}|$ during stress (or $|\delta V_{th}^S|$ in the later cycles) and during recovery (or $|\delta V_{th}^R|$) is governed by a tunneling mechanism, indicated by the relatively small activation energy of ~ 45 meV (Fig. 4.8(a)). In the next section, further insights will be presented to provide a sound understanding of the hypothetical two mechanisms for the NBTI framework as well as the understanding of these phenomena and its implication on NBTI .

4.2 Reassessing NBTI Mechanism : Two Mechanism framework

In the previous section, we discussed the discrepancies of the hydrogen transport R-D model as observed from UFS experimental results. This “non-hydrogen transport” characteristic was substantiated by the “*non-decreasing*” cyclical $|\Delta V_{th}|$ component and inability to demonstrate N_{it} recovery over an extended relaxation period. In addition, the observation of two distinct activation energies (~ 45 meV and 110 meV) through Arrhenius experiment shed light that supports the notion of two different degradation mechanisms explaining NBTI. Based on the hindsight from experimental results, NBTI is postulated to be explained by a spontaneous hole trapping/detrapping

mechanism acting in-situ with a background of interface traps generation. In the next few sections, discussion on the origin and the characteristics of hole trapping/detrapping and interface trap mechanisms will be presented to reassess NBTI mechanism from a two mechanism proposed framework.

4.2.1 Origin of Hole Trapping/Detrapping in NBTI

A cyclic behavior for the increase and decrease in the magnitude of the threshold voltage (V_{th}) was revealed in section 4.1.2 as evident in dynamic NBTI (DNBTI) experiment [18, 70]. This cyclic behavior observation was shown earlier to be inconsistent with the hydrogen transport model [6, 7], which stipulated that the decrease or recovery of V_{th} must reduce progressively under dynamic NBTI. This repetitive behavior of the dynamic NBTI is strikingly similar to that of a class of switching hole traps observed following *X*-ray irradiation (see [69], Fig. 9) and substrate hot-hole injection stress (see [71], Fig. 4). A widely proposed candidate for a switching hole trap at/near the Si/SiO₂ interface is the E' center ($O_3 \equiv Si^+ \cdot Si \equiv O_3$), arising from the oxygen vacancy defect [69]. Electron spin resonance (ESR) measurement on NBTI [47, 48, 72-74] has shown that oxygen vacancy defects (or hole traps) are the origin of this cyclic threshold voltage shift ($|\Delta V_{th}|$).

According to the first principle calculations of oxygen vacancies in amorphous SiO₂, the vast majority of oxygen vacancies (90%) do not pucker after capturing a hole, which is denoted as E'_δ centers [75, 76]. The most favorable E'_δ centers located at shallow energy band regime and have dimer configuration at positively charged state,

makes the formation and recovery of defects very fast. The two Si atoms adjoining a pre-existing oxygen vacancy rebond and form a neutral dimer in amorphous SiO₂. Applying high electric field stress, the neutral dimers may lose one electron and leave the other electrons shared by the two Si atoms, which is the E'_δ center.

Figure 4.9: Schematic of neutral oxygen vacancy (left) and the positively charged E'_δ center (right), which is also called as shallow state in energy level point of view.

E'_δ centers can be classified into two sets depending on whether they can or cannot at least metastably capture an electron in the neutral state. The set of E'_δ centers that cannot capture an electron in the neutral state introduce energy levels ranging from 0.5 eV to 1 eV about the valence band. On the other hand, the set of E'_δ centers that can capture electrons introduce slightly higher energy levels ranging from 1.9 eV to 2.1 eV about the valence band. The two sets of E'_δ centers can also be distinguished in terms of the Si-Si bond distances of the E'_δ centers. The E'_δ centers that cannot capture an electron in the neutral states have shorter Si-Si bond than those that can capture an electron in the neutral states.

Figure 4.10: The electrical state energy distribution of shallow hole traps (E'_δ) when p-MOSFET is under negative gate voltage stress.

The shallow E'_δ centers provide temporal dispersion of the hole transportation through the SiO_2 because of the shallow energy in the dimer structure. The shallow energy level of E'_δ centers, especially energy levels ranging from 0.5 eV to 1 eV above the valence band, makes the capturing/ releasing holes extremely fast under excessive tunneling holes from substrate to gate. The neutral Si-Si bond captures a hole and forms E'_δ center without a Si atom puckered. An E'_δ center has no energy barrier to capture an electron and become a neutral Si-Si bond because of its dimer structure. The emission time for the trapped holes are supposed to be on the order of $\exp(-E_T/kT)$ [77, 78], where k is Boltzmann's constant, T is the temperature in Kelvin, and the E_T is the trap energy in eV. Even though the UFS fast measurement data might not be able to capture all the positive charges of E'_δ centers, the large gap between the UFS fast measurement data and the conventional slow measurement data is the strong experimental evidence of the E'_δ centers in first principle calculation, as shown in Fig.

3.9. The large transient $|\Delta V_{th}|$ upon application of stress shows the immediate formation of E'_δ centers after high-electrical field stress is applied.

Another form of oxygen vacancy, known as E'_γ centers, are formed with significant structural relaxation when the neutral dimer Si-Si bonds capture a hole. This accounts for only a minority (~10%) of the oxygen vacancy as a prerequisite condition for localized void space to be present in order for significant structural relaxation to occur. E'_γ centers can be further classified into $E'_{\gamma4}$ and $E'_{\gamma5}$ centers depending on the bonding configuration of the puckered Si atom upon structural relaxation.

For an $E'_{\gamma5}$ center to be formed, one of the Si atoms will suffer significant relaxation and relaxes back past the virtual plane defined by its three oxygen neighbors and bonds with another network oxygen atom and a network Si atom. The puckered Si atom now becomes five-fold-coordinated with a localized positive state at ~3.8 eV above the valence band. However, the $E'_{\gamma5}$ centers are examined to be unstable in this positively charged state [75]. When an electron is given back to this center, it collapses immediately to the former dimer configuration without an energy barrier. So $E'_{\gamma5}$ centers can also be the origin of fast hole traps.

In conjunction with $E'_{\gamma5}$, an $E'_{\gamma4}$ center also suffers significant relaxation with one of the Si atoms relaxing back past the virtual plane defined by its three oxygen neighbors and bonds with another network oxygen atom. This oxygen atom now becomes four-fold-coordinated with a positive charged state at ~4.3 eV from the valence band.

Unlike $E'_{\gamma 5}$ center, when given an electron to its center, an $E'_{\gamma 4}$ center still remains in its puckered configuration with an energy barrier for Si–Si bond reformation spanning 0.5 to 1.2 eV. The electron, in fact, goes to fill the dangling bond, whereas the hole is still left on the puckered Si side. This neutral dipole configuration, upon capturing a hole, provides a positive defect state around 3.5 eV above the valence band.

The present hole-trap model for NBTI [79] is based on the Harry Diamond Laboratories (HDL) framework [69]. The framework behind the HDL (or $E'_{\gamma 4}$ center) model is that strained Si-Si bonds broken upon the capture of holes structurally relax into the stable $\equiv\text{Si}\cdot^+\text{Si}\equiv$ configuration (Fig. 4.11), which in turn acts as a switching hole trap causing the V_{th} fluctuations of DNBTI. It stipulates that a switching hole trap is formed from the structural relaxation that follows the capture of a hole at the oxygen vacancy defect ($1 \rightarrow 2$). In this case, the strained Si-Si bonds in E'_{γ} center will be broken upon a hole capture and the resultant back-bonding of the Si^+ to a neighboring O atom has stabilized the defect cluster, acting as a switching hole trap precursor that explains the alternate transition of between a positive state (2) and a charge-compensated state (3) (see Fig. 4.11). Reformation of the Si-Si bond ($3 \rightarrow 1$) is deemed unlikely and would perhaps only account for the recovery over a very long term. Alternate transitions of the switching hole trap between the positive (2) and charge-compensated (3) states appear to explain the cyclical behavior (Fig. 4.12 (b); $|\delta V_{\text{th}}^{\text{R}}| \approx |\delta V_{\text{th}}^{\text{S}}|$) and highly repetitive evolution (Fig. 4.12(c)) of $|\Delta V_{\text{th}}|$ under alternate stressing and relaxation cycles.

Figure 4.11: Schematic illustration of the $E_{\gamma 4}'$ center model for NBTI [79]. An important feature of this model is the significant asymmetrical structural relaxation which occurs upon hole capture ($1 \rightarrow 2$), which results in one of the Si atom moving past the plane containing the oxygen atoms and getting back-bonded to a neighboring oxygen atom. This significant structural relaxation reduces the likelihood of the reformation of Si-Si bond ($3 \rightarrow 1$). Transition from state 3 to 2 (during stressing) and from state 2 to 3 (during relaxation) are believed to explain the cyclic shifts observed in the threshold voltage $|\Delta V_{th}|$.

4.2.2 Observation of Hole Trapping/Detrapping in NBTI

It was highlighted that the test devices (DUTs) were p^+ polysilicon gate p-MOSFETs employing SiON with gate oxide thickness of 1.8nm (4% nitrogen concentration). Channel length and width is 120 nm and 10 μm . DNBTI test was carried out at 100 $^{\circ}\text{C}$. An ultrafast switching method, based on the pulsed current-voltage setup, was used in the measurement of the linear I_d - V_d curve [25, 64]. The narrow timing (100 ns) and short duty-cycle (0.1 %) gate pulses ensured that recovery was minimized. $|\Delta V_{th}|$ was extracted by a constant subthreshold drain current method.

To further examine the nature of the constant $|\delta V_{th}^R|$ which is postulated to be hole trapping/detrapping arising from oxide traps, we compare $|\delta V_{th}^R|$ to the amount of $|\Delta V_{th}|$ increased in the prior stress cycle ($|\delta V_{th}^S|$) as a function of the number of S/R cycles in Fig. 4.12(b). Although $|\delta V_{th}^S| > |\delta V_{th}^R|$ in the initial few cycles, the former soon approaches the latter as dynamic stressing continued, i.e., in the later stage of dynamic stress, the amount of $|\Delta V_{th}|$ increased during the stress cycle is almost entirely recovered in the ensuing relaxation cycle. At this juncture, the author would like to highlight that an analysis based on a limited number of stress/recovery cycles (particularly the first few ones) would easily mislead one into concluding that the result seems to conform to the understanding of R–D model (on the basis that $|\delta V_{th}^S| > |\delta V_{th}^R|$, which appeared to suggest a self-limiting recovery behavior). However, data from subsequent S/R cycles, which shown $|\delta V_{th}^S| \approx |\delta V_{th}^R|$, in conjunction with a constant $|\delta V_{th}^R|$ right from the first S/R cycle, clearly indicate that this conclusion is incorrect. The evolution of $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$, was examined from the 10th to the 30th S/R cycles and for all the S/R cycles, respectively, in Fig. 4.12(c) (thin lines). The symbols denoted the average characteristics. The closely clustered thin lines, with less than 5% deviation from the mean, affirmed the repetitive nature of $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$. In addition, further evidence of a relatively small activation energy of ~ 45 meV (Fig. 4.8) revealed by Arrhenius experiment, postulates that the repetitive nature arises from a hole trapping/detrapping mechanism (via tunneling). Based on the first principle calculation of oxygen vacancy, this understanding also coincided with the presence of shallow E'_δ centers (0.5 eV to 1 eV above valence band) [75], which favors the rapid switching of charged state to charge compensated state (short tunneling constant) into and out of the ultra-thin gate oxide.

Figure 4.12: Typical evolution of threshold voltage shift, $|\Delta V_{th}|$ during the stress (-2 V) and relaxation cycles of dynamic NBTI. (b) $|\Delta V_{th}|$ at the end-of-stress (EoS), end-of-recovery (EoR), and the amount of $|\Delta V_{th}|$ shift during each stress and relaxation cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$, respectively as a function of the number of stress/relaxation (S/R) cycles, n . The constant $|\delta V_{th}^R|$ in conjunction with $|\delta V_{th}^S| \approx |\delta V_{th}^R|$, ($N > 10$) clearly depict a cyclic $|\Delta V_{th}|$ shift behavior. (c) This is confirmed by the highly repetitive $|\Delta V_{th}|$ shift over many stress and relaxation cycles (thin lines). Symbols denote the average.

4.2.3 Shallow-Level (Fast) Versus Deep-Level (Slow) Hole Traps

In the previous section, we discuss about the repetitive evolution of $|\Delta V_{th}|$ in dynamic NBTI experiment (Fig. 4.12), attributing it to hole trapping/detrapping mechanism, as evident by the relative small activation energy of ~ 45 meV. This also coincides with the shallow-level energy state of E'_8 center (0.5 eV to 1.0 eV above valance band) [75, 76] which favors the rapid switching of charged state to charge compensated state and vice versa in an ultra-thin gate oxide. To facilitate discussion, we will term the $|\Delta V_{th}|$ component attributed to hole trapping as $|\Delta V_{th}^h|$.

However, one must also consider the wide spread of the defect energy state and tunneling time constant in an amorphous nature gate oxide. If we revisit the evolution of $|\Delta V_{th}^h|$ in Fig. 4.13 (open symbol), it is observed that 80 % of the overall $|\Delta V_{th}^h|$ of the 1000s stress cycle is reached in less than 10s. Similarly, ~90 % of the overall $|\Delta V_{th}^h|$ induced by the stress cycle is recovered in less than 10s. These observations are evidence of very rapid trapping/detrapping of holes [10], in general agreement with the very short direct-tunneling time constant in the ultrathin gate oxide. Nevertheless, evolution of $|\Delta V_{th}^h|$ did not saturate but continues to evolve by about 20 % and 10 %, respectively, for the remaining stress and recovery periods. Considering that the ultrathin oxide limits the range of trapping time constants “spatially”, such a wide distribution of trapping time constants revealed experimentally must necessarily point to a considerable spread in the energy states of the trapped holes.

As compared to the former dynamic NBTI case at relaxation gate voltage, ($V_g^{relax} = 0$ V), it was observed that the application of positive relaxation voltage ($V_g^{relax} = +1.5$ V) during dynamic NBTI had enhanced the magnitude of recovery $|\Delta V_{th}^h|$ by ~10% as compared to former case (open symbol; Fig. 4.13). However, the intriguing fact about this component is that it displays a similar repetitive nature like the former case. However, one may question that the application of positive gate voltage may have shortened the detrapping time constant of shallow-level hole traps, which allows more recovery of a charged states to its charged compensated state for the given time period of 1000 s.

In a separate experiment, two devices were stressed and relaxed at $V_g^{\text{relax}} = 0$ V. In one scenario, one of the device was further relaxed with a positive $V_g^{\text{relax}} = +1.5$ V for 2000s, while the other was left in the room ambient and monitored for 30 days. It was observed that the device that relaxed at +1.5V exhibited an abrupt recovery of $|\Delta V_{\text{th}}|$ and remained constant (~ 17 mV shift) when back at 0V. This seems to suggest that all previously charged hole traps are fully discharged by the presence of positive gate voltage. However, the device which was left in room ambient exhibits an extreme slow recovery but later recovered to the similar magnitude (~ 17 mV) after ~ 10 days. Unlike shallow-level hole traps that have a short emission constant, the experimental result suggests that there is another group of hole traps that exhibit a long emission constant that could take up days, as evident from Fig. 4.14. However, we need to emphasize that the experiment was carried out on an ultra-thin gate oxide (1.8 nm) whereby such long tunneling time constant is considered not feasible. Thus, it is postulated that these positive charged states are “locked in” at deep-level where their energy states are “pinned” by the Si/SiON band discontinuity, hence hindering the emission of the charges state with an energy barrier to overcome.

Experimental evidence in support of holes trapped at deep-level energy states, with characteristically long emission time under low negative oxide field, was reported previously [80]. The inference on the existence of deep-level hole traps was based on a larger $|\Delta V_{\text{th}}|$ observed for static stress, as compared to bipolar stress, for a given density of interface traps [80, 81]. For bipolar stress, alternate application of positive gate voltage pulses would have most of the deep-level hole traps (DLHTs) “discharged” through lowering of the trap states below the Si conduction band. Non-negligible

spontaneous acceleration of $|\Delta V_{th}^h|$ recovery is evident when a +1.5 V (instead of 0 V) gate voltage was applied during relaxation (filled triangle; Fig. 4.13), further affirming the inference of “deep-level hole trapping” during NBTI stress. It should be mentioned that $|\Delta V_{th}^h|$ subsequently changes by less than 5% for the remaining 1000 s period; indicating almost complete annihilation of DLHTs by the positive gate voltage; remnant DLHTs may be ascribed to those located very close to the interface and as a consequence, their energy states are “pinned” by the Si/SiON band discontinuity.

Figure 4.13: Thin line: Evolution of the hole trapping component of threshold voltage shift, $|\Delta V_{th}^h|$ during the stress and recovery cycles of dynamic NBTI (from the 10^{th} to 30^{th}). Symbols denote average characteristics. Only the average characteristics are shown for the case of 1.5 V gate recovery voltage. The gate stress voltage (-2.2 V) was the same for both cases.

Figure 4.14: Evolution of $|\Delta V_{th}|$ of two devices undergoing NBTI stress (-2.6 V) for 8×10^3 s. One device is then relaxed at 0 V and monitored for 30 days. Another was relaxed at a positive voltage ($+1.5$ V) for 2000 s and relaxed back to 0 V for another 1000 s. Evidence showed that recovery of deep level hole traps is accelerated by a positive gate bias.

A multiphonon-field-assisted tunneling (MPFAT) theory [79], whereby subsequent structural relaxation moves the defect state to a deeper level, was recently proposed to explain the relatively wide distribution of hole trapping time constants in the ultrathin gate oxide. The MPFAT process is characterized by a relatively strong temperature activation ($\sim T^2$ dependence on temperature), as shown in ref. [79]. On the other hand, $|\Delta V_{th}^h|$ is found, in our work, to have relatively small activation energy of 40 meV or $\sim T^{1.3}$ dependence on temperature (Fig. 4.15(b)). A possible reason for the discrepancy lies in the different measurement methods employed, which result in different amount of fast hole trapping/detrapping being measured. In ref. [79], the “on-the-fly” method [11] was used where the initial delay of 1 ms was believed to have underestimated the

fast hole trapping component. Assuming logarithmic time dependence down to much shorter times (based on Fig. 1 of ref. [79]), we estimate $\sim 27\%$ of $|\Delta V_{th}^h|$ (at the end of the 1000s stress period) is reached 1 ms after stress is applied (refer to the stress-cycle data (filled circle) for the case of +1.5 V recovery in Fig. 4.13). Subtracting this component of fast $|\Delta V_{th}^h|$ from the overall measured value (at 1000s) yields a stronger temperature dependence ($\sim T^{-1.6}$) for the remaining “slower” $|\Delta V_{th}^h|^*$ (Fig. 4.15(b)).

