Oxide-defects driven study on reliability and synaptic response characterization of logic devices

Ju, Xin

2021

Ju, X. (2021). Oxide-defects driven study on reliability and synaptic response characterization of logic devices. Doctoral thesis, Nanyang Technological University, Singapore. <https://hdl.handle.net/10356/152681>

<https://hdl.handle.net/10356/152681>

<https://doi.org/10.32657/10356/152681>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 15:51:51 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**OXIDE-DEFECTS DRIVEN STUDY ON RELIABILITY
AND SYNAPTIC RESPONSE CHARACTERIZATION
OF LOGIC DEVICES**

JU XIN

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2021

**OXIDE-DEFECTS DRIVEN STUDY ON
RELIABILITY AND SYNAPTIC RESPONSE
CHARACTERIZATION OF LOGIC DEVICES**

JU XIN

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2021

STATEMENT OF ORIGINALITY

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarized materials, and has not been submitted for a higher degree to any other University or Institution.

05-Aug-2021

Date

NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU

JU Xin

JU Xin

SUPERVISOR DECLARATION STATEMENT

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

05-Aug-2021

Date _____

Assoc Prof. ANG Diing Shenp

AUTHORSHIP ATTRIBUTION STATEMENT

This thesis contains material from five papers published in the following peer-reviewed journals and conference proceedings where I was the first and/or corresponding author.

Chapter 4 is published as X. Ju, D. S. Ang, and C. J. Gu, “Impact of Channel Hot-Hole Stressing on Gate-Oxide Trap’s Emission,” *IEEE Transactions on Electron Devices*, vol. 67, no. 11, pp. 4720–4727, Sept. 2020, doi: 10.1109/TED.2020.3019982, X. Ju, and D. S. Ang, “Alteration of Gate-Oxide Trap Capture/Emission Time Constants by Channel Hot-Carrier Effect in the Metal-Oxide-Semiconductor Field-Effect Transistor,” *IEEE Access*, vol. 8, pp. 14048-14053, Jan. 2020, doi: 10.1109/ACCESS.2020.2966577, and X. Ju, and D. S. Ang, “Impact of Hot-Carrier Stress on Oxide Trap Switching Activity in Nanoscale FETs,” *2019 China Semiconductor Technology International Conference (CSTIC)*, July 2019, doi: 10.1109/CSTIC.2019.8755672.

The contributions of the co-authors are as follows:

- Assoc. Prof. D. S. Ang suggested the initial project direction, provided guidance in the interpretation of the measurement data, as well as edited the manuscript.
- I prepared the experiment set up, collected all the data measured in the experiments, and assisted in establishing a model to help an understanding of the main physical mechanism. Also, I analyzed the data collected as well as drafted the manuscript.

- Dr. C. J. Gu carried out the first-principles simulation, providing guidance in the analysis of simulation results collected.

Chapter 5 is published as X. Ju, and D. S. Ang, “Response of Switching Hole Traps in the Small-Area P-MOSFET Under Channel Hot-Hole Effect,” in *2019 IEEE International Reliability Physics Symposium (IRPS)*, May 2019, pp. 1–4, doi: 10.1109/IRPS.2019.8720476.

The contributions of the co-authors are as follows:

- Assoc. Prof. D. S. Ang provided the initial project direction, supervised the data analysis, and edited the manuscript.
- I conducted the experiment and collected all the measurement data in this project. Also, I interpreted the data collected as well as drafted the manuscript.

Chapter 6 is published as X. Ju, and D. S. Ang, “Gate-Oxide Trapping Enabled Synaptic Logic Transistor,” in *2020 IEEE International Reliability Physics Symposium (IRPS)*, Jun. 2020, pp. 1–6, doi: 10.1109/IRPS45951. 2020.9129338.

The contributions of the co-authors are as follows:

- Assoc. Prof. D. S. Ang suggested the initial project direction, provided guidance in the interpretation of the measurement data, and edited the manuscript.
- I designed the experimental methods and prepared characterization setup in this project. Also, I collected all the data measured in this experiment as well as analyzed the experimental data measured and drafted the manuscript.

05-Aug-2021

Date

NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU
NTU NTU NTU NTU NTU NTU NTU NTU

JU Xin

ACKNOWLEDGEMENT

Time flies, the past four years of my doctoral study have given me many exciting memories and at the same time all sorts of mixed feelings well up in my heart. The journey of conducting and articulating the research works is full of arduousness and it is undoubtedly the most challenging task I have undertaken to date. Fortunately, the selfless dedication and cordial help from professors, research members, friends, and family members have given me the strength to overcome the challenges that arise during a PhD course. This work would not have been possibly completed without their gracious assistance and constant encouragement over the years.

First and foremost, I would like to take this opportunity to express my utmost appreciation to my project advisor and lifetime mentor – Assoc. Prof. Ang Diing Shenp from Nanyang Technological University (NTU), for his tremendous support, invaluable experience, and continuous guidance throughout the course of my doctoral study. His immense scientific knowledge and valuable experience in device physics and characterizations along with passion in research have motivated me to improve my research competency through extensive reading, independent thinking, and teamwork mindset. Furthermore, he had guided me very well in delivering research findings through writing and presentation. I am grateful to him for the numerous discussions in research work and life which had made my research experience productive and gratifying. Most importantly, I am grateful to have such a precious opportunity to work with Prof. Ang who had instilled in me with passion in work, freedom of thinking, and an intriguing way of life.

Furthermore, I am extremely privileged to have worked with Prof Ang's Research Group whom I feel like my second family with irreplaceable friendship, constructive

collaborations, and fruitful discussions. I would like to genuinely thank all other lab mates Dr. Tung Zhi Yan, Dr. Zhang Haizhong, Dr. Kwang Sing Yew, Dr. Zhou Yu, and Mr. Pranav Sairam Kalaga for their technical conversations and knowledge sharing. In particular, I would also like to thank Dr. Lin Ye for being very supportive in many aspects during my doctoral study. Senior Sing Yew, appreciate your patience in helping to proofread the thesis. Your feedback is particularly helpful! Greatest gratitude to their assistance, I can carry out experimental research so smoothly and thoroughly.

I would like to owe a vote of thanks to the well-experienced technicians in NTU Semiconductor Characterization Lab, Mr. Seow Yong Hing, Ms. Seet Lye Ping, and Mr. Ng Yong Chiang, for their technical and maintenance support in the prober stations. I also extend my gratefulness and cherish the friendship from Zhiyi, Xinhua, Yuxi, Brain Jia Xu, Fang Qiang, Yuanbo, and all my friends in NTU for their help and lovely companions.

In addition, I would like to extend my gratitude to the thesis advisory committee members – Prof. Kim Tae Hyoung and Prof. Tang Xiaohong, for their constructive reviews, sharp proposals, and timely suggestions which have helped me to be more focused on the project direction and enriched my research insight. I am grateful to the Ministry of Education Singapore and Nanyang Technological University for a 4-year generous co-funding research scholarship.

Last but not least, I would like to dedicate this dissertation to my family – particularly my dearest parents Ju Zhoufeng and Liu Yaling. They have always been by my side whenever I need it. Their everlasting love and unconditional care have been constant supports for me to keep moving forward in my life and, I am forever thankful.

THANK YOU to all of you – this journey would not have happened without you!

TABLE OF CONTENT

STATEMENT OF ORIGINALITY	i
SUPERVISOR DECLARATION STATEMENT	ii
AUTHORSHIP ATTRIBUTION STATEMENT	iii
ACKNOWLEDGEMENT	vi
TABLE OF CONTENT	viii
ABSTRACT.....	xi
LIST OF FIGURES	xv
LIST OF TABLES	xxvii
LIST OF ABBREVIATIONS	xxviii
LIST OF SYMBOLS	xxxi
CHAPTER 1: INTRODUCTION	1
1.1 Motivation.....	1
1.1.1 Downscaling of CMOS Technology.....	1
1.1.2 Reliability Challenges in Nanoscale MOS Devices.....	3
1.1.3 Artificial Synapse for Brain-Inspired Computing.....	8
1.2 Objectives	11
CHAPTER 2: LITERATURE REVIEW	13
2.1 Introduction to Bias Temperature Instability	13
2.1.1 Revival of Investigation on BTI	13
2.1.2 Recent Studies on Single Charge Trapping on Individual Defects...	17
2.2 Recent Progress on Hot Carrier Degradation Study in MOSFETs.....	22
2.2.1 Introduction to HCD	22
2.2.2 Physical Mechanisms for HCD.....	23

2.2.3	Impact of HCD on Large and Nano-Scale Devices	26
CHAPTER 3: CHARACTERIZATION TECHNIQUES		29
3.1	Introduction.....	29
3.2	Ultra-Fast Measurement Techniques	29
CHAPTER 4: IMPACT OF CHANNEL HOT-HOLE STRESSING ON GATE- OXIDE TRAP'S EMISSION		37
4.1	Introduction.....	37
4.2	Defects Identification Method	39
4.3	Experimental Details.....	41
4.4	Results and Discussion	45
4.4.1	Channel Hot-Hole effect	46
4.4.2	Effect of Rest Period	56
4.4.3	Statistical Variations	63
4.4.4	Trap's Position Along the Channel.....	65
4.4.5	Possible Physical Mechanism(s).....	69
4.5	Chapter Summary	76
CHAPTER 5: RESPONSE OF SWITCHING HOLE TRAPS IN THE SMALL- AREA P-MOSFET UNDER CHANNEL HOT-HOLE EFFECT		77
5.1	Introduction.....	77
5.2	Experimental Details.....	79
5.3	Results and Discussion	82
5.4	Chapter Summary	93
CHAPTER 6: GATE-OXIDE TRAPPING ENABLED SYNAPTIC LOGIC TRANSISTOR		95
6.1	Introduction.....	95

6.2	Experimental Details.....	101
6.3	Results and Discussion	102
6.4	Chapter Summary	120
CHAPTER 7: CONCLUSION and RECOMMENDATION.....		122
7.1	Conclusion	122
7.2	Recommendations for Future Work.....	125
LIST OF PUBLICATIONS		127
REFERENCES		129

ABSTRACT

In the modern information technology era, the rapid evolution of very-large-scale integration (VLSI) technology benefited significantly from the aggressive downscaling of geometry of the fundamental building block, that is metal-oxide-semiconductor (MOS) transistor. New architectures and materials have been introduced into the advanced complementary metal-oxide semiconductor (CMOS) device for continuous downscaling in order to boost the state-of-the-art integrated circuits' (ICs') performances. At the same time, however, the understanding of the reliability of CMOS is increasingly challenging and transformed by the rapid development of technology. With the dimension of a device shrinking into nanoscale regions, the dramatic increase in gate oxide electric field and device operation temperature becomes a critical limiting factor for the continual downscaling of transistors. In addition, the impact of random capture and emission by a single defect in gate oxide becomes more prominent, causing the device reliability to be more stochastic and time-variant. As a result, this brings severe challenges to the assessment of device performance, reliability characterization, and underlying mechanisms analysis.

Bias temperature instability (BTI) and hot carrier degradation (HCD) are deemed as the two most critical reliability issues in a wide range of degradation phenomena. Both aging mechanisms remarkably affect the performance of MOS transistors due to fluctuation of parametric shifts, and therefore, must be properly understood and modeled. To date, numerous experimental and theoretical studies had been carried out to understand the effect of gate voltage (V_g) stress on oxide traps. The study on the effect of drain voltage (V_d) stress on oxide traps is quite limited although it is similarly important. In this work, we aim to systematically examine the effect of channel hot-hole

(CHH) stress on oxide traps in nanoscale p-MOS devices. Generally, the emission time of hole traps is presumed to be time-invariant, but our studies suggest otherwise. A change in trap's emission time constants under operating gate voltage is observed after CHH stress. Traps which are relatively inactive (capture once and do not emit within the measurement time frame) may be rendered much more active (exhibit many capture and emission events) after the CHH stress due to a decrease in the emission time constant. On the other hand, traps which are active may become inactive after the CHH stress due to an increase in the emission time constant. Reversion to the pre-CHH-stress behavior is observed after extended rest under unbiased condition. The traps could then undergo similar changes in emission behavior when the CHH stress is reapplied.

We proposed a three-states defect model with nitrogen atom in vacancy-interstitial (V_o-O_i) defect based on ab-initio atomic simulation. We also discussed the role of hydrogen atoms in hot carrier degradation. The three defect states include permanent trap state ($S1$), shallow trap state ($S2$), and metastable relatively permanent state ($S3$). This model provides a further understanding of emission behavior in an active trap.

Apart from the study of the impact of CHH stress, we also reported two distinct peculiarities of defect behavior in small-area MOS devices when subjected to mixed-stress conditions. First, in all devices tested, the occupancy rate (OR) of a trap could either be increased or decreased as the drain voltage $|V_d|$ is increased towards the gate voltage $|V_g|$. A further increase of $|V_d|$ would eventually lead to a decrease in OR. Second, the reduced emission probability is typically reversed upon removing the drain bias but in ~11% of the traps, this is not the case, i.e., the carrier-trapping became more permanent. In order to obtain a compact device model, it is essential to have a thorough understanding of these new behaviors which arise from BTI combined with drain (V_d)

bias stressing. In addition, these new behaviors should also be accounted for in circuit-level reliability modeling.

To meet the increasing demand for computation, bandwidth, and latency in applications such as cloud, network, and edge computing, circumventing the inherent limitations of von-Neumann bottleneck in conventional computation systems is desperately needed. Neuromorphic computing is considered to be the most promising approach for the development of low-power circuits and systems due to its distributed memory/processor architecture and massive connectivity. The fundamental of building a bio-mimetic system is successful implementation of simple, tunable, and repeatable synaptic element. This has spurred the burgeoning attention of exploring diverse materials and devices architecture for possible candidates of synaptic devices. We successfully demonstrated that a normal logic CMOS transistor is capable of emulating the versatile bio-synaptic-like responses such as tuning of short-term and long-term plasticity in excitatory and inhibitory processes by modulating the filled traps in channel conductance of MOS device.

The higher-order functionalities like meta-plasticity and spike-timing-dependent plasticity are also demonstrated by encoding relative timing. These findings provide an encouraging solution for hardware development of neuro-realistic spiking neural networks due to its advantages in manufacturability, scalability, as well as voltage requirement compatible with CMOS logic.

In summary, this thesis discusses our studies on the key characterization issues and physical mechanisms of channel-hot carrier (CHC) stressing in reliability and variability in nanoscale MOS devices. This provides physical understanding of HCD in reliability qualification and helps reliability-aware circuits design in nanoscale CMOS technology. In addition, successful demonstration of synaptic characteristics with a logic transistor

by modulating charge capture/emission kinetics in gate oxide shows the potential of realizing neuromorphic systems capable of encoding and processing complex spatiotemporal information.

LIST OF FIGURES

Figure		Page
1-1	The International Technology Roadmap for Semiconductors (ITRS) trend lines for the prediction of feature size of integrated circuit metal-oxide-semiconductor devices, data obtained from ITRS 2013 and ITRS 2.0 (ITRS 2015) [7].	2
1-2	Charging and discharging of single defects in metal-oxide-semiconductor transistors, leading to various reliability and fluctuation problems that greatly decreases the performance of integrated devices.	4
1-3	(Left) Schematic of a classical 6T static random-access memory (SRAM) cell during standard operation, showing the negative-bias temperature instability (NBTI) and positive-bias temperature instability (PBTI) stresses for p- and n-type metal-oxide-semiconductor devices, respectively. (Right) The variation in the design margin caused by different fluctuation problems in complementary metal-oxide-semiconductor (CMOS) SRAM technology, due to a reduction in device feature size [26].	5
1-4	With the shrinkage of device area, the variations (e.g., ΔV_t) in single oxide trap-induced device performance become more pronounced, despite the trap number being substantially reduced.	7
1-5	The concept of neuromorphic computing, which is designed to imitate the unique and massive parallelism of biological neural networks and would achieve truly cost-efficient computing [45].	9
1-6	Typical schematic of biological neural and artificial synapses [49].	11
2-1	Illustration of typical bias temperature instability and hot-carrier degradation mechanisms for n- and p-type metal-oxide-semiconductor transistors in actual complementary metal-oxide-semiconductor circuit operating conditions.	14
2-2	Schematic of a transistor cross-section, illustrating a representation of the negative-bias temperature instability stress effect that occurs in p-type metal-oxide-semiconductor devices.	15
2-3	Variation in maximum gate overdrive voltage with a decrease in the thickness of the equivalent gate oxide layer in advanced nodes, to meet the requirement of a 10-year lifetime for devices [70].	16

2-4	In large-area devices, a continuous recovery characteristic is typically observed. This originates from the superposition of the charge-transition events of many single defects. When the geometry of metal-oxide-semiconductor transistors is downscaled to that of nanoscale devices, the number of active defects is also substantially reduced, whereas the effect of a single defect on the device behavior becomes more pronounced. Consequently, scaled nodes exhibit individual charge-transition events during recovery, which are visible as discrete steps in their recovery profiles [96].	17
2-5	Comparison of two different oxide thickness regimes of metal-oxide-semiconductor field-effect transistor degradation with stress time. The changes in interface trap density and bulk trap density show that the bulk defect dominates the negative-bias temperature instability degradation in a sub-nanometer equivalent-oxide-thickness (EOT) regime [70].	19
2-6	A non-radiative multiphonon charge-capture model is proposed to describe the charge-trapping behaviors of single defects. Specifically, the four states of the defect model, comprising two neutral defect states 1 and 1', and two charged defect states 2 and 2', account for the existence of a metastable state. Together with E' center properties, this model explains the bias-dependent charge-capture and emission behavior of single defects [86], [91].	20
2-7	(a) Bias and temperature dependence of the charge-capture and charge-emission times of a switching oxide trap. A four-state defect model is used to explain the bias- and temperature-dependent charge-trapping/detrapping behavior, which facilitate different transition paths. (b) The corresponding approximate potential energy surfaces along a configuration coordinate are depicted [98], [99].	21
2-8	(Left) With downscaling of the power supply voltage, hot-carrier degradation may become a more serious reliability-limiting factor. (Right) Transistors in a real-circuit-application scenario exhibit diverse bias conditions with associated aging mechanisms [15], [100].	22
2-9	(Left) Schematic diagram of hot carrier degradation that occurs in metal-oxide-semiconductor field-effect transistors. (Right) Sketch of the process of Si-H bond dissociation induced by the successive bombardment of two hot carriers [105].	23
2-10	The interface coordinate distribution and energy distribution of interface trap density due to hot-carrier stress degradation of a metal-oxide-semiconductor field-effect transistor [105].	24
2-11	Schematic of a single energetic carrier and multi-vibrational excitation carrier mechanisms [110].	25

2-12	A schematic representation of combinations of MVE- and AB-mechanisms. It means that the bond can be firstly excited to an intermediate level i by several colder carriers and then dissociated by a solitary high energetical carrier [112].	25
2-13	A representative testing procedure of positive-bias temperature instability–hot-carrier degradation–positive-bias temperature instability, showing the effect of hot-carrier degradation on devices [114].	26
2-14	The ΔV_t shift under alternating hot-carrier degradation-positive-bias temperature instability stress conditions in short- and long-channel devices, where two different stress sequences are hot-carrier degradation-positive-bias temperature instability-hot-carrier degradation-positive-bias temperature instability and positive-bias temperature instability-hot-carrier degradation-positive-bias temperature instability [113].	27
2-15	Recovery part of degradation under the application of mixed negative-bias temperature instability/hot-carrier degradation stress under different V_d bias conditions [117].	28
2-16	Modified characteristics of emission time constant during mixed NBTI-HCD stress, the value of which tends to decrease as drain ($ V_d $) bias increases [117].	28
3-1	Module for testing ultra-fast bias temperature instability that performs various measurement types, including spot, step sweep, smooth sweep, and sample measurements. The timing of each type of measurement is flexibly modified by adjusting the test sample rate and independent measurement settings [118], [119].	30
3-2	Illustration of the window for spot measurements during a specific pulse waveform. The start- and stop-sampling points are denoted as 75% and 90% of each step width, respectively [118].	31
3-3	Schematic of actual connections of a four-terminal local sensor to test devices during the experimental characterization, where the 4225-RPM module is connected to the device under test via an SMA attached to a dual SSMC adapter cable (Y-cable). The customized bias and waveform functions are applied via the 4225-RPM module to the gate and drain terminals, whereas the source and bulk terminals are grounded via the line of Y-cable labeled with black tape [118], [119].	33
3-4	Schematic of Model 4225-RPM local sensing connection. The center conductor of the force triaxial connector is connected to a device under test (DUT) high, whereas its outer shield is connected to DUT low. In this connection, the sense connector is not used without any influence [118], [119].	35

4-1	Construction of a trap spectral map (TSM) (bottom) by the extraction of individual trap emission events from three representative drain current degradation, $ \Delta I_d $, recovery traces (top) ensuing BTI stress. The emission time constant, τ_e , of a defect is obtained by tracking the discrete step height, h , from the single trap emission events, i.e., each defect is identified by a cluster of (τ_e, h) points as shown in the TSM (bottom). Plot of $ \Delta I_d $ recovery step versus emission time yields clusters of data points which identify traps that are active (i.e., repeatedly capture and emit) under the recurring measurement cycles.	40
4-2	A schematic illustration of the electrical testing sequence. Monitoring steps under operating condition at $V_g = -1$ V performed before and after the CHH stress allow the emission characteristics of active oxide traps to be compared under similar measurement conditions.	41
4-3	An example of I_d evolution monitored under the operating condition at V_g of -1 V using the same DUT as in following testing. The dotted lines are eye guides, attempting to delineate what appears to be an ΔI_d recovery step, which is related to a hole captured by an oxide trap at $\sim 10^{-4}$ s amidst the “noisy” measurement data.	42
4-4	Four typical $ \Delta I_d $ recovery traces during 100 s relaxation interval at $V_g = -0.6$ V ensuing 10 s biasing interval at operation condition of $V_g = -1$ V. The step-like recovery of degraded drain current ($ \Delta I_d $) corresponds to discharging of pre-charged oxide traps during the emission interval. $ \Delta I_d ^{\text{eoc}}$ labeled represents the remnant of $ \Delta I_d $ at the end of each cycle.	43
4-5	A plot of $ \Delta I_d $ recovery step heights versus the corresponding emission times (τ_e) gives a trap spectral map (TSM) comprising clusters of data points. (a) and (b) TSMs ($ \Delta I_d $ (I_d recovery due to trapped-hole emission) versus emission time) obtained from the first and 25 th monitoring steps, respectively. The percentages show the switching rates of the active traps.	44
4-6	A breakdown of the percentages of traps affected or unaffected by CHH stress. 10 identical test devices with a total of 21 active traps were examined over a period of year.	45
4-7	(a)(i) and b(i) Selected $ \Delta I_d $ traces measured during the monitoring step made before and after CHH stress at $V_d = V_g = -1.5$ V. $ \Delta I_d $ recovery steps due to the emission of trapped holes can be seen. In b(i), the cycle numbers are indicated in the same colors as the traces. (a)(ii) and b(ii) Corresponding trap spectral maps (with occupancy rates as indicated) for two active oxide traps D1 and D2.	46
4-8	$ \Delta I_d $ at the end of each cycle, $ \Delta I_d ^{\text{eoc}}$ (blue line) as a function of monitoring cycle number. Cycle numbers before and after the step	47

decrease due to trap D3 emission are shown in red (middle panel). $|\Delta I_d|$ recovery steps due to active D3 switching after the $V_d = -2$ V CHH stress are shown as green circles (right panel). Vertical lines mark cycles in which D3 was inactive (i.e., no hole capture thus no emission was observed).

4-9	(a)(i) and b(i) Selected $ \Delta I_d $ traces measured during the monitoring step, carried out after CHH stress at $V_d = -2$ V and -2.5 V, respectively (V_g fixed at -1.5 V). (a)(ii) and b(ii) Corresponding trap spectral maps for with occupancy rates as indicated. A new switching trap D3 appeared while trap D2 disappeared after the CHH stress.	48
4-10	$ \Delta I_d ^{\text{eoc}}$ (blue line) as a function of monitoring cycle number. Cycle numbers before and after the step increase due to D2 being charged are shown in red. D2 emissions are shown as purple circles (left panel).	49
4-11	(a)(i) TSM showing three active traps in another pristine device. (a)(ii) TSM after CHH stress at $V_g = V_d = -1.5$ V showing a marked reduction in X1's SR and an increase in X3's SR. (b) $ \Delta I_d ^{\text{eoc}}$ versus cycle number showing the 6 cycles in which X1 was switching in the beginning. X1 showed no emission after it became charged in cycle 13.	50
4-12	TSMs (a) showing two active traps (E1 and E2) in another pristine device. (b) and (c) After stressing the device under channel hot-hole condition at $V_d = -1.5$ V and -2.0 V, respectively with V_g fixed at -1 V. (d) Reapplying the CHH stress produced a similar persistent increase of E1's cluster, after a one-month rest was restored to the pristine state; The capture time constant τ_c for the traps, extracted based on Eqn. 4-2, are indicated below the respective clusters.	52
4-13	The probability density function for emission characteristics of the two defects, B1 and B2 in test device B at $V_g = -0.6$ V during the pre-stress monitoring step. (a) Emission time distributions of active traps fitted according to Eqn. 4-1, which follows an exponential distribution, consistent with a first-order reaction-limited process [139]. Extremely good fitting of the measured data to the exponential function can be achieved. Cumulative density function for an active traps' capture characteristic at $V_g = -1$ V. (b) Occupancy versus capture interval characteristics of active traps, fitted according to Eqn. 4-2.	53
4-14	(a)(i) Emission time (t_e) distributions of trap E1 before and after CHH stresses. Lines are curve fitting based on the expression shown. The extracted emission time constant (τ_e) are as indicated. (a)(ii) "Before 2 nd CHH": E1's t_e distribution after a three-day rest was restored to the pristine state in (a)(i). Reapplying the CHH stress produced a similar persistent increase of E1's τ_e , which was again restored to the	55

pristine case after a one-month rest (orange sphere). (a)(iii) Transfer curves before and after CHH stress at different V_d 's (V_g fixed at -1 V), after rest and after the 2nd CHH stress and rest. (b) Trap F3 in another test device also exhibited a gradual, persistent increase in τ_e after CHH stresses [133].

4-15	(a)(i), (b)(i) Selected $ \Delta I_d $ recovery traces obtained from a monitoring step performed on the device in Figure 4-7 and 4-9 after specific rest periods. (a)(ii), (b)(ii) Corresponding trap spectral maps.	57
4-16	Evolution of $ \Delta I_d ^{\text{eoc}}$ as a function of monitoring cycles. Vertical orange lines (right panel) denote cycles in which trap D2 was inactive.	58
4-17	Seven $ \Delta I_d $ relaxation traces out of a total of a hundred recorded after the two-month rest, each showing both D2 and D3 emissions (marked by cyan and magenta arrows, respectively) within the same trace.	59
4-18	(a) Another example of both D2 and D3 emissions occurring in the same $ \Delta I_d $ recovery trace measured in cycle 1 of the monitoring step performed after reapplication of the $V_d = -2.5$ V CHH stress following the two-month rest. (b) TSM taken after reapplication of the $V_d = -2.5$ V stress showing a marked decrease and increase in D2's and D3's SR, respectively.	62
4-19	A charged oxide or interface trap in the vicinity of a current bottleneck can induce a large change in drain current through an electrostatic effect [143].	64
4-20	(a) $ \Delta I_d $ recovery step versus applied drain bias V_d from which the positions of three active switching traps along the channel are extracted and expressed as fractions of the channel length L ($L = 0$ corresponds to the source end). (b) TSM of the three traps (@ $V_d = -0.1$ V) with the individual switching rates (SR's) as indicated. (c) SR of the three traps as a function of V_d , taken during the first V_d sweep from -0.1 V to -1.5 V.	66
4-21	(a) Remeasured SR of the three traps as a function of V_d remeasured showing a marked change in trap Y3's SR. (b) TSM retaken at $V_d = -0.1$ V showing the absence of Y3 cluster. (c) TSM taken subsequently at 150°C showing the shift of emission times to shorter time scale and the appearance of the Y3 cluster. (d) SR as a function of V_d taken with the source (S) and drain (D) connections interchanged. (e) SR as a function of V_d taken subsequently with S and D interchanged back.	67
4-22	(a) Dependence of Y2's SR on V_g , giving evidence of a threshold oxide-voltage for Y2 activation which is about -0.9 V. The inset is lateral trap distribution, the schematic sketch of the relative positions of the three characterized traps within the oxide. (b) A schematic	68

cross-sectional diagram of an p-MOSFET, depicting two general hot-carrier (red filled circle) trajectories. It also shows the associated secondary impact ionization events, yielding energetic tertiary holes (open circles) at depths (lower) and (deeper) along trajectory 2.

4-23	(a) Schematic energy band diagram of the structure under positively charged condition. The release and restoration of H atom induces transformation in defect state. (b) A three-states defect model with a N atom in vacancy-interstitial (V_o-O_i) defect. H atoms play important role in hot carrier degradation. The transition between different defect states can be observed. (c) The possible atomic structure of shallow hole trap and deep hole trap.	71
4-24	The initial position of H was placed in four different locations near the N of V_o-O_i defect with 2-fold coordinated N. This is to account for the diffusion path of H atom into the two-fold coordinated N defect. The MEP and CTL of trap state transition from $S1$ to $S3$ for each case show relatively deep CTL ($S3$, e.g., CTL=5.589 eV), considering it is at the meta-stable state.	72
4-25	The initial position of H atom was placed in four different locations near the N of V_o-O_i defect with 2-fold coordinated N considering the diffusion path of H atom into the two-fold coordinated N defect. The MEP and CTL of trap state transition from $S1$ to $S2$ for each case.	73
4-26	(Top) Atomic structure of V_o-O_i defect network. An obvious atomic structural change due to the release or restoration of H atom during structural relaxation is observed. (Bottom) Corresponding MEP curve and CTL of the structures shown above. The energy barrier during structural transition from deep hole trap state to a stable shallow trap state is around ~0.6 eV. A change in the structure will induce a drop in CTL.	75
5-1	A schematic of stress conditions region which illustrates different stress conditions, including NBTI, CHH, and mixed NBTI and CHH. The region of stress conditions describes the boundaries of the two-dimensional parameter space (gate stressing and drain stressing biases) applied in this chapter.	78
5-2	The schematically detailed illustration of measurement sequence used to investigate the response behavior of switching oxide traps that already active in operation condition ($V_{DD} = -1$ V) to a varying drain bias (V_d) and fixed gate bias (V_g) stressing.	80
5-3	Representative example of a defect spectral map, derived by plotting the $ \Delta I_d $ recovery step versus emission time of all hole-detrapping events in the 100 recovery traces sensed at $V_g = -0.6$ V, $V_d = -0.1$ V. Left and right are defect spectral map for trap D1 and D2 in device #D measured at fresh condition and remeasured @ $V_g = -1$ V, $V_d =$	81

–0.1 V, following an earlier series of mixed NBTI/CHH testing which saw V_d being varied from –0.1 V to –1.5 V, respectively.

- 5-4 (a) Representative examples of the two kinds of V_d -induced occupancy-rate variation trends seen across all SOTs tested under mixed NBTI/CHH stressing. (b) Occupancy rate (OR) versus V_d for trap D1 and D2 in device #D which remeasured at 150 °C ((a) measurement under 100 °C). 83
- 5-5 Obtain capture (τ_c) and emission (τ_e) time constants through the time dependent defect spectroscopy measurements of trap D1 (left) and D2 (right) as a function of V_d bias. The V_g is fixed at constant –1 V and $T = 100$ °C. τ_c (τ_e) extracted according to the description illustrated in chapter 4 behave increase (decrease) when only considering inhomogeneous E_{ox} induced by the increasing V_d bias. 84
- 5-6 DUT #D measured under mixed NBTI/CHH at constant $V_g = -1.5$ V, with V_d varying. A newly activated trap D3 has a typically similar V_d dependence of its OR as D1. The τ_e and τ_c for two V_d values are labeled, clearly showing the larger increase of τ_e with $|V_d|$ as the reason for the rising D3's OR even though τ_c is simultaneously increase slightly (similar to D1 (Figure 5-5(a), left)). 85
- 5-7 Arrhenius plot for the τ_c of D1 and D2. Based on NMP theory, inset is the simplistic two-dimensional potential energy surface of the trap along the reaction coordinates between neutral and charged state. Green parabola represents the neutral state, another red one acts the charged state. It indicates charge trapping process has to overcome a certain capture barrier (E_{AE}) for the transition from neutral to charged state. 86
- 5-8 Three active SOTs (X1, X2 and X3) are exposed in another DUT #X, showing a similar traps' OR dependence on V_d , as those in DUT #D (cf. Figure 5-4(a)). 87
- 5-9 (a)(i) Defect spectral map of DUT #X measured @ $V_g = -1$ V, $V_d = -0.1$ V under fresh state, with the ORs indicated. (a)(ii) During Mixed NBTI/CHH stressing, defect spectral map for X1, X2 and X3 measured at condition of $V_g = V_d = -1.5$ V. A significant drop in the OR of X1 can be observed. (a)(iii) Defect spectral map measured @ $V_g = -1$ V, $V_d = -0.1$ V, after the measurement @ $V_g = V_d = -1.5$ V. It shows no emission is detected for X1, while the OR of X2 and X3 are similar to those before the high $|V_d|$ biasing was applied (b) After 10 s @ $V_g = V_d = -1.5$ V, DUT #X was relaxed @ $V_g = -0.6$ V, $V_d = -0.1$ V for 100 s; this sequence was repeated 100 times. $|\Delta I_d|^{cor}$ refers to remnant (or unrecovered) I_d degradation at the end of the relaxation interval. The light orange bands marked the cycles in which X1 captured a hole during the CHH stress and emitted it during relaxation; the dark bands denote otherwise. The light read circles 88

show I_d degradation or $|\Delta I_d|$ arising from X1 capturing a hole, given by the recovery in I_d observed when the captured hole was emitted in the relaxation interval. Upward arrows indicate the generation of other permanent trapped holes after X1 became permanently charged at cycle 13.

5-10	(a) Dependence of X2's and X3's OR on V_d , with the source and drain connections interchanged during measurement, i.e., V_d was applied to the original source terminal and the original drain terminal was grounded (cf. Figure 5-8, Figure 5-9(a)(i)) (b) Dependence of X2's and X3's OR on V_g , giving evidence of a threshold oxide-voltage for X3 activation. The threshold for X2 is more positive than -0.8 V and cannot be measured since the relaxation $V_g = -0.6$ V. Use of more positive relaxation V_g compromises current measurement speed. Inset shows that X2 and X3 might be situated close to each other.	89
5-11	(a) Schematic diagram depicts the impact of the occupation of trap X3 on the potential energy of trap X2. (b) A simplistic 2D potential energy surface to describe the impact of the Coulomb repulsion effect on the energy barrier of nearby traps. Upon trap X3 capture a hole, the energy barrier for trap X2 to transition from neutral to charged state is now increased (i.e., a larger capture barrier E_2) owing to the Coulomb repulsion effect, resulting in a larger trap capture time constant.	91
5-12	Capture emission time (CET) map correlation of traps from several devices which were examined over a period of time. τ_c and τ_e are directly extracted from Eqn. 4-1 and Eqn. 4-2. A comparison between the traps that are sensitive to (a) CHH and (b) E_{ox} change shows that, τ_c and τ_e of the former are much smaller than those of the latter.	93
6-1	An analogy between a biological synapse and an artificial synapse based on a logic n-FET: (a) Diagrammatic of biological neurons and the direct gap between the axon end of the previous neuron and the dendrites of the next neuron formed synapse, where the synaptic connection strength is reflective of the number of Ca^{2+} ion channels formed by the difference in the action potentials from pre- and post-synaptic neurons as the brain adapt to new information; This characteristic is emulated in the later in the form of charge trapping in the gate oxide, controlled by voltage pulses applied at gate and drain terminals. (b) 3D schematic of our logic synaptic transistor response to the different polarity pulse signal. Hole trapping (top) increases the channel conductance and can be used to represent an increase in synaptic strength. Electron trapping (bottom), on the other hand, decreases the channel conductance and the synaptic strength.	97
6-2	Gate potential stimuli inducing the corresponding excitatory and inhibitory behaviors by the hole trapping and electron trapping effect, imitating the process of updating weights during learning epochs. Specifically, transfer curves of an n-FET test device showing a	100

reduced threshold voltage or increased drain current due to hole trapping after a negative gate voltage pulse and an increased threshold voltage or reduced drain current due to electron trapping after a positive gate voltage pulse. Inset depicts the energy band diagrams corresponding to hole or electron trapping in the gate stack caused by a negative or positive stimulus applied on the presynaptic neuron (gate electrode).

6-3	The fundamental synaptic plasticity for both excitatory and inhibitory mode in an n-FET, where excitatory response caused by a negative pulse and inhibitory response arose from a positive pulse: (a) and (c) Dependence of the excitatory/inhibitory drain (post-synaptic) current response on the gate (pre-synaptic) voltage spike duration (50 to 500 μ s); Progressive modulation of the drain (post-synaptic) current evidence the spike number-dependent plasticity (SNDP): (b) and (d) Excitatory/inhibitory post-synaptic current response with varying number of negative/positive gate (pre-synaptic) spikes, and corresponding fitting, indicating the stretched-exponential modulation process in both potentiation and depression. Inset in each figure gives the details of the voltage waveforms. A wait period of 40 s (sufficient for complete recovery of the current response (cf. Figure 6-5) at the pulse voltages applied) is inserted in-between measurements involving different pulse widths.	103
6-4	(a) Comparison of the response of the device to a successive pulse train comprising 10 identical pulses (-2.0 V, 10 ms) with different frequencies or pulse intervals (10, 100, 1000 Hz). (b) The synaptic plasticity change index or gain, which depends on $(A_{10}-A_1)/A_1$, increases from 9.6% to 17.3%, suggesting the existence of spiking-rate-dependent plasticity or a dynamic high-pass filtering characteristic.	105
6-5	(a) and (b) The decay evolution of excitatory/inhibitory drain (post-synaptic) current that triggered by the negative/positive gate (pre-synaptic) spikes with different amplitudes and the same width of 10 ms, resembling the “forgetting” curves of biological synapses following pre-synapse firing. Inset in each figure gives detailed setting of the spike waveforms. All the current was monitored after each stimulus at pulsed V_g of 0.85V and V_d of 0.1 V (0.1 μ s).	106
6-6	(a) Short- and long-term synaptic weight changes as a function of the gate (pre-synaptic) spike voltage. The short- and long-term weight change is computed as the difference in the drain current before pre-synaptic spike and the average current after spike taken before 500 μ s and between 1 and 40 s, respectively. The lines are modeled with an exponential function. (b) Stability of the oxide traps of a n-FET: excitatory current response at selected repeated spiking numbers (5000). Inset shows the typical pre-synaptic excitatory pulses (10 ms) for the endurance testing.	107

6-7	Meta-plasticity behavior of a TiN/HfO ₂ gated n-FET. (a) IPSC responses with and without different amplitudes of priming pulses. Inset of (a): Holes trapping in HfO ₂ , induced by the priming pulse, leads to a greater trapping of electrons during the pre-synaptic spike due to the enhanced surface electric field and reduced the capture energy barrier. IPSC is enlarged by an excitatory priming pulse applied prior to the inhibitory pre-synaptic pulse (b) EPSC responses with and without various amplitudes of priming pulses. Inset of (b) shows reduction of hole trapping during the excitatory pre-synaptic pulse due to holes already trapped in the HfO ₂ during the prior priming pulse. EPSC is expected to be suppressed by a prior excitatory priming pulse.	109
6-8	(a) and (b) Weight percentage changes (%) as a function of the priming pulse voltages. Inset schematic diagram depicts the details on the voltage waveforms setting to trigger meta-plasticity. The readout current was collected at pulsed V_g of 0.85V and V_d of 0.1 V (0.1 μ s).	110
6-9	Different forms of spike-time dependent plasticity or STDP characteristic learning in a TiN/HfO ₂ gated n-FET, where the weight change Δw as a function of relative arrival timing difference Δt between the pre- and post-synaptic spikes: (a) Symmetrical anti-Hebbian rule, showing tunable STDP learning behavior, as a demonstration by changing pulse parameters; (b) Symmetrical Hebbian rule; The solid lines in the left and right diagram are both expressed by Gaussian fitting.	111
6-10	Asymmetrical Hebbian rule; Insets in the above figure depict the shape of pre-synaptic and post-synaptic voltage pulses waveform. The solid lines inside are expressed by exponential fitting.	112
6-11	Schematic of the non-overlapping pulse-pairs for STDP implementation, as well as manifesting the adjustability of the STDP behavior by terms of varying pulse-pairs amplitude. As schematized in the left, pre-synaptic and post-synaptic spikes are both applied on the gate electrode. The solid lines inside the right diagram are used exponential fitting.	116
6-12	Comparison of transfer curve before (1 st) and after alternating the (a) p-FET and (b) n-FET between excitatory and inhibitory responses for 9 times (10 th), demonstrating that the erase operation helps devices to return to their initial state for subsequent alternative measurements.	142
6-13	Fundamental synaptic signatures for both excitatory and inhibitory mode in a logic p-type transistor. Excitation (a) and (c) inhibition behavior triggers an EPSC and IPSC, in which spike-duration-dependent plasticity caused by a positive and negative presynaptic spike of 1.2 V and -1.2 V, respectively. Pulse number facilitation responses, i.e., spike-number-dependent plasticity, resulted from applying increasing number of positive and negative presynaptic	117

spike. The lines are also expressed well by stretched-exponential fitting (Eqn. 6-1 and 6-2).

6-14	Decay evolution of the excitatory/inhibitory drain (postsynaptic) current in case of a p-Si logic transistor, which is triggered by positive/negative gate (presynaptic) spikes with different amplitudes and the same widths (50 μ s). The inset in each figure shows detailed settings of the spike waveforms. EPSC/IPSC is monitored at V_g of -0.85 V/ -0.9 V and V_d of -0.1 V (0.1 μ s).	118
6-15	Two forms of STDP are implemented in present p-FET device, (a) symmetric Hebbian learning rule and (b) antisymmetric Hebbian learning rule. The lines in (a) and (b) are in accordance with the Eqn. 6-5 and 6-6, respectively.	119
7-1	Capture and emission time constants τ_c and τ_e , respectively for two active electron traps extracted from the measurement phase performed after each channel hot-electron (CHE) stress. The test device is a TiN/HfO ₂ -gated n-channel transistor.	126

LIST OF TABLES

Table		Page
6-1	The fitting values of scaling factor (A) and time constant (τ) for the symmetric Hebbian STDP learning rule.	113
6-2	The fitting values of scaling factor (A_+/A_-) and time constant (τ_+/τ_-) for the positive and negative Δt in non-overlapping pulse pairs enabled asymmetric Hebbian STDP learning rule.	115

LIST OF ABBREVIATIONS

AC	alternating current
ACS	Automated Characterization Suite
ANN	artificial neural network
AI	artificial intelligence
ALD	atomic layer deposition
BTI	bias temperature instability
BTS	bias-temperature stress
CDF	cumulative density function
CMOS	complementary metal-oxide semiconductor
CTL	charge transition level
CPU	central processing unit
CET	capture emission time
CHH	channel hot hole
CHC	channel hot carrier
DC	direct current
DNBTI	dynamic negative-bias temperature instability
DPBTI	dynamic positive-bias temperature instability
DUT	device under test
EOC	end of each cycle
EOS	end of stress
EOT	equivalent oxide thickness
EES	electron-electron scattering
ESR	electronic spin resonance
ET	electron trap
FeFET	ferroelectric field-effect transistor
FinFET	fin field-effect transistor
FETs	field-effect transistors
HCI	hot carrier injection
HCD	hot carrier degradation
HKMG	high- k /metal gate
IC	integrated circuit

LTP	long-term plasticity
MOSFET	metal-oxide-semiconductor field-effect transistor
MVE	multiple vibrational excitations
MEP	minimum energy path
MRAM	magnetoresistive random access memory
NBTI	negative-bias temperature instability
NMP	non-radiative multiphonon
NEB	nudged elastic band
OR	occupancy rate
PBTI	positive-bias temperature instability
POA	post-oxidation annealing
PUF	physically unclonable function
PSC	post synaptic current
PCRAM	phase change random access memory
PMU	pulse measurement unit
RTN	random telegraph noise
RTF	random trap fluctuation
RDD	random dopant distribution
RPM	remote preamplifier
RRAM	resistive random access memory
RF	radio frequency
SCS	semiconductor characterization system
SMU	source measurement unit
SNM	static-noise margin
STDP	spike-timing-dependent plasticity
SRDP	spike rate-dependent plasticity
SNDP	spike number-dependent plasticity
SOT	switching oxide trap
SPCT	single pulse charge trapping
SRAM	static random-access memory
STP	short-term plasticity
SR	switching rate
TRNG	true random number generator
TDDB	time-dependent dielectric breakdown

TSM	trap spectral map
UFS	ultra-fast switching
VASP	Vienna <i>ab initio</i> simulation package
VBE	valence-band electron

LIST OF SYMBOLS

I_d	drain current
I_g	gate current
V_g	gate voltage
V_t, V_{th}	threshold voltage
V_d	drain voltage
V_{DD}	supply voltage
SS	subthreshold slope
G_m	transconductance
L	channel length
W	channel width
T_{ox}	gate dielectric thickness
C_{ox}	gate oxide capacitance per unit area
E_{ox}	gate oxide electric field
t_{inv}	channel inversion layer thickness
$E_{coulomb}$	coulomb energy
μ	channel mobility
T	temperature
k	Boltzmann constant
q	electrical charge, 1.602×10^{-19} C
n_c	channel carrier concentration
E_f	fermi level
ϵ_{Si}	permittivity of semiconductor substrate
ϵ_{SiON}	permittivity of SiON oxide
E_{AE}, V_e	activation energy
Q_{it}, N_{it}	interface trap density
Q_{ot}, N_{ot}	oxide charge density
C_g	gate capacitance
R	recoverable component
P	permanent component
V_O	oxygen-vacancy
h	recovery step height

τ_c	capture time constant
τ_e	emission time constant

CHAPTER 1: INTRODUCTION

1.1 Motivation

1.1.1 Downscaling of CMOS Technology

Integrated circuits (ICs), which are often referred to as microelectronic circuits, chips, or simply microchips, comprise a set of electronic circuits, which includes a series of transistors, built upon a small, flat piece usually of silicon. ICs have arguably always been the focal point of the information technology revolution. Their advent has revolutionized the electronics industry and paved the way for the development of devices such as mobile phones, personal computers, televisions, and many appliances found around the home. In addition, the widespread use of ICs has led to the global proliferation of modern electronic devices.

After entering the era of intelligence, advanced ICs technology continually enables and also benefits from the rapid technological advances made in other traditional industries, as well as contributes to their vigor and vitality. IC industries thus form a basic, leading, and strategic framework that propel the rapid advancement and innovation of emerging information technologies such as fifth-generation mobile communications, big data, automotive electronics, aerospace, and artificial intelligence (AI).

Complementary metal-oxide-semiconductor (CMOS) technology is deemed the key foundation for a vast number of ICs, which typically comprises an immense number of silicon-based metal-oxide-semiconductor field-effect transistors (Si-MOSFETs). Over

Figure 1-1. The International Technology Roadmap for Semiconductors (ITRS) trend lines for the prediction of feature size of integrated circuit metal-oxide-semiconductor devices, data obtained from ITRS 2013 and ITRS 2.0 (ITRS 2015) [7].

the past five decades, the exponential growth in the number of metal-oxide-semiconductor (MOS) devices per chip has significantly empowered the fast progress experienced in the semiconductor industry. The Moore's law accurately predicts the evolution of semiconductor technology, as outlined by the International Technology Roadmap for Semiconductors, shown in Figure 1-1. That is, the number of transistors per unit area on a microchip is expected to double every 18 months, or the minimum feature size of devices is estimated to reduce by $1/\sqrt{2}$ of the latest technology node every 18 months [1], [2]. To achieve this inherent goal, researchers, and engineers endeavour to miniaturize transistors for ever better performance of ICs technology with a smaller form factor [3]. The benefits of down-scaling come from two-fold. On one hand, the geometry reduction of a single device enables more devices to be fitted into the same-sized chips, thus significantly increasing the density and complexity of ICs, and meeting ultimately cost requirements of various advanced applications [1], [4], [5]. On the other

hand, owing to the relentless shrinking of channel width and length, each device can switch faster with much lower power consumption [6].

Recently, intensive efforts have been made to surmount the limitations in the performance and manufacturability of ICs, and to maintain continual downscaling in state-of-the-art ICs manufacturing process. For instance, the introduction of new materials and newfangled technologies, such as strain engineering [8], high-k/metal gate technology [9], and multi-gate device structures [10] further boots the performance of ICs. Given the benefits gained from drastic improvements in the finesse of the equipment and processes involved in transistor fabrication, the feature size of MOS devices has now been pushed to the atomic scale. Most leading semiconductor companies, such as Taiwan Semiconductor Manufacturing Company, Samsung, and Intel have reached sub 10-nm technology nodes and released 5-nm fin field-effect transistor-based technology [11-14]. They have also mapped out achievement routes of 3-nm and 2-nm technology nodes for the next few years.

1.1.2 Reliability Challenges in Nanoscale MOS Devices

Despite this encouraging progress so far, though, the reliable operation of transistors under their nominal bias condition, which is of utmost significance for all such technologies, remains to be achieved. The determination of an accurate risk evaluation has attracted widespread attention as a means to achieve this objective. The transition to ultra-scale CMOS devices is part of the downscaling into the nanoscale region to enable ever-increasing integration density, but also results in certain drawbacks. In particular, increased gate-oxide electric field and elevated operating temperature, creates large challenges to device reliability. The following two factors are also severe obstacles that nullify the performance benefits gained by conventional CMOS scaling [15-17]. First,

Figure 1-2. Charging and discharging of single defects in metal-oxide-semiconductor transistors, leading to various reliability and fluctuation problems that greatly decreases the performance of integrated devices.

view from oxide electric field, slowed downscaling of the power supply voltage (V_{DD}) in proportion to the linear dimensions of transistors does not safeguard the oxide electric field-related reliability issues of high-density ICs [18-20]. The theoretical limitation placed on the subthreshold slope (SS) must satisfy the standby power requirements, and it is therefore difficult to proportionally downscale the threshold voltage (V_t) of highly-scaled field-effect transistors (FETs) [21]. Moreover, leaving a certain design margin is needed to ensure safe operation of the circuits in the presence of emerging non-ideal factors, such as soft errors and dynamic variation. Thus, a drastic increase in the electric field with a shrinkage of oxide thickness, even under normal operating conditions, will significantly enhance the severity of aging issues and shorten device lifetime [15], [22]. These adverse aging phenomena necessitate increased design complexity, which results in a reduced design margin for ICs. On the other hand, the continual increase in the density of transistors and the complexity of ICs also leads to an ever-increasing power

density of a chip over time. However, their heat dissipation capacity does not correspondingly improve, and therefore, the heat production cannot be appropriately downscaled. As a result, the operating temperature of devices or ICs in real-time applications continually increases [19], [23]. Furthermore, most device degradation problems are accelerated under high-temperature operating conditions. Based on afore discussed factors, the combined influence of higher electric stresses and elevated temperatures means that device aging is an increasingly significant problem, which must be fully understood and modeled, as well as robust design consideration in superior performance systems.

Figure 1-3. (Left) Schematic of a classical 6T static random-access memory (SRAM) cell during standard operation, showing the negative-bias temperature instability (NBTI) and positive-bias temperature instability (PBTI) stresses for p- and n-type metal-oxide-semiconductor devices, respectively. (Right) The variation in the design margin caused by different fluctuation problems in complementary metal-oxide-semiconductor (CMOS) SRAM technology, due to a reduction in device feature size. [24]

The parametric fluctuations in device performance are another prominent concern resulting from the miniaturization of transistors. This phenomenon originates from static variation induced by intrinsic process fluctuations and also dynamic variation induced by charge transition under various bias conditions [24-27]. According to the literature,

oxide traps are considered to be the central physical factor responsible for several effects detrimental to the operating characteristics and reliability of metal-oxide-semiconductor transistors. Figure 1-2 presents the typical reliability and fluctuation problems that occurred in miniaturized devices, such as random telegraph noise (RTN) [28], [29], random trap fluctuation (RTF) [30], bias-temperature instability (BTI), and BTI-induced variability [31-33], all of which are associated with complex dynamic capture/emission characteristics within gate-oxide traps. The undesirable changes in transistor characteristics posed by these aging mechanisms are of prime practical interest for precise reliability assessment of realistic circuit usage. This is particularly true for timing-sensitive circuits, for example, static random-access memory (SRAM) is highly susceptible to timing jitter caused by parametric shifts. Reliable operation requires that the V_t for nominal devices in an SRAM cell is highly reproducible, to avoid transistor pairs mismatch [34]. Figure 1-3 depicts an indication of typical BTI that occurs in a conventional six-transistor (6T) SRAM cell. As expected, the design headroom that is required to guarantee the desired stability tightens with reduced device feature size, as do the noise tolerance criteria. Consequently, a sharp increase in non-ideal effects in the SRAM of CMOS ICs significantly increases the probability of their transient failure [28], [35].

Last but not least, it is eminently difficult to control the uniformity of thickness of a gate oxide, which only comprises a few atomic layers, during advanced fabrication processes [36]. Each device has a distinct oxide thickness with a unique interface-roughness profile, which is attributable to the randomly distributed characteristics of highly scaled FETs. [37]. On top of this, several random dopant atoms typically underlie a channel. This means the drain current density is no longer uniformly distributed, and instead, multiple percolation paths are observed along the channel. [38-40]. Intriguingly,

the influence subjected to a single oxide trap on device behavior is considerably

Figure 1-4. With the shrinkage of device area, the variations (e.g., ΔV_t) in single oxide trap-induced device performance become more pronounced, despite the trap number being substantially reduced.

pronounced in scaled nodes, as demonstrated in Figure 1-4 [41]. Hence, despite the substantial decrease in the number of traps, though, the V_t shift or noise power is not decreased. Indeed, the transient nature of carrier trapping, which presents one of the major challenges in modeling BTI, has been shown to be primarily associated with the capture and emission of holes by switching oxide traps (SOTs) under pulsed gate stressing. With the continual miniaturization of transistors toward devices a few tens of nanometers dimension, the V_t shift proceeds in discrete steps. Although the discrete and stochastic nature of defect creation and charge trapping is challenging to overcome, investigating how to do so offers a unique opportunity to study the physics behind these phenomena at a single-defect level. The charge transition behaviors of small numbers of single defects have evolved as the main root of reliability challenges in modern transistor technologies. This trend drives reliability problems to show increasingly stochastic and time-varying characteristic, leading to the exploration of dynamic responses to device performance degradation.

The abovementioned problems are serious challenges to device-lifetime prediction, reliability characterization, and analysis of the underlying physical mechanisms of devices. An in-depth understanding of the intricate charge-trapping dynamics of gate-oxide defects is essential for modeling device behavior and developing accurate models to predict circuit aging amongst others. In particular, defect properties vary greatly, and many of their peculiarities are lost in averaged analyses of large-area device degradation behavior. In nano-scale devices, however, the charge transition events of individual defects can be traced and studied in detail. Therefore, studies centered on effects of individual oxide defects on the reliability and variability of nanoscale MOS devices resulting are meticulously analyzed in this thesis. This includes an analysis of the effects of channel hot carrier (CHC) stress and the combined effect of BTI and CHC on the degradation of device performance.

1.1.3 Artificial Synapse for Brain-Inspired Computing

The investigations discussed above cover the key aspects of characterization, physical mechanism analysis, and the resulting variability in nanoscale MOS devices. Thus, they provide guidelines for reliability qualification and reliability-aware IC design for such devices. Furthermore, the implications of charge-trapping kinetics in oxide defects, as detailed in this part, will aid in understanding the following chapter, which presents logic transistors for potential neuromorphic applications.

The recent surge in global information science and technology is credited to the enormous advancements in silicon IC-based computing platforms. The amount of data available globally is currently exploding at an astounding rate and the processing of such big data to yield societal and economic benefits calls for a massive volume of computing resources. Notably, within approximately 2 years it is expected that both technical and

Figure 1-5. The concept of neuromorphic computing, which is designed to imitate the unique and massive parallelism of biological neural networks and would achieve truly cost-efficient computing [42].

economic challenges will slow the advances in the development of current classical computing technologies. The downscaling of CMOS transistors is already slowing down and gradually approaching physical limitations. Moreover, we are on the cusp of a revolution in the ubiquity of AI and cognitive computing, and yet the large energy and throughput requirement for spanning cloud-based servers to cutting-edge devices presents a substantial challenge for current computer processors. This has created an urgent requirement for alternative solutions to complex real-world problems, to enable innovative and intelligent applications based on artificial neural networks [43]. Accordingly, the prospect of AI has fueled a rising demand for generic cognitive computing, which offers considerable promise for deriving intelligence and knowledge from vast amounts of data in an energy-efficient manner. In contrast, conventional computers are based on von Neumann architecture, and their areas for memory and computation are physically separated. As a result, they must shuttle data back and forth at high speeds, which consumes substantial time and power, which means that conventional computers are inefficient at processing vast amounts of data. To go beyond

this bottleneck and reduce the energy consumed in training and deploying neural network algorithms, computers must transgress from being based on von Neumann architectures to non-von Neumann architectures, such that memory and computational areas are co-located in some form. Non-von Neumann-based computing devices comprising neural networks would enable the use of extremely powerful software-based AI algorithms, which would benefit mankind worldwide.

The human brain is the most powerful and energy-efficient processor on Earth, and particularly excels at distinguishing patterns or inferring information under indistinct conditions. Hence, guided by the hierarchical structure and neuro-synaptic framework of the human nervous system, the new paradigm of brain-inspired or neuromorphic computing (as illustrated in Figure 1-5) is of widespread interest. This approach has been investigated owing to its distributed memory and processors, massive connectivity, extensive ability to perform calculations in parallel, and power-efficient operation [44], [45].

To fully exploit the unique capabilities of the human brain, we must faithfully emulate its functionality in hardware. A key element is a powerful artificial synapse that possesses plasticity akin to its biological counterpart, which enables local storage of memory and determination of the connectivity weight between two-linked neurons [47]. Moreover, a plastic artificial synapse can functionally and structurally modify itself in response to an external stimulus. Numerous components can achieve the required synaptic plasticity, such as resistive random-access memory (ReRAM) [48-50], phase-change memory (PCRAM) [51-53], magnetoresistive random-access memory (MRAM) [54], [55] ferroelectric field-effect transistor (FeFET) [56-58], and flash memory [59-62]. Figure 1-6 depicts an analogy between a brain synapse and a representative artificial synapse. However, very few attempts have been made to use logic transistors as a key

Figure 1-6. Typical schematic of biological neural and artificial synapses [46].

building block pertinent to neuromorphic applications hitherto. Given the CMOS-compatibility of MOS devices and their greater manufacturability than memory implementations, we used a MOS device to mimic synaptic weight-change plasticity. In particular, given the charge modulation effects and analog programming characteristics of single MOSFETs, the augmentative functionalities such as weight-dependent plasticity, spike-timing-dependent plasticity, and endurance of synaptic response were comprehensively investigated.

1.2 Objectives

This study is aimed to analyze the unique characteristic of single-oxide traps under various stressing conditions that significantly affect the reliable operation of nanoscale transistors, to further improve the physical understanding of oxide defect-centered hot-carrier degradation (HCD) mechanisms; and explore the new functionality of logic devices enabled by the modulation of charge-trapping kinetics. The detailed objectives are listed below:

- (1) To investigate the effect of CHC stress on the nature and properties of individual active traps in nanoscale devices. This will aid the counterchecking of the

experimental observations previously reported for large-area transistors and the analysis of the fundamental roots behind the degradation caused by individual traps under HCD conditions.

- (2) To examine the correlation between the effect of CHC stress and typical locations of different oxide traps along the channel direction (i.e., spatial locations), where the relative location is deduced by exploiting the readout drain bias dependence of the current recovery step heights (ΔI_{ds}) caused by the defects.
- (3) To derive an atomic-level picture of a change in the switching (capture/emission) behaviors of traps in small-area p-type MOS transistors under CHC stress, via simulations of defect-structure relaxation.
- (4) To explore the response behaviors of switching hole traps under the mixed stressing conditions of BTI and channel hot-hole stress, resulting from non-idealized BTI or HCD stress in real mixed-signal circuit applications.
- (5) To scrutinize the capability of a logic FET to emulate the key behaviors of biological synapses. This will require modification of the population of filled traps in the MOS structure itself, by interaction with neuron spikes.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction to Bias Temperature Instability

2.1.1 Revival of Investigation on BTI

MOSFETs exhibit a limited lifetime due to inevitable performance degradation during their operation. Thus, it is essential to evaluate the effect of degradation induced by various aging mechanisms on device parameters, such as drain current (I_d), threshold voltage (V_t), subthreshold slope (SS), transconductance (G_m), and channel mobility (μ), etc. [63]. Among them, V_t is the dominant parameter and allows for controlling the on/off switching by the applied gate voltage in MOSFETs. The value of V_t in a MOSFET is dependent on several parameters. However, only oxide charge density (Q_{ot}) and interface trap density (Q_{it}) can engender a drift of this value occurring in MOSFETs, and thus threaten the reliable operation of MOSFETs. This can be expressed as follows:

$$\Delta V_t = -\frac{\Delta Q_{it}(t) + \Delta Q_{ot}(t)}{C_{ox}} \quad (\text{Eqn. 2-1})$$

Based on the above equation, an increase in either Q_{it} or Q_{ot} , or both, causes a V_t shift as the bias stress progresses. The interface traps are relatively permanent, i.e., they exhibit almost no post-stress relaxation effect [64], [65] (although they may be eliminated to some extent by annealing at a higher temperature) and account for the steady drift of important transistor parameters, such as V_t , away from their nominal values. Charge

Figure 2-1. Illustration of typical bias temperature instability and hot-carrier degradation mechanisms for n- and p-type metal-oxide-semiconductor transistors in actual complementary metal-oxide-semiconductor circuit operating conditions.

trapping at oxide defects, on the other hand, is generally more transient or “volatile”, i.e., exhibits a significant relaxation effect, resulting in the extensive recovery of parametric shifts [64], [66]. To a further specification, their generation kinetics under accelerated stress conditions were intensively studied for many years usually along two major directions, namely hot-carrier injection (HCI) and time-dependent dielectric breakdown (TDDB) tests. With the negative-bias temperature instability (NBTI) effect in the p-MOSFET becoming the foremost front-end reliability issue in the early 2000s [67], efforts from both academia and industries in the past decades were centered mainly on this issue, particularly when moving to scaled technology nodes. Furthermore, the high-k/metal gate stacks are now widely used for sub-45 nm technology nodes. More recent studies have focused on the positive-bias temperature instability (PBTI) effect, which also severely decreases the performance of high-k gated n-MOSFETs [68-70]. Overall, BTI is one of the most prominent contributors to the alteration of device behavior, and studies on both NBTI and PBTI aging mechanisms in p- and n-MOSFETs, respectively, are equally vital. Figure 2-1 depicts the typical bias conditions of NBTI and PBTI in n-MOS and p-MOS transistors, respectively, under actual circuit operation. In general, set in the range of 50–250 °C, BTI typically manifests as a drift of V_t or I_d

Figure 2-2. Schematic of a transistor cross-section, illustrating a representation of the negative-bias temperature instability stress effect that occurs in p-type metal-oxide-semiconductor devices.

over time and increased performance variability when constant biases are applied across the gate oxide of a MOSFET, and electric field is studied up $E_{ox} \leq 8$ MV/cm. Figure 2-2 draws a representation of the NBTI stress progress that occurs in PMOS devices.

Thereafter, many studies have been performed in order to reveal the mechanism and impact of BTI and determine approaches to reduce its effect [71-75]. (1) BTI may cause long-term lifetime degradation of FETs. As shown in Figure 2-3, when the equivalent oxide thickness (EOT) is reduced to less than 1-nm, the ideal maximum overdrive voltage required to meet the 10-year lifetime requirement of devices is less than 0.4 V. Thus, the drastically increased electric field will significantly worsen the BTI degradation. (2) Degradation due to BTI may be rapidly reversed once the bias stress is removed, which means that the precise characterization of BTI is very difficult. During BTI measurement, the crucial concern is the corruption by the spontaneous recovery effect of V_t upon the interruption of stress, that is, gate bias stress is constantly interrupted to measure transfer curve by conventional current-gate (I_d - V_g) characterization. The measurement delay occurs after removing the gate bias stress, could range from a few milliseconds to seconds, which can lead to an amount of ΔV_t

Figure 2-3. Variation in maximum gate overdrive voltage with a decrease in the thickness of the equivalent gate oxide layer in advanced nodes, to meet the requirement of a 10-year lifetime for devices [70].

recovery and render the inaccurate (underestimated) extraction of ΔV_t . Alternatively, an ultra-fast V_t characterization method is employed, allowing for direct measurement of the ΔV_t to reduce the impact of sampling process on BTI measurement results [76], [77]. Furthermore, the fast recovery characteristic also makes BTI effect having rather complicated frequency-dependence and duty-cycle dependence [78], [79], raising more questions for device-level modeling and circuit-level reliability simulation. (3) In general, an IC is deemed to be malfunctioning when the V_t shift exceeds 50 mV or the reduction in G_m exceeds 10% [80]. ICs may not function if the value of V_t exceeds the designed tolerance. Hence, BTI-induced degradation of device parameters may cause semiconductor circuits to malfunction. For instance, an increase in the gate and path delays of logic circuits increases the risk of timing violations during operation. In addition, gate overdrive degradation, caused by an increase in V_t , leads to static-noise margin (SNM) reduction in standard RAM, and I_d reduction and frequency degradation in ring oscillators [81]. There is much controversy about the physical mechanism of the degradation caused by BTI, and a unified theoretical explanation is yet to be reached.

2.1.2 Recent Studies on Single Charge Trapping on Individual Defects

Figure 2-4. In large-area devices, a continuous recovery characteristic is typically observed. This originates from the superposition of the charge-transition events of many single defects. When the geometry of metal-oxide-semiconductor transistors is downscaled to that of nanoscale devices, the number of active defects is also substantially reduced, whereas the effect of a single defect on the device behavior becomes more pronounced. Consequently, scaled nodes exhibit individual charge-transition events during recovery, which are visible as discrete steps in their recovery profiles [97].

An analysis of the degradation of nanoscale MOS-device performance shows that NBTI exhibits step-like recovery characteristics, rather than the continuous recovery exhibited by traditional large-scale devices, thereby leading to time-dependent instability in MOS device characteristics [33], [37], [41], [82]. There are only a handful of defects or SOTs present in the gate oxide layer above the channel region in small-area MOSFETs [86]. Unlike large-area devices, each small-area MOSFET behaves differently due to random dopant distribution (RDD). This RDD causes a non-uniform

potential profile in the channel, leading to carrier conduction through the percolation paths. In addition, a single trapping event of a defect located above the constriction point may block the percolation path of a carrier conduction, resulting in a significant shift in the V_t and severe I_d degradation [83].

Although the same physical mechanism is responsible for charge trapping in both large-area and miniaturized devices, the profile of I_d drift differs for scaled MOS transistors (Figure 2-4 (right)). Figure 2-4 (left) indicates the number of defects that affect device behavior. Large-area transistors, which degrade in a relatively uniform manner, tend to exhibit an “average”, identical degradation behavior. However, the continual downscaling of transistor geometries means that single-charge transitions have evolved as discrete steps. The number of traps in each device and the magnitudes of drift in device performance drifts caused by different traps are based on Poisson and exponential distributions, respectively [33], [37]. This results in significant differences in BTI degradation and recovery among different devices, which is denoted device-to-device variation [82]. In addition, step-like degradation occurs during the V_g stress phase [84]. Furthermore, Cho *et al.* [70], [85] reported that border traps near the gate oxide interface dominate NBTI degradation in a sub-nanometer equivalent-oxide-thickness (EOT) regime, as illustrated in Figure 2-5.

Previous studies of the complex defect dynamics of the charge-transfer process between the substrate and oxide defects revealed the following intriguing feature: the oxide defects often exhibited a metastable state in both neutral and positively charged states [86-88]. In addition, the observed transient RTN, defect disappearance, and frequency-dependent behaviors, also supported the existence of metastability [89]. Based on these findings, Grassler *et al.* [90] generated an advanced physical four-state non-radiative multiphonon (NMP) model, which comprises two relatively stable states

Figure 2-5. Comparison of two different oxide thickness regimes of metal-oxide-semiconductor field-effect transistor degradation with stress time. The changes in interface trap density and bulk trap density show that the bulk defect dominates the negative-bias temperature instability degradation in a sub-nanometer equivalent-oxide-thickness (EOT) regime [70].

(1 and 2) and two metastable states (1' and 2'). In each charge state, the defect is represented by a double well. Between these two states, the stable state exhibits lower energy, and the other exhibits the closest metastable minimum. The transitions in which a charge exchange occurs between states 1 and 2' or 2 and 1' are described by an NMP process. On the other hand, the defects undergo structural relaxation between two defect configurations via a thermal barrier (i.e., $1 \rightarrow 1'$ or $2 \rightarrow 2'$), but do not alter the charged state. A significant difference between the barriers is that the charge-transfer reaction leads to bias-dependent behavior, whereas the thermally activated barriers produce bias-insensitive transition responses. The atomic configurations of the so-called E' center verify the physical accuracy of the model, calculated using *ab initio* approaches, as shown in Figure 2-6 [86], [91]. This defect class has been proposed as a likely hole-trap candidate in p-MOS transistors [92]. Further plausible trap candidates involving hydrogen have also been introduced to clarify the BTI phenomenon and the observed

Figure 2-6. A non-radiative multiphonon charge-capture model is proposed to describe the charge-trapping behaviors of single defects. Specifically, the four states of the defect model, comprising two neutral defect states 1 and 1', and two charged defect states 2 and 2', account for the existence of a metastable state. Together with E' center properties, this model explains the bias-dependent charge-capture and emission behaviour of single defects [86], [91].

experimental data (i.e., the hydroxyl- E' centers and defects in the hydrogen bridge configuration) [93], [94]. Recently, an examination of the effect of dynamic NBTI on small-area devices by Tung *et al.* [95] has revealed that a defect configuration may transform from a recoverable component (R) of total degradation to a more permanent positively charged state (P). SOTs may also exhibit intermittent charging under the applied stress, in addition to an increased emission time at random time periods. These observations were also verified by a NBTI study conducted by Kaczer *et al.* [88], which showed that the oxide defects can temporarily lose their electrical activity and were reactivated at the same site, with the same properties. This discovery of the volatility of these defects led to the proposal of a four-state model of hydroxyl- E' centers over the hydrogen bridges. In addition, several studies centered on the PBTI response in n-MOS

Figure 2-7. (a) Bias and temperature dependence of the charge-capture and charge-emission times of a switching oxide trap. A four-state defect model is used to explain the bias- and temperature-dependent charge-trapping/detrapping behavior, which facilitate different transition paths. (b) The corresponding approximate potential energy surfaces along a configuration coordinate are depicted [98], [99].

devices have proposed that elongated oxygen bonds can act as a suitable electron trap (ET) candidate in charge transition [96], [97].

The charge trapping/detrapping kinetics observed in previous studies [98], [99], as shown in Figure 2-7, demonstrated a typical SOTs bias-dependent charge-transition characteristic. This observation agrees with the defect of multi-states proposed by the NMP model. Figure 2-7(a) shows that capture time-constants decrease with increasing V_g bias, whereas the emission time increases. In addition, Figure 2-7(b) describes the possible charge-transition paths in a two-dimensional energy-coordinate diagram. The charge-capture event occurs along the transition path $1 \rightarrow 2' \rightarrow 2$ when a carrier possesses the energy to surpass the barrier between two different states. Similarly, a charge-emission event occurs along the $2 \rightarrow 1' \rightarrow 1$ path, where the emission time-constant is determined by the barrier between the states 2 and 1'. Furthermore, a relative shift in

the position of the trap energy level, which is a direct result of the applied V_g bias, is indicated by the dashed lines in Figure 2-7(b).

2.2 Recent Progress on Hot Carrier Degradation Study in MOSFETs

2.2.1 Introduction to HCD

Another non-negligible reliability-limiting factor resulting from the downscaling of V_{DD} is the drain (V_d) bias-induced HCI, as shown in Figures 2-1 and 2-8 [15], [100]. Given the need to consider various stress conditions in real circuits to achieve a more accurate predictive capability, all gate/drain bias (V_g/V_d) operating conditions should be considered, i.e., not only the “worst stress” conditions but also the interaction between

Figure 2-8. (Left) With downscaling of the power supply voltage, hot-carrier degradation may become a more serious reliability-limiting factor. (Right) Transistors in a real-circuit-application scenario exhibit diverse bias conditions with associated aging mechanisms [15], [100].

BTI and HCD under mixed BTI/HC stress conditions. Thus, studies on HCD in n-MOSFETs have revealed new degradation mechanisms, such as electron–electron scattering (EES) [101], [102] and multiple vibrational excitations (MVEs) [103]. Furthermore, to understand the BTI–HCI interaction in detail, an analysis of the

degradation of individual defects at a microscopic scale is required. This will enable predictive modeling at a transistor level, and ultimately allow an accurate predictive assessment of reliability in real-time circuit applications.

2.2.2 Physical Mechanisms for HCD

The modern CMOS fabrication process requires a post-oxidation annealing (POA) step. During this process, the interaction between highly energetic carriers and atoms leads to structural disorder at the insulator–silicon substrate interface, thus introducing many electrically active, dangling Si–H bonds [104]. As mentioned above, if a MOSFET is subjected to a drain (V_d) bias, the channel carriers are accelerated by the lateral electric field and gain substantially higher energy, which makes them hot carriers.

Figure 2-9. (Left) Schematic diagram of hot carrier degradation that occurs in metal-oxide-semiconductor field-effect transistors. (Right) Sketch of the process of Si–H bond dissociation induced by the successive bombardment of two hot carriers [105].

The energy of these carriers depends on the applied bias amplitude. When such carriers collide with the dielectric–Si interface of the MOSFET, they deposit energy and dissociate the Si–H bonds, thereby causing degradation at or near this interface and altering device performance. This dangling bond-rupture mechanism leads to the creation of interface states, and is denoted as HCD, as shown in Figure 2-9 [105]. In

Figure 2-10. The interface coordinate distribution and energy distribution of interface trap density due to hot-carrier stress degradation of a metal-oxide-semiconductor field-effect transistor [105].

addition, the density of interface traps varies with the energy across the Si bandgap, and the created traps are mostly located around the drain side of the device (Figure 2-10). Under the applied stress, the newly generated interface traps may be charged by capturing carriers. These charged traps distort device electrostatics, which is observable as a shift in V_t . Moreover, the increase in the number of dangling bonds at the interface is observed as a decrease in device μ , which results from an increase in the interface scattering of the carriers. The reduced μ leads to a reduction in SS and may be evaluated when the I_d-V_g measurements are performed [106], [107] and by the alteration of C-V characteristics of the device [108]. Thus, the MOSFET parameters are distorted.

To explain HCD in miniaturized devices, its physical origin has been recently extended to cold carriers, in which a series of collisions with low-energy particles also can create interfacial states [109]. As illustrated in Figure 2-11, the transistors under HCI stress conditions undergo various degradation mechanisms. At a high V_d/V_g value,

Figure 2-11. Schematic of a single energetic carrier and multi-vibrational excitation carrier mechanisms [110].

the channel carriers bombarding the interface are already rather energetic, because of lateral field acceleration. This triggers single-collision Si–H bond rupture, which is described as a single-energetic-carrier mechanism. In contrast, at a high V_g/V_d value, the increased carrier density triggers multiple interactions with several colder channel carriers, which excites and ultimately breaks a Si–H bond. This is referred to as the MVE process [105], [110], [111]. With a modification in the model, the real scenario of the Si-H bond breakage comprises a combination of the single- and multiple-carrier

Figure 2-12. A schematic representation of combinations of MVE- and AB-mechanisms. It means that the bond can be firstly excited to an intermediate level i by several colder carriers and then dissociated by a solitary high energetical carrier [112].

processes, as shown in Figure 2-12 [112].

Figure 2-13. A representative testing procedure of positive-bias temperature instability–hot-carrier degradation–positive-bias temperature instability, showing the effect of hot-carrier degradation on devices [114].

2.2.3 Impact of HCD on Large and Nano-Scale Devices

According to Duan *et al.*[113-115], HCD generates new cyclic electron traps (ETs) that respond to V_g stress. They speculated that newly activated ETs are created by the bombardment of hot carriers; once generated, these ETs may be switch-activated in response to PBTI. As shown in Figure 2-13, cyclic ETs clearly increase under HCD (i.e., cyclic ETs > pre-HCD cyclic ETs). Higher cyclic ETs can only be explained by the generation of more traps, as the charging conditions in the first and third stages are identical. Thus, additional cyclic ETs are generated by HCD, which can be quantitatively characterized by comparison with the number of cyclic and pre-existing cyclic ETs. Figure 2-13 also shows that the overall HCD is much larger than the degradation induced by pure PBTI, which agrees with a previously reported result [116].

Figure 2-14. The ΔV_t shift under alternating hot-carrier degradation-positive-bias temperature instability stress conditions in short- and long-channel devices, where two different stress sequences are hot-carrier degradation-positive-bias temperature instability-hot-carrier degradation-positive-bias temperature instability and positive-bias temperature instability-hot-carrier degradation-positive-bias temperature instability [113].

Figure 2-14 compares the ΔV_t shift under alternating PBTI–HCD stress conditions between short- and long-channel devices [113]. This shows in a short-channel FET, although the HCD in the third stage follows the HCD dynamics of the first stage, the PBTI in the third stage does not follow the PBTI dynamics of the first stage. In contrast, in a long-channel FET, although the PBTI in the third stage follows the PBTI dynamics of the first stage, the HCD in the third stage does not follow the HCD dynamics of the first stage. This observation clearly shows the interplay between HCD- and PBTI-induced degradation (i.e., insertion of HCD alters the PBTI dynamics in short-channel FETs, whereas insertion of PBTI alters the HCD dynamics in long-channel devices). Thus, device degradation cannot be assessed by examining only the independent effects of HCD and PBTI degradation mechanisms and the dynamics: their interaction must also be examined.

Figure 2-15. Recovery part of degradation under the application of mixed negative-bias temperature instability/hot-carrier degradation stress under different V_d bias conditions [117].

Grasser *et al.* [117] illustrated that oxide defects along all lateral positions of a channel remain in the neutral state under a mixture of NBTI and HCD, as reflected by the absence of emission-related events during post-stress relaxation. This is the primary cause of the discrepancy between experimental and simulation results at high V_d stress values, as shown in Figure 2-15. Furthermore, Grasser *et al.* [117] found that the characteristic time constants of emission (τ_e) and capture (τ_c) are distorted under non-equilibrium stress conditions, consequently, during a stress phase, a defect emits a charge carrier immediately after capturing it. Thus, defects exhibit neutral state under mixed NBTI–HCD conditions, which decreases trap occupancy (Figure 2-16).

Figure 2-16. Modified characteristics of emission time constant during mixed NBTI–HCD stress, the value of which tends to decrease as drain ($|V_d|$) bias increases [117].

CHAPTER 3: CHARACTERIZATION TECHNIQUES

3.1 Introduction

In this chapter, ultra-fast BTI measurement, which supports sub-microsecond pulse characterization and is performed using Keithley Automated Characterization Suite (ACS) software, is thoroughly discussed. In addition, the specific characterization methods used in each experiment are explained in Chapters 4–6.

3.2 Ultra-Fast Measurement Techniques

The ultra-fast BTI package integrates Keithley’s advanced DC I–V and ultrafast pulse I–V measurement capabilities with automatic test-executive software. For the characterization of reliability and novel applications in cutting-edge semiconductor technology, it is desirable to obtain the most powerful and sophisticated ultrafast measurement platform. Thus, this platform comprises a Keithley 4200-SCS parameter analyzer equipped with a 4225-PMU ultrafast I–V module, two 4225-RPM remote amplifier/switches, ACS software, an ultrafast BTI test-project module, and effective cables [118], [119].

In particular, the 4225-PMU ultrafast I–V module forms the hardware core of the ultrafast I–V measurement and is essential for high-speed electrical characterization at the microsecond level [118], [119]. This module, which incorporates a sophisticated two-channel waveform generator with high-speed voltage and current-measurement capabilities, a deep measurement buffer, and a real-time test-execution engine, supports

Figure 3-1. Module for testing ultra-fast bias temperature instability that performs various measurement types, including spot, step sweep, smooth sweep, and sample measurements. The timing of each type of measurement is flexibly modified by adjusting the test sample rate and independent measurement settings [118], [119].

complex waveform output and specification of the measurement settings. A customized test-sequence is formed by linking multiple measurement types, as exemplified in Figure 3-1. The measurements are automatically performed by the real-time test execution engine of the system through the mathematical means of the multiple samples. This reduces the volume of data that must be transferred and parsed throughout the test. Furthermore, the outcomes are cascaded back to a high-level test module to realize near-real-time analysis and automatic test-termination.

The remote amplifier/switch 4225-RPM module is designed to maximize the suitable current-measurement sensitivity of the 4225-PMU for ultra-fast I–V testing [118], [119]. Combining the remote pre-amplifier (RPM) and pulse measurement unit (PMU) reduces the relaxation effects, and thus the speed requirements for ultra-fast BTI characterization are met. The RPM helps to locate the fast-speed measurement circuits and pulse source close to the device under test (DUT), minimizing the cable length and thus decreasing

Figure 3-2. Illustration of the window for spot measurements during a specific pulse waveform. The start- and stop-sampling points are denoted as 75% and 90% of each step width, respectively [118].

parasitic cable-related effects, such as source overshoot, load-line effects, and cable capacitance. In addition, a reduction in cable length renders the round-trip signal-propagation delay shorter than the rise or fall time-interval of a defined pulse. Moreover, the 4225-RPM module also functions as an external switch matrix that is capable of automatically switching among direct current (DC), capacitance–voltage (C–V), and pulsed types of measurement, which greatly simplifies the connections of the DUT. It also provides various signals, which satisfies the requirement for multiple types of electrical characterization.

Thus, ACS is a programmable measurement platform used with the ultrafast BTI module that flexibly engineers various stress environments and stress timings, and enables various types of measurement to be performed, such as spot, sweep, and sample measurements [118], [119]. Moreover, ACS enables the characterization of BTI degradation and recovery behavior under an applied DC or (alternating current) AC stress and the incorporation of single-pulse charge trapping measurements into longer stress-measure sequences. In particular, high-accuracy, low-level measurements of degradation and recovery behaviors under various stress conditions, as well as before and after stress applications, are reliably performed by the ultrafast BTI module.

The critical advantages of the integrated test module applied during our experimental measurement are stated below [118], [119]:

(1) Enhancement in sensitivity and speed

As BTI is a highly dynamic phenomenon rapid and sensitive measurements are essential for its accurate characterization. As such, an integrated measurement system exhibits an optimal combination of measurement sensitivity and speed for ultra-fast BTI characterization. This system yields accurate ultrafast results without the use of interconnects and radiofrequency (RF) structures.

(2) Minimization of unwanted drain-to-source fields

To eliminate undesirable charge displacement and HCI effects, it is important to minimize the drain-to-source electric field in BTI degradation-recovery experiments. As the drain current is proportional to the drain-to-source field, the more sensitive the drain current measurement is, the lower is the required drain voltage. The ACS ultra-fast BTI package generates superior low-current measurements (i.e., measurements with signal and noise resolutions of less than 10 nA), facilitating the use of lower drain voltages and thereby yielding more accurate results.

(3) Reduction in relaxation time

Measurement of degradation requires high speed and highly sensitive techniques. This system meets this requirement, as single-point I_d spot measurements are performed within 1 μ s and 10-point I_d - V_g step sweeps are completed in less than 10 μ s.

(4) Ease of test-sequence development

A programmable measurement sequence of up to 20 linked tests may be intuitively created using the appropriate values of the measurement parameters. This customized

Figure 3-3. Schematic of actual connections of a four-terminal local sensor to test devices during the experimental characterization, where the 4225-RPM module is connected to the device under test via an SMA attached to a dual SSMC adapter cable (Y-cable). The customized bias and waveform functions are applied via the 4225-RPM module to the gate and drain terminals, whereas the source and bulk terminals are grounded via the line of Y-cable labeled with black tape [118], [119].

setting enables the selection of one or more measurement types, such as spot, step sweep, smooth sweep, or sample measurements, thus facilitating multifunctional

characterization via the ultrafast BTI test module [118], [119]. The details of the four measurement types are as follows:

(1) Spot

The spot measurement result is a single mean value calculated from the average of the samples used within the measurement window during the gate and drain pulsing. Figure 3-2 shows a defined measurement window that may be used for a spot measurement during a pulse, where the start- and stop-sampling points are denoted as 75% and 90% of the timeline at both the amplitude and base levels, respectively. The size of the portions to be sampled, together with the sampling frequency, contribute to the determination of how many samples must be measured at the amplitude and base levels. In general, the programmable sampling interval is set as fast as 5 ns (i.e., the sampling rate is considered to be 200 MHz).

(2) Step Sweep

The step sweep is similar to a conventional DC source measure unit (SMU) sweep, in that each step comprises an integration (or averaging) period and a settling period. Similarly, the measurement window for each step follows the rule of the measurement window used in the spot measurement, where the start- and stop-sampling points are determined at 75% and 90% of each step range, respectively.

(3) Smooth Sweep

The signal in a smooth sweep is continually sampled for the entire sweep, without including any settling time.

(4) Sample

The measurement of a sample is similar to that of a smooth sweep, except that the signal in a sample measurement is continuously sampled at constant gate and drain voltages.

The actual connections of four-terminal local sensing are schematically illustrated in Figure 3-3, in which an SMA is attached via a dual SSMC adapter cable (Y-cable) to ensure the 4225-RPM module is credibly connected to the DUT throughout the experimental characterizations. Figure 3-4 depicts the details of a local sensing connection, where each force-and-sense triaxial connector comprises a central conductor and an outer shield. The central conductor of the force connector is connected to the DUT to obtain the input signal, whereas its outer shield is connected to the DUT to obtain a grounded signal. Two sense connectors are not used in this study. In case of four-terminal MOSFET testing, the gate and drain terminals are independently connected to two pairs of central conductors of the force connector, whereas the source

Figure 3-4. Schematic of Model 4225-RPM local sensing connection. The center conductor of the force triaxial connector is connected to a device under test (DUT) high, whereas its outer shield is connected to DUT low. In this connection, the sense connector is not used without any influence [118], [119].

and bulk terminals are connected to the outer shield of each force connector for grounding. The user-designed bias and waveform is easily applied with DUTs via the

RPM module, which is controlled by the ACS module. Thus, various testing waveforms may be obtained with the ultrafast BTI module, enabling maximum flexibility and ease of use, even when customizing complex test sequences (as shown in Figure 3-1). In the experiments conducted in this study, the CHC stress and relaxation voltages were applied synchronously at the gate and drain terminals, and the behavior of the single-oxide trap was characterized by monitoring the trace of the drain current during the relaxation period. In addition, all experiments were conducted on an SUSS 200-mm probe station equipped with an auto-controllable chuck heater. The details of each experiment are discussed in the following sections.

CHAPTER 4: IMPACT OF CHANNEL HOT-HOLE STRESSING ON GATE-OXIDE TRAP'S EMISSION

4.1 Introduction

Gate oxide and interface traps pose a serious threat to the reliable operation of the MOS transistor. Their generation kinetics under accelerated aging tests have been well characterized for large-area transistors. Recent attention is mainly centered on the response of individual traps in small-area transistors, in terms of their capture/emission (τ_c/τ_e) time constants after bias-temperature instability (BTI) stress [120-122]. As revealed by simulation studies [123], the stochastic capture/emission of inversion charges by individual oxide traps can produce significant fluctuations in the drain current (I_d) of a small gate-area MOSFET, when coupled with the random dopant fluctuation effect. Hence, there have also been numerous experimental studies, mostly BTI-oriented, targeted at characterizing individual oxide traps (also known as switching traps) in small gate-area transistors [33], [120], [122], due to the importance of the parametric fluctuations on the operation of some circuits, e.g. SRAM [124], [125]. The general understanding is that τ_c and τ_e are functions of applied biases and temperature only and are invariant under a given operating condition. Interestingly, it has been found in some studies [88], [95], [126], [127] that accelerated gate stressing, typical during BTI testing, could change the capture/emission characteristics of some switching traps, showing a volatile nature not only in terms of charge occupancy, but also in terms of the

atomic structure of the underlying defects. These effects may impact models or applications that presume trap properties are stress invariant.

In analog and mixed-signal circuits, the transistor is also continuously subjected to the channel hot-carrier (CHC) effect due to a dynamic workload condition. To-date, studies have only reported the effect of gate voltage stress on oxide traps, while the effect of drain voltage stress is equally important but is seldom studied. Although a finite drain voltage (V_d) may lessen the BTI effect near the drain side of the channel due to a decrease in the effective oxide field, severe BTI degradation could still occur in the rest of the channel, especially at a high gate voltage (V_g) and temperature, and this has motivated several studies on combined BTI/CHC stressing [113], [128-131]. In contrast to the expected lessening of HCI degradation at high temperature, a past study observed worsening HCI degradation with increase in temperature in high-k gated n-MOSFETs stressed at $V_g \sim V_d$ and ascribed the observation to PBTI's positive temperature dependence. On the other hand, reports on the impact of CHC stress on the traps' properties are limited and only received attention lately [132], [133]. In that study, switching traps which were active under NBTI-only stress (i.e., capture a hole during NBTI stress and emit it during post-NBTI relaxation) were found to be inactive during combined NBTI and channel hot-hole (CHH) stress, as reflected by the absence of related emission events during post-stress relaxation. It was shown that under the combined stress, the trap's emission time was reduced more than its capture time, which resulted in a decrease of the trap's occupancy. However, it was unclear whether this was caused by the decrease in oxide electric field across the channel (due to the non-zero V_d), thus making emissions happen more easily or whether it was a consequence of a hot-hole effect. It has been shown that the emission times of oxide traps could be

substantially reduced under low oxide electric field [120], [122]. Therefore, a further examination of the CHH effect is warranted.

This chapter examines the impact of channel hot-hole (CHH) stress on oxide traps in a nanoscale p-MOSFET. A change in the traps' emission time constants under operating gate voltage is observed after the CHH stress. Herein, we present further experimental evidence showing a persistent change in a trap's τ_e after channel hot-hole (CHH) stress [133]. Besides evidence of τ_e decrease, we show evidence of τ_e increase with a probability about three times the former. It is also shown that the τ_e changes are uncorrelated with the traps' spatial positions along the channel (i.e., traps near the source end could be more affected than those near the drain end) [132]. This suggests possible roles by other subtle factors (e.g., local defect chemistry), in addition to the proposed defect structural change arising from interaction with hydrogen species released by the CHH stress [133], [134].

4.2 Defects Identification Method

In small-area (i.e., ultra-scaled) devices, only a handful of traps are present. Each of them can contribute to a significant amount of degradation in the I_d during stress. The degraded drain current ($|\Delta I_d|$) was plotted with respect to the pre-stress value during each 10 ks recovery. Three typical drain current recovery traces of a p-MOSFET is shown in Figure 4-1. Recovery of $|\Delta I_d|$ is due to discharging of pre-charged oxide traps during the emission interval. The measured data are given by the thin yellow and green lines (slightly noisy) in the top part of the figure. Each recovery step in the relaxation traces carries a discrete I_d step height, h , of the carrier-detrapping event and an emission time, τ_e , corresponding to the detrapping event. In the bottom part, the automatically extracted

Figure 4-1. Construction of a trap spectral map (TSM) (bottom) by the extraction of individual trap emission events from three representative drain current degradation, $|\Delta I_d|$, recovery traces (top) ensuing BTI stress. The emission time constant, τ_e , of a defect is obtained by tracking the discrete step height, h , from the single trap emission events, i.e., each defect is identified by a cluster of (τ_e, h) points as shown in the TSM (bottom). Plot of $|\Delta I_d|$ recovery step versus emission time yields clusters of data points which identify traps that are active (i.e., repeatedly capture and emit) under the recurring measurement cycles.

emission times and step heights which are nearly unambiguous fingerprints of each defect is marked by the thick lines together with the symbols, and the construction of trap spectral map (TSM) is built by extraction of single detrapping events during I_d relaxation traces. In order to reliably characterize the stochastic charging/discharging process, a number of stress/recovery experiments have to be performed. The statistical properties of the discrete steps in the I_d recovery traces are then analyzed by using TSM. It concludes that each cluster of (τ_e, h) points located on the TSM denotes an individual active defect in the devices (upper).

Figure 4-2. A schematic illustration of the electrical testing sequence. Monitoring steps under operating condition at $V_g = -1$ V performed before and after the CHH stress allow the emission characteristics of active oxide traps to be compared under similar measurement conditions.

4.3 Experimental Details

The DUTs used were nanoscale p-MOSFETs with drawn channel length and width of 60 nm and 120 nm, respectively. The gate stack consists of a p^+ -doped polysilicon gate and a 1.7-nm thick SiON gate dielectric developed by decoupled plasma nitridation. The operating voltage was 1 V, and the threshold voltage of DUTs was -0.3 V.

The electrical testing sequence applied to a p-channel transistor is illustrated in Figure 4-2. The temperature was 100 °C unless stated otherwise. A pristine test device was first subjected to a monitoring step, during which a gate voltage (V_g) of -1 V was applied for 10 s to induce the capture of holes by traps, resulting in a degradation (Δ) of the drain current (I_d). V_g was then switched to -0.6 V for 100 s to allow for emission of the trapped holes. While it is desirable to directly measure a trap's emission under $V_g = -1$ V, practical difficulty has, however, prevented this theoretically sound method. We have found that current measurement at a larger $|V_g|$ is typically “noisier” (Figure 4-3,

Figure 4-3. An example of I_d evolution monitored under the operating condition at V_g of -1 V using the same DUT as in following testing. The dotted lines are eye guides, attempting to delineate what appears to be an ΔI_d recovery step, which is related to a hole captured by an oxide trap at $\sim 10^{-4}$ s amidst the “noisy” measurement data.

shows an instance of the I_d evolution trace measured directly under operating condition at V_g of -1 V.) due to other random telegraphic noises; it is unclear whether they originate from the device or the measurement equipment, making the step degradation in the I_d arising from a hole-capturing event barely visible and difficult to identify. In addition, the larger current requires the use of the next higher measurement range, resulting in a decrease in measurement resolution. This compromises measurement accuracy of smaller current steps associated with some traps of interest (e.g., trap D3 in Figure 4-7 which is characterized by a current step of ~ 20 – 25 nA). To a first order, a lower $|V_g|$ (i.e., a relaxation voltage of -0.6 V in this experiment) was used to avoid the difficulty of observing trapped-hole emissions due to random telegraphic noise present at -1 V and for improving the I_d measurement resolution [126]. Since a trap’s emission time constant scales exponentially with the applied V_g , as is widely demonstrated in the literature [90], [93], a similar conclusion should apply at other V_g biases based on the emission study made at $V_g = -0.6$ V. In the 100-s interval, I_d was continuously measured (at a drain voltage $V_d = -0.1$ V) to detect trapped-hole emissions, which result in the

recovery of the degraded I_d ($|\Delta I_d|$), manifested as “downward” steps in the $|\Delta I_d|$ versus time curve (e.g., Figure 4-4) [135]. The above cycle was repeated 100 times to collect the emission statistics of traps that were active (i.e., repeatedly capture and emit holes) under the measurement conditions. The time delay between two measurement cycles

Figure 4-4. Four typical $|\Delta I_d|$ recovery traces during 100 s relaxation interval at $V_g = -0.6$ V ensuing 10 s biasing interval at operation condition of $V_g = -1$ V. The step-like recovery of degraded drain current ($|\Delta I_d|$) corresponds to discharging of pre-charged oxide traps during the emission interval. $|\Delta I_d|^{\text{eoc}}$ labeled represents the remnant of $|\Delta I_d|$ at the end of each cycle.

was of the order of several milliseconds. Active traps were identified via a plot of $|\Delta I_d|$ versus emission time which yielded the trap spectral map or TSM (e.g., Figure 4-5). Each distinct cluster of data points in the TSM corresponds to an active trap [135]. The switching rate (SR) of an active trap is defined as the number of hole-emission events observed out of the 100 applied cycles. This number approximates the probability of a net capture of a hole at the end of the 10-s interval and subsequent emission within the 100-s measurement window.

After the monitoring step, the device was subjected to a CHH stress at $V_g = V_d = -1.5$ V for 100 s. Following this, the same monitoring step was reapplied to check the

Figure 4-5. A plot of $|\Delta I_d|$ recovery step heights versus the corresponding emission times (τ_e) gives a trap spectral map (TSM) comprising clusters of data points. (a) and (b) TSMs ($|\Delta I_d|$ (I_d recovery due to trapped-hole emission) versus emission time) obtained from the first and 25th monitoring steps, respectively. The percentages show the switching rates of the active traps.

impact of the CHH stress on the emission characteristics of the traps. Since V_d is the main driving force for the CHH effect, the testing sequence was repeated for different stress V_d 's (-2.0 V, -2.5 V) with V_g fixed at either -1 or -1.5 V.

The traps' τ_c and τ_e are known to be strongly influenced by the oxide electric field [90], [120], [122]. The large $|V_d|$ applied during CHC stress reduces the oxide field along the channel, which in turn affects the capture/emission characteristics of the traps. This makes it very difficult to separate the CHC effect on the traps' time constants from the effect due to the change of the oxide field from post-CHH emission measurement. Here, we compare the emission characteristics of active traps under the same measurement conditions before and after the CHH stress. Any effect of the CHC stress may then be unambiguously revealed.

Notably, to check whether the monitoring step itself would change the emission behavior of the traps [95], [136], the trap spectral map was measured with the monitoring step repeated numerous times. No evident change in the traps' SR and emission time distribution can be observed even after repeating the monitoring steps

many times. Figure 4-5 shows an example where the comparison between the first and 25th TSM's reveals no apparent difference. The results imply that the possible role of a monitoring V_g bias on the emission characteristics may be ruled out.

4.4 Results and Discussion

In this section, all the observations from the experiment will be elaborated, followed by a discussion about the possible mechanism behind the observations. Altogether, 10 identical test devices with a total of 21 active traps were examined over a period of more than one year. After about 15 traps, a preliminary statistical trend has emerged where ~60% of the traps were unaffected by the CHH stress, ~30% and 10% had their τ_e 's increased and decreased, respectively (Figure 4-6). In the following, representative

Figure 4-6. A breakdown of the percentages of traps affected or unaffected by CHH stress. 10 identical test devices with a total of 21 active traps were examined over a period of year.

results of some test devices are shown. The outcomes clearly show that the CHH effect can alter the switching traps' emission responsible for the dynamic NBTI effect and such interaction may have to be included in the performance prediction of circuits. Typical results are discussed in the following sections.

Figure 4-7. (a)(i) and b(i) Selected $|\Delta I_d|$ traces measured during the monitoring step made before and after CHH stress at $V_d = V_g = -1.5$ V. $|\Delta I_d|$ recovery steps due to the emission of trapped holes can be seen. In b(i), the cycle numbers are indicated in the same colors as the traces. (a)(ii) and b(ii) Corresponding trap spectral maps (with occupancy rates as indicated) for two active oxide traps D1 and D2.

4.4.1 Channel Hot-Hole effect

In one test device, two active traps (D1 and D2) can be identified during the monitoring step measurement in the pristine state, as is apparent from the two distinct $|\Delta I_d|$ recovery steps in some of the selected relaxation traces (Figure 4-7(a)(i)), and the two distinct clusters in the TSM (Figure 4-7(a)(ii)) [135]. It should be noted that $|\Delta I_d|$ at the end of the first cycle, $|\Delta I_d|^{\text{eoc}}$ is above zero (~ 120 nA), indicating a remnant I_d degradation. For the remaining cycles, their $|\Delta I_d|^{\text{eoc}}$'s are the same as that of the first cycle (Figure 4-8). This shows that besides D1 and D2, other traps also captured a hole during the first 10-s interval and the holes remained trapped throughout the monitoring

Figure 4-8. $|\Delta I_d|$ at the end of each cycle, $|\Delta I_d|^{\text{eoc}}$ (blue line) as a function of monitoring cycle number. Cycle numbers before and after the step decrease due to trap D3 emission are shown in red (middle panel). $|\Delta I_d|$ recovery steps due to active D3 switching after the $V_d = -2$ V CHH stress are shown as green circles (right panel). Vertical lines mark cycles in which D3 was inactive (i.e., no hole capture thus no emission was observed).

step, i.e., the emission times of the traps concerned were much longer than the total measurement time of the monitoring step.

After the CHH stress at $V_d = -1.5$ V, no additional $|\Delta I_d|$ recovery step or cluster appeared (Figures 4-7(b)(i), (ii)), and $|\Delta I_d|^{\text{eoc}}$ for cycles 1 to 22 remained around 120 nA (Figure 4-8), like the pristine case. These observations show that the CHH stress did not generate any new active trap (within the measurement window). Both D1 and D2 appeared unaffected by the CHH stress. Interestingly, from cycle 22 to 23, a decrease in $|\Delta I_d|^{\text{eoc}}$ by ~ 20 nA occurred (Figure 4-7(b) (i)). In the following cycles, $|\Delta I_d|^{\text{eoc}}$ remained unchanged at the new lower level ~ 100 nA (Figure 4-8). The decrease of $|\Delta I_d|^{\text{eoc}}$ below the 120 nA level in the monitoring step after the CHH stress signifies the occurrence of a hole-emission event from one of the traps (referred as DX) that was charged during the very first $V_g = -1$ V interval. This should be contrasted against the case before the

Figure 4-9. (a)(i) and b(i) Selected $|\Delta I_d|$ traces measured during the monitoring step, carried out after CHH stress at $V_d = -2$ V and -2.5 V, respectively (V_g fixed at -1.5 V). (a)(ii) and b(ii) Corresponding trap spectral maps for with occupancy rates as indicated. A new switching trap D3 appeared while trap D2 disappeared after the CHH stress.

CHH stress, where DX retained the trapped hole throughout the monitoring step. However, the $|\Delta I_d|$ recovery step corresponding to DX's emission is not observed in the $|\Delta I_d|$ trace of cycle 23. Since the average $|\Delta I_d|$ at the beginning of cycle 24 is lower than that of cycle 22 by ~ 20 nA, this suggests that the emission might have occurred during the transition from cycle 22 to 23, when V_g was returned to 0 V momentarily. After this, DX was not active in the remaining cycles in view of its absence from the TSM and the unchanged $|\Delta I_d|^{\text{eoc}}$ after cycle 22.

A salient observation after the $V_d = -2$ V CHH stress pertains to the appearance of a new emission event in the $|\Delta I_d|$ traces (Figure 4-9(a)(i)) and the corresponding cluster

Figure 4-10. $|\Delta I_d|^{\text{eoc}}$ (blue line) as a function of monitoring cycle number. Cycle numbers before and after the step increase due to D2 being charged are shown in red. D2 emissions are shown as purple circles (left panel).

in the TSM (Figure 4-9(a)(ii)). This shows that a new trap (D3) was “activated” by the CHH stress. The almost identical $|\Delta I_d|$ recovery steps (20 nA) of D3 and DX (Figure 4-8) should be noted. Considering the random distribution of channel dopants, the likelihood of two different traps having almost identical $|\Delta I_d|$ recovery steps is extremely low [123]. The similar $|\Delta I_d|$ recovery steps strongly suggest that D3 is DX or is situated near DX (hence both are structurally very similar). The short τ_e of D3 is a huge contrast from that before the CHH stress. This observation of a decrease in a trap’s τ_e after CHH stress is, however, relatively rare and applies to 2 out of 21 or ~10% of the traps.

D1’s SR remained at ~50%, unchanged even after more severe CHH stresses $V_d = -2$ and -2.5 V (Figure 4-9). Its emission time distribution was also unaffected, as is evident from the almost identical position of the cluster in the TSM. D1 is an example of traps that are immune to the CHH effect (13 out of 21 or ~60%). The $|\Delta I_d|$ recovery step was, however, decreased after the $V_d = -2.5$ V CHH stress. This may be ascribed to the degradation in I_d due to the generation of interface traps and deep-level hole

trapping by the CHH stress [137], as is manifested by the upshift in $|\Delta I_d|^{\text{eoc}}$ after each stress. Simulation study [123] has shown that a trap's $|\Delta I_d|$ recovery step is dependent on the local inversion charge density, which is reduced after the generation of more positive oxide/interface trapped charges by the CHH stress.

Figure 4-11. (a)(i) TSM showing three active traps in another pristine device. (a)(ii) TSM after CHH stress at $V_g = V_d = -1.5$ V showing a marked reduction in X1's SR and an increase in X3's SR. (b) $|\Delta I_d|^{\text{eoc}}$ versus cycle number showing the 6 cycles in which X1 was switching in the beginning. X1 showed no emission after it became charged in cycle 13.

On the other hand, a marked change in D2's emission behavior is observed after the $V_d = -2$ V CHH stress. D2 totally "disappeared" from the $|\Delta I_d|$ traces and TSM in the post-CHH monitoring step. (Figures 4-9(a)(i), (ii)). Two possible scenarios may account for this result: (1) D2 is not charged during the $V_g = -1$ V interval due to either a substantial increase of τ_c and/or a decrease of τ_e ; (2) D2's τ_e is increased, hence the hole remained trapped throughout the monitoring step. An interesting change in $|\Delta I_d|^{\text{eoc}}$ (Figure 4-10) supports the second scenario. From cycle 11 to 12, $|\Delta I_d|^{\text{eoc}}$ was increased

by ~ 50 nA, equal to D2's $|\Delta I_d|$ recovery step, and it remained at the new higher level for the rest of the cycles. This signifies that D2 captured a hole in cycle 12 and retained the hole for a substantial period, i.e., its τ_e is longer than the rest of the measurement time. Before the $V_d = -2$ V CHH stress, D2's τ_e is ~ 20 ms. The results show that the τ_e was significantly increased by the CHH stress. This observation applies to 6 out of 21 or $\sim 30\%$ of the traps measured. Figure 4-11 presents the observations for a trap X1 from another test device where a transition to a more permanently charged state was seen during the monitoring phase after CHH stress, just like D2.

CHH stress at $V_d = -2.5$ V did not result in any further apparent change in the traps' activity (Figures 4-9(b)(i), (ii)). D1 remained unaffected (except for the decrease in $|\Delta I_d|$ recovery step as noted before); no emission was recorded for D2 and D3 was as active as before the -2.5 V stress. But an intriguing evolution of $|\Delta I_d|^{\text{eoc}}$ in the post-CHH monitoring step should be highlighted (Figure 4-10). $|\Delta I_d|^{\text{eoc}}$ showed two step-like increases (the first from cycle 7 to 8 and the second from cycle 87 to 88), both equal to D2's $|\Delta I_d|$ recovery step of 50 nA. A similar observation is made after the $V_d = -2$ V CHH stress, albeit there is only one step increase in $|\Delta I_d|^{\text{eoc}}$ from cycle 11 to 12. It is possible that D2 was discharged during the $V_d = -2.5$ V CHH stress and was later charged again in the post-CHH monitoring step. The two $|\Delta I_d|^{\text{eoc}}$ increases indicate suggest a neutral to +1 and then to +2 charge state transitions. *Ab-initio* simulation has revealed the negative-U nature of a charged trap [138], which allows the trap to gain an additional charge of the same polarity if the energy cost is outweighed by the decrease in system energy due to lattice relaxation.

It is known that a comprehensive capture/emission time constant extraction of all CHH-affected traps would strengthen the observation in alteration on active traps'

Figure 4-12. TSMs (a) showing two active traps (E1 and E2) in another pristine device. (b) and (c) After stressing the device under channel hot-hole condition at $V_d = -1.5$ V and -2.0 V, respectively with V_g fixed at -1 V. (d) Reapplying the CHH stress produced a similar persistent increase of E1's cluster, after a one-month rest was restored to the pristine state; The capture time constant τ_c for the traps, extracted based on Eqn. 4-2, are indicated below the respective clusters.

emission characteristic. For a given active trap, the statistical variation in t_e of each cluster (e.g., Figure 4-12) can be extracted by using the exponential distribution fitting according to [133], [134]:

$$p = \frac{t_e}{\tau_e} \cdot \exp\left(-\frac{t_e}{\tau_e}\right) \quad (\text{Eqn. 4-1})$$

where t_e is the relaxation time upon stress termination while τ_e is the emission time constant of an SOT during relaxation. From which the τ_e (@ $V_g = -0.6$ V in this case) may be determined; p is the probability density function. For instance, the emission times of traps B1, B2 before the CHH stress step, as depicted in Figure 4-13(a), follow

an exponential distribution, consistent with a first-order reaction limited process [139]. The extracted emission time constants for each trap is stated above the fitting lines. The τ_e of an active trap is a stochastic event, and it is one of the key characteristics of an active trap. Different traps have different unique τ_e . The τ_e of an active trap is the average time used by a trap to release a captured charge at the condition of a given gate voltage.

Figure 4-13. The probability density function for emission characteristics of the two defects, B1 and B2 in test device B at $V_g = -0.6 \text{ V}$ during the pre-stress monitoring step. (a) Emission time distributions of active traps fitted according to Eqn. 4-1, which follows an exponential distribution, consistent with a first-order reaction-limited process [139]. Extremely good fitting of the measured data to the exponential function can be achieved. Cumulative density function for an active traps' capture characteristic at $V_g = -1 \text{ V}$. (b) Occupancy versus capture interval characteristics of active traps, fitted according to Eqn. 4-2.

The probability for the emission of a captured charge is high around this τ_e .

Since an emission observed during relaxation indicates that the trap involved must be previously charged during t_c , the number of data points in a cluster expressed as a percentage of the total measurement cycles gives the trap's switching rate or probability P that a net capture of a hole would occur during t_c . By varying t_c , a trap's τ_c (@ $V_g =$

–1 V) may be extracted from its P versus t_c curve, that is, the fitting of a cumulative density function for an active trap's capture characteristic at $V_g = -1$ V (Figure 4-13(b)) according to the following expression: [93], [134]

$$P = \frac{\tau_e}{\tau_e + \tau_c} \cdot \left[1 - \exp\left(-\frac{t_c}{\tau_c}\right) \right] \quad (\text{Eqn. 4-2})$$

where τ_e is the corresponding emission time constant at $V_g = -1$ V. However, a monitoring step includes 100 cycles, each monitoring cycle consists of a 10-s capture interval and a 100-s emission interval. This means every single data point on the cumulative density function plot takes more than 3 hours to collect. To achieve an accurate cumulative density plot for each trap, one typically has to lengthen the capture interval to 500 s or longer. Thus, the measurement time increases significantly, and it typically needs ~3–4 days on average to draw a complete trap occupancy fitting curve. To keep the total experimental time within a reasonable level, as well as repeating the measurements to ascertain the traps' behaviors before and after channel hot-hole stress, post-stress relaxation and resuming the measurements to check for consistency of behaviors, not every DUT was fully characterized for capture time constant. After the above measurement phase, the device is subjected to a channel hot-hole (CHH) stress phase for 100 s. During CHH stress, the applied V_d reduces the oxide field along the channel and this affects the traps' τ_c and τ_e [90], [135]. Thus, trapping characteristics during the CHH stress are not studied. However, it should be noted such τ_c and τ_e changes which arise out of a change in the oxide field are temporary. If the CHC stress did not affect the trap properties, then reapplying the measurement phase after the CHC stress should reveal similar trap behaviors. However, this is not the case for some traps.

Figure 4-14. (a)(i) Emission time (t_e) distributions of trap E1 before and after CHH stresses. Lines are curve fitting based on the expression shown. The extracted emission time constant (τ_e) are as indicated. (a)(ii) “Before 2nd CHH”: E1’s t_e distribution after a three-day rest was restored to the pristine state in (a)(i). Reapplying the CHH stress produced a similar persistent increase of E1’s τ_e , which was again restored to the pristine case after a one-month rest (orange sphere). (a)(iii) Transfer curves before and after CHH stress at different V_d ’s (V_g fixed at -1 V), after rest and after the 2nd CHH stress and rest. (b) Trap F3 in another test device also exhibited a gradual, persistent increase in τ_e after CHH stresses [133].

Figures 4-12 and 4-14 show further examples of traps (E1 and F3 from two other test devices) whose τ_e ’s were persistently increased after CHH stress [133]. Unlike D2 where the significant τ_e increase has made it impractical to quantify the change within a reasonable experimental timeframe, the changes related to E1 and F3 are much smaller and thus can be measured. After a 100-s CHH stress at $V_g = -1$ V, $V_d = -1.5$ V, a re-run of the measurement phase clearly reveals a shift of the emission cluster for trap E1 towards a longer time relative to the pristine case (Figure 4-12(a)(i)). A further 100-s

CHH stress at $V_d = -2$ V (with V_g fixed at -1 V) yielded more shift (Figure 4-12(a)(i)). To quantify the observed changes, the τ_e (@ $V_g = -0.6$ V) was extracted by fitting the respective t_e emission time distributions according to Eqn. 4-1 (Figure 4-14(a)(i)). The traps' emission time distributions were fitted to the expression shown in the insets. E1's τ_e was increased from 9.6 ms (pristine) to 74 ms after the $V_d = -1.5$ V CHH stress and was further increased to 150 ms after the $V_d = -2$ V stress (with V_g fixed at -1.5 V in both cases). Similar changes apply to F3's τ_e (Figure 4-14(b)). F3's τ_e was increased from 0.17 s (pristine) to 0.32 s after the $V_d = -1.5$ V CHH stress and was further increased to 1.02 s after the $V_d = -2$ V stress (with V_g fixed at -1.0 V in both cases). In addition to the τ_e increase, a decrease of E1's occupancy rate is also observed after the CHH stress (Figures 4-12(a)-(c)). To understand this, E1's τ_c (@ $V_g = -1.0$ V) was extracted by a probability fitting plot based on Eqn. (2). Interestingly, τ_c is also increased after CHH stress, which translates to a decrease of capture probability and hence the occupancy rate for a given t_c . In this example, it can be seen that the CHH stress could make an oxide trap less likely to capture a hole but upon a successful capture, the trapped hole would be retained for a longer time.

4.4.2 Effect of Rest Period

Next, the DUTs that had undergone the application of CHH stress previously were stored at room temperature. After storage for a period of time, the DUTs were again subjected to 100 monitoring steps to examine the effect of long rest periods on traps' switching activities. Interestingly, after being stored for an adequate relaxation period,

Figure 4-15. (a)(i), (b)(i) Selected $|\Delta I_d|$ recovery traces obtained from a monitoring step performed on the device in Figure 4-7 and 4-9 after specific rest periods. (a)(ii), (b)(ii) Corresponding trap spectral maps.

some of the defects demonstrate that the emission characteristic can be restored back to their state prior to CHH stress. After testing, the device in Figure 4-15 was kept unbiased at room temperature for a period of two months. The monitoring step was applied at the one-month and two-month intervals to check the traps' behaviors. After one-month rest (Figure 4-15(a) and 4-16), no obvious change from the behaviors obtained after the last $V_d = -2.5$ V CHH stress can be observed. D1 maintained a similar SR and reduced $|\Delta I_d|$ recovery step. No emissions were recorded for D2, indicating that it remained charged as previously concluded. D3 is as active as it was after the last CHH stress.

On the other hand, significant changes are observed after two months of rest (Figure 4-15(b) and 4-16). D2 was active again, with a SR comparable to the cases (Figure 4-7) before it became more permanently charged. Similarity between the

Figure 4-16. Evolution of $|\Delta I_d|^{\text{eoc}}$ as a function of monitoring cycles. Vertical orange lines (right panel) denote cycles in which trap D2 was inactive.

emission time distribution after rest and those in the previous cases indicate a restoration of D2's τ_e . A similar observation applies to E1's τ_e , the t_e distribution remeasured after leaving the device unbiased for three days (open square) can be seen to have shifted back and almost coincides with that of the pristine state (filled square). E1's τ_e was reduced from 150 ms after the $V_d = -2$ V stress (Figure 4-14(a)(i)) back to 10 ms after a three-day rest (Figure 4-14(a)(ii)). This is also accompanied by a restoration of the trap's occupancy rate, indicating that τ_c has also recovered. A marked reduction in D3's SR, from the previous over 90% to ~8% is also evident. Since the emission times fall in a similar range as those before the rest, this drastic decrease in SR may be ascribed to an increase in the trap's capture time. More importantly, the occurrence of both D2 and D3 emissions in numerous $|\Delta I_d|$ relaxation traces should be highlighted. Prior to this observation, D2 and D3 might have looked like they were the same trap. When D2 disappeared after the $V_d = -2$ V CHH stress, D3 happened to appear with a similar emission time distribution but a smaller $|\Delta I_d|$ recovery step height, which could be ascribed to I_d degradation like in the case of D1. However, brief arguments against this

Figure 4-17. Seven $|\Delta I_d|$ relaxation traces out of a total of a hundred recorded after the two-month rest, each showing both D2 and D3 emissions (marked by cyan and magenta arrows, respectively) within the same trace.

assumption based on the observation of both D2 and D3 emissions in some relaxation traces. To further check this inference, since the current step of D3's emissions is close to the measurement noise using a computer program, we manually inspected every one of the 100 measured $|\Delta I_d|$ recovery traces obtained after the two-month rest. It was found that D3 recorded eight emissions, with seven of them happening in the same traces where D2's emissions were also seen. The eight traces are shown in Figure 4-17, with cyan and magenta arrows marking D3's and D2's emissions, respectively; these emission events are also updated in the Figure 4-15(b)(ii). Thus, the multiple instances of both D2 and D3 emitting in the same $|\Delta I_d|$ relaxation trace tend to rule out the scenario that D2 and D3 are the same trap.

It is interesting to also point out that reapplying the same CHH stress after the rest period could increase the recovered τ_e of the previously affected traps again. This is evident for E1 (Figure 4-14(a)(ii)), albeit the τ_e increments are different from those in the first round of CHH stresses (Figure 4-14(a)(i)). Applying the CHH stress a second

time again shifts E1's t_e distribution towards a longer time and decreases the occupancy rate. E1's τ_e was increased from 10 ms (before 2nd CHH) to 44 ms after the $V_d = -1.5$ V CHH stress and was further increased to 54 ms after the $V_d = -2$ V stress (with V_g fixed at -1.0 V in both cases). A similar restoration back to the pristine state can be seen after the device was left unbiased for one month. The t_e distribution remeasured after leaving the device unbiased for one-month (filled orange ball) can be seen to have shifted back and almost coincides with that of the pristine state (filled square).

Furthermore, it has been shown that CHH stress generates additional $\text{SiO}_x\text{N}_y/\text{Si}$ interface states [98], [139], [140] which increase V_t and degrade I_d , as validated in Figure 4-14(a)(iii). There is a need to clarify whether these changes have affected E1's τ_c and τ_e , thus we compare the changes in the time constants to in the transfer curves. After the first and second round of CHH stress, a near-complete restoration of E1's τ_c and τ_e is seen after three days and one month, respectively as noted above. However, there is no noticeable recovery in the transfer curve for both cases can be detected. From these observations, it would be reasonable to exclude the role of V_t and I_d shifts on E1's τ_c and τ_e .

Moreover, previous simulation results [123] have demonstrated that any variation occurred in the local inversion charge density may affect a trap's switching behavior. This can be reflected by a change in the I_d recovery step height upon emission of a charged carrier. In our experiments, it is observed that some active traps display visible change in switching activity after the CHH stress, while the average I_d recovery steps associated with these affected traps remains almost identical before and after the CHH stress (Figures 4-9 and 4-11). This suggests that the possible effect of local inversion electrostatic variation on the trap's capture or emission characteristics may be ruled out.

In addition, fluctuations of the I_d recovery step of an active trap with or without alteration in switching activity (e.g., trap E1 in Figure 4-12) are detected at operating condition after subjected to CHH stress. In details, we compared the capture and emission time as well as the $|\Delta I_d|$ recovery step height of cluster ‘E1’ obtained during before-CHH stress monitoring step and after one-month storage re-monitoring step (Figure 4-12(a) and (d)). The capture and emission times of these two clusters extracted are found to be almost the same. However, the $|\Delta I_d|$ recovery step height of cluster X obtained during the monitoring step after one-month recovery is higher than that of cluster ‘E1’ obtained during the before- and after-CHH stress monitoring step (Figure 4-12). Remarkably, the capture and emission behaviors and the related $|\Delta I_d|$ recovery step height of individual traps are stochastic events [86], [141]. Thus, it is deduced that the possibility of different defects have the same capture and emission time in an identical device is extremely low. Due to the same capture time constants as well as similar emission times distribution for clusters ‘E1’ and X (Figure 4-12), it is quite unlikely these two clusters originate from different active traps although the $|\Delta I_d|$ recovery step heights are different.

In the period of one-month restorage, a changed electrostatic (and/or a modified percolation path) may happen due to the recovery effects of additional damage triggered by hot-holes (interface states or deep-level traps). And also, other traps, which are near the trap ‘E1’ or share the same percolation path with trap ‘E1’, may be activated and charged during the CHH stress step. This induces variations or reorganization of the percolation path upon emission a trapped charge, and hence, increasing or reducing the I_d degradation contribution of trap ‘E1’ at the operating condition [123]. As a result, multiple observations indicate the clusters ‘E1’ and X most likely related to the same

defect, called trap E1, which exhibits different $|\Delta I_d|$ recovery step height after the application of CHH stress.

Resting the device at room temperature after the second round of CHH stresses results in τ_e recovering back to the pre-stress state again. Reapplying the CHH stress also changes D2's and D3's behavior along similar trends as seen previously, i.e., a marked decrease and increase in SR, respectively (Figure 4-18) compared to the case after the two-month rest (Figure 4-15(b)). However, D2 did not completely “disappear”

Figure 4-18. (a) Another example of both D2 and D3 emissions occurring in the same $|\Delta I_d|$ recovery trace measured in cycle 1 of the monitoring step performed after reapplication of the $V_d = -2.5$ V CHH stress following the two-month rest. (b) TSM taken after reapplication of the $V_d = -2.5$ V stress showing a marked decrease and increase in D2's and D3's SR, respectively.

from the TSM, although two out of the seven emissions clearly show increased emission times not seen before the stress was reapplied. These qualitatively similar changes after reapplying the CHH stress confirm the presence of some definite mechanisms behind the observed changes although the changes differ quantitatively each time due to the stochastic nature of the underlying atomic processes. For instance, the release of H at the interface by applying different stress bias conditions possibly triggers additional defects/re-structuring of defects, however, in a more stochastic way.

4.4.3 Statistical Variations

Although the small-area devices tested are on the same wafer with closely matched pristine I – V curves, significant variations in post-stress behavior are observed. These latent variations revealed by electrical stress may pose a greater challenge to circuit operation than those time-zero variations arising from processing, which could be mitigated via variation-aware circuit design strategies [142]. For instance, D2's τ_e was not but those of E1 and F3 were clearly affected after the $V_d = -1.5$ V stress. D2's τ_e was significantly increased after the $V_d = -2$ V stress but the impact on E1 and F3 was incremental. Reapplication of the CHH stress after recovery also did not lead to similar changes. As shown later, a trap's response to CHH stress is also not correlated to its position along the channel, i.e., traps located near the drain are not necessarily more affected than those near the source (where the CHH effect is less severe). These observations suggest the nature of the defect as a key factor determining how it would respond to the CHH stress.

In addition, varying I_d degradations are observed despite closely matched pristine I – V curves. The device in Figure 4-7 and Figure 4-8 showed no I_d degradation whereas that in Figure 4-14(a) showed a substantial degradation of more than 10% after the $V_d = -1.5$ V stress (Figure 4-14(a)(iii)). For cases where a relatively large degradation occurs, we believe another effect—random channel dopant fluctuation—may be at play, which results in percolation current paths in the channel region of small-area devices [83], [143]. And thus, an obvious discrepancy in I_d degradation between different devices under the same CHH stress is observed. As shown in recent studies [143], this effect gives rise to “bottlenecks” in channel conduction which enlarge the spread in the parametric distribution of a small-area transistor. Stress-generated trap situated near or

directly above these percolation paths could induce significant change in I_d when they are charged, as schematically illustrated in Figure 4-19 [143].

It should also note that the I_d recovery behavior of the small-area device used in this study is different from the large-area counterpart, which is typically characterized by an initial fast recovery phase followed by a slower phase that tends towards quasi-

Figure 4-19. A charged oxide or interface trap in the vicinity of a current bottleneck can induce a large change in drain current through an electrostatic effect [143].

saturation [66]. This recovery behavior may be ascribed to the effect of a large ensemble of traps with a broad emission time distribution [83]. Traps with short emission times dominate the initial fast recovery. Once this phase is over, subsequent recovery is controlled by traps having much longer emission times [83]. Since the I_d perturbation arising a trap's occupancy change is small in a large area device, recovery progresses gradually over an extended period without any abrupt step changes in I_d . On the other hand, the recovery of a small-area device is determined by a few traps. The traps' energy levels in the oxide bandgap (which determine the emission times) and their positions in the channel (which determine the recovery step-sizes) strongly influence the recovery evolution.

4.4.4 Trap's Position Along the Channel

Given the localized nature of the channel hot-carrier effect, a relevant question would be whether a trap's position along the channel would determine the extent by which its switching behavior is altered by the stress. To probe this, we extracted the traps' positions along the channel by the approach described in [132] and correlate this information to the traps' SR measured as a function of V_d . The relative lateral position X_T/L (0 is source side and 1 is drain side) of the traps is extracted by exploiting the recovery drain bias dependence of the ΔI_d step heights (see Figure 4-9(a)), during which the gate recovery bias keeps constant at -0.6 V. Referenced from Illarionov *et al.*'s simulation result [145] in consideration with different random dopant configurations, this dependence can be linearly fitted in a simplified way. The X_T/L is expressed with the following formula [145]:

$$\frac{X_T}{L} = \frac{1}{2} - \text{sign}(P_1) \sqrt{2\alpha^2 \log(P_{0\max} / P_0)} \quad (\text{Eqn. 4-3})$$

where P_1 is the slope and P_0 is the intercept of the linear fitting with Y-axis, $P_{0\max}$ is the largest step height observed corresponding to a trap in the middle of the channel, which is Y1. In our calculations, the value of α is approximately 0.17 based on the simulation outcomes in [145]. The extraction results of traps' lateral position can be found in Figure 4-20. Measurement data and linear fittings are labeled with the traps' name and the extracted relative lateral positions X_T/L , for example, Y1/0.55L. The inset of Figure 4-22(a) depicts the trap distribution in the oxide schematically.

Figure 4-20 illustrates an example in detail, where the traps' location is extracted by application of above-discussed method. In this device, there were three active switching traps, Y1, Y2 and Y3 situated near the middle of the channel, close to the drain side and close to the source side, respectively (Figure 4-20(a)). The traps' SRs

were stable even after repeating the monitoring steps many times (Figure 4-20(b)). This may be explained by the low V_g bias at operating condition; therefore, any impact on traps' switching activity caused by monitoring V_g bias may happen beyond the

Figure 4-20. (a) $|\Delta I_d|$ recovery step versus applied drain bias V_d from which the positions of three active switching traps along the channel are extracted and expressed as fractions of the channel length L ($L = 0$ corresponds to the source end). (b) TSM of the three traps ($@ V_d = -0.1$ V) with the individual switching rates (SR's) as indicated. (c) SR of the three traps as a function of V_d , taken during the first V_d sweep from -0.1 V to -1.5 V.

experimental time frame. Figure 4-20(c) depicts the traps' SRs as V_d was varied from -0.1 V to -1.5 V with V_g fixed at -1 V. The relative independence of Y1's SR on V_d is consistent with its mid-channel position, as the moderate V_d would not significantly reduce the oxide field along the channel. While the variation of Y2's SR with V_d is expected by virtue of its near-drain position, the non-negligible variation of Y3's SR is, however, surprising due to its near-source position. More interestingly, a repeat of the measurements with V_d scanned across the same range yielded similar results for Y1's and Y2's SR, but a marked decrease in Y3's SR as compared to the first measurement (Figure 4-21(a) and (b)). In subsequent measurements, Y3's SR remained around 0 or 1%, implying that its switching behavior had been affected by the applied V_d in the first

Figure 4-21. (a) Remeasured SR of the three traps as a function of V_d remeasured showing a marked change in trap Y3's SR. (b) TSM retaken at $V_d = -0.1$ V showing the absence of Y3 cluster. (c) TSM taken subsequently at 150 °C showing the shift of emission times to shorter time scale and the appearance of the Y3 cluster. (d) SR as a function of V_d taken with the source (S) and drain (D) connections interchanged. (e) SR as a function of V_d taken subsequently with S and D interchanged back.

measurement. To check for a possible increase in Y3's τ_e beyond the relaxation time window, the monitoring step was performed at a higher temperature of 150 °C. Since trapped charge emission is thermally activated, a higher temperature would result in a smaller τ_e , allowing Y3's emission to be again observed. The TSM in Figure 4-21(c) indeed shows the reappearance of Y3's cluster, confirming that Y3's τ_e was increased by the V_d biasing. It should be noted that Y2's cluster is shifted to a shorter time range and Y1's cluster is not seen as it has shifted below the lower time limit.

We further studied Y2's SR variation by exchanging the source and drain connections, which placed Y2 near the source side. After the exchange, Y2's SR became

Figure 4-22. (a) Dependence of Y2's SR on V_g , giving evidence of a threshold oxide-voltage for Y2 activation which is about -0.9 V. The inset is lateral trap distribution, the schematic sketch of the relative positions of the three characterized traps within the oxide. (b) A schematic cross-sectional diagram of an p-MOSFET, depicting two general hot-carrier (red filled circle) trajectories. It also shows the associated secondary impact ionization events, yielding energetic tertiary holes (open circles) at depths (lower) and (deeper) along trajectory 2.

relatively independent of V_d , like the case of Y1, shown in Figure 4-21(d). Interchanging back the source and drain connections produced a similar decreasing SR trend as seen before (Figure 4-21(e)). In addition, a V_g bias dependence of Y2's SR is also measured in Figure 4-22(a), giving evidence of a threshold oxide-voltage for Y2 activation which is about -0.9 V. Thus, these observations imply that Y2's SR variation is a result of the reduction in oxide field near the drain as $|V_d|$ is increased. The reduced oxide field lowers the probability that Y2 is charged during the capture interval [131].

In summary, while the results of Y1 and Y2 with respect to V_d biasing could be explained in terms of their positions along the channel, the apparent increase of Y3's τ_e is surprising, given the trap's near-source position. Since the source region is subjected to the BTI effect during V_d stress, the possibility of Y3 being affected through this mechanism needs to be considered [95], [126]. However, experimental results from the

monitoring step (@ $V_g = -1$ V, $V_d = -0.1$ V), even after repeated measurements, have shown that traps' SRs were stable before V_d stress (e.g., Figure 4-5 and Figure 4-20(b)). This may be attributed to the low V_g used; therefore, any impact on traps' emission may happen beyond the experimental timeframe. On the other hand, an obvious change in Y3's emission behavior is observed after a high V_d bias (with V_g maintained at a low bias) even if it is near the source side. The above leads us to hypothesize that Y3's nature (since traps close to the drain may not always be affected), coupled with non-local mechanisms related to the CHH effect (e.g., Figure 4-22(b)), may have caused the change in its emission behavior.

4.4.5 Possible Physical Mechanism(s)

A pertinent question at this stage would be why the τ_e 's of some traps are altered by the CHH stress (e.g., D2, D3, E1, F1) while other traps (e.g., D1) are not affected. The exact physical mechanism(s) may require some time to decipher since there is a lack of experimental methods that could directly probe the chemical nature of the affected traps. However, due to the importance of oxide defects on the operation of MOSFETs, intensive studies made in the past have resulted in some widely accepted theories for the CHC degradation mechanism as well as the atomic origins of the oxide and interface traps. It would thus be useful to propose possible explanations for the experimental results obtained in our study.

Detailed studies on CHC stress induced MOSFET degradation have consistently pointed to the dissociation of interfacial Si-H bonds by energetic carriers as the main reason for interface state generation [140], [146], [147]. The rate of interface trap generation is dependent on several key factors, namely carrier energy and the rate of hydrogen migration away from the broken bond sites. CHC stress tests involving

devices annealed in deuterium, a heavy hydrogen, saw an obvious slowing down of the interface trap generation rate [146]. This provides a strong support for the proposed mechanism.

On the other hand, there is a consensus that oxygen vacancy (V_O) defects ($O_3\equiv Si$ $Si\equiv O_3$) or E' centers are a major source of hole traps in the SiO_2 gate oxide [148-150]. Several first-principles simulation studies have shown that the properties of the V_O defect could be changed by hydrogen species migrating through the oxide network [98], [139], [151]. When the hydrogen species are negatively charged, they tend to bond with Si atoms. However, hydrogen species tend to bond with the oxygen atoms when the former are positively charged [152]. Such bonding of hydrogen to the SiO_2 lattice atoms could cause large shifts in the trap level and thus change the trap's emission behavior [153].

With the above considerations, we outline possible hypotheses for the observed τ_e changes following the CHH stress. During the CHH stress, hot holes dissociate the interfacial Si-H bonds. The liberated hydrogen species migrate through the oxide and change the traps' properties by re-bonding with the lattice silicon/oxygen atoms. From *ab-initio* simulation, Goes *et al.* [154] showed that the hole-trap level of a V_O defect in SiO_2 is shifted towards the conduction band edge when a hydrogen atom is attached to it, making the defect a deeper hole trap with a longer emission time. Since the incorporation of nitrogen in SiO_2 increases the density of hole traps [155], it would be relevant to also consider the interaction between hydrogen and nitrogenated V_O defect or k_N center [156], where an oxygen atom is substituted by a nitrogen atom (i.e., O_2NSi $Si\equiv O_3$). From atomistic simulation, Gu *et al.* [157] found that a k_N center with a two-fold coordinated nitrogen gave rise to a deep-level hole trap in the SiO_2 bandgap. However, when a hydrogen is attached to the nitrogen atom, converting the latter to the

Figure 4-23. (a) Schematic energy band diagram of the structure under positively charged condition. The release and restoration of H atom induces transformation in defect state. (b) A three-states defect model with a N atom in vacancy-interstitial (V_o-O_i) defect. H atoms play important role in hot carrier degradation. The transition between different defect states can be observed. (c) The possible atomic structure of shallow hole trap and deep hole trap.

three-fold coordination, the resultant hole trap is much shallower. Thus, a shallow hole trap could be changed into a deeper one when the hydrogen atom is detached, via interaction with another nearby hydrogen atom released by the CHH stress. Conversely, when a hydrogen atom from the CHH stress bonds with a two-fold coordinated nitrogen of the k_N center, the hole-trap level is changed from deep to shallow. Such interactions should be dictated by the existence of energetically favorable pathways, which in turn depend on the local defect chemistries, and may account for the observation that not all traps are affected. The affected traps may revert to the pre-stress state when further changes to the local hydrogen bonding arise from phononic perturbations in the oxide

Figure 4-24. The initial position of H was placed in four different locations near the N of *Vo-Oi* defect with 2-fold coordinated N. This is to account for the diffusion path of H atom into the two-fold coordinated N defect. The MEP and CTL of trap state transition from *S1* to *S3* for each case show relatively deep CTL (*S3*, e.g., CTL=5.589 eV), considering it is at the meta-stable state.

network during the rest period. The associated changes in trap levels may be checked by a detailed voltage bias and temperature dependence study in a future work.

Besides defect structural change arising from interaction with hydrogen species, we further point out that localized heating at the drain could also lead to non-local hot-carrier injection through secondary heating pathways, e.g., further heating of energetic carriers that made up the substrate current [158]. An earlier work [159] has shown that source-side gate oxide degradation became worse when a reverse substrate bias was applied, corroborating the proposed role of substrate current heating. Hence, the

Figure 4-25. The initial position of H atom was placed in four different locations near the N of V_o-O_i defect with 2-fold coordinated N considering the diffusion path of H atom into the two-fold coordinated N defect. The MEP and CTL of trap state transition from $S1$ to $S2$ for each case.

consideration of non-local hot-carrier injection, and the resultant defect structural change arising from electron-phonon interaction [160], may also be relevant in understanding why the τ_e of a source-side trap situated can be affected by the CHH stress (cf. Figure 4-22(b)).

Three-states defect model with nitrogen atom in vacancy-interstitial (V_o-O_i) defect is proposed. A dynamic structural relaxation method, based on density functional theory, is implemented via the first-principles simulation [161], [162]. A hydrogen atom is attached to the neighboring lattice void of (V_o-O_i) defect in the model and then subjected to structural relaxation under positively charged condition (+1 charge state) until it

reaches a new stable state. Thereafter, a nudged elastic band (NEB) method is applied to search for the transition structures and the corresponding minimum energy path (MEP); the corresponding charge transition levels (CTLs) are calculated when the initial defect state transitions into a stable network state [163].

As shown in Figure 4-23, the defective *Vo-Oi* structure with 2-fold coordinated nitrogen atom (N) results in a trapped-hole site with a deep CTL in the SiON bandgap. This state is named *S1* with typical CTL at around +6 eV. In contrast, the 3-fold coordinated N with dangling nitrogen bond terminated by a hydrogen atom (H) result in a trapped-hole site with a much shallower CTL. This state is called *S2* with a typical CTL at around +2.5 eV. Both defect states were obtained through structural relaxation. The H atom plays a paramount role in the variations of emission behavior in an active trap after CHH stress.

To diffuse a H atom into the two-fold coordinated N defect through different paths, the initial position of H atom was placed in four different locations (noted as H1, H2, H3, and H4) near the N of *Vo-Oi* defect with 2-fold coordinated N. After the relaxation of atomic structure under +1, the structure would change but still quite similar to the defective structure of *Vo-Oi* defect with 2-fold coordinated N. This state has a relatively deep CTL (e.g., CTL=5.589 eV), considering it is in metastable state (*S3*). The calculation of the MEP between *S1* and *S3* for each case is shown in Figure 4-24. The results show that no barriers are found in 3 out of 4 cases. However, a relatively high barrier is found in case 4. The atomic structure shows that the positions of Si, O, and N atoms have not changed significantly. Therefore, the barrier was caused by the change of H position. For the first three cases, no barrier was found, indicating that the initial position of H atom is in stable state. In contrast, the H atom in case 4 needs to overcome barrier to reach the final stable structure.

Figure 4-26. (Top) Atomic structure of $Vo-Oi$ defect network. An obvious atomic structural change due to the release or restoration of H atom during structural relaxation is observed. (Bottom) Corresponding MEP curve and CTL of the structures shown above. The energy barrier during structural transition from deep hole trap state to a stable shallow trap state is around ~0.6 eV. A change in the structure will induce a drop in CTL.

After inserting a H atom into N, not all of the $Vo-Oi$ defects with 2-fold coordinated N remain defective as $S1$ with deep CTL. Some of them straightforwardly relax to $S2$ with a shallow CTL. The MEP and CTL of the four cases are shown in Figure 4-25. When a severe atomic structural relaxation occurs (i.e., significant change in the position of Si and O), multi-barriers are observed in the trap state during transition from $S1$ to $S2$. Generally, in the case of deep-level to deep-level trap state transition, the Si and O atoms positions in the supercell would not change significantly. Figure 4-26 depicts an atomic structure change in $Vo-Oi$ defect network due to the release or restoration of H atom through structural relaxation. When a defect re-attaches with a H atom, the energy

barrier due to the structural transition from deep hole trap state to a stable shallow trap state is around ~ 0.6 eV. A change in the structure induces a huge drop in CTL.

4.5 Chapter Summary

In this study, it is shown that the τ_e of trapped holes in the SiO₂ gate oxide could be changed by a CHH stress. During the post-stress rest period, the affected traps may revert to their pre-stress behaviors and be affected again when the CHH stress is reapplied, although the outcome from the re-stress generally differs quantitatively from the preceding cases due to the stochastic nature of the underlying processes. It is proposed that hydrogen species released from the CHH stress may interact with the trap sites and thus change the energy levels and the resultant emission times. Further defect structural changes to that occur during the post-stress rest period may account for the reversion to pre-stress behaviors. Non-local hot-carrier injection arising from secondary heating pathways could explain the observation that traps situated near the source side are also affected. The findings of this study may impact aging models and reliability-aware circuit designs, as well as attempts that made use of the stochastic capture/emission processes for implementing physically unclonable functions, where stress-invariant trap properties are an important premise.

CHAPTER 5: RESPONSE OF SWITCHING HOLE TRAPS IN THE SMALL-AREA P-MOSFET UNDER CHANNEL HOT-HOLE EFFECT

5.1 Introduction

As discussion explicated previously, bias-temperature instability (BTI) and hot carrier degradation (HCD) are among most crucial reliability challenges that severely degrade the device performance and thereby impact the static logic CMOS circuits. To this end, the nature and behavior of oxide traps that contribute to the dynamics of BTI and HCD should be properly understood and modeled. Meanwhile, as the device dimension scales down, considerable attention has been paid to the single oxide traps since it can significantly impact the performance of nanoscale MOSFETs through the random telegraphic noise generated in the drain current (I_d).

To date, much understanding on the characteristics of discrete switching oxide traps (SOTs) that in a small-area MOSFET has been obtained from the perspective of BTI or gate stressing (V_g) studies [86], [88], [93], [126], [127]. On the other hand, comparable to BTI effect, the hot carrier injection (HCI) degradation is also a serious concern [129], [164]. Consequently, a further consideration on the drain stressing (V_d) is crucial since it is more relevant to actual mixed analog-digital circuit condition. Based on extensive theoretical and experimental studies on the aforementioned two degradation mechanisms, physical models to describe them independently have already been introduced in condition of an idealized setting. However, in realistic circuits scenario,

Figure 5-1. A schematic of stress conditions region which illustrates different stress conditions, including NBTI, CHH, and mixed NBTI and CHH. The region of stress conditions describes the boundaries of the two-dimensional parameter space (gate stressing and drain stressing biases) applied in this chapter.

MOSFETs are rarely subjected to exclusive BTI or HCD conditions, and only a limited number of investigations are available with regard to the interaction between BTI and HCI degradation on device performance [117], [132]. Therefore, a comprehensive understanding of the response behaviors of SOTs under the mixed gate and drain stressing is imperative and worthwhile.

In this chapter, it is focused on single oxide defects which make the contribution to the recoverable component of total degradation, i.e., the pre-existing switching oxide traps. A thorough experimental examination of the impact of mixed NBTI and Channel hot-hole (CHH) effect on the SOTs in small-area p-MOSFETs is presented, where the typical stress conditions applied are schematically depicted in Figure 5-1. Different colors separate the bias combinations which would trigger various degradation mechanisms, and the increasing color intensity of different regions implies the increasing impact of the stressing on the device performance degradation.

After further examination, two distinct behaviors of switching hole traps in nanoscale, SiON gated p-MOSFETs subjected to channel hot-hole (CHH) effect are broadly observed and elucidated. First, in all devices tested, the occupancy rate (OR) of a trap could show either increased or decreased as the drain voltage $|V_d|$ is increased towards the gate voltage $|V_g|$. A further increase of $|V_d|$ would eventually lead to a decrease trend of OR. Second, the reduced emission probability is typically reversed upon removing the drain bias but in $\sim 11\%$ of the traps, this is not the case, i.e., the hole-trapping became more permanent, or the trap's emission time is elongated. These observations may be explained in terms of two competing effects: (1) tendency of the trap to retain the trapped hole as revealed by a V_d -induced increase of the traps' emission time, that can be understood as a volatile trap becoming electrically inactive result in a contribution to permanent degradation; (2) increase (decrease) in capture (emission) time what is commonly described by the decreasing oxide field or trap deactivation due to the effective oxide voltage ($\sim V_g - V_d$) falling below the threshold oxide-voltage for switching. Therefore, apart from those characteristics of single oxide traps already revealed through gate biasing studies, these new behaviors arising from drain biasing should also be accounted for in circuit reliability compact modeling.

5.2 Experimental Details

Eight similar p⁺ polysilicon gate p-MOSFETs of a 45-nm commercial technology, each of 60 nm drawn gate length and 120 nm gate width and having a total of 19 active SOTs under the measurement condition used, were probed. The gate oxide is a 1.7-nm thick silicon oxynitride, achieved by decoupled plasma nitridation. In all measurements, the source and bulk electrode were connected to the ground, in order to avoid the introduction of the floating body effect. In addition, the testing temperature is controlled

Figure 5-2. The schematically detailed illustration of measurement sequence used to investigate the response behavior of switching oxide traps that already active in operation condition ($V_{DD} = -1$ V) to a varying drain bias (V_d) and fixed gate bias (V_g) stressing.

by the thermo chuck of probe station. To gain insight into the physic picture of capture and emission processes during and after mixed NBTI/CHH stressing, τ_c (τ_e), the OR and the I_d step height behavior of individual traps for different combinations of stress biasing are fully characterized. Here, each device-under-test (DUT) was subjected to a 10-s mixed NBTI/CHH stressing, followed by a 100-s relaxation (always maintained @ $V_g = -0.6$, $V_d = -0.1$ V) to monitor the recovery component of degradation, with the characterization sequence repeated 100 times. The detailed description of the measurement setting sequence is outlined in Figure 5-2. During the measurements, the 100-cycle test was reapplied several times on the same device, with V_g fixed as before but $|V_d|$ of the mixed NBTI/CHH stress increased each time. Furthermore, since the typical ΔV_t extraction during the recovery sequences needs the method of constant drain

Figure 5-3. Representative example of a defect spectral map, derived by plotting the $|\Delta I_d|$ recovery step versus emission time of all hole-detrapping events in the 100 recovery traces sensed at $V_g = -0.6 \text{ V}$, $V_d = -0.1 \text{ V}$. Left and right are defect spectral map for trap D1 and D2 in device #D measured at fresh condition and remeasured @ $V_g = -1 \text{ V}$, $V_d = -0.1 \text{ V}$, following an earlier series of mixed NBTI/CHH testing which saw V_d being varied from -0.1 V to -1.5 V , respectively.

current (normally near the threshold voltage) to record and subsequently convert into ΔV_t via using the initial transfer (I_d - V_g) characteristics. To reduce the extraction complexity and ensure the experimental data accuracy, and therefore, we apply an easier approach through plotting of the I_d recovery step (due to the emission of a hole trapped during the mixed stress) versus emission time (Figure 5-3(a), (b)), helping identify SOTs that are active under a given test condition intuitively [86]. Take a representative device #D as an example, one can clearly see two traps were utterly characterized, called D1 and D2, where each single trap individually identified by a cluster in above spectral map. The percentages followed by the trap number label give the occupancy rates for these two traps activated @ $V_g = -1 \text{ V}$, $V_d = -0.1 \text{ V}$.

5.3 Results and Discussion

Each trap causes I_d recovery steps with a particular step height at a specific mean value of emission time. By assigning the unique I_d steps to a trap (implied by dashed line ellipses in Figure 5-3), the parameters such as trap characteristic τ_c (τ_e) and the OR can be extracted as described in chapter 4. In the following, typical variation of the occupancy rate (OR) of two SOTs D1 and D2 in DUT #D as a function of V_d bias are shown in Figure 5-4(a). OR is the percentage of emission events observed in the 100 recovery steps, giving the likelihood for the net capture of a hole by a SOT in the CHH interval. It is remarkable to see from Figure 5-4(a), D1's OR first increases and then decreases with the increase of $|V_d|$. On the other hand, D2's OR decreases steadily with the increase of $|V_d|$. In order to scrutinize that the V_g bias of -1 V applied during mixed stress did not affect the traps' OR, re-measurement was made after the various CHH stress at the same V_g bias but with V_d set back to -0.1 V. In light of comparison results before and after mixed stressing, as exemplified in Figure 5-3(a), (b), one can see almost unchanged OR's of the two traps, confirming the proposition that the OR fluctuations exposed in Figure 5-4(a) are due to V_d stressing variation and not the constantly applied $V_g = -1$ V during 10-s mixed NBTI/CHH stress interval. The result is that such representative V_d -induced fluctuations of SOTs' OR can be observed in all DUTs that share the same attributes.

It is commonly accepted that a trap can capture a charge carrier at a specific capture time when shifting its energy level above the E_f . Moreover, it can emit it at a particular emission time when moving the energy level below the E_f again. Such transitions occur by charge trapping or detrapping have been accurately described by non-radiative multiphonon process (NMP) theory [90]. For example, by decreasing V_g bias the probability that a trap captures a charge carrier during stress interval would reduce, and

Figure 5-4. (a) Representative examples of the two kinds of V_d -induced occupancy-rate variation trends seen across all SOTs tested under mixed NBTI/CHH stressing. (b) Occupancy rate (OR) versus V_d for trap D1 and D2 in device #D which remeasured at 150 °C ((a) measurement under 100 °C)

thus the OR reduces while τ_c (τ_e) increases (decreases). Increasing $|V_d|$ leads to a decrease of the oxide field of inhomogeneous E_{ox} along the channel additionally to a heating of holes in the channel. To explain this complex response, the variation of OR is considered to have originated from the interaction between these two effects. Specifically, SOTs are able to capture and emit a hole at certain characteristic capture and emission times [86], [98]. As a result, to distinguish the impact of hot holes and oxide-field reduction on the SOTs' OR, the capture (τ_c) and emission (τ_e) time constants of D1 and D2 are extracted as a function of V_d (Figure 5-5). For D2, τ_c (τ_e) increases (decreases) with increase in $|V_d|$. These changes are typical of a decreasing oxide field and reduce the likelihood of a net hole-capture in the CHH interval [93], [120]. Since the decreasing oxide electrical field would reduce (increase) the probability that a trap capture (emits) a charge carrier during stress. The steady decrease in D2's OR (Figure 5-4(a)) may thus be ascribed to the oxide-field effect being more important than the hot-

Figure 5-5. Obtain capture (τ_c) and emission (τ_e) time constants through the time dependent defect spectroscopy measurements of trap D1 (left) and D2 (right) as a function of V_d bias. The V_g is fixed at constant -1 V and $T = 100$ °C. τ_c (τ_e) extracted according to the description illustrated in chapter 4 behave increase (decrease) when only considering inhomogeneous E_{ox} induced by the increasing V_d bias.

hole effect. On the other hand, D1's τ_e first increases with $|V_d|$ and then decreases with further $|V_d|$ increase, while τ_c increases steadily (like that of D2). Therefore, the initial increase of D1's OR (Figure 5-4(a)) is due to D1 being able to trap the hole for a longer time as $|V_d|$ is increased, i.e., hot-hole effect enhances the likelihood of a net hole capture despite the decreasing oxide field. In addition, the eventual decrease in OR may be attributed to the increasing importance of the oxide-field effect. As a further evidential support for the significance of hot-hole effect, the measurements shown in Figure 5-4(a) were repeated at 150 °C (Figure 5-4(b)). Since drain-induced hole heating would be reduced at a high temperature, according to the early result, the initial increase of D1's OR should be actually affected [165]. Indeed, the initial increase is suppressed; the OR now decreases steadily with increase of $|V_d|$, like that of D2. These observations indicate

Figure 5-6. DUT #D measured under mixed NBTI/CHH at constant $V_g = -1.5$ V, with V_d varying. A newly activated trap D3 has a typically similar V_d dependence of its OR as D1. The τ_e and τ_c for two V_d values are labeled, clearly showing the larger increase of τ_e with $|V_d|$ as the reason for the rising D3's OR even though τ_c is simultaneously increase slightly (similar to D1 (Figure 5-5(a), left)).

that the two competing mechanisms: oxide-field effect and hot-hole effect are responsible for the behavior of D1 under CHH.

Measurements taken at a more negative $V_g = -1.5$ V yielded similar variations of D1's and D2's OR with $|V_d|$ increase (Figure 5-6). This variation of OR appears consistent with the trend portrayed in Figure 5-4(a). However, unlike in Figure 5-4(a), decrease in D1's OR is not seen. This is because with a more negative V_g , the impact of a decreasing oxide field is "postponed" to more negative V_d values (not measured to avoid severe degradation). It can be observed that a new SOT D3 is activated under conditions of negative $V_g = -1.5$ V with V_d stressing and its OR's dependence on $|V_d|$ is like that of D1. Furthermore, the τ_e and τ_c for two V_d values of D3 are also indicated in Figure 5-6, showing the considerable increase of τ_e with $|V_d|$ as the reason for the rising OR (similar to D1 (Figure 5-5(a), left)). Notably, similar results are obtained on another

Figure 5-7. Arrhenius plot for the τ_c of D1 and D2. Based on NMP theory, inset is the simplistic two-dimensional potential energy surface of the trap along the reaction coordinates between neutral and charged state. Green parabola represents the neutral state, another red one acts the charged state. It indicates charge trapping process has to overcome a certain capture barrier (E_{AE}) for the transition from neutral to charged state.

DUT X (Figure 5-8). Here, the response of SOT X2 and X3 to a varying V_d are similar to those of D1 and D2, respectively.

Hole heating varies along the channel and thus one would expect the trap's lateral position to impact its OR's dependence on V_d . To inspect this inference, intentionally interchanging the source and drain i.e., V_d was applied to the original source terminal and the original drain terminal was grounded, however, yielded similar dependence of the SOTs' OR on $|V_d|$ in some cases (e.g., DUT X, before and after interchanging the source and drain, shown in Figure 5-8 and Figure 5-10(a), respectively). This implies that besides the trap's position, its characteristic nature could be another important factor. As a detailed factorial study is quite impossible, a lumped parameter that describes a SOT's overall response to a changing V_d would be useful. In view of this, we extracted the activation energy of hole capture for D1 and D2, at condition of $V_g = V_d = -1$ V

Figure 5-8. Three active SOTs (X1, X2 and X3) are exposed in another DUT #X, showing a similar traps' OR dependence on V_d , as those in DUT #D (cf. Figure 5-4(a)).

(Figure 5-7) [86], [166]. Considering the observation from the experimental results, it is intriguing to note that D1's activation energy (E_{AE}) for the vibronic transition of charge capture is about half that of D2, clearly showing a smaller hole-trapping barrier in the former. Such a transition process was described accurately by application of the simplistic 2D energy-coordinate schematic inset in Figure 5-7 based on the NMP theory. Then, this indicates that the trapping activity of a hot hole in the CHH interval may be more likely to occur in D1 than in D2, leading to the increase in D1's OR with $|V_d|$ or the CHH effect (Figure 5-4(a)).

On top of the afore cases, it is also observed that the switching activity of a SOT at operating gate condition ($V_g = -1$ V) could totally disappear after CHH stressing. SOT X1 of DUT #X shows a representative example. Initially, X1's OR (@ $V_d = -0.1$ V; $V_g = -1$ V) is 60% (Figure 5-9(a)(i)). During CHH stress @ $V_d = V_g = -1.5$ V, a marked extent of reduction of X1's OR to 6% is observed (Figure 5-9(a)(ii)). Remeasurement @ $V_d = -0.1$ V; $V_g = -1$ V yielded 0% for X1's OR (Figure 5-9(a)(iii)), indicating that X1 is no longer active in the measurement window or fail to discharge the trapped hole. To probe this drastic change in X1's behaviour and explicate the physics behind it,

Figure 5-9. (a)(i) Defect spectral map of DUT #X measured @ $V_g = -1$ V, $V_d = -0.1$ V under fresh state, with the ORs indicated. (a)(ii) During Mixed NBTI/CHH stressing, defect spectral map for X1, X2 and X3 measured at condition of $V_g = V_d = -1.5$ V. A significant drop in the OR of X1 can be observed. (a)(iii) Defect spectral map measured @ $V_g = -1$ V, $V_d = -0.1$ V, after the measurement @ $V_g = V_d = -1.5$ V. It shows no emission is detected for X1, while the OR of X2 and X3 are similar to those before the high $|V_d|$ biasing was applied (b) After 10 s @ $V_g = V_d = -1.5$ V, DUT #X was relaxed @ $V_g = -0.6$ V, $V_d = -0.1$ V for 100 s; this sequence was repeated 100 times. $|\Delta I_d|^{eoc}$ refers to remnant (or unrecovered) I_d degradation at the end of the relaxation interval. The light orange bands marked the cycles in which X1 captured a hole during the CHH stress and emitted it during relaxation; the dark bands denote otherwise. The light red circles show I_d degradation or $|\Delta I_d|$ arising from X1 capturing a hole, given by the recovery in I_d observed when the captured hole was emitted in the relaxation interval. Upward arrows indicate the generation of other permanent trapped holes after X1 became permanently charged at cycle 13.

$|\Delta I_d|^{eoc}$, i.e., I_d degradation at the end of each relaxation interval of the CHH-relax test was scrutinized, as exemplified in Figure 5-9(b). $|\Delta I_d|^{eoc}$ provides an indication whether a SOT could emit the trapped hole in a relaxation interval. Any failure to emit would

Figure 5-10. (a) Dependence of X2's and X3's OR on V_d , with the source and drain connections interchanged during measurement, i.e., V_d was applied to the original source terminal and the original drain terminal was grounded (cf. Figure 5-8, Figure 5-9(a)(i)) (b) Dependence of X2's and X3's OR on V_g , giving evidence of a threshold oxide-voltage for X3 activation. The threshold for X2 is more positive than -0.8 V and cannot be measured since the relaxation $V_g = -0.6$ V. Use of more positive relaxation V_g compromises current measurement speed. Inset shows that X2 and X3 might be situated close to each other.

lead to increase of $|\Delta I_d|^{eor}$ relative to the previous cycle. For cycles in which X1 was still active @ $V_d = V_g = -1.5$ V (red hollow circles, Figure 5-9(b)), $|\Delta I_d|^{eor}$ remained constant, i.e., all three SOTs could emit during relaxation. From cycle 12 to 13, an increase of $|\Delta I_d|^{eor}$ by ~ 0.28 μ A, which corresponds to I_d degradation due to X1 charging, can be observed. This signifies that X1 failed to emit in cycle 12. After cycle 13, no decrease of $|\Delta I_d|^{eor}$ can be observed, indicating that X1 remained charged for the rest of the cycles. This shows that during the CHH stress, X1 was transformed into a more permanent hole trap with emission time far exceeding the measurement window, rather than loss of capture activity, i.e., deactivation of SOTs. Such a fascinating outcome is found in two SOTs. This result is consistent with a previous study which reported an increase in deep-level hole trapping by combined NBTI and CHH stresses [137].

Further studies were conducted to examine other possible causes of the decreasing trend of OR when $|V_d|$ increases to a relatively large value (e.g., D1 in Figure 5-4(a)). It should be noted that V_g -dependence study has revealed a significant change in a trap's OR over a narrow voltage range, suggesting the existence of a "threshold" oxide voltage for switching for individual oxide traps (Figure 5-10(b)). Below this threshold, the trap is not activated, therefore the corresponding OR rapidly drops to zero. This may help explain the rather steep decrease in OR as $|V_d|$ approaches or exceeds $|V_g|$ (e.g., D1, X2), leading to the oxide field falling below the switching activation "threshold". In addition, Figure 5-10(b) shows another interesting observation: there appears to be an anti-correlated change of X2's and X3's OR as a function of $|V_g|$, as indicated, when X3's OR increases, that of X2 decreases and vice versa. This suggests that the two traps are near to each other as illustrated in the inset of Figure 5-10, the capture of a hole by a trap changes the local channel electrostatics, which in turn reduce the likelihood of the capture of a hole by the other trap. The possible mechanism can be further explained with NMP theory [90], in which gate oxide trap's capture characteristic is affected by the interplay of two factors: one is the change in channel carrier energy itself, and the other is the variation in lattice thermal vibration energy that associated with the relaxation of surrounding lattice upon capture a hole. Figure 5-11(b) shows a simplistic configuration coordinate to illustrate a change in the total energy barrier resulted from the coupling effects discussed above. Initially, a relatively small energy barrier is here needed to be overcome for capturing charge. This allows the trap X2 to switch easily from the neutral state to the charged state, leading to a short capture time constant (arrow 1). However, under the influence of the Coulomb repulsion effect caused by the adjacent charged trap, it requires a longer capture time since a larger trapping energy barrier

Figure 5-11. (a) Schematic diagram depicts the impact of the occupation of trap X3 on the potential energy of trap X2. (b) A simplistic 2D potential energy surface to describe the impact of the Coulomb repulsion effect on the energy barrier of nearby traps. Upon trap X3 capture a hole, the energy barrier for trap X2 to transition from neutral to charged state is now increased (i.e., a larger capture barrier E2) owing to the Coulomb repulsion effect, resulting in a larger trap capture time constant.

exists between the neutral state to the charged state (arrow 2). As described in [98], the trap's capture time constant can be modelled as the form of:

$$\frac{1}{\tau_c} = C_p \cdot n_c \cdot \exp\left(-\frac{E_{AE}}{kT}\right) \quad (\text{Eqn. 5-1})$$

where C_p is the prefactor, n_c is the channel carrier concentration, and E_{AE} is the activation energy for the transition involving charge exchange between the substrate and oxide traps. In Figure 5-11(a), polarized charges present in the substrate when trap X3 is occupied by a hole. It is believed that the polarized charge is equivalent to the image charge Q' in the channel, denoted by [167]:

$$Q' = Q \cdot \frac{\epsilon_{SiON} - \epsilon_{Si}}{\epsilon_{SiON} + \epsilon_{Si}} \quad (\text{Eqn. 5-2})$$

where Q is the trap X3 charge, ϵ_{SiON} and ϵ_{Si} are the dielectric constants of SiON and Si substrate, respectively. From Eqn. 5-2, the image-charge Q' has the same polarity as charged trap X3 since ϵ_{SiON} is larger than ϵ_{Si} . These two positive charges supply an

additional potential around trap X2, which results in an increase in activation energy for charge vibronic transition. The $E_{Coulomb}$ is defined as [168]:

$$E_{coulomb} \approx \frac{Q^2}{4\pi\epsilon_{SiON}\gamma_0} \left[1 + \frac{Qt_{inv}C_g(V_g - V_t)}{4kT\epsilon_{SiON}WL} \right]^{-1} \quad (\text{Eqn. 5-3})$$

Among them, γ_0 is the distance between two traps, t_{inv} is the thickness of channel inversion layer, C_g is the gate capacitance, V_t is the threshold voltage, W and L are the channel width and length, respectively. Calculation from the equations above, the E_{AE} of trap X2 increases by the value of $E_{Coulomb} = E_2 - E_1$. As a consequence, the capture time constant of trap X2 will increase apparently when trap X3 is occupied, which can be described with the equation as follow:

$$\frac{1}{\tau_{c_{X_2X_3}}} \approx C_p \cdot n \cdot \exp\left(-\frac{E_1 + E_{Coulomb}}{kT}\right) \quad (\text{Eqn. 5-4})$$

When the probability of capturing a hole by trap X3 increases, the local channel electrostatics will change, causing the oxide field of trap X2 decreases to the value below switching activation “threshold”. Remarkably, in contrast to typical SOTs’ behavior which shows easier capture-activity after being triggered by a successively higher V_g , the X2’ OR presents an abnormal reduction trend with the increasing of V_g . This observation supports the inference that two adjacent oxide traps may affect each other’s capturing probability due to coulomb effect and thereby, leading to a reverse correlation in OR.

In addition, according to the results shown in Figure 5-6, we deduce that the lowering of energy barrier for hole switching in shallow-level defects, allows these oxide defects to respond sensitively to CHH effect during mixed NBTI/CHH stressing due to the lower capture and emission time constants. This proposition is coherent with the trend outlined in Figure 5-12, in which the effective CET maps correlated with

Figure 5-12. Capture emission time (CET) map correlation of traps from several devices which were examined over a period of time. τ_c and τ_e are directly extracted from Eqn. 4-1 and Eqn. 4-2. A comparison between the traps that are sensitive to (a) CHH and (b) E_{ox} change shows that, τ_c and τ_e of the former are much smaller than those of the latter.

experimental results supported by application of Eqn. 4-1 and Eqn. 4-2. The observations indicate that the τ_c (τ_e) of a trap which is susceptible to inhomogeneous E_{ox} is about two orders higher than the other trap which is more vulnerable to CHH effect.

5.4 Chapter Summary

In summary, a detailed examination on the response of switching hole traps under mixed NBTI/CHH stress conditions had been carried out and discussed in this chapter. Experimental results show that the OR of a switching hole trap could be affected in several interesting ways when subjected to V_d stress. The change in a trap's OR may be ascribed to several factors as summarized below: 1) A competing effect between the decreasing oxide field along the channel due to the applied V_d and the CHH effect. The former decreases the likelihood of a net hole capture by the trap while the latter increases such likelihood, as revealed by the extracted capture and emission time constants. 2) An

interaction between the CHH effect and the trap that leads to the latter being able to retain the trapped hole more permanently, through supposedly a structural change that arises from the enhanced phonon effect under hot hole condition. 3) The existence of a trap's switching threshold which leads to a steep fall in the trap's OR when the oxide field falls below this threshold due to the applied V_d . The possibly of two nearby traps affecting each other's capture activity, leading to an anti-correlated change in OR is also observed.

CHAPTER 6: GATE-OXIDE TRAPPING ENABLED SYNAPTIC LOGIC TRANSISTOR

6.1 Introduction

Electronic traps in the gate oxide of the MOS field-effect transistor are known to have a profound influence on the transistor operation and they have been studied as a reliability issue, as well as a possible avenue for novel applications by exploiting the randomness of charge trapping and detrapping. Random telegraphic drain-current (I_d) noise [169-171] and threshold-voltage (V_t) instability [144], [172-176] that arise from charge capture/emission by oxide traps have been intensively studied due to their potential impact on timing-sensitive circuits, e.g., causing timing jitter in the SRAM [23], [177], [178]. Furthermore, some studies have leveraged the stochastic nature of the capture/emission processes for realizing physically unclonable functions (PUFs) and a true random number generator (TRNG) [179-181], a key primitive for future data and communication secured applications on the internet of things era. Others have explored the possibility of using a charge-trap memory transistor for the implementation of artificial synapse [182-186] that emulates the plasticity behavior of biological synapses in response to neuron spikes, due to the capability of consecutively modulating the population of filled traps and hence the channel conductance of the MOS structure. In addition to the HCD related reliability studies reported in Chapters 4 and 5, given the mature transistor technology, the exploration of utilizing a logic transistor to function as

suitable artificial synapses for neuromorphic applications was thoroughly investigated as well.

A dramatic increase in the need to extract general relations from large sets of high-dimensional data has driven the use of machine learning and multidimensional optimization algorithms in fields such as pattern recognition, computer vision, natural language processing, and autonomous systems. The unprecedented booming of intelligent technologies is undoubtedly poised to bring an impressive influence on diverse sectors of everyday work life. Nonetheless, the execution of data transmission between the logic-memory and dedicated central processing units (CPUs) based on classical computing architectures puts an intrinsic cap on the speed of computations, which is fundamentally incompatible with the massive parallelism computation feature associated with flexible software-based algorithms, such as the advanced construction of artificial neural networks (ANNs). In light of this discrepancy, these ANNs will consume significant computer resources and electrical power, particularly during training and deploying neural network algorithms, thereby igniting a strong interest in bringing these requirements down and in enabling continuous on-line learning to take place in self-contained, mobile configurations. Taking inspiration from the most efficient known computational entity, the human brain, which superior in complex cognitive tasks with low energy consumption and high speed, massively parallel computing has emerged as a leading paradigm designed to provide a shift beyond traditional computer processing bottleneck. In particular, it would achieve speed-ups, and multivariate computation, giving rise to a truly extraordinary efficient device. Generally, this type of parallel computing is called neuromorphic computing, a growing expectation is that neuromorphic technologies would have a pronounced role to play in delivering these efficiency gains.

Figure 6-1. An analogy between a biological synapse and an artificial synapse based on a logic n-FET: (a) Diagrammatic of biological neurons and the direct gap between the axon end of the previous neuron and the dendrites of the next neuron formed synapse, where the synaptic connection strength is reflective of the number of Ca^{2+} ion channels formed by the difference in the action potentials from pre- and post-synaptic neurons as the brain adapt to new information; This characteristic is emulated in the later in the form of charge trapping in the gate oxide, controlled by voltage pulses applied at gate and drain terminals. (b) 3D schematic of our logic synaptic transistor response to the different polarity pulse signal. Hole trapping (top) increases the channel conductance and can be used to represent an increase in synaptic strength. Electron trapping (bottom), on the other hand, decreases the channel conductance and the synaptic strength.

Neurologically, the nervous system can accomplish many perceptual intelligence functions and responds precisely to external environmental stimuli. These complex tasks, such as linguistic comprehension, abstract reasoning, motor control, are efficiently performed by a dense network of neurons (computing element) $\approx 10^{12}$ with corresponding trillions of chemical synapses (versatile memory element) $\approx 10^{15}$. A

biological synapse is a gapped junction between two neurons, where interneural signaling happens through releasing diverse excitatory and inhibitory neurotransmitters [187]. Most notably, the connection strength, i.e., synaptic weight is reflective of the number of Ca^{2+} ion channels formed by difference in action potentials between pre- and post-synaptic neurons, which can be modulated into either excitatory (potentiation) or inhibitory (depression) post-synaptic potential/current by consecutive incentives [188-190]. The process of reconfiguring and memorizing signal conductivity between two-linked neurons is commonly termed functional plasticity (Figure 6-1(a)). It is widely believed this feature dictates that the learning and adaptability activities in the brain. Thus, a critical unit towards the hardware realization of neuromorphic systems is the artificial synapse, which implements diverse synaptic plasticity for mimicking features of biological synapses by modulating the strength of the connection between neurons as a function of activity.

To this end, neuronal spiking and synaptic response have been constructed relying on the specific functionality and robustness of the circuit design based on conventional silicon complementary metal-oxide-semiconductor (CMOS) technology [191-193]. Additionally, to reduce system complexity and solve the suffering from the issue of scalability due to area and energy inefficiencies, the incorporation of augmentative functionalities into a single synaptic device has been actively committed. As a result, various types of devices have been successfully attempted as possible artificial synapse candidates for the implementation of synaptic emulation. Thanks to the advantages of nonvolatile resistance modulation, low power consumption, good retention, and endurance, the two-terminal memristor is accepted as the most popular synaptic simulator. Especially, a simple but efficient sandwich structure facilitates its excellence in three-dimensional (3D) integration. The existing memristive adoptions typically have

abundant ion dynamics, such as resistive random access memory (RRAM) [48], [50], [194-197], phase-change random access memory (PCRAM) [198], [199], ferroelectric random access memory (FeRAM) [200], [201], etc. Temporal and spatial randomness in the movement of ions and nanoparticles usually exist in these materials and structures, in which brings a key challenge in controlling device-to-device variability. Thus, despite heartening progress so far, applicability in large-scale neuromorphic systems remains to be further improved. On the other hand, transistor-type devices are also being ignited to mimic synaptic operations, e.g., floating-gate non-volatile memory transistor. This device, capitalizing on the charge trapping/de-trapping mechanism within the gate oxide stack, exhibits enduring replication of neuro-biological behaviors [202], [203]. However, complex structural design for reliable controllability has caused incompatibility issue, thereby limiting its integration capability in logic circuitry. The manufacturing-available Si-based logic transistor shows the preponderance in this regard, due to its CMOS-compatibility, excellent scalability for large-scale integration, higher accuracy, and greater stability for synaptic weight adjustment. Therefore, logic transistor is an alternative promising manner for implementing synaptic functions. However, not many research works have been reported to date [204].

With the motivations aforementioned, we aim to study the synaptic responses of a normal logic transistor due to its commercial fabrication capability. Hafnium dioxide/silicon oxide ($\text{HfO}_2/\text{SiO}_x$) stack with sub-2 nm equivalent thickness is used as the gate insulator in our synaptic device since it has excellent stability and has been widely deployed in the mature CMOS manufacturing process. Metal gate TiN is selected due to its compatibility with CMOS technology [205]. Asynchronous plasticity in both potentiation and depression is crucial for robust learning process, tunable information transmission, and network dynamics. Excitatory behavior is obtained by negative pulse

Figure 6-2. Gate potential stimuli inducing the corresponding excitatory and inhibitory behaviors by the hole trapping and electron trapping effect, imitating the process of updating weights during learning epochs. Specifically, transfer curves of an n-FET test device showing a reduced threshold voltage or increased drain current due to hole trapping after a negative gate voltage pulse and an increased threshold voltage or reduced drain current due to electron trapping after a positive gate voltage pulse. Inset depicts the energy band diagrams corresponding to hole or electron trapping in the gate stack caused by a negative or positive stimulus applied on the presynaptic neuron (gate electrode).

stimuli which is ascribed to the persistent trapping dynamics of carrier holes. Inhibitory synaptic function is realized by positive pulse stimuli which is attributed to the trapping effect of electrons from inversion layer (Figure 6-1(b)). The positive (negative) shift of transfer characteristics of an n-FET device in Figure 6-2 elucidates the modulation of I_d by positive (negative) gate voltage pulse. The inset depicts the energy band diagrams corresponding to electron or hole trapping in gate stack. Our results show various synaptic behaviors including short-term/long-term plasticity (STP/LTP), spike amplitude, and action number dependent plasticity in our logic transistors. Furthermore, built on intrinsic charge trapping kinetics within oxide defects, the artificial synapses are able to naturally realize advanced functions such as metaplasticity and classical

asynchronous conditioning spike-timing-dependent plasticity (STDP), as well as implementation of STDP without the need for intentionally designed overlapping pulses. In addition, results on neural network simulations show that the logic synaptic transistor can be programmed for recognition tasks. Implementation of reliable and reproducible synapse functions by capitalizing on matured manufacturing technology, which shows a promising way for cost-saving and shorter development cycles to realize the goal of highly intelligent and robust neuromorphic systems.

6.2 Experimental Details

The test devices (DUTs) used in this study were fabricated utilizing the 0.13 μm High-k/Metal Gate (HKMG) CMOS commercial technology. The gate dielectric stack is composed of 5-nm HfO_2 (ALD) and 1-nm interfacial SiO_2 layer which is chemically grown. An equivalent oxide thickness here is about 1.7-nm. Excellent CMOS performance and manufacturability is demonstrated earlier in Semtech products. The n-channel devices used show 200 nm channel length and 1 μm channel width, and p-channel devices show 250 nm channel length and 10 μm channel width. Here, ultra-fast current-voltage (I - V) measurements (via transient mode) throughout the experiment were carried out on a Keithley model 4200-SCS parameter analyzer equipped with model 4225-RPM remote amplifier/switches, and all the characterizations are performed in an ambient atmosphere with room temperature. Furthermore, in all measurements the source and bulk electrode were connected to ground, engineered to eliminate floating body effect. Here, n-FET, as a representative example, emulated a synaptic cleft that controls the flow of ionic charge between two neighboring neurons to implement synaptic weight updating (Figure 6-1(a)). In addition, as illustrated in Figure 6-1(b), the

gate and drain electrodes of the transistor were connected to the pre- and post-synapses, respectively, which are designed to receive an external stimulus.

6.3 Results and Discussion

In biological synapses, the transfer of action potentials from presynaptic neuron to postsynaptic neuron induces postsynaptic signals. In particular, the postsynaptic current (PSC) would afford either excited or inhibited response, which results in the creation of either excitatory or inhibitory (positive or negative) weight parameter, i.e., synaptic plasticity. With the same analogy, the electrical signals, applied to the gate electrode of our device corresponds to the presynaptic stimuli whereas the channel current, flowing through the source and drain (I_{ds}) corresponds to the PSC. A negative (positive) pre-synaptic (gate) spike will modulate filled traps population in the gate insulator layer of a logic transistor and attract (deplete) additional electrons from the conductive channel, giving rise to a slight increase (decrease) in I_{ds} , i.e., excitatory/inhibitory postsynaptic current (EPSC/IPSC). This conductance difference induced by external signal stimuli reflects the connection strength between two logic nodes. A single presynaptic spike (-1.6 V or $+1.0$ V, 50 μ s) is observed to trigger a peak EPSC or IPSC of $+2.13$ μ A or -0.74 μ A at pulsed gate voltage of $+0.85$ V and drain voltage of $+0.1$ V, before decaying to a resting current level after a few milliseconds (Figure 6-3(a), (c)). The short-term strength of PSCs can be modulated by varying the duration of presynaptic spike. It is discernible that an evident positive correlation between the PSCs and the width of the applied spike because more charges will be driven to trapping in gate oxide defects with increasing spike duration. In addition, an increase in the number of spikes can imitate the repetitive rehearsal processes in brain to tune synaptic plasticity [206], [207]. This basic characteristic accord closely with our experimental results, as representation in

Figure 6-3. The fundamental synaptic plasticity for both excitatory and inhibitory mode in an n-FET, where excitatory response caused by a negative pulse and inhibitory response arose from a positive pulse: (a) and (c) Dependence of the excitatory/inhibitory drain (post-synaptic) current response on the gate (pre-synaptic) voltage spike duration (50 to 500 μs); Progressive modulation of the drain (post-synaptic) current evidence the spike number-dependent plasticity (SNDP): (b) and (d) Excitatory/inhibitory post-synaptic current response with varying number of negative/positive gate (pre-synaptic) spikes, and corresponding fitting, indicating the stretched-exponential modulation process in both potentiation and depression. Inset in each figure gives the details of the voltage waveforms. A wait period of 40 s (sufficient for complete recovery of the current response (cf. Figure 6-5) at the pulse voltages applied) is inserted in-between measurements involving different pulse widths.

Figure 6-3(b), (d). When the action number is increased, the summation of PSCs value increases and gets saturated eventually, manifesting the spike number-dependent plasticity (SNDP). If the next stimulus arrives before the effect of the earlier one subsides, the influence of the consecutive pulse train will be accumulated. Thus, in the beginning, although accompanied by the detrapping effect, more and more carriers will

be injected into the oxide defects when increasing spike numbers, which results in an accumulated modification of channel conductivity. Despite of this, there are only a limited number of trap levels below the Fermi level E_f due to the energy band bending at a specific gate voltage. Consequently, as a greater number of traps are populated, gradually all or most of the traps that can be affected by the spike amplitude are influenced and the PSCs will eventually saturate. For further examination, we use the following device behavioral model to study the non-linear weight dependent plasticity [208]. The correlation between PSC and the number of pulses (P) is illustrated with the equation below:

$$PSC = \alpha \times \left(1 - \exp\left(\frac{-P}{A}\right) \right) + PSC_{\min} \quad (\text{Eqn. 6-1})$$

$$\alpha = (PSC_{\max} - PSC_{\min}) / \left(1 - \exp\left(\frac{-P_{\max}}{A}\right) \right) \quad (\text{Eqn. 6-2})$$

$PSC_{\max}$, $PSC_{\min}$, and $P_{\max}$ represent the maximum PSC, the minimum PSC, and the maximum pulse number required to switch the device between the minimum and maximum conductance states, respectively. These three parameters are directly extracted from experimental data. A is the fitting parameter that controls the non-linear weight dependent plasticity.

When a consecutive voltage train containing ten presynaptic action potentials of pulse (-2.0 V, 10 ms) is applied, the synaptic plasticity then is observed highly dependent upon the pulse interval or stimulation rate, exhibiting dynamic high pass temporal filtering characteristic, as shown in Figure 6-4. By way of illustration, the amplitude of EPSC explicitly increases with a higher facilitation index when the pulse rate increases from 10 to 1000 Hz. The observation results can be explicable by the volatile modulation in channel conductance with frequency-dependent gain, which

Figure 6-4. (a) Comparison of the response of the device to a successive pulse train comprising 10 identical pulses (-2.0 V, 10 ms) with different frequencies or pulse intervals (10, 100, 1000 Hz). (b) The synaptic plasticity change index or gain, which depends on $(A_{10} - A_1)/A_1$, increases from 9.6% to 17.3%, suggesting the existence of spiking-rate-dependent plasticity or a dynamic high-pass filtering characteristic.

again bears similarity to the spiking-rate dependent plasticity (SRDP) or dynamic filtering characteristic in the neuronal system. It is commonly accepted that these short-term features are essential to enable information transmission and very rich network dynamic characteristics, such as adaptation, temporal filtering, damped oscillation, traveling front and pulse, state hopping with transient population spike, robust self-organized critical activity, etc. For instance, short-term depression provides a mechanism that automatically disables network activity to achieve iconic sensory memory [209]. Apart from the modulation of EPSCs and IPSCs with varying durations or numbers of presynaptic stimuli, we also show that channel conductance in our device is adjustable with spike amplitude, which produces amplitude-dependent plasticity. A small negative presynaptic spike yields a fast decaying excitatory response with no remnant ΔI_{ds} degradation, giving a short-term working memory. The physical origin is

Figure 6-5. (a) and (b) The decay evolution of excitatory/inhibitory drain (post-synaptic) current that triggered by the negative/positive gate (pre-synaptic) spikes with different amplitudes and the same width of 10 ms, resembling the “forgetting” curves of biological synapses following pre-synapse firing. Inset in each figure gives detailed setting of the spike waveforms. All the current was monitored after each stimulus at pulsed V_g of +0.85 V and V_d of +0.1 V (0.1 μ s).

due to the “filling” of only shallow oxide traps which emits the trapped holes rapidly during relaxation [66]. On the other hand, with a more negative presynaptic spike, the I_{ds} recovery reaches stagnant to a non-zero remnant ΔI_{ds} , indicating the existence of a relatively long-term memory effect. The storage of a non-negligible remnant charge in the thin, direct tunneling gate oxide is possible as a result of the activation of deep-level hole traps holding much longer emission time by the higher negative gate pulse [66]. Similar observations as well apply to a positive presynaptic spike, which generates an IPSC.

A general trend of increased for both the potentiation and depression is observed as the amplitude of incident action potentials increase. After-pulse I_d measurement

Figure 6-6. (a) Short- and long-term synaptic weight changes as a function of the gate (pre-synaptic) spike voltage. The short- and long-term weight change is computed as the difference in the drain current before pre-synaptic spike and the average current after spike taken before 500 μs and between 1 and 40 s, respectively. The lines are modeled with an exponential function. (b) Stability of the oxide traps of a n-FET: excitatory current response at selected repeated spiking numbers (5000). Inset shows the various presynaptic excitatory pulses (10 ms) selected for the endurance testing.

reveals a PSC relaxation behavior (Figure 6-5(a), (b)), each case can be modeled with the bi-exponential function in the form [210]:

$$\Delta I(t) = A_1 \times e^{-t/\tau_1} + A_2 \times e^{-t/\tau_2} + \Delta I_0 \quad (\text{Eqn. 6-3})$$

where the time constant τ_1 describes the initial rapid decay right after the spike. The other time constant τ_2 fits the comparatively slow decay that settles to a quasi-steady-state value ΔI_0 . For example, fitting values of τ_1 and τ_2 are 28.21 μs, and 1.62 ms (1.83 ms, and 57.54 ms) for -1.6 (+1.2) V gate spike, respectively. It shows that the evolution of PSCs relaxation resembles the dynamic “forgetting” curve of a biological synapse [188]. These behaviors can be fully explained by the internal charge relaxation dynamics, based on an appreciably high density of pre-existing structural defects existing in the

transition metal oxide. Synaptic weight change (Δw) as a function of presynaptic pulse amplitude is depicted in Figure 6-6(a). Short-term plasticity (STP) is defined as the difference between the I_{ds} averaged over the first 500 μs after a presynaptic spike, and that before the spike. We define long-term plasticity (LTP) as the difference between the average I_{ds} between 1 s and 40 s after a presynaptic spike, and that before the spike. The magnitude of different plasticity increases with increasing presynaptic spike amplitude and can be described by an exponential function of the form:

$$\Delta w = A \times e^{V_{gs}/V_e} + \Delta w_o \quad (\text{Eqn. 6-4})$$

where A is the prefactor, V_{gs} is the applied spike voltage, V_e is related to the activation energy for the oxide traps and Δw_o depicts change induced by the read voltage.

Synaptic plasticity is modulated by a learning action stimulus [211]. In addition to simulating the basic behavior of a biological synapse, a neuromorphic device has to exhibit advanced forms of synaptic plasticity, such that it is sufficiently flexible for real-time neural network applications. The synaptic weight change for the same neural activities may depend on the initial condition of the synapse; that is, the historical activity of the synapse will largely dictate its subsequent synaptic responses and ultimate neural functions. This characteristic is called metaplasticity, i.e., the plasticity of synaptic plasticity [212]. A logic n-FET mimics this functionality by applying a “priming” pulse at the gate, before applying the main learning action to evoke synaptic plasticity. The first priming pulse represents a prior pre-synaptic neuronal activity, which is equivalent to the activity of an internal modulator that regulates the subsequent synaptic efficacy. To demonstrate a typical feature of metaplasticity, the normal +1.6-V 50- μs and -1.6-V 500- μs gate voltage spikes for triggering the IPSC and EPSC responses, respectively, were preceded by different potentiating priming pulses to mimic

Figure 6-7. Meta-plasticity behavior of a TiN/HfO₂ gated n-FET. (a) IPSC responses with and without different amplitudes of priming pulses. Inset of (a): Holes trapping in HfO₂, induced by the priming pulse, leads to a greater trapping of electrons during the pre-synaptic spike due to the enhanced surface electric field and reduced the capture energy barrier. IPSC is enlarged by an excitatory priming pulse applied prior to the inhibitory pre-synaptic pulse (b) EPSC responses with and without various amplitudes of priming pulses. Inset of (b) shows reduction of hole trapping during the excitatory pre-synaptic pulse due to holes already trapped in the HfO₂ during the prior priming pulse. EPSC is expected to be suppressed by a prior excitatory priming pulse.

prior neuronal activity. The residual effect of the preceding pulse interacted with or counteracted the contribution of the subsequent pulse. Figure 6-7(a) shows that a negative priming pulse enlarged the depression or inhibitory response generated by a subsequent positive spike, due to the hole trapping that occurred during the priming pulse. The trapped holes increased the interface field, which increased the electron trapping. The inset in Figure 6-7(a) shows the energy band diagram; it can be seen that the potential built-in by the trapped holes induced a greater downward bending of an energy band than that in the absence of trapped holes and thus reduced the energy barrier for subsequent electron trapping. In contrast, a negative priming pulse suppressed the

Figure 6-8. (a) and (b) Weight percentage changes (%) as a function of the priming pulse voltages. Inset schematic diagram depicts the details on the voltage waveforms setting to trigger meta-plasticity. The readout current was collected at pulsed V_g of $+0.85\text{ V}$ and V_d of $+0.1\text{ V}$ ($0.1\text{ }\mu\text{s}$).

potentiation or excitatory response of a subsequent negative spike (Figure 6-7(b)), which is attributable to the fact that the holes trapped during priming weakened the interface field and impeded further trapping of the holes during the negative spike. The inset in Figure 6-7(b) shows the energy band diagram under an applied negative-learning action spike. Figure 6-8 quantitatively represents the synaptic-weight changes extracted from the PSCs of the above two cases, which increased from approximately 3% to 9% with the priming pulse amplitude, highlighting the role of a priming pulse in the internal state of a device. This process facilitated a highly nonlinear optimization of the synaptic weights based on the signal sequence they transformed, which in turn resulted in optimal selectivity for these signals. The details of the voltage waveform setting required to trigger metaplasticity are shown in the inset of Figure 6-8.

In addition, to support the robustness of the implementation, we examined the stability of the oxide-trapping kinetics under repeated spiking stimulation. An example of the process used to select a normal excitatory pulse is shown in Figure 6-6(b), as a scenario of the endurance test, and depicts the average and standard deviation of the

Figure 6-9. Different forms of spike-time dependent plasticity or STDP characteristic learning in a TiN/HfO₂ gated n-FET, where the weight change Δw as a function of relative arrival timing difference Δt between the pre- and post-synaptic spikes: (a) Symmetrical anti-Hebbian rule, showing tunable STDP learning behavior, as a demonstration by changing pulse parameters; (b) Symmetrical Hebbian rule; The solid lines in the left and right diagram are both expressed by Gaussian fitting.

evolution of the excitatory current response with relaxation time in 5,000 iterations of the spiking numbers. As can be seen, there was not a substantial variation in the PSC response, whose relaxation function agrees well with that in Eqn. 6-3. This indicated that a reproducible synaptic response was achieved with fewer cycle-to-cycle variations and may overcome some of the drawbacks associated with memristive implementation.

The synaptic plasticity resulting from the high correlation between the temporal order and interval (Δt) of the pre- and post-synaptic neuron activities was introduced by Hebbian [213]. One of the most important Hebbian learning rules is spike-timing-dependent plasticity (STDP) with precise timing control, by which the magnitude and polarity of the synaptic weight can be modulated. STDP is thus a key synaptic modification rule, as it enables the transformation of the time information within neural networks into memory storage. The change in biological synaptic weight can be

Figure 6-10. Asymmetrical Hebbian rule; Insets in the above figure depict the shape of pre-synaptic and post-synaptic voltage pulses waveform. The solid lines inside are expressed by exponential fitting.

expressed in the form of relative spike timing $\Delta t (= t_{\text{post}} - t_{\text{pre}})$, where t_{post} and t_{pre} , indicate the arrival times of post- and pre-synaptic neuron spikes, respectively. Here, we show that a logic n-FET may be used to mimic this functionality of biological synapses and enable the effective translation of temporal spike correlations into voltage amplitude differences created by long-term potentiation/depression, and thus exhibit different STDP behaviors. The drain current was recorded under the conditions of $V_g = +0.75$ V and $V_d = +0.3$ V before and after the spiking, to determine the change in the synaptic weight. In the logic transistor, we used the change in conductance ($\Delta G/G$) of the device to quantify the alteration of synaptic efficacy. As shown in Figure 6-9(a) and (b), a typical symmetric learning rule was followed by engineered overlapping pre- and post-synapse spike pulses. The variation in the arrival times of the pre- and post-synaptic spikes finely adjusted the Δw value: a larger value of Δw was obtained when the time interval Δt approached 0, whereas a smaller Δw value was obtained when Δt moved away from 0. The symmetric weight changes obtained as a function of the interval between the two spikes are generalized using the Gaussian function, as follows [214]:

		Fitting Parameters	
		A (%)	τ (ms)
Symmetric Hebbian		28.8	9.7
Symmetric Anti- Hebbian	$V_{pre-peak}=1.2$ V	-19.8	28.1
	$V_{pre-peak}=1.4$ V	-23.7	30.1
	$V_{pre-peak}=1.6$ V	-30.1	32.6

Table 6-1. The fitting values of scaling factor (A) and time constant (τ) for the symmetric Hebbian STDP learning rule.

$$\Delta w = A \times \exp\left(\frac{-\Delta t^2}{\tau^2}\right) + \Delta w_0 \quad (\text{Eqn. 6-5})$$

where Δw , A , Δt , τ , and Δw_0 denote the change in synaptic weight, fitting constant, time interval, fitting time value, and synaptic weight constant, respectively. Figure 6-9(a) demonstrates the achievement of a tunable STDP learning activity by varying the spike parameters, where both A and τ increase with the waveform voltage. In practice, these can be modified to obtain the desired STDP behavior. Table 6-1 shows that the fitting time values of our synaptic device agree well with the range of values observed in biological synapses, indicating our successful reproduction of symmetric STDP [215]. In addition, Figure 6-10 depicts another type of asymmetric STDP learning rule. Synaptic weight potentiates (i.e., stimulates) an LTP response when the postsynaptic spike fires before the presynaptic spike (positive spike-timing interval, $\Delta t > 0$). In contrast, the synaptic weight depresses an LTP response (i.e., an LTD response is triggered) in a reversed sequence in which a presynaptic spike fires briefly after the postsynaptic spike (negative spike-timing interval, $\Delta t < 0$). A shorter Δt yielded a larger change in the synaptic weight, whereas a longer Δt had an insignificant effect on synaptic

Figure 6-11. Schematic of the non-overlapping pulse-pairs for STDP implementation, as well as manifesting the adjustability of the STDP behavior by terms of varying pulse-pairs amplitude. As schematized in the left, pre-synaptic and post-synaptic spikes are both applied on the gate electrode. The solid lines inside the right diagram are used exponential fitting.

weight. Thus, the law of an antisymmetric STDP form is expressed using the following exponential function [216]:

$$\Delta w = \begin{cases} A_+ \times e^{-|\Delta t|/\tau_+}, & \Delta t > 0 \\ A_- \times e^{-|\Delta t|/\tau_-}, & \Delta t < 0 \end{cases} \quad (\text{Eqn. 6-6})$$

where $A_{\pm}$ and $\tau_{\pm}$ are represent the scaling factors and time constants, respectively. The values of τ_+/τ_- were obtained as 51.4/41.9 ms from the fitting curve presented in Figure 6-10. The timed encodings and spatial summations of PSCs indicate that the artificial synapse has immense potential to enable homeostatic modulation of its excitability.

Despite the STDP behaviors being validated by manipulation of the spike forms, as in many previous studies, all of these traditional techniques inevitably require strict timing mechanisms and additional peripheral circuits. The latter require large physical chip areas and considerable memory, which renders them impractical for scaling and limits their applications. As such, there is a strong demand for a new type of synaptic device that can inherently incorporate STDP-like characteristics and is thus suitable for

		Fitting Parameters			
		A_+ (%)	τ_+ (μ s)	A_- (%)	τ_- (μ s)
Asymmetric Hebbian	+1.8/-1.8 V	10.6	41.2	-13.5	-41.7
	+1.6/-1.6 V	9.2	38.5	-11.5	-34.9
	+1.4/-1.4 V	7.0	31.3	-9.0	-31.8

Table 6-2. The fitting values of scaling factor (A_+/A_-) and time constant (τ_+/τ_-) for the positive and negative Δt in non-overlapping pulse pairs enabled asymmetric Hebbian STDP learning rule.

implementing biologically equivalent large-scale spiking neural networks (SNNs). Figure 6-11 shows the simple achievement of STDP in our device, due to a natural timing mechanism of non-overlapping pre- and post-synaptic pulse pairs, both of which were applied to the gate electrode. The nature and extent of the effective weight change/conductance depended on the correlation between the pulse pairs. As the preceding positive (negative) pulse causes short-term depression (potentiation) when $\Delta t > 0$ ($\Delta t < 0$), the residual effect caused by the preceding pulse-introduced trapping of negative (positive) carriers determined the net effect of the pulse pairs. Taking the case of $\Delta t > 0$ as an example, in longer intervals (absolute value of Δt increases) a prolonged gap between two adjacent pulse pairs suppressed the conductance change, because the residual effect of the preceding pulse decayed with time (due to the detrapping of charged carriers). However, when Δt decreased, the stronger residual effect of the preceding pulse enhanced the synaptic connection change induced by the second pulse, resulting in the device following an asymmetric Hebbian learning rule. The dependence of PSC on the spike amplitude played a significant role in this scenario.

Figure 6-12. Comparison of transfer curve before (1st) and after alternating the (a) p-FET and (b) n-FET between excitatory and inhibitory responses for 9 times (10th), demonstrating that the erase operation helps devices to return to their initial state for subsequent alternative measurements.

Consequently, we implemented tunable symmetric STDP rules by the configuration of pulse pair amplitudes, in which Δt alone modulated the synaptic weight. The curve fitting result agreed well with the biphasic decaying exponential form (Eqn. 6-6), and the fitting parameters extracted for three different pulse pairs are displayed in Table 6-2. In addition, when using such nonoverlapping short pulse pairs, the accompanying recovery time decreased accordingly, which reduced Δt during STDP learning, and thereby, remarkably increased the training speed [195]. The above results indicate that the current-enriched STDP characteristics enabled the development of varying learning rules. If different learning rules are feasible in an SNN, more sophisticated scenarios can be addressed by the SNN, with improved learning efficiency.

Given our achievement of synaptic behaviors on a logic n-Si transistor, naturally, we examined whether p-Si logic devices could also perform the same function. Motivated by this, an artificial p-Si transistor-based synapse was realized. The concept of this synapse is as follows. Using the flexibility of filled traps in an MOS structure

Figure 6-13. Fundamental synaptic signatures for both excitatory and inhibitory mode in a logic p-type transistor. Excitation (a) and (c) inhibition behavior triggers an EPSC and IPSC, in which spike-duration-dependent plasticity caused by a positive and negative presynaptic spike of +1.2 V and -1.2 V, respectively. Pulse number facilitation responses, i.e., spike-number-dependent plasticity, resulted from applying increasing number of positive and negative presynaptic spike. The lines are also expressed well by stretched-exponential fitting (Eqn. 6-1 and 6-2).

and under the negative spike excitation, holes are trapped with a decrease in channel conductance, leading to IPSC. In contrast, excitation of the positive presynaptic spike induced EPSC, as electrons are trapped with an increase in channel conductance. It is noteworthy to mention that if the device starts from a potentiated state (long-term plasticity), rather than a fresh state, an applied negative-voltage pulse drives the trapped deep-level charge back to the channel. This is akin to erase-processing, and thus converts the transfer recovery curve into the pristine state without a significant change. Figure 6-12 compares the transfer curve before (first) and after alternating the n-FET and p-FET between the excitatory and inhibitory responses nine times (tenth), indicating that these FETs may be erased back to their initial state for subsequent alternative measurements

Figure 6-14. Decay evolution of the excitatory/inhibitory drain (postsynaptic) current in case of a p-Si logic transistor, which is triggered by positive/negative gate (presynaptic) spikes with different amplitudes and the same widths (50 μs). The inset in each figure shows detailed settings of the spike waveforms. EPSC/IPSC is monitored at V_g of -0.85 V/ -0.9 V and V_d of -0.1 V (0.1 μs).

between the excitatory and inhibitory responses, without evident degradation. On top of this, as representation in Figure 6-13(a) and (c), Δ PSCs increased progressively with the duration of the presynaptic stimulus signals, which matched the spike duration-dependent plasticity of a biological synapse. Normally, the postsynaptic responses (EPSCs or IPSCs) triggered by presynaptic signals are recorded as a function of the pulse repetition number, which plays an important role in information processing in biological systems. To demonstrate this, Figure 6-13(b) and (d) depicts an operational schematic of our artificial synapse with SNDP modulation. Furthermore, the spike amplitude-dependent functionality emulated by the proposed p-Si transistor is shown in Figure 6-14. From the perspective of the timing modulation rule, empirical STDP characteristics were verified using the method of encoded pre- and postsynaptic spiking stimulus shown in the inset of Figure 6-15, where the relative spike timing Δt was defined as $t_{\text{pre}} - t_{\text{post}}$. Such reverse STDP phenomena are computationally pronounced

Figure 6-15. Two forms of STDP are implemented in present p-FET device, (a) symmetric Hebbian learning rule and (b) antisymmetric Hebbian learning rule. The lines in (a) and (b) are in accordance with the Eqn. 6-5 and 6-6, respectively.

and useful for associative learning in concert with post-spike hyperpolarization. The experimental results of the two different STDP learning rules are credibly validated in Figure 6-15, where the fitting lines used to describe the relation between Δw and Δt under the above conditions agree well with Eqns. 6-5 and 6-6, respectively. The fitting parameter τ in the symmetric Hebbian learning rule was 13.8 ms, whereas the fitting parameters τ_+/τ_- in the antisymmetric Hebbian learning rule were 31.5/40.8 ms. All of these values lay within a reasonable range of those observed in their biological synapse counterparts [217].

Finally, it is worth noting the responses of some synapses are not completely fixed and can be reconfigured between different synaptic responses obtained by the modulator signals from the central nervous system. This endogenous modulation is the basis for the fine-tuning of tactile sensations in biological systems [218]. To achieve this functionality, the ability of n- and p-Si transistors to exhibit different modes of synaptic responses for the same-polarity stimulus signal was examined, the signal selector could be embedded into pairs of n- and p-Si transistors. This would facilitate a selective determination of the flow direction of the excited pulses. In our device, we could readily

configure the artificial synaptic responses to the same-polarity stimulus to an excitatory or inhibitory mode, to realize dynamic reconfigurability of the artificial synapse. Such reconfigurable synaptic responses are crucial to effective imitation of the subjective modulation of a tactile perception in brain nervous systems (e.g., the cognitive regulation of pain, which is affected by attention, emotion, and other cognitive variables) [219].

6.4 Chapter Summary

In summary, we have experimentally explored the potential of a commercially manufacturable titanium nitride/hafnium oxide gated logic Si-MOSFET to function as an artificial synapse. This MOSFET successfully simulated the diverse synaptic functionalities of its biological counterpart, including short- and long-term bilingual synaptic (excitatory and inhibitory) processes, spike number-dependent plasticity, and spike amplitude-tunable plasticity. Specifically, the rich internal dynamics of oxide trapping in the thin-gate dielectric of a logic transistor enables a native mechanism to encode the synaptic-weight plasticity through gate and drain voltage pulses. Here, the EPSC and IPSC responses originated from the charge–discharge kinetics of the different polarity carriers in the gate insulator of our device. By introducing priming spikes, the effects of metaplasticity on EPSC or IPSC were examined, and different STDP behaviors were achieved by engineering various spiking pulses.

This study illustrates the potential of logic transistors as units for the construction of neural networks. Our results may facilitate the development of a large-scale neural network based on logic transistors that exhibit less device-to-device variability. The viable application of logic transistors as artificial synapses will create new opportunities

for the development of highly intelligent neuromorphic systems. This may include a bio-faithful route toward large-scale hardware neural networks.

CHAPTER 7: CONCLUSION AND RECOMMENDATION

7.1 Conclusion

Electrical-stress-invariant gate-oxide trap's capture and emission time constants have been the basis of device aging models as well as novel applications that leverage the stochastic nature of the capture and emission processes, such as the true random number generator. An examination regarding the impact of channel hot-hole (CHH) stress on active oxide traps monitored at operating condition is conducted, revealing that the CHH stress can either increase or decrease the likelihood of an active oxide trap emission a charged hole at the operating condition. In this work, we show that this general presumption is only valid for about three-fifths of the oxide-trap population studied. For the remaining two-fifths, the traps' capture and/or emission time constants could be changed by the CHH effect. These findings can be explained by possible phonon-facilitated local defect atomic structural changes, resulting in modifying the trap's switching behavior. Moreover, a reversion of the altered trap time constant to the value before the CHH-stress is also observed if given a substantial recovery period, but the period varies significantly for different traps (from several hours to months). Additionally, CHH stress can change the $|\Delta I_d|$ recovery step height without modification in the switching activity of traps during the monitoring cycle. This may be due to the electrostatic effect inducing local changes on channel current density. Since the channel hot-carrier stress (CHC) effect is present in all scaled transistors, the findings would have important implications for models/applications that presume oxide-trap properties to be stress-invariant.

The importance of nitrogen in SiO_xN_y gate stack engineering has prompted a study on how nitrogen impacts the energetics of $V_O\text{-}O_i$ defects in SiO_2 . Based on the first principles simulation, it was found that a nitrogenated $V_O\text{-}O_i$ defect (i.e., $\text{NO}_2\text{Si Si}\equiv\text{O}_3$), where one of the oxygen atoms is substituted by a nitrogen atom, gives rise to a deep-level hole trap when the nitrogen exists in a two-fold coordination. However, the deep hole trap could be changed into a shallow one when an H atom is attached to the nitrogen's lone pair, showing the 3-fold coordinated N. With the above considerations, we propose a plausible three-state defective model for explanation of the changes in τ_e observed after CHH stress. During CHH stress, hydrogen released from hot-carrier induced dissociation of hydrogen bonding migrate through the oxide, disrupting /changing the local lattice environment of the $V_O\text{-}O_i$ defect. This in turn affects the energy levels of the hole traps and the corresponding time constants. The affected traps may revert to the state before CHH stress when further changes to the local environment arise from phononic perturbations causing hydrogen to leave or reattach the defect site.

After analyzing the experimental results of the impact of CHC stressing on the gate-oxide trap's emission, it is intriguing to examine the interaction of BTI and CHC on nanoscale-device degradation. This work elucidates two distinct behaviors of switching hole traps in small, SiON gated p-MOSFETs subjected to mixed CHH and NBTI stressing. Case (1) in all devices tested, the switching rate (SR) of a trap could either be increased or decreased as the drain voltage $|V_d|$ is increased towards the gate voltage $|V_g|$. Then, a further increase of $|V_d|$ eventually gives rise to a decrease of SR. Case (2), after mixed CHC and NBTI stressing, the reduced switching probability is typically reversed upon removing the drain bias. However, in ~11% of the traps, this is not the case, i.e., emit the trapped hole at a longer time frame, meaning the hole-trapping became more permanent. Thus, besides those characteristics of single oxide traps already revealed

through gate biasing studies, these new behaviors arising from drain biasing should also be accounted for in mixed-signal circuits reliability compact modeling.

Brain-inspired neuromorphic systems have attracted much attention as a new computing paradigm for energy-efficient computation by enabling massive parallelism in artificial neural networks. The successful realization of a large-scale manufacturable artificial synapse holds the key to a full-fledged neuromorphic hardware application. We have experimentally explored the possibility of a typical TiN/HfO₂ gated logic Si MOSFET (with $E_{OT} < 2$ nm) functioning as an artificial synapse, where the gate oxide-semiconductor channel plays a role as synapse cleft. Specifically, the internal rich dynamics of oxide trapping in the thin gate dielectric of logic transistor enables a native mechanism to encode the synaptic-weight plasticity through gate and drain voltage pulse. The charging and discharging of oxide traps in the MOS structure is controlled to modulate the channel conductance interpreted as the synaptic weight. A successful demonstration of basic synaptic responses is achieved in both n- and p-channel MOSFETs. In addition, metaplasticity as well as spiking time-dependent plasticity, the higher-order synaptic responses, are also emulated by encoding relative timing of spike waveforms. Furthermore, when exposed to simple spike pairs only on the presynaptic terminal, these devices exhibit the ability to learn without the need for a complex peripheral circuit, mimicking Hebbian learning by internal charge kinetics. Given the mature transistor technology with less consideration of device-to-device variability, as well as without utilization of external factors (e.g., surface adsorption of gas molecules) that are similar to novel structure-based implementations, the feasible application of logic transistors as artificial synapse would be beneficial for the process towards commercial neuromorphic systems.

7.2 Recommendations for Future Work

In the above-discussed work, the stressed samples were rested at room temperature for a long time. The measurement of temperature dependence may help to shed light on the physical mechanism of the traps created in CHH stressing. In fact, we have started to do further experiments including the temperature dependence. In addition, more studies are needed to validate the hypothesis and to understand why certain traps are immune to the CHC stress as well as the different lengths of recovery time of affected traps. It is hoped to obtain fascinating results in the coming future work.

Similar experiments are suggested to be performed on nano-scale HKMG n/p-MOSFETs, and the extent of the switching activity observed in HKMG n/p-MOSFETs can be compared to the observations in SiON p-MOSFETs. In fact, we have started to do further CHC stressing experiments on devices that are HKMG n-MOSFETs with a gate stack comprising TiN/HfO₂ (3.5 nm). The drawn channel length and width of 32 nm and 80 nm, respectively. In some cases, a decrease is observed. Figure 7-1 shows the example of a TiN/HfO₂-gated small-area n-channel transistor studied by a measurement-stress-measurement sequence. After channel hot-electron (CHE) stress, a persistent decrease of trap G1's and G2's τ_c and τ_e (@ $V_g = 1.8$ V)) can be observed, with the decrease becoming more significant after a more severe CHE stress at a higher V_d . The decrease of τ_c and τ_e would mean that the traps can now capture and emit charges more frequently, increasing the frequency of I_d fluctuations in a given time period. For this device, the decrease in τ_c and τ_e persisted for a couple of days before recovering back to the pristine values, similar to the earlier observations found in SiON p-MOSFETs.

The rapid development of neuromorphic electronics has been followed by improved synaptic devices, which can memristively process input signals. Future work may

Figure 7-1. Capture and emission time constants τ_c and τ_e , respectively for two active electron traps extracted from the measurement phase performed after each channel hot-electron (CHE) stress. The test device is a TiN/HfO₂-gated n-channel transistor.

involve exploring the possibility of a single logic device function as a spiking neuron (i.e., showing the leaky integrate-and-fire (LIF) behavior). Successful implementation based on the same device, integration of the neuron and synapse arrays to form a full CMOS-compatible on-chip compact neuromorphic circuit is attractive. Further innovation in the modulation of the synaptic spike on drain terminal to realize tunable spiking neuron responses. Different synaptic device architecture may also be done (e.g., flexible organic transistor or transparent RRAM) to expand the synaptic behavioral space such as to include tunability and responsivity of artificial synapses to different type synaptic activity (e.g., light signal).

LIST OF PUBLICATIONS

- **X. Ju**, D. S. Ang, and C. J. Gu, “Impact of Channel Hot-Hole Stressing on Gate-Oxide Trap’s Emission,” *IEEE Transactions on Electron Devices*, vol. 61, no. 11, pp. 4720–4727, Sep. 2020, doi: 10.1109/TED.2020.3019982.
- **X. Ju** and D. S. Ang, “Alteration of Gate-Oxide Trap Capture/Emission Time Constants by Channel Hot-Carrier Effect in the Metal-Oxide-Semiconductor Field-Effect Transistor,” *IEEE Access*, vol. 8, pp. 14048–14053, Jan. 2020, doi: 10.1109/ACCESS.2020.2966577.
- H. Z. Zhang*, **X. Ju***, K. S. Yew, and D. S. Ang, “Implementation of Simple but Powerful Trilayer Oxide-Based Artificial Synapses with a Tailored Bio-Synapse-Like Structure,” *ACS applied materials & interfaces*, vol. 12, no. 1, pp. 1036-1045, Dec. 2019, doi: <https://doi.org/10.1021/acsami.9b17026>. (***equally contributed**)
- **X. Ju**, and D. S. Ang, “Gate-Oxide Trapping Enabled Synaptic Logic Transistor,” in *Proceedings IEEE International Reliability Physics Symposium (IRPS)*, June 2020, pp. 1–6, doi: 10.1109/IRPS45951.2020.9129338.
- **X. Ju** and D. S. Ang, “Impact of Hot-Carrier Stress on Oxide Trap Switching Activity in Nanoscale FETs,” in *2019 China Semiconductor Technology International Conference (CSTIC)*, July 2019, pp. 1–3, doi: 10.1109/CSTIC.2019.8755672.
- **X. Ju** and D. S. Ang, “Response of Switching Hole Traps in the Small-Area P-MOSFET Under Channel Hot-Hole Effect,” in *Proceedings IEEE International Reliability Physics Symposium (IRPS)*, May 2019, pp. 1–4, doi: 10.1109/IRPS.2019.8720476.
- H. Z. Zhang, **X. Ju**, Y. Zhou, C. J. Gu, J. S. Pan, and D. S. Ang, “Realization of Self-Compliance Resistive Switching Memory via Tailoring Interfacial Oxygen,” *ACS applied materials & interfaces*, vol. 11, no. 44, pp. 41490–41496, Oct. 2019, doi: <https://doi.org/10.1021/acsami.9b11772>.
- C. J. Gu, C. L. Zhou, D. S. Ang, **X. Ju**, R. Y. Gu, and T. L. Duan, “The role of the disordered HfO₂ network in the high- κ n-MOSFET shallow electron trapping,” *Journal of Applied Physics*, vol. 125, no. 2, pp. