

CHARACTERIZATION OF PMOSFET DEGRADATION IN  
NEGATIVE BIAS TEMPERATURE INSTABILITY TEST

SOON FOO YEW

FACULTY OF ENGINEERING  
UNIVERSITY OF MALAYA  
KUALA LUMPUR, MALAYSIA

2012

**CHARACTERIZATION OF pMOSFET DEGRADATION IN  
NEGATIVE BIAS TEMPERATURE INSTABILITY TEST**

**SOON FOO YEW**

DISSERTATION SUBMITTED IN FULFILMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF MASTER OF  
ENGINEERING SCIENCE

FACULTY OF ENGINEERING  
UNIVERSITY OF MALAYA  
KUALA LUMPUR, MALAYSIA

2012

# UNIVERSITY of MALAYA

## ORIGINAL LITERARY WORK DECLARATION

**Name of Candidate:** SOON FOO YEW (IC/Passport No.:

**Registration/Matric No.:** KGA 100018

**Name of Degree:** MASTER OF ENGINEERING SCIENCE

**Title of Project Paper / Research Report / Dissertation / Thesis ("This Work"):**

CHARACTERIZATION OF pMOSFET DEGRADATION IN NEGATIVE BIAS

TEMPERATURE INSTABILITY TEST

**Field of Study:** MICROELECTRONICS

I do solemnly and sincerely declare that:

- (1) I am the sole author / writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate Signature

Date:

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name:

Designation:

## Abstract

Threshold voltage instability has become a major IC reliability concern for sub-micron CMOS process technology. In the past,  $V_{TH}$  Stability test is commonly used by the wafer fabrication plant to have a quick assessment on this reliability concern during process qualification, due to its simple test procedure. In recent years, Negative Bias Temperature Instability (NBTI) test has been used extensively to characterize not only the threshold voltage, but also other transistor parameters. The NBTI test consists of interim measurements at high temperature during stress, allowing degradation behavior to be studied in more detail than  $V_{TH}$  Stability test. This experimental study has demonstrated the capability in measuring transistor parametric at high temperature for NBTI characterization. The conventional DC NBTI stress tests are performed on  $0.18\ \mu\text{m}$  pMOSFET with a gate oxide thickness of  $2.9\ \text{nm}$ , fabricated on  $0.18\ \mu\text{m}$  Dual Gate CMOS process. Data analysis on parametric degradation behavior indicates that increase in interface states play a dominant role and the extracted  $n$  exponent matches the analytically derived value from the Reaction-Diffusion model. Also, the results from this conventional DC NBTI stress show that degradation seen on saturation mode is less severe than the degradation on linear mode of pMOSFET operation. This is further confirmed by analysis of lifetime extrapolation which shows the  $I_{DSAT}$  (saturation mode) lifetime is 2 order of magnitudes higher than  $V_{TEXT}/V_{TCI}$  (linear mode) lifetime. Additional analysis shows that  $V_{TCI}$  is the most sensitive parameter to be monitored during NBTI stress and will be used as the key parameter in the later stage of this experimental study. From further experimental work, it can be seen that NBTI test allows a deeper understanding of pMOSFET parametric degradation behavior than  $V_{TH}$  Stability test, that shows a strong dependence of NBTI degradation on temperature,

channel length and gate oxide thickness variation. In addition, the proposed Optimized  $I_D$ - $V_G$  sweep measurement method is shown to generate a more accurate lifetime extraction and can easily be implemented in wafer fabrication plant without additional hardware or software. This method is applicable for process development, process qualification and periodical monitoring for NBTI degradation performance during mass production.

## Abstrak

Ketidakstabilan voltan ambang telah menjadi kebimbangan utama kebolehpercayaan IC untuk proses teknologi sub-mikron CMOS. Biasanya, ujian Kestabilan  $V_{TH}$  ( $V_{TH}$  Stability) digunakan oleh pengeluar wafer terfabrikasi semasa proses kelayakan untuk mendapatkan penilaian yang cepat pada peringkat kebolehpercayaan, disebabkan oleh prosedur ujian yang mudah. Kebelakangan ini, ujian Bias Ketakstabilan Suhu Negatif (NBTI) telah digunakan secara meluas untuk mencirikan bukan sahaja voltan ambang, tetapi juga parameter transistor lain. Ujian NBTI terdiri daripada ukuran interim pada suhu yang tinggi semasa tekanan voltan, membenarkan tingkah laku degradasi untuk dikaji dengan lebih terperinci daripada ujian Kestabilan  $V_{TH}$ . Kajian eksperimen ini telah menunjukkan keupayaan dalam mengukur parameter transistor pada suhu yang tinggi untuk pencirian NBTI. Ujian tekanan DC konvensional NBTI telah dilaksanakan ke atas pMOSFET 0.18  $\mu\text{m}$  dengan ketebalan oksida Get sebanyak 2.9 nm, yang dibentuk dengan proses CMOS 0.18  $\mu\text{m}$  Dwi-Get. Analisis data tentang tingkah laku degradasi berparameter menunjukkan bahawa peningkatan dalam ruang antara-muka memainkan peranan yang dominan dan eksponen  $n$  yang diekstrak sepadan dengan nilai analitis yang diperolehi daripada model Reaksi-Resapan. Selain itu, hasil daripada NBTI DC konvensional ini menunjukkan bahawa degradasi yang dilihat pada mod ketepuan kurang daripada degradasi pada mod linear operasi pMOSFET. Ini adalah disahkan dengan analisis penentuluaran jangka hayat yang menunjukkan hayat  $I_{DSAT}$  (mod ketepuan) adalah 2 tertib magnitud lebih tinggi daripada hayat  $V_{TEXT} / V_{TCI}$  (mod linear). Analisis tambahan menunjukkan bahawa  $V_{TCI}$  merupakan parameter yang paling sensitif yang dipantau semasa tekanan NBTI dan akan digunakan sebagai parameter utama dalam peringkat kajian lanjutan eksperimen ini. Dari kerja uji kajian yang lanjut,

ia boleh dilihat bahawa ujian NBTI membolehkan pemahaman yang lebih mendalam untuk tingkah laku degradasi berparameter pMOSFET daripada ujian Kestabilan  $V_{TH}$ , yang menunjukkan pergantungan kuat degradasi NBTI kepada perubahan suhu, kepanjangan saluran transistor dan ketebalan oksida Get. Di samping itu, cadangan kaedah pengukuran dengan mengoptimumkan sapuan  $I_D-V_G$  menunjukkan penentuluan jangka hayat yang lebih tepat dan boleh dilaksanakan dengan mudah dalam pengeluaran wafer terfabrikasi tanpa perkakasan atau perisian tambahan. Kaedah ini boleh digunakan untuk proses pembangunan, proses kelayakan dan pemantauan berkala untuk prestasi degradasi NBTI semasa pengeluaran secara besar-besaran.

## **Acknowledgments**

I would like to express my sincere thanks and gratitude to my adviser, Assoc. Prof. Dr. Norhayati Soin for her guidance and support through out this project. At the same time, I would like to thank Assoc. Prof. Dr. Zahirul Alam at International Islamic University, Malaysia (IIUM) that supported the project fully and allowed us to use the lab equipments freely. Special thanks go to my fellow researchers, S.F.Wan Muhamad Hatta and Dayanasari Abd. Hadi for their assistance, supportive ideas and constructive discussion. Also, I wish to express my sincere thanks to Lai Kin On and Sharifah Shafini from Silterra (M) Sdn. Bhd., for providing the wafer sample and information related to the wafer sample. Not to be left out, I would like to give a big thumb up to IIUM Electronics System Lab technician, Mohd. Saiful for his help in keeping the lab equipment in good working conditions.

Finally, I am grateful to my family for their love and support. Especially to my mother who took care of the daily routine, allowing me to concentrate on this project.

# Contents

	Page
<b>Abstract</b>	ii
<b>Abstrak</b>	iv
<b>Acknowledgements</b>	vi
<b>Contents</b>	vii
<b>List of Figures</b>	xi
<b>List of Tables</b>	xv
<b>List of Symbols and Abbreviations</b>	xvi
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Motivation	2
1.3 Objectives	2
1.4 Scope of this Research	3
1.5 Organization of the Dissertation	3
<b>Chapter 2 Degradation in CMOS</b>	<b>6</b>
2.1 Introduction	6
2.2 Hot Carrier Injection	6
2.2.1. Modeling of Hot Carrier Injection	8
2.3 Time-Dependent Dielectric Breakdown	9
2.3.1. E-Model or Constant Field/Voltage Acceleration Exponential Model	11
2.3.2. 1/E-Model or Anode Hole Injection Model	11

2.3.3. V-Model (exponential with voltage)	12
2.3.4. V-Model (Power law)	12
2.4 Electromigration	12
2.4.1. Modeling of Electromigration	15
2.5 Negative Bias Temperature Instability	15
2.5.1. Brief History of NBTI	16
2.5.2. Interface Trap Charge	18
2.5.3. Mechanism and Modeling of NBTI	20
2.5.3(a) Hole Trapping Model	20
2.5.3(b) Thermally Assisted Electron Tunneling Model	20
2.5.3(c) Reaction-Diffusion Model	21
2.5.4. Recovery of NBTI	26
<b>Chapter 3 Measurement Methods of CMOS Degradation</b>	<b>27</b>
3.1 Introduction	27
3.2 MOSFET Saturation Drain Current and Threshold Voltage	27
3.2.1. Threshold voltage measurement	29
3.2.1(a) Maximum Transconductance ( $g_{m,max}$ ) Method, $V_{TEXT}$	29
3.2.1(b) Constant Current Method, $V_{TCI}$	30
3.2.2. Saturation drain current measurement	31
3.3 $I_{Sub,max}$ Stress Measurement for HCI	32
3.4 Charge Pumping for Interface States Measurement	34
3.5 Constant Voltage Stress Measurement for TDDB	36
3.6 Constant Current Stress Measurement for EM	37
3.7 NBTI Measurement Techniques	39
3.7.1. Conventional DC Stress-Measure-Stress method	39

3.7.2. On-The-Fly Measurement Method	41
3.7.3. Fast Switching Method	42
3.7.4. Other Advanced Measurement Methods	45
<b>Chapter 4 Research Methodology</b>	<b>48</b>
4.1 Introduction	48
4.2 Research Project Flow Chart	48
4.3 Wafer Sample	51
4.3.1. Pad group design	52
4.3.2. Selection criteria for pMOSFET	54
4.4 Hardware: Keithley Measurement System	55
4.4.1. Correlation of Keithley 4200 SCS to Industry	55
4.4.2. Keithley 4200 SCS correlation results	57
4.5 Hardware: Everbeing Probe Station	59
4.5.1. Temperature profiling on Everbeing probe station	59
4.5.2. Results of temperature profiling	60
4.5.3. Constraint in using Everbeing hot chuck	62
4.6 NBTI Study with Conventional DC Stress-Measure-Stress Method	63
4.6.1. Experimental set-up and conditions	63
4.7 The Effect of Temperature, Gate oxide Thickness and Channel Length on NBTI Degradation	65
4.8 Alternative to Advance Fast Measurement Technique	65
4.8.1. Optimized $I_D$ - $V_G$ sweep ( $V_{THopt}$ )	66
<b>Chapter 5 NBTI Experimental Results and Discussion</b>	<b>69</b>
5.1 Introduction	69

5.2 Measurement Accuracy and Repeatability	69
5.3 $I_D$ - $V_D$ and $I_D$ - $V_G$ Degradation Curves	70
5.4 Parameter degradation-vs-Stress Time	73
5.5 NBTI Lifetime Extraction	76
5.6 The Effect of Temperature on NBTI Degradation	78
5.6.1. Arrhenius Plot on $\Delta V_{TCI}$	79
5.6.2. Arrhenius Model and Activation Energy	80
5.6.3. The Effect of Thermal Acceleration in NBTI Mechanism	81
5.7 The Effect of Channel Length (L) on NBTI Degradation	82
5.8 The Effect of Gate oxide Thickness on NBTI Degradation	89
5.9 Optimized $I_D$ - $V_G$ ( $V_{THopt}$ ) to Reduce 'Dead' Time	94
5.9.1. Optimized $I_D$ - $V_G$ Sweep Measurement Results and Discussion	95
5.9.2. Statistical Analysis on Comparing $\Delta V_{THopt}$ and $\Delta V_{TCI}$	97
5.9.3. Lifetime Extraction on $V_{THopt}$	102
5.10 Implication of this Research Work to the Industry	103
<b>Chapter 6 Conclusion and Future Work</b>	<b>107</b>
6.1 Recommendations on Future work	109
<b>Bibliography</b>	<b>111</b>
<b>List of Publications</b>	<b>118</b>

## List of Figures

<b>Figure</b>	<b>Caption</b>	<b>Page</b>
Figure 2.1	Hot carrier injection mechanism for nMOSFET (Keithley, 2000).	7
Figure 2.2	Traps and defects formed a conductive path across the gate oxide (Vollertsen, 2004).	10
Figure 2.3	Drift (diffusion) of metal atom under the influence of “electron wind” (Ho, 2004).	13
Figure 2.4	Formation of void and hillock as a result of electromigration. (Ohring, 1998)	14
Figure 2.5	CMOS circuit in logic gate representation, showing pMOSFET in inverter gate under negative gate bias with reference to source and drain (Alam, 2005).	16
Figure 2.6	Interface states on the Si-SiO <sub>2</sub> interface (Schroder and Babcock, 2003).	18
Figure 2.7	Interface states in the band diagram (Schroder and Babcock, 2003).	19
Figure 2.8	Diffusion of hydrogen (Alam et al., 2007).	24
Figure 3.1	pMOSFET operated in the linear region (low drain voltage) (Acuthan and Bhat, 2007).	28
Figure 3.2	pMOSFET operated in saturation region, after pinch off (Acuthan and Bhat, 2007).	28
Figure 3.3	I <sub>D</sub> -V <sub>G</sub> plot with the transconductance, g <sub>m</sub> (blue dotted plot) showing a maximum point. The red plot is I <sub>D</sub> and the black line is the tangent at the maximum point of g <sub>m</sub> . Also shown is the V <sub>TCI</sub> determination through I <sub>0</sub> .	30
Figure 3.4	I <sub>D</sub> -V <sub>D</sub> curves with V <sub>G</sub> at different levels, showing I <sub>D</sub> is controlled by V <sub>G</sub> in saturation region.	31
Figure 3.5	HCI stress test procedure (JESD28-A, 2001).	32
Figure 3.6	I <sub>SUB</sub> is substrate current showing the maximum point (Keithley, 2000).	33
Figure 3.7	nMOSFET under charge pumping setup. PGU = pulse generator unit, SMU = source measurement unit (Keithley, 2007).	35
Figure 3.8	I-t trace showing jump in gate leakage current due to breakdown. Dotted line indicates the failing criteria (Vollertsen, 2004).	37
Figure 3.9	Conventional DC Stress Measure Stress flow (JESD90, 2004).	40

Figure 3.10	Measurement schematics for (a) drain current (Denais et al., 2004), (b) drain current and transconductance without interrupting stress (Huard et al., 2005).	42
Figure 3.11	(a) The pMOSFET is biased as shown. $V_G$ is varied between the measurement and stress values while $I_D$ is monitored. (b) $V_G$ is converted to $\Delta V_{TH}$ by horizontally shifting the initial $I_D$ - $V_G$ curve (Kaczer et al., 2005).	43
Figure 3.12	(a) $I_D$ - $V_G$ plotted in semi-log plot. The subthreshold slope is the linear portion of the curve. (b) The subthreshold slope (SS) can be easily extracted by solving the exponential equation on the linear portion (Schlunder et al., 2008).	44
Figure 3.13	Test configuration of Ultra Fast $V_{TH}$ Method (Reisinger et al., 2006).	45
Figure 3.14	(a) Test configuration of Fast $V_G$ -Ramp introduced by Kerber et al. (2003). (b) An improved version with extra circuitry and same length coaxial cable, being introduced by Shen et al. (2006).	46
Figure 3.15	(a) Ultra Fast Switching test configuration utilizing pulsed I-V measurement scheme. (b) Waveform showing the stress and measurement. (Du et al., 2009).	47
Figure 4.1	Flow chart showing the feasibility study phase of the project. The flow chart will continue to Figure 4.2.	49
Figure 4.2	Continuation of project flow chart showing experimental phase.	50
Figure 4.3	Wafer map showing the location of test structures of interest.	51
Figure 4.4	Description of pad group ET02 (left, pMOSFET thin oxide) and ET04 (right, pMOSFET thick oxide).	52
Figure 4.5	Keithley 4200 SCS front panel (Keithley, 2009).	55
Figure 4.6	Typical configuration of a SMU (Keithley, 2009).	56
Figure 4.7	$I_D$ - $V_D$ sweep with $V_G$ stepping from 0 V $\rightarrow$ -1.8 V at step of -0.3 V. The red color square is the $I_{DSAT}$ data from wafer fabrication plant.	58
Figure 4.8	$I_D$ - $V_G$ sweep at $V_D = -0.1$ V, plotted on semi-log scale. The $I_D$ is in absolute form. For 10/0.18 $\mu$ m pMOSFET, $V_{TCI}$ is $V_G @ I_D = 5.556E-6$ A. The red color square is the $V_{TCI}$ data from wafer fabrication plant.	58
Figure 4.9	Everbeing probe station. (Everbeing, 2002)	59
Figure 4.10	Everbeing probe station hot chuck profile with an 8" wafer sample.	61

Figure 4.11	Bias configuration during NBTI stress of a pMOSFET (Yamamoto, 1999).	64
Figure 4.12	$I_D$ - $V_G$ sweep showing the minimum range of measurement to determine $V_{TEXT}$ and $V_{TCI}$ . This data is measured at room temperature.	68
Figure 4.13	A portion of the $I_D$ - $V_G$ sweep measurement from conventional DC SMS, showing the $V_{TCI0}$ , and the subsequent increase in $V_{TCI}$ as the stress progresses.	68
Figure 5.1	$I_D$ - $V_D$ ( $V_G = -1.8$ V) sweep degradation for $V_{G, stress} = -3.4$ V, 125 °C.	71
Figure 5.2	$I_D$ - $V_G$ ( $V_D = -0.1$ V) sweep degradation at $V_{G, stress} = -3.4$ V, 125 °C. .	71
Figure 5.3	$I_{DSAT}$ degradation-vs-stress time, plotted on log-log scale. The equation shown is in the form of Power Law equation. The black horizontal line indicates the 10% shift in $I_{DSAT}$ .	73
Figure 5.4	$V_{TEXT}$ degradation-vs-stress time, plotted on log-log scale.	74
Figure 5.5	$V_{TCI}$ degradation-vs-stress time, plotted on log-log scale.	75
Figure 5.6	Comparison of degradation between $I_{DSAT}$ , $V_{TEXT}$ and $V_{TCI}$ at stress voltage of -3.4 V, 125 °C.	76
Figure 5.7	TTF-vs- $V_{G, stress}$ plotted on semi-log plot to extract the lifetime at 125 °C. The equations shown are in the form of Exponential Law equation.	78
Figure 5.8	$\Delta V_{TCI}$ -vs- $1/k_B T$ plotted on semi-log plot to extract the $E_a$ at $V_G = -3.0$ V, -3.2 V. The equations shown are in the form of Exponential law equation.	79
Figure 5.9	(a) pMOSFET at the point of pinch-off. (b) pMOSFET beyond pinch-off with short channel effect (Acuthan and Bhat, 2007).	83
Figure 5.10	Comparison of $V_{TCI}$ degradation between 10/0.18 $\mu\text{m}$ and 10/10 $\mu\text{m}$ at stress voltage of -2.8 V & -3.2 V, 125 °C.	84
Figure 5.11	Oxide charge in gate oxide (Taur and Ning, 1998).	86
Figure 5.12	$\Delta V_{TH}$ -vs- $L$ at different $V_{G, stress}$ , temperature and stress time (Cellere et al., 2004).	87
Figure 5.13	TTF-vs- $ V_{G, stress} $ for lifetime extraction on $I_{DSAT}$ and $V_{TCI}$ , comparing 10/0.18 $\mu\text{m}$ and 10/10 $\mu\text{m}$ pMOSFET.	88
Figure 5.14	Comparison of $V_{TCI}$ degradation between 10/0.18 $\mu\text{m}$ (thin oxide) and 10/0.30 $\mu\text{m}$ (thick oxide) for $E_{ox}$ of 9.5 MV/cm and 10.85 MV/cm.	91

Figure 5.15	Interface trap generation for pMOSFET with pure and nitrided oxides over a wide range of gate oxide thickness (Huard et al., 2006).	92
Figure 5.16	TTF-vs- $E_{G, stress}$ for lifetime extraction on $I_{DSAT}$ and $V_{TCI}$ , comparing 10/0.18 $\mu m$ and 10/0.30 $\mu m$ pMOSFET.	93
Figure 5.17	Shift in $V_{THopt}$ vs Stress Time on a log-log plot. The time exponent, $n$ ranged from 0.186 to 0.192.	96
Figure 5.18	$V_{TCI}$ vs $V_{THopt}$ degradation at -3.4 V and -3.0 V.	96
Figure 5.19	F distribution showing example of rejection area for hypothesis testing.	98
Figure 5.20	t distribution showing example of rejection area for hypothesis testing.	100
Figure 5.21	TTF-vs- $V_{G, stress}$ for lifetime extraction on $V_{TCI}$ and $V_{THopt}$ .	102

## List of Tables

<b>Table</b>	<b>Caption</b>	<b>Page</b>
Table 4.1	Correlation data on $I_{DSAT}$ and $V_{TCI}$ showing the data from wafer fabrication plant matching very well with Keithley 4200 SCS.	57
Table 4.2	Wafer surface temperature measured with Fluke 568 thermometer at different locations of the hot chuck, with Temperature Controller set at 145 °C.	61
Table 4.3	Additional set of measurement at 3 locations with Temperature Controller set at 125 °C.	62
Table 4.4	Stress conditions for conventional DC SMS.	64
Table 4.5	Stress conditions for thick oxide pMOSFET.	66
Table 5.1	Comparison between thin oxide and thick oxide.	90
Table 5.2	Test statistic, $F_o$ for the specified time interval at -3.4 V and -3.0 V.	99
Table 5.3	Test statistic, $t_o$ for the specified time interval at -3.4 V and -3.0 V.	101

## List of Symbols and Abbreviations

$\beta$	Voltage acceleration factor in TDDB V-model (exponential)
$\Delta E$	Difference between inversion Fermi and accumulation Fermi
$\Phi_F$	Fermi potential
$\epsilon_0$	Permittivity of free space / vacuum
$\epsilon_{ox}$	Permittivity (dielectric constant) of SiO <sub>2</sub>
$\epsilon_S$	Permittivity (dielectric constant) of silicon
$\gamma$	Field acceleration factor in TDDB E-model
$\mu$	Mobility for charge carrier
$\mu_H$	Mobility of hydrogen
$\mu_1$	Mean of population (first group) in hypothesis testing
$\mu_2$	Mean of population (second group) in hypothesis testing
$\mu A$	micro-ampere
$\mu m$	micrometer
$\mu sec$	micro-second
$\sigma_1^2$	Variance of population (first group) in hypothesis testing
$\sigma_2^2$	Variance of population (second group) in hypothesis testing
$V_g$	Pulse at gate, supply by pulse generator
$\tau$	Lifetime
$a$	Order of the reaction in Reaction-Diffusion model
$A$	Constant in HCI $1/V_{DS}$ model, EM Black's equation, NBTI Lifetime model, Arrhenius model
$A_0$	Fit parameter in TDDB E-model, 1/E-model, V-model (exponential)
$\overline{A}_1$	Mean of sample (first group) in hypothesis testing

$\bar{A}_2$	Mean of sample (second group) in hypothesis testing
$A_r$	Area of gate
$A_T$	Thermal acceleration factor
$B$	Fit parameter in HCI $I_{Sub}/W$ model
$B_0$	Fit parameter in TDDB V-model (Power law)
$BV_{ox}$	Oxide Breakdown voltage
$C$	Fit parameter in HCI $I_{Sub}/W$ model
$C_{ox}$	Gate oxide capacitance per unit area
$D$	Fit parameter in TDDB 1/E-model
$D_H$	Diffusion constant of hydrogen
$D_{it}$	Density of interface trap charge
$E$	Electric field
$E_a$	Activation energy
$E_{a-cht}$	Activation energy of chemical reaction
$E_{crit}$	Critical Electric Field
$E_{G, stress}$	Electric field across gate oxide during stress
$E_{ox}$	Electric field across gate oxide
$f$	Frequency
$F_o$	Test statistic for F distribution
$F_{crit}$	Critical level in F distribution
$G$	Fit parameter in HCI $I_{Sub}/I_D$ model
$\text{GHz}$	Giga Hertz
$g_m$	Transconductance
$H$ or $H^\circ$	Hydrogen atom
$H_2$	Hydrogen molecule

$H^+$	Hydrogen with a proton
$h^+$	Trapping holes
$H_o$	Null hypothesis
$H_A$	Alternative hypothesis
$I_0$	Gate leakage before TDDB stress
$I_{cp}$	Charge Pumping current
$I_D$	Drain current
$I_{DLIN}$	Drain current in Linear mode
$I_{DSAT}$	Drain current in Saturation mode, also called Drive current
$I_{Gstress}$	Gate leakage current during stress
$I_o$	$I_D$ that defined the onset of $V_{TH}$ for $V_{TCI}$
$I_{off}$	MOSFET off current
$I_{Sub}$	Substrate current
$I_{Sub,max}$	Maximum substrate current
$J$	Current density
$J_c$	Critical current density
$k$ or $k_B$	Boltzmann constant
$k_f$	Interface trap generation rate
$k_r$	Interface trap annealing rate
$L$	Gate length
$m$	Power law exponent in TDDB V-model (Power law)
$mA$	Milli-ampere
$mV$	milli-volt
$M$	Technology scaling factor for dielectric in HCI $I_{Sub}/I_D$ model
$MA$	Mega-ampere

$n$	Time exponent for NBTI degradation
$n^+$	Highly doped n-type region
$n_1$	Sample size (first group) in hypothesis testing
$n_2$	Sample size (second group) in hypothesis testing
$N_0$	Number of electrically inactive Si-H bonds at the interface
$N_a$	Concentration of acceptor
$N_f$	Number of fixed oxide charge
$N_H$	Number of diffusing hydrogen as a function of position and time
$N_{it}$	Number of interface trap charge
$N_{ot}$	Number of oxide trap charge
nm	nanometer
nsec	nano-second
$p$	Electromigration current exponent
$p^+$	Highly doped p-type region
$P$	Fit parameter in HCI $1/V_{DS}$ model
$P_b$	Interface trap state
$q$	Charge of carrier
$R$	Reaction rate
$R_L$	Load resistance
$R_{ref}$	Reaction rate at reference temperature
$S_1^2$	Variance for sample (first group) in hypothesis testing
$S_2^2$	Variance for sample (second group) in hypothesis testing
Si	Silicon
$SiO_2$	Silicon dioxide
$SiO_xN_y$	Silicon oxynitride

$S_p^2$	Pooled variance in hypothesis testing
t	Time
$t_{50}$	Time to 50.0% failure in Lognormal plot
$t_{63.2}$	Time to 63.2% failure in Weibull plot
$t_o$	Test statistic in t distribution
$t_{crit}$	Critical value in t distribution
T	Temperature in Kelvin
$T_{bd}$ or $t_{bd}$	Time to breakdown
$T_{op}$	Operating Temperature
$T_{ox}$	Thickness of the gate oxide
$T_{st}$	Stress Temperature
$V_D$	Drain voltage
$V_{DD}$	Operating voltage
$V_{DS}$	Drain to Source voltage
$V_{DSAT}$	Drain to Source voltage in Saturation mode
$V_{FB}$	Flat band voltage
$V_{G,meas}$	Gate voltage at Measurement level
$V_G$	Gate voltage
$V_{G,stress}$	Gate voltage at Stress level
$V_{nom}$	Nominal operating voltage, 10% higher than $V_{DD}$
$V_{ox}$	Voltage across gate oxide
$V_{SB}$	Source-Substrate voltage
$V_{TCI}$	$V_{TH}$ extracted with Constant Current method
$V_{TCI-0}$	$V_{TCI}$ measured before stress
$V_{TEXT}$	Extrapolated $V_{TH}$ using $G_{m,max}$ method

$V_{TH}$	Threshold Voltage
$V_{TH0}$	Threshold Voltage before stress
$V_{THn}$	nMOSFET Threshold Voltage
$V_{THopt}$	$V_{TCI}$ measured from Optimized $I_D$ - $V_G$ sweep
$V_{THp}$	pMOSFET Threshold Voltage
W	Gate width
x	Distance from Si-SiO <sub>2</sub> interface
X	Exponential parameter in NBTI Lifetime model
CMOS	Complementary MOSFET
DC	Direct Current
EM	Electromigration
ET	Prefix for Pad Group name
FN	Fowler-Nordheim
GNDU	Ground Unit
HCI	Hot Carrier Injection
IC	Integrated Circuit
LOCOS	Local Oxidation of Silicon
MOSFET	Metal-Oxide-Semiconductor Field-effect Transistor
NBTI	Negative Bias Temperature Instability
nMOSFET	n-type MOSFET
Opamp	Operation Amplifier
OTF	On-The-Fly
PGU	Pulse Generator Unit
pMOSFET	p-type MOSFET
RT	Room Temperature

SCE	Short Channel Effect
SCS	Semiconductor Characterization System
SMS	Stress-Measure-Stress
SMU	Source Measure Unit
SS	Subthreshold Slope
STI	Shallow Trench Isolation
TCR	Temperature Coefficient of Resistance
TDDB	Time Dependent Dielectric Breakdown
TTF	Time-to-Fail
WfrFab	Wafer Fabrication plant

# Chapter 1. Introduction

The invention of transistor in 1947 at Bell Labs and followed by the invention of integrated circuit (IC) in the 1950s had set the stage for the booming development in the semiconductor industry, especially in the past several decades. In the quest for improved performance and driving down the cost, the semiconductor industry has been involved in a race to scale down the feature size of devices on the IC.

## 1.1 Background

Metal-Oxide-Semiconductor Field-Effect-Transistor (MOSFET) is the core component in modern IC. Down scaling of MOSFET increases manufacturing process variation and leakage current, and makes the devices less reliable. One of the reliability degradation mechanism that has gained tremendous scientific and industrial interest is Negative Bias Temperature Instability (NBTI). This degradation mechanism primarily affects the p-channel MOSFET (pMOSFET), which is stressed with negative gate voltages at elevated temperatures. The resulting effect is the shift of important pMOSFET parameters, such as threshold voltage or the drain current and degrades the performance of the IC.

With higher packing density, heat generated by the IC during operation increased significantly, aggravating the NBTI phenomenon. This is due to the fact that pMOSFET is found in every logic gate on IC fabricated by Complementary MOSFET (CMOS) process, which is the current mainstream manufacturing technology. The NBTI degradation is further worsen by scaling down the thickness of gate oxide and changing gate oxide from silicon dioxide to silicon oxynitride. Hence, NBTI has become the limiting factor in IC reliability.

## 1.2 Motivation

NBTI Characterization method has been evolving ever since NBTI degradation begin to catch the attention of the industry, partly due to the availability of high-end measurement tools and largely due to the recovery effect of NBTI degradation. Presently, many existing wafer fabrication plants with mature non-leading edge process technology such as  $0.18\ \mu\text{m} \rightarrow 0.13\ \mu\text{m}$  CMOS process, typically run the  $V_{\text{TH}}$  Stability test to assess threshold voltage shift.  $V_{\text{TH}}$  Stability test is run with similar stress bias as NBTI, at similar temperature range, but without interim characterization. In a way it is related to NBTI but does not provide insight on the transistor degradation behavior and does not allow for lifetime extrapolation.  $V_{\text{TH}}$  Stability test requires simple equipment setup and so there is no urgent need for these wafer fabrication plants to perform in-depth NBTI characterization using more sophisticated setup that is able to perform interim high temperature measurement during stress. Very often, this kind of sophisticated setup needs heavy investment in term of cost. However, without an in-depth understanding on the transistor degradation behavior and its impact on transistor parametric performance, it will be difficult for these wafer fabrication plants to improve its manufacturing process.

## 1.3 Objectives

With these motivations in mind, the objectives of this research are as follow:

- i) To establish the transistor parametric measurement and NBTI testing capability in order to perform a characterization (including effect on temperature, channel length and gate oxide thickness) on NBTI degradation behavior and validate the Reaction-Diffusion model, which cannot be achieved through  $V_{\text{TH}}$  Stability test.
- ii) To evaluate the different NBTI characterization methods in minimizing the measurement “dead” time and propose a method that can be implemented on

current equipment set.

## **1.4 Scope of this Research**

The NBTI degradation characterization performs in this research will be based on experimental study. Therefore, the capability of the available equipment set will pose some challenges on this experimental study. For example, the present measurement instrument consists of 3 DC SMUs (Direct Current Source Measure Unit) where as for transistor parametric measurement, typically 4 DC SMUs are required. Nevertheless, this experimental study will still be able to proceed, by using a GNDU (Ground Unit) as a replacement for the 4<sup>th</sup> DC SMU, to be connected to the substrate. This will hinder the measurement of substrate current. The experimental study will focus on conventional DC NBTI stress, that can be supported by the present equipment set. The conventional DC NBTI stress will be applied to the existing test structures (various design of pMOSFETs) on the 0.18  $\mu\text{m}$  dual gate CMOS wafer, to study the dependence of NBTI degradation on temperature, channel length and gate oxide thickness. This research will also attempt to explore other measurement methods to complement the drawback of using conventional DC NBTI stress.

## **1.5 Organization of the Dissertation**

This dissertation is organized into 6 chapters, starts with the introduction, follows by the different type of degradations commonly seen on CMOS IC that include NBTI, their test methods, current understanding of NBTI degradation and various measurement techniques. It then moves on to discuss the systematic methodology in carrying out this characterization study and other measurement methods that will be explored, follows by analysis of results and discussion. The dissertation will end with a conclusion and recommendation on future works. Below is an overview on each chapter:

- Chapter 1 is the introduction of the dissertation, explaining the motivation in carrying out this research. The objectives are shown in this chapter, together with the organization of the dissertation.
- Chapter 2 provides an overview on the key degradation mechanisms commonly seen on CMOS IC. This includes the degradation mechanisms related to the transistor, such as Time Dependent Dielectric Breakdown (TDDB), Hot Carrier Injection (HCI) and Negative Bias Temperature Instability (NBTI). Also shown is the degradation mechanism related to the metal interconnection, Electromigration (EM).
- Chapter 3 details the common methods used to characterize those degradation mechanisms described in Chapter 2. It starts with the methods to measure the transistor threshold voltage in linear mode of operation and drain current in saturation mode of operation. This is followed by the  $I_{\text{Sub,max}}$  stress for HCI, Charge Pumping to characterize interface states, Constant Voltage stress for TDDB, Constant Current stress for EM and the various methods used for NBTI that include Conventional DC Stress-Measure-Stress, On-The-Fly, Fast Switching, Ultra Fast  $V_{\text{TH}}$ , Fast  $V_{\text{G}}$  Ramp and Ultra Fast Switching.
- Chapter 4 proposes a systematic methodology in performing an extensive characterization on NBTI degradation. A flow chart is used to explain the necessary experiments in this research. It describes the considerations in selecting the suitable test structures for NBTI stress, the approach in correlating the existing measurement system against the industry production tester and also the correlation results. It ends with the experimental plan with the appropriate biasing conditions in performing the extensive set of experiments and also explains the proposed fast measurement method.
- Chapter 5 shows the experimental results and analysis methods on the data. The

NBTI degradation is explained in various aspects, justifies with semiconductor device physics and published results in the literature. Also shown is the proposed fast measurement method and discussion on the significance of the results from this research to the industry.

- Chapter 6 is the conclusion of this dissertation, highlighting the results that meet the objectives that had been set in Section 1.3. The recommendations for future work is also being discussed in this chapter.

## **Chapter 2. Degradation in CMOS**

### **2.1 Introduction**

Complementary Metal Oxide Semiconductor (CMOS) has become the dominant technology in the electronic industry for the past several decades, and continues to dominate in the coming years. The scaling of CMOS technology into deep sub-micron regimes has brought about new reliability challenges in MOSFET device. These key reliability challenges are Hot Carrier Injection (HCI), Negative Bias Temperature Instability (NBTI), Time-Dependent Dielectric Breakdown (TDDB) and Electromigration (EM), which can pose a limit to the device scaling, and cause circuit performance degradation. These are intrinsic wear out mechanisms that is inherent in the design and materials used and will be reviewed in this chapter, with the emphasis being given to NBTI.

### **2.2 Hot Carrier Injection**

Hot carrier injection describes the degradation of intrinsic MOSFET device characteristics due to the trapping of charge in the gate dielectric. Hot carriers are those carriers (electrons or holes) that are not in thermal equilibrium with the rest of the semiconductor crystal lattice. The energy of these carriers does not correspond to that of the conduction band and valence band. This situation arises when the electric field seen by the carriers is sufficiently high to change their average energy, causing the energized carriers being accelerated between the collisions with the lattice (Shah, 1992).

In metal oxide semiconductor field effect transistor (MOSFET) devices, electric fields are utilized to control the flow of electrons or holes through a device. The electric fields within the device increase with the continuous reduction in device dimension, while the operating voltage reduction does not follow in tandem. This is the approach of

constant voltage scaling, popularly used by the industry. The source-to-drain electric field, specifically the channel region near the drain has been increased in every successive generation of technology. The result of this increase in the number of energized carriers and high electric field leads to:

- Impact ionization and avalanche multiplication near the drain region, resulting in increased drain current and substrate current.
- Injection of carriers into the gate dielectric near the drain region, resulting in gate current.

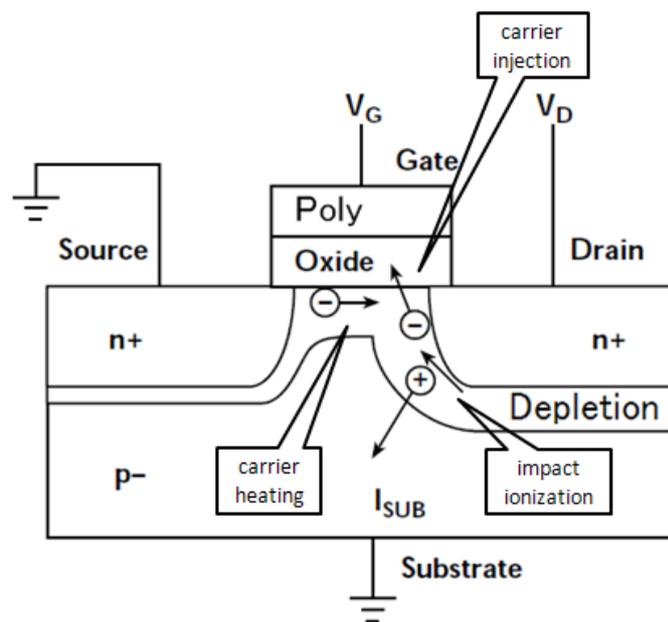


Figure 2.1: Hot carrier injection mechanism for nMOSFET (Keithley, 2000).

The injected hot carrier into the gate dielectric may become trapped in the bulk of the gate dielectric, in the interface of Si-SiO<sub>2</sub> and/or within the sidewall spacer. These trapped carriers disturb the normal distribution of electric fields within the device, which results in a shift of the threshold voltage, device on-resistance (transconductance) and driving capability (saturation current). Over time, as more and more carriers are trapped, degradation of the device characteristics can be significant, and impact overall

circuit performance (Dikmen et al., 1994).

### 2.2.1 Modeling of Hot Carrier Injection

There are three models commonly used, with each model has equal validity and are based on the same carrier heating mechanism, where the probability of a carrier achieving enough energy to cause device damage is directly related to the exponent of the lateral electric field. These three models are: Substrate/drain current ratio model, Substrate current model and Drain-source voltage acceleration model (JESD28-1, 2001).

The substrate current ( $I_{Sub}$ ) is a direct measure of the impact ionization in the drain region, and is hence used as a measure of carrier heating. Hot carrier injection can be modeled as a function of the ratio of the substrate current to drain current (Hu et al., 1985), which is expressed in this form:

$$\tau = M (I_{Sub})^{-2.9} (I_D)^{1.9} (\Delta V_{TH})^{1.5} W \quad (2.1)$$

where  $\tau$  is the device lifetime defined as the stress time required to achieve a targeted threshold voltage shift,  $I_{Sub}$  is substrate current,  $I_D$  is drain current,  $V_{TH}$  is the threshold voltage,  $W$  is the channel width, and  $M$  is the dielectric technology scaling factor. This relationship is derived from fundamental device physics and is calibrated using empirical data. However, for current sub-micron MOSFET, Equation (2.2) is used more commonly (JESD28-1, 2001) in the industry ( $I_{Sub}/I_D$  model):

$$\tau * I_D = M \left( \frac{I_{Sub}}{I_D} \right)^{-G} W \quad (2.2)$$

where  $G$  and  $M$  need to derive from experimental data.

Another model that uses substrate current is the Substrate current model, where the device lifetime is proportional to the normalized  $I_{Sub}$  with  $W$ , so that the degradation seen by the transistor is independent of the device width ( $I_{Sub}/W$  model).

$$\tau = C \left( \frac{I_{Sub}}{W} \right)^{-B} \quad (2.3)$$

where  $C$  and  $B$  are fit parameters that can be determined from experimental data.

The third model is Drain-source voltage acceleration model ( $1/V_{DS}$  model). It is based on the carrier heating that is due to the voltage applied on the drain and it is also related to the substrate current changes (Takeda et al., 1983).

$$\tau = A \exp\left(\frac{P}{V_{DS}}\right) \quad (2.4)$$

where  $A$  is a constant and  $P$  is fit parameter that can be determined from experimental data.

### 2.3 Time-Dependent Dielectric Breakdown

Time-dependent dielectric breakdown is a failure mechanism which occurs in thin dielectric film, which is the gate oxide of MOSFET. TDDB is a wear out mechanism of gate oxide, under the influence of sufficiently high applied electric field and over a period of time, the gate oxide will breakdown and irreversibly damage. Inevitably, the gate oxide breakdown will depend on the defects introduced during the manufacturing process. Therefore, TDDB is usually performed on area intensive capacitor to assess the intrinsic quality of the gate oxide. Also, gate-edge intensive (poly finger) capacitor and field-oxide edge intensive (LOCOS/STI finger) capacitor will be used to assess the interaction of gate poly process and isolation (LOCOS/STI) process with gate oxide (JESD92, 2003).

The TDDB time-to-fail is dependent on the rate of redistribution of the electric field, where it is the time taken to reach the critical electric field ( $E_{crit}$ ). The  $E_{crit}$  is usually related to the intrinsic quality of gate oxide. For the breakdown voltage,  $BV_{ox}$ , it is closely correlated to  $E_{crit}$  if the gate oxide is defect free. However, the presence of a defect can act to focus and enhance local electric field, thereby reducing the apparent  $BV_{ox}$ , but not the  $E_{crit}$  value. Hence,  $BV_{ox}$  is often used in the wafer fabrication as a quick way to gauge the defectivity level of gate oxide.

The generic characteristic of TDDB is the dielectric fails when a conductive path forms in the dielectric, leading to a short between anode and cathode (namely a runaway phenomenon). This is usually accelerated by elevated temperature and under either constant voltage or constant current stress. When sufficient current is flowing between anode and cathode to cause Joule heating, the conducting path is formed (Figure 2.2). There are four models being used by the industry for  $SiO_2$  based gate dielectric, developed on the basis of large amount of experimental data over the years. These four models are based on the assumption of two mechanisms: field-driven and current-driven.

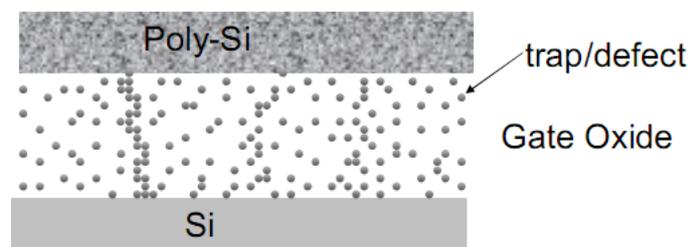


Figure 2.2: Traps and defects formed a conductive path across the gate oxide

(Vollertsen, 2004).

### 2.3.1 E-Model or Constant Field/Voltage Acceleration Exponential Model

The E-Model is based on thermochemical model for oxide breakdown (McPherson & Baglee, 1985) under the assumption of field driven mechanism for gate oxide thickness  $> 4$  nm. This model describes the oxide breakdown is due to coupling between the applied electric field ( $E_{ox}$ ) with the oxide dipole moments (which is the field enhanced bond breakage at the Si-SiO<sub>2</sub> interface). The  $E_{ox}$  serves to reduce the activation energy required for thermal bond breakage and therefore exponentially increases the reaction rate for failure. It predicts that the time to breakdown ( $T_{bd}$ ) has an exponential linear dependence on applied electric field (JEP122F, 2010).

$$T_{bd} = A_0 \exp(-\gamma E_{ox}) * \exp\left(\frac{E_a}{k_B T}\right) \quad (2.5)$$

where  $A_0$  and  $\gamma$  (field acceleration factor) are temperature dependent fit parameters.

### 2.3.2 1/E-Model or Anode Hole Injection Model

The 1/E-Model is based on anode hole injection model (Moazzami et al., 1989; Degraeve et al., 1995) under the assumption of current driven mechanism. In this model, damage is assumed to be due to current flow through the dielectric due to Fowler-Nordheim (FN) conduction. Electrons inject from the cathode is accelerated to anode, and resulting in impact ionization that produces holes which in turn may tunnel back to dielectric at anode (hot hole anode injection mechanism). Since both electrons (from the cathode) and the holes (from the anode) are the result of FN conduction, then the time to breakdown is expected to show an exponential dependence on the reciprocal of the electric field,  $1/E_{ox}$ .

$$T_{bd} = A_0 \exp\left(\frac{D}{E_{ox}}\right) \quad (2.6)$$

where  $A_0$  and  $D$  (field acceleration factor) are temperature dependent fit parameters.

### 2.3.3 V-Model (exponential with voltage)

The V-Model (exponential) is based on the E-Model and it is for gate oxide thickness  $< 4$  nm. Nicollian et al. (2000) showed experimental data for  $T_{ox} = 2.7$  nm to be fitting better with exponential V-Model than E-Model. Similarly, Alam et al. (2000) found that at low stress voltage, the hole generation process is dominated by the impact ionization process which is a strong function of stress voltage. A model of exponential with voltage rather than electric field represents the reliability performance well. Therefore, the time to breakdown ( $T_{bd}$ ) is expressed as:

$$T_{bd} = A_0 \exp(-\beta V) * \exp\left(\frac{E_a}{k_B T}\right) \quad (2.7)$$

where  $A_0$  and  $\beta$  (voltage acceleration factor) are temperature dependent fit parameters.

### 2.3.4 V-Model (Power law)

The power law V-Model is based on analysis of breakdown data that suggests the breakdown mechanism is a voltage-driven process as oxide thickness is decreased ( $< 2$  nm). It is therefore appropriate to analyze the breakdown data in terms of  $V_{ox}$  dependent and not  $E_{ox}$ . This is due to the direct tunneling of current through the oxide in the ballistic transport manner (no scattering or energy loss). The amount of energy delivers to the anode is (electron)  $\times$  (V).

$$T_{bd} = B_0 V^{-m} \quad (2.8)$$

where  $m$  is power law exponent which is independent of thickness and  $B_0$  is the prefactor (Wu et al., 2000).

## 2.4 Electromigration

Electromigration is a failure mechanism which occurs in the interconnect metallization lines. Electromigration is a diffusion process influenced by the current density carried

by the interconnect lines. It is a momentum exchange between current carrying electrons and host metal lattice, when the current is sufficiently high to cause drift (diffusion) of metal atoms in the direction of the electron flow (Figure 2.3). The number of metal atom that migrates in the “electron wind” (flux density), is dependent on:

- The magnitude of forces that tend to hold the atoms in place, and thus the elemental nature of the conductor, crystal size and grain boundary chemistry.
- The magnitude of forces that tend to dislodge the atoms, including the electron current density, temperature and mechanical stresses.

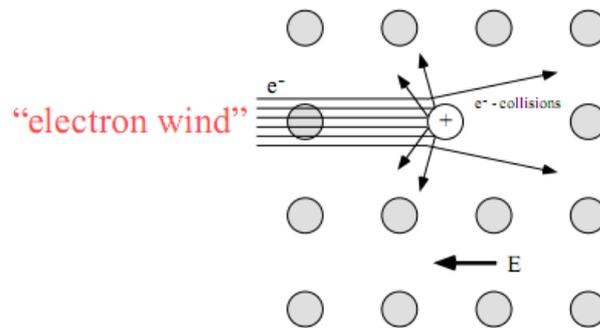


Figure 2.3: Drift (diffusion) of metal atom under the influence of “electron wind” (Ho, 2004).

The aluminum interconnect lines used in IC is polycrystalline in nature, and therefore is prone to electromigration due to the relatively loose binding of the aluminum atoms, especially at the surface of the individual crystal grains. Electromigration failures occur if a void, created at a point where the flux of the outgoing ions exceeds the incoming flux, becomes large enough to cause an open in the metal line (Kraft et al., 1997). On the other hand, when aluminum atoms are piled up at a point (hillock) where the incoming ion flux exceeds the outgoing flux, a metal line will short to the adjacent or overhead metal lines (Figure 2.4).

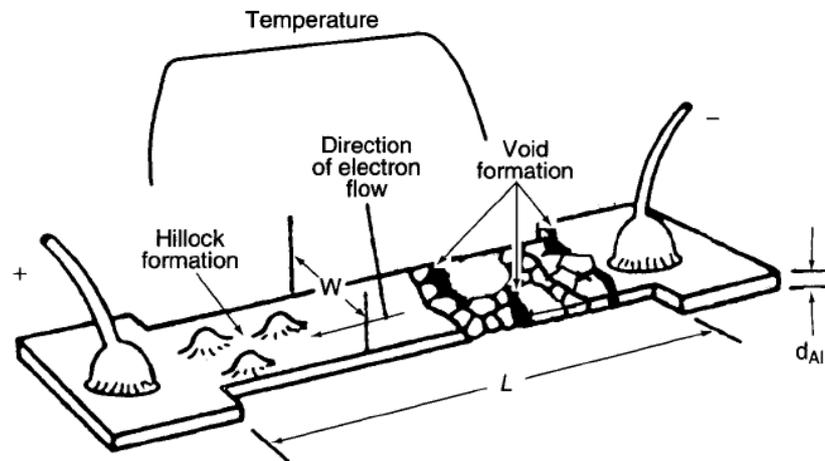


Figure 2.4: Formation of void and hillock as a result of electromigration. (Ohring, 1998)

All IC that use polycrystalline aluminum metallization are susceptible to electromigration failures. The older generation of CMOS IC is less affected by this failure mechanism, as the maximum operating current in each metal line is inherently lower. With the continue scaling down of feature size into deep sub-micron regions, and couple with operating frequencies reaching into GHz range, reliability of modern CMOS devices is critically impacted by electromigration. Two key factors affecting electromigration behavior are temperature and current density (Ohring, 1998). Temperature may often be controlled after fabrication through the use of packages and heat sinks. On the other hand, current density must be controlled at the design stage. IC designer often needs to carry out thorough preventive analysis in order to map out potential locations within the design that carries high current density. These “hot spot” locations may need to be redesigned in order to lower the reliability risk.

Electromigration can not be screened out through production testing, as this is an intrinsic wear out mechanism. Electromigration usually does not manifest itself until an IC has been operated for months, or even years. Therefore, deep sub-micron IC will be failures waiting to occur, unless the problem is eliminated during the design phase and

also taking into account the defects introduced in wafer fabrication process that may induce electromigration mechanism.

#### **2.4.1 Modeling of Electromigration**

The modeling of electromigration is based on the mass transport mechanism of metal atom under the influence of temperature and current. Historically, the mean time to failure for a group of metal test lines under the same stress condition can be modeled by this modified Black's equation (Filippi et al., 1993):

$$t_{50} = \frac{A}{(J - J_c)^p} * \exp\left(\frac{E_a}{k_B T}\right) \quad (2.9)$$

where  $A$  is a constant,  $J$  is current density in metal test lines,  $J_c$  is the critical current density below which no electromigration take place,  $p$  is the model parameter for current density (current exponent).

#### **2.5 Negative Bias Temperature Instability**

Negative Bias Temperature Instability (NBTI) is a wear out mechanism experienced by pMOSFET under inversion for a prolonged period of time at either room or elevated temperature. The damage created may lead to substantial pMOSFET parameters shift, that caused an absolute increase in threshold voltage and reduction in mobility, drain current and transconductance. It becomes one of the most critical device reliability issues as the technology advances with scaling and introduction of new material.

Both the negative gate voltages and/or elevated temperature can generate NBTI effect, and a combination of both leads to a larger magnitude of device characteristics shift. NBTI occurs primarily in pMOSFETs with negative gate voltage bias and appears to be negligible for positive gate voltage. For nMOSFETs, the parameters shift is

negligible for either positive or negative gate voltages (Makabe et al., 2000). In CMOS circuits, NBTI occurs most commonly when the pMOSFET in the logic gate of inverter is at ground state (Figure 2.5). This logic state can be very common when the IC is in standby and NBTI effect can manifest itself for long period of time. The gradual shift in device characteristics can lead to timing shifts and increases the spread in signal arrival in logic gates, thereby causing circuit failures either in product testing or field application.

The following sections will review a brief history of NBTI, the models that had been used to explain this phenomenon and also recent research done on this wear out mechanism.

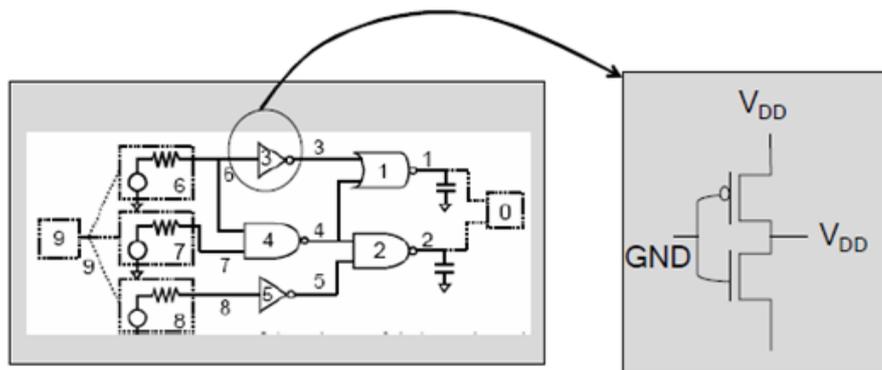


Figure 2.5: CMOS circuit in logic gate representation, showing pMOSFET in inverter gate under negative gate bias with reference to source and drain (Alam, 2005).

### 2.5.1 Brief History of NBTI

NBTI has a complicated history and first observed in the 60s by Deal et al. (1967), when MOSFET devices were subjected to negative gate bias at elevated temperature (300-450 °C), with both interface trap charge ( $N_{it}$ ) and oxide trap charge ( $N_{ot}$ ) showing positive charge build-up at the Si-SiO<sub>2</sub> interface. Later in the 70s, Breed (1975) reported relaxation phenomenon in MOSFET devices under bias temperature stress, which was ascribed as hole trapping/detrapping. This is followed by Jeppson and Svensson (1977)

that modeled NBTI in Reaction-Diffusion framework, where reaction at Si/SiO<sub>2</sub> interface forms N<sub>it</sub> and N<sub>ot</sub> and presumably a hydroxyl group which diffuses away.

However, the NBTI issue “disappeared” when nMOSFET was adopted as the dominant manufacturing technology during the 80s. When the industry moved to CMOS as the preferred manufacturing technology during the late 80s, the industry and research community were focusing on dealing with HCI issue on nMOSFET. Comparatively, the pMOSFET NBTI phenomenon is not as serious as the nMOSFET HCI, due to the fact that pMOSFET is a buried channel device during this period of early adoption of CMOS (Taur and Ning, 1998). Nevertheless, Blat et al. (1991) reported NBTI results on pMOSFET and showed the involvement of water related species in transistor parameters degradation under NBTI stress.

When the industry continue to scale down the transistor feature length during the late 90s, the need to maintain the threshold voltage of the pMOSFET due to the short channel effect results in using p<sup>+</sup> polysilicon as the gate material for pMOSFET. This changed the CMOS process technology from n<sup>+</sup> poly gate to n<sup>+</sup>/p<sup>+</sup> poly gate (Taur and Ning, 1998). As the need to lower IC operating power, voltage scaling is used with the scaling down of gate oxide thickness, which reduces HCI but increases the electric field and temperature (heat from dissipated power during IC operation). This in effect reintroduces NBTI (Kimizuka et al., 2000). The NBTI phenomenon is further aggravated by the fact that the pMOSFET becomes a surface channel device (due to the change to p<sup>+</sup> polysilicon gate, Taur and Ning (1998)), resulting in higher probability of interaction of the carriers with the Si-SiO<sub>2</sub> interface. Not only is digital CMOS being affected, analog transistor fabricated on CMOS process is being affected as well (Thewes et al., 1999).

As the industry went on to further down scaling as per Moore's Law (Moore, 1975), the gate oxide started to suffer high leakage at about 2.2 nm (moving into direct tunneling regime), which is at technology node with gate length lower than 0.18  $\mu\text{m}$ . With that, the industry moved on to introduce lightly nitrided oxide for logic devices, not only to contain the gate leakage issue, but also to prevent boron penetration from the p+ polysilicon gate as well. However, adding nitrogen made NBTI effect getting worse, as shown by Kimizuka et al. (2000). Since then, the interest on NBTI has grown, as the partial recovery mechanism has complicated the measurement and data analysis.

### 2.5.2 Interface Trap Charge

Interface trap plays a key role in NBTI degradation and its characteristic is reviewed in this section. The generation of interface trap charges is an artifact of the oxidization process. The oxygen provided via steam bonds with the silicon as it is exposed to the crystal, but due to differences in lattice constants and bonding orientations, there will be trivalent Si atoms at the Si-SiO<sub>2</sub> interface with unpaired valence electrons in a dangling orbital. These dangling orbital states are often referred to as P<sub>b</sub> centers (Schroder and Babcock, 2003). Different crystal orientations give rise to different types of interface traps, and a graphical representation of possible orientations is shown in Figure 2.6.

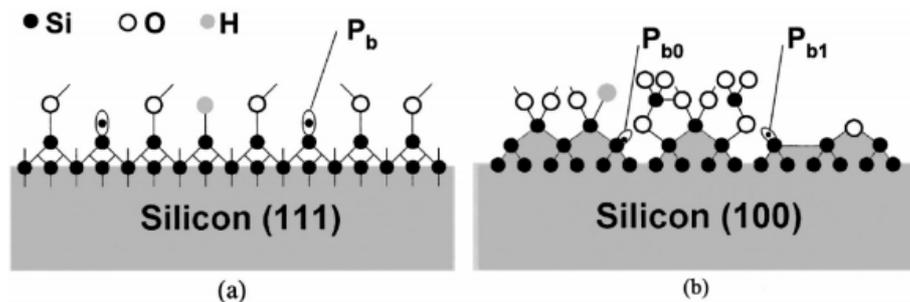


Figure 2.6: Interface states on the Si-SiO<sub>2</sub> interface (Schroder and Babcock, 2003).

These  $P_b$  states have an amphoteric nature which means the dangling bonds can be occupied by zero, one or two electrons. In a way, the same defect can be positively charged, neutral or negatively charged. Those  $P_b$  states that are electrically active (positive or negative), provide a trap that can capture holes or electrons. In other words, the effect of these mechanisms creates a distribution of possible states that can be occupied in the exclusion regions of the band gap. In the upper part of the band gap, the traps act as acceptors so that they will be negative when filled and neutral when empty (Pierret, 1996). In the lower region of the band gap, the traps act as donors. As such, they will be neutral when filled and positive when empty (Pierret, 1996). Figure 2.7 shows the representation of the interface states in the band gap diagram and the charge contribution to the device.

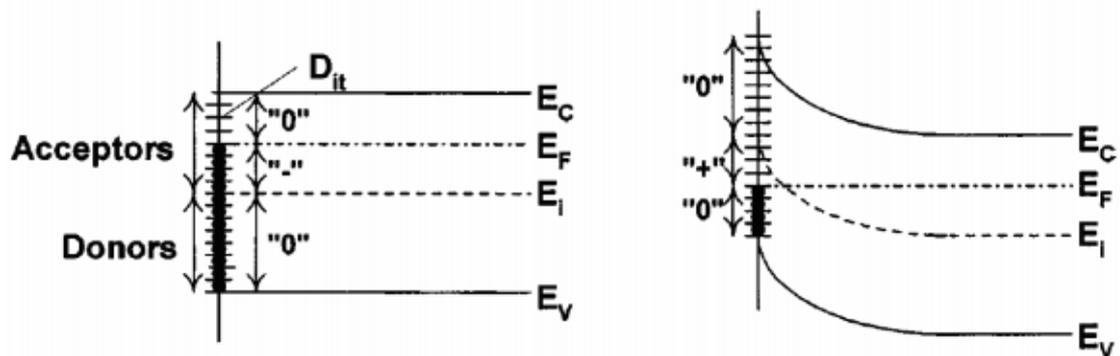


Figure 2.7: Interface states in the band diagram (Schroder and Babcock, 2003).

During subsequent process steps (high temperature anneal), these dangling bonds are generally passivated by hydrogen atoms, which create Si-H bonds at the Si-SiO<sub>2</sub> interface. The Si-H bonds will improve the parametric performance of transistor during the initial stage of operation. However, as Si-H bonds are weak, they may be broken during the operating lifetime of the transistor, which in turn lead to a degradation of its parameters.

### **2.5.3 Mechanism and Modeling of NBTI**

After first being studied by Deal et al. (1967), many models have been proposed for the physical mechanisms on NBTI. Among those models, holes injected into the oxide, tunneling electrons and electrochemical reactions were the main subjects.

#### **2.5.3(a) Hole Trapping Model**

The hole trapping model is based on avalanche hole injection measurements on unstressed MOS capacitors with the NBTI stress (Haywood et al., 1985; Lu et al., 1988). This model proposes that the negative mid-gap voltage shift, which is believed to be a measure of the change of positive oxide charge without the contribution from interface states, is due to the filling of intrinsic hole traps. All the positive charge generated by preceding negative bias stress can be removed by the positive bias stress. But the exact mechanism for hole injection into the oxide is still unknown.

#### **2.5.3(b) Thermally Assisted Electron Tunneling Model**

The thermally assisted electron tunneling model was established by Breed (1975). According to this model, the neutral or positive centers which cause the charge trapping, are located near the interface of Si-SiO<sub>2</sub>. Under negative bias stress, these charge trapping centers are excited. The electrons in the excited states then tunnel into empty states of the conduction band of the silicon and become available for conduction as leakage current. These empty states are positively charged, thereby changing the parametric characteristic of the transistor. Therefore, this process is a thermally assisted tunneling process.

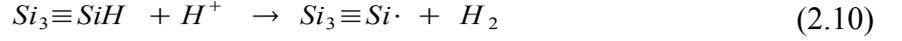
### 2.5.3(c) Reaction-Diffusion Model

The Reaction-Diffusion model is the most prevalent model used by many researchers. This model explains the NBTI effect in terms of electrochemical reactions. Due to the crystal mismatch, the oxide that grows upon the silicon crystal leaves traps for holes or electrons along the Si-SiO<sub>2</sub> interface. However hydrogen is introduced into the interface areas of the transistor during annealing, which is one of the process step in the fabrication of gate stack. The hydrogen bonds with the available dangling silicon bonds and thus effectively passivated the traps. Most experimentally observed data indicates that the instability arises as a result of a chemical reaction of the hydrogen at the interface and a subsequent diffusion of it through the oxide (Alam et al., 2007). As the hydrogen disassociates from the silicon bonds (Si-H), the interface traps are exposed. The shift in the threshold voltage ( $\Delta V_{TH}$ ) is dependent on both the stress time and temperature, but it specifically relates to the diffusion of the hydrogen from the Si-SiO<sub>2</sub> interface. It is also suggested that similar effects can be observed from the breaking of silicon and oxygen bonds in the oxide.

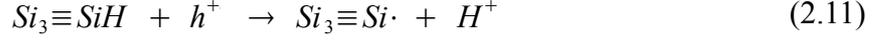
This forms the basic framework for Reaction-Diffusion model. The reaction portion of the model deals with the generation of interface states through the interaction of channel holes and the applied stress (electric field). The other portion relates to the transport of the hydrogen from the interface. There are multiple methods to describe the reaction and diffusion model. Different approaches use different species of hydrogen ( $H^+$ ,  $H^o$ ,  $H_2$ ), and there is still a decent amount of controversy over which one is correct.

The generation of traps can occur in multiple ways. For the generation of the neutral  $H_2$ , a positively charged hydrogen atom will react with a hydrogen bonded to a silicon dangling bond. The reaction is shown in the following equation (Schroder and

Babcock, 2003):



The tetrahedrally bonded silicon with the hydrogen termination is de-passivated with the  $H^+$  (attributed as proton). A source of  $H^+$  can involve trapping holes ( $h^+$ ) from the channel with the following equation (Schroder and Babcock, 2003):



In addition, there may be another reaction that can lead to the generation of neutral  $H_2$ . This involves the creation of neutral  $H^0$  as it disassociates in the electric field, and is shown in the equation below:



where  $H^0$  is a neutral interstitial hydrogen atom or atomic hydrogen (Schroder and Babcock, 2003). Other contributions can be added from fixed charge interactions, but the previous equations represent the majority of reactions involving hydrogen species that contributed to the Reaction-Diffusion model.

If the main hydrogen species for NBTI mechanism are neutral, their flux will be governed by diffusion. This occurs due to a difference in concentration and is related to the mobility of hydrogen through the different materials. As they depart from the interface, they leave behind a concentration of interface states. This process at the interface is modeled by a rate equation as (Jeppson and Svensson, 1977):

$$\frac{\partial N_{it}(t)}{\partial t} = \underbrace{k_f(N_0 - N_{it}(t))}_{\text{generation}} - \underbrace{k_r N_{it}(t) N_H(0,t)^{\frac{1}{a}}}_{\text{annealing}} \quad (2.13)$$

where  $N_{it}$  is the number of interface state,  $k_f$  is the interface trap generation and  $k_r$  is the annealing rate. The symbol  $N_0$  denotes the number of electrically inactive Si-H bonds at the interface and  $N_H(0, t)$  is the surface concentration of the diffusing species. The value

of  $a$  gives the order of the reaction. In the original publication by Jeppson & Svensson (1977), neutral hydrogen,  $H^0$  was proposed, which is obtained with  $a = 1$ . For molecular hydrogen ( $H_2$ ),  $a = 2$ .

The first term (generation) on the right hand side of Equation (2.13) denotes the forward reaction (i.e. breaking of bonds to increase the trap density) has been set proportional to the density of the unbroken bond. The second term (annealing) denotes the reverse reaction (restoration of the bond) had been set proportional to the trap density (broken bond density) and the available hydrogen at the interface.

On the other hand, the hydrogen species may diffuse (or drift) away from the interface as follow (Alam et al., 2007):

$$\frac{\partial N_{it}}{\partial t} = -D_H \frac{\partial N_H(x, t)}{\partial x} + N_H(x, t) \mu_H E_{ox} \quad (2.14)$$

where  $D_H$  is the diffusion constant of hydrogen,  $N_H$  is the number of hydrogen as a function of position and time,  $\mu_H$  is the mobility of hydrogen and  $E_{ox}$  is the electric field in the oxide. The hydrogen atom may diffuse with diffusion constant ( $D_H$ ) if the atoms are neutral, or drift with mobility  $\mu_H$  if they are charged.

The Reaction-Diffusion model assumes that  $N_{it}$  generation is the primary cause of NBTI. Typically the rate of trap generation in Equation (2.13) is usually small compares to the dissociation rate of Si-H bond and the annealing rate. With  $N_{it} \ll N_0 \sim 5 \times 10^{12} \text{ cm}^{-2}$ , Equation (2.13) can be written as (Alam et al., 2007):

$$\left( \frac{k_F N_0}{k_R} \right) \approx N_H(0, t) N_{it} \quad (2.15)$$

Also, Equation (2.14) may be restated as a conservation equation which requires that the number of broken Si-H bonds equal that of total H concentration in the gate stack.

Therefore, the number of interface states  $N_{it}$  can be calculated from the total amount of hydrogen produced using the following relationship (Alam et al., 2007):

$$N_{it}(t) = \int_0^{x(t)} N_H(x, t) dx \quad (2.16)$$

where  $x = 0$  is defined at the Si-SiO<sub>2</sub> interface and  $x(t)$  defines the tip of the diffusion or drift front (Figure 2.8).

The current explanation to the diffusion assumes production of neutral hydrogen at the interface that combines to form molecular hydrogen. From this point the hydrogen concentration decreases in a linear fashion as the distance increases from the interface.

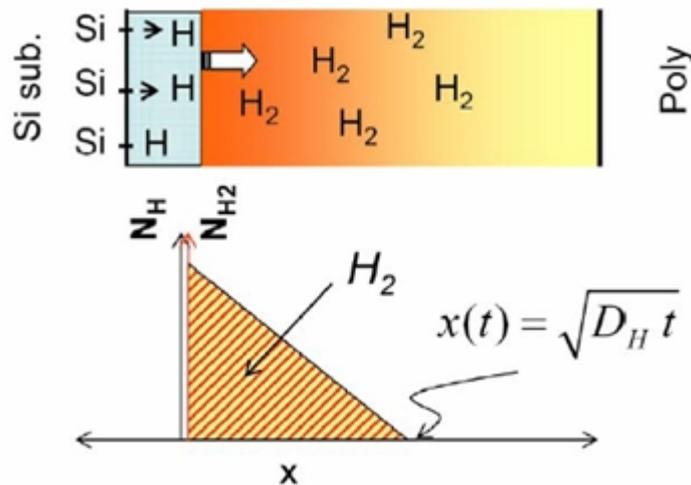


Figure 2.8: Diffusion of hydrogen (Alam et al., 2007).

Due to the fact that these are both neutral species as mentioned before, the drift portion of Equation (2.14) is removed. The actual number of interface states is then calculated using the diffusion coefficient of hydrogen,  $D_H$ . The diffusion distance at a given time  $t$  is  $x(t) \sim (D_H t)^{1/2}$ . Therefore the right hand side of Equation (2.16) can be interpreted as the area under the triangle (sometimes referred to as 'Triangle Method'). This yields the following equation (Alam et al., 2007):

$$N_{it}(t) = \int_0^{\sqrt{D_H t}} N_H(x, t) dx = \frac{1}{2} N_H(0, t) \sqrt{D_H t} \quad (2.17)$$

By inserting the expression in Equation (2.17) into Equation (2.15), the number of interface states becomes (Alam et al., 2007):

$$N_{it}(t) = \sqrt{\frac{k_F N_0}{2 k_R}} (D_H t)^{\frac{1}{4}} \quad (2.18)$$

Equation (2.18) demonstrates the dependence of interface states density with the diffusion of hydrogen species with the exponent on the time value,  $n = 1/4$ . This is what can be used to track the deterioration of the device as the stress time increases. The above analytical solution of Reaction-Diffusion model matches many of the recent published experimental data on NBTI degradation with an exponent of  $n = 1/4$  (Kimizuka et al., 2000; Schroder and Babcock, 2003; Alam et al., 2007). It interprets the prevailing view that NBTI degradation is characterized by diffusion of atomic hydrogen in the gate dielectric. In addition, Equation (2.18) follows the general form of the equation that relates the activation energy of degradation ( $E_a$ ), and the temporal component to the number of interface states. The form is shown in the following equation (Schroder and Babcock, 2003):

$$N_{it}(t) = A e^{-\frac{E_a}{k_B T}} t^n \quad (2.19)$$

This equation, displays in a logarithmic plot, shows the change in the interface states with time. This information is easily relating to the change in the threshold voltage and also to other circuit parameters. This also indicates that the device degradation in time can be magnified with higher temperatures.

#### **2.5.4 Recovery of NBTI**

Under the NBTI stress, interface traps are created at the Si-SiO<sub>2</sub> interface, where the Si-H bonds at this interface become broken during the stress (as explained in detail in 2.5.3). The hydrogen atoms that are released during the Si-H breakage tend to drift away from the Si-SiO<sub>2</sub> interface into the SiO<sub>2</sub> bulk, especially when there is bias on the gate. When the negative gate voltage is removed, some of the released hydrogen atoms in the bulk of SiO<sub>2</sub> (at higher concentration) are expected to diffuse back to the Si-SiO<sub>2</sub> interface. Thus a fraction of NBTI degradation can be recovered by annealing at high temperature with the negative gate bias removed (Ershov et al., 2003). Complete recovery is not expected to take place because some of the released hydrogen atoms may undergo a reduction reaction while in the SiO<sub>2</sub> bulk, unless a positive gate bias is applied to induce a complete recovery.

## Chapter 3. Measurement Methods of CMOS Degradation

### 3.1 Introduction

CMOS degradation is evaluated by using discrete test structures such as MOSFET, capacitor or metal line (Bordelon et al., 1999). To evaluate the degradation, key parameters that represent the performance of the test structure will be measured. These key parameters (drain saturation current, threshold voltage) will be measured during the stress or when the stress is stopped temporarily. The following sections will review the measurement of these parameters and the key reliability stress methods, with the emphasis being given to NBTI stress measurement.

### 3.2 MOSFET Saturation Drain Current and Threshold Voltage

Saturation drain current and threshold voltage are the key parameters for evaluating the performance of MOSFET. Threshold voltage is the point when the transistor is turn on, where the channel inversion layer transits from weak to strong inversion. As the inversion layer of the channel is formed, allowing current to flow from drain to source, the MOSFET operates like a resistor (ohmic), and the drain current is controlled by the gate voltage relatively linear to both the source and drain voltages (Figure 3.1).

Typically, the drain current in linear region is expressed as (Taur and Ning, 1998):

$$I_D = \mu C_{ox} \frac{W}{L} [(V_G - V_{TH})V_D - \frac{V_D^2}{2}], \quad \text{for } V_D < V_G - V_{TH} \quad (3.1)$$

From Equation (3.1), with small  $V_D$ , the term  $V_D^2/2$  is negligible and  $I_D$  will be controlled by  $V_G$  linearly.

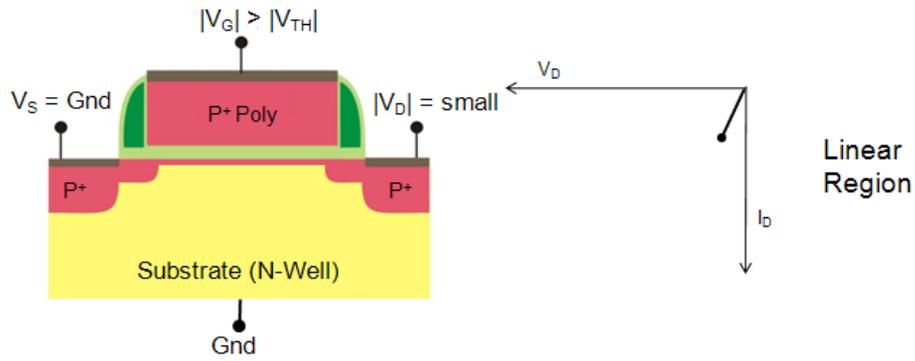


Figure 3.1: pMOSFET operated in the linear region (low drain voltage) (Acuthan and Bhat, 2007).

The linear region will continue to extend further when the  $V_D$  increases, to a point when the inversion layer is pinched off at the drain side (Figure 3.2). This happens when the difference between the gate voltage and the drain voltage is equal to the threshold voltage ( $V_D = V_G - V_{TH}$ ). With the magnitude of the drain voltage moving higher, the current in the channel is saturated, which means the drain current is now weakly dependent upon drain voltage and controls primarily by the gate voltage ( $V_G$ ).

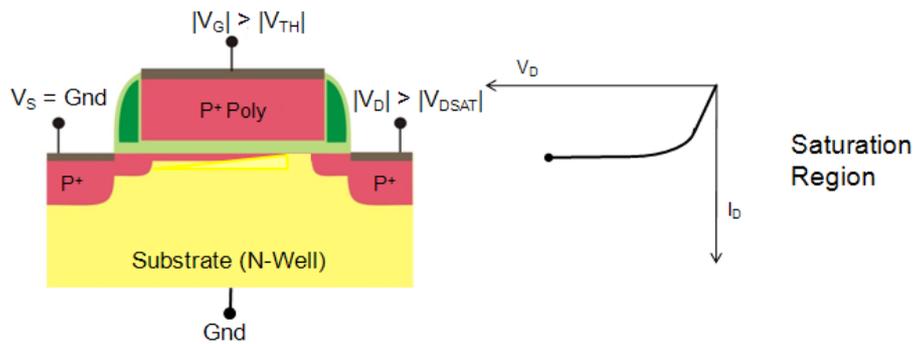


Figure 3.2: pMOSFET operated in saturation region, after pinched off (Acuthan and Bhat, 2007).

Based on  $V_D = V_G - V_{TH}$  and Equation (3.1), the saturation drain current,  $I_{DSAT}$  can be expressed as below:

$$I_{DSAT} = \mu C_{ox} \frac{W}{L} \left[ \frac{(V_G - V_{TH})^2}{2} \right], \quad \text{for } V_D > V_G - V_{TH} \quad (3.2)$$

Equation (3.2) shows that the  $I_{D_{SAT}}$  is not dependent on  $V_D$  but is controlled by  $V_G$ . Therefore, once the  $I_D$  reaches the saturation state, it will remain constant when  $V_D$  increases, as shown in Figure 3.2.

### 3.2.1 Threshold voltage measurement

Threshold voltage is commonly being written in this form (Taur and Ning, 1998), due to its ease of incorporation into analytical solution:

$$V_{TH} = V_{FB} + 2\Phi_F + \frac{\sqrt{2\varepsilon_s q N_a (2\Phi_F + V_{SB})}}{C_{ox}} \quad (3.3)$$

where  $V_{FB}$  is the flat band voltage,  $V_{SB}$  is the source-substrate voltage,  $\varepsilon_s$  dielectric constant of silicon,  $q$  is charge of carrier,  $N_a$  is the acceptor density,  $\Phi_F$  is the bulk Fermi potential,  $C_{ox}$  is the oxide capacitance. However, Equation (3.3) is not directly measurable from experimental I – V characteristics, such as using standard parameter analyzer (with Source Measurement Unit, SMU). Therefore, other definition of  $V_{TH}$  has been developed to allow  $V_{TH}$  to be able to determine from standard parameter analyzer measurement. Most of these definitions are using methods that are based on the  $I_D$ - $V_G$  curve with small  $V_D$  at  $\sim 100$  mV (linear region) to extract the threshold voltage. These methods are Maximum Transconductance Method and Constant Current Method, and commonly being used by the industry also.

#### 3.2.1(a) Maximum Transconductance ( $g_{m,max}$ ) Method, $V_{TEXT}$

Transconductance ( $g_m$ ) is defined as the ratio of the current change at the output,  $I_D$  to the voltage change at the input,  $V_G$  (Taur and Ning, 1998):

$$g_m = \frac{d I_D}{d V_G} \quad (3.4)$$

By applying a small  $V_D$ , sweeping the gate voltage from low to high and measuring the  $I_D$ , the  $I_D$ - $V_G$  plot can be obtained as in Figure 3.3.  $g_m$  will be the slope of the curve and

its value changes at every  $V_G$  step. There is a point of maximum slope ( $g_m$ ) at about 0.5 V above the threshold voltage. Therefore, it is conventional to define the linearly extrapolated threshold voltage ( $V_{TEXT}$ ), by the intercept of a tangent through this point. For a second order correction in  $V_D$ ,  $V_{TEXT}$  is obtained by subtracting  $0.5V_D$  from the intercept (Taur and Ning, 1998):

$$V_{TEXT} = V_G(I_D = 0) - 0.5V_D \quad (3.5)$$

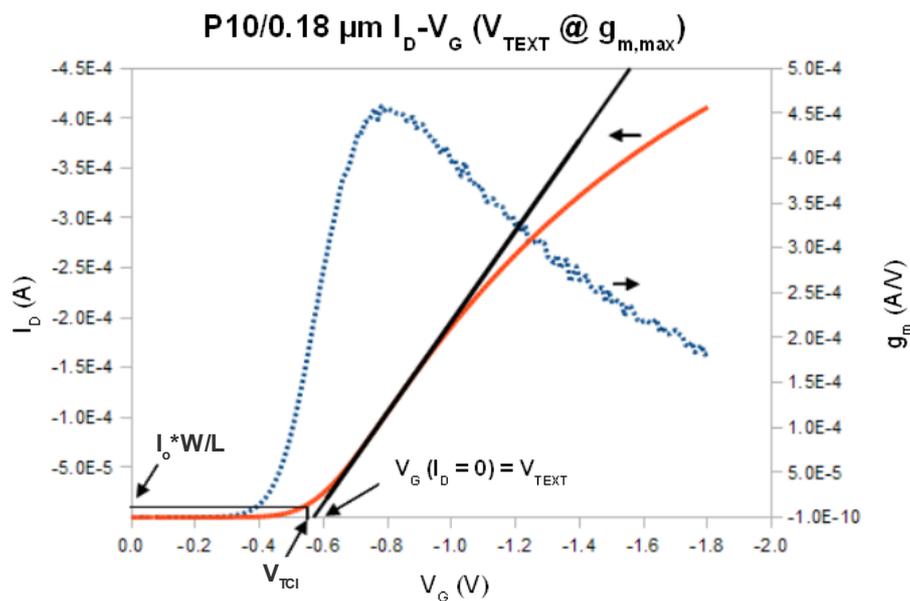


Figure 3.3:  $I_D$ - $V_G$  plot with the transconductance,  $g_m$  (blue dotted plot) showing a maximum point. The red plot is  $I_D$  and the black line is the tangent at the maximum point of  $g_m$ . Also shown is the  $V_{TCI}$  determination through  $I_o$ .

### 3.2.1(b) Constant Current Method, $V_{TCI}$

Another commonly employed definition is based on the drain to source current at the linear region ( $|V_D| = 0.1$  V), such as in Equation (3.1). From Figure 3.3, it is obvious that the drain current at the threshold voltage is higher than zero. This is utilized in the Constant Current method where the gate voltage at a specified threshold drain current is taken to be the threshold voltage. For a given constant current level  $I_o$  (e.g.  $0.1 \mu A$ ), one can define a constant current threshold voltage,  $V_{TCI}$  such that (JESD90, 2004):

$$V_{TCI} = V_G @ I_D = I_o(W/L) \quad (3.6)$$

Figure 3.3 shows graphically how the  $V_{TCI}$  can be determined. The advantage of such a threshold voltage definition is two-fold. First, it is easy to extract from standard measurement instrument and is therefore suitable for automated measurement of a large number of devices during mass production. Second, the MOSFET off current,  $I_{off} = I_D$  ( $V_G = 0$ ) can be directly calculated from  $I_o$ ,  $V_{TCI}$  and the subthreshold slope (Taur and Ning, 1998). This reduces test time in mass production, as a single measurement can yield two results.

### 3.2.2 Saturation drain current measurement

The saturation drain current,  $I_{DSAT}$  is defined as the drain current when the gate and drain is biased at nominal operating voltage with the source and substrate grounded. This means that from the  $I_D$ - $V_D$  sweep with the gate biased at operating voltage, the  $I_{DSAT}$  will be the point where  $I_D$  at  $V_D =$  operating voltage. Figure 3.4 shows the  $I_D$ - $V_D$  plot at different  $V_G$  with the  $I_{DSAT}$  is shown as the red square for this thin oxide pMOSFET with operating voltage of -1.8 V.

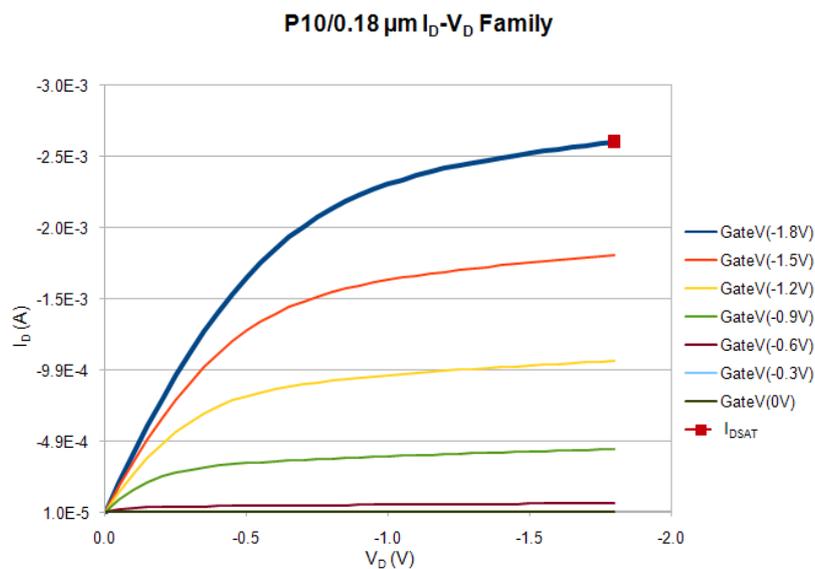


Figure 3.4:  $I_D$ - $V_D$  curves with  $V_G$  at different levels, showing  $I_{DS}$  is controlled by  $V_G$  in saturation region.

### 3.3 $I_{Sub,max}$ Stress Measurement for HCI

HCI stress is typically performed on the minimum geometry device for a particular process technology (JP-001, 2002). As an example, in the case of a dual voltage process (1.8 V and 3.3 V), the HCI should be performed on both type of transistors with minimum channel length at  $L = 0.18 \mu\text{m}$  and  $L = 0.30 \mu\text{m}$  respectively. HCI stress will require test system with four SMUs, as all the four terminals of the transistor need to be connected. HCI Stress is typically performed at room temperature and the voltage bias is DC condition. Figure 3.5 shows the flow chart of the HCI stress test.

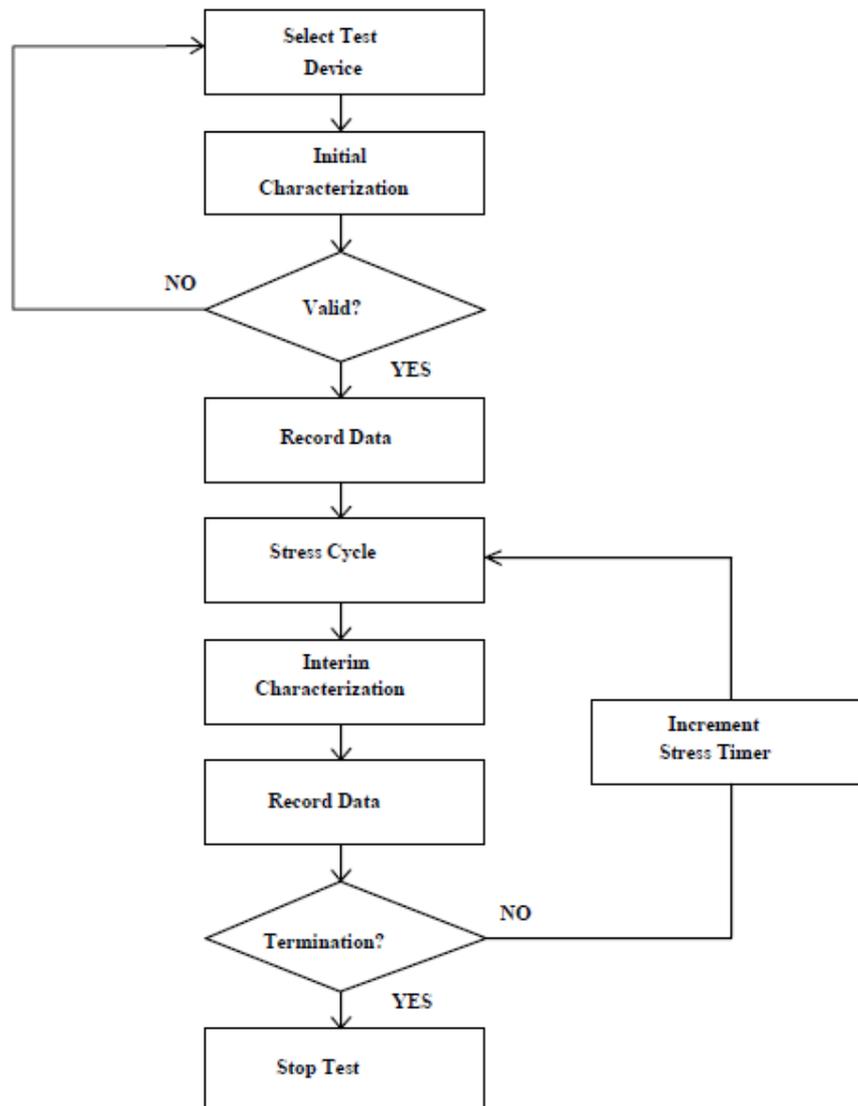


Figure 3.5: HCI stress test procedure (JESD28-A, 2001).

The most common method for HCI stress is to bias the device at maximum substrate current ( $I_{\text{Sub,max}}$ ), as this is where impact ionization is the highest, indicating high number of hot carrier near the drain. At this state, the transistor is expected to suffer the most damage. The substrate current depends on the channel lateral electric field. At low  $V_G$ , with the transistor in saturation, the lateral electric field increases with increasing gate voltage until  $V_G \approx V_D/3 - V_D/2$ . At this point,  $I_{\text{Sub}}$  increases to a maximum level (Figure 3.6). For higher gate voltages, the transistor enters its linear region, the lateral electric field decreases as does the substrate current.

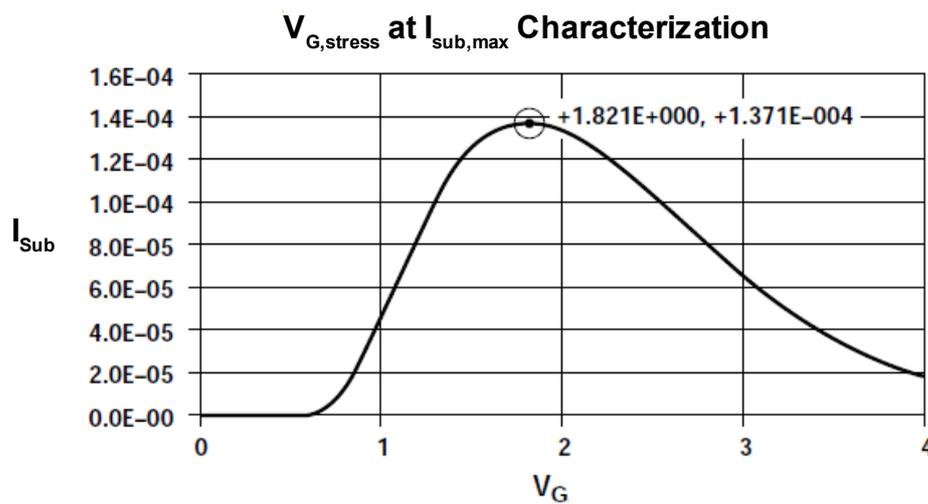


Figure 3.6:  $I_{\text{Sub}}$  is substrate current showing the maximum point (Keithley, 2000).

During the initial characterization and interim characterization, the MOSFET parameters such as  $I_{\text{DSAT}}$ ,  $V_{\text{TH}}$  will be measured. The MOSFET will be biased at  $I_{\text{Sub,max}}$  condition for a period of time, the stress will be removed and switched to interim characterization. This cycle is repeated until the device reaches the target amount of degradation (e.g. 10% shift in  $I_{\text{DSAT}}$ ). To extract the HCI lifetime, a minimum of three set of  $V_D$ - $V_G$  at  $I_{\text{Sub,max}}$  condition is required for any of the models described in Section 2.2. The lifetime extrapolation is done at  $1.1V_{\text{DD}}$ , nominal operating temperature and minimum channel length per design rule.

### 3.4 Charge Pumping for Interface States Measurement

The main HCI degradation mechanism for MOSFETs is believed to be interface trap generation, and  $I_{\text{Sub}}$  is a good monitor of such damage but is not able to quantify the interface trap generated. Interface trap is strongly related to interface states. A more common measurement that can be used for interface states is by charge pumping (Acovic et al., 1996). Charge pumping utilizes a MOSFET as the test structure, making it suitable for interface states measurements on small geometry MOSFETs (Groeseneken et al., 1984). Other method of measuring interface states is Capacitance-Voltage (C-V) measurement, using large area MOS capacitor structure.

Charge pumping technique was introduced by Brugler and Jaspers (1969) and this technique was further enhanced by Groeseneken et al. (1984). It becomes popular among many researchers as it can measure on the MOSFET device directly, requires basic equipment that is easy to set up and data analysis is straight forward. This technique is based on the recombination process at the Si-SiO<sub>2</sub> interface involving the interface states. The biasing condition applied induces a substrate current which can be directly related to the number of interface states.

The basic experimental setup for the charge pumping technique is depicted in Figure 3.7. The source and drain of the transistor are connected together and a certain reverse bias voltage is applied with respect to the substrate. The gate is pulsed between accumulation and inversion conditions by the pulse generator unit (PGU), while the charge pumping current ( $I_{\text{cp}}$ ) is measured at the substrate by the SMU. When the gate is pulsed into inversion, the surface becomes deeply depleted and electrons will flow from the source and drain regions into the channel, where some of them will be captured by the interface states. When the gate pulse reverses to drive the surface back into

accumulation, the mobile electrons drift back to the source and drain under the influence of the reverse bias, but the charges trapped in the interface states will recombine with the majority carriers from the substrate. This mechanism results in a net flow of negative charge into the substrate, which is the charge pumping current ( $I_{cp}$ ). The  $I_{cp}$  is related to the number of interface states ( $N_{it}$ ) through Equation (3.7) (Keithley, 2007):

$$I_{cp} = N_{it} q A_r f \Delta E \quad (3.7)$$

where  $q$  is the electron charge,  $A_r$  is the area of gate,  $f$  is the frequency and  $\Delta E$  is the difference between the inversion Fermi level and the accumulation Fermi level (Groesenekan et al., 1984).  $\Delta E$  is related to the thermal velocity of carrier, intrinsic carrier concentration, rise time and fall time of the pulse applied. By measuring the  $I_{cp}$ , it is able to provide an average value of  $N_{it}$  over the energy interval  $\Delta E$ . And by calculating the shift on  $I_{cp}$  in successive stress in term of percentage, the shift in  $N_{it}$  can be monitored as  $\Delta N_{it} (\%) = \Delta I_{cp} (\%)$ .

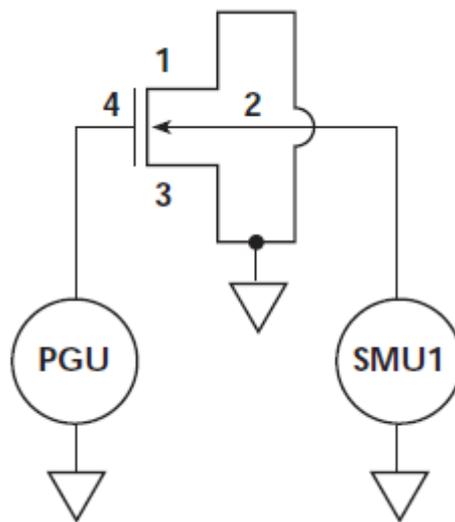


Figure 3.7: nMOSFET under charge pumping setup. PGU = pulse generator unit, SMU = source measurement unit (Keithley, 2007).

### 3.5 Constant Voltage Stress Measurement for TDDB

TDDB stress is typically performed on the minimum gate oxide thickness area capacitor for a particular process technology. As an example, in the case of a dual voltage process, the TDDB stress test should be performed on area capacitor with both gate oxide thickness ( $T_{ox} = 2.9 \text{ nm}, 5.6 \text{ nm}$ ) for the respective operating voltages of 1.8 V and 3.3 V. TDDB stress will require test system with SMU connected to the gate of capacitor, to bias the gate with voltage and monitor the leakage current during the stress. TDDB stress is typically performed at elevated temperatures and the biasing condition is constant voltage stress on gate of capacitor. This is the most common method, as recommended by JESD92 (2003).

Unlike HCI stress, TDDB stress test is non-stop until the gate oxide breakdown. Therefore the test flow of TDDB is simple: (a) measures the gate leakage current before stress ( $I_0$ ), (b) applies constant stress voltage on gate of capacitor and monitors and records the gate leakage current ( $I_{Gstress}$ ) during stress. If the  $I_{Gstress}$  has reached the failure criterion (e.g. 10 times increase compare to  $I_0$ ), then the gate oxide is considered as breakdown and records the time of breakdown. Figure 3.8 shows an example of the current-time (I-t) trace of a group of capacitors under TDDB constant voltage stress.

The time-to-breakdown ( $t_{bd}$ ) for each capacitor under the same gate stress voltage and temperature is plotted on the Weibull plot to obtain the  $t_{63.2}$  (time for 63.2% of capacitors breakdown). Therefore, under low stress voltage, it is not necessary to have all the sample (with the same stress condition) to reach the state of breakdown. If there is more than 63.2% of sample reached breakdown, then the stress test for that stress condition can be stopped.

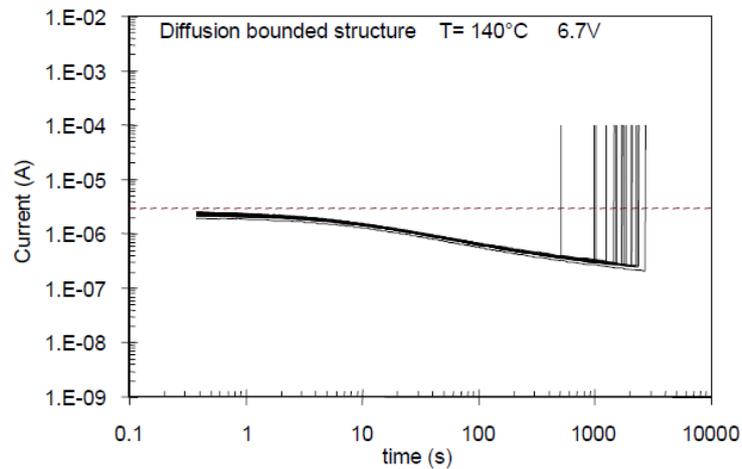


Figure 3.8: I-t trace showing jump in gate leakage current due to breakdown. Dotted line indicates the failing criteria (Vollertsen, 2004).

To extract the model parameters described in Section 2.3, a test matrix that consists of three stress voltages ( $V_1$ ,  $V_2$ ,  $V_3$ ) and three stress temperatures ( $T_1$ ,  $T_2$ ,  $T_3$ ) is needed. Typically, when testing with  $V_1$ ,  $V_2$ ,  $V_3$  (below oxide breakdown voltage that determined from voltage ramp measurement),  $T_2$  will be used. This allows the voltage acceleration factor ( $\gamma$ ) to be determined. In the same manner, when  $T_1$ ,  $T_2$ ,  $T_3$  (in the range of 25-200 °C) are applied, the stress voltage of  $V_2$  will be used. Thus, the activation energy ( $E_a$ ) that describes this temperature dependence can be determined. Of course, the structure of the test matrix is not mandatory, can be varied depending on the need of the evaluation. With the model parameters (such as  $\gamma$ ,  $E_a$  for E-Model) determined, the gate oxide lifetime at nominal operating voltage and temperature can be extrapolated (Martin et al., 1998).

### 3.6 Constant Current Stress Measurement for EM

EM stress testing is performed on each metal layer and via layer with dimension as per minimum design rule. If there are metal layers with the same design rule and go through the same fabrication process, testing on one of the metal layer will be sufficient. Stress

test shall be performed at elevated temperature (e.g. 150 °C – 250 °C for aluminum) to accelerate the mass transport of atoms. EM stress test requires test system with SMU to source constant current on the test metal line and monitors the voltage shift over time to check for change in resistance (JESD202, 2006).

Similar to TDDB, EM stress test is a non-stop test until the resistance change reaches the target value. The time to reach this target resistance change is the time-to-fail (TTF). A constant current density of several MA/cm<sup>2</sup> is usually applied with a constant oven temperature (>150 °C). The electrical resistance of every sample is monitored continuously. Due to the high current density being applied, Joule heating on the metal line structure is inevitable, changing the actual temperature of the sample during stress test. This additional temperature rise is estimated by using the temperature coefficient of resistance (TCR) data obtained from the same type of sample under test (JESD33-B, 2004). The TCR measurement is performed before the actual EM stress starts, by sending a short current pulse of magnitude ~ 0.1 mA (Tan et al., 2007) to measure the sample resistance at pre-selected temperatures (3 temperatures that are < stress temperature). The TCR value is used to adjust the actual current applied (still considered to be constant current stress as the variation of input current is small), so that the actual temperature at stress seen by the sample is the target stress temperature. This adjustment will take into account the Joule heating effect due to high current density.

EM failure mechanism is modeled by the Black's equation, as depicted by Equation (2.8). The constant current stress test allows all the Black's parameters ( $E_a$ ,  $p$  and  $A$ ) to be determined. Hence, this requires the EM tests to be performed on at least two different temperatures and two different current densities. One set of test will be at  $T_1$ - $J_1$ ; another set at  $T_2$ - $J_1$ ; and the last set at either  $T_1$ - $J_2$  or  $T_2$ - $J_2$ . Thus a minimum of

three different stress conditions are required. By plotting the TTF for each of the sample with the same stress condition in lognormal plot, the  $t_{50}$  (time for 50% of sample reaching failure criteria) for each stress conditions can be determined. With these  $t_{50}$  from each of the stress condition, the activation energy ( $E_a$ ), current exponent ( $p$ ) and constant  $A$  can be determined (JESD63, 1998).

### **3.7 NBTI Measurement Techniques**

The initial reported NBTI results were based on conventional DC Stress-Measure-Stress method, essentially the same approach as HCI testing. During the Measure phase, the MOSFET I-V characteristics are measured with a semiconductor parameter analyzer or the change in interface states is measured using charge pumping technique (with pulse generator). As the recovery effect is found to be playing a significant role in measured data, new techniques were developed, in an attempt to minimize the recovery effect, such as On-The-Fly (OTF) technique and its variants. In recent years, other advanced NBTI measurement techniques were developed (such as Fast Switching, Ultra Fast  $V_{TH}$ , Fast  $V_G$ -Ramp) in a continue quest to deal with the recovery effect.

#### **3.7.1 Conventional DC Stress-Measure-Stress method**

Conventional DC Stress-Measure-Stress method has been used initially to characterize NBTI degradation mechanism (La Rosa et al., 1997; JESD90, 2004; etc.). This is the similar approach being used to perform HCI stress, where transistor characterization (measuring parameters like  $I_D$ ,  $V_{TH}$  or  $N_{it}$ ) is performed before applying stress bias. It then follows with stress for a period of time, stops the stress, repeats the same parametric measurement and continues the stress again. The stress-measure-stress cycle will be repeated and all the parametric characterization will be done at the same temperature. The whole process is depicted in Figure 3.9.

Typically, the  $I_D$  is determined from  $I_D$ - $V_D$  measurement and the  $V_{TH}$  is determined from  $I_D$ - $V_G$  measurement. This can be achieved through standard semiconductor parameter analyzer with SMU and also a wafer probe station with hot chuck (if the sample is in wafer form). The interface traps,  $N_{it}$  is popularly measured with charge pumping method, which required a pulse generator on top of the semiconductor parameter analyzer and wafer probe station with hot chuck.

There is a period of time (“dead” time) between “stress” and “measure” phase, which depends on equipment type and configuration of SMUs connection. This “dead” time varied from researcher to researcher and is frequently not given in publications. When it is discovered that the “dead” time is quite important, due to the recovery mechanism, various techniques were developed to minimize the “dead” time (in other words, to minimize the recovery).

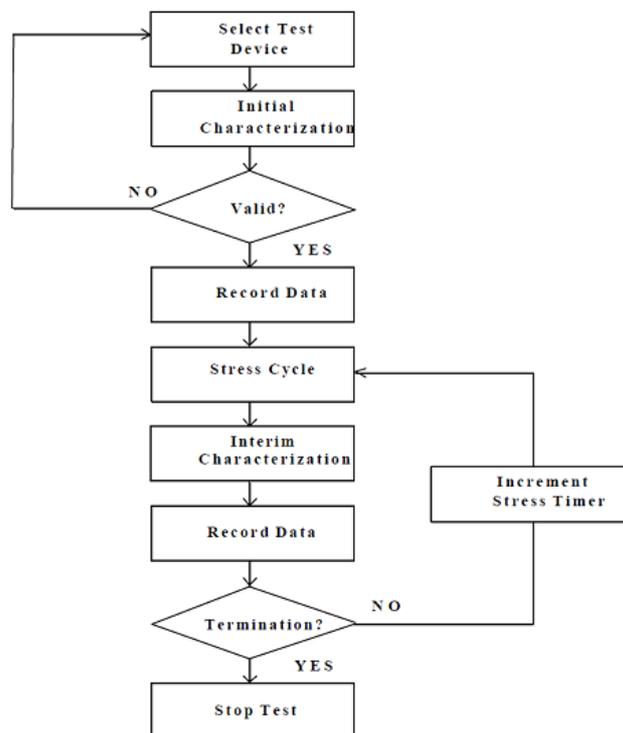


Figure 3.9: Conventional DC Stress Measure Stress flow (JESD90, 2004).

### 3.7.2 On-The-Fly Measurement Method

Rangan et al. (2003) uses  $V_D = 50$  mV and monitors the linear drain current ( $I_{DLIN}$ ) during stress and recovery, in order to minimize recovery due to the switching from stress to measurement.  $I_{DLIN}$  shifts are proportional to  $V_{TH}$  shifts and therefore proportional to defect generation. Time delay (“dead” time) between stress and recovery measurements is minimized by quickly switching the gate voltage from stress level to recovery level using a function generator unit, through a trigger from a semiconductor parameter analyzer.

Denais et al. (2004) and Huard et al. (2005) propose the On-The-Fly method, which is a modification based on Rangan et al. (2003) method. A voltage pulse of  $\sim 50$  mV is applied to the drain and  $I_D$  is being recorded at the end of the stress period, without interrupting the stress (as shown in Figure 3.10(a)). This is repeated and in this way  $I_D$  degradation is recorded with minimal “dead” time. Since the  $I_D$  current depends on threshold voltage, mobility and other device parameters, it is not possible to extract the  $\Delta V_{TH}$ . In a modification of this method (coined as second level On-The-Fly), as shown in Figure 3.10(b), a small gate voltage pulses are applied on the DC gate stress voltage, at the same time of the drain voltage pulse. The  $I_D$  is measured at the three different gate voltages, allowing the transconductance ( $g_m$ ) at stress level to be determined. With both  $I_D$  and  $g_m$  at stress available, the  $\Delta V_{TH}$  can be determined.

The OTF method seem to be able to eliminate the “dead” time, but there are limitations to get a correct extraction of  $\Delta V_{TH}$ . The measurement of the first  $I_D$  data point after the initiation of the stress requires a certain amount of time from the moment the stress is started. Significant trapping may take place and it is highly dependent on instrumentation used (measurement delay) and spontaneous interface trap generation

(Shen et al., 2006). The shift in  $I_D$  measured with OTF is sensitive to the hole traps that are activated at  $V_{G,\text{stress}}$ . In addition, the error due to mobility degradation is not accounted for in the extracted  $\Delta V_{\text{TH}}$  (Hu et al., 2008) for OTF method.

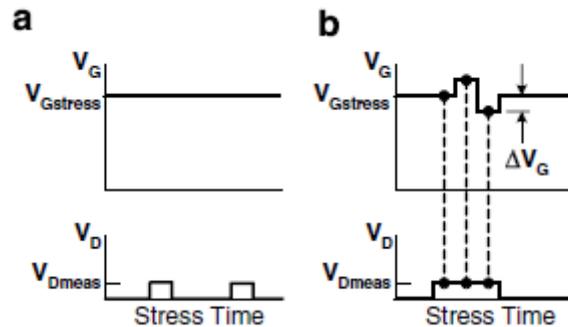


Figure 3.10: Measurement schematics for (a) drain current (Denais et al., 2004), (b) drain current and transconductance without interrupting stress (Huard et al., 2005).

### 3.7.3 Fast Switching Method

An improvement to the OTF method was proposed by Kaczer et al. (2005) by using fast switching of gate voltage from stress level ( $V_{G,\text{stress}}$ ) to measurement level ( $V_{G,\text{meas}}$ ). The  $I_D$ - $V_G$  characteristics of an unstressed device is measured initially so that  $V_{\text{TH}}$  before stress ( $V_{\text{TH}0}$ ) can be determined. Then the gate voltage is lowered briefly from its stress value to a voltage near  $V_{\text{TH}}$  before stress ( $V_{G,\text{meas}} \approx V_{\text{TH}0}$ ) and the  $I_D$  is measured. This  $V_{G,\text{meas}}$  will be used for all the measurement phase. The drain is kept constant at -50 mV throughout the stress and measurement phase. The gate voltage shift of the stressed  $I_D$  is converted to  $\Delta V_{\text{TH}}$  by horizontally shifting the initial  $I_D$ - $V_G$  curve. This can be achieved by solving the parabolic equation that describes the portion of initial pre-stress  $I_D$ - $V_G$  curve near the  $V_{G,\text{meas}}$  and assuming the shape of that portion of  $I_D$ - $V_G$  curve do not changed after stress.

Measuring  $I_D$  at gate voltage near  $V_{\text{TH}0}$  is essentially the approximation of extracting threshold voltage through constant current method (see Section 3.2.1(b)). The

advantage of measuring  $I_D$  near  $V_{TH0}$  is that no hole trapping in pre-existing traps takes place during measurement as gate voltage is quite low and the measurement can be done very fast as this is just a 1-point measurement.

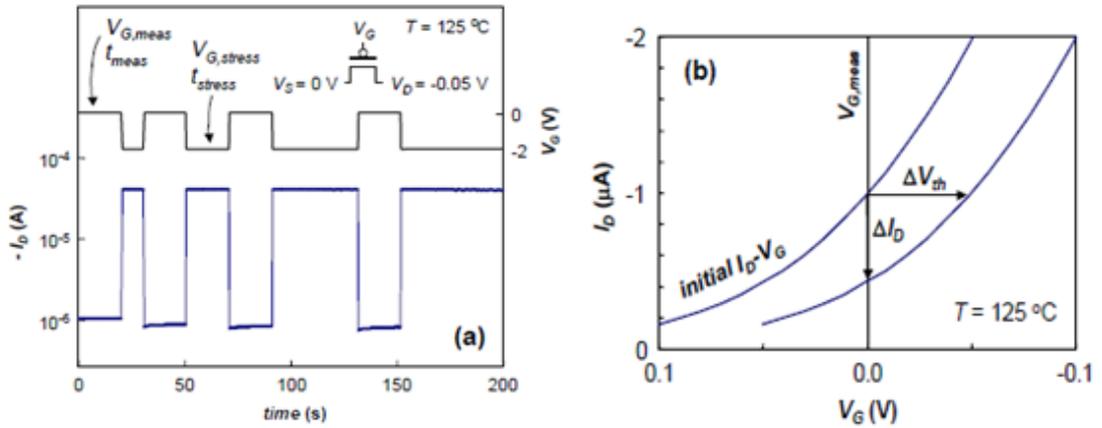


Figure 3.11: (a) The pMOSFET is biased as shown.  $V_G$  is varied between the measurement and stress values while  $I_D$  is monitored. (b)  $V_G$  is converted to  $\Delta V_{TH}$  by horizontally shifting the initial  $I_D$ - $V_G$  curve (Kaczer et al., 2005).

Another way of extracting  $\Delta V_{TH}$  from the measured  $I_D$  in between stress cycle is using the extracted subthreshold slope (SS) from the initial pre-stress  $I_D$ - $V_G$  curve plotted on semi-log plot as shown in Figure 3.12 (a). The  $I_D$  being measured in this biasing condition is essentially the subthreshold current of the pMOSFET. The subthreshold current, even though is small, flows through the entire channel and will potentially be able to detect charge exchange between the oxide and the substrate all along the channel. By assuming that the subthreshold slope does not change significantly with stress, the  $\Delta V_{TH}$  can be calculated through the following equation (Huard et al., IRW 2005):

$$\Delta V_{TH}(t) = \frac{\log(I_D(0)) - \log(I_D(t))}{SS} \quad (3.8)$$

This Fast Switching method with 1-point  $I_D$  measurement had been drafted to the JEDEC Committee 14.2.2, to be approved as a NBTI measurement standard, in addition to current existing test standard JESD90 (Conventional DC Stress-Measure-Stress).

Schlunder et al. (2008) introduce an improved version of this Fast Switching method. Instead of performing a 1-point  $I_D$  measurement after stress, 2-point  $I_D$  measurements are made (at 2 different  $V_{G,meas}$ ). This allows SS between stress cycles to be determined, as it shows that the SS is not constant when the stress time increases. The change in SS affects the extracted  $\Delta V_{TH}$ , and Equation (3.8) can be rewritten as:

$$\Delta V_{TH}(t) = \frac{\log(I_D(0)) - \log(I_D(t))}{SS(t)} \quad (3.9)$$

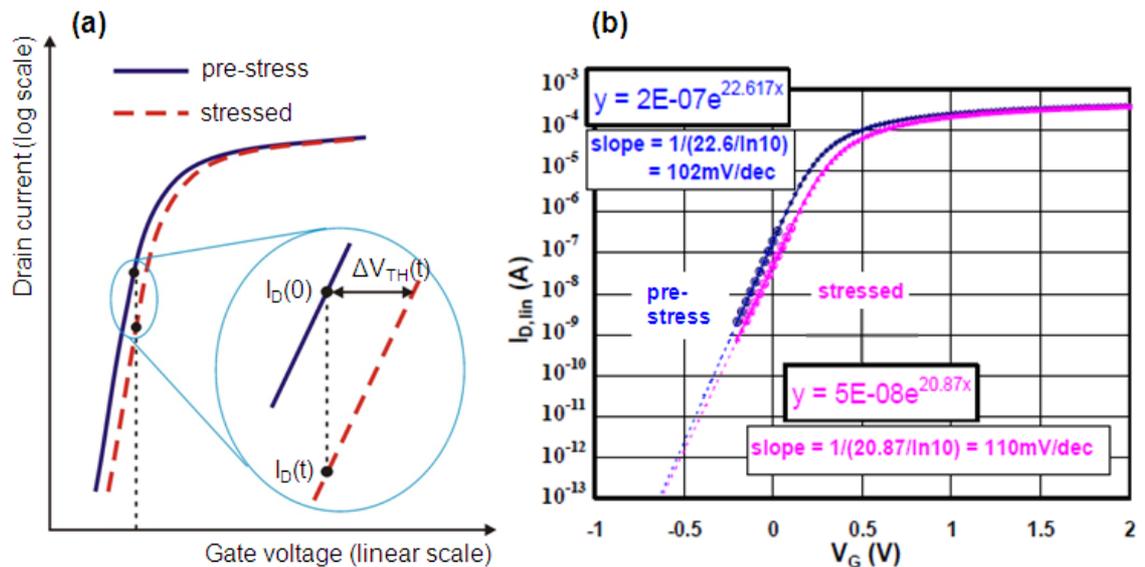


Figure 3.12: (a)  $I_D$ - $V_G$  plotted in semi-log plot. The subthreshold slope is the linear portion of the curve. (b) The subthreshold slope (SS) can be easily extracted by solving the exponential equation on the linear portion (Schlunder et al., 2008).

### 3.7.4 Other Advanced Measurement Methods

Other advanced measurement methods were developed to further reduce the “dead” time. These methods deviate from the standard parameter analyzer type of equipment by adding extra circuitry or other instruments. Very often, these advanced methods are either being patented, require custom build circuit board which is not commercially available or require expensive new measurement unit which is capable of fast measurement. Therefore, these techniques will be costly and the old wafer fabrication plant will be reluctant to use such techniques.

The Ultra Fast  $V_{TH}$  method was introduced by Reisinger et al. (2006) utilizing a feedback loop to force a constant current across the source and drain of pMOSFET. An Opamp is then used to try to zero the difference between actual current and current setpoint (usually the constant current to extract  $V_{TH}$ ), by applying appropriate  $V_G$ . With the test configuration depicted in Figure 3.13, the  $V_G$  can be recorded after the settling time of the feedback loop (about 1  $\mu$ sec after stress). This  $V_G$  value will be the  $V_{TH}$  after stress, as the  $I_D =$  current setpoint.

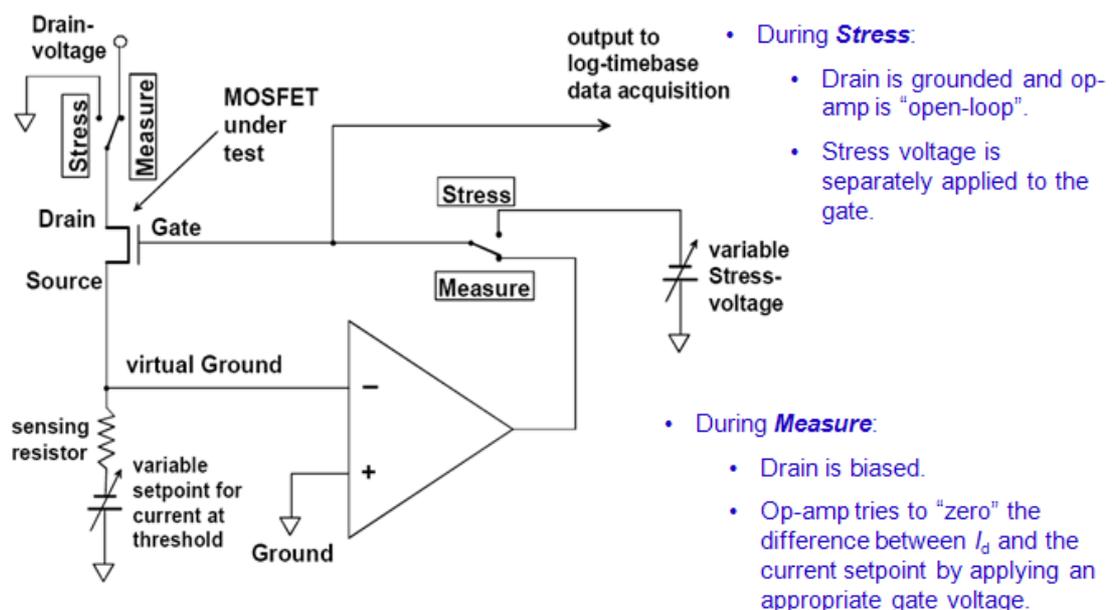


Figure 3.13: Test configuration of Ultra Fast  $V_{TH}$  Method (Reisinger et al., 2006).

Kerber et al. (2003) introduced the fast  $I_D$ - $V_G$  measurement method with  $V_G$  Ramp (from Pulse Generator) and utilized an oscilloscope for measurement (Figure 3.14(a)). The ramp in  $V_G$  will cause fluctuation in  $V_D$  (as a result of fluctuation in the  $I_D$ ) and records by the oscilloscope.  $I_D$  is then calculated as follows:

$$I_D = \frac{100\text{mV}}{V_D} \cdot \left( \frac{100\text{mV} - V_D}{R_L} \right) \quad (3.10)$$

With the value of the calculated  $I_D$ , the  $I_D$ - $V_G$  curve can be reconstructed to determine the  $V_{TH}$ . However, there is a relative delay between  $I_D$  and  $V_G$  waveforms (due to parasitic capacitance, highlighted in red color in Figure 3.14(a)), resulting in the  $I_D$ - $V_G$  curve distortion. Shen et al. (2006) introduced an improved version of this Fast  $V_G$ -Ramp method by adding extra circuitry and proper cabling to ensure same signal propagation delay (see Figure 3.14(b)) between the drain and gate. With this setup, the measurement delay (“dead” time) can be reduced down to 1  $\mu\text{sec}$ .

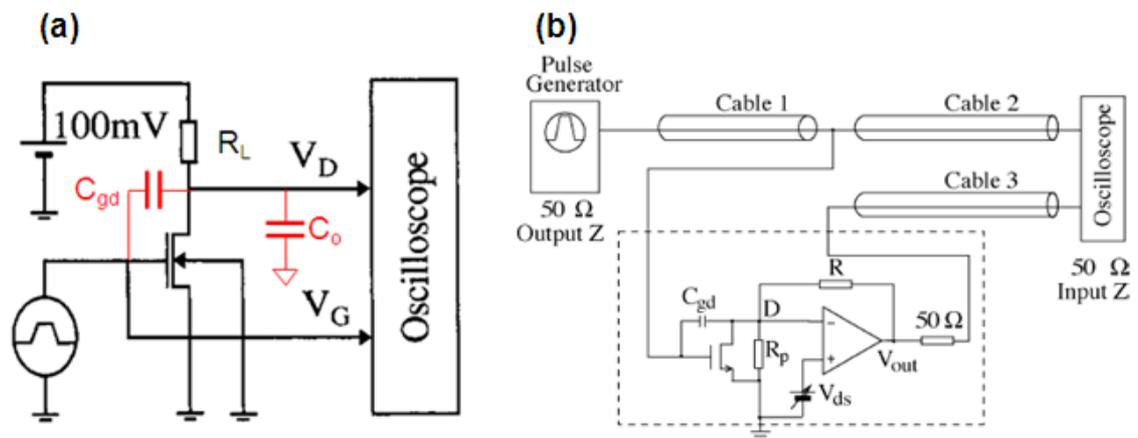


Figure 3.14: (a) Test configuration of Fast  $V_G$ -Ramp introduced by Kerber et al. (2003).

(b) An improved version with extra circuitry and same length coaxial cable, being introduced by Shen et al. (2006).

Further reduction in the “dead” time is achieved with the Ultra Fast Switching method introduced by Du et al. (2009), which is based on pulsed I-V measurement. This test configuration requires new measurement instrument that incorporates Pulse

Generator Unit, Source Measurement Unit, and Oscilloscope (see Figure 3.15(a)). Pulsed I-V is a measurement technique aimed at minimizing self heating of transistor during I-V measurement. With the gate of pMOSFET constantly being stressed, measurement at a specific interval is achieved by injecting a narrow positive pulse of  $v_g$  which momentarily brings the overall gate voltage down to measurement level (Figure 3.15(b)). Corresponding change in  $I_D$  is converted into voltage (via an internal resistance) and measures by a high speed digital oscilloscope. The “dead” time for this technique can be reduced down to  $< 100$  nsec.

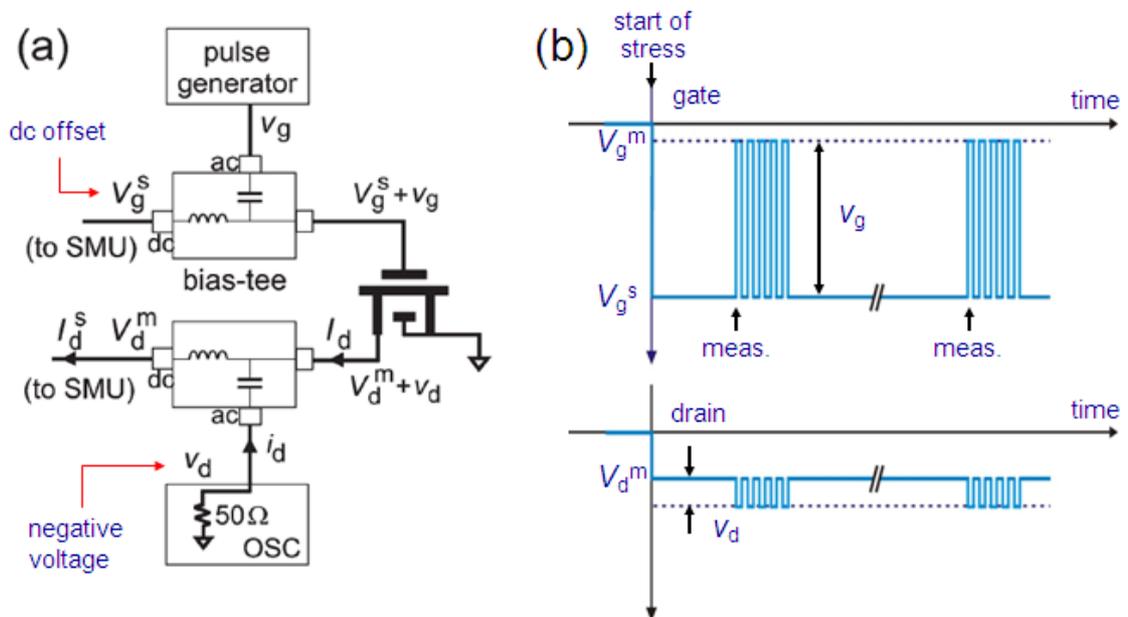


Figure 3.15: (a) Ultra Fast Switching test configuration utilizing pulsed I-V measurement scheme. (b) Waveform showing the stress and measurement. (Du et al., 2009).

## **Chapter 4. Research Methodology**

### **4.1 Introduction**

This chapter will provide an overview of the project methodology by using a flow chart, and follows by the detailed description of the flow chart. The flow chart will serve as the guideline for the activities carry out on this project to achieve the objectives specified in Chapter 1. In order to achieve these objectives, the flow chart will lay out the available options to be evaluated and making decision to proceed to next step. The following sections will cover the considerations in selecting the sample, equipment, measurement methods, biasing condition and experimental setup for performing characterization on NBTI degradation.

### **4.2 Research Project Flow Chart**

The project flow chart is mainly divided into 2 major phases: the feasibility study phase and the experimental phase. The feasibility study phase looks into the selection of the correct pMOSFET to be used for performing experiment on the available wafer sample and to ensure the existing hardware (measurement system and probe station) is capable of making the correct and accurate measurement. Before embarking onto the experimental phase, the measurement obtains from this feasibility study phase will be used to benchmark with other test system to make sure that the hardware performance is up to the industry standard. The experimental phase will cover the experiments to characterize the NBTI degradation on the selected pMOSFET using conventional method. This phase also includes other characterization experiments by varying the stress temperature, channel length and gate oxide thickness. The last part of this phase will evaluate the fast measurement methods that can be implemented on existing hardware and performs experiments on the selected fast method to compare with the conventional method.

Figure 4.1 and 4.2 show the project flow chart for the feasibility study phase and experimental phase respectively.

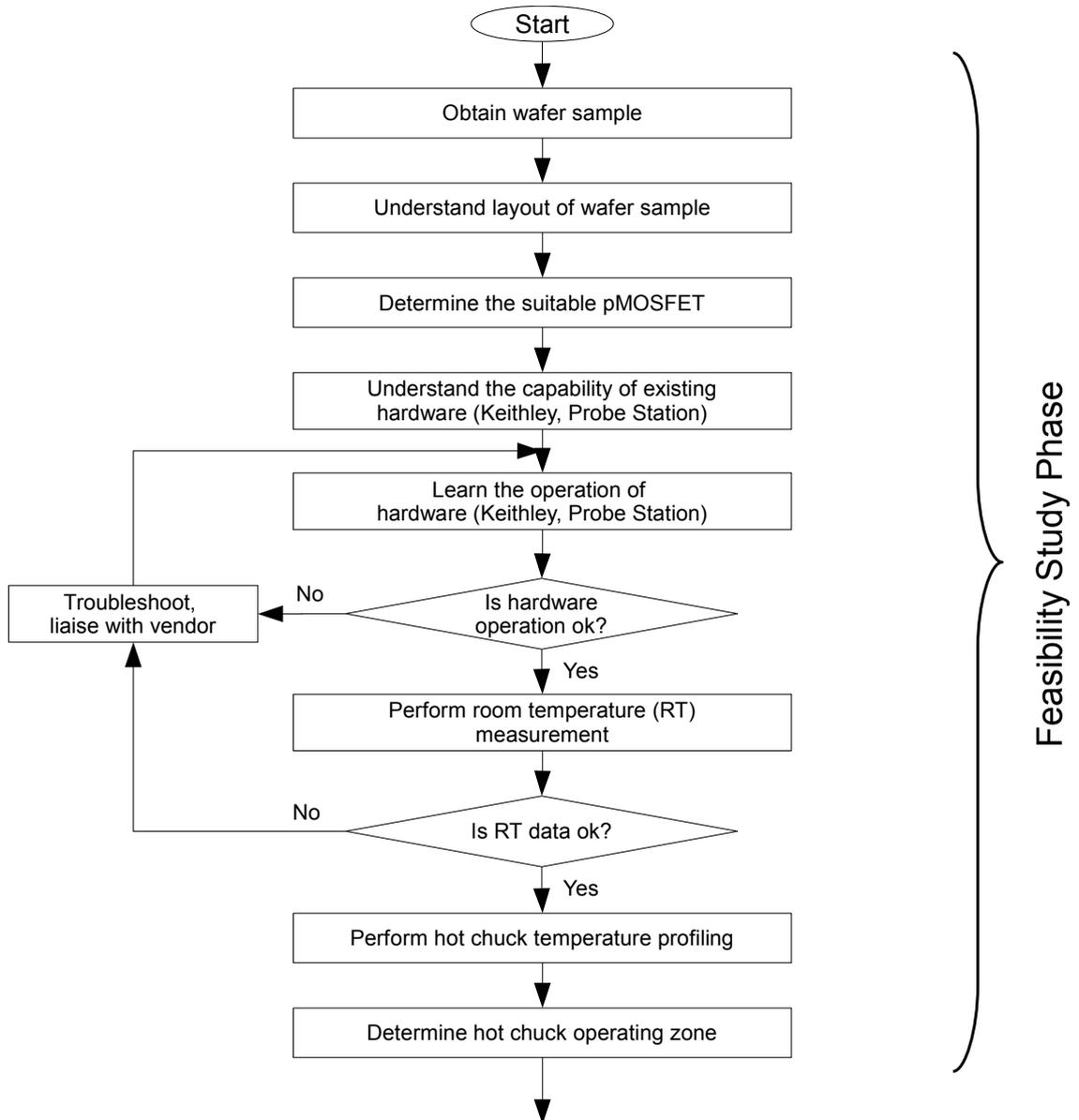


Figure 4.1: Flow chart showing the feasibility study phase of the project. The flow chart will continue to Figure 4.2.

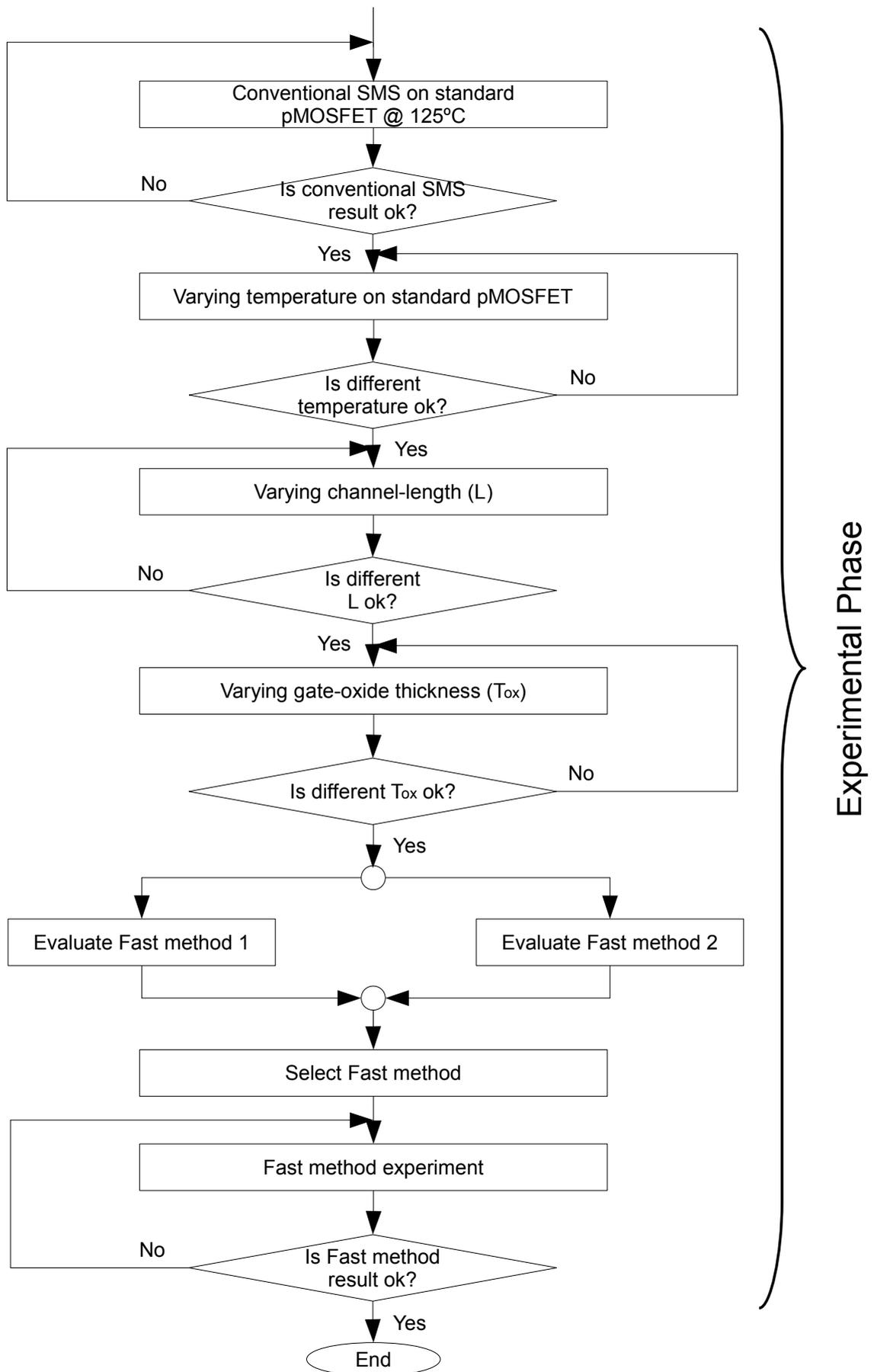


Figure 4.2: Continuation of project flow chart showing experimental phase.

### 4.3 Wafer Sample

The wafer sample provided by the wafer fabrication plant is an 8" wafer, fabricated on a 0.18  $\mu\text{m}$  dual gate oxide CMOS process. To characterize the NBTI degradation, the stress test needs to be performed on isolated pMOSFET. This specifically designed test structure allows the pMOSFET to be probed by a probe station, enabling voltage bias to be applied and current being measured during stress and measurement cycles of NBTI test.

Based on the information provided by the wafer fabrication plant, the wafer sample is a production wafer, with test structures located on the scribe lane (Figure 4.3). The scribe lane is the area in between product IC and this area will be destroyed when the wafer is being saw during assembly into packaged IC. Therefore, all the measurement collected from test structures need to be done at wafer level.

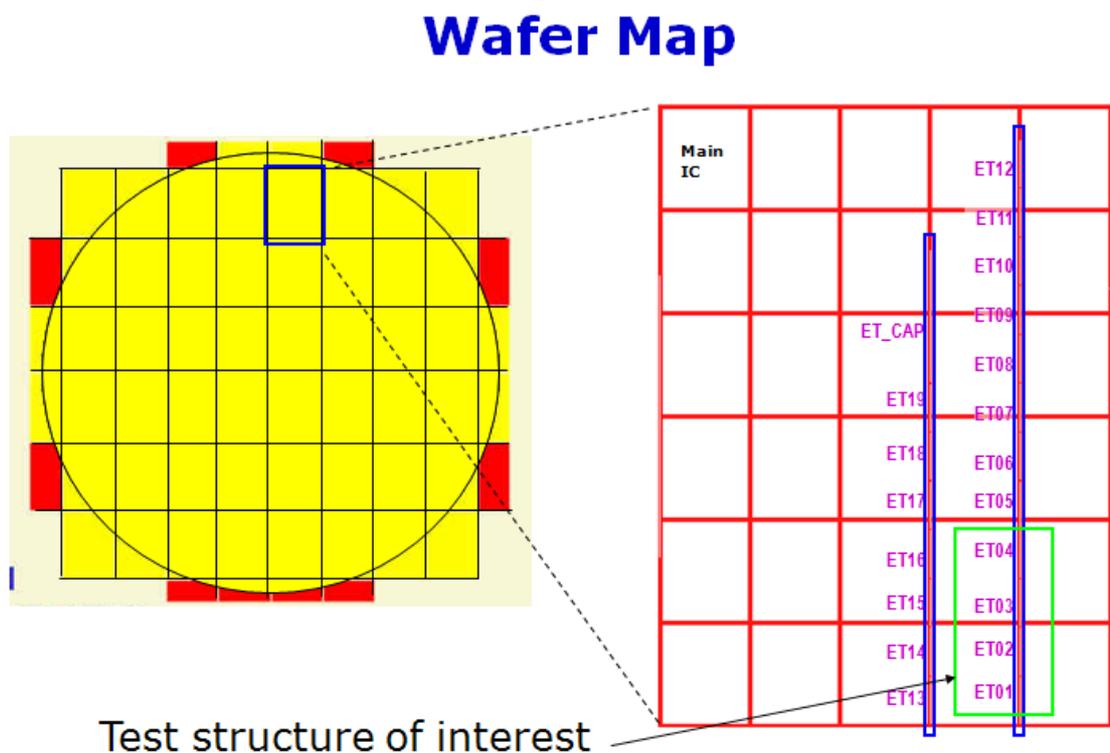


Figure 4.3: Wafer map showing the location of test structures of interest.

### 4.3.1 Pad group design

As described by the documentation from wafer fabrication plant, similar type of test structures with different designs are grouped together with naming started with “ET” (Figure 4.3), with each test group consists of 18 probe pads. For example, ET01 groups all the thin oxide nMOSFETs together, and the transistors are connected with common gate, common source and common substrate pad. For thin oxide pMOSFETs, they are found in ET02 and thick oxide pMOSFETs are found in ET04 (Figure 4.4) with similar style of connection.

ET02 – CMOS18 Pchannel Thingate		ET04 – CMOS18 Pchannel Thickgate	
1	10/10 drain	1	10/10 drain
2	10/0.18 drain	2	10/0.30 drain
3	10/0.18 drain rotated	3	P+Island blocked resistor 99/2.7
4	10/0.16 drain	4	P+Island blocked resistor 99/0.45
5	0.22/0.18 drain	5	P+Poly blocked resistor 99/2.7
6	0.22/10 drain	6	P+Poly blocked resistor 99/0.45
7	Array drain, 1000 parallel 2.7/0.18	7	Array drain, 1000 parallel 2.7/0.31
8	Common gate	8	Common gate
9	Common Source	9	Common Source
10	Common NWell	10	Common NWell
11	10/10 drain	11	10/0.37 drain
12	10/0.17 drain	12	10/0.34 drain
13	10/0.17 drain rotated	13	10/0.34 drain rotated
14	10/0.19 drain	14	N+Poly blocked resistor 99/2.7
15	0.22/0.17 drain	15	N+Poly blocked resistor 99/0.45
16	Array drain, 1000 parallel 2.7/0.17	16	Array drain, 1000 parallel 2.7/0.34
17	P+ Island salicided resistor 99.99/0.22	17	P+ Poly salicided resistor 99.99/0.18
18	P-channel Thingate MOSCAP Uses common source for majority carriers, Gate area = 16587um <sup>2</sup>	18	P-channel Thickgate MOSCAP Uses common source for majority carriers, Gate area = 16587um <sup>2</sup>

Figure 4.4: Description of pad group ET02 (left, thin oxide pMOSFET) and ET04 (right, thick oxide pMOSFET).

With this common gate, common source and common substrate style of connection, 15 transistors can be grouped together under one pad group, allowing more transistors to be placed into the limited space of the scribe lane. This scheme will allow the production automatic tester to perform testing more efficiently and to reduce the test time. However, the common gate connection among different transistors is not desirable for NBTI stress test. For example, if 10/0.18  $\mu\text{m}$  pMOSFET is under NBTI stress test, pad 8 is biased at stress voltage; pad 2, 9 and 10 will be biased to ground. Due to pad 8 (gate pad) is being biased at stress voltage, all other transistors with the gate connected to this pad will be stressed to some extent, even though their respective drain pad is floating. For production testing, this will not be an issue as the applied bias during production testing is typically at nominal operating voltage. In the case when there are test items in the production testing that required applied voltage that is higher than nominal operating voltage, those test items will be tested last for that particular pad group, in order not to interfere with the results of other test items. Hence, when one of the transistors in a particular pad group has been stressed for NBTI, other transistors in the same pad group will not be eligible for NBTI stress test. This will limit the number of sample that can be used on each wafer for this project.

As the wafer samples provided are production wafer, it is the standard procedure for the wafer fabrication plant to perform production testing on scribe lane test structures, to determine whether the wafer produced is within specification limit. An inspection under high magnification optical microscope revealed that the wafer sample had been tested for at least 5 locations across the wafer, with some wafers being tested on more locations. These tested locations have visible probe mark under the optical microscope. As the voltage level being used during production testing is unknown and for the reason not to introduce unfavorable factor into the experimental study, these

tested locations will be excluded from this NBTI study.

Another consideration that needs to be taken is the design of transistor, which is represented by W/L in Figure 4.4. As per the design rule for this process technology, the minimum channel length (L) for thin oxide pMOSFET is 0.18  $\mu\text{m}$  and for thick oxide pMOSFET is 0.30  $\mu\text{m}$ . To reduce the shallow trench isolation (STI) effect on transistor performance, transistors with larger channel width (W) are preferred. From Figure 4.4, the 10/0.18  $\mu\text{m}$  and 10/0.30  $\mu\text{m}$  transistor are selected for NBTI testing. There are also transistors with a 10% reduction in channel length available in ET02 and ET04, and they are not expected to show any significant difference in NBTI performance as compared with minimum channel length transistor (per design rule), hence will not be used in this experimental study. To study the variation of channel length on NBTI degradation, pMOSFET with 10/10  $\mu\text{m}$  design will be used.

#### **4.3.2 Selection criteria for pMOSFET**

The summarized list of pMOSFET selection criteria for this NBTI experimental study is as shown below:

- Any pad group that had been probed by the wafer fabrication plant will not be used anymore.
- When one of the transistors in a pad group had been subjected to NBTI stress, other transistors in the same pad group will not be used for NBTI stress.
- pMOSFET with minimum channel length and wide channel width as per design rule will be used, e.g. 10/0.18  $\mu\text{m}$  for thin oxide, 10/0.30  $\mu\text{m}$  for thick oxide.
- To study the channel length variation on NBTI degradation, pMOSFET with 10/10  $\mu\text{m}$  design will be used.

## 4.4 Hardware: Keithley Measurement System

The common practice to measure transistor parameter is to apply voltage bias on the transistor terminals and measures the current on those terminals at the same time (Ketchen and Bhusan, 2011). This needs the voltage source and current meter connected to the same terminal. To simplify this setup, commercially available Source Measurement unit (SMU) is often used. The currently available system, Keithley 4200 Semiconductor Characterization System (SCS) consists of three SMUs and a ground unit (GNU). This is not a preferable configuration as usually four SMUs are needed for transistor I-V measurement. To circumvent this short coming, the connection to Source or Bulk can be substituted by the GNDU, as during transistor parameter measurement and NBTI stressing, Source and Bulk are grounded and the current on Source and Bulk is not required to be measured. Figure 4.5 shows the front panel of Keithley 4200 SCS.



Figure 4.5: Keithley 4200 SCS front panel (Keithley, 2009).

### 4.4.1 Correlation of Keithley 4200 SCS to Industry

Before the start of NBTI stressing experiments, the accuracy and repeatability of SMUs in Keithley 4200 SCS need to be correlated to be at least on par with the industry production tester. The SMU operates as a voltage or current source (depending on

source function) connected in series with an I-Meter, and connected in parallel with a V-Meter (Figure 4.6). The voltage limit (V-limit) and current limit (I-limit) circuits limit the voltage or current to the programmed compliance value. The best way to complete this verification is to run the same measurement on the wafer sample as those being performed by the production test from wafer fabrication plant.  $I_D$ - $V_D$  and  $I_D$ - $V_G$  sweep measurement at room temperature will be performed on a 10/0.18  $\mu\text{m}$  pMOSFET from the wafer sample. The saturation drain current and threshold voltage will be extracted respectively. The saturation drain current and threshold voltage are the key parameters for any transistor and will be benchmarked against the same data from the wafer fabrication plant.

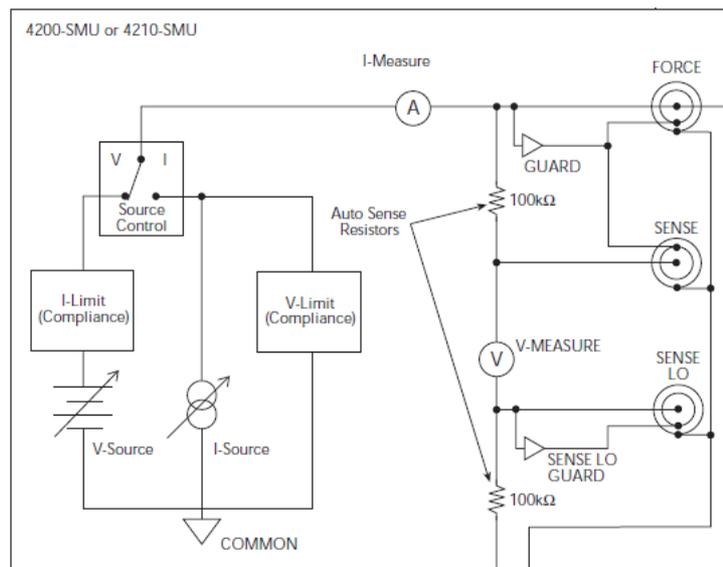


Figure 4.6: Typical configuration of a SMU (Keithley, 2009).

The 3 SMUs on the Keithley 4200 SCS will be connected to Drain, Gate and Source pad of the pMOSFET and the GNDU will be connected to the Bulk pad.  $I_D$ - $V_D$  sweep at different  $V_G$  will be performed to check the pMOSFET transfer characteristic and to extract saturation drain current,  $I_{DSAT}$  at this biasing condition:  $V_D = V_G = -1.8$  V,  $V_S = V_B = 0$  V. In addition,  $I_D$ - $V_G$  sweep at  $V_D = -0.1$  V (linear region) is performed to

extract the constant current threshold voltage ( $V_{TCI}$ ). These measurements will be repeated 5 times to check the repeatability of the SMUs. The  $V_{TCI}$  is defined as in Equation (4.1), in line with the definition from the wafer fabrication plant and JESD90 (2004).

$$V_{TCI} = V_G @ I_D = -0.1 \mu A * \frac{W}{L} \quad (4.1)$$

#### 4.4.2 Keithley 4200 SCS correlation results

Table 4.1 below shows the measured data at Keithley 4200 SCS and data from the wafer fabrication plant production tester. Figure 4.7 and 4.8 are showing the transfer characteristic of  $I_D$ - $V_D$  and  $I_D$ - $V_G$  sweep, with data from wafer fabrication plant (WfrFab) plotted on the same chart.

Table 4.1: Correlation data on  $I_{DSAT}$  and  $V_{TCI}$  showing the data from wafer fabrication plant matching very well with Keithley 4200 SCS.

$I_D$ - $V_D$ Sweep	$I_{DSAT}$ (A)	$I_D$ - $V_G$ Sweep	$V_{TCI}$ (V)
Meas1	-2.5931E-3	Meas1	-0.5000
Meas2	-2.5921E-3	Meas2	-0.5001
Meas3	-2.5924E-3	Meas3	-0.5002
Meas4	-2.5910E-3	Meas4	-0.5001
Meas5	-2.5929E-3	Meas5	-0.5001
Mean	<b>-2.5923E-3</b>	Mean	<b>-0.5001</b>
Std. dev.	8.2993E-7	Std. dev.	7.0711E-5
WfrFab	<b>-2.5975E-3</b>	WfrFab	<b>-0.5070</b>

This data shows that the measurement made by Keithley 4200 SCS is correlated to the data from the production tester at the wafer fabrication plant. Both  $I_{DSAT}$  and  $V_{TCI}$  from the production tester overlap exactly on the plot in Figure 4.7 and Figure 4.8. This shows that the Keithley 4200 SCS is capable of measuring current at linear region (low current) and saturation region (high current) of the transistor. Also, the Keithley 4200 SCS is capable of making accurate repeatable measurements, as indicated by the small

standard deviation value in Table 4.1.

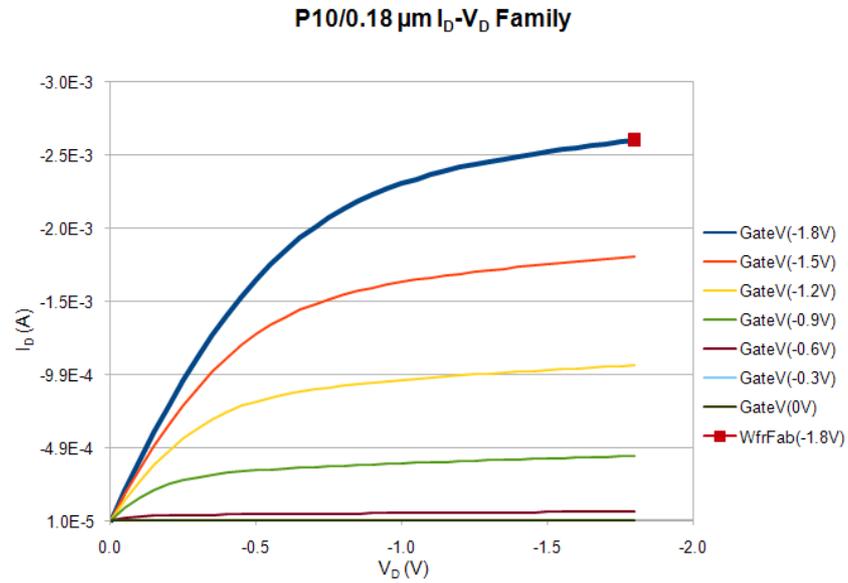


Figure 4.7:  $I_D$ - $V_D$  sweep with  $V_G$  stepping from 0 V  $\rightarrow$  -1.8 V at step of -0.3 V. The red color square is the  $I_{DSAT}$  data from wafer fabrication plant.

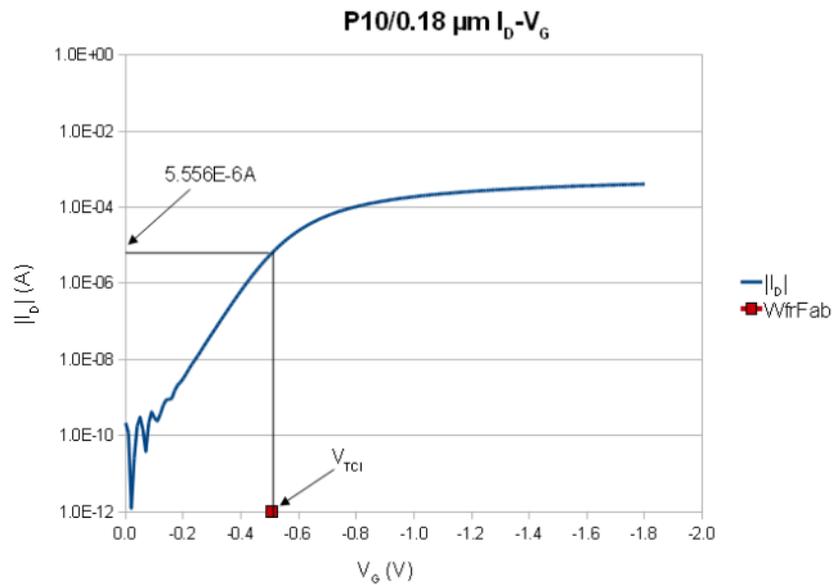


Figure 4.8:  $I_D$ - $V_G$  sweep at  $V_D = -0.1$  V, plotted on semi-log scale. The  $I_D$  is in absolute form. For 10/0.18  $\mu\text{m}$  pMOSFET,  $V_{TCI}$  is  $V_G @ I_D = 5.556\text{E-}6$  A. The red color square is the  $V_{TCI}$  data from wafer fabrication plant.

## 4.5 Hardware: Everbeing Probe Station

For this project, the sample for experimental study is in wafer form. Therefore, a probe station with hot chuck is required, as the NBTI stress will require high temperature. The probe station is to hold the wafer (with vacuum) and provides connection to the SMUs and GNU in Keithley 4200 SCS. The hot chuck, where the wafer sample will be placed on top of it will provide the temperature to the wafer sample, through the temperature controller. This probe station is made by Everbeing, with a 6" hot chuck (mounted on anti-vibration stage) and 4 micro-positioners that hold the probe needle (Figure 4.9). The micro-positioners are connected to the SMUs of Keithley 4200 SCS with triaxial cable. Care must be taken when adjusting the position of micro-positioner as the base is magnetized. The micro-positioner is being hold firmly on the platen of the probe station by magnetic force.

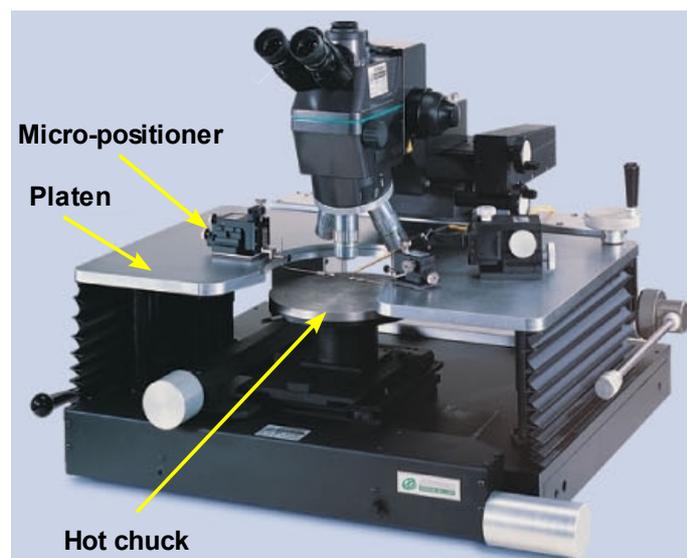


Figure 4.9: Everbeing probe station (Everbeing, 2002).

### 4.5.1 Temperature profiling on Everbeing probe station

NBTI stress needs to be done at elevated temperature (JESD90, 2004), and the hot chuck on the Everbeing probe station needs to deliver accurate and consistent

temperature control. The wafer sample which is 8" in diameter has to be placed on the smaller 6" hot chuck. Due to the difference in size, the temperature on the wafer surface needs to be verified first. High temperature (e.g. 125 °C) will be set on the Temperature Controller of the hot chuck, and allows it to stabilize for at least 10 minutes. Then the temperature on the wafer surface will be measured by Fluke 568 thermometer, using a K-type thermocouple at different locations across the wafer surface, according to the 6" diameter of the hot chuck, to check for the profile of temperature gradient. Due to the larger size of the 8" wafer, the area that is not in contact with the 6" hot chuck will not be measured. The temperature at the same location will be measured for at least 5 minutes, to check for the temperature fluctuation.

#### **4.5.2 Results of temperature profiling**

Due to the heat dissipation from the larger wafer sample under the lab ambient condition, the temperature across the wafer surface is not even. In fact, none of the measured temperatures across the different locations on the wafer shows the same value as the setting on the Temperature Controller. After several trial runs, the Temperature Controller has to set at 145 °C in order to achieve a value of 125 °C on the location of wafer surface that is correlated to the center of the hot chuck. Table 4.2 shows the minimum, maximum and mean value of the temperature at different locations with the measurement locations depicted in Figure 4.10. To assess the relationship between the setting on Temperature Controller and actual wafer surface temperature, another set of measurements with Temperature Controller set at 125 °C is performed with the results shown in Table 4.3.

Table 4.2: Wafer surface temperature measured with Fluke 568 thermometer at different locations of the hot chuck, with Temperature Controller set at 145 °C.

Chuck Location	Min (°C)	Max (°C)	Mean (°C)
Top1	119.7	120.1	119.93
Top2	120.4	120.9	120.67
Top3	122.5	123.8	123.1
Center	125	125.6	125.3
Bottom3	122.8	123.1	122.97
Bottom2	121.7	122.4	121.97
Bottom1	119.8	120.4	120.13
Left1	121.4	121.8	121.57
Right1	120.6	121.5	121.07

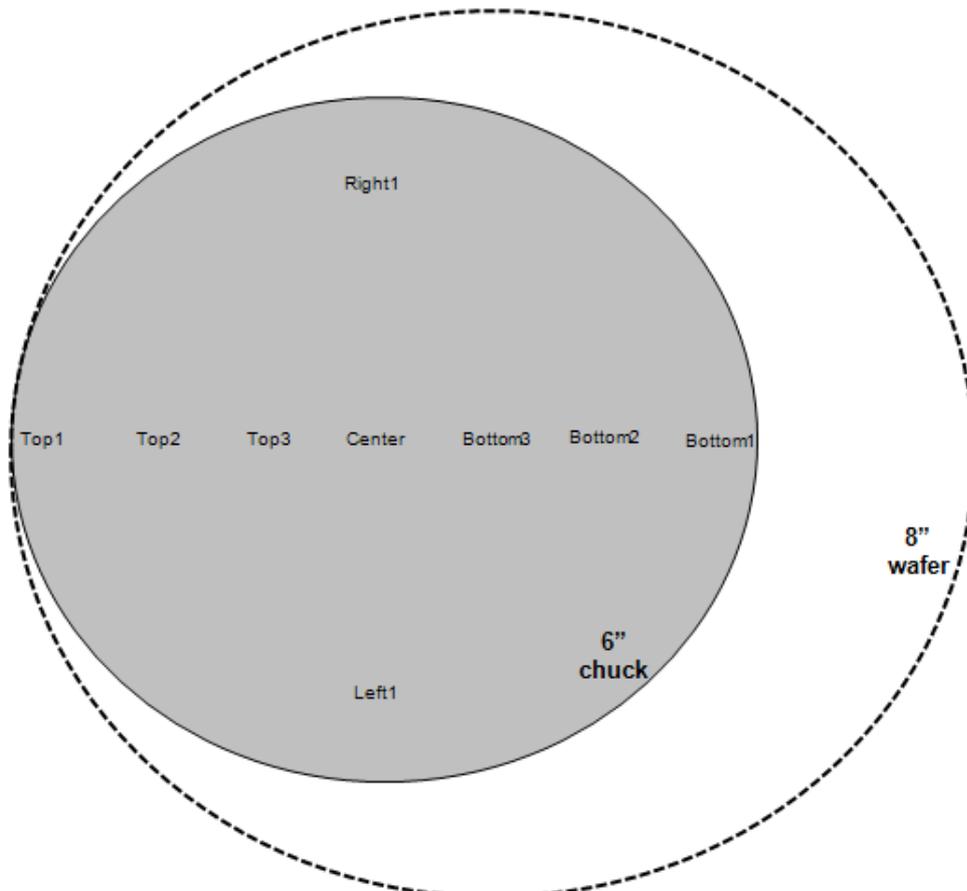
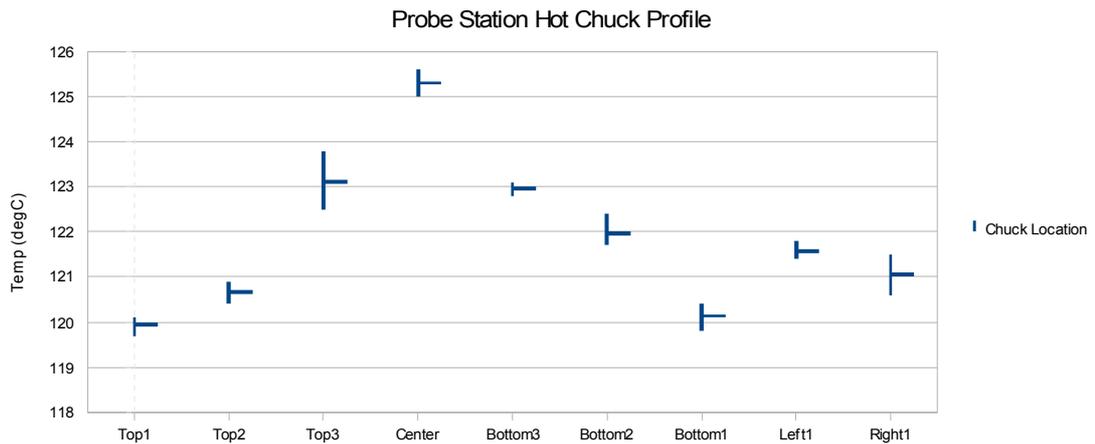


Figure 4.10: Everbeing probe station hot chuck profile with an 8" wafer sample.

Table 4.3: Additional set of measurement at 3 locations with Temperature Controller set at 125 °C.

<b>Chuck Location</b>	<b>Min (°C)</b>	<b>Max (°C)</b>	<b>Mean (°C)</b>
Top2	100.9	101.9	101.4
Center	104.6	105.2	104.9
Bottom2	101.3	102.1	101.8

#### **4.5.3 Constraints in using Everbeing hot chuck**

The results from Table 4.2 show there is a difference between the temperature set at the Temperature Controller and the temperature on the wafer surface. The difference is about 20 °C, e.g. when the target is to achieve 125 °C on the wafer surface that is corresponded to the center of the hot chuck, the Temperature Controller needs to set to 145 °C. This correlation is further confirmed by an additional set of measurements with Temperature Controller set at 125 °C, as shown in Table 4.3.

The temperature profile across the hot chuck varies by about 5 °C, which is unavoidable, considering the sample (in this case an 8” wafer) is much larger than the 6” hot chuck and the hot chuck is not enclosed. This creates a constraint in using the Everbeing hot chuck. To ensure the consistency in performing the NBTI experimental study, the temperature gradient effect need to be minimized. Hence, it is advisable to position the wafer in such a way that the transistor targeted for stressing will always be close to the center of the hot chuck when performing NBTI stress, to maintain the temperature to be consistent across different transistors for the same stress temperature.

If a different size or type of sample is being used, a new temperature profiling needs to be carried out, as the rate of temperature dissipation will be different.

## **4.6 NBTI Study with Conventional DC Stress-Measure-Stress Method**

With the Keithley 4200 SCS correlated to the production tester from the wafer fabrication plant and the probe station hot chuck temperature profile determined, the project can move on to the experimental phase. The NBTI experimental study will commence with the conventional DC Stress-Measure-Stress method. The main purpose of performing the NBTI characterization using conventional DC Stress-Measure-Stress method is to establish a baseline for this 0.18  $\mu\text{m}$  CMOS process technology and also builds an understanding on NBTI degradation. The experimental steps for this method have been described in Section 3.7.1. During initial characterization and interim characterization,  $I_D$ - $V_D$  sweep and  $I_D$ - $V_G$  sweep will be performed and the  $I_{\text{DSAT}}$ ,  $V_{\text{TEXT}}$  and  $V_{\text{TCI}}$  will be extracted.

### **4.6.1 Experimental set-up and conditions**

The transistor selected for conventional DC SMS method will be the 10/0.18  $\mu\text{m}$  thin oxide ( $T_{\text{ox}} = 2.9 \text{ nm}$ ) pMOSFET, as this is the core transistor being used in this 0.18  $\mu\text{m}$  CMOS process. The 8" wafer will be placed on the hot chuck of the probe station. The temperature controller will be set at 20  $^{\circ}\text{C}$  higher than target temperature on center of hot chuck. The bond pads for the targeted pMOSFET will be probed using the micro-positioners, which are connected to the SMUs and GNUs of the Keithley system. The temperature on the wafer will be stabilized for 5 minutes after the temperature controller reached the set temperature. The temperature on the target pMOSFET will be measured with Fluke 568 thermometer before commencement of measurement and stress tests.

The stress voltage applied to the gate needs to be lower than the gate oxide breakdown voltage ( $\sim -5.9 \text{ V}$ , as per production data from wafer fabrication plant). In any reliability stress test with elevated stress condition, it is always advisable to use

stress condition as close to the use condition as possible, by taking into account the time requires to reach the target level of degradation. Trial experiment runs are carried out on 10/0.18  $\mu\text{m}$  pMOSFET on current equipment setup to estimate the stress voltage requires to complete the experimental study within a reasonable time frame. From the trial experiment runs, the stress voltage applied to the gate of pMOSFET will be -2.8 V, -3.0 V, -3.2 V & -3.4 V at temperature of 125  $^{\circ}\text{C}$ . The biasing configuration during stress is illustrated in Figure 4.11 and the stress conditions with sample size are summarized in Table 4.4.

Table 4.4: Stress conditions for conventional DC SMS.

pMOSFET	10/0.18 $\mu\text{m}$	10/0.18 $\mu\text{m}$	10/0.18 $\mu\text{m}$	10/0.18 $\mu\text{m}$
Temperature	125 $^{\circ}\text{C}$	125 $^{\circ}\text{C}$	125 $^{\circ}\text{C}$	125 $^{\circ}\text{C}$
Gate (SMU1)	-2.8 V	-3.0 V	-3.2 V	-3.4 V
Drain (SMU2)	0 V	0 V	0 V	0 V
Source (SMU3)	0 V	0 V	0 V	0 V
Bulk (GNDU)	GND	GND	GND	GND
Sample size	5	5	5	5

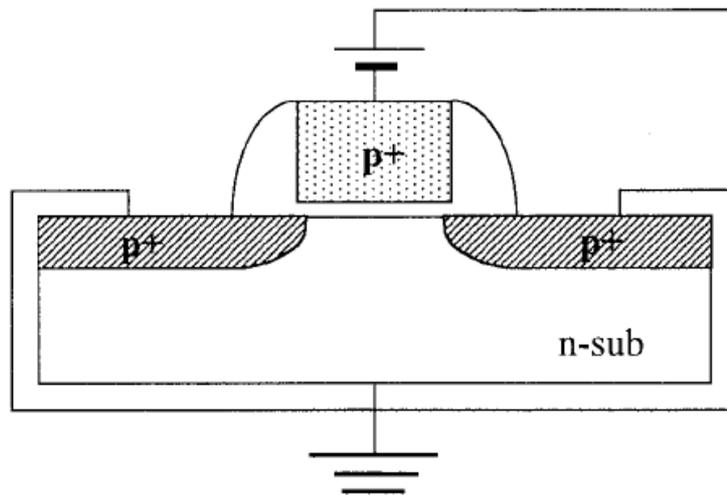


Figure 4.11: Bias configuration during NBTI stress of a pMOSFET (Yamamoto, 1999).

## 4.7 The Effect of Temperature, Gate oxide Thickness and Channel Length on NBTI Degradation

NBTI degradation is highly dependent on temperature and IC in normal operating condition will dissipate power, raising the temperature of the IC and also the transistor. Hence, it is important to understand the temperature effect on NBTI degradation. To study the temperature effect, experiments conducted at other temperatures (105 °C and 145 °C) will be carried out. This enables the extraction of activation energy ( $E_a$ ) that is due to the temperature effect. These experiments will use the same stress voltage applied to the gate as in Table 4.4, with the same sample size. The variation of temperature (105 °C and 145 °C.) is achieved by setting the Temperature Controller to 125 °C and 165 °C , taking into account the offset determined in Section 4.5.2.

Typically in IC fabricated on 0.18  $\mu\text{m}$  dual gate oxide CMOS process, there will be some transistors designed with channel length larger than 0.18  $\mu\text{m}$ , as required by some of the circuitry. The NBTI degradation of this larger channel length pMOSFET needs to be assessed. For this experiment, the variation of channel length is achieved by stressing thin oxide pMOSFET with larger channel length. Based on the pMOSFET selection criteria determined in Section 4.3, the 10/10  $\mu\text{m}$  thin oxide pMOSFET (see Figure 4.4, pMOSFET in pad 1 of ET02) will be subjected to the NBTI stress, using the stress conditions and sample size found in Table 4.4. This will enable a good comparison with 10/0.18  $\mu\text{m}$  thin oxide pMOSFET results.

Another aspect of transistor geometric effect to NBTI degradation is the gate oxide thickness (Alam and Mahapatra, 2005). The input/output circuitry is often designed with transistors operating at higher voltage, which is 3.3 V in this 0.18  $\mu\text{m}$  dual gate oxide CMOS process. To operate transistor at such voltage level, thicker gate

oxide ( $T_{ox} = 5.6 \text{ nm}$ ) is required. This variation of gate oxide thickness can be achieved by using  $10/0.30 \text{ }\mu\text{m}$  pMOSFET (minimum channel length for thick oxide) in ET04 in Figure 4.4. The stress voltage applies to the gate of thick oxide pMOSFET will be set at  $-6.07 \text{ V}$ ,  $-5.70 \text{ V}$ ,  $-5.32 \text{ V}$  and  $-5.0 \text{ V}$ , lower than the gate oxide breakdown voltage of  $-11.4 \text{ V}$  (as per production data from wafer fabrication plant). These voltage levels are selected based on the need to maintain the same electric field across the gate oxide as those stress voltages in  $10/0.18 \text{ }\mu\text{m}$  thin oxide pMOSFET stress. The stress conditions for thick oxide  $10/0.30 \text{ }\mu\text{m}$  pMOSFET is summarized in Table 4.5. Similarly, the plan is to apply each stress condition to 5 thick oxide pMOSFETs.

Table 4.5: Stress conditions for thick oxide pMOSFET.

pMOSFET	10/0.30 $\mu\text{m}$	10/0.30 $\mu\text{m}$	10/0.30 $\mu\text{m}$	10/0.30 $\mu\text{m}$
Temperature	125 °C	125 °C	125 °C	125 °C
Gate (SMU1)	-5.0 V	- 5.32 V	- 5.70 V	- 6.07 V
Drain (SMU2)	0 V	0 V	0 V	0 V
Source (SMU3)	0 V	0 V	0 V	0 V
Bulk (GNDU)	GND	GND	GND	GND
Sample size	5	5	5	5

#### 4.8 Alternative to Advance Fast Measurement Technique

The last phase of this NBTI experimental study will look into the fast measurement techniques during the interim characterization to reduce the “dead” time. As described in Section 3.7, various advance fast measurement techniques had been demonstrated, by employing new hardware, additional proprietary circuitry and/or customized software developed in-house, to reduce the “dead” time down to micro-second or nano-second. The implementation of these advance fast measurement techniques on standard test system with DC SMUs (such as the current equipment setup on Keithley 4200 SCS) is not feasible. Hence, these fast measurement techniques such as On-The-Fly, Fast Switching, Ultra Fast  $V_{TH}$ , Fast  $V_G$  Ramp and Ultra Fast Switching will not be evaluated in this experimental study.

#### 4.8.1 Optimized $I_D$ - $V_G$ sweep ( $V_{THopt}$ )

In an attempt to assess the effect of “dead” time on NBTI degradation characteristic, an optimized  $I_D$ - $V_G$  sweep for interim characterization is proposed here. This optimized  $I_D$ - $V_G$  sweep makes use of the characteristic of extracting  $V_{TCI}$  from  $I_D$ - $V_G$  sweep data. Section 3.2 explained in detail the measurement and extraction of threshold voltage,  $V_{TEXT}$  and  $V_{TCI}$ . Comparing between the two threshold voltage measurement methods,  $V_{TCI}$  can be determined faster than  $V_{TEXT}$ , as  $V_{TCI}$  is determined from interpolation of two  $V_G$  value at the fixed  $I_0 \cdot W/L$ . Therefore, only a short range of  $I_D$ - $V_G$  sweep measurement that centers around the  $V_{TCI-0}$  ( $V_{TCI}$  before stress) is required. As a comparison,  $V_{TEXT}$  required a full  $I_D$ - $V_G$  sweep in order to generate the maximum point in transconductance to perform the extrapolation to obtain threshold voltage. This involves more measurement points being made, and the total measurement time to obtain  $V_{TEXT}$  will be longer than  $V_{TCI}$ . This difference is described graphically in Figure 4.12.

For this Optimized  $I_D$ - $V_G$  sweep, a short sweep with  $V_G$  that centers at  $V_{TCI-0}$ , ( $V_G = V_{TCI-0} \pm 0.1$  V) is used during the interim characterization. This short range of  $\pm 0.1$  V is selected to provide enough room to capture all the changes in  $V_{TCI}$  during NBTI stress, based on the results of Conventional DC Stress-Measure-Stress method (Figure 4.13). The extracted threshold voltage from the Optimized  $I_D$ - $V_G$  sweep will be named as  $V_{THopt}$ , to distinguish it from  $V_{TCI}$  measured in conventional DC SMS method.

During the conventional DC SMS, the interim characterization (full sweep on  $I_D$ - $V_D$  and  $I_D$ - $V_G$ ) takes about 9 seconds. By eliminating the  $I_D$ - $V_D$  sweep and using Optimized  $I_D$ - $V_G$  sweep measurement, the interim characterization “dead” time is reduced to about 1 second, a 9 times improvement over the conventional DC SMS

method.

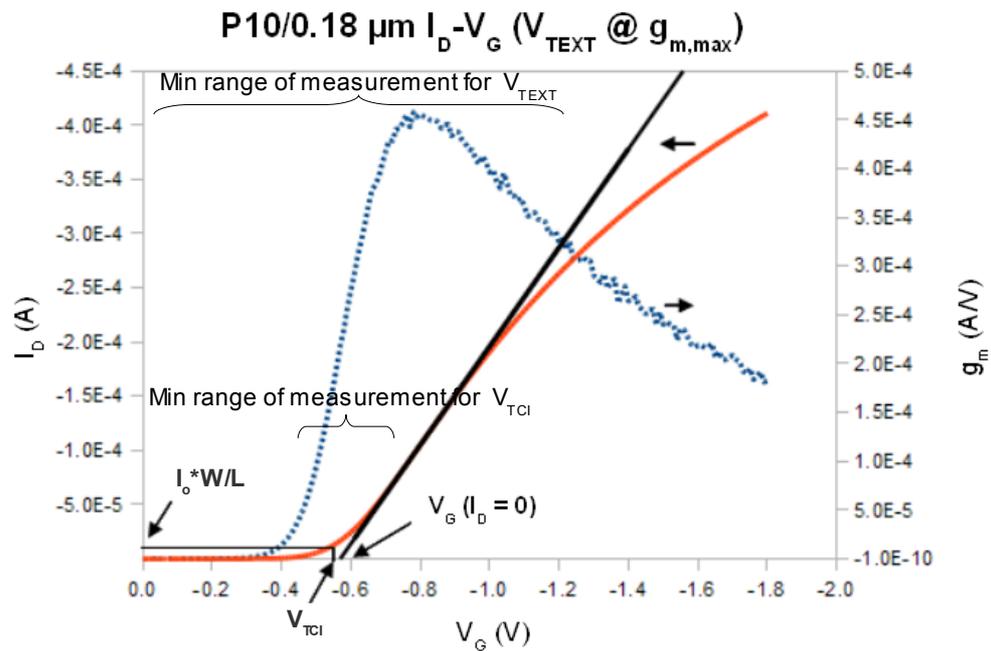


Figure 4.12.  $I_D$ - $V_G$  sweep showing the minimum range of measurement to determine  $V_{TEXT}$  and  $V_{TCI}$ . This data is measured at room temperature.

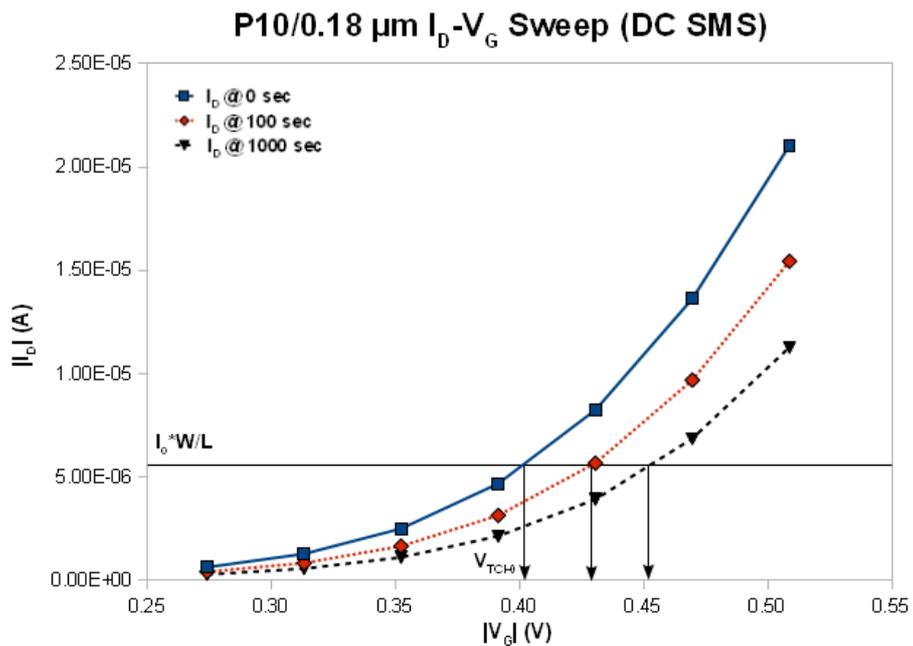


Figure 4.13: A portion of the  $I_D$ - $V_G$  sweep measurement from conventional DC SMS, showing the  $V_{TCI,0}$ , and the subsequent increase in  $V_{TCI}$  as the stress progresses.

## **Chapter 5. NBTI Experimental Results and Discussion**

### **5.1 Introduction**

The Conventional DC Stress-Measure-Stress NBTI experimental data will be reviewed and analyzed in this chapter. The raw data ( $I_D$ - $V_D$  and  $I_D$ - $V_G$  degradation curves) collected on the 0.18  $\mu\text{m}$  thin oxide pMOSFET by the Keithley 4200 SCS will be analyzed. This will be followed by analysis of shift in pMOSFET parameters, such as  $I_{\text{DSAT}}$ ,  $V_{\text{TEXT}}$  and  $V_{\text{TCI}}$ . Next, the NBTI lifetime extraction will be discussed, which will be of interest to the wafer fabrication plant. Data from the experiments on temperature effect, channel length effect and gate oxide thickness effect will be analyzed and reviewed based on parameter degradation and lifetime comparison with 0.18  $\mu\text{m}$  thin oxide pMOSFET. Also, the  $V_{\text{THopt}}$  data from optimized  $I_D$ - $V_G$  sweep will be analyzed, showing its advantage over Conventional DC Stress-Measure-Stress method. Lastly, a discussion on the potential of applying the optimized  $I_D$ - $V_G$  sweep method to the wafer fabrication plant will end this chapter.

### **5.2 Measurement Accuracy and Repeatability**

In any experiment where measurement is required, the data obtained needs to be accurate, stable and repeatable. In the discussion of Keithley measurement system in Section 4.4, it was shown that the measurement data obtained is accurate and repeatable. In fact, the  $I_{\text{DSAT}}$  and  $V_{\text{TCI}}$  data are correlating well to the production data from the wafer fabrication plant. In addition, the hot chuck temperature profiling in Section 4.5 has established a relationship between the temperature controller and target temperature on center of hot chuck. It also identifies the central zone on the hot chuck for the sample to be located, in order to maintain a stable and accurate target temperature on the sample. These initial works in Section 4.4 and 4.5 have ensured the integrity of data obtained subsequently.

### 5.3 $I_D$ - $V_D$ and $I_D$ - $V_G$ Degradation Curves

The raw data collected from Keithley 4200 SCS during conventional DC SMS experiments will be the  $I_D$ - $V_D$  curves and  $I_D$ - $V_G$  curves. When these curves are plotted together, it show visually the degradation of the transfer characteristic of pMOSFET operation under NBTI stress. Figure 5.1 and 5.2 show the example of  $I_D$ - $V_D$  (at  $V_G = -1.8$  V) and  $I_D - V_G$  (at  $V_D = -0.1$  V) curves for  $V_{G, \text{stress}} = -3.4$  V @ 125 °C respectively. Other stress voltages will produce similar type of degradation plot, with different magnitude of degradation.

Plotting the degraded  $I_D$ - $V_D$  and  $I_D$ - $V_G$  curves together will allow a quick visual assessment of the behavior of pMOSFET operation under NBTI stress. It can be seen from Figure 5.1 and 5.2 that degradation on saturation drain current and linear drain current are of different magnitude when the voltage is sweeping from low to high. This indicates that the interface states generated by the NBTI stress is affecting the pMOSFET saturation and linear mode of operation differently . This correlates well to the nature of the flow of charge carriers from source to drain, where less charge carriers are involved in linear mode of operation than saturation mode of operation.

At the point of saturation of the drain current and with the gate of pMOSFET fully turns on ( $V_G = -1.8$ V), the reduction in  $I_{DSAT}$  is seen to increase with stress time. Based on the Reaction-Diffusion model, the reduction in  $I_{DSAT}$  shows the increase in interface states density as the NBTI stress applied longer (Alam et al., 2007). The increase in interface states density is a result of higher number of Si-H bond breakage at the interface, creating more silicon dangling bond. The majority of released hydrogen species will diffuse away from the Si-SiO<sub>2</sub> interface under the influence of gate electric field. From the Reaction-Diffusion model, the released hydrogen species may

recombine with silicon dangling bond, but the probability for this to happen is low when the stress bias remained. As stress time increases, more hydrogen species diffuse away from the Si-SiO<sub>2</sub> interface, creating higher number of the silicon dangling bond. The silicon dangling bond has the tendency to capture holes, which is the majority carrier of conduction in pMOSFET. With the reduction of holes in the inversion channel, the mobility ( $\mu$ ) of holes is reduced and the end result is reduction in  $I_{DSAT}$ . The mobility has been shown to be linearly correlated to  $I_{DSAT}$  in Equation (3.2). This will cause the pMOSFET to lose current drive capability in saturation mode, which is a key feature for transistor in driving the next stage in logic gate circuit.

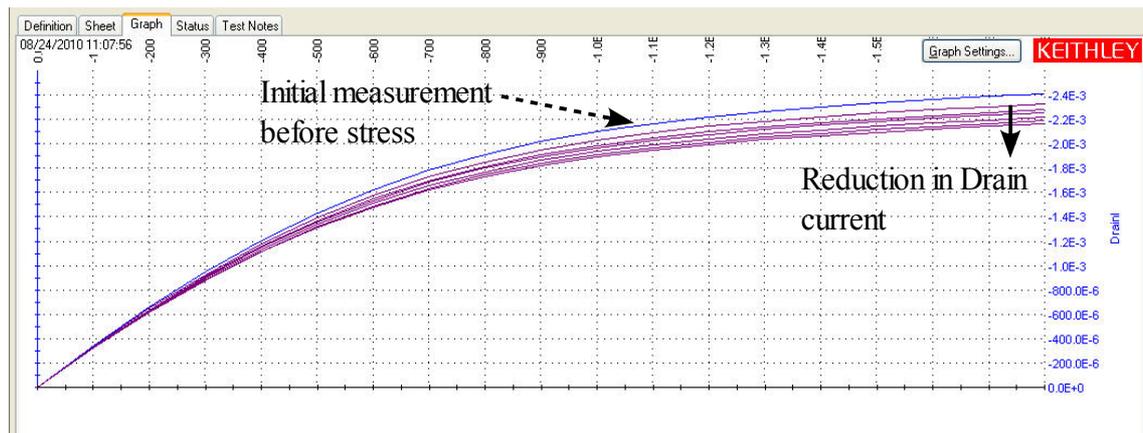


Figure 5.1:  $I_D$ - $V_D$  ( $V_G = -1.8$  V) sweep degradation for  $V_{G, stress} = -3.4$  V, 125 °C.

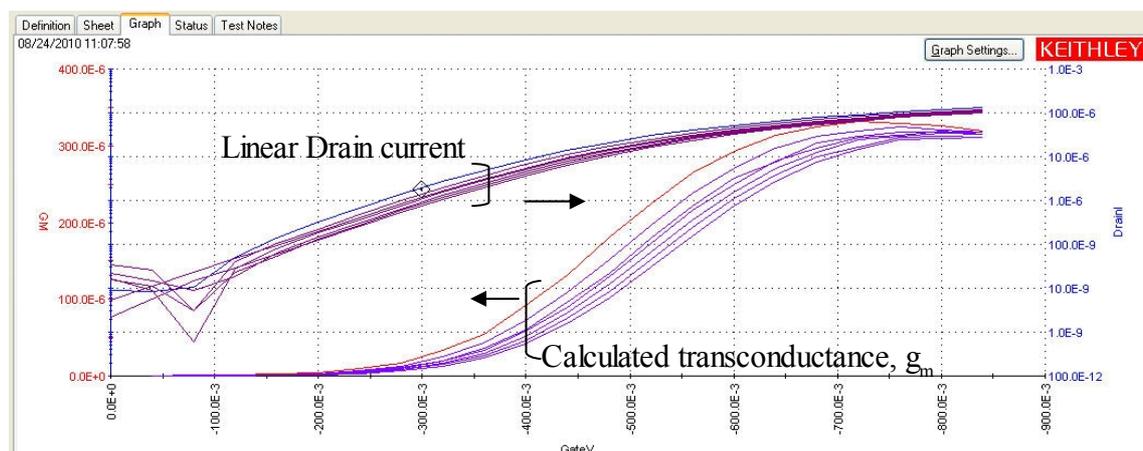


Figure 5.2:  $I_D$ - $V_G$  ( $V_D = -0.1$  V) sweep degradation at  $V_{G, stress} = -3.4$  V, 125 °C.

In the linear mode of operation of pMOSFET (where  $V_D = -0.1$  V), weak inversion is formed in the channel. The linear drain current shows a similar pattern as the saturation drain current, where the  $I_D$ - $V_G$  curves show a uniform reduction as the stress time progresses. This is expected as the mechanism in creating silicon dangling bond and capturing of holes in the inversion channel remains the same in linear mode of operation. The mobility of holes in linear mode of operation is being reduced and so is the linear drain current (Equation (3.1)). The linear mode of operation defines the threshold voltage of the pMOSFET and with the reduction in the linear drain current, the  $V_{TCI}$  will increase, as it takes higher gate voltage to reach the same drain current that will turn on the pMOSFET. In the similar manner, the transconductance will be reduced after the NBTI stress, as can be seen in Figure 5.2. This will result in the increment of  $V_{TEXT}$  after extrapolation. The increase in threshold voltage and decrease in transconductance are attributed to the increase in interface states density and the fixed charge at the Si-SiO<sub>2</sub> interface respectively (Schroder and Babcock, 2003). This will affect the switching speed of the pMOSFET.

#### 5.4 Parameter degradation-vs-Stress Time

The degradation curves in Figure 5.1 and 5.2 show the degradation behavior visually, do not allow the formulation of mathematical model. By using the changes in key parameters ( $I_{DSAT}$ ,  $V_{TEXT}$  and  $V_{TCI}$ ), the shift in these parameters can be plotted versus the stress time for all stress voltages to model the degradation mechanism. Figure 5.3 shows the  $\Delta I_{DSAT}$ -vs-stress time plot from  $I_D$ - $V_D$  sweep for  $V_{G, stress} = -2.8$  V,  $-3.0$  V,  $-3.2$  V and  $-3.4$  V. The parameter degradation (e.g.  $\Delta I_{DSAT}$ ) is expressed in the form of shift in percentage:

$$\Delta I_{DSAT} = \frac{I_{DSAT}(0) - I_{DSAT}(t)}{I_{DSAT}(0)} * 100 \% \quad (5.1)$$

where  $I_{DSAT}(0)$  is before stress,  $I_{DSAT}(t)$  is measured during interim characterization

(JESD90, 2004).

The  $\Delta I_{DSAT}$ -vs-stress time is plotted in a log-log scale plot as per JESD90, 2004. Since not all the stress voltage had been stressed with five 10/0.18  $\mu\text{m}$  pMOSFETs (as per the original plan), the average value of  $\Delta I_{DSAT}$  is used. The fact that all the  $\Delta I_{DSAT}$  data points for each stress voltages line up in a straight line shows that the degradation behavior follows the Power law. The equations in Figure 5.3 are in the same form as Equation (2.19) and also Equation (5) found in JESD90 (2004). This indicates that the interface states,  $N_{it}$  play a dominant role in the degradation of  $I_{DSAT}$ . The time exponent,  $n$  for each of the stress voltage ranges from 0.21  $\rightarrow$  0.23, very close to the  $n$  value of 0.25 in Equation (2.18). This again proves that interface states are mainly responsible for  $I_{DSAT}$  degradation. Similar  $n$  value had been reported in the literature (e.g. Kimizuka et al., 2000; Schroder and Babcock, 2003), in agreement with these results.

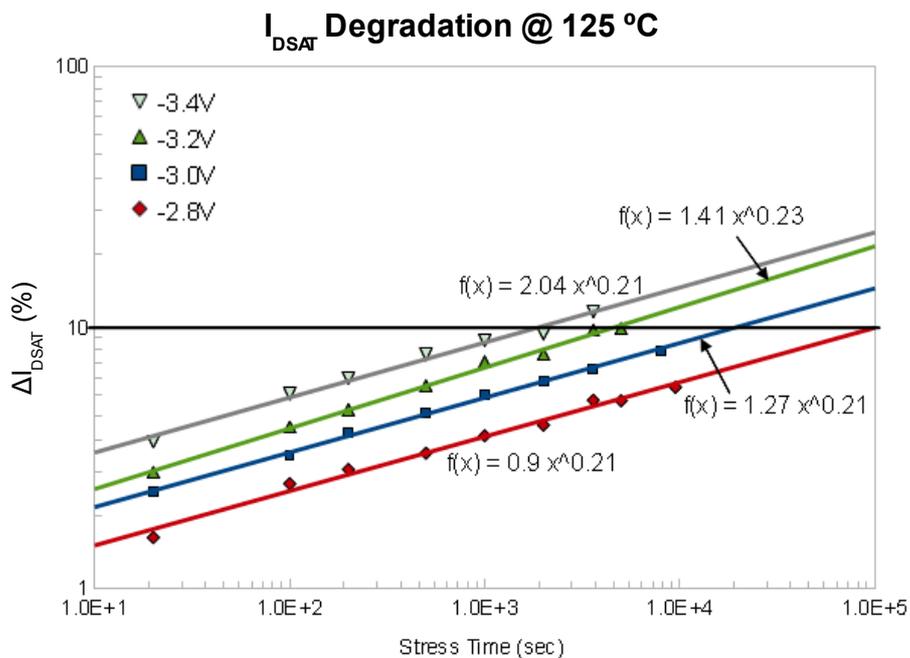


Figure 5.3:  $I_{DSAT}$  degradation-vs-stress time, plotted on log-log scale. The equation shown is in the form of Power law equation. The black horizontal line indicates the 10% shift in  $I_{DSAT}$ .

In the similar manner, for linear mode of operation, the shift of threshold voltage,  $\Delta V_{\text{TEXT}}$  and  $\Delta V_{\text{TCI}}$  versus stress time can be plotted on the log-log scale. Similar to  $\Delta I_{\text{DSAT}}$ , the average value of  $\Delta V_{\text{TEXT}}$  and  $\Delta V_{\text{TCI}}$  for each stress voltage are used. This is shown in Figure 5.4 and 5.5 respectively. The data point for each stress voltage for  $\Delta V_{\text{TEXT}}$ -vs-stress time and  $\Delta V_{\text{TCI}}$ -vs-stress time line up in a straight line, showing the degradation of threshold voltage follows the Power law as in  $\Delta I_{\text{DSAT}}$ .  $\Delta V_{\text{TEXT}}$  has time exponent,  $n$  value that ranges from 0.24  $\rightarrow$  0.25 and  $\Delta V_{\text{TCI}}$  has time exponent,  $n$  value that ranges from 0.22  $\rightarrow$  0.24. Both show  $n$  value that is closer to 0.25 as compared to  $\Delta I_{\text{DSAT}}$ , indicating  $N_{\text{it}}$  is even more dominant in the linear mode than in saturation mode.

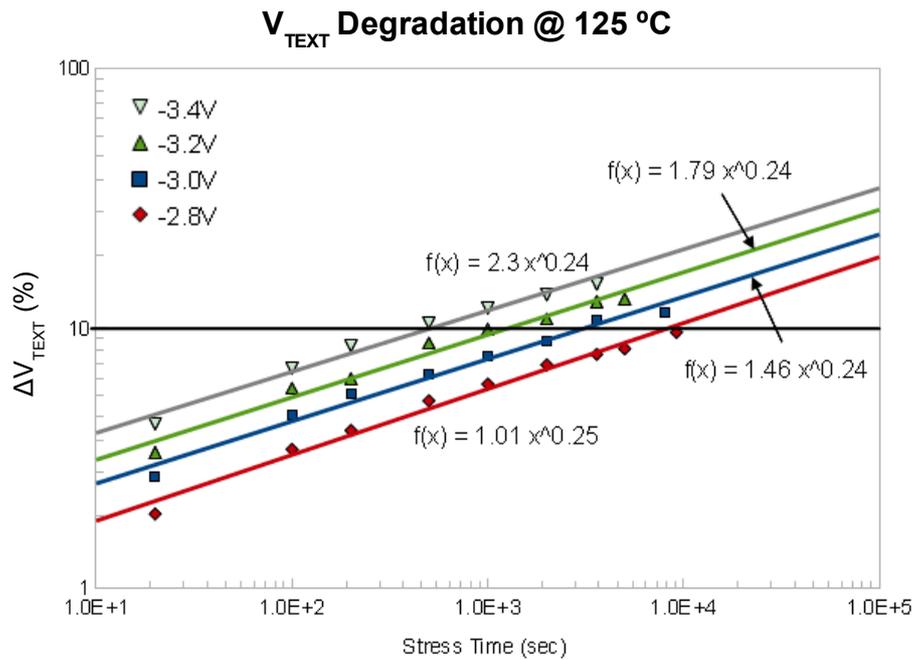


Figure 5.4:  $V_{\text{TEXT}}$  degradation-vs-stress time, plotted on log-log scale.

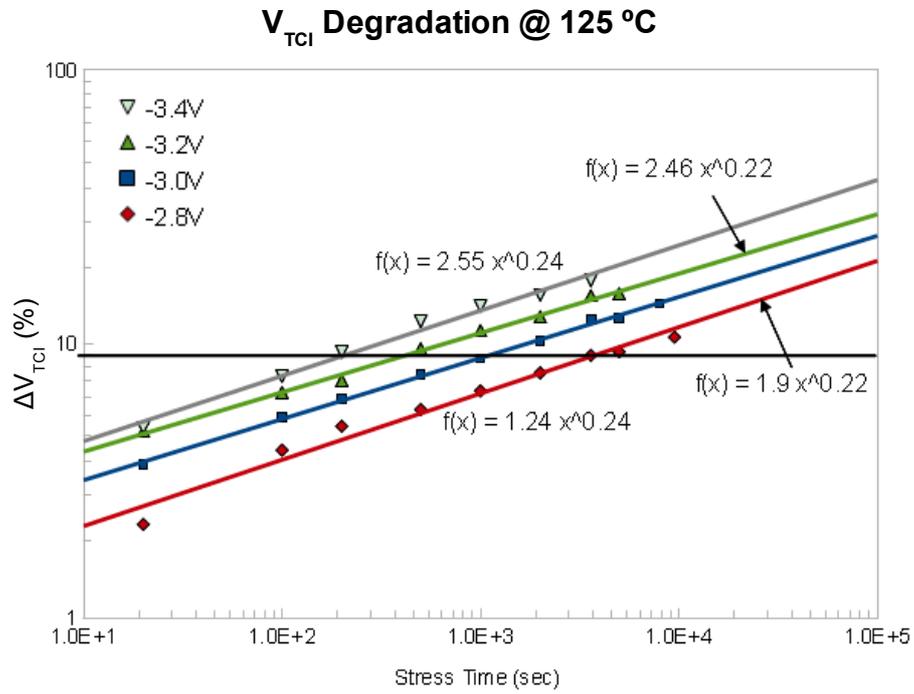


Figure 5.5: V<sub>TCI</sub> degradation-vs-stress time, plotted on log-log scale.

In Section 2.5.3, the time exponent,  $n = 0.25$  had been derived analytically in detail using Reaction-Diffusion model. With  $n$  value close to 0.25 obtained from this experiment, the degradation mechanism seen on  $I_{DSAT}$ ,  $V_{TEXT}$  and  $V_{TCI}$  is dominated by the interface states generation, which had been explained by the Reaction-Diffusion model. Thus, the NBTI degradation behavior on this 0.18  $\mu\text{m}$  CMOS process has been modeled into the form that is derived analytically using the Reaction-Diffusion model.

Comparing between the three parameters mentioned above, the degradation on  $I_{DSAT}$  is the slowest as compared to  $V_{TEXT}$  and  $V_{TCI}$  for the same stress voltage. This can be seen in Figure 5.6 and is in accordance to the  $n$  value of  $\Delta I_{DSAT}$  which is slightly smaller than  $n$  value of  $\Delta V_{TEXT}$  and  $\Delta V_{TCI}$ . This shows that threshold voltage in the linear mode of operation is more sensitive to the NBTI stress than drive current ( $I_{DSAT}$ ) in the saturation mode of operation. This is in agreement with the literature where majority of the reported data are on threshold voltage degradation.

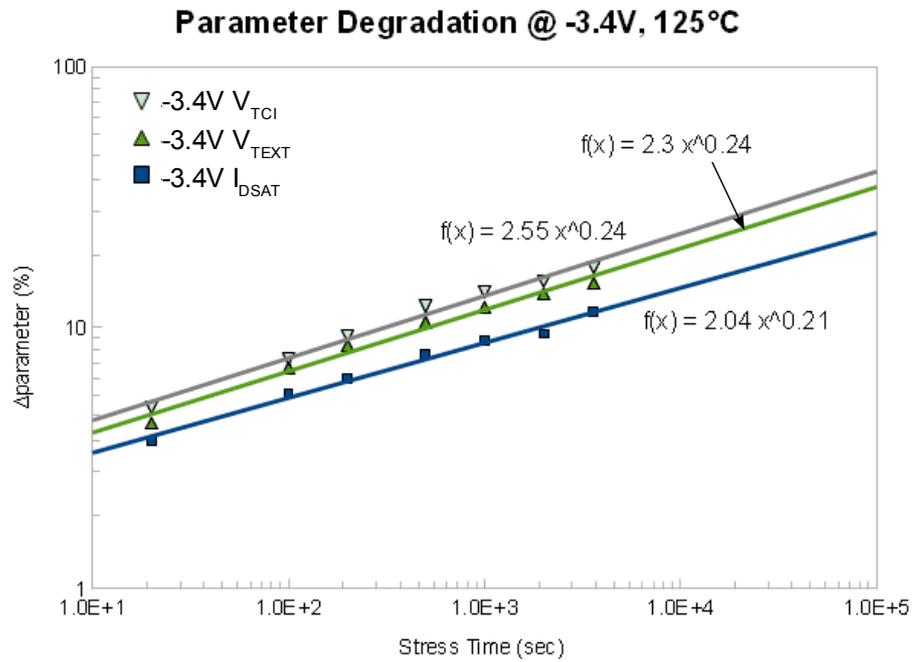


Figure 5.6: Comparison of degradation between  $I_{DSAT}$ ,  $V_{TEXT}$  and  $V_{TCI}$  at stress voltage of -3.4 V, 125 °C.

Among the two methods of extracting threshold voltage from  $I_D$ - $V_G$  sweep,  $V_{TCI}$  shows a slightly shorter time to reach 10% degradation than  $V_{TEXT}$ , at the similar rate as  $V_{TEXT}$  ( $n$  exponent is the same). This shows that  $V_{TCI}$  is the most sensitive parameter to be used to monitor pMOSFET degradation in NBTI stress. This can be seen in recent published results on NBTI, where  $V_{TCI}$  is the preferred method to extract threshold voltage as compared to  $V_{TEXT}$  (Ang et al., 2005; Brisbin et al., 2008; Ho et al., 2009).

## 5.5 NBTI Lifetime Extraction

In addition to parameter degradation behavior, it is of the wafer fabrication plant interest to gauge the lifetime of pMOSFET under NBTI stress. To extract the lifetime, Time-to-Fail (TTF) for each  $V_{G, stress}$  needs to be determined on the parameter degradation-vs-stress time plot. TTF is the time a pMOSFET is considered failing when any of the critical parameters exceed specification limit. The general practice for TTF extraction is to set a failing criteria (specification limit) for each parameter (JESD90, 2004; Ershov et

al., 2003; Tsujikawa et al., 2003). An example is a 10% shift in a particular parameter as the device failing criteria. Another approach is to set a certain absolute shift for each parameter, e.g.  $\Delta I_{\text{DSAT}} = 0.3 \text{ mA}$  or  $\Delta V_{\text{TH}} = 50 \text{ mV}$  as the failing criteria. This approach needs an in-depth understanding on the operation of the IC and the technology node where it is sensitive to such parameter shift. By using a generic approach in this analysis, the TTF extraction will be based on the 10% shift on  $I_{\text{DSAT}}$ ,  $V_{\text{TEXT}}$  &  $V_{\text{TCI}}$  as the failing criteria.

In Figure 5.3, 5.4 and 5.5, for each stress voltage, the value where the degradation fit line crosses 10% will be plotted on the TTF-vs- $V_{\text{G, stress}}$  semi-log plot as shown in Figure 5.7. The TTF data lines up in a straight line with y-axis in log scale, and this shows the TTF can be related to stress voltage with an exponential dependence which can be fitted with an analytical model in the following form (Ershov et al., 2003):

$$\tau = A * e^{(-X * V_{\sigma})} \quad (5.2)$$

where  $\tau$  is the lifetime,  $A$  is a constant and  $X$  is the exponential parameter for stress voltage.

As the TTF data is generated from experiment performed at 125 °C, this lifetime model is valid for this temperature only. The fitting equations found in Figure 5.7 are in the same form as Equation (5.2), validating the exponential dependence of TTF to stress voltage. The TTF-vs- $V_{\text{G, stress}}$  plot will be extrapolated back to the nominal gate operating voltage ( $V_{\text{nom}}$ ). For safeguarding, this  $V_{\text{nom}}$  is typically set at 10% higher than the operating voltage ( $V_{\text{DD}} = -1.8 \text{ V}$ ) for this thin oxide pMOSFET.

$$V_{\text{nom}} = 1.1 * V_{\text{DD}} \quad (5.3)$$

Hence, the lifetime will be extrapolated back to  $V_{\text{nom}} = -1.98 \text{ V}$ .

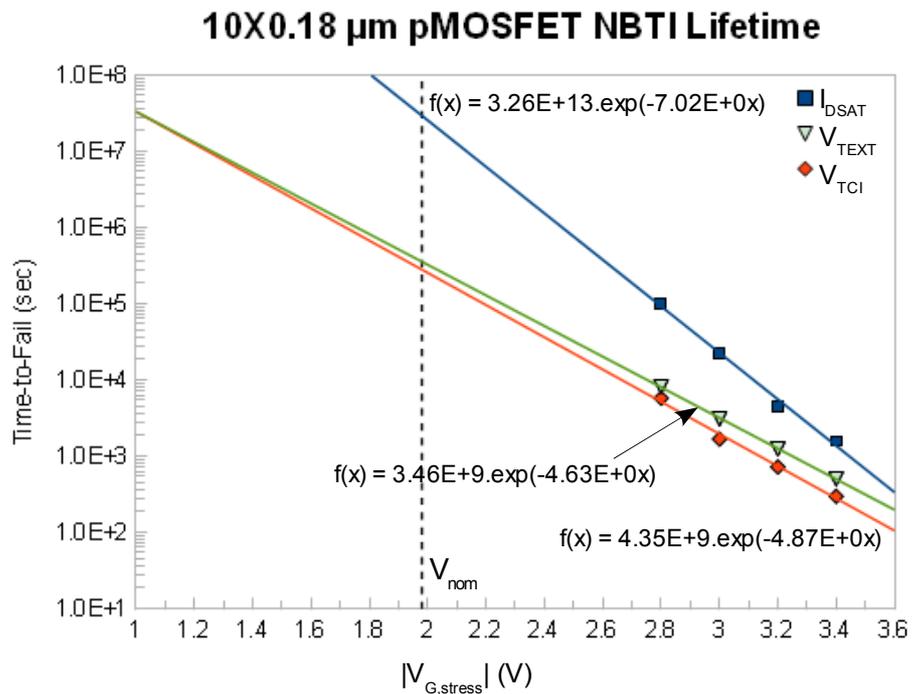


Figure 5.7: TTF-vs- $V_{G, stress}$  plotted on semi-log plot to extract the lifetime at 125 °C. The equations shown are in the form of Exponential law equation.

From Figure 5.7,  $I_{DSAT}$  has a lifetime that is about 2 order of magnitudes larger than  $V_{TEXT}$  and  $V_{TCI}$ . This again confirms that the degradation during saturation mode is less severe than linear mode, indicating the generated interface states have less effect on the drain current when the drain current magnitude is high (saturation drain current is 6 times larger than linear drain current). Though there is a small difference between the  $V_{TEXT}$  and  $V_{TCI}$  lifetime, where  $V_{TCI}$  is slightly lower than  $V_{TEXT}$ . This indicates that  $V_{TCI}$  is the most sensitive parameter to NBTI degradation and will be used as the key monitoring parameter in the later phase of this experimental study.

## 5.6 The Effect of Temperature on NBTI Degradation

Modern IC is being used at various conditions in actual field application, and may have hot spots in circuit design leading to large temperature gradients across a chip, such as those on microprocessor or System-On-a-Chip IC. In addition, change in environment temperature is also known to affect IC performance and reliability. It will be the interest

of the wafer fabrication plant to understand the effect of temperature on this degradation mechanism. This can be achieved by extracting the best estimation of thermal activation energy to determine the acceleration factor under the common understanding of the intended failure mechanism. The acceleration factor will allow the estimation of device failure rate in system application (JEP122F, 2010).

### 5.6.1 Arrhenius Plot on $\Delta V_{TCI}$

Figure 5.8 shows the Arrhenius plot of  $\Delta V_{TCI}$ , which is plotted against the  $1/k_B T$  where  $k_B$  is the Boltzmann constant and  $T$  is the temperature in Kelvin. As indicated in Section 4.7, the range of temperature used is 105 °C, 125 °C and 145 °C, which are shown in Figure 5.8. The  $\Delta V_{TCI}$  for both stress voltages of -3.0 V and -3.2 V fit nicely into a straight line, showing a strong dependence of NBTI degradation on temperature variation. The extraction of activation energy will be discussed in the following section.

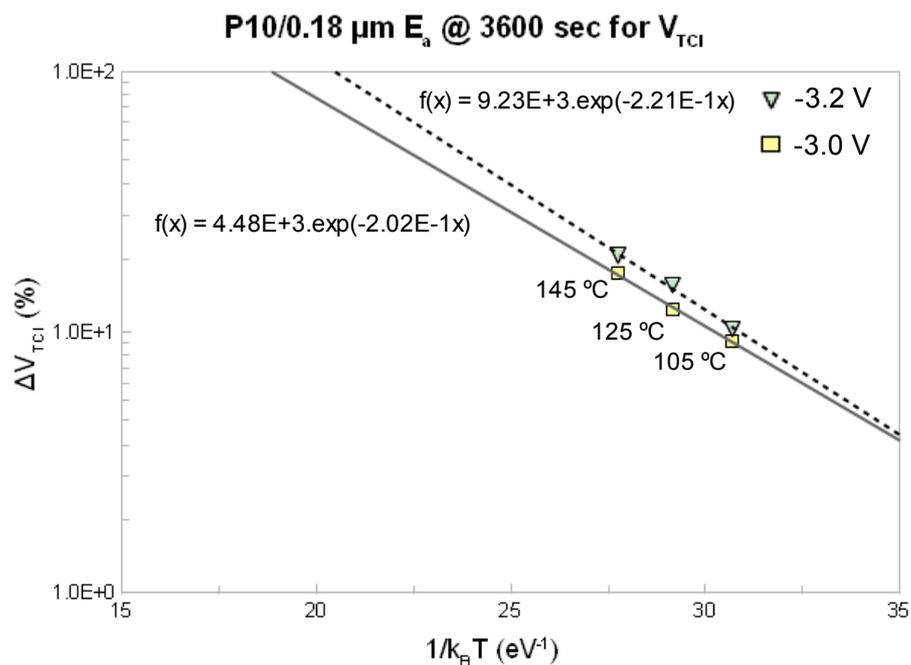


Figure 5.8:  $\Delta V_{TCI}$ -vs- $1/k_B T$  plotted on semi-log plot to extract the  $E_a$  at  $V_G = -3.0$  V, -3.2 V. The equations shown are in the form of Exponential law equation.

### 5.6.2 Arrhenius Model and Activation Energy

The Arrhenius model is built on the theory that a state of equilibrium exists between the reactants and product of a reaction, which is separated by a finite energy difference. This difference in energy is referred to as the activation energy. The Arrhenius model was originally formulated in the field of chemistry reaction study by Dr. Savante Arrhenius (Lall et al., 1995) in the following form:

$$R = R_{ref} * \exp\left(\frac{-E_{a-chrt}}{k_B T}\right) \quad (5.4)$$

where  $R$  is the reaction rate,  $R_{ref}$  is the reaction rate at a reference temperature,  $E_{a-chrt}$  is the activation energy of chemical reaction,  $k_B$  is the Boltzmann constant and  $T$  is steady state temperature. Equation (5.4) has been adopted by other studies in different area such a kinetic theory, thermodynamics, diffusion and etc, in the forms that contain the exponential term similar to Arrhenius model in Equation (5.4).

Similarly, the semiconductor industry has adopted Arrhenius model to characterize the effect of temperature on IC reliability. The basic thermal acceleration equation that describes the relationship between acceleration factor and activation energy is given as below (JEP122F, 2010):

$$A_T = \exp\left[\left(\frac{-E_a}{k_B}\right)\left(\frac{1}{T_{st}} - \frac{1}{T_{op}}\right)\right] \quad (5.5)$$

where  $A_T$  is the thermal acceleration factor,  $E_a$  is the activation energy,  $T_{st}$  is the stress temperature,  $T_{op}$  is the operating temperature. Equation (5.5) has been modified further to suit the different type of failure mechanisms, such as those discussed in Section 2.3 (TDDB) and Section 2.4 (EM).

### 5.6.3 The Effect of Thermal Acceleration in NBTI Mechanism

Similar to TDDB and EM, NBTI degradation is strongly influenced by temperature. This can be seen from the discussion of analytical solution of Reaction-Diffusion model in Section 2.5.3, where both the reaction process (generation of interface states) and diffusion process (hydrogen species) are known to be dependent on temperature (Huard et al., 2003; Ang et al., 2005). Therefore, the NBTI stress experiments are conducted at elevated temperature (e.g. 105 °C, 125 °C, 145 °C) on the 10/0.18 μm pMOSFET, much higher than the normal operating temperature (typically quoted at 55 °C for most IC data sheet). This is to accelerate the degradation mechanism, so that the NBTI stress experiments will complete in a reasonable amount of time. At elevated temperature, the bond breaking process and diffusion of hydrogen species response to this thermal activation according to the Arrhenius model. As such, the transistor parameter shift under the influence of elevated temperature can be expressed as (Ohring, 1998):

$$\Delta parameter \propto A * \exp\left(\frac{-E_a}{k_B T}\right) \quad (5.6)$$

where  $A$  is a constant,  $E_a$  is the activation energy,  $k_B$  is the Boltzmann constant and  $T$  is temperature in Kelvin. By stressing the pMOSFET at the same gate voltage at different temperatures, Equation (5.6) can be rewritten as:

$$\Delta parameter = A * \exp\left(\frac{-E_a}{k_B T}\right) \quad (5.7)$$

With Equation (5.7), the  $\Delta V_{TCI}$  vs  $1/k_B T$  can be plotted on a semi-log plot and the activation energy ( $E_a$ ) will be easily extracted based on the data from three temperatures, collected for the same stress time. The activation energy is essentially the energy that must be overcome in order for the traps to be generated, leading to the shift in transistor parameters.

From Figure 5.8, it can be seen that  $\Delta V_{\text{TCI}}$  strongly depends on temperature as it follows the Arrhenius law. The equations shown in Figure 5.8 are in the same form as Equation (5.7). Therefore, the  $E_a$  at -3.0 V is 0.202 eV and at -3.2 V is 0.221 eV. These values are close to each other and can be considered to be in the same range, in agreement with Mitani et al. (2002), Schroder and Babcock (2003) and Chakravarthi et al. (2004). It shows that regardless of  $V_{\text{G, stress}}$ , the same mechanism is responsible for the  $V_{\text{TCI}}$  degradation, given the similar  $E_a$  value (Strong et al., 2009). This  $E_a$  value consists both the energy required to dissociate the Si-H bond and the energy for the hydrogen species to diffuse from the interface.

### **5.7 The Effect of Channel Length (L) on NBTI Degradation**

The experiments performed so far had been focused on pMOSFET with minimum channel length,  $L = 0.18 \mu\text{m}$ . This device dimension inevitably falls under the Short Channel Effect (SCE). Generally, the description for SCE phenomenon found in the literature showed the following characteristics (Acuthan and Bhat, 2007):

- Threshold voltage becomes smaller with channel length reduction.
- Drain current does not saturate.
- Transconductance becomes independent of gate voltage.

The 10/0.18  $\mu\text{m}$  pMOSFET measured at room temperature shows such characteristic, where the drain current does not saturate after inversion channel pinch-off as per Equation 3.2. In fact, the drain current is increasing linearly for  $V_{\text{D}} < -1.0 \text{ V}$ , at a much smaller slope than the linear region (Figure 4.7). This is due to the channel length modulation, where the effective channel length is reduced after the transistor is pinched-off. In the case of short channel device, this reduction is not negligible and accounts for a higher percentage of effective channel length loss as compared to long channel length device. Thus, for short channel length device, the electric field from the drain has a

larger influence on the charge in the inversion channel under the gate. The pinch-off point (indicated as  $V_{DSAT}$  in Figure 5.9) is moving towards the source when the drain voltage continue to increase, creating a modulated distance of  $\Delta L$ . By substituting  $L$  with  $(L-\Delta L)$  into Equation (3.2), the drain current for short channel device becomes (Acuthan and Bhat, 2007):

$$I_{D-SCE} = \mu C_{ox} \frac{W}{L-\Delta L} \left[ \frac{(V_G - V_{TH})^2}{2} \right] \quad (5.8)$$

Combining Equation (3.2) with Equation (5.8),  $I_{D-SCE}$  can be related to  $I_{DSAT}$  as below:

$$I_{D-SCE} = \frac{I_{DSAT}}{1-(\Delta L/L)} \quad (5.9)$$

With  $\Delta L$  directly proportional to  $V_D$ , the denominator of Equation (5.9) will be smaller when  $V_D$  increase, hence the drain current will be increasing after pinch-off for short channel device.

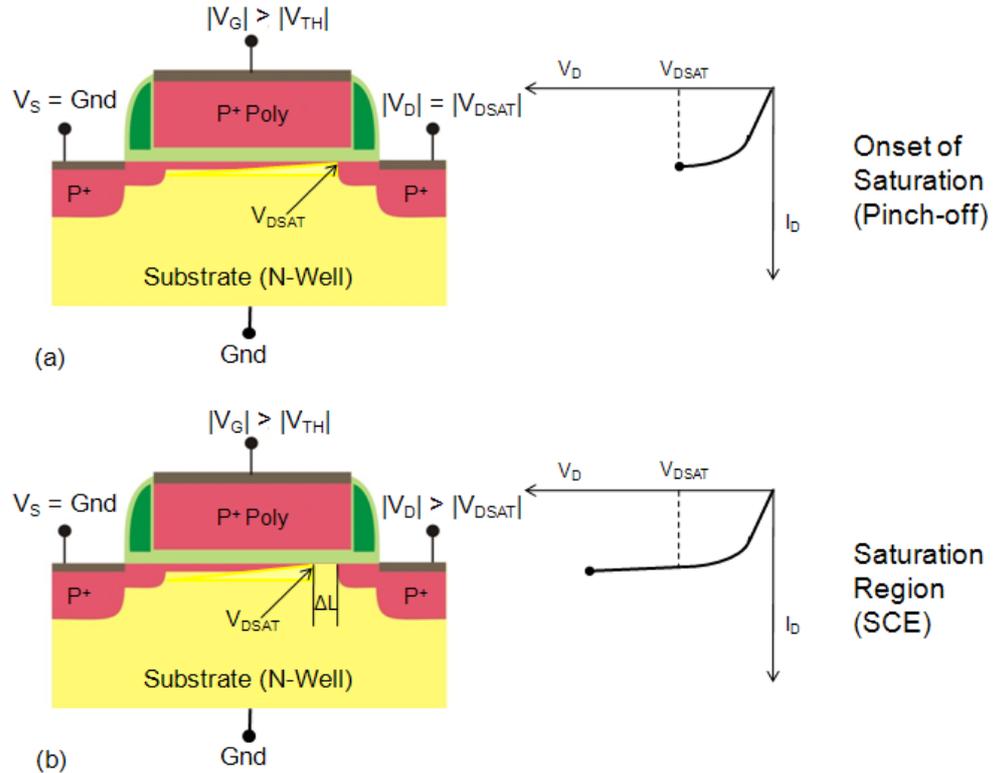


Figure 5.9: (a) pMOSFET at the point of pinch off. (b) pMOSFET beyond pinch off with short channel effect (Acuthan and Bhat, 2007).

The results shown in Section 5.4 & 5.5 established a baseline for NBTI degradation for short channel device, which formed the majority of transistors in a typical IC for such technology node. Long channel length device is rarely used, but in certain circuit functionality, such as analog circuit, transistors are designed at larger than minimum design rule to minimize the process variation effect on transistor parameter (e.g. threshold voltage mismatch). Therefore, an understanding of channel length effect on NBTI degradation will be of interest to the wafer fabrication plant. Due to the limitation on the design of the test structure on current wafer sample, the experiment on channel length effect can only be extended to 10/10  $\mu\text{m}$  pMOSFET (as explained in Section 4.3).

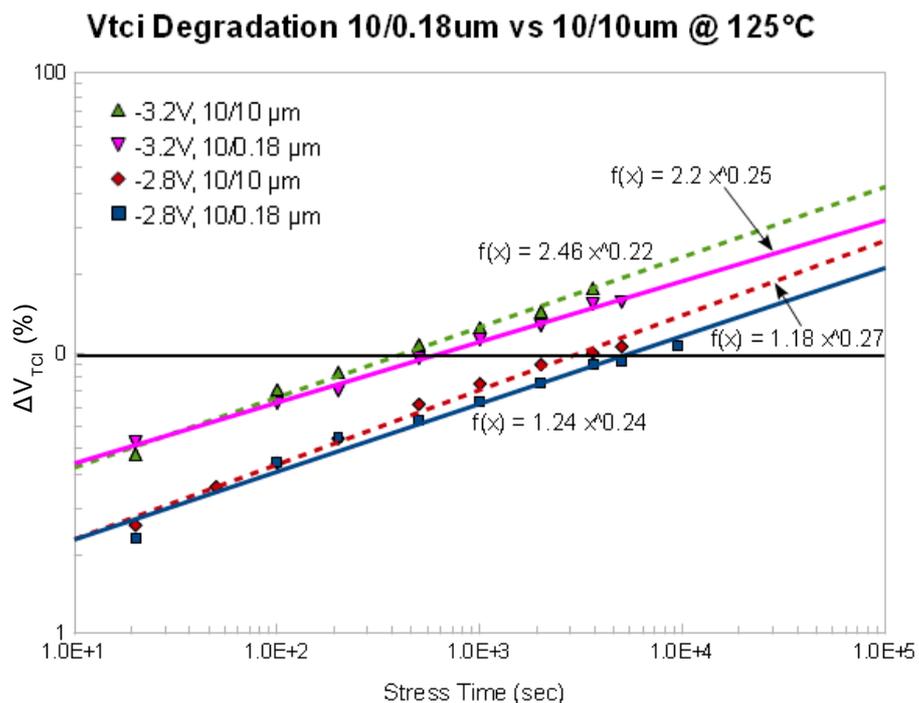


Figure 5.10: Comparison of  $V_{TCI}$  degradation between 10/0.18  $\mu\text{m}$  and 10/10  $\mu\text{m}$  at stress voltage of -2.8 V & -3.2 V, 125  $^{\circ}\text{C}$ .

Figure 5.10 shows the  $\Delta V_{TCI}$ -vs-stress time degradation plot for 10/0.18  $\mu\text{m}$  and 10/10  $\mu\text{m}$  at stress voltage of -2.8 V and -3.2 V. The degradation of 10/10  $\mu\text{m}$  follows the Power law, similar to 10/0.18  $\mu\text{m}$ . The extracted time exponent,  $n$  is 0.27 and 0.22

respectively, not very far off from the  $n$  value obtained for 10/0.18  $\mu\text{m}$ . From the degradation plot, it is observed that the 10/10  $\mu\text{m}$  transistor degrades faster than the 10/0.18  $\mu\text{m}$ , with the TTF at 10% shift being smaller for 10/10  $\mu\text{m}$ . In other words, for the same amount of stress time, the 10/10  $\mu\text{m}$  pMOSFET shows a larger  $\Delta V_{\text{TCI}}$  than 10/0.18  $\mu\text{m}$ . This result is in agreement with Cellere et al. (2004) and Pan (2006), but contradictory to the majority reported results in the literature (e.g. Yamamoto et al., 1999; Sekhar et al., 2004; Ho et al., 2009).

As 10/0.18  $\mu\text{m}$  is suffering from SCE, the NBTI degradation on 10/0.18  $\mu\text{m}$  is expected to be more severe. However, comparing with the case of 10/10  $\mu\text{m}$  pMOSFET, the degradation of 10/10  $\mu\text{m}$  pMOSFET is more severe. The difference in degradation behavior can be explained by the role played by fixed oxide charge ( $N_f$ ) and oxide trapped charge ( $N_{\text{ot}}$ ) within the gate oxide (Figure 5.11). The fixed oxide charge is the charge that is trapped near the Si-SiO<sub>2</sub> interface, but not directly adjacent to it. They are primarily due to excess silicon species introduced during oxidation and during post oxidation heat treatment. Fixed oxide charge represents the transition region from Si to SiO<sub>2</sub>, and this thin region of non-stoichiometric oxide typically consists of SiO<sub>2</sub>, SiO, SiO<sub>x</sub>. These charges are positive charge and they are called fixed charges because the applied gate voltage does not affect their occupancy. The charge density depends on the oxidation ambient and temperature, cooling conditions and orientation of silicon surface where oxidation occur (Taur and Ning, 1998).

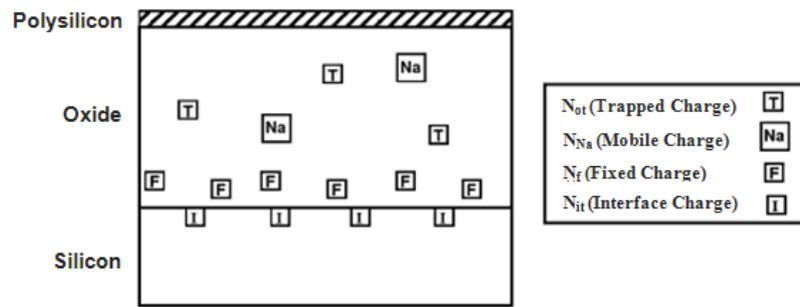


Figure 5.11: Oxide charge in gate oxide (Taur and Ning, 1998).

On the other hand, oxide trapped charges are positive or negative charge that may be due to holes or electrons trapped in the oxide. When electrons are injected (such as avalanche injection, Fowler-Nordheim tunneling, direct tunneling, hot carrier injection) and trap in the bulk of oxide, then  $N_{ot}$  is negative. If electron-hole pairs being generated within the oxide due to ionizing radiation and result in  $N_{ot}$ , then the  $N_{ot}$  will be positive because electrons will move through the oxide and out to the silicon substrate (in the case of pMOSFET) faster than holes towards the polysilicon where as the holes will be trapped within the oxide. Thus the  $N_{ot}$  is positively charged. In addition, the oxidation condition may also introduce electron traps, such as the moisture containing ambient during oxidation process, as reported by Nicollian et al. (1971). The presence of fixed oxide charges and oxide trap charges affect the potential in the channel and also act as charged scattering centers. These charges reduce the mobility of carriers thus altering the drain current and threshold voltage.

Both  $N_f$  and  $N_{ot}$  have become more significant in long channel length pMOSFET due to their larger quantity. Chung et al. (2002) showed that  $N_{ot}$  dominates in the channel region and  $N_{it}$  is responsible for damage at the interface near the Gate-to-Source/Drain overlap region. For 10/0.18  $\mu\text{m}$  pMOSFET, the damage by  $N_{it}$  near the Gate-to-Source/Drain overlap region has a greater impact on the transistor parametric

degradation. However, for large channel length device like 10/10  $\mu\text{m}$  pMOSFET,  $N_f$  and  $N_{ot}$  become dominant and even contribute more damage (as the total  $N_f$  and  $N_{ot}$  become larger) to the degradation than the  $N_{it}$  (near the Gate-to-Source/Drain overlap region). Both  $N_f$  and  $N_{ot}$  is dependent on process condition and its quantity will increase with larger geometry on the channel region (Taur and Ning, 1998).

The majority of the published results on channel length dependence in the literature (such as Yamamoto et al., 1999; Sekhar et al., 2004) show that large channel length has smaller degree of degradation. However all these data come from measurement on pMOSFET with  $L < 1 \mu\text{m}$ . Cellere et al. (2004) perform measurements on a full range of  $L$  ( $0.16 \mu\text{m} \rightarrow 10 \mu\text{m}$ ) and show that  $\Delta V_{TH}$  decreases when  $L$  increases from  $0.16 \mu\text{m}$  to  $0.34 \mu\text{m}$ . As  $L$  increases further from  $1 \mu\text{m}$ ,  $\Delta V_{TH}$  increases rapidly and the  $10 \mu\text{m}$  data shows even larger  $\Delta V_{TH}$  (Figure 5.12).

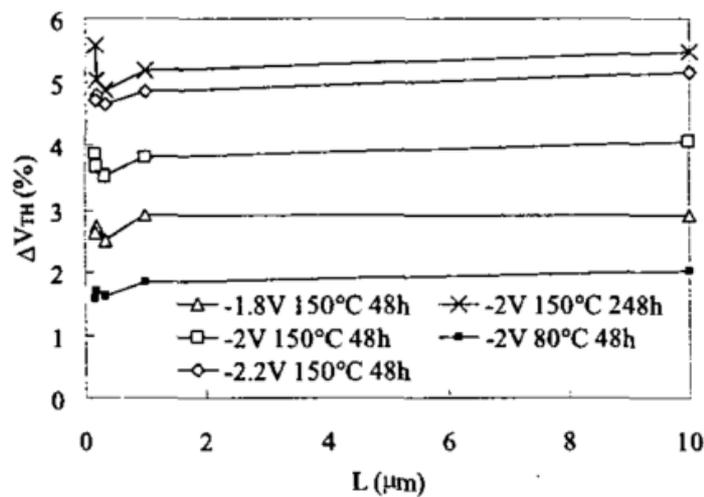


Figure 5.12:  $\Delta V_{TH}$ -vs- $L$  at different  $V_{G, stress}$ , temperatures and stress time (Cellere et al., 2004).

The 10/10  $\mu\text{m}$  TTF data for  $I_{DSAT}$  and  $V_{TCI}$  are extracted the same way as 10/0.18  $\mu\text{m}$  and TTF-vs- $V_{G, stress}$  is plotted on the semi-log plot (Figure 5.13). Both the  $I_{DSAT}$  and  $V_{TCI}$  lifetime for 10/10  $\mu\text{m}$  are lower than 10/0.18  $\mu\text{m}$ , in agreement with Mooraka et al.

(2007). This result confirms the significant role play by  $N_f$  and  $N_{ot}$  in degradation of large channel length pMOSFET.

Similar to 10/0.18  $\mu\text{m}$ , the  $I_{\text{DSAT}}$  lifetime on 10/10  $\mu\text{m}$  is higher than  $V_{\text{TCl}}$  lifetime. However, the  $I_{\text{DSAT}}$  lifetime on 10/10  $\mu\text{m}$  is  $\sim 1$  order of magnitude higher than the  $V_{\text{TCl}}$  lifetime. This is different than 10/0.18  $\mu\text{m}$ , where  $I_{\text{DSAT}}$  lifetime is 2 order of magnitudes higher than the  $V_{\text{TCl}}$  lifetime. This shows that  $N_f$  and  $N_{ot}$  play a stronger role during the saturation mode of operation of pMOSFET as compared to linear mode of operation. Therefore, both type of trap charges are more profound during the saturation mode of operation of the pMOSFET.

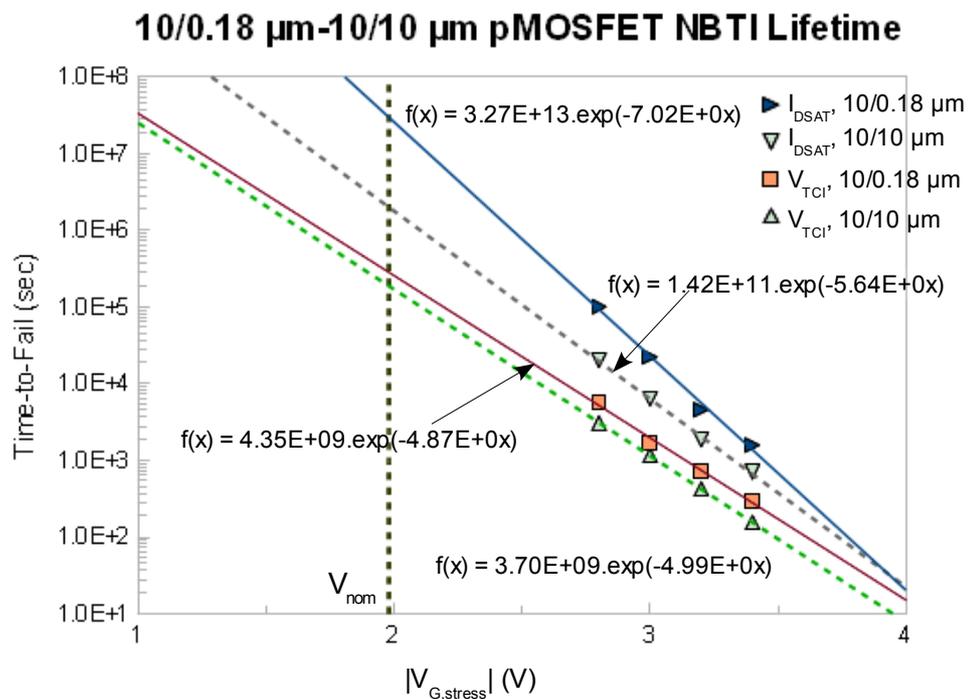


Figure 5.13: TTF-vs- $|V_{\text{G,stress}}|$  for lifetime extraction on  $I_{\text{DSAT}}$  and  $V_{\text{TCl}}$ , comparing 10/0.18  $\mu\text{m}$  and 10/10  $\mu\text{m}$  pMOSFET.

From the above results, care has to be taken by the IC designer for analog circuit to factor in the higher degree of NBTI degradation on long channel length pMOSFET. For the wafer fabrication plant, a thorough characterization on channel length

dependence needs to be performed on a wide range of channel length, in order to determine the worst case channel length for a particular technology node. This will greatly help the IC designer in deciding the design of circuit with long channel length pMOSFET.

## 5.8 The Effect of Gate oxide Thickness on NBTI Degradation

In modern CMOS IC, due to the requirement for the IC to operate in optimum speed for a particular technology node and also to maintain the robustness in interfacing with other IC components, the core transistor is typically designed to the minimum channel length ( $L_{\min}$ ) allowable by the technology node. The Input/Output circuitry which is at higher operating voltage will be designed with thicker gate oxide at channel length  $\sim 2$  times of  $L_{\min}$ . This is to ensure the Input/Output circuitry will be able to operate at the same voltage level as what other IC components are using. As such, the operating voltage of thick oxide pMOSFET will be higher than thin oxide pMOSFET (Annema et al., 2005).

The gate oxide thickness is directly affecting the operation of pMOSFET, where both threshold voltage and drain current are related to  $C_{ox}$  (see Equation (3.1), (3.2), (3.3) and (5.8)). The capacitance of the gate oxide is inversely dependent on the gate oxide thickness as follow:

$$C_{ox} = \frac{\epsilon_{ox}\epsilon_0}{T_{ox}} \quad (5.10)$$

where  $\epsilon_{ox}$  is the silicon oxide permittivity,  $\epsilon_0$  is the vacuum permittivity and  $T_{ox}$  is the gate oxide thickness. By lowering the gate oxide thickness, the drain current will be increased and threshold voltage will be reduced, allowing the transistor to achieve higher current driving capability and switching faster. This will improve the performance of IC operation if the majority of transistors are designed with minimum

channel length at thin gate oxide as allowable per the technology node.

Table 5.1 shows the comparison between thin oxide and thick oxide pMOSFET for this wafer sample. As can be seen from Table 5.1, the operating voltage of thick oxide pMOSFET has been set at -3.3 V, so that the effective electric field is about equivalent to thin oxide pMOSFET. Therefore, to do an accurate comparison on NBTI degradation between the two type of transistors with different oxide thicknesses and different processes, it can only be done on the basis of equal electric field across the gate oxide ( $E_{ox}$ ) (Reisinger et al., 2009). Hence the stress voltage for thick oxide is -5.0 V, -5.32 V, -5.7 V and -6.07 V, where the electric field on three of its  $V_{G, stress}$  are closed to those used in 10/0.18  $\mu\text{m}$  tests.

Table 5.1: Comparison between thin oxide and thick oxide.

Parameters	Thin oxide	Thick oxide
Minimum Channel length, $L_{min}$	0.18 $\mu\text{m}$	0.30 $\mu\text{m}$
Gate oxide Thickness, $T_{ox}$	2.9 nm	5.6 nm
Operating Voltage, $V_{DD}$	-1.8 V	-3.3 V
Electric Fields at $V_{DD}$	6.1 MV/cm	5.89 MV/cm
$V_{G, stress,1}$	-2.8 V (9.49 MV/cm)	-5.0 V (8.93 MV/cm)
$V_{G, stress,2}$	-3.0 V (10.17 MV/cm)	-5.32 V (9.5 MV/cm)
$V_{G, stress,3}$	-3.2 V (10.85 MV/cm)	-5.7 V (10.18 MV/cm)
$V_{G, stress,4}$	-3.4 V (11.53 MV/cm)	-6.07 V (10.84 MV/cm)

The results of NBTI stress for thick oxide pMOSFET is shown in Figure 5.14. The 10/0.30  $\mu\text{m}$  pMOSFET  $V_{TCI}$  degradation follows the Power law, with  $n$  exponent = 0.25 for both  $V_{G, stress}$ . This shows that the same phenomenon observed on 10/0.18  $\mu\text{m}$  is responsible for the NBTI degradation on 10/0.30  $\mu\text{m}$  pMOSFET. However, 10/0.30  $\mu\text{m}$  pMOSFET shows a faster degradation for both  $V_{G, stress}$  as compared to 10/0.18  $\mu\text{m}$  at the same  $E_{ox}$ , in agreement with Huard et al. (2006). This indicates that other factor, such as

difference in process condition (which results in changes of  $N_f$  and  $N_{ot}$  level), might also be contributing to the 10/0.30  $\mu\text{m}$  pMOSFET NBTI degradation. Huard et al. (2006) pointed out that  $N_{it}$  generation is identical for the same oxide electric field when the gate oxide thickness is increased ( $T_{ox} = 2.3 \text{ nm} \rightarrow 10 \text{ nm}$ ) and  $N_{it}$  is not the sole root cause of pMOSFET parameters shift under NBTI stress (Figure 5.15).

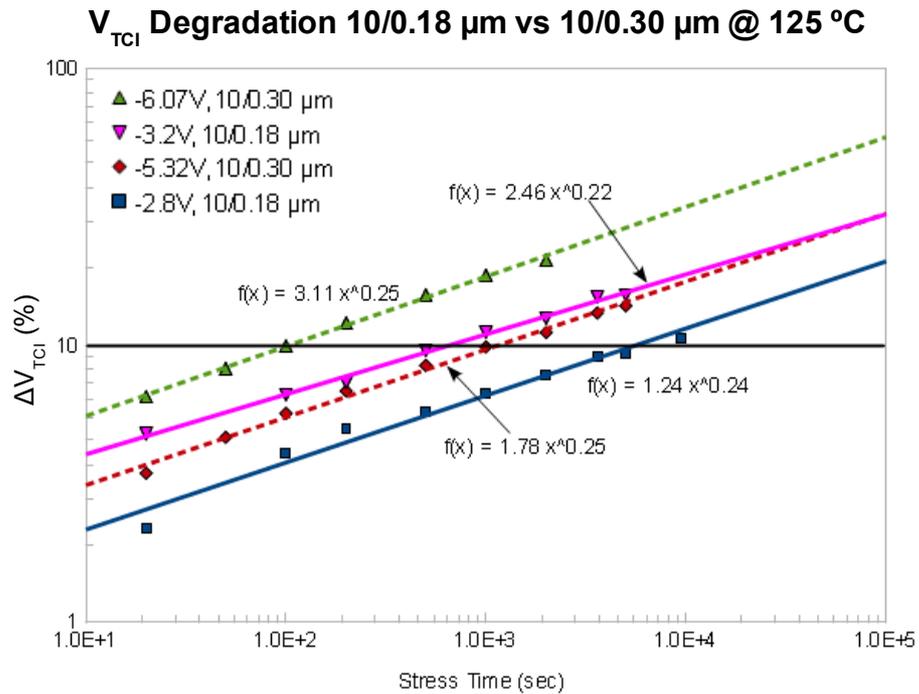


Figure 5.14: Comparison of  $V_{TCI}$  degradation between 10/0.18  $\mu\text{m}$  (thin oxide) and 10/0.30  $\mu\text{m}$  (thick oxide) for  $E_{ox}$  of 9.5 MV/cm and 10.85 MV/cm.

The additional mechanism involved in NBTI degradation can be attributed to the positive charges induced during the stress by holes trapping into the oxide. This is because under NBTI stress condition, the holes will be pulled towards the gate oxide from the substrate, interacting with Si-H to generate the  $N_{it}$  and some may gain enough thermal energy to tunnel into gate oxide under the elevated thermal ambient. These holes may be trapped in the oxide by existing traps (created during fabrication process) or traps created due to the tunneling mechanism. By applying a varying frequency on

the gate voltage, Huard et al. (2006) demonstrate that majority of the trapped holes are due to existing traps prior to stress and traps generation during stress is relatively small for pure SiO<sub>2</sub>.

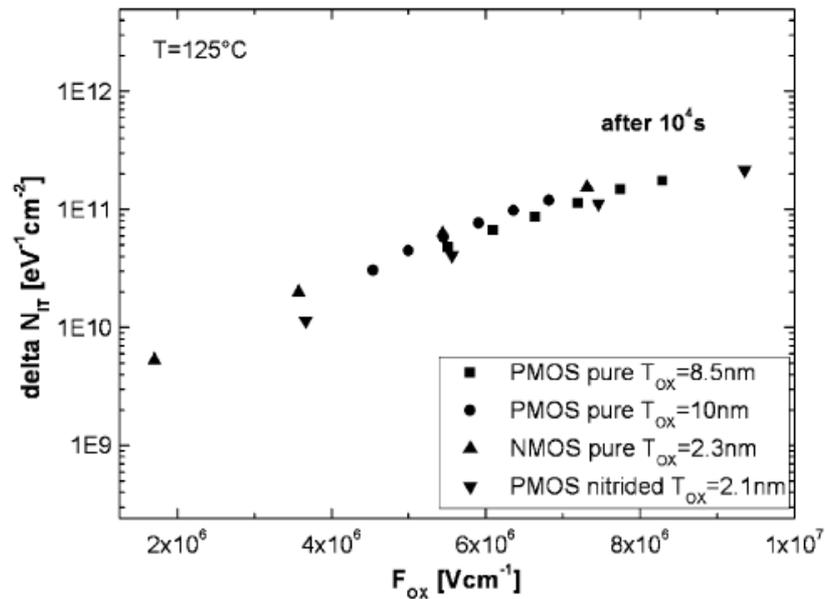


Figure 5.15: Interface trap generation for pMOSFET with pure and nitrided oxides over a wide range of gate oxide thickness (Huard et al., 2006).

For this wafer sample with dual gate, the gate oxide for thin oxide and thick oxide is grown separately. The thick oxide area is grown with Wet Oxidation while the thin oxide area is grown with Dry Oxidation. This strategy is being employed because the process time for Wet Oxidation is about 10 times shorter than Dry Oxidation for the same thickness. By using Wet Oxidation on thick oxide, the cost of manufacturing is reduced, as well as the cycle time to process the wafer. However, the Wet Oxidation process that used to grow the thick oxide is known to enhance NBTI degradation. Kimizuka et al. (2000) show results that Wet Oxidation has higher  $\Delta V_{TH}$  as compared to Dry Oxidation. Furthermore, Sasada et al. (1998) perform experiments using silicon nitride layer to block migration of moisture from subsequent process (backend

metallization processes) into the gate oxide. pMOSFET without the silicon nitride blocking layer shows a much larger  $\Delta V_{TH}$  than those with silicon nitride blocking layer. These results point to the water species have a tendency of creating traps within the gate oxide and enhance NBTI degradation. These reported findings from the literature correlate well to the results shown in Figure 5.14. Also, with faster reaction rate during Wet oxidation, there will be more imperfections or crystal lattice defects in the gate oxide that will interact with the water species to create oxide traps. Therefore, the enhanced degradation seen on 10/0.30  $\mu\text{m}$  pMOSFET is primarily due to the pre-existing traps created during Wet oxidation and also thicker gate oxide that results in higher  $N_{ot}:N_{it}$  ratio as compared to 10/0.18  $\mu\text{m}$ .

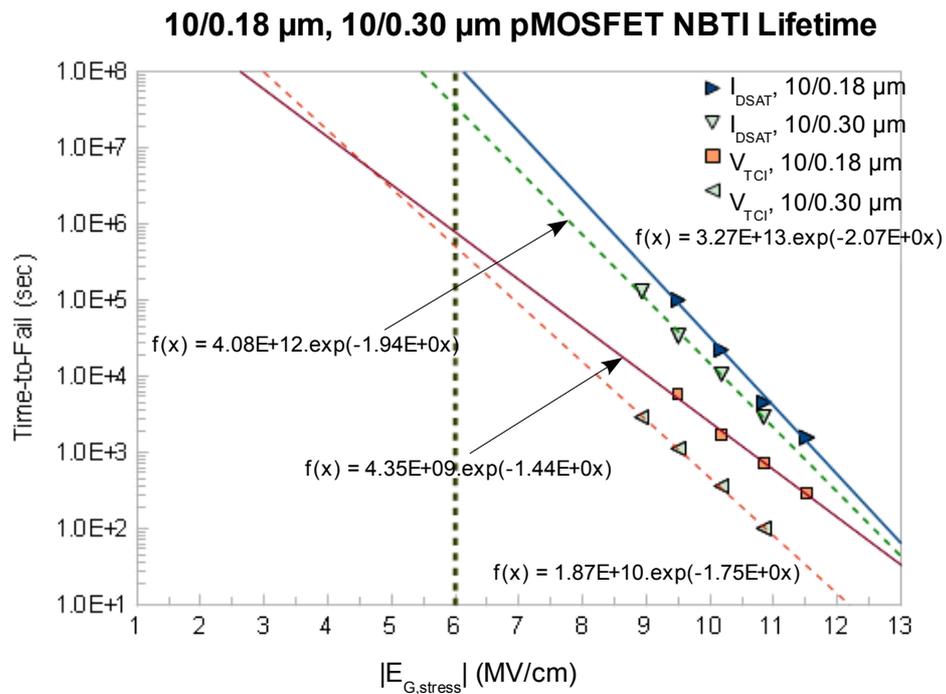


Figure 5.16: TTF-vs- $E_{G,stress}$  for lifetime extraction on  $I_{DSAT}$  and  $V_{TCI}$ , comparing 10/0.18  $\mu\text{m}$  and 10/0.30  $\mu\text{m}$  pMOSFET.

For lifetime extraction, the TTF of the 10/0.30  $\mu\text{m}$  thick oxide pMOSFET is plotted against the electric field at stress voltage ( $E_{G,stress}$ ) in Figure 5.16. This is to facilitate the comparison with 10/0.18  $\mu\text{m}$ , where its TTF is also plotted against the

$E_{G, \text{stress}}$ . Similar to 10/0.18  $\mu\text{m}$ , the lifetime for  $I_{\text{DSAT}}$  is higher than  $V_{\text{TCl}}$  for 10/0.30  $\mu\text{m}$ . This means that regardless the thickness of gate oxide, the linear operation of the transistor is always more sensitive than the saturation operation in NBTI degradation. Due to the faster degradation seen on 10/0.30  $\mu\text{m}$ , both the  $I_{\text{DSAT}}$  and  $V_{\text{TCl}}$  lifetime is lower than those of 10/0.18  $\mu\text{m}$ . With the lower extracted lifetime on 10/0.30  $\mu\text{m}$ , this further confirms that  $N_{\text{ot}}$  plays a more significant role than  $N_{\text{it}}$  in thick oxide NBTI degradation as compared to thin oxide, primarily due to the higher number of pre-existing oxide traps. Hence the wafer fabrication plant has to control the Wet oxidation process to minimize trap generation and the IC designer needs to take extra care in designing the Input/Output circuitry to avoid excessive/prolong high voltage level on the gate of thick oxide pMOSFET.

### **5.9 Optimized $I_{\text{D}}-V_{\text{G}}$ ( $V_{\text{THopt}}$ ) to Reduce “Dead” Time**

As discussed in Section 3.7.1, the Conventional DC Stress-Measure-Stress method exhibits a rather long “dead” time, due to the complete I-V characteristics ( $I_{\text{D}}-V_{\text{D}}$ ,  $I_{\text{D}}-V_{\text{G}}$ ) being measured with voltage sweeping from  $0 \rightarrow -1.8$  V. The time taken to perform  $I_{\text{D}}-V_{\text{D}}$  and  $I_{\text{D}}-V_{\text{G}}$  sweep measurements can be up to  $\sim 9$  sec on the Keithley 4200 SCS. Such a long “dead” time in between the stress cycle will result in recovery of the degradation and may cause an over estimation or under estimation of lifetime extraction. Ershov et al. (2003) show a 3 times increase in lifetime extraction with long “dead” time measurement. A long “dead” time allows more recovery to take place and this recovery is caused by the repassivation of some of the Si dangling bond ( $N_{\text{it}}$ ) by hydrogen. The hydrogen diffuses back into the Si dangling bonds as per the Reaction-Diffusion model to form Si-H bonds. Also the detrapping of holes from oxide traps ( $N_{\text{ot}}$ ) when the stress is removed during interim monitoring measurement in between stress cycles (Tsujikawa et al., 2003) may contribute to the recovery as well.

The results in Section 5.4 show that threshold voltage (linear mode of operation) is the critical parameter to be monitored for NBTI degradation. By omitting the  $I_D$ - $V_D$  sweep during the monitoring measurement in between stress cycles, this “dead” time can be reduced by half. To reduce the “dead” time further, a short range of  $I_D$ - $V_G$  sweep that centers on the extracted  $V_{TCI}$  before stress (instead of sweeping from 0 V up to nominal operating voltage) is employed. By omitting the full  $I_D$ - $V_D$  sweep and using this optimized  $I_D$ - $V_G$  sweep, the “dead” time during the monitoring measurement phase is reduced to approximately 1 sec.

The following sections will show the Optimized  $I_D$ - $V_G$  sweep experimental results with shorter “dead” time. Optimized  $I_D$ - $V_G$  sweep is being used due to its simple setup and easy implementation and also the fact that  $V_{TCI}$  is the most sensitive parameter, due to the lowest lifetime obtained in Section 5.4. For clarity sake, the  $V_{TCI}$  degradation obtained from Optimized  $I_D$ - $V_G$  sweep will be named as  $V_{THopt}$  and will compare to  $V_{TCI}$  degradation from the Conventional DC Stress-Measure-Stress method. The extracted lifetime between the two will also be compared.

### **5.9.1 Optimized $I_D$ - $V_G$ Sweep Measurement Results and Discussion**

The  $V_{THopt}$  degradation data obtained from Optimized  $I_D$ - $V_G$  sweep method is plotted on the log-log scale (Figure 5.17). The  $V_{THopt}$  degradation follows the Power law, same as those seen in Conventional DC Stress-Measure-Stress method. However, the time exponent,  $n$  is lower, falls into the range of 0.186 to 0.192. This results are in agreement with Varghese et al. (2005), where shorter “dead” time will reduce the  $n$  exponent. To show the comparison with Conventional DC Stress-Measure-Stress, the  $V_{TCI}$  degradation at -3.2 V and -2.8 V for both method is shown in Figure 5.18.

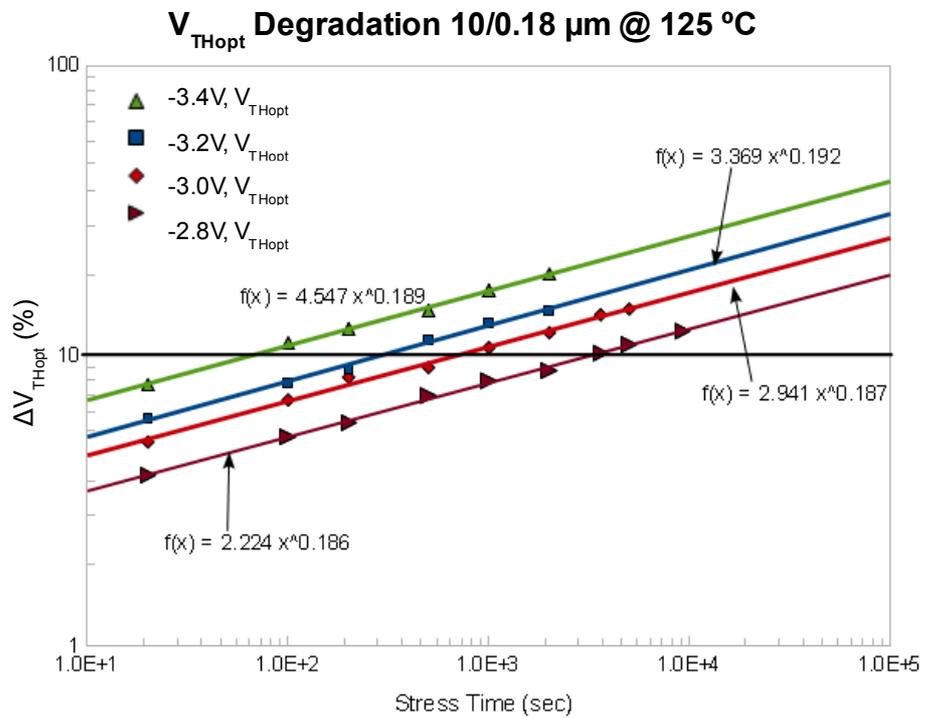


Figure 5.17: Shift in V<sub>THopt</sub> vs Stress Time on a log-log plot. The time exponent, *n* is ranged from 0.186 to 0.192.

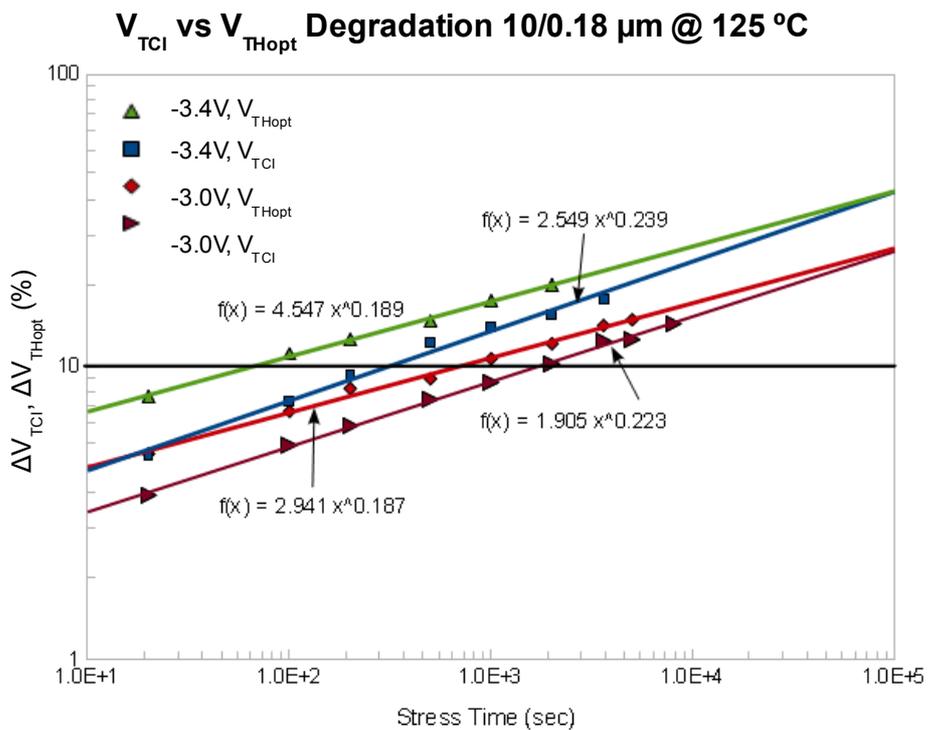


Figure 5.18: V<sub>TCI</sub> vs V<sub>THopt</sub> degradation at -3.4 V and -3.0 V.

The  $V_{\text{THopt}}$  shows a faster degradation as compared to the Conventional  $V_{\text{TCI}}$  degradation, especially during the initial phase of stress. This difference can be attributed to the recovery that is taking place when the stress was removed to perform the  $I_{\text{D}}-V_{\text{G}}$  measurement. In Optimized  $I_{\text{D}}-V_{\text{G}}$  sweep where the recovery is reduced, higher shift in  $V_{\text{THopt}}$  can be seen as most of the traps ( $N_{\text{it}}$  and  $N_{\text{ot}}$ ) created are still present. This is especially true during the initial phase of the stress, where shift in  $V_{\text{THopt}}$  is mainly responsible by the reaction process (breaking of Si-H bond to create interface traps), as per the Reaction-Diffusion model. When the stress progresses longer, the rate of shift in  $V_{\text{THopt}}$  is reduced as the reaction process becomes saturated and the shift of  $V_{\text{THopt}}$  becomes a diffusion limited process of hydrogen species.

In the case of Conventional  $V_{\text{TCI}}$ , the long “dead” time allows more of the generated  $N_{\text{it}}$  and  $N_{\text{ot}}$  to be annealed, resulting in smaller shift in  $V_{\text{TCI}}$  especially during the initial phase of stress (Parthasarathy et al., 2005). However, in the later phase of the stress, more and more  $N_{\text{it}}$  and  $N_{\text{ot}}$  are generated and the repassivation and detrapping process become limited by the “dead” time, which is constant for all monitoring measurement phase. Hence the shift in  $V_{\text{TCI}}$  is slowly catching up with those in  $V_{\text{THopt}}$ . As the reaction process becomes saturated like the  $V_{\text{THopt}}$  (at a later time), the shift of  $V_{\text{TCI}}$  is expected to be comparable with  $V_{\text{THopt}}$ .

### **5.9.2 Statistical Analysis on Comparing $\Delta V_{\text{THopt}}$ and $\Delta V_{\text{TCI}}$**

To answer the question whether  $\Delta V_{\text{THopt}}$  and  $\Delta V_{\text{TCI}}$  will converge if longer stress time is applied, statistical analysis employing hypothesis testing ( $t$ -test) on the mean of  $\Delta V_{\text{THopt}}$  and  $\Delta V_{\text{TCI}}$  at specific time interval (e.g. 100 sec, 1000 sec, 2000 sec or 5000 sec) will be performed. For this experimental study, the actual variance of the  $\Delta V_{\text{THopt}}$  and  $\Delta V_{\text{TCI}}$  population is unknown. Therefore, the hypothesis testing will need to be proceeded in

two stages: Stage 1 is the  $F$ -test on sample variance and Stage 2 is the  $t$ -test on sample mean (Montgomery and Runger, 2010). The  $F$ -test is to allow the proper equations to be used in  $t$ -test according to the assumption chosen, based on the result from  $F$ -test.

Stage 1: Assuming the population variance for  $\Delta V_{\text{THopt}}$  is  $\sigma_1^2$  and  $\Delta V_{\text{TCI}}$  is  $\sigma_2^2$ , the hypothesis will be as below:

$$\text{Null hypothesis, } H_o : \sigma_1^2 = \sigma_2^2 \quad (5.11)$$

$$\text{Alternative hypothesis, } H_A : \sigma_1^2 \neq \sigma_2^2 \quad (5.12)$$

The test statistic,  $F_o$  is the ratio of sample variance,  $S_1^2$  and  $S_2^2$  (Montgomery and Runger, 2010):

$$F_o = \frac{S_1^2}{S_2^2} \quad (5.13)$$

With 3 runs for each stress voltage for  $V_{\text{THopt}}$  and  $V_{\text{TCI}}$ , and the significance level ( $\alpha$ ) of 0.05 in making an error, the  $F_{\text{crit}}$  is 0.026 for lower tail and 39 for upper tail of the  $F$  distribution. If the  $F_o$  value is in between these two  $F_{\text{crit}}$  values ( $0.026 < F_o < 39$ , Figure 5.19), the null hypothesis will not be rejected, at the 0.05 level of significance. Table 5.2 showed the  $F_o$  results for the selected time intervals, based on the  $\Delta V_{\text{THopt}}$  and  $\Delta V_{\text{TCI}}$  data in Figure 5.18.

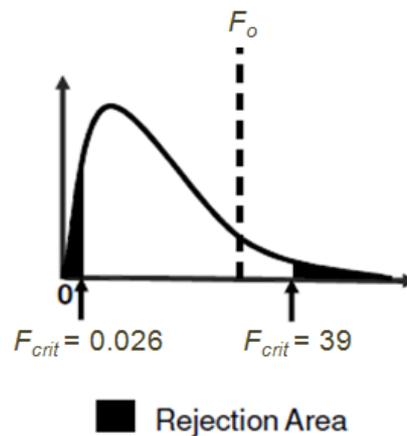


Figure 5.19:  $F$  distribution showing example of rejection area for hypothesis testing.

Table 5.2. Test statistic,  $F_o$  for the specified time interval at -3.4 V and -3.0 V.

$V_{G, stress}$	Time	$V_{TH}$	Variance	$F_o$
-3.4V	100 sec	$\Delta V_{THopt}$	0.071	0.32
		$\Delta V_{TCI}$	0.224	
	1000 sec	$\Delta V_{THopt}$	0.614	1.43
		$\Delta V_{TCI}$	0.430	
	2000 sec	$\Delta V_{THopt}$	1.003	0.99
		$\Delta V_{TCI}$	1.011	
-3.0V	100 sec	$\Delta V_{THopt}$	0.421	1.25
		$\Delta V_{TCI}$	0.337	
	1000 sec	$\Delta V_{THopt}$	0.684	1.34
		$\Delta V_{TCI}$	0.509	
	5000 sec	$\Delta V_{THopt}$	1.156	2.12
		$\Delta V_{TCI}$	0.544	

The  $F_o$  value in Table 5.2 is out of the rejection area of the  $F$ -distribution. This means that the null hypothesis cannot be rejected at the 0.05 level of significance. In other words, there is no strong evidence to indicate that the variance between  $\Delta V_{THopt}$  and  $\Delta V_{TCI}$  are significantly different.

Stage 2: The hypothesis testing in Stage 1 indicates the variance of  $\Delta V_{THopt}$  and  $\Delta V_{TCI}$  can be assumed to be approximately equal. Therefore, the test statistic ( $t_o$ ) to compare the mean of two sample with the assumption that the variance are unknown but assumed to be approximately equal is (Montgomery and Runger, 2010):

$$t_o = \frac{\bar{A}_1 - \bar{A}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (5.14)$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad (5.15)$$

where  $\bar{A}_1$  and  $\bar{A}_2$  are mean of the sample,  $n_1$  and  $n_2$  are sample size,  $S_p^2$  is pooled variance.

With this assumption, the hypothesis to assess the mean of  $\Delta V_{\text{THopt}}$  is significantly larger than the mean of  $\Delta V_{\text{TCl}}$  is:

$$\text{Null hypothesis, } H_o : \mu_1 \leq \mu_2 \quad (5.16)$$

$$\text{Alternative hypothesis, } H_A : \mu_1 > \mu_2 \quad (5.17)$$

With 3 runs for each stress voltage for  $V_{\text{THopt}}$  and  $V_{\text{TCl}}$ , and the significance level ( $\alpha$ ) of 0.05 in making an error, the  $t_{\text{crit}}$  is 2.13 for right tail of the  $t$  distribution. If the  $t_o$  value is larger than this  $t_{\text{crit}}$  value (Figure 5.20), the null hypothesis will be rejected, at the 0.05 level of significance. Table 5.3 shows the  $t_o$  results for the selected time intervals, based on the  $\Delta V_{\text{THopt}}$  and  $\Delta V_{\text{TCl}}$  data in Figure 5.18.

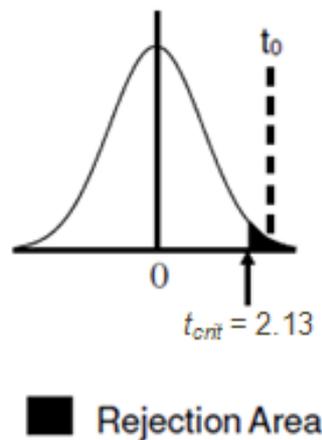


Figure 5.20:  $t$  distribution showing example of rejection area for hypothesis testing.

The  $t_o$  value in Table 5.3 is within the rejection area of the  $t$ -distribution. This means that the null hypothesis is rejected at the 0.05 level of significance. In other words, there is strong evidence to indicate that the mean of  $\Delta V_{\text{THopt}}$  is significantly larger than mean of  $\Delta V_{\text{TCl}}$ .

Table 5.3. Test statistic,  $t_o$  for the specified time interval at -3.4 V and -3.0 V.

$V_{G, stress}$	Time	$V_{TH}$	Mean	Variance	$S_p$	df	$t_o$
-3.4V	100 sec	$\Delta V_{THopt}$	10.85	0.071	0.147	4	10.18
		$\Delta V_{TCI}$	7.66	0.224			
	1000 sec	$\Delta V_{THopt}$	16.78	0.614	0.522	4	5.90
		$\Delta V_{TCI}$	13.3	0.430			
	2000 sec	$\Delta V_{THopt}$	19.13	1.003	1.007	4	4.21
		$\Delta V_{TCI}$	15.68	1.011			
-3.0V	100 sec	$\Delta V_{THopt}$	6.96	0.421	0.379	4	3.26
		$\Delta V_{TCI}$	5.32	0.337			
	1000 sec	$\Delta V_{THopt}$	10.7	0.684	0.597	4	2.86
		$\Delta V_{TCI}$	8.89	0.509			
	5000 sec	$\Delta V_{THopt}$	14.46	1.156	0.850	4	2.31
		$\Delta V_{TCI}$	12.73	0.544			

The  $t_o$  value in Table 5.2 is an indication of the magnitude of difference between  $\Delta V_{THopt}$  and  $\Delta V_{TCI}$ . If the  $t_o$  value is further away (to the right side of the  $t$  distribution) from the  $t_{crit}$ , then the  $\Delta V_{THopt}$  is much larger than  $\Delta V_{TCI}$ . For both  $V_{G, stress}$  of -3.4 V and -3.0 V, the  $t_o$  value is inversely proportional to stress time, i.e.  $t_o$  is smaller with increasing time. Therefore, statistically, it can be concluded that the  $\Delta V_{THopt}$  and  $\Delta V_{TCI}$  will converge if the stress time is large (e.g. 100,000 sec). The extrapolated degradation line (fitted with Power law) in Figure 5.18 also projects that the convergence of both type of degradation. This observation is in agreement with Reisinger et al. (2009) where  $V_{TH}$  degradation on fast measurement (with 1  $\mu$ sec “dead” time) and slow measurement (with 1 sec “dead” time) will coincide at 1,000,000 sec.

This result shows that by using measurement method that shorten the “dead” time, experimental stress time can be reduced as either fast or slow measurement will yield approximately the same shift in  $V_{TCI}$  and  $V_{THopt}$  when stress time is long. This is the reason why the research community is always striving to bring down the “dead” time in order to quantify the true degradation in  $V_{TH}$  (minimizing the recovery effect), so that an accurate model can be developed. This in turn will allow the accelerated

stress time to reduce, thereby saving time and money in running reliability stress test. As such, due to the recovery effect on NBTI degradation, many different measurement methods have been developed (as described in Section 3.7) by the research community.

### 5.9.3 Lifetime Extraction on $V_{THopt}$

The difference in  $V_{TCI}$  and  $V_{THopt}$  degradation between the two methods will affect the lifetime extraction. The lifetime extraction will be based on the TTF criteria (as described in Section 5.4), and typically the stress test will not stop until the device reaches the failing criteria, which usually takes very long time especially at low stress voltage. With the faster degradation on  $V_{THopt}$ , its TTF will be smaller comparatively. This result is shown in Figure 5.21.

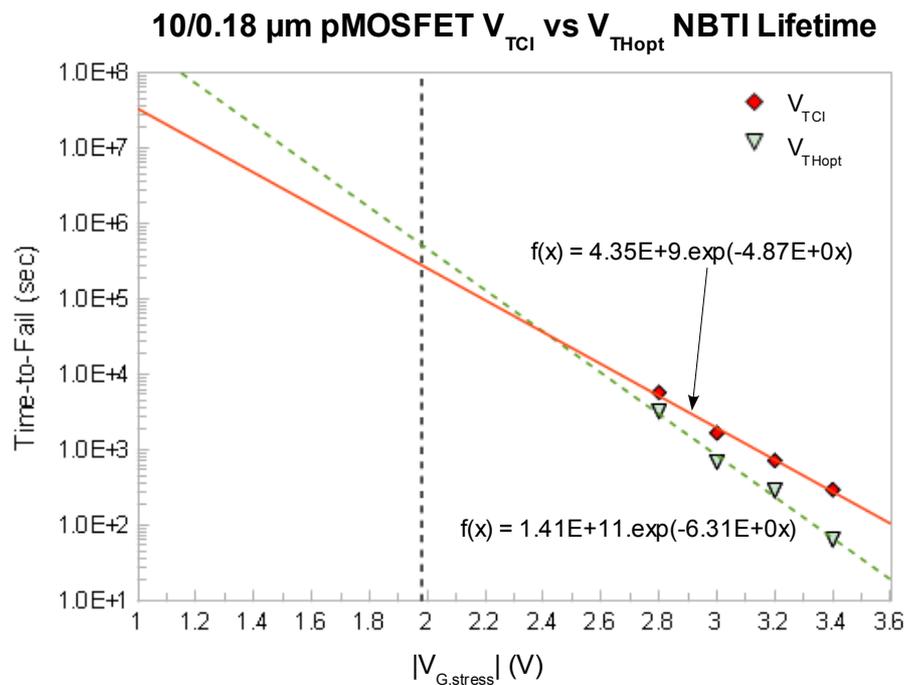


Figure 5.21: TTF-vs- $V_{G, stress}$  for lifetime extraction on  $V_{TCI}$  and  $V_{THopt}$ .

From the extrapolation to the operating voltage, the lifetime for  $V_{THopt}$  is larger than  $V_{TCI}$ , even though the TTF for all  $V_{G, stress}$  is smaller. This is different than the results

in thick gate oxide (Section 5.7), where its TTF is smaller, resulting in smaller extracted lifetime. Comparing between Figure 5.14 and Figure 5.17, the big difference lies in the  $n$  exponent. For thick oxide pMOSFET (Figure 5.14), the  $n$  exponent is  $\sim 0.25$  (similar to  $V_{TCI}$  for thin oxide pMOSFET). Where as for  $V_{THopt}$ , the  $n$  exponent is  $\sim 0.19$  (Figure 5.17). The  $n$  exponent value has a strong effect in lifetime extrapolation and can be an indication on lifetime extraction as well.

The  $V_{THopt}$  lifetime is preferable than the  $V_{TCI}$  lifetime due to the fact that this measurement method reflects closer to the actual degradation seen by the pMOSFET, with less recovery. This results show that measurement technique with long “dead” time will result in a significant under estimation of lifetime. This is another reason why the research community is striving to develop measurement technique with the shortest “dead” time. The  $V_{THopt}$  lifetime obtained from this experiment is in agreement with Reisinger et al. (2009), where the lifetime extracted from fast measurement (with 1  $\mu$ sec “dead” time) is at least 1 order of magnitude higher than slow measurement (with 1 sec “dead” time). In summary, measurement with long “dead” time is not desirable in NBTI stress test, due to the under estimation of device operating lifetime and allowing more recovery to take place (Ershov et al., 2003).

## **5.10 Implication of this Research Work to the Industry**

To date, this research work has performed an experimental characterization on the 0.18  $\mu$ m dual gate CMOS process wafer through a systematic methodology and devises a method to perform NBTI with shorter “dead” time (Optimized  $I_D$ - $V_G$  sweep). Such extensive study can be adopted by the wafer fabrication plant during the process development and qualification of new wafer fabrication process as well as monitoring the NBTI performance during mass production.

An extensive characterization as the one performed in this research work during the development of new wafer fabrication process will allow the wafer fabrication plant to determine which transistor parameters are more sensitive to NBTI stress. This is important as it provides valuable information to the wafer fabrication process designer on the transistor current driving capability, switching speed and tolerance to leakage under NBTI condition. With that knowledge, the wafer fabrication process designer can decide which area in the wafer fabrication process that needs further improvement.

From the discussion in Section 5.9.1, NBTI stress test that utilizes Optimized  $I_D$ - $V_G$  sweep can be applied to the wafer process qualification, which is the last phase of the process development. A smaller set of experiments that consist of the minimum channel length transistor for each of the operating voltages can be subjected to NBTI stress at one elevated temperature, on wafer with the new improvement. This will allow the wafer fabrication plant to validate the improved wafer process at a much shorter time than tests done during process development. Subsequent to this, when the wafer fabrication process enters into mass production, monitoring of NBTI performance can also be carried out using this Optimized  $I_D$ - $V_G$  sweep method, where the following considerations will be discussed.

First and foremost, the saturation drain current ( $I_{DSAT}$ ) degradation had been shown to be much less severe as compared to threshold voltage. This means that the wafer fabrication plant do not need to perform the  $I_{DSAT}$  measurement when monitoring the NBTI performance during mass production, as the results in earlier sections showed that  $V_{TCI}$  is the most sensitive parameter towards NBTI degradation. Secondly, Optimized  $I_D$ - $V_G$  sweep measurement method can be used during interim characterization of the monitoring measurements, as threshold voltage is the only

parameter required to be measured. This is coupled with the fact that Optimized  $I_D$ - $V_G$  sweep measurement can be easily set up in any tester with parametric measurement capability. Furthermore,  $V_{THopt}$  extracted from Optimized  $I_D$ - $V_G$  sweep method produces more accurate lifetime extraction. With its capability to capture more actual degradation during the initial phase of the stress, Optimized  $I_D$ - $V_G$  sweep allows the choice of setting lower  $V_{G, stress}$ , lower temperature or even cutting down the total stress time (with its characteristic of reaching TTF criteria faster). Lowering the  $V_{G, stress}$  to the level that is closed to the operating voltage has the advantage of getting closer to the actual lifetime (Reisinger et al., 2009).

One aspect that the wafer fabrication plant needs to consider is the large sample size for monitoring the NBTI degradation performance during mass production. To do that, one needs to run the NBTI test on a group of pMOSFET (e.g. 10 transistors) on one wafer from a lot at a specific time frequency (e.g. 1 lot/week). This will result in long test time if Conventional DC Stress-Measure-Stress method is used. By employing Optimized  $I_D$ - $V_G$  sweep method, the NBTI test can be done at lower temperature and the test time can be reduced by selecting higher  $V_{G, stress}$ . This short stress test can be run on production wafer (where pMOSFET test structure is available on the scribe lane) using production tester (which typically capable of testing at temperature as high as 70 °C – 80 °C). Therefore, by incorporating Optimized  $I_D$ - $V_G$  sweep into production test program, the engineer can configure the test flow in such a way (with the option to turn on or off the NBTI test) that a number of lot will be tested in one week or one month (e.g. 1 wafer/lot on a weekly basis). In this manner, the wafer fabrication plant will be able to monitor the NBTI degradation performance on a periodical basis, with just investing a small amount of production tester time. Furthermore, by having NBTI test (with Optimized  $I_D$ - $V_G$  sweep method) as part of the production test program, the

engineer can easily use it as part of the analysis for any abnormal lot that fails the production test (if it has not been tested for NBTI).

In a nutshell, the wafer fabrication plant can implement the NBTI test (with Optimized  $I_D$ - $V_G$  sweep) into the production tester for periodical monitoring and also on the bench parametric tester for engineering study and qualification, without investing any additional hardware and software. The results from this research have demonstrated the feasibility of performing Optimized  $I_D$ - $V_G$  sweep measurement on 0.18  $\mu\text{m}$  pMOSFET and it is believed that this method can be applied to newer process technology node, like 0.13  $\mu\text{m}$ , 0.11  $\mu\text{m}$  or smaller.

## Chapter 6. Conclusion and Future Work

This NBTI characterization study achieved its initial goal of setting up NBTI testing capability, with measured data correlated with those from the industry. The same equipment setup can be used to perform other wafer level reliability stress tests, such as Time-Dependent Dielectric Breakdown (TDDB), Hot Carrier Injection (HCI), Electromigration (EM) and etc.

The NBTI characterization study had built up a baseline degradation data on 0.18  $\mu\text{m}$  thin oxide pMOSFET using Conventional DC Stress-Measure-Stress method. The results showed that increase in interface states generation and holes trapping by pre-existing oxide traps were mainly responsible for the degradation seen on key transistor parameters, such as drive current in saturation mode ( $I_{\text{DSAT}}$ ) and threshold voltage in linear mode ( $V_{\text{TEXT}}$ ,  $V_{\text{TCI}}$ ). The degradation in these parameters showed the time exponent,  $n$  value was very close to 0.25, a typical value seen on the shift in interface states and this value had been derived analytically using the Reaction-Diffusion model. Thus, the degradation behavior on this 0.18  $\mu\text{m}$  dual gate CMOS wafer was validated to be able to explain by the widely used Reaction-Diffusion model, which cannot be achieved through  $V_{\text{TH}}$  Stability test.

Further analysis on the data showed that the degradation seen on  $I_{\text{DSAT}}$  in saturation mode was less severe than degradation seen on threshold voltage in linear mode, as  $I_{\text{DSAT}}$  lifetime was 2 order of magnitudes larger than threshold voltage lifetime ( $V_{\text{TEXT}}$ ,  $V_{\text{TCI}}$ ). Therefore, the wafer fabrication plant needs to focus on improving the robustness of pMOSFET linear mode of operation by finding ways to reduce interface states and pre-existing oxide traps generation. Also, among the threshold voltage measurement method,  $V_{\text{TCI}}$  was shown to be most sensitive for NBTI degradation.

The experiments performed with different temperature had showed the strong dependence of NBTI degradation on temperature and similar to other reliability failure mechanisms, it followed the Arrhenius law. Contrary to most reported work in the literature, experimental results on long channel length showed a worst than expected NBTI degradation. This can be attributed to the position of interface trap which is concentrated near the drain side and the higher contribution of fixed oxide charge and oxide trapped charge for long channel length pMOSFET. In addition, experiments on thicker gate oxide pMOSFET showed a much worst NBTI degradation and can be concluded as the difference in gate oxidation process, where higher number of oxide trapped charge is found on thick gate oxide (Wet oxidation) as compared to thin gate oxide (Dry oxidation).

Lastly, the evaluation to minimize the measurement “dead” time using the advance methods was found to be not feasible to implement on current equipment set. In an attempt to assess the impact of measurement “dead” time on NBTI degradation, an Optimized  $I_D$ - $V_G$  sweep ( $V_{THopt}$ ) measurement method was proposed here, without additional hardware or software introduced into the Keithley 4200 SCS. The experimental results showed that  $V_{THopt}$  produced faster degradation as less recovery was taking place. The extracted lifetime was surprisingly higher than the  $V_{TCI}$  lifetime obtained from Conventional DC Stress-Measure-Stress method, due to the lower time exponent value. The characteristics of this Optimized  $I_D$ - $V_G$  sweep method allow it to be easily incorporated into wafer fabrication plants' production tester and also bench tester. This enables periodical monitoring of NBTI degradation performance and also performs engineering and qualification study.

With the results presented above, this research work had achieved all the objectives outlined in Chapter 1, in setting up the NBTI testing capability, gaining an in depth understanding of NBTI degradation, validated the 0.18  $\mu\text{m}$  dual gate CMOS wafer sample to be adhered to Reaction-Diffusion model, assessed the effect of temperature, channel length and gate oxide thickness and also proposed the Optimized  $I_D$ - $V_G$  sweep measurement method to be used by wafer fabrication plants, without additional cost incurred. The results of this research had been published in ASQED 2011 and a journal paper had been submitted to International Journal of Electronics.

## 6.1 Recommendations on Future Work

Below are some of the recommendations on future experimental study that can be carried out, to further expand the scope of this research work which may include other methods and availability of wafer sample at more advanced technology node.

1. This characterization study can be extended to wafers fabricated at more advanced technology node, for example 0.13  $\mu\text{m}$ , 0.11  $\mu\text{m}$  and beyond. The industry has made a major change on these technology nodes, namely changing the gate oxide from pure  $\text{SiO}_2$  to silicon oxynitride ( $\text{SiO}_x\text{N}_y$ ). This change is to mitigate the short channel effect when the channel length is scaled down further for higher switching speed. With shorter channel length, the gate oxide thickness has to be thinner so that the threshold voltage can be lower to improve the switching speed. With pure  $\text{SiO}_2$ , thinner oxide (typically  $< 2$  nm) will result in higher gate leakage as direct tunneling will take effect. By introducing nitrogen into the gate oxide, the effective dielectric constant of  $\text{SiO}_x\text{N}_y$  will increase, allowing the physical  $\text{SiO}_x\text{N}_y$  thickness to increase so that the electrical  $\text{SiO}_x\text{N}_y$  thickness is thinner. In this way, the threshold voltage will be lowered at these

advance nodes and the physically thicker  $\text{SiO}_x\text{N}_y$  will minimize the gate leakage. Because of this gate oxide material change, the NBTI degradation mechanism may be different than pure  $\text{SiO}_2$ . An extensive characterization such as the one carried out in this research will enable a full understanding on this gate material change.

2. As pointed out in the results and discussion of channel length dependence and gate oxide thickness dependence in previous sections, the NBTI degradation is contributed by interface states generation and holes trapped in pre-existing oxide traps. To quantify the contribution by interface states, charge pumping measurement method can be used. Section 3.4 described the charge pumping measurement method and this will require a pulse generator unit to be connected to the Keithley 4200 SCS. By quantifying the  $N_{it}$  during the stress, the contribution of  $N_{ot}$  will be known and its behavior can be studied.
3. The Keithley 4200 SCS has the software option to perform On-The-Fly (OTF) measurement. Unfortunately, this is not available in the current system and will require additional cost to acquire this software option. Alternatively, one can customize the test flow by learning the programming aspect of Keithley 4200 SCS. By customizing the test flow that follows the way OTF is performed, the comparison can be made between Conventional DC SMS, Optimized  $I_D$ - $V_G$  sweep and OTF. Similarly, the Fast Switching measurement method as described in Section 3.7.3 may also be customized into the Keithley 4200 SCS.

## Bibliography

- Achuthan, M. K., & Bhat, K. N. (2007). *Fundamentals of semiconductor devices*. New Delhi: Tata McGraw-Hill.
- Acovic, A., LaRosa, G., & Sun, Y. C. (1996). A review of hot-carrier degradation mechanisms in MOSFETs. *Microelectronics Reliability*, 36(7-8), 845-869.
- Alam, M. A. (2005). Negative Bias Temperature Instability Basics/Modeling. *IEEE International Reliability Physics Symposium Tutorials*.
- Alam, M. A., Bude, J., & Ghetti, A. (2000). Field acceleration for oxide breakdown-can an accurate anode hole injection model resolve the E vs. 1/E controversy? *IEEE International Reliability Physics Symposium Proceedings*, 21-26.
- Alam, M. A., Kufluoglu, H., Varghese, D., & Mahapatra, S. (2007). A comprehensive model for PMOS NBTI degradation: Recent progress. *Microelectronics Reliability*, 47(6), 853-862.
- Alam, M. A., & Mahapatra, S. (2005). A comprehensive model of PMOS NBTI degradation. *Microelectronics Reliability*, 45(1), 71-81.
- Ang, D. S., Wang, S., & Ling, C. H. (2005). Evidence of two distinct degradation mechanisms: From temperature dependence of negative bias stressing of the ultrathin gate p-MOSFET. *IEEE Electron Device Letters*, 26(12), 906-908.
- Annema, A. J., Nauta, B., VanLangevelde, R., & Tuinhout, H. (2005). Analog circuits in ultra-deep-submicron CMOS. *IEEE Journal of Solid-State Circuits*, 40(1), 132-143.
- Blat, C. E., Nicollian, E. H., & Poindexter, E. H. (1991). Mechanism of negative-bias-temperature instability. *Journal of Applied Physics*, 69(3), 1712-1720.
- Bordelon, J., DeShazo, D., Tranchina, B., & Yeric, G. (1999). Proper test structure design for reliability Physics evaluation. *IEEE International Reliability Physics Symposium Tutorials*.
- Breed, D. J. (1975). A new model for negative voltage instability in MOS devices. *Applied Physics Letter*, 26(3), 116-118.
- Brisbin, D., & Chaparala, P. (2008). The Effect of the Subthreshold Slope Degradation on NBTI Device Characterization. *IEEE International Integrated Reliability Workshop*, 96-99.
- Brugler, J. S., & Jespers, P. G. A. (1969). Charge pumping in MOS devices. *IEEE Transaction on Electron Devices*, 16(3), 297-302.
- Cellere, G., Valentini, M. G., & Paccagnella, A. (2004). Effect of channel width, length and latent damage on NBTI. *IEEE International Conference on Integrated Circuit Design and Technology*, 303-306.

- Chakravarthi, S., Krishnan, A. T., Reddy, V., Machala, C. F., & Krishnan, S. (2004). A comprehensive framework for predictive modeling of negative bias temperature instability. *IEEE International Reliability Physics Symposium Proceedings*, 79-86.
- Chung, S. S., Lo, D. K., Yang, J. J., & Lin, T. C. (2002). Localization of NBTI-induced oxide damage in direct tunneling regime gate oxide pMOSFET using a novel low gate-leakage gated-diode (L<sup>2</sup>-GD) method. *IEEE International Electron Devices Meeting Technical Digests*, 13-16.
- Deal, B. E., Sklar, M., Grove, A. S., & Snow, E. H. (1967). Characteristics of surface-state charge (Q<sub>ss</sub>) of thermally oxidized silicon. *Journal of Electrochemical Society*, 114, 266-274.
- Degraeve, R., Groeseneken, G., Bellens, R., Depas, M., & Maes, H. E. (1995). A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides. *IEEE International Electron Devices Meeting*. 863-866.
- Denais, M., Bravaix, A., Huard, V., Parthasarathy, C., Ribes, G., Perrier, F., et al. (2004). On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's. *IEEE International Electron Devices Meeting Technical Digests*, 109-112.
- Dikmen, C. T., Dogan, N. S., Osman, M., & Bhattacharyya, A. (1994). Hot-carrier degradation in LDD-MOSFETs at high temperatures. *Solid-State Electronics*, 37(12), 1993-1995
- Du, G. A., Ang, D. S., Teo, Z. Q., & Hu, Y. Z. (2009). Ultrafast measurement on NBTI. *IEEE Electron Device Letters*, 30(3), 275-277.
- Ershov, M., Saxena, S., Karbasi, H., Winters, S., Minehane, S., Babcock, J., et al. (2003). Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field-effect transistors. *Applied Physics Letters*, 83(8), 1647-1649.
- Everbeing (2002). *Model : BD-6 Analytical Probe Station*. Hsinchu, Taiwan: Everbeing International Group.
- Filippi, R. G., Biery, G. A., & Wachnik, R. A. (1993). The electromigration short-length effect in Ti-AlCu-Ti metallization with tungsten studs. *Journal of Applied Physics*, 78(6), 3756-3768.
- Groeseneken, G., Maes, H., Beltran, N., & DeKeersmaecker, R. F. (1984). Reliable approach to charge-pumping measurements in MOS transistors. *IEEE Transaction on Electron Devices*, 31(1), 42-53.
- Haywood, S. K., & De Keersmaecker, R. F. (1985). Hole trapping and interface state generation during bias-temperature stress of SiO<sub>2</sub> layers. *Applied Physics Letter*, 47(4), 381-383.
- Ho, P. S. (2004). Electromigration reliability of Cu interconnects and impact of low k dielectrics. *IEEE International Reliability Physics Symposium Tutorials*.

- Ho, T. J. J., Ang, D. S., Tang, L. J., Phua, T. W. H., & Ng, C. M. (2009). Role of nitrogen on the gate length dependence of NBTI. *IEEE Electron Device Letters*, 30(7), 772-774.
- Hu, C., Tam, S. C., Hsu, F. C., Ko, P. K., Chan, T. Y., & Terril, K. W. (1985). Hot-electron-induced MOSFET degradation—Model, monitor, and improvement. *IEEE Transactions on Electron Devices*, 32(2), 375-385.
- Hu, Y. Z., Ang, D. S., & Du, G. A. (2008). An improved methodology for monitoring NBTI induced threshold voltage shift of scaled p-MOSFETs. *IEEE International Reliability Physics Symposium Proceedings*, 743-744.
- Huard, V., Denais, M., & Parthasarathy, C. (2006). NBTI degradation: From physical mechanisms to modelling. *Microelectronics Reliability*, 46(1), 1-23.
- Huard, V., Monsieur, F., Ribes, G., & Bruyere, S. (2003). Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs. *IEEE International Reliability Physics Symposium Proceedings*, 1-5.
- Huard, V., Parthasarathy, C., & Denais, M. (2005). Single-hole detrapping events in pMOSFETs NBTI degradation. *IEEE International Integrated Reliability Workshop*, 5-9.
- JEP122F (2010). *Failure mechanisms and models for semiconductor devices*. Arlington, VA: JEDEC Solid State Technology Association.
- Jeppson, K. O., & Svensson, C. M. (1977). Negative bias stress of MOS devices at high electric-fields and degradation of MNOS devices. *Journal of Applied Physics*, 48(5), 2004-2014.
- JESD202 (2006). *Method for characterizing the electromigration failure time distribution of interconnects under constant-current and temperature stress*. Arlington, VA: JEDEC Solid State Technology Association.
- JESD28-A (2001). *Procedure for measuring n-channel MOSFET hot-carrier-induced degradation under DC stress*. Arlington, VA: JEDEC Solid State Technology Association.
- JESD28-1 (2001). *N-Channel MOSFET hot carrier data analysis*. Arlington, VA: JEDEC Solid State Technology Association.
- JESD33-B (2004). *Standard method for measuring and using the temperature coefficient of resistance to determine the temperature of a metallization line*. Arlington, VA: JEDEC Solid State Technology Association.
- JESD63 (1998). *Standard method for calculating the electromigration model parameters for current density and temperature*. Arlington, VA: JEDEC Solid State Technology Association.
- JESD90 (2004). *A procedure for measuring p-channel MOSFET negative bias temperature instabilities*. Arlington, VA: JEDEC Solid State Technology Association.

- JESD92 (2003). *Procedure for characterizing time-dependent dielectric breakdown of ultra-thin gate dielectrics*. Arlington, VA: JEDEC Solid State Technology Association.
- JP-001 (2002). *Foundry process qualification guidelines*. Arlington, VA: JEDEC Solid State Technology Association.
- Kaczer, B., Arkhipov, V., Degraeve, R., Collaert, N., Groeseneken, G., & Goodwin, M. (2005). Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification. *IEEE International Reliability Physics Symposium Proceedings*, 381-387.
- Keithley (2000). *Evaluating hot carrier induced degradation of MOSFET devices [AN-2197]*. Cleveland, OH: Keithley Instruments Inc.
- Keithley (2007). *Making charge-pumping measurements with the model 4200-SCS semiconductor characterization system and series 3400 pulse/pattern generator [AN-2851]*. Cleveland, OH: Keithley Instruments Inc.
- Keithley (2009). *Model 4200-SCS Semiconductor Characterization System: Reference Manual [4200-901-01]*. Cleveland, OH: Keithley Instruments, Inc.
- Kerber, A., Cartier, E., Pantisano, L., Rosmeulen, M., Degraeve, R., Kauerauf, T., et al. (2003). Characterization of the  $V_T$ -instability in  $\text{SiO}_2/\text{HfO}_2$  gate dielectrics. *IEEE International Reliability Physics Symposium Proceedings*, 41-45.
- Ketchen, M. B., & Bhusan, M. (2011). *Microelectronic Test Structures for CMOS Technology*. New York: Springer.
- Kimizuka, N., Yamaguchi, K., Imai, K., Iizuka, T., Liu, C. T., Keller, R. C., et al. (2000). NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10-micron gate CMOS generation. *Proceeding of Symposium on VLSI Technology*, 92-93.
- Kraft, O., Sanchez, Jr., Bauer, M., & Arzt, E. (1997). Quantitative analysis of electromigration damage in Al-based conductor lines. *Journal of Material Research*, 12(8), 2027-2037
- Lall, P., Pecht, M., & Hakim, E. (1995). Characterization of functional relationship between temperature and microelectronic reliability. *Microelectronics Reliability*, 35(3), 377-402.
- La Rosa, G., Fuarin, F., Rauch, S., Acovic, A., Lukaitis, J., & Crabbe, E. (1997). NBTI – Channel hot carrier effects in PMOSFETS in advanced CMOS technologies. *IEEE International Reliability Physics Symposium Proceedings*, 282-286.
- Lu, D., Ruggles, G. A., & Wortman, J. J. (1988). Effects of processing conditions on negative bias temperature instability in metal-oxide-semiconductor structures. *Applied Physics Letter*, 52(16), 1344-1346.

- Makabe, M., Kubota, T., & Kitano, T. (2000). Bias-temperature degradation of pMOSFETs: mechanism and suppression. *IEEE International Reliability Physics Symposium Proceedings*, 205-209.
- Martin, A., O'Sullivan, P., & Mathewson, A. (1998). Dielectric reliability measurement methods: A review. *Microelectronics Reliability*, 38(1), 37-72.
- McPherson, J., & Baglee, D. A. (1985). Acceleration factors for thin gate oxide stressing. *IEEE International Reliability Physics Symposium Proceedings*, 1-5.
- Mitani, Y., Nagamine, M., Satake, H., & Toriumi, A. (2002). NBTI mechanism in ultra-thin gate dielectric - Nitrogen-originated mechanism in SiON -. *IEEE International Electron Devices Meeting Technical Digests*, 3-6.
- Moazzami, R., Lee, J., & Hu, C. (1989). Temperature acceleration of time-dependent dielectric breakdown. *IEEE Transactions on Electron Devices*, 36(11), 2462-2465.
- Montgomery, D. C., & Runger, G. C. (2010). *Applied Statistics and Probability for Engineers*. New Jersey: John Wiley & Son Inc.
- Mooraka, R., Manjularani, K. N., Ho, D., Mathur, N., & Puchner, H. (2007). A simulation methodology for the operational life time study of deep sub-micron technology integrated circuits using channel length dependent NBTI model. *International Workshop on Physics of Semiconductor Devices*, 146-149.
- Moore, G. E. (1975). Progress in digital integrated electronics. *IEEE International Electron Devices Meeting Technical Digests*, 11-13.
- Nicollian, E. H., Berglund, C. N., Schmidt, P. F., & Andrews, J. M. (1971). Electrochemical charging of thermal SiO<sub>2</sub> films by injected electron currents. *Journal of Applied Physics*, 42(13), 5654-5664.
- Nicollian, P. E., Hunter, W. R., & Hu, J. C. (2000). Experimental evidence for voltage driven breakdown models in ultrathin gate oxides. *IEEE International Reliability Physics Symposium Proceedings*, 7-15.
- Ohring, M. (1998). *Reliability and failure of electronic materials and devices*. San Diego, CA.: Academic Press.
- Pan, T-M (2006). Channel length dependence of negative bias temperature instability on pMOSFETs with either B- or BF<sub>2</sub>-implanted source/drain. *Journal of Electrochemical Society*, 153(7), G707-G711.
- Parthasarathy, C., Denais, M., Huard, V., Ribes, G., Vincent, E., & Bravaix, A. (2005). Characterization and modeling NBTI for design-in-reliability. *IEEE International Integrated Reliability Workshop*, 158-162.
- Pierret, R. F. (1996). *Semiconductor Fundamentals*. Reading, MA: Addison-Wesley.

- Rangan, S., Mielke, N., & Yeh, E. C. C. (2003). Universal recovery behaviour of negative bias temperature instability. *IEEE International Electron Devices Meeting Technical Digests*, 341-344.
- Reisinger, H., Blank, O., Heinrigs, W., Muhlhoff, A., Gustin, W., & Schlunder, C. (2006). Analysis of NBTI degradation- and recovery-behavior based on ultra fast  $V_T$ -measurements. *IEEE International Reliability Physics Symposium Proceedings*, 448-453.
- Reisinger, H., Vollertsen, R.-P., Wagner, P.-J., Huttner, T., Martin, A., Aresu, S., et al. (2009). A study of NBTI and short-term threshold hysteresis of thin nitrided and thick non-nitrided oxides. *IEEE Transactions on Device and Materials Reliability*, 9(2), 106-114.
- Sasada, K., Arimoto, M., Nagasawa, H., Nishida, A., Aoe, H., Dan, T., et al. (1998). The influence of SiN films on negative bias temperature instability and characteristics in MOSFET's. *IEEE International Conference on Microelectronic Test Structures*, 207-210.
- Schlunder, C., Hoffmann, M., Vollertsen, R.-P., Schindler, G., Heinrigs, W., Gustin, W., et al. (2008). A novel multi-point NBTI characterization methodology using Smart Intermediate Stress (SIS). *IEEE International Reliability Physics Symposium Proceedings*, 79-86.
- Schroder, D. K., & Babcock, J. A. (2003). Negative bias temperature instability: Road to cross at deep submicron silicon semiconductor manufacturing. *Journal of Applied Physics*, 94(1), 1-18.
- Sekhar, D. C., Ray, P. P., De Souza, M. M., & Chaparala, P. (2004). Edge effect under temperature bias stress of 0.18 $\mu$ m PMOS technology. *International Conference on Microelectronics*, 2, 645-648.
- Shah, J. (1992). *Hot carriers in semiconductor nanostructures : physics and applications*. Boston, MA: Academic Press.
- Shen, C., Li, M.-F., Foo, C. E., Yang, T., Huang, D. M., Yap, A., Samudra, G. S., et al. (2006). Characterization and physical origin of fast  $V_{th}$  transient in NBTI of pMOSFETs with SiON dielectric. *IEEE International Electron Devices Meeting Technical Digests*, 333-336.
- Shen, C., Li, M.-F., Wang, X. P., Yeo, Y.-C., & Kwong, D.-L. (2006). A fast measurement technique of MOSFET Id-Vg characteristics. *IEEE Electron Device Letters*, 27(1), 55-57.
- Strong, A. W., Wu, E. Y., Vollertsen, R. P., Sune, J., & Rosa, G. L. (2009). *Reliability wearout mechanisms in advanced CMOS technologies*. New York: John Wiley & Sons.
- Takeda, E., Nakagome, Y., Kume, H., & Asai, S. (1983). New hot-carrier injection and device degradation in submicron MOSFETs. *IEE Solid-State and Electron Devices Proceedings I*, 130(3), 44-150.

- Tan, C. M., & Roy, A. (2007). Electromigration in ULSI interconnects. *Materials Science and Engineering: R: Reports*, 58(1-2), 1-75.
- Taur, Y., & Ning, T. H. (1998). *Fundamentals of modern VLSI devices*. Cambridge: Cambridge University Press.
- Thewes, R., Brederlow, R., Schlunder, C., Wieczorek, P., Hesener, A., Ankele, B., et al. (1999). Device reliability in analog CMOS applications. *IEEE International Electron Devices Meeting Technical Digests*, 81-84.
- Tsujikawa, S., Mine, T., Watanabe, K., Shimamoto, Y., Tsuchiya, R., Ohnishi, K., et al. (2003). Negative bias temperature instability of pMOSFETs with ultra-thin SiON gate dielectrics. *IEEE International Reliability Physics Symposium Proceeding*, 6-11.
- Varghese, D., Saha, D., Mahapatra, S., Ahmed, K., Nouri, F., & Alam, M. (2005). On the dispersive versus arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: Measurements, theory, and implications. *IEEE International Electron Devices Meeting Technical Digests*, 11-14.
- Vollertsen, R. P. (2004). Gate oxide reliability methodology and models. *IEEE International Reliability Physics Symposium Tutorials*.
- Wu, E. Y., Aitken, J., Nowak, E., Vayshenker, A., Varekamp, P., Hueckel, G., et al. (2000). Voltage-dependent voltage-acceleration of oxide breakdown for ultra-thin oxides. *IEEE International Electron Devices Meeting Technical Digests*, 541-544.
- Yamamoto, T., Uwasawa, K., & Mogami, T. (1999). Bias temperature instability in scaled p-polysilicon gate p-MOSFET 's. *IEEE Transactions on Electron Devices*, 46(5), 921-926.

## List of Publications

1. **Soon, F. Y., & Soin, N.** (2011). NBTI degradation with Optimized  $I_D$ - $V_G$  sweep for wafer fabrication monitoring. *3<sup>rd</sup> Asia Symposium on Quality Electronic Design Proceedings*, 272-276.
2. (Submitted) **Soon, F. Y., Hatta, S. F. W. M., Hadi, D. A., Soin, N., Alam, A. H. M., & Nordin, A. N.** (2011). Negative bias temperature instability characterization and lifetime evaluations of submicron pMOSFET. *International Journal of Electronics*.