Figure 4.15: (a) A comparison of the activation energy of the hole trapping component of threshold voltage shift, $|\Delta V_{th}^h|$ and the relatively permanent component at the end of the recovery interval, $|\Delta V_{th}^{EoR}|$ (due mainly to interface state generation). $|\Delta V_{th}^h|$ was determined using a 1.5 V gate recovery voltage (cf. Fig. 4.13; filled triangle). (b) The corresponding power-law dependence of $|\Delta V_{th}^h|$ on temperature T , giving an exponent of ~ 1.3 . Also shown is the result for $|\Delta V_{th}^h|^*$, an underestimated

form of $|\Delta V_{th}^h|$ assuming that the V_{th} estimated at 1 ms after the start of stress is the pre-stress V_{th} . Reducing the contribution of the fast hole trapping component clearly reveals a stronger dependence of slower hole trapping, which can be understood in terms of the multiphonon field-assisted tunneling model.

Figure 4.16: A comparison of the gate stress voltage dependence of fast and slow threshold voltage shift, $|\Delta V_{th}|$. The fast $|\Delta V_{th}|$ component corresponds to $|\Delta V_{th}|$ taken 1 s after stress has commenced while the slow $|\Delta V_{th}|$ component is the subsequent $|\Delta V_{th}|$ shift for the remaining stress period (5×10^3 s).

The stress gate voltage (V_g^{stress}) dependence of fast $|\Delta V_{th}^h|$ is also different from that of the slow $|\Delta V_{th}^h|$ and $|\Delta V_{th}^{it}|$ ($|\Delta V_{th}|$ due to ΔN_{it}), as evidenced in Fig. 4.16. The former follows quasi-linear dependence while the latter exhibits quadratic dependence on V_g^{stress} [79, 82]. The quasi-linear gate voltage dependence, in conjunction with the relatively weak thermal activation, suggest elastic tunneling as the mechanism for fast hole trapping. The experimental results depicted in Fig. 4.17 and 4.18 further support the presence of fast hole trapping that exhibits a quasi-linear dependence on the gate voltage. In the experiment, a relatively low V_g^{stress} (-1.2 V) was first applied to the test

device for 5×10^3 s. V_g^{stress} was subsequently decreased by -0.2 V and the procedure repeated as shown in Fig. 4.17. The initial “jump” in $|\Delta V_{\text{th}}|$ (measured after ~ 1 s) was recorded and ascribed as $|\Delta V_{\text{th}}|$ due predominantly due to fast hole trapping each time V_g^{stress} was decreased by 0.2 V. This $|\Delta V_{\text{th}}|$ component (denoted by the filled rectangle) is compared against the subsequent increase in $|\Delta V_{\text{th}}|$ for the remaining stress period of $\sim 5 \times 10^3$ s (denoted by the open rectangle) in Fig. 4.18. Interestingly, the initial change in $|\Delta V_{\text{th}}|$ following each V_g^{stress} decrement of -0.2 V was found to be almost the same throughout the range of V_g^{stress} investigated, regardless of the “stress history” of the device. This observation may be explained by the simple “trap-filling” model (based on the assumption of uniform hole trap density) depicted in Fig. 4.19 — each equal decrement of V_g^{stress} would raise an equal amount of initially neutral hole trapping centers (see shaded regions) above the Fermi potential, resulting in them becoming positively charged.

Figure 4.17: Evolution of threshold voltage shift, $|\Delta V_{\text{th}}|$ during NBTI. Starting from -1 V, the gate stress voltage was decreased, in step of -0.2 V, each time after the device had been stressed at a given voltage for 5×10^3 s.

A current viewpoint regards fast hole trapping as a “parasitic” effect arising from preexisting oxide traps and it should therefore not be considered part of NBTI [14]. However, more experimental investigation may be necessary before one entirely excludes the role of fast hole trapping on NBTI. As shown in Fig. 4.12(b) and 4.13, the conclusion of a cyclic switching hole trap mechanism remains unchanged regardless whether the fast hole trapping component is included or not. This result suggests a single hole trapping mechanism but with a wide dispersion of time constants, in accordance to an earlier inference of a wide distribution in the energy states of the trapped holes [80, 81]. Observations of different V_g^{stress} and temperature dependence “embedded” in an “aggregate” $|\Delta V_{\text{th}}^{\text{h}}|$ is not surprising given the amorphous nature of the gate oxide, which necessarily implies a spread in the activation energy and hole trapping time constant. On the same note, a question now arises: Is the larger activation energy (110 meV) of $|\Delta V_{\text{th}}^{\text{EoR}}|$, i.e. the relatively permanent component of $|\Delta V_{\text{th}}|$, also a consequence of the spread in activation energy of a common mechanism (Fig. 4.15(a))? We will revisit this point in the section 4.2.5 of discussion.

Figure 4.18: A comparison of fast and slow threshold voltage shift, $|\Delta V_{th}|$ resulting from the stress methodology shown in Fig. 4.17. The almost equal increase in fast $|\Delta V_{th}|$ for a given increase in $|V_g^{stress}|$ from one stress cycle to the next confirms the linear dependence of fast $|\Delta V_{th}|$ on V_g^{stress} . The dashed line is an eye-guide only.

Figure 4.19: A possible explanation for the almost identical increment of $|\Delta V_{th}|$ when the magnitude of the gate stress voltage was successively increased by a given step size (Fig. 4.18). Assuming a uniform energy distribution of hole trap precursors in the ultrathin gate oxide, a given negative change in the gate voltage, ΔV_g ($V_g \ll V_{th}$) will raise an equal amount of trap precursors above the Fermi energy, resulting in them becoming positively charged through spontaneous hole capture from the channel.

4.2.4 Correlation Between Oxide Trap Generation and Hole Trapping/De trapping Mechanism

It was earlier discussed that the observation of hole trapping/detrapping phenomena from dynamic NBTI arose from the origin of E'_{δ} center [75]. These are often classified as oxide vacancy defects (also known as hole traps), which are responsible for the transient fluctuations in V_{th} . However, the origin of these switching hole traps precursors responsible for the transient charging/discharging effect remains unclear. Was the origin of hole traps precursors an inherent property of the amorphous gate oxide, i.e. is present during fabrication process. Or if it is due to an extrinsic man made factor, whereby stressing the device could generate more of such hole traps center precursors, which may make create pathways for more trapping and detrapping activity. In this section, we will provide a systematic analysis based on UFS experimental results which reveals that the E'_{δ} center precursors, responsible for hole trapping/detrapping mechanism is inherent to the nature of the amorphous gate oxide.

To examine the postulation of stress-induced generation of hole traps precursors, dynamic NBTI was typically carried out in two phases in a continuous manner with stress and relaxation period at 1000 s as depicted in Fig. 4.20. The first phase (C1) involves three dynamic NBTI cycles carried out at a lower stress (relax) voltage of -1.9 V (0 V), followed by the second phase (C2) which involves a higher stress (relax) voltage of -2.3 V (0 V). The purpose of the two phases, involving a lower and higher stress voltage is to allow an intentional generation of hole traps precursors during higher stress (C2) phase, as observed from increase in $|\Delta V_{th}^{EoR}|$. If the origin of hole

trap precursors are stress-induced, the higher stress would anticipate generation of more hole trap precursors. These will create pathways for increased hole trapping/detrapping activity, which can be signified by a rising $|\delta V_{th}^R|$ trend.

Figure 4.20: For the range of negative V_g^{stress} commonly employed for NBTI testing, we found that $|\Delta V_{th}^R|$ at a given V_g^{stress} is unaffected by a preceding stress at a more negative V_g^{stress} . (a) A fresh DUT was first subjected to dynamic NBTI (DNBTI) stress at $V_g^{\text{stress}} = -1.9$ V. This was followed by a -2.3 V DNBTI stress. This sequence was repeated several times and evolution of $|\Delta V_{th}|$ captured. (b) Despite several more negative V_g^{stress} at -2.3 V, $|\delta V_{th}^R|$ at $V_g^{\text{stress}} = -1.9$ V remains the same as the very first time the DUT was stressed at -1.9 V.

However, the experiment result shown in Fig. 4.20 does not support the notion that hole traps precursors are stress-induced. It was observed after the application of the higher stress (after first cycle of C2 phase), a significant amount of positive traps have been generated as observed from the increased $|\Delta V_{th}^{\text{EoR}}|$ (after 6000 s - Fig.5.4(a)). The fact that the extracted $|\delta V_{th}^R|$ component registered at both -2.3 V and -1.9 V remains comparable over 4 repeated cycles (Fig. 4.20(b)) suggests that it is the same pool of hole traps precursors responsible for the cyclic $|\Delta V_{th}|$ for the respective given

stress conditions and uncorrelated to the notion of more hole trap precursors being generated at the higher stress voltage.

To further substantiate this finding, a separate experiment was carried out where a device was first stressed at $V_g^{\text{stress}} = -1.15\text{V}$ ($\sim 4\text{MV/cm}$) and relaxed at 0V repeatedly for 4 cycles. The test was then repeated in a continual manner with V_g^{stress} increased (negatively) by -0.25V each time until it reached -2.65V ($\sim 10\text{MV/cm}$). This represents the “ramp-up” phase. This was followed by “ramp-down” phase where V_g^{stress} is decreased (positively) by the same 0.25V step back to -1.15V , and similar four S/R cycles applied at each V_g^{stress} step. Fig. 4.21(a) shows the evolution of $|\Delta V_{\text{th}}|$ for the test sequence. By taking the difference in the $|\Delta V_{\text{th}}|$ registered during end of stress ($|\Delta V_{\text{th}}^{\text{EoS}}|$) and end of relaxation ($|\Delta V_{\text{th}}^{\text{EoR}}|$) for a particular cycle, the amount of cyclic $|\Delta V_{\text{th}}|$ component, $|\delta V_{\text{th}}^{\text{R}}|$ would be reflected on Fig. 4.21(b). It is observed from the evolution during the ramp-up phase that additional hole trapping occurs, as evident from the larger cyclic $|\delta V_{\text{th}}^{\text{R}}|$ as well as the increase of the relatively permanent $|\Delta V_{\text{th}}|$. However, it is highlighted that the cyclic $|\delta V_{\text{th}}^{\text{R}}|$ at lower V_g^{stress} did not register an increase on the ramp-down phase (Fig. 4.21(b)).

Figure 4.21: (a) A fresh DUT was subjected to 4 stress (@ -1.15 V) and relaxation (@ 0 V) cycles. Each S/R cycle lasted 1×10^3 s. The test was then repeated on the same DUT with V_g^{stress} increased (negatively) by -0.25 V each time, until $V_g^{\text{stress}} = -2.65$ V. The V_g^{stress} was then decreased (positively) each time by the same voltage step and the test repeated. Evolution of $|\Delta V_{th}|$ for the entire test sequence is shown. (b) Despite having subjected the DUT to a higher stress oxide field, there is no change in the $|\delta V_{th}^R|$ at less negative gate stress voltages.

There are two possible explanations for the observation. First, new stress-induced hole trap precursors were generated but they had fully recovered and hence were not manifested in the latter “ramp-down” phase. We used a relaxation period of 1×10^3 s; it is inevitably arguable whether this is long enough for full recovery to happen, but the non-negligible increase in the relatively permanent part of $|\Delta V_{th}|$ suggest that recovery is incomplete. In imposing full recovery of the switching hole traps, one must necessarily presume a different defect generation mechanism for the permanent $|\Delta V_{th}|$ increase, which is speculated to be interface trap generation. This inference is supported by temperature dependence study [18], which showed that the activation energy for the permanent $|\Delta V_{th}|$ increase was different from that of the cyclic $|\Delta V_{th}|$. An alternative explanation is that no new switching hole traps were generated, i.e. the increase of the cyclic $|\Delta V_{th}|$ in the ramp-up phase was due to the activation of a greater number of *preexisting* charge-compensated switching hole traps (Fig. 4.11; state 3 $\rightarrow$ state 2).

If we revisit Fig. 4.12(b), the observation of a gradually increasing $|\Delta V_{th}^{EoR}|$ with ensuing dynamic NBTI cycles indicated that positive traps are being generated. However, in a separate experiment which monitored the gate leakage current (I_g), revealed an insignificant change in gate leakage current after 20 cycles of dynamic NBTI cycles. This strongly implied the nonexistence of any new permanent oxide traps or hole traps precursors generation throughout 30 dynamic NBTI cycle at low field stressing (> 9 MV/cm). This suggest that the defects responsible for the relative permanent shift must be situated not in the oxide bulk but very close to, or at the Si/SiO₂ interface. Although this study did not give any direct evidence on the nature

of these defects, it was widely speculated to be caused by the Si-H bond dissociation [18], which will be further discussed in Section 4.2.7. But nevertheless, the results clearly show that a substantial part of the interface degradation under NBTI stress does not correlate with oxide trap precursors generation.

Figure 4.22: (a) $|\Delta V_{th}|$ recovery per cycle or $|\delta V_{th}^R|$ as a function of the number of DNBTS or CVS-DNBTS cycles. The inset shows the evolution of $|\Delta V_{th}|$ in a typical DNBTS cycle. (b) Corresponding gate leakage current after 20 DNBTS or 19 CVS-DNBTS cycles. CVS: $V_g^{\text{stress}} = 2.7$ V and time = 50 s.

While a constant cyclic $|\Delta V_{th}|$, $|\delta V_{th}^R|$ was consistently observed on our devices for the stress conditions studied, it was not always the case for other devices or stress conditions [83-85]. Gao *et al.* [83] reported a decrease in $|\delta V_{th}^R|$ for HfO₂ gate p-MOSFET, whereas HfSiON and SiON counterparts exhibited constant $|\delta V_{th}^R|$ under similar stress conditions. Grasser *et al.* [85] also observed a decrease in the $|\delta V_{th}^R|$ of the non-nitrided SiO₂ gate p-MOSFET at higher temperatures ($\sim 200^\circ\text{C}$). A TDDS

study on a SiON device at elevated temperatures revealed a halt in the hole emission event at a given defect for an extended period [84].

In the following experiments, it can be shown that a similar decrease in $|\delta V_{th}^R|$ can be observed for SiON gate p-MOSFET stressed under a larger oxide field or at extended stress period (without interruptions) at lower field. Test devices were p+ polysilicon gate p-MOSFETs with 1.7 nm oxynitride (SiON) gate dielectrics prepared by decoupled plasma nitridation. To examine the correlation between ΔN_{ot} and NBTI, the following test sequence was adopted. A pristine p-MOSFET was first stressed at a gate voltage $V_g^{stress} = -1.8$ V (~ 8.5 MV/cm oxide field) and relaxed at $V_g^{relax} = 0$ V. The stress and relaxation phases each lasted 1×10^3 s. The dynamic NBTI (DNBTI) test was followed by a constant voltage stress (CVS) at a larger V_g^{stress} magnitude for a given time. No measurement was made during this time. The DNBTI test was then resumed to examine the impact of the prior CVS on the $|\Delta V_{th}|$ evolution. The CVS-DNBTI sequence was repeated numerous times and V_g^{stress} transitions between the two were made on-the-fly. In a separate experiment, the DNBTI test was replaced by gate current I_g measurement to gauge the extent of ΔN_{ot} increase after CVS. For all experiments, the temperature was set at 100 °C.

In the case of constant voltage stress – dynamic NBTI stress (CVS–DNBTS), a substantial increase in I_g was evident after 20 cycles (Fig. 4.22(b)). Since no increase in I_g was observed for the case of DNBTS, the observed increased I_g or oxide trap generation is attributed to CVS. This observation tend to show that the oxide trap generation observed by some studies [86, 87] is the effect of a high oxide field and is

not related to NBTI [88]. However, the reduction in $|\delta V_{th}^R|$ (Fig. 4.22(a)) warrant further examination. Fig. 4.23(a) shows that the decrease in $|\delta V_{th}^R|$ tracks the fractional increase in I_g closely; the two exhibited an excellent linear correlation (Fig. 4.23(b)). This means that, with the occurrence of oxide trap generation by CVS, trapped-hole relaxation in the subsequent DNBTS cycle would be reduced. It should be emphasized that the decrease in $|\delta V_{th}^R|$ is unexpected since one would expect hole trapping and thus, detrapping under DNBTS would be expected to increase instead of decrease with oxide trap generation. A more substantial decrease in $|\delta V_{th}^R|$ is observed with a larger V_g^{stress} of CVS (Fig. 4.24(a)). Fig. 3(b) compares the evolution of $|\Delta V_{th}^{EoS}|$ and $|\Delta V_{th}^{EoR}|$, the respective $|\Delta V_{th}|$ at the end of stress and relaxation phases for CVS–DNBTS and DNBTS, as a function of the number of cycles. Interestingly, when a more significant $|\delta V_{th}^R|$ decrease is observed in the late stage of the CVS–DNBTS (Fig. 4.24(a)), the $|\Delta V_{th}^{EoS}|$ is seen to increase at a rate comparable to that of the DNBTS. This means that, in both cases, the total positive trapped charge ($|\Delta V_{th}^{EoS}|$ probes both oxide and interface trapped charges) is increasing at a similar rate to the late stage. However, the $|\Delta V_{th}^{EoR}|$ of the CVS–DNBTS is observed to rise faster compared to $|\Delta V_{th}^{EoS}|$. This observation implies that, in the DNBTS cycle following the CVS, the proportion of more permanent trapped holes is increased, i.e., the prior CVS had transformed a fraction of the recoverable trapped holes into a more permanent form. The correlated increase in I_g (Fig. 4.23) implies that the underlying oxide traps also stem from this transformation mechanism. On the other hand, $|\Delta V_{th}^{EoS}|$ and $|\Delta V_{th}^{EoR}|$ for the case of DNBTS both increase at a similar rate, yielding a constant $|\delta V_{th}^R|$ in the absence of oxide trap generation (Fig. 4.22(b)). Fig. 4 shows that, with a smaller V_g^{stress} of CVS, apparent reduction of $|\delta V_{th}^R|$ could only be

observed at a much longer CVS time. This implies that transformation of the trapped holes into permanent oxide traps could occur under DC stress after a sufficiently long time. However, under AC stress, this transformation is inhibited by a nearly complete relaxation of the trapped holes (cf. constant $|\delta V_{th}^R|$; Fig. 4.22(a)). It should be mentioned that the first few applications of CVS always produced an evident “jump” in the $|\Delta V_{th}^{EoR}|$, relative to that of the first DNBTs cycle (cf. Fig. 4.24(b)). The increase is often more significant than the increase of I_g or the decrease in R . As shown in Fig. 2(a), both ΔI_g and change in $|\delta V_{th}^R|$ are approximately zero, whereas the $|\Delta V_{th}^{EoR}|$ at the second DNBTs cycle (i.e., after the first CVS) is increased by 20 mV (not shown). The latter corresponds to a non-negligible effective trap density of $2.5 \times 10^{11} \text{ cm}^{-2}$. A similar observation applies to the case of the 3.1V CVS (Fig. 4.24(b)); $|\Delta V_{th}^{EoR}|$ was increased by 50 mV, whereas the $|\delta V_{th}^R|$ was reduced by only 10 mV after the first CVS. This discrepancy requires us to presume that the defects responsible for the $|\Delta V_{th}^{EoR}|$ increase are situated at the Si/SiO₂ interface, implying that part of the interface degradation is not correlated with oxide trap generation. The oxygen vacancy defect or the E' center [89] is a major source of trapped-hole sites in the gate oxide. A recent work employing on-the-fly ESR measurement showed that NBTI recovery was correlated to the E' center [90]. On this basis, we provide a possible explanation for the transformation of trapped holes into more permanent oxide traps, notwithstanding that other defect types may also account for the observation. *Ab initio* simulation studies have shown that the E'_δ center accounts for more than 90% of the total E' center in SiO₂ [76]. When a hole is captured at the E'_δ center, the change in the distance between two Si atoms is marginal, i.e., there is marginal structural relaxation and thus, there is a high tendency for the Si–Si bond to

reform. Hence, the trapped-hole site is annihilated when the V_g^{stress} is removed [76, 91]. On the other hand, a high oxide field can facilitate the transformation, via distortion of the oxide network and/or enhanced interaction between energetic holes and the lattice of the E'_δ center into the E'_γ center. The latter is a relatively permanent form involving significant structural relaxation [76, 91]. As a result, the tendency for the trapped-hole site to revert to the Si–Si dimer precursor state is reduced. These more permanent trapped-hole sites, in turn, increase $|\Delta V_{\text{th}}^{\text{EoR}}|$ and contribute an additional gate leakage current. It should be noted that the categorization of trapped holes as E'_δ or E'_γ center is broad; one should expect the structural recovery time of a given trapped hole site to be determined by the local atomic arrangement and random interaction with phonons. Thus, “intra-category” transformations and considerable spread of recovery time in each category are likely scenarios.

Figure 4.23: (a) Fractional increase in the gate tunneling current $\Delta I_g/I_{go}$ (at $V_g = 1.5$ V) or stress-induced leakage current (SILC) is observed to “track” the decrease in $|\delta V_{th}^R|$ closely. (b) Excellent linear correlation between $\Delta I_g/I_{go}$ (extracted at various voltages) and change in $|\delta V_{th}^R|$.

Figure 4.24: (a) With a larger V_g of CVS [cf. Fig. 4.22(a)], a more substantial decrease of $|\delta V_{th}^R|$ is observed. (b) Evolution of $|\Delta V_{th}|$ at the end of the stress cycle ($|\Delta V_{th}^{EoS}|$) and at the end of the relaxation cycle ($|\Delta V_{th}^{EoR}|$) as a function of the number of DNBTI or CVS-DNBTI cycles, for the same devices in (a).

4.2.5 Observation of Interface Traps Mechanism in NBTI

In the previous discussion on the UFS experimental result in comparison with the understanding of R-D model, it can clearly be shown that interface state generation and passivation based on hydrogen transport kinetics, described in the context of R-D model, cannot be used solely to explain NBTI. Instead, the observation of two distinct activation energies (~ 45 meV and 110 meV) through Arrhenius experiments shed light that supported the notion of two different degradation mechanisms explaining NBTI [14, 46]. One of the mechanisms is hole trapping/detrapping, which has been previously discussed with the UFS experimental results. This hole trapping/detrapping mechanism was also reported in numerous studies as well [14, 15], which attributed the rapid “degradation transient” observed in dynamic NBTI to the positive charging and recovery of pre-existing oxide trap (also known as E' center). The other

mechanism is the interface state generation (N_{it}), which is still postulated by some that the generation of N_{it} is governed by physics of R-D model [7-10]. The traditional charge pumping (CP) [62] requires stress interruption and thus entails a considerably long measurement delay. It was deemed earlier as inadequate for the NBTI characterization in the view of the possible interface state repassivation based on the understanding of R-D model [23]. To address this issue, the developed UFS-CP methodology which is a modification of our ultrafast switching method [18], was used. At a given interval, a series of positive 100-ns gate pulses of amplitude $|V_g^{\text{stress}}| + 1$ V are injected, at periods of 10 μ s each, at the gate bias-tee. The pulses switched the DUT between inversion and accumulation repeatedly (Fig. 4.25), inducing a dc surface recombination current measured with sub-pA resolution, at the source/drain terminal. Through this UFS-CP methodology, new insights on the understanding and characteristics of N_{it} generation via charge pump current (I_{cp}) results will be presented. Test devices (DUT) were p^+ polysilicon gate p-MOSFETs with 2.8 nm SiO_2 gate dielectrics. Channel length and width was 120nm and 10 μ m, respectively. Gate stress voltage V_g^{stress} ranged from -2.6 to -1.8 V; gate voltage (V_g) at recovery was -1 V and the temperature was 100°C . An ultrafast CP method (Fig. 4.25), which is a modification of our ultrafast switching method [18], was used. At a given interval, a series of positive 100-ns gate pulses of amplitude $|V_g^{\text{stress}}| + 1$ V was injected, at periods of 10 μ s each, at the gate bias-tee. The pulses switched the DUT between inversion and accumulation repeatedly (Fig. 3), inducing a dc surface recombination current measured, with sub-pA resolution, at the source/drain terminal.

Figure 4.25: (left) Schematic illustration of the ultra-fast switching-charge pumping (UFS-CP) measurement setup, adapted from ultra-fast switching method. (right) Schematic illustration of the gate waveforms during dynamic NBTI. For pre-stress and recovery, the DUT is biased at -1 V. The gate bias is switched to V_g^{stress} during stress. The top of the gate pulse for all cases is $+1$ V, except stated otherwise. V_{inv} (-0.3 V) and V_{acc} ($+0.5$ V) refer to the threshold voltage and flat band voltage, respectively of charge pumping. The pulse period was fixed at $10\text{ }\mu\text{s}$ and the pulse width ranged from 100 - 900 ns; the corresponding duty cycle ranged from 1 - 9% .

Figure 4.26: Increase in the charge pumping current (ΔI_{cp}), measured by an ultrafast 100 -ns pulse, as a function of stress time. Gate stress voltages are as indicated. Power-law dependence of ΔI_{cp} on stress time is evident and the exponent n ranges from 0.145 - 0.171 . This observation has been regarded as an experimental proof for a H_2 diffusion-limited N_{it} generation mechanism, which has a theoretical exponent n of 0.167 .

4.2.6 Cyclic Behavior of ΔI_{cp}

Similar dynamic NBTI experiments were carried out in comparison with the simulated results based on the understanding of the R-D model. It has been highlighted that the selection of test device used is p-MOSFET with dielectric thickness of 2.8nm. Channel length and width are 120nm and 10 μ m, respectively. This is to suppress the high tunneling leakage current at stress voltage, which may potentially corrupt the measurement results. All UFS-CP experiments are conducted at 100°C with measurements sampled at drain/source. Fig. 4.27 depicts the evolution of ΔI_{cp} under dynamic NBTI. Abrupt increase/decrease in ΔI_{cp} , upon transition from recovery-to-stress/stress-to-recovery is apparent. This behavior was ascribed to spontaneous dissociation of Si-H bonds by holes/re-passivation of dangling Si bonds by interfacial hydrogen species, i.e. the reaction regime of the R-D model. Following the abrupt changes, the relatively slow evolution of ΔI_{cp} was believed to be determined by the relatively slow hydrogen diffusion process. To facilitate further explanation, three parameters are defined (Fig. 4.27): (i) ΔI_{cp}^{EoS} (ΔI_{cp}^{EoR}) denotes cumulative ΔI_{cp} at the end of the stress (recovery) cycle; (ii) the amount of ΔI_{cp} recovered per cycle, δI_{cp}^r ($= \Delta I_{cp}^{EoS} - \Delta I_{cp}^{EoR}$)

Figure 4.27: UFS-CP experimental results showing the evolution of ΔI_{cp} under dynamic NBTI stress and relaxation for 30 cycles at $V_g^{stress} = -2.6V$ and $V_{relax} = -1V$ for similar time duration of 1000s at $100^\circ C$. Illustration showing definition of (i) δI_{cp}^R denotes the change in I_{cp} shift at the end of a particular recovery phases; (ii) ΔI_{cp}^{EOS} (ΔI_{cp}^{EOR}) denotes the cumulative I_{cp} after a given number of stress (recovery) phases.

In the context of the R-D model, it stipulated that δI_{cp}^R would decrease with increased dynamic stress/recovery (S/R) cycling, since the amount of hydrogen diffusing away from the Si/SiO₂ interface increases with the number of S/R cycles (Fig. 4.1(b)). However, these principles of the R-D model were not borne out experimentally. Fig. 4.27 shows the evolution of ΔI_{cp}^{EOS} , ΔI_{cp}^{EOR} and δI_{cp}^R as a function of dynamic S/R cycles. It is highlighted that δI_{cp}^R observed remains relatively constant and is independent of S/R cycling. This experimental finding contradicts the understanding of the R-D model.

Figure 4.28: (a) Evolution of ΔI_{cp}^{EoS} , ΔI_{cp}^{EoR} and δI_{cp}^R , as a function of stress/recovery cycles for the device depicted in Fig. 4.27. The amount of I_{cp} recovered per cycle, δI_{cp}^R is observed to be consistent throughout the dynamic NBTI cycles, (b) Normalized δN_{it}^R or δI_{cp}^R recovered per cycle with respect to 1st cycle based on R-D model simulation and experimental is compared.

To further affirm the experimental observation against the validity of the R-D model, the time ratio of stress to recovery interval (t_{stress}/t_{relax}) is varied. Similar dynamic NBTI experiments were repeated for 14 dynamic cycles with similar stress (relax) conditions at $-2.6V$ ($-1V$) and t_{stress} (t_{relax}) were varied over 2 orders of magnitude from 100 s to 10000 s. If the physics of the R-D model is true, longer t_{stress} (i.e. 10000 s) would have postulated more hydrogen to diffuse away from interface as compared to shorter t_{stress} (i.e. 1000 s). As a result, a lower concentration of interfacial hydrogen would be available to passivate the interface traps, corresponding to lesser δI_{cp}^R during recovery phase. On the contrary, a device subjected to shorter t_{stress} would have a higher concentration of interfacial hydrogen which will correspond to more δI_{cp}^R during recovery phase. Fig. 4.29 plots the extracted parameters ΔI_{cp}^{EoS} , ΔI_{cp}^{EoR} and δI_{cp}^R of the last cycle for the given time ratio (t_{stress}/t_{relax}) varied over a wide range (10

to 0.1). According to the R-D model, one would expect a corresponding significant change in ΔI_{cp}^{EoR} , since the amount of ΔN_{it} relaxation varies with time ratio [29]. This inference is, however, not borne out by the result depicted in Fig. 4.29.

Figure 4.29: δI_{cp}^R (circle) is observed to be independent of $t_{\text{stress}}/t_{\text{relax}}$. This result does not corroborate a transport-limited mechanism of N_{it} generation and recovery as stipulated by the R-D model. It can also be observed that ΔI_{cp}^{EoR} (open triangle) is insensitive to t_{relax} and appears to be only determined by t_{stress} . This implies that a component of stress induced ΔI_{cp} is relatively permanent, i.e. irrecoverable.

One salient observation noticed in Fig. 4.29, is that the magnitude of ΔI_{cp}^{EoS} is higher as t_{stress} increases from 1000 s to 10000 s. However, for a given scenario when t_{stress} is 5000 s, the magnitude of ΔI_{cp}^{EoR} for different t_{relax} of 100 s and 1000 s are similar. Recalling the understanding of R-D model, for a given t_{relax} of 100 s, the hydrogen which is located away from the interface would have a limited time to back-diffuse and passivate with the N_{it} . And in the experiment, this t_{relax} is extended to 1000 s

whereby more hydrogen would be expected to back-diffuse to interface and result in more N_{it} being passivated. However, this experimental observation of a consistent δI_{cp}^R implies that the amount of N_{it} recovery is not enhanced by the extended t_{relax} and this finding does not corroborate with the transport-limited NBTI mechanism. This conclusion is similar for rest of the time ratio cases. Therefore, it is concluded that the R-D model is still not able to explain the N_{it} generation and passivation in the two mechanisms NBTI framework proposed.

4.2.7 Negligible Recovery of Interface Traps

The next fundamental question that one may ask, is what governs the stress induced N_{it} when the R-D model fails to explain the N_{it} generation in NBTI. With hindsight from the UFS-CP experiment, some key characteristics of this stress induced N_{it} mechanism were identified. Recalling in Fig. 4.29, it is highlighted that the amount of ΔI_{cp}^{EoR} remains consistent for a fixed t_{stress} (i.e. 1000 s and 5000 s) and the amount of recovery δI_{cp}^R remains consistent despite t_{relax} being varied from 1000 s to 10000 s in the case of $t_{stress} = 1000$ s. These observations can be portrayed as two distinct components in UFS-CP measurement data; one is always fully recoverable (observed as δI_{cp}^R) while the other is permanent (observed as ΔI_{cp}^{EoR}).

To further probe the issue on ΔN_{it} relaxation, ΔI_{cp}^{EoR} at the end of a given *cumulative* stress time of 2×10^4 s is examined as a function of time ratio. It should be noted that because the time ratio is different in each case, the device was stressed and relaxed for different number of cycles to achieve the given cumulative stress time. For instance,

for $t_{\text{stress}}/t_{\text{relax}} = 5000/100$ (both numbers in seconds), four S/R cycles were needed and the cumulative relaxation time was 400 s. On the other hand, for $t_{\text{stress}}/t_{\text{relax}} = 1000/10000$, twenty S/R cycles were needed and the corresponding cumulative relaxation time was 2×10^5 s, 500 times longer than the former.

From Fig. 4.30(a), it is clearly shown that for a given *cumulative* stress time, $\Delta I_{\text{cp}}^{\text{EoR}}$ was observed to be comparable, regardless of the cumulative recovery time and recovery period examined. This implies that stress induced ΔN_{it} does not recover or recovers extremely slowly such that it may be regarded as constant even during quasi-static measurement. To affirm this proposition, a device was initially stressed and left to recover in a ambient condition for a long term scale of up to 50 days as shown in Fig. 4.30(b). It was clearly shown that the amount of stress induced ΔI_{cp} after the short term recovery deviates from the long term recovery by $\sim 1.5\text{pA}$ which is considered negligible, as this corresponds to a very minor change in ΔN_{it} ($\sim 0.6 \times 10^{10} \text{ cm}^{-2}$). Therefore, it is affirmed that the permanent component of stress induced ΔN_{it} is relatively permanent (i.e. irreversible process) and does not recover significantly given an extremely long recovery period. From the above inference, it is deduced that the permanent component of N_{it} is likely to be dependent on t_{stress} only, and it does not recover given prolonged t_{relax} .

Figure 4.30: (a) Identical ΔI_{cp}^{EoR} (open triangle) is observed at a given cumulative stress time (2×10^4 s), regardless of the cumulative recovery time which ranges from 400 to 2×10^5 s. This shows that a component of ΔI_{cp} which depends only on the stress time but not on the recovery time, i.e. irrecoverable. $V_g^{stress} = -2.6V$; temperature was $100^\circ C$. (b) ΔI_{cp}^{EoR} observed on a long term scale for up to 50 days

If we recall our earlier discussion on UFS experimental results, we postulate about a possible relative permanent component generated, amidst the hole trapping/detrapping component to explain NBTI. Similar conclusion was also observed from UFS experiments, where dynamic NBTI experiments were carried out at a fixed cumulative stress time of 3×10^4 s for various time ratios (Fig. 4.31). To achieve that, device for time ratio = 30 (given $t_{stress}/t_{relax} = 3000$ s/100 s) would undergo ten dynamic cycles to achieve the fixed cumulative stress time of 3×10^4 s but only clocking a cumulative relaxation time of 1000 s. On the other hand, device for time ratio = 0.33 (given $t_{stress}/t_{relax} = 1000$ s / 3000 s) would undergo thirty cycles to achieve the fixed cumulative stress time of 3×10^4 s but clock a *longer* cumulative relaxation time of 9×10^4 s. In spite of very different total “clock” times (due to a widely varying

t_{relax}), the striking identical $|\Delta V_{\text{th}}^{\text{EoR}}|$ (open triangle) across all time ratio affirms the non-recovery of N_{it} generation in UFS data.

Figure 4.31: $|\delta V_{\text{th}}^{\text{S}}|$, $|\delta V_{\text{th}}^{\text{R}}|$, $|\Delta V_{\text{th}}^{\text{EoS}}|$, and $|\Delta V_{\text{th}}^{\text{EoR}}|$ as a function of the time ratio of stress to recovery interval, $t_{\text{stress}}/t_{\text{relax}}$. The triangle denotes $|\Delta V_{\text{th}}^{\text{EoS}}|$ and $|\Delta V_{\text{th}}^{\text{EoR}}|$ extracted after a given cumulative stress time (3×10^4 s). Stress intervals of 1000s undergo 30 dynamic cycles while stress intervals of 3000 s undergo 10 dynamic cycles in order to achieve the defined cumulative stress time. As a result, the cumulative recovery time is varied for all cases.

After understanding the permanent stress induced ΔN_{it} , the recoverable component of ΔN_{it} was investigated. It is consistently observed that the recoverable component $\delta I_{\text{cp}}^{\text{R}}$ is always superimposed onto the permanent $\Delta I_{\text{cp}}^{\text{EoR}}$ during stress and is always recovered fully within the shortest recovery interval of 100 s during relaxation. Longer recovery interval of up to 1×10^4 s does not stipulate further ΔI_{cp} recovery which implies that this recovery process is a fast process whereby most of the

recoverable component is spontaneously recovered upon removal of stress. Further investigations through Arrhenius experiments (Fig. 4.32) have revealed a rather small activation energy (~ 60 meV), suggesting that this recoverable N_{it} component δI_{cp}^R is a rather temperature insensitive process and impossible to be explained by the hydrogen transport mechanism. Instead, *Ab initio* simulation has shown that H^+ near the Si interface could behave as a hole trap [10] which may be linked to hole trapping/detrapping mechanism as observed in similar UFS experimental results. On the contrary, stronger thermal activation (~ 144 meV) of the permanent N_{it} component ΔI_{cp}^{EoR} shows that a larger energy barrier must be overcome to create this component of I_{cp} shift. The larger energy barrier is probably needed to initiate some local bond breakage at the interface states (i.e. Si-H bonds), thus the immunity of the resultant defects to recovery.

Figure 4.32: Distinct activation energies of δI_{cp}^R and ΔI_{cp}^{EoR} imply that two different mechanisms are responsible for the measured ΔI_{cp} . This finding shows that ΔI_{cp} is made up of two different components.

Figure 4.33: Change in the charge pumping current, ΔI_{cp} at three different stages of dynamic NBTI as a function of the gate pulse width used during measurement. The dashed circle highlights the comparison among the three parameters at timing comparable to state-of-the-art integrated circuits.

One is able to readily observe the same behavior between δI_{cp}^R and $|\delta V_{th}^R|$ or $|\Delta V_{th}^h|$ (Fig. 4.28(b)). In addition, the amount of ΔI_{cp} increased per stress cycle, δI_{cp}^S is also seen to become almost equal to δI_{cp}^R after a few S/R cycles. These characteristics of δI_{cp}^S and δI_{cp}^R suggests that they also arise from the same switching hole traps responsible for the cyclic behavior of $|\Delta V_{th}^h|$. To probe further, pulse width dependence of δI_{cp}^S (or δI_{cp}^R) is examined for a given stress condition in Fig. 4.33. It has been reported that the amount of charge recombined per gate pulse cycle (I_{cp}/f) is increased when either the period of the inversion or the accumulation cycle is increased [92]. This increase in I_{cp}/f is ascribed to hole traps located further from the Si/SiO₂ interface being able to take part in the recombination process, since a longer pulse width would mean greater probability for electrons/holes to tunnel in/out of

traps deeper into the gate oxide [92]. Spatial profiling of oxide trap density based on this principle has been demonstrated [92]. In Fig. 4.33, ΔI_{cp} at three different points in a given S/R cycle — end of stress (EoS); ~ 300 ms after termination of stress or beginning of recovery (BoR), and 1000 s after termination of stress or end of recovery (EoR) — are examined as a function of pulse width (duration when the gate is at +1 V). ΔI_{cp}^{EoS} is observed to increase with pulse width whereas ΔI_{cp}^{BoR} and ΔI_{cp}^{EoR} are independent of pulse width. This observation strongly implies that ΔI_{cp} measured on-the-fly, i.e. by switching the gate from V_g^{stress} to +1 V, incorporates a hole trap recombination current (ΔI_h) in addition to the interface trap recombination current (ΔI_{it}). The absence of pulse width dependence for the case of ΔI_{cp}^{BoR} and ΔI_{cp}^{EoR} (i.e. after changing the gate from V_g^{stress} to -1 V) suggests that the ΔI_h observed does not simply arise from the “filling” of preexisting oxide traps but from hole traps induced on-the-fly by the V_g^{stress} applied. On the other hand, if the evolution of ΔI_{cp} under dynamic stress were to be determined by a reaction-transport mechanism (i.e. the R-D model), an increase in pulse width should have led to a smaller ΔI_{cp} measured during stress (due to greater recovery caused by a longer delay) and a correspondingly smaller step-like change at S/R transition points. However, this inference is clearly not borne out by the experimental results depicted in Fig. 4.33.

Fig. 4.33 clearly shows that ΔI_h reduces at short pulse widths. Its value at pulse timing typical of circuit operation ($\sim 10^{-10}$ s) may be estimated as shown. One may argue the outcome does not offer much practical value, since in actual circuits, devices would not be operated in the exact manner they were stressed and measured in our experiments. Nonetheless, the much reduced “gap” between ΔI_{cp}^{EoS} and ΔI_{cp}^{BoR}

indicates that (i) a large part of the abrupt change in ΔI_{cp} observed at S/R transition points arises from a non-negligible ΔI_h at relatively long pulse widths and (ii) interface states do not exhibit the fast passivation or generation stipulated by the reaction phase of the R-D model. In conjunction with the very small difference between ΔI_{cp}^{BoR} and ΔI_{cp}^{EoR} (measured 1000s apart), the result further indicates that there is no substantial relaxation of stress induced interface states.

4.3 Summary

A major intrinsic limitation of the reaction-diffusion model for negative-bias temperature instability is revealed through dynamic stress experiments. We found no evidence of self-limiting recovery, one of the key features of the transport-based R-D model, after repeating the stress and relaxation cycles alternately for many times. UFS experimental results revealed that the amount of $|\Delta V_{th}|$ induced in a given fixed stress interval is almost completely removed in a following similar recovery interval. This cyclic behavior, observed consistently over many stress/recovery cycles, does not conform to the current R-D model notion that interface state generation, rate-limited by hydrogen transport, is responsible for the V_{th} degradation/recovery typically observed. A hydrogen transport model requires the amount of recovery to decrease progressively with stress/recovery cycling, which is not observed.

Further insights through Arrhenius experiment revealed two distinct degradation mechanisms, evident by a small and large activation energy (~ 45 meV and ~ 110 meV). The relatively small activation energy of 45 meV suggests a cyclical $|\Delta V_{th}|$ component arises from a hole trapping/detrapping mechanism (via tunneling). The

stronger thermal activation of 0.11 eV is perceived as a more energy-intensive process involving the breaking of Si-H bonds for creation of N_{it} .

The hole trapping/detrapping phenomenon, perceived to have resulted in the rapid cyclical $|\Delta V_{th}|$ evolution in dynamic NBTI, may be ascribed to the well known oxygen vacancy defect (E' centers) and its variants (E'_δ centers) in the SiO_2 gate dielectric. As evident from the first principal calculation, the shallow energy state of E'_δ center is calculated to be 0.5 eV to 1 eV above the SiO_2 valance band, which favors the condition for a rapid charge exchange and short tunneling constant in a ultra-thin gate oxide. However, one must also consider the wide spread of the defect energy state and tunneling time constant in an amorphous nature gate oxide. Further evidence from a positive relaxation gate voltage revealed another important class of hole traps, classified as deep level hole trap (DLHT). A DLHT is ascribed to a defect located very close to the interface and as a consequence, their energy states are “pinned” by the Si/ SiO_2 band discontinuity. Due to the higher (deeper) energy state, a charged DLHT requires a much longer emission time (which could take up days) for it to be discharged. However, the recovery of DHLT can be accelerated by a positive gate bias.

Through a combination of dynamic NBTI testing at high and low voltages, experimental results show that the number of hole traps that can be activated under a given stress condition is fixed, although this number can vary with the exact stress voltage applied. Even when the applied stress resulted in an increase of stress induced leakage current (SILC) or bulk oxide trap density, the number of hole traps activated

under dynamic NBTI testing did not increase but exhibited a decrease instead. Analysis of $|\Delta V_{th}|$ measured immediately upon the termination of stress and after a substantial recovery shows interestingly a faster increase of the latter. As the $|\Delta V_{th}|$ measured at the end of stress probes the total positive oxide trapped charge while that measured after relaxation only probes the unrecovered portion of the positive oxide trapped charge, the results show that the decrease in the number of hole traps activated during dynamic NBTI test is a consequence of a portion of the hole traps being transformed into a more permanent form. This transformation is shown to always correspond to the increase of SILC, i.e. when the transformation is absent, so is the SILC increase, implying that the transformation is responsible for the generation of more permanent bulk traps that are responsible for the observed SILC. The transformation is consistent with the irreversible structural change that may occur at an E' center upon the capture of a hole.

Further presumption that the R-D model could still be used to explain the N_{it} generation in the two-mechanism NBTI model was also ruled out. UFS-CP measurement of dynamic NBTI shows that the amount of ΔI_{cp} induced in a given stress interval is completely removed in a following recovery interval, which disagrees with the notion of a hydrogen transport model that depicts the amount of recovery to decrease progressively with stress/recovery cycling. Instead, it was found that stress induced ΔN_{it} comprises of two distinct characteristics coupled together; (i) one is a consistent fast recoverable N_{it} component which is spontaneously recovered upon termination of stress (ii) while the other is the permanent N_{it} component. The fast recoverable N_{it} component, identified as a parasitic signal which originates from

the N_{ot} was inevitably captured in the UFS-CP measurement. The other permanent component, identified as the generation of N_{it} , is shown to be an irreversible process, whereby the generation of N_{it} is solely dependent on stress period and the generated defects show no apparent recovery even after an extremely long recovery period (up to 50 days).

5 SEPARATION OF OXIDE TRAP AND INTERFACE TRAP IN UFS AND UFS-CP MEASUREMENT METHODOLOGY

5.1 Separation of Parasitic Oxide Trap Signal in UFS-CP

Recent developments in NBTI studies have suggested that the fast transient behavior ($< 1\text{s}$) of threshold voltage shift ($|\Delta V_{\text{th}}|$) is attributed to hole trapping, (N_{ot}) (believed to be caused by pre-existing oxide traps) while the generation of Si/SiO₂ interface states (N_{it}), (rate-limited by hydrogen diffusion in the context of R-D model) is considered to be responsible for the subsequent slow increase in $|\Delta V_{\text{th}}|$. However, the fast transient $|\Delta V_{\text{th}}|$ has been seen as a parasitic hole trapping component as it will saturate with time. Instead, it is the generation of N_{it} which will increase with time and determines the NBTI lifetime. Therefore, a thorough understanding of N_{it} generation is crucial for NBTI.

5.1.1 Introduction

In Chapter 4.2.5, we discussed the observation of an abrupt change in ΔI_{cp} , typically at the start of a stress or recovery cycle in the as-measured UFS-CP data. This is attributed to the charge interaction between the fast responsive oxide trap with the channel, reacting in sync with the charge pumping pulse waveform. Consequently, it

became an inevitable measurement artifact near the Si/SiO₂ interfacial which is superimposed with the interface state charge pumping current (I_{cp}). As this is inherent to the UFS-CP measurement, the presence of the parasitic oxide trap signal will distort the result of the interface state generation signal. In order to understand the evolution of interface state generation during stress, it is essential to separate the parasitic oxide trap signal from the as-measured ΔI_{cp} data. Initial studies have shown that a small portion of oxide traps are able to respond within 100 ns, thus making it practically challenging for separation to take place. Nevertheless, these measurement artifacts must be eliminated in order to truly understand N_{it} behavior and the co-development of a reliable model. In this section, a systematic approach to separate the parasitic oxide trap signal based on the phenomenological cyclic nature is presented.

5.1.2 Experimental

Test devices (DUT) were p⁺ polysilicon gate p-MOSFETs with 2.8 nm SiO₂ gate dielectrics. Channel length and width are 120 nm and 10 μ m, respectively. Gate stress voltage V_g^{stress} ranged from -2.6 to -1.8 V; gate voltage (V_g) at recovery was -1 V and the temperature was 100 °C. An ultrafast CP method, which is a modification of our ultrafast switching method [18], was used. At a given interval, a series of positive 100-ns gate pulses of amplitude $|V_g^{stress}| + 1$ V was injected, at periods of 10 μ s each, at the gate bias-tee. The pulses switched the DUT between inversion and accumulation repeatedly, inducing a dc surface recombination current, measured with sub-pA resolution at the source/drain terminal.

5.1.3 Separation of Oxide Trap and Interface Traps in UFS-CP

Ultra-Fast Charge Pumping (UFS-CP) built upon charge pumping (CP) theory has been traditionally thought to yield interface states (N_{it}) signals only by pulsing the gate between stress and accumulation. But latest findings have suggested possibilities of an additional oxide trap component (N_{ot}) sampled by I_{cp} , which have corrupts the extraction of N_{it} from present UFS-CP results [18].

The initial observation on the change in the ΔI_{cp} under static NBTI stress which to a first order, is commonly believed to be a direct manifestation of ΔN_{it} , and is deemed key supporting evidence: (i) stress induced increase in ΔI_{cp} exhibits a power-law exponent $n \sim 0.145 - 0.171$ (Fig. 5.1) [16]. (ii) ΔI_{cp} exhibits a step-like change followed by quasi-saturation under dynamic NBTI. This seems to agree well with the R-D model, explained by a H_2 diffusion kinetics which will exhibit a power-law exponent of $n = 0.167$. However, this presumption based on static NBTI was misinterpreted, as the presence of N_{ot} component was not taken into consideration in as-measured I_{cp} measurement.

The existence of N_{ot} contribution in UFS-CP was first observed from dynamic NBTI experiment. It was shown that the amount of ΔI_{cp} recovered per cycle, $\delta I_{cp}^R (= \Delta I_{cp}^{EoS} - \Delta I_{cp}^{EoR})$ is constant, independent of stress/recovery (S/R) cycling. This exhibits strikingly similar striking cyclic characteristics to the earlier observation of hole trap made on UFS experimental data (Fig. 4.4).

Figure. 5.1: (a) Increase in the charge pumping current (ΔI_{cp}), measured by an ultrafast 100-ns pulse, as a function of stress time. Gate stress voltages are as indicated. Power-law dependence of ΔI_{cp} on stress time is evident and the exponent n ranges from 0.145-0.171. This observation has been regarded as an experimental proof for a H_2 diffusion-limited N_{it} generation mechanism, which has a theoretical exponent n of 0.167. (b) Evolution of ΔI_{cp} under dynamic NBTI (with equal stress and recovery intervals). ΔI_{cp}^{EoS} (ΔI_{cp}^{EoR}) denotes ΔI_{cp} at the end of the stress (recovery) cycle. The typical 2-stage ΔI_{cp} evolution, i.e. abrupt initial change followed by quasi-saturation is believed to correspond to the reaction and diffusion limited phases of the R-D model.

Further findings by constant base UFS-CP (Fig. 5.2) have revealed an increased I_{cp} with pulse width (from 100 ns to 900 ns), even when the DUT is pulsed from V_g^{stress} (-2.6V) to a level V_g^T (-0.6 V) in a strong inversion regime. This phenomenon contradicts with the understanding of CP theory as the charge pumping voltage (V_{cp}) is pulsed within strong inversion. CP theory [93] stipulates that when $V_g^T < V_{th}$, occupancy of N_{it} is unchanged and surface electron density is negligible, no electron-hole recombination is possible at the interface states. A possible explanation for this phenomenon is that oxide band bending is lowered, thus the occupancy of hole trap is changed, causing a hole trap to lose its electronic charges to the channel and thus contributing an additional component to I_{cp} (Fig. 5.3).

Figure 5.2: Additional I_{cp} can be observed even by pulsing the gate for a longer duration between V_g^{stress} and $V_g^T < V_{th}$, i.e. always maintaining the device in strong inversion. In this regime, no electron-hole recombination can occur at interface states since their occupancy is not changed and the surface electron concentration is negligible. However, occupancy of hole traps in the oxide bulk would change in tandem with oxide band bending, modulated by the gate pulse (Fig. 5.3). Thus, the additional I_{cp} is due to “charge pumping” of hole traps.

Figure 5.3: Schematic illustration of change in oxide band bending when the gate is pulsed from a negative stress voltage (V_g^{stress}) to a less negative level $V_g^T (< V_{th})$, threshold voltage), during constant base charge pumping. (a) During stress ($V_g^T = V_g^s$), interface states and interfacial hole traps quasi Fermi energy (dotted line) are charged by holes from the inversion layer. (b) When V_g is pulsed to $V_g^T (< V_{th})$, interface trap occupancy does not change and surface electron concentration is negligible (surface

potential is “pinned”). Hence, no net electron-hole recombination occurs at interface states. However, hole traps “falling” below the hole quasi Fermi energy are neutralized via electronic charge exchange with the gate and/or substrate valence band (dashed lines depicts oxide band bending at $V_g = V_g^{\text{stress}}$). A longer pulse width results in more hole traps being neutralized, and this is in turn registered as an increased supply of holes from the drain/source to recharge the traps when V_g is pulsed back to V_g^{stress} . When majority of hole traps are neutralized, further increasing the pulse width yields little ΔI_{cp} increase, giving an impression that ΔI_{cp} is independent on pulse timing [16]. This in turn leads to an incorrect inference of negligible hole trap contribution to the measured I_{cp} .

Further evidence to ascertain the role of oxide trap contribution in UFS-CP is revealed by varying the pulse width in the similar sampling frequency (Fig. 5.4). According to the CP theory [93], the charge pumping current is represented by $I_{\text{cp}} = q.A.f.N_{\text{it}}$. (where $q = 1.602 \times 10^{-19}\text{C}$, A is the area of the device, f is the frequency of the charge pumping and N_{it} is the interface trap density). It stipulates that I_{cp} is independent of pulse duty cycle for a given frequency and N_{it} . However, in the following dynamic NBTI experiment where the duty cycle is varied by increasing the pulse width for a given fixed sampling frequency, the ΔI_{cp} measured during stress actually becomes larger when the CP pulse width is increased from 100 ns to 900 ns (Fig. 5.4). A larger ΔI_{cp} measured during stress which is detected at longer pulse width (Fig. 5.4) does not conform to CP theory. On the other hand, the increase in CP pulse timing has no apparent effect on ΔI_{cp} during recovery.

Figure 5.4: (a) ΔI_{cp} measured during stress increases with CP pulse timing (or duty cycle). ΔI_{cp} under recovery is, however, is not affected by the change in duty cycle. The latter conforms to the CP theory [93], which stipulates that I_{cp} is independent of pulse duty cycle for a given N_{it} density. The former implies contribution of oxide traps at longer pulse timing. (b) Amount of ΔI_{cp} recovered per cycle, δI_{cp}^R is increased at a longer CP pulse timing, but it remains independent of stress/recovery cycling. The increase is due to more hole traps contributing to ΔI_{cp} measured during stress (as confirmed in Fig. 5.1 and explained in Fig. 5.2). The resultant more abrupt ΔI_{cp} decrease at the onset of recovery would potentially lead to a wrong conclusion of fast ΔN_{it} repassivation.

Since the dependence of ΔI_{cp} on pulse width “disappears” right at the onset of recovery, the result implies that the abrupt decrease in ΔI_{cp} is an artifact due to charge pumping of oxide traps during stress but not during recovery. These oxide traps are annihilated (Fig. 5.4(a)), hence ΔI_{cp} becomes independent of the pulse duty cycle after stress is removed. Therefore, the abrupt decrease in ΔI_{cp} cannot be regarded as an experimental proof for spontaneous repassivation of ΔN_{it} . Identical ΔI_{cp}^{EoR} observed over a wide range of recovery period (Fig. 4.29 and Fig. 4.30) implies that stress induced ΔN_{it} does not recover or recovers at such an extremely slow rate that it may be regarded as constant even during quasi-static measurement. Any increase in I_{cp} with pulse duty cycle has been ascribed to oxide traps [94].

A possible hypothesis based on the understanding of N_{ot} is formulated to explain the phenomenon as illustrated in Fig. 5.5. During gate pulsing at stress, pulse width of 100ns have resulted in a portion of this fast responsive positive N_{ot} located near the interface to discharge concurrently with N_{it} to drain; generating an additional I_{ot} component which is expected to increase with longer pulse width/discharge time. But when at relaxation phase, these N_{ot} is a neutral state and does not contribute to I_{cp} . Hence, it explains the comparable ΔI_{cp} regardless of pulse width variation. Therefore, the above hypothesis suggests that I_{ot} is present at stress phase ($I_{cp}^{stress} = I_{it} + I_{ot}$) which can be determined by subtracting it with relaxation ($I_{cp} = I_{it}$) phase.

Figure 5.5: (a) Illustration showing combination of charged N_{ot} and N_{it} contributing to I_{cp} during stress. During charge pumping measurement, the gate pulse will allow charged N_{ot} near the interface to discharge, thus generating a parasitic component, I_{ot} which superimposes with I_{it} . Increasing the gate pulse from 100ns to 900ns will allow more time for charged N_{ot} located deeper into the gate oxide to discharge, which explains the increased I_{cp} observed during stress. (b) Illustration showing absence of charged N_{ot} during relaxation. Only contribution from N_{it} is recorded in I_{cp} measurement. Increasing the gate pulse will have no impact on I_{it} as stipulated by CP theory [93].

Since ΔI_{cp} is contaminated by the parasitic oxide-trap signal (ΔI_{ot}), it must be corrected before the true exponent n for ΔI_{it} (CP current due to ΔN_{it}) can be revealed. To separate them, we first estimated ΔI_{ot} from δI_{cp}^R , derived from dynamic NBTI (Fig. 5.4(b)), based on the earlier observation that ΔN_{it} does not recover (Fig. 4.30(b)). Hence, δI_{cp}^R is due to charge pumping of oxide traps during stress. Subtracting ΔI_{ot} from ΔI_{cp} gives an exponent n of ~ 0.27 for ΔI_{it} (or ΔN_{it}) (Fig. 5.6(a)). Consistent exponents for ΔN_{it} ($n \sim 0.27 - 0.28$) are obtained when the separation method was performed at different V_g^{stress} (Fig. 5.6(b)).

Figure 5.6: (a) Power-law exponent n for ΔN_{it} generation is shown to be ~ 0.28 , after subtracting a constant hole trap component (ΔI_{ot}) from the measured ΔI_{cp} . ΔI_{ot} is estimated from δI_{cp}^R obtained by dynamic NBTI (Fig. 5.4(b)). ΔI_{it} denotes the component of ΔI_{cp} due to ΔN_{it} . (b) Comparable exponents of ~ 0.27 are observed over a wide V_g^{stress} range. Thus, an exponent of ~ 0.16 cannot be regarded as an experimental proof of H_2 diffusion kinetics since the measured ΔI_{cp} is “contaminated” by a hole-trap signal. The relatively permanent nature of ΔN_{it} (Fig. 4.30(b)), however, does not suggest that ΔN_{it} arises from atomic hydrogen diffusion kinetics [7].

5.1.4 Temperature-Dependent of Power-Law Exponent on Interface Trap

In a recent work [16], the exponent n of the ΔI_{cp} drift curve was observed to increase with temperature. This observation was deemed the experimental proof for dispersive hydrogen transport [30]. It was argued that hydrogen transport across an amorphous oxide is frequently “disrupted” by the random capture and emission of hydrogen species by oxide traps. Dispersive transport is particularly prominent at low temperature. The transport mechanism is expected to approach classical diffusion as temperature increases (i.e. increasingly energetic hydrogen species are less prone to capture by oxide traps). An increase in exponent n towards 0.16 (the signature of H_2

diffusion) with temperature (see Fig. 12(b) of ref. [16]) appeared to support the dispersive transport model.

We also observe from our experiments that the exponent n of the ΔI_{cp} drift curve is increased with temperature (Fig. 5.7(a)). However, in view that as-measured ΔI_{cp} comprises a parasitic oxide trap current ΔI_{ot} , the exponent n of ΔI_{it} , i.e. the component of ΔI_{cp} due to ΔN_{it} can only be unambiguously examined after removal of ΔI_{ot} . Based on the observation that ΔI_{ot} is approximately constant for t_{stress} greater than 1000 s (cf. Fig. 4.30(b)), it was first determined from dynamic NBTI experiments performed at different temperatures and then subtracted from the ΔI_{cp} drift curve for stress time greater than 1000 s; an example is illustrated in Fig. 5.6(a).

Figure 5.7: (a) Exponent n of the as-measured ΔI_{cp} increases with temperature. This may be explained by the superposition of two mechanisms, i.e. ΔI_{ot} and ΔI_{it} , the latter having a stronger temperature dependence than the other (Fig. 5.8) (b) Power-law exponent n of ΔI_{it} is consistently $\sim 0.26 - 0.27$ over a wide temperature range, after subtracting the oxide-trap component ΔI_{ot} from the measured ΔI_{cp} .

Figure 5.8: (a) Exponent n of the as-measured ΔI_{cp} increases with temperature before and after decoupling. This may be explained by the superposition of two mechanisms, i.e. ΔI_{ot} and ΔI_{it} , the latter having a stronger temperature dependence than the other (b) Distinct activation energies of δI_{cp}^R (ΔI_{ot}) and ΔI_{cp}^{EoR} (ΔI_{it}) imply that two different mechanisms are responsible for the measured ΔI_{cp} . This finding corroborates the results in Fig. 4.32, showing that ΔI_{cp} is made up of two different components.

Evolution of ΔI_{it} at different temperatures is depicted in Fig. 5.7(b). Interestingly, after having eliminated ΔI_{ot} from the as-measured ΔI_{cp} , the resultant exponent n for ΔI_{it} is independent of temperature. This result implies that variation of the exponent n of as-measured ΔI_{cp} with temperature is a consequence of the superposition of two mechanisms, i.e. hole trapping and ΔN_{it} having different thermal activation. A comparison of the activation energy of ΔI_{ot} and ΔI_{it} clearly reveals different thermal activation of the two components (Fig. 5.8(b)). A larger activation energy of ΔI_{it} means that the relative contribution of ΔI_{it} to ΔI_{cp} increases more significantly with temperature, thus inducing an increase in exponent n of ΔI_{cp} from ~ 0.1 towards ~ 0.18 (Fig. 5.8(a)). The results strongly imply that temperature dependence of exponent n is

an experimental artifact arising from the superposition of two mechanisms having different thermal activations.

5.1.5 Summary

New experimental evidence (from ultrafast CP) which does not conform to the R-D model is presented. It is shown that (i) amount of ΔI_{cp} recovery per cycle is independent of stress/recovery cycling; (ii) the abrupt change in ΔI_{cp} typically observed at the transition between stress/recovery is an artifact caused by charge pumping of oxide traps during stress but not during recovery; (iii) the power-law time exponent of ~ 0.16 for ΔI_{cp} is due to ΔI_{cp} being contaminated by ΔI_{ot} (after eliminating ΔI_{ot} , the actual exponent for ΔN_{it} is ~ 0.27); (v) the power-law time exponent of ΔN_{it} is independent of temperature. In view of these new evidences, previous claims of consistency between generation/recovery of ΔN_{it} and the R-D model or its dispersive counterpart must be reviewed.

5.2 Separation of Hole Trapping and Interface-State Generation by UFS

Hole-trapping and interface-state components of NBTI induced threshold voltage shift are separated via ultrafast switching measurement (UFS). Based on the phenomenological observation that dynamic NBTI is determined by a cyclic hole trapping/detrapping mechanism and that interface state generation is relatively permanent, the time dependence of hole trapping during stress is precisely determined and then subtracted from the overall degradation to yield the time dependence of interface state generation. Interface state generation is observed to exhibit power-law time dependence with an initial exponent of ~ 0.5 , which subsequently decreases to a steady value of ~ 0.25 after $\sim 1 \times 10^3$ s at the stress condition studied (oxide field ~ 9 MV/cm). This evolution is shown to be consistent with that obtained independently via the UFS-CP method discussed in the previous section.

5.2.1 Introduction

Experimental results have shown that hole-trap (ΔN_h) and interface state (ΔN_{it}) generation are responsible for negative-bias temperature instability (NBTI) [81, 95]. Ascertaining the impact of individual defects on the evolution of transistor degradation is of paramount importance to the understanding and modeling of NBTI. Earlier controversy on the recovery mechanism had cast doubt on the validity of conventional methods, such as charge pumping (CP), for ΔN_{it} measurement. An ultrafast charge pumping (UFS-CP) method was then proposed for ΔN_{it} measurement [16] but an oxide-trap current still has to be corrected before information on ΔN_{it} can be obtained [18].

However, the charge pumping method only provides a fraction of ΔN_{it} since a portion of the Si bandgap is scanned and the application on very thin gate devices (~ 1.5 nm) is limited by the large gate tunneling leakage. Moreover, the CP method does not provide comprehensive information on ΔN_h — a deficiency that is particularly critical in light of recent evidence showing prolonged evolution of hole trapping in an ultrathin gate oxide. Some authors have assumed rapid saturation of ΔN_h and subtracted the threshold voltage shift ($|\Delta V_{th}|$) measured at ~ 1 s from the overall degradation to determine ΔN_{it} [14]. However, it is earlier understood that hole-trap generation can continue for prolonged periods (~ 1000 s) even in an ultrathin gate oxide [81] – this phenomenon has been explained by multiphonon-field-assisted tunneling [79]. In this light, (i) the time dependence of ΔN_{it} determined by assuming a constant ΔN_h is not accurate; (ii) precise separation of ΔN_h and ΔN_{it} as a function of time is difficult since both are continuously being generated. We demonstrate, for the *first time*, rigorous separation of ΔN_h and ΔN_{it} from the as-measured $|\Delta V_{th}|$ in UFS. The approach is based upon the insight that dynamic NBTI is solely determined by repetitive hole trapping/detrapping [70] and that ΔN_{it} is permanent [18, 81]. Measurement of hole-trapping kinetics is achieved by repeatedly “cycling” the device between stress and relaxation (under a positive gate voltage) without an increase of ΔN_{it} , and the data used subsequently to delineate the evolution of ΔN_h and ΔN_{it} in the first stress cycle.

5.2.2 Experimental

Test devices were p^+ polysilicon gate p-MOSFETs with drawn channel length and width of 120 nm and 2 μm , respectively, from a 65 nm CMOS technology. The oxynitride (SiON) gate oxide was prepared by decoupled plasma nitridation and the nitrogen concentration at the gate/SiON interface was ~ 13 at. %. Devices were stressed at an oxide field of ~ 9 MV/cm and a temperature of 100 $^{\circ}\text{C}$. Recovery was performed at a 0 V gate voltage (V_g), unless stated otherwise, at the same temperature. Device degradation was monitored in terms of ΔV_{th} , extracted via a constant sub-threshold drain current ($I_d = 20 \mu\text{A}$) method from I_d - V_g curves measured by an UFS methodology. I_d - V_g measurement was achieved, at selected stress intervals, by pulsing the gate from stress to individual measurement levels for a very short period of 100 ns. The short delay improves the measurement precision of the hole trapping effect, known to exhibit very rapid relaxation [10].

5.2.3 Separation of Hole Trapping and Interface Trap

The methodology for separating ΔN_h and ΔN_{it} from $|\Delta V_{th}|$ is based on the insight that dynamic NBTI is solely determined by a cyclic hole-trap mechanism (Refer to Section 4.2.2) and the steady increase of $|\Delta V_{th}|$ is a consequence of the relatively permanent ΔN_{it} (Refer to Section 4.2.7). In Fig. 5.9, the change in $|\Delta V_{th}|$ after each stress and recovery cycle, $|\delta V_{th}^S|$ and $|\delta V_{th}^R|$ respectively, as well as $|\Delta V_{th}|$ at the end of each recovery period ($|\Delta V_{th}^{EoR}|$) (inset) are plotted as a function of the number of

stress/recovery (S/R) cycles, N . As discussed in Section 4.2.2, a constant $|\delta V_{th}^R|$, independent of N , does not conform to a hydrogen transport model. In conjunction with an equal $|\delta V_{th}^S|$ (for $N > 10$), the observations indicate that dynamic NBTI is governed by a cyclic hole-trap mechanism, superimposed on a progressively increasing $|\Delta V_{th}^{EoR}|$ due to the relatively permanent ΔN_{it} (and some contribution of deep-level hole traps). This inference is further supported by the relatively small activation energy of 40 meV for $|\delta V_{th}^S|$ (or $|\delta V_{th}^R|$) (Fig. 5.10(b)), which implies hole tunneling as the underlying mechanism [81].

Figure 5.9: Amount of threshold voltage shift during the stress period ($|\delta V_{th}^S|$), the recovery period ($|\delta V_{th}^R|$), and the threshold voltage shift at the end of the recovery period ($|\Delta V_{th}^{EoR}|$) as a function of the number of stress/recovery (S/R) cycles (N). Inset shows evolution of $|\Delta V_{th}|$ during stress and recovery.

Figure 5.10: (a) Evolution of $|\Delta V_{th}^h|$ during stress and recovery, showing highly repetitive hole trapping and detrapping characteristic, respectively, over many stress/recovery cycles (thin lines; $N = 10$ to 30). Symbols denote average characteristics. For clarity, only average characteristics are shown for the 1.5 and 1.8 V recovery voltage. (b) Arrhenius relations for $|\delta V_{th}^S|$ (or $|\delta V_{th}^R|$) and $|\Delta V_{th}^{EoR}|$, showing a much smaller activation energy of the former. Inset shows $\sim T^{1.3}$ temperature dependence of $|\delta V_{th}^S|$.

Trapping/detrapping of holes over extended period is apparent in Fig. 5.10(a) – $|\Delta V_{th}^h|$ continues to evolve logarithmically with time by more than 20% and 10% during stress and recovery (0 V) respectively, from 10 to 1000 s. This observation clearly shows that the classical assumption of saturation of hole trapping at ~ 1 s [14] is inaccurate. The large dispersion in hole detrapping time constant could be ascribed to a relatively broad energy distribution of hole traps [81]. Since $|\Delta V_{th}^h|$ recovery (0 V) shows no sign of saturation up to 1000 s, it implies that the interval is not long enough for complete hole detrapping, i.e. there are remnant holes trapped at deep energy levels which require considerably long periods to detrapp (Refer to Section 4.2.3). To ascertain the extent of contribution to $|\Delta V_{th}|$ from slow deep-level hole traps (DLHTs), a gate voltage of +1.5 V (instead of 0 V) was applied during recovery to neutralize trapped holes by substrate electron injection. Initial recovery of $|\Delta V_{th}^h|$ is clearly accelerated in this case (filled triangles in Fig. 5.10(a)). The subsequent almost “flat” recovery characteristics ($\sim 2\%$ change in $|\Delta V_{th}^h|$ over the 1000-s period) implies near complete recovery of $|\Delta V_{th}^h|$. Further increasing the recovery voltage by 20% (1.5 to 1.8 V) yielded only $\sim 1.5\%$ increase in $|\Delta V_{th}^h|$ recovery (dashed line; Fig. 5.10(a)), implying that majority of DLHTs were detrapped. As will be shown later, it is essential for DLHTs to be removed as failure to do so would lead to a severe underestimation of hole trapping kinetics (apparent from a comparison of stress induced $|\Delta V_{th}^h|$ denoted by the open circle (for 0 V recovery) and filled circle (for +1.5 V recovery)). Evolution of ΔN_{it} , extracted based on the underestimated hole trapping data, would then be distorted.

The fact that $|\delta V_{th}^S|$ becomes almost equal to $|\delta V_{th}^R|$ after repeated S/R cycling (Fig. 5.10(a); arrow A) provides a means for determining the time dependence of hole trapping (Fig. 5.10(a); filled circle), with negligible contribution from ΔN_{it} in each stress cycle. The average stress induced $|\Delta V_{th}^h|$ characteristic is redrawn in Fig. 5.11(a) on a log-log scale. Power-law time dependence is evident, with an exponent of ~ 0.07 [10]. Also shown is the evolution of $|\Delta V_{th}|$ for the first stress cycle (open circle). A larger exponent of 0.112 for the latter may be attributed to non-negligible generation of interface states when stress is first applied. The difference between the open and filled circles yields the evolution of $|\Delta V_{th}|$ due to interface states ($|\Delta V_{th}^{it}|$), which also exhibits power-law time dependence within the interval (1000 s) probed, but with a much larger exponent of 0.493. Although the larger exponent of 0.5 was not reported in NBTI literature previously, the similar time exponent was reported by Ang *et. al* when the p-MOSFET was stressed with a positive gate bias [96]. Under the positive stress condition, N_{it} generation in absence of hole trapping was postulated. Thus, supporting the notion for the larger exponent observed for $|\Delta V_{th}^{it}|$ generation. On the other hand, if majority of DLHTs are not eliminated (0 V recovery), the resultant exponent of $|\Delta V_{th}^{it}| + |\Delta V_{th}^{DLHT}|$ becomes smaller (~ 0.392). Thus, a misguided inference on the evolution of ΔN_{it} would arise if the effect of DLHTs is not properly accounted for [97].

It should be mentioned that an alternative means which first measures $|\Delta V_{th}|$ ($= |\Delta V_{th}^h| + |\Delta V_{th}^{it}|$) and then subjects the device to relaxation at a positive gate voltage before a second $|\Delta V_{th}|$ measurement ($= |\Delta V_{th}^{it}|$) would not yield correct data on ΔN_h evolution due

to relaxation and the subsequent non-negligible period to “rebuild” hole trapping to the level before stress interruption (Fig. 5.10(a)). In contrast, our approach relies on the phenomenological behavior of cyclic hole trapping in dynamic NBTI for independent determination of hole trapping kinetics. The data is then used to delineate the contribution of $|\Delta V_{th}^h|$ to $|\Delta V_{th}|$ measured under minimal delay in the first stress cycle.

To further validate the principal assumption of our approach, i.e. hole trapping is cyclic and ΔN_{it} is permanent, we compare the evolution of $|\Delta V_{th}^{it}|$ to that calculated from ΔN_{it} measured by the CP method (after oxide trap current correction [18]) over a wide range of stress time in Fig. 5.11(b). Excellent agreement is apparent. To obtain $|\Delta V_{th}^{it}|$ by our method, the stress cycle was extended to 3×10^4 s but the recovery (+1.5 V) period was kept at 1000s (on basis that $|\Delta V_{th}^h|$ recovery is almost complete; Fig. 5.10(a)). The device was cycled between stress and recovery for 15 times to ascertain the evolution of $|\Delta V_{th}^h|$ in later cycles. On the other hand, CP current measurement was achieved within a fraction of a second [18]). Despite the much longer recovery in the former, $|\Delta V_{th}^{it}|$ is higher and exhibits nearly identical evolution as that obtained through the CP method. This observation confirms the permanence of ΔN_{it} . The smaller $|\Delta V_{th}^{it}|$ obtained through the CP method is principally due to a limited interval of the Si bandgap scanned.

Subsequent decrease in the exponent of $|\Delta V_{th}^{it}|$ from ~ 0.5 to a steady value of 0.25 is apparent. Similar evolution has been observed [67, 98]. On the other hand, a smaller exponent (~ 0.16) for “ $|\Delta V_{th}^{it}|$ ” is obtained when one assumes saturation of hole trapping at ~ 1 s (Fig. 3 of ref.[14]; Fig. 1(a) of ref. [70]), i.e. by neglecting further hole trapping.

For the stress condition studied here (~ 9 MV/cm), a steady value of the exponent is attained only after ~ 1000 s. Moreover, the “transition” from a high to low exponent occurs at a relatively small $|\Delta V_{th}^{it}| \sim 10$ mV. This implies that accurate assessment of the impact of ΔN_{it} on V_{th} shift must necessarily entail substantial time, due to subsequent decrease in the exponent.

Figure 5.11: (a) The average stress induced $|\Delta V_{th}^h|$ characteristic of Fig. 5.10(a) is redrawn on a log-log plot. As-measured $|\Delta V_{th}|$ of the first stress cycle is also shown. Difference between $|\Delta V_{th}|$ and $|\Delta V_{th}^h|$ yields $|\Delta V_{th}^{it}|$ (the component of $|\Delta V_{th}|$ due to interface state generation). Also shown is the evolution of $|\Delta V_{th}^{it}| + |\Delta V_{th}^{DLHT}|$ for the case of 0 V recovery. (b) Evolution of $|\Delta V_{th}^{it}|$ (circle), as extracted by the approach

outlined in Fig. 5.11(a). Also shown is $|\Delta V_{th}^{it}|$ based on ΔN_{it} measured by the CP method.

5.2.4 Summary

The components of $|\Delta V_{th}|$ due to hole trapping and interface state generation are consistently extracted from the as-measured $|\Delta V_{th}|$. Based on the phenomenological observation that dynamic NBTI is determined by a cyclic hole trapping/detrapping mechanism, the time dependence of hole trapping is reliably determined from repeated stress/recovery cycling and then the hole trapping component is subtracted from the overall $|\Delta V_{th}|$ of the first stress cycle. The underlying principle of the approach is verified by the excellent agreement between the extracted $|\Delta V_{th}^{it}|$ and that obtained from the CP method.

5.3 Impact of Hole Trapping Studied on a Seasoned Device

A “seasoned” device, stressed at prolong period (1×10^5 s) to wither out interface trap generation is demonstrated. Interface trap generation is shown to exhibit a power law exponent of ~ 0.25 after $\sim 1 \times 10^3$ s at the stress condition studied (oxide field ~ 10 MV/cm), but slowly evolves to quasi-constant power law exponent of ~ 0.084 when stress is extended beyond 1×10^5 s. Based on the earlier inference that interface traps is relatively permanent, any further degradation to device shows negligible increase/decrease in ΔI_{cp} , reaching a quasi-saturation stage. In conjunction with observation on hole trapping in dynamic NBTI, prolonged stressing at moderate gate voltage will not affect spontaneous charging nature of hole trap. In this light, an unique characterization on hole trapping in the absence of interface trap generation is demonstrated for the first time. Hole Trapping is shown to exhibit a power law time exponent of ~ 0.05 and an activation energy of 60 meV. This evolution is shown to be consistent with that obtained independently via the separation method discussed in the previous section.

5.3.1 Introduction

Owing to its potential impact on circuit operation, intense effort to understand and model the threshold voltage (V_{th}) fluctuations caused by dynamic NBTI has been made. As confirmed in the earlier discussion, V_{th} fluctuations are predominantly a hole trapping/detrapping process caused by the electronic charging/discharging of E' center, as evident from the consistent cyclic $|\Delta V_{th}|$ evolution in dynamic NBTI.

Owing to the wide distribution of energy states of E' center in the amorphous gate oxide, the ability to charge up more hole-traps (N_h) will increase with stress voltage due to oxide band lowering. (Refer to Section 4.2.4). However, characterization of hole trapping is previously determined from dynamic NBTI, superimposed with a background of relative permanent interface trap (N_{it}) generation. Ascertaining the impact of N_h solely on the evolution of transistor degradation is practically challenging since both N_h and N_{it} are continuously being generated. The critical question is: how could one ascertain the relative importance of hole trapping, especially at device operating condition, in relation to the relatively permanent interface state generation in NBTI. We demonstrate, for the *first time*, a sole characterization of ΔN_h in the absence of N_{it} , from the as-measured $|\Delta V_{th}|$ in UFS. The approach is based on a “seasoned” device where device is stressed for a long term (1×10^5 s) at a moderate stress condition (oxide field ~ 10 MV/cm), rendering further generation of N_{it} to be negligible. Based on the insight that ΔN_{it} is permanent [18, 81] and N_h remains unaffected by prolonged moderate stressing, measurement of hole-trapping kinetics can be extracted from the as-measured ΔV_{th} by UFS.

5.3.2 Experimental

Test devices were p^+ polysilicon gate p-MOSFETs employing 1.6 and 2.4 nm plasma-nitrided SiO_2 gate dielectrics. Stressing was carried out at 100 °C unless stated otherwise. An ultrafast switching method, based on the pulsed current-voltage setup, was used for the measurement of the linear I_d - V_d curve [25]. The extremely narrow (100 ns) and short duty cycle (0.1 %) measurement pulses ensured that recovery effect

was minimized. V_{th} shift ($|\Delta V_{th}|$) was extracted by a constant sub-threshold drain current (20 μA) method.

5.3.3 Preparation of a “Seasoned” Device

The methodology for a sole characterization of ΔN_h in the absence of N_{it} generation, from the as-measured $|\Delta V_{th}|$ in UFS, is demonstrated on a “seasoned” device. This “seasoned” device was prepared based on the insight that ΔN_{it} is permanent and prolonged stress period is to allow a substantial N_{it} generation, rendering the device degradation to a stage whereby further N_{it} generation is considered negligible. In conjunction with the earlier inference that N_h remains unaffected by prolonged moderate stressing, measurement of hole-trapping kinetics can be extracted from the as-measured $|\Delta V_{th}|$ by UFS.

Figure 5.12: A device stressed at a moderately negative V_g^{stress} for an extended period beyond 1×10^5 s. Power law exponent of ΔI_{cp} evolves from ~ 0.266 in earlier stage of

stress and ~ 0.084 in the latter stage of stress. Demonstrating the quasi-saturation of N_{it} . (inset) Generation rate of ΔI_{cp} decreases progressively with stress time, rendering further generation of N_{it} to be negligible

The method for preparing a “seasoned” device is demonstrated by stressing the device for a prolonged stress period ($\sim 1 \times 10^5$ s) at a moderate constant gate stress (oxide field ~ 10 MV/cm). Note that the selection of gate stress plays a critical role to prevent any catastrophic failure like dielectric breakdown to happen during the process of preparation. A hindsight that a catastrophic failure had occurred can be detected from I_{cp} in UFS-CP measurement. The gate leakage current will increase exponentially to a few orders higher ($>10^{-8}$ A range) than former before breakdown. This will potentially overwhelm the I_{cp} measurement (in order of $\sim 10^{-12}$ A range), rendering ΔI_{cp} to fluctuate in a disorderly manner. However, this was not observed in the ΔI_{cp} measurement as evident in Fig. 5.12, affirming the device functionality after prolonged stressing.

An attempt to generate substantial N_{it} generation to a quasi-constant was observed at stress time of 1×10^5 s. As shown in Fig. 5.12, ΔI_{cp} is shown to exhibit a initial power law exponent of ~ 0.25 after stressing for $\sim 1 \times 10^3$ s but slowly evolves to quasi-constant power law exponent of ~ 0.084 when stress is extended beyond 1×10^5 s. This small power law exponent ~ 0.084 observed, implies that generation of N_{it} had reached a quasi saturation stage whereby any further degradation would yield a negligible amount of increase in I_{cp} . This is evident in inset of Fig. 5.12 where the generation rate of N_{it} has decreased progressively to a insignificant rate of $< 10^{-17}$ cm $^{-2}$

$^2/s$. In conjunction with the earlier inference that N_{it} is relatively permanent, any further V_{th} degradation or recovery due to N_{it} is regarded as negligible.

Figure 5.13: The “seasoned” device in this case was already stressed at a moderately negative V_g^{stress} for some time and deep level hole traps fully recovered using a positive gate voltage. The same “seasoned” device was then stressed again at a less negative V_g^{stress} to reactivate hole trapping for 8×10^3 s. A full recovery under a positive gate voltage indicates that only N_h is reactivated and there is no further generation of ΔN_{it} . This allows a precise examination of the time dependence of N_h at various less negative V_g^{stress} , independent of the influence of ΔN_{it} .

Although prolonged stress at moderate stress voltage has exhausted further N_{it} generation on the “seasoned” device, the ΔN_h which is solely determine by the cyclic hole-trap mechanism in dynamic NBTI, remains unaffected by the moderate stress. This was evident by the consistent cyclic component, R which displayed repetitive

behavior in dynamic NBTI stress (Fig. 4.6). In addition, experimental observation by inducing a combination of higher and lower stress voltage also yielded a consistent cyclic component. Affirming the immunity of spontaneous charging nature of N_h to moderate stress. To validate the assumption that only N_h is reactivated and no further generation of ΔN_{it} occurs at the less negative V_g^{stress} , the device was stressed for 80×10^3 s and a full recovery under a positive V_g^{stress} was observed which confirms this assumption. On this basis, measurement of hole-trapping kinetics can be extracted from the as-measured ΔV_{th} by UFS.

5.3.4 Impact of Hole Trapping/Detrapping (HTD)

An approach is outlined to study the V_g^{stress} dependence of hole trapping independent from that of ΔN_{it} is done on a “seasoned” device. To prepare the sample, a fresh device was first stressed at a moderately negative V_g^{stress} , to generate substantial ΔN_{it} such as subsequent reactivation of hole trapping at less negative V_g^{stress} yields no further generation of ΔN_{it} . This can be achieved on the device sample when stressed beyond 1×10^5 s at moderate stress condition (oxide field ~ 10 MV/cm). The remnants of any deep level hole trap (DLHT) components were then fully recovered with the aid of a positive V_g . Completing the preparation of a “seasoned” device. The “seasoned” device was subsequently used to characterize the evolution of hole trapping at less negative V_g^{stress} , where ΔN_{it} remained essentially negligible.

Using this method, we obtained the time dependence of hole trapping over a broad V_g^{stress} range (Fig. 5.14). Logarithmic evolution, spanning many orders of time

magnitude (i.e. no sign of saturation) is evident. The power law exponent over a broad V_g^{stress} range is shown to be ~ 0.05 which clearly shows that the classical assumption of saturation of hole trapping at ~ 1 s [14] is inaccurate. The large spread in the hole trapping time constant of an ultrathin gate oxide has been explained by MPFAT [25].

Figure 5.14: Evolution of hole trapping during stress, extracted from the as measured $|V_{\text{th}}|$ in UFS on a “seasoned” device (a) For various gate stress voltages down to as low as -1.4 V; Logarithmic evolution, spanning many orders of time magnitude (i.e. no sign of saturation) is evident (b) Power law time exponent of ~ 0.05 across various gate stress shows a clear indication that classical assumption of saturation of hole trapping at ~ 1 s is inaccurate.

An Arrhenius experiment was conducted with temperature ranging from $50^\circ\text{C} \sim 150^\circ\text{C}$ in Fig. 5.15. The small activation energy ~ 60 meV is in agreement with the earlier activation energy of ~ 45 meV reported from the cyclical $|\Delta V_{\text{th}}|$ component arising from dynamic NBTI (Fig. 4.8). However, there is still a small difference observed, attributed to the minority presence of deeper state hole traps which require a longer time to build up but may not be significantly observed in earlier reported cyclical $|\Delta V_{\text{th}}|$ components in dynamic NBTI which has a stress period of 1×10^3 s.

Owing to the wide distribution of defect states of E' center in the amorphous gate oxide, the process of charging up deeper level hole traps require a longer time constant and a more energetic charge which accounts for the slightly higher activation energy observed. Unlike shallow level hole traps which can be fast responsive and have a small tunneling constant in a ultra-thin gate oxide.

Figure 5.15: (a) Evolution of hole trapping across various temperature, extracted from the as measured $|V_{th}|$ in UFS on a “seasoned” device (b) Arrhenius plot showing a relatively small activation energy of ~ 60 meV for HTD process.

In a separate experiment, the HTD data at a given V_g^{stress} was subtracted from the overall $|\Delta V_{th}|$ which is measured on another identical fresh device to obtain the corresponding evolution of ΔN_{it} . An example is depicted in Fig. 5.16. The extracted $|\Delta V_{th}|$ due to ΔN_{it} exhibits a power-law exponent of 0.25, in excellent agreement with that determined independently using an ultrafast charge pumping method, with appropriate oxide trap current correction as shown in Fig. 5.6.

Figure 5.16: (a) An illustration of how the contribution of ΔN_{it} towards the as-measured $|\Delta V_{th}|$ may be readily extracted based on the time dependence data for hole trapping determined in Fig. 5.14. (b) The extracted $|\Delta V_{th}|$ due to ΔN_{it} exhibits a power-law exponent of 0.25, in excellent agreement with that determined independently using an ultrafast charge pumping method, with appropriate oxide trap current correction as shown in Fig. 5.6.

HTD and ΔN_{it} , extracted using this method are compared in Fig. 5.17(a) as a function of V_g^{stress} at a given stress time. For the devices studied, HTD remains predominant over the entire range of V_g^{stress} investigated, even down to a level close to the operating voltage, whereas ΔN_{it} practically vanishes as V_g^{stress} is reduced. The relative importance of HTD over ΔN_{it} is observed to increase as V_g^{stress} is scaled down for the devices studied. (Fig. 5.18). HTD and ΔN_{it} extracted by assuming rapid saturation (in ~ 1 s) are compared to those extracted using the method outlined in this work in Fig. 5.17(b). Gross underestimation of HTD based on the “1-s” assumption is apparent.

Figure 5.17: (a) Gate stress voltage, V_g^{stress} dependence of HDT and ΔN_{it} , extracted after 8×10^4 s stressing at 100°C , using the method outlined in Figs. 8-10). (b) A comparison of HDT extracted in this work to that obtained by a "1-s" assumption (i.e. assuming that hole trapping saturates in ~ 1 s). Significant underestimation of hole trapping using the 1-s assumption is apparent across the entire gate stress voltage range. This in turn overestimates the importance of interface trap generation.

Figure 5.18: For the devices examined, the importance of HTD, relative to ΔN_{it} , is observed to increase as V_g^{stress} scales towards the operating condition.

5.3.5 Summary

An approach for the precise separation of HTD and ΔN_{it} is proposed. Through intentional stressing at prolonged periods on a device, substantial interface traps are generated to a degradation state where further N_{it} generation is considered negligible. In conjunction with the immunity of HTD to degradation, a unique characterization on hole trapping in the absence of interface trap generation is demonstrated for the first time. The method was successfully applied to examine the absolute as well as relative variations in HTD and ΔN_{it} over a broad range of V_g^{stress} for the first time. A series of key experimental characteristics of HTD is presented. Based on these phenomenological characteristics of HTD, a method for separating HTD and ΔN_{it} from as-measured V_{th} drift curves is described and used to investigate the relative change of these two components as a function of V_g^{stress} .

6 CONCLUSION AND RECOMMENDATION

6.1 Conclusion

In this report, the fundamentals of the well-known Reaction-Diffusion (R-D) model used to explain the NBTI phenomenon has been discussed. This model, based on a hydrogen transport mechanism, which describes the evolution of Si/SiO₂ interface states (N_{it}) contributing to NBTI, was widely accepted because it was able to match the evolution of the p-MOSFET parameter drift, e.g., threshold voltage shift $|\Delta V_{th}|$, transconductance degradation $|\Delta g_m|$, etc., which typically exhibited power-law time dependence with a ubiquitous exponent n of 0.25.

However, with the recent development of fast measurement methods, more insights, (e.g. exhibited power-law time exponent 0.167 or smaller), that do not conform to this R-D model is revealed. Many discussions have been sparked regarding the commonly observed instantaneous threshold voltage (V_{th}) shift during stress and relaxation phase. Several believe that this instantaneous V_{th} shift is solely explained by the generation and passivation of N_{it} at the Si/SiO₂ interface based on the framework of R-D model. On the other hand, some think that NBTI has two mechanisms, involving both hole trapping (N_{ot}) and N_{it} generation, which are used to explain the fast and slow $|V_{th}|$ shift.

In this work, a detailed study of dynamic negative bias temperature instability (NBTI) in P-MOSFETs with the ultrathin SiO₂ and SiON gate dielectrics is presented in this thesis. In order to resolve the key challenge in the measurement delay affecting the

characterization of NBTI, the early stage of the research focused on the development of a new ultra-fast switching (UFS) method and its variant ultra-fast charge pumping (UFS-CP). Both UFS and UFS-CP techniques have the capability to characterize ΔV_{th} and ΔN_{it} within 100 ns measurement time and the hardware setup and working principles are presented as well.

The first discussion focuses on the understanding and validation of the R-D model. New insights demonstrated through dynamic NBTI experimental results measured by UFS have revealed a specific fast responsive cyclical $|\Delta V_{th}|$ component which can be generated during stress and fully recovered upon relaxation, regardless of the stress and recovery intervals. This behavior, observed over many stress/recovery cycles does not corroborate with the R-D model, which stipulates that interface state relaxation should decrease progressively with stress/recovery cycling. Further insights through Arrhenius experiment revealed two distinct degradation mechanisms, evident by a small and large activation energy (~ 45 meV and ~ 110 meV). The relatively small activation energy of 45 meV suggests a cyclical $|\Delta V_{th}|$ component arising from a hole trapping/detrapping mechanism (via tunneling). The stronger thermal activation of 0.11 eV is perceived as a more energy-intensive process involving the breaking of Si-H bonds for creation of N_{it} . Detailed testing also revealed different time dependencies for hole trapping and N_{it} generation, implying that these two mechanisms are distinct.

Knowing that there is a physical limitation of the R-D model that cannot explain the dynamic NBTI results, a reassessment of the NBTI degradation mechanism based on a two mechanisms framework was proposed. To establish this framework, intense

research and discussions focusing on these two distinct mechanisms was carried out. The first degradation mechanism, identified as a hole trapping/detrapping phenomenon, was perceived to have resulted in the rapid cyclical $|\Delta V_{th}|$ evolution in dynamic NBTI, ascribed to the well known oxygen vacancy defect (E' centers) and its variants (E'_δ centers) in the SiO_2 or an additional k_N center in SiON gate dielectric. As evident from the first principal calculation, the shallow energy state of E'_δ center is calculated to be 0.5 eV to 1 eV above the SiO_2 valance band, which favors the condition for a rapid charge exchange and short tunneling constant in a ultra-thin gate oxide. However, one must also consider the wide spread of the defect energy state and tunneling time constant in an amorphous-nature gate oxide. Further evidence from a positive relaxation gate voltage revealed another important minority class of hole traps, classified as deep level hole trap (DLHT). A DLHT is ascribed to a defect located very close to the interface and as a consequence, their energy states are “pinned” by the Si/ SiO_2 band discontinuity. Due to the higher (deeper) energy state, a charged DLHT requires a much longer emission time (which could take several days) for it to be discharged. However, the recovery of DHLT can be accelerated by a positive gate bias.

The very origin of this hole traps precursor which led to the subsequent degradation of NBTI remains unclear. Through a combination series of dynamic NBTI testing at high and low voltages, it is understood that the number of hole traps precursors that can be activated under a given stress condition is fixed, although this number can vary with the exact stress voltage applied. Even when a higher oxide field is applied to intentionally create more hole trap precursors, denoted by the increase of stress

induced leakage current (SILC) or bulk oxide trap density, the number of hole traps activated under dynamic NBTI testing did not increase but exhibited a decrease instead. Analysis of $|\Delta V_{th}|$ measured immediately upon the termination of stress and after a substantial recovery interestingly shows a faster increase of the latter. As the $|\Delta V_{th}|$ measured at the end of stress probes the total positive oxide trapped charge, while that measured after relaxation only probes the unrecovered portion of the positive oxide trapped charge, the results show that the decrease in the number of hole traps activated during dynamic NBTI test is a consequence of a portion of the hole traps being transformed into a more permanent form. This transformation is shown to always correspond to the increase of SILC, i.e. when the transformation is absent, so does the SILC increase, implying that the transformation is responsible for the generation of more permanent bulk traps that are responsible for the observed SILC. The transformation is consistent with the irreversible structural change that may occur at an E' center upon the capture of a hole.

In the next stage of discussion, the second degradation mechanism, identified as a interface trap generation, is shown to be an irreversible process. Unlike the R-D model that depicts N_{it} to be recoverable by hydrogen passivation, the generation of N_{it} is solely dependent on stress period and the generated N_{it} defects show no apparent recovery even after an extremely long recovery period (up to 50 days). Earlier observation of the abrupt increase and decrease ΔI_{cp} was found to be a parasitic signal which originated from the N_{ot} and was inevitably captured in the UFS-CP measurement.

In order to establish the time evolution contribution from two distinct degradation mechanisms, a systematic separation approach is outlined to decouple both components to evaluate the impact on NBTI. Based on the phenomenological observation that NBTI is determined by a cyclic hole trapping/detrapping mechanism and that interface-state generation is relatively permanent, the time dependence of hole trapping during stress is precisely determined through dynamic NBTI in UFS measurement, and then subtracted from the overall degradation of the first cycle to yield the time dependence of N_{it} generation. N_{it} generation is shown to exhibit power-law time dependence with an initial exponent of ~ 0.5 , which subsequently decreases to a steady value of 0.25 after ~ 1000 s at the stress condition studied. This evolution is shown to be consistent to that obtained independently via the UFS-CP separation methodology, which separates the parasitic oxide trap signal from the as-measured signal in UFS-CP, yielding a power-law time exponent of ~ 0.27 for ΔN_{it} .

In the last section of discussion, a novel method to solely characterize the impact of hole trapping in the absence of N_{it} generation is demonstrated on a “seasoned” device. By intentionally stressing a device for a prolonged period of time, the N_{it} generation is observed to reach a quasi-constant stage, exhibiting negligible increase in N_{it} generation. It is shown that hole trapping does not saturate but continues to evolve at a power law time exponent of ~ 0.05 which is consistent with the observed ~ 0.07 extracted from the earlier separation approach. The impact of hole trapping is also examined over a broad range of stress voltage and extrapolated to operating bias condition of 1.2V for the first time. By comparing the magnitude of the two degradation mechanisms obtained separately, it is worth mentioning that hole trapping

has a relative importance as compared to interface trap generation at chip operating condition.

6.2 Recommendation

The key purpose for an acceleration test is to accurately estimate the lifetime of the device during real operations within an acceptable testing time. In order to do that, it is essential to establish a robust NBTI model. In this thesis, the fundamental understanding of a two mechanism NBTI framework and the respective characteristics and contributions have been intensively discussed. In closing, the author would like to highlight several recommendations.

As highlighted earlier, NBTI experimental data is very sensitive to the recovery effect which will have a significant impact on extrapolated lifetime. However, due to the lack of a common industrial standard in terms of fast measurement methodologies, the various fast measurement techniques presented till date have their respective measurement delays ranging from 1 μ s to 100 ns. This will inevitably yield a different power law time exponent which may consequently distort the extrapolation of NBTI. To establish an accurate lifetime estimation, the key focus relies mainly on a fast characterization methodology with minimum disruption during accelerated stressing. In this thesis, the two fast measurement methodologies developed (UFS and UFS-CP), have a measurement delay of 100 ns which poses as a good candidate for future fast measurement characterization.

In addition, it is understood in this thesis that the origin of hole traps precursors is inherent to the quality of gate oxide. And with the increased complexity of nitrogen incorporated gate oxide processes, the magnitude of the cyclic hole trapping component is reported to increase with nitrogen concentration, which also explains the worsened NBTI effect observed for SiON gate oxide. Hence, the variation in terms of gate oxide fabrication process flow has to be taken into consideration for NBTI modeling and the extrapolation of the NBTI lifetime.

Lastly, higher stress voltage is often used to accelerate the failure mechanism (shorter testing time) to determine the lifetime and then extrapolate down to operating conditions. However, the author would like to emphasize the importance of the stress voltage for the characterization of NBTI. As highlighted in section 4.2.4, it is observed that at a higher stress voltage, the hole trap precursors that previously participated in the hole trapping/detrapping mechanism had transformed into a permanent form of oxide trap, as evident by the decreasing amount of cyclic threshold voltage shift. However, this transformation was not observed at lower stress voltage. Hence, the selection of stress voltage is critical to avoid triggering an additional degradation mechanism which may not play a role in NBTI degradation mechanisms at operating conditions.

LIST OF PUBLICATIONS

- **Z. Q. Teo**, D. S. Ang, and G. A. Du, “Observation of Two Gate Stress Voltage Dependence of NBTI Induced Threshold Voltage Shift of Ultra-Thin Oxynitride Gate P-MOSFET”, in *Proc. IRPS*, 2009.
- **Z. Q. Teo**, D. S. Ang, K. S. See and P. Z. Yang, “Effect of Mechanical Strain on the NBTI of Short-Channel P-MOSFETS : Role of Impact Ionization”, in *Proc. IRPS*, 2009.
- **Z. Q. Teo**, D. S. Ang and K. S. See, “Can the Reaction-Diffusion Model Explain Generation and Recovery of Interface States Contributing to NBTI?”, in *IEEE IEDM*, 2009.
- **Z. Q. Teo**, D. S. Ang and C. M. Ng, “Non-Hydrogen-Transport” Characteristics of Dynamic Negative-Bias Temperature Instability”, in *IEEE Elec. Dev. Lett.* 2010
- **Z. Q. Teo**, D. S. Ang and C. M. Ng, “Separation of Hole Trapping and Interface-State Generation by Ultrafast Measurement on Dynamic Negative-Bias Temperature Instability”, in *IEEE Elec. Dev. Lett.* 2010

- **Z. Q. Teo**, A. A. Boo, D. S. Ang and K. C. Leong, “On the Cyclic Threshold Voltage Shift of Dynamic Negative-Bias Temperature Instability”, in *Proc. IRPS*, 20011.
- D. S. Ang, **Z. Q. Teo** and C. M. Ng, “Reassessing NBTI Mechanism by Ultrafast Charge Pumping Measurement”, in *IEEE IIRW*, 2009
- D. S. Ang, **Z. Q. Teo**, T. J. J. Ho and C. M. Ng, “Reassessing the Mechanism of Negative-Bias Temperature Instability by Repetitive Stress/Relaxation Experiments”, in *IEEE Trans. on Dev Mater. Reliab.*, 2011
- G. A. Du, D. S. Ang, **Z. Q. Teo** and Y. Z. Hu, “Ultrafast Measurement on NBTI”, in *IEEE Elec. Dev. Lett.*, 2009.
- D. S. Ang, Y. Z. Hu, **Z. Q. Teo**, G. A. Du, C. S. Lai and J. J. Ho, “Negative-Bias Temperature Instability: Measurement and Degradation Mechanisms”, in *Trans. Electrochem.*, 2009.
- D. S. Ang, Y. Z. Hu, **Z. Q. Teo** and G. A. Du, and J. J. Ho, “Negative-Bias Temperature Instability of the Ultra-thin Gate p-MOSFET: Measurement and Physical Mechanisms”, in *Electrochem. Soc.*, 2009.

- Y. Z. Hu, D. S. Ang, **Z. Q. Teo** and G. A. Du, “A Robust Ultrafast Switching Methodology for Device Parameter Characterization of Bias-Temperature Instability”, in *Proc. IRPS*, 2010.
- Y. Z. Hu, D. S. Ang and **Z. Q. Teo**, “Threshold Voltage and Mobility Extraction by Ultrafast Switching Measurement on NBTI”, in *IEEE Elec. Dev. Lett.* 2010
- A. A. Boo, D. S. Ang, **Z. Q. Teo** and K. C. Leong, “Correlation Between Oxide Trap Generation and Negative-Bias Temperature Instability”, in *IEEE Elec. Dev. Lett.* 2012
- Y. Gao, A. A. Boo, **Z. Q. Teo** and D. S. Ang, “On the Evolution of the Recoverable Component of SiON, HfSiON and HfO₂ P-MOSFETs under Dynamic NBTI”, in *Proc. IRPS*, 2011.
- D. S. Ang, C. S. Lai, G. A. Du, **Z. Q. Teo**, J. J. Ho and Y. Z. Hu, “Effect of Hole-Trap Distribution on the Power-Law Time Exponent of NBTI”, in *IEEE Elec Dev Lett*, 2009.
- Y. Gao, D. S. Ang, A. A. Boo and **Z. Q. Teo**, “On the dynamic NBTI of the HfO₂ and HfSiON P-MOSFET”, in *Microelectro. Eng*, 2011.

- T. J. J. Ho, D. S. Ang, A. A. Boo, **Z. Q. Teo** and K. C. Leong, “Are Interface State Generation and Positive Oxide Trapping under Negative-Bias Temperature Stressing Correlated or Coupled?”, in *Trans. Elec Dev.*, 2012.

REFERENCES

- [1] Y. Miura and Y. Matukura, "Investigation of silicon–silicon dioxide interface using MOS structure," *Jpn J Appl Phys*, p. 5:180, 1966.
- [2] G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng, and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling,," *IEEE Trans. Electron Devices*, vol. Vol. 23, pp. pp. 734-736, 2002.
- [3] Y.-H. Lee, S. Jacobs, S. Stadler, N. Mielke, and R. Nachman, "The impact of PMOST bias-temperature degradation on logic circuit reliability performance.,," *Microelec Reliab*, pp.;45:107–14., 2005.
- [4] R. Kuroda, K. Watanabe, A. Teramoto, M. Mifuji, T. Yamaha, and S. Sugawa, "Acculate circuit performance prediction model and lifetime prediction method of NBT stressed devices for highly reliable ULSI circuits.,," *Int Electron Dev Meeting Techn Digest* pp.:717–20, 2005.
- [5] Y. Mitani, M. Nagamine, H. Satake, and Toriumi, "NBTI mechanism in ultra-thin gate dielectric—nitrogen-originated mechanism in SiON," *Int Electron Dev Meeting Techn Digest* pp. p:509–12, 2002.
- [6] K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," *Journal of Applied Physics*, vol. v 48, n 5, pp. 48, 2004, 1977.
- [7] Alam, "A critical examination of the mechanisms of dynamic NBTI for p-MOSFETs," *IEDM*, pp. pp. 345-348,, 2003.
- [8] M. A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, "A Comprehensive Model of PMOS NBTI Degradation: Recent progress," *Microelectronics Reliability*, 2007.
- [9] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability*, vol. v 45, pp. p 71-81, 2005.
- [10] M. F. L. C. Shen, C. E. Foo, T. Yang, D. M. Huang, A. Yap, G. S. Samudra and Y. C. Yeo, "Characterization and Physical Orgin of Fast Vth Transient in NBTI of pMOSFETs with SiON Dielectric," *IEEE Electron Device Meeting*, pp. 333-336, 2006.
- [11] M. Denais, C. Parthasarathy, G. Ribes, Y. Rey-Tauriac, N. Revil, A. Bravaix, *et al.*, "On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's," *IEDM Tech. Dig.*, pp. 109–112, 2004.
- [12] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam, "On the dispersive versus arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: measurements, theory, and implications," *IEDM Tech. Dig.*, pp. 684–687, 2005.
- [13] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI Degradation- and Recovery-Behavior Based on Ultra Fast VT-Measurements," *Proc. Int. Rel. Phys. Symp.*, pp. 448–453, 2006.
- [14] J. H. Lee, W. H. Wu, A. E. Islam, M. A. Alam, and A. S. Oates, "Separation method of hole trapping and interface trap generation and their roles in NBTI reaction-diffusion model," *Proc. Int. Rel. Phys. Symp.*, pp. 745–746, 2008.
- [15] V. Huard and M. Denais, "Hole trapping effect on methodology for DC and AC negative bias temperature instability measurements in PMOS transistors," *Proc. Int. Rel. Phys. Symp.*, pp. 40-45, 2006.
- [16] W. J. Liu, Z. Y. Liu, D. Huang, C. C. Liao, L. F. Zhang, Z. H. Gan, *et al.*, "On-The-Fly Interface Trap Measurement and Its Impact on the Understanding of NBTI Mechanism for p-MOSFETs with SiON Gate Dielectric," *Int. Electron Devices Meeting*, pp. 813-816, 2007.
- [17] Z. Q. Teo, A. A. Boo, D. S. Ang, and K. C. Leong, "On the cyclic threshold voltage shift of dynamic negative-bias temperature instability," in *49th International Reliability Physics Symposium, IRPS 2011, April 10, 2011 - April 14, 2011*, Monterey, CA, United states, 2011, pp. XT.10.1-XT.10.5.
- [18] Z. Q. Teo, D. S. Ang, and K. S. See, "Can the Reaction-Diffusion Model Explain Generation and Recovery of Interface States Contributing to NBTI?," in *2009 IEEE International Electron Devices Meeting (IEDM 2009)*, 7-9 Dec. 2009, Piscataway, NJ, USA, 2009, p. 4 pp.
- [19] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, "Characteristics of surface-state charge (Qss) of thermally oxidized silicon," *J. Electrochem. Soc.*, vol. 114, pp. 266-274, 1967.

-
- [20] N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai, and T. Horiuchi, "The impact of bias temperature instability for direct tunneling ultra-thin gate oxide on MOSFET scaling," *Symp. VLSI Tech.*, pp. 73-74, 1999.
 - [21] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *J. Appl. Phys.*, vol. 94, pp. 1-18, 2003.
 - [22] Y. Mitani, "Influence of nitrogen in ultra-thin SiON on negative bias temperature instability under AC stress," *Int. Electron Devices Meeting*, pp. 117-120, 2004.
 - [23] G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng, and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling," *IEEE Trans. Electron Devices*, vol. 23, pp. 734-736, 2002.
 - [24] H. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, and J. Backcock, "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field effect transistors," *Appl. Phys. Lett.*, vol. 83, pp. 1647-1649, 2003.
 - [25] G. A. Du, D. S. Ang, Z. Q. Teo, and Y. Z. Hu, "Ultrafast measurement on NBTI," *IEEE Electron Device Letters*, vol. 30, pp. 275-7, 03/ 2009.
 - [26] S. Chakravarthi, A. T. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability," *Proc. IRPS*, pp. 273-282, 2004.
 - [27] A. E. Islam, H. Kufluoglu, D. Varghese, and M. A. Islam, "Critical analysis of short-term negative bias temperature instability measurements: Explaining the effect of time-zero delay for on-the-fly measurements," *Appl. Phys. Lett.*, vol. 90, 2007.
 - [28] S. A. Mahapatra, K. ; Varghese, D. ; Islam, A.E. ; Gupta, G.; Madhav, L. ; Saha, D. ; Alam, M.A. , "On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?," *Proc. IRPS*, pp. 1-9, 2007.
 - [29] T. Grassler, W. Göss, V. Sverdlov, and B. Kaczer, "The universality of NBTI relaxation and its implications for modeling and characterization," *Proc. IRPS*, pp. 268-280, 2007.
 - [30] B. Kaczer, V. Arkhipov, M. Jurczak, and G. Groeseneken, "Negative Bias Temperature Instability (NBTI) in SiO₂ and SiON gate dielectrics understood through disorder-controlled kinetics," *Microelectronic Engineering*, vol. 80, pp. 122-125, 2005.
 - [31] M. Houssa, M. Aoulaiche, S. D. Gendt, G. Groeseneken, M. M. Heyns, and A. Stesmans, "Reaction-dispersive proton transport model for negative bias temperature instabilities," *Appl. Phys. Lett.*, vol. 86, 2005.
 - [32] D. S. Ang and S. Wang, "On the non-Arrhenius behavior of negative-bias temperature instability," *Applied Physics Letters*, vol. 88, p. 093506, 2006.
 - [33] D. S. Ang and K. L. Pey, "Evidence for two distinct positive trapped charge components in NBTI stressed p-MOSFETs employing ultrathin CVD silicon nitride gate dielectric," *IEEE Electron Device Letters*, vol. 25, pp. 637-639, 2004.
 - [34] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "A finite-oxide thickness-based analytical model for negative bias temperature instability," *IEEE Trans. Dev., Mat. Reliab.*, vol. 9, pp. 537-556, 2009.
 - [35] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, A. Varghese, *et al.*, "Material dependence of hydrogen diffusion: Implications for NBTI degradation," *IEDM Tech. Dig.*, pp. 688-691, 2005.
 - [36] H. Reisinger, R. P. Vollersten, P. J. Wagner, T. Huttner, A. Martin, S. Aresu, *et al.*, "A study of NBTI and short-term threshold hysteresis of thin nitrided and thick non-nitrided oxides," *IEEE Trans. Mat. Dev. Reliab.*, vol. 9, pp. 106-114, 2009.
 - [37] V. Huard, "Two independent components modeling for Negative Bias Temperature Instability," *Proc. IRPS*, pp. 33-42, 2010.
 - [38] S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, *et al.*, "A Comparative Study of Different Physics-Based NBTI Models," *IEEE Trans. Electron Devices*, vol. 60, pp. 901 - 916, 2013.
 - [39] A. Stesmans, "Dissociation Kinetics of Hydrogen-Passivated Pb Defects at the (111)Si/SiO₂ Interface," *Phys. Rev. B*, vol. 61, pp. 8393-8403, 2000.
 - [40] P. Gray and D. Brown, "Density of SiO₂-Si interface states," *Appl Phys Lett*, pp. 8:31-3;, 1966.
 - [41] E. Poindexter and P. Caplan, "Characterization of Si/SiO₂ interface defects by electron spin resonance," *Progr Surf Sci*, pp. 14:201-94, 1983.
-

-
- [42] M. Reed and J. Plummer, "Chemistry of Si-SiO₂ interface trap annealing," *J Appl Phys*, pp. 63:5776-93, 1988.
 - [43] A. Stirling, A. Pasquarello, J. Charlier, and R. Car, "Dangling bond defects at SiO₂/Si interfaces: atomic structure of the Pb1 center.," *Phys Rev Lett*, pp. 85:2773-6., 2000.
 - [44] P. M. Lenahan, "Atomic scale defects involved in MOS reliability problems," *Micro. Reliab.*, vol. 69, pp. 173-181, 2003.
 - [45] M. L. Reed and J. D. Plummer, "Chemistry of Si-SiO₂ interface trap annealing," *J. Appl. Phys.*, vol. 63, pp. 5776-5793, 1988.
 - [46] S. Mahapatra, V. D. Maheta, A. E. Islam, and M. A. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole Trapping Components in PNO p-MOSFETs," *IEEE Trans. Electron Devices*, vol. 56, pp. 236-242, 2009.
 - [47] J. F. Conley and P. M. Lenahan, "Electron spin resonance evidence for the structure of a switching oxide trap: Long term structural change at silicon dangling bond sites in SiO₂," *Appl Phys Lett*, vol. 67, 1995.
 - [48] J. F. Conley and P. M. Lenahan, "Electron Spin Resonance Evidence that E'γ₁ Centers Can Behave as Switching Oxide Traps," *Trans on Nuclear Science*, vol. 42, pp. 1744-1749, 1995.
 - [49] W. L. Warren, J. Kanicki, and E. H. Pondexter, "Paramagnetic point defects in silicon nitride and silicon oxynitride thin films on silicon," *Colloids and Surfaces A*, pp. 311-317, 1996.
 - [50] Y. H. Poon, H. Wong, and K. F. Man, "Deep level transient spectroscopy of hole emission from nitrided oxide/silicon interface traps," *Int. Electron Devices Meeting*, pp. 81-86, 1996.
 - [51] S. Ogawa, M. Shimaya, and N. Shiono, "Interface-trap generation at ultra-thin SiO₂ (4-6 nm)-Si interfaces during negative-bias temperature aging," *J. Appl. Phys.*, vol. Vol. 77, pp. pp.1137-1148, 1995.
 - [52] S. Mahapatra, P. B. Kumar, and M. A. Alam, "Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs," *IEEE Trans. Electron Devices*, vol. vol. 51, pp. pp. 1371-1379, 2004.
 - [53] S. S. H. Ershov, H. Karbasi, S. Winters, S. Minehane, and J. Backcock, , "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field effect transistors," *Appl. Phys. Lett.*, , vol. vol. 83, , pp. pp. 1647-1649, , 2003.
 - [54] S. Zafar, Y. H. Lee, and J. Stathis, "Evaluation of NBTI in HfO₂ gate-dielectric stacks with tungsten gates," *IEEE Electron Devices Lett.*, vol. 25, pp. 153-155, 2004.
 - [55] H. Ershov, H. Karbasi, S. Winters, S. Minehane, and J. Backcock, "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field effect transistors," *Appl. Phys. Lett.*, vol. 83, 2003.
 - [56] S. Rangan, N. Mielke, and E. C. C. Yeh, "Universal recovery behavior of negative bias temperature instability.," *IEDM*, pp. pp. 341-344, 2003.
 - [57] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent issues in negative-bias temperature instability - Initial degradation, field dependence of interface trap generation, hole trapping effects, and relaxation," *IEEE Trans. Electron Devices*, vol. 54, pp. 2143-2154, 2007.
 - [58] H. Kufluoglu and M. A. Alam, "A theory of interface trap induced NBTI degradation for reduced cross-section MOSFETs," *IEEE Trans Electron Dev*, vol. 53, pp. 20-30, 2005.
 - [59] S. Zafar, Y. H. Lee, and J. Stathis, "Evaluation of NBTI in HfO₂ gate-dielectric stacks with tungsten gates.," *IEEE Electron Devices Lett*, vol. vol. 25, , pp. pp. 153-155, 2004.
 - [60] C. Shen, M. F. Li, X. P. Wang, Y. C. Yeo, and D.-L. Kwong, "A fast measurement technique of MOSFET Id-Vg Characteristics," *IEEE Electron Device Letters*, vol. 27, 2006.
 - [61] T. Yang, M. F. Li, C. Shen, C. H. Ang, C. Zhu, Y.-C. Yeo, *et al.*, "Fast and slow dynamic NBTI components in p-MOSFET with SiON dielectric and their impact on device life-time and circuit application," *Symp VLSI Tech.*, pp. 92-93, 2005.
 - [62] D. S. Ang, "Observation of Suppressed Interface State Relaxation Under Positive Gate Biasing of Ultrathin Oxynitride Gate PMOS Subjected to NBTI Stressing," *IEEE Electron Device Letters*, vol. 27, pp. 412-415, 2006.
 - [63] H. Youzhou, "Improved Electrical Characterization of Negative-Bias Temperature Instability " *Ph.D NTU*, 2013.
 - [64] Y. Z. Hu, D. S. Ang, and Z. Q. Teo, "Threshold voltage and mobility extraction by ultrafast switching measurement on NBTI," *IEEE Transactions on Electron Devices*, vol. 57, pp. 2027-31, 08/ 2010.
-

-
- [65] T. J. J. Ho, D. S. Ang, A. A. Boo, Z. Q. Teo, and K. C. Leong, "Are interface state generation and positive oxide charge trapping under negative-bias temperature stressing correlated or coupled?," *IEEE Transactions on Electron Devices*, vol. 59, pp. 1013-1022, 2012.
 - [66] A. E. Islam, "Modeling Interface-defect Generation (MIG)," DOI: 10254/nanohub-r1647.2, 2006
 - [67] T. Grassler, B. Kaczer, P. Hehenberger, W. Göss, R. O'Connor, H. Reisinger, *et al.*, "Simultaneous extraction of recoverable and permanent components contributing to bias-temperature instability," *Int. Electron Dev. Mtg. Tech. Dig.*, pp. 801-804, 2007.
 - [68] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "Location, structure, and density of states of NBTI-induced defects in plasma nitrided pMOSFETs," *Proc. Int. Reliab. Phys. Symp.*, pp. 503-510, 2007.
 - [69] A. J. Leis and T. R. Oldham, "Time dependence of switching oxide traps," *IEEE Trans. Nucl. Sci.*, vol. 41, pp. 1835-1843, 1994.
 - [70] Z. Q. Teo, D. S. Ang, and C. M. Ng, "'Non-hydrogen-transport' characteristics of dynamic negative-bias temperature instability," *IEEE Electron Device Letters*, vol. 31, pp. 269-271, 2010.
 - [71] J. F. Zhang, C. Z. Zhao, A. H. Chen, G. Groeseneken, and R. Degraeve, "Hole traps in silicon dioxides – Part I: Properties," *IEEE Trans. Electron Devices*, vol. 51, pp. 1267-1273, 2004.
 - [72] J. P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, and S. Krishnan, "Atomic-scale defects involved in the negative-bias temperature instability," *IEEE Trans. Dev. & Mat. Reliab.*, vol. 7, pp. 540-557, 2007.
 - [73] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "Identification of the atomic-scale defects involved in the negative bias temperature instability in plasma-nitrided p-channel metal-oxide-silicon field-effect transistors," *J. Appl. Phys.*, vol. 103, 2008.
 - [74] J. T. Ryan, P. M. Lenahan, T. Grassler, and H. Enichlmair, "Observations of negative bias temperature instability defect generation via on the fly electron spin resonance," *Appl. Phys. Lett.*, vol. 96, 2010.
 - [75] C. J. Nicklaw, Z. Y. Lu, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "The structure, properties, and dynamics of oxygen vacancies in amorphous SiO₂," *IEEE Transactions on Nuclear Science*, vol. 49, pp. 2667-2673, Dec 2002.
 - [76] Z. Y. Lu, C. J. Nicklaw, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Structure, properties, and dynamics of oxygen vacancies in amorphous SiO₂," *Physical Review Letters*, vol. 89, Dec 2002.
 - [77] D. Goguenheim and M. Lannoo, "THEORETICAL AND EXPERIMENTAL ASPECTS OF THE THERMAL-DEPENDENCE OF ELECTRON-CAPTURE COEFFICIENTS," *Journal of Applied Physics*, vol. 68, pp. 1059-1069, Aug 1990.
 - [78] D. Goguenheim and M. Lannoo, "THEORETICAL CALCULATION OF THE ELECTRON-CAPTURE CROSS-SECTION DUE TO A DANGLING BOND AT THE Si(111)-SiO₂ INTERFACE," *Physical Review B*, vol. 44, pp. 1724-1733, Jul 1991.
 - [79] T. Grassler, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelheibel, "A two-stage model for negative bias temperature instability," *Proc. IRPS*, pp. 33-44, 2009.
 - [80] D. S. Ang and S. Wang, "Recovery of the NBTI stressed ultra-thin gate P-MOSFET: The role of deep-level hole traps," *IEEE Electron Dev. Lett.*, vol. 27, pp. 914-916, 2006.
 - [81] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative bias temperature instability," *IEEE Trans. Dev. Mat. Reliab.*, vol. 8, pp. 22-34, 2008.
 - [82] Z. Q. Teo, D. S. Ang, and G. A. Du, "Observation of two gate stress voltage dependence of NBTI induced threshold voltage shift of ultra-thin oxynitride gate p-MOSFET," in *2009 IEEE International Reliability Physics Symposium (IRPS)*, 26-30 April 2009, Piscataway, NJ, USA, 2009, pp. 1002-4.
 - [83] Y. Gao, A. A. Boo, Z. Q. Teo, and D. S. Ang, "On the evolution of the recoverable component of the SiON, HfSiON and HfO₂ P-MOSFETs under dynamic NBTI," in *49th International Reliability Physics Symposium, IRPS 2011, April 10, 2011 - April 14, 2011*, Monterey, CA, United states, 2011, pp. XT.8.1-XT.8.6.
 - [84] T. Grassler, P.-J. Wagner, H. Reisinger, T. Aichinger, G. Pobegen, M. Nelhiebel, *et al.*, "Analytic modeling of the bias temperature instability using capture/emission time maps," *IEDM Tech. Dig.*, pp. 618-621, 2011.
-

-
- [85] T. Grassler, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time-dependent defect spectroscopy for characterization of border traps in metal–oxide–semiconductor transistors," *Phys. Rev. B*, vol. 82, 2010.
- [86] V. Huard, F. Monsieur, G. Ribes, and S. Bruyere, "Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs," *Proc. IRPS*, pp. 178–182, 2003.
- [87] S. Tsujikawa and J. Yugami, "Evidence for bulk trap generation during NBTI phenomenon in pMOSFETs with ultrathin SiON gate dielectrics," *IEEE Trans. Electron Devices*, vol. 53, pp. 51–55, 2006.
- [88] S. Mahapatra, A. E. Islam, S. Deora, V. D. Maheta, K. Joshi, A. Jain, *et al.*, "A critical re-evaluation of the usefulness of R-D framework in predicting NBTI stress and recovery," *Proc. IRPS*, pp. 614–623, 2011.
- [89] W. L. Warren, E. H. Poindexter, M. Offenberger, and W. Müller-Warmuth, "Paramagnetic point defects in amorphous silicon dioxide and amorphous silicon nitride thin films I. a-SiO₂," *J. Electrochem. Soc.*, vol. 139, pp. 872–880, 1992.
- [90] J. T. Ryan, P. M. Lenahan, T. Grassler, and H. Enichlmair, "Recovery free electron spin resonance observations of NBTI degradation," *Proc. Int. Rel. Phys. Symp.*, pp. 43–49, 2010.
- [91] T. Uchino and T. Yoko, "Density functional theory of structural transformations of oxygen-deficient centers in amorphous silica during hole trapping: Structure and formation mechanism of the E' center," *Phys. Rev. B*, vol. 74, 2006.
- [92] R. E. Paulsen, R. R. Siergiej, M. L. French, and M. H. White, "Observation of near-interface oxide traps with the charge-pumping technique," *IEEE Electron Dev. Lett.*, vol. 13, 1992.
- [93] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. De Keersmaecker, "A reliable approach to charge-pumping measurements in MOS transistors," *IEEE Transactions on Electron Devices*, vol. ED-31, pp. 42–53, 01/ 1984.
- [94] D. Bauza, "A general and reliable model for charge pumping-Part I: model and basic charge-pumping mechanisms," *IEEE Transactions on Electron Devices*, vol. 56, pp. 70–7, 01/ 2009.
- [95] M. Denais, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, *et al.*, "Interface trap generation and hole trapping under NBTI and PBTI in advanced CMOS technology with a 2-nm gate oxide," *IEEE Trans. Mat. Dev. Reliab.*, vol. 4, pp. 715–722, 2004.
- [96] G. A. D. D.S. Ang, S. Wang, "Mechanism of Bias-Temperature Instability: Results from Positive Gate Stress," *Physical and Failure Analysis of Integrated Circuits, 2007. IPFA 2007. 14th International Symposium on the*, 2007.
- [97] D. S. Ang, S. C. S. Lai, G. A. Du, Z. Q. Teo, T. J. J. Ho, and Y. Z. Hu, "Effect of hole-trap distribution on the power-law time exponent of NBTI," *IEEE Electron Device Letters*, vol. 30, pp. 751–3, 07/ 2009.
- [98] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectron. Reliab.*, vol. 46, pp. 1–23, 2006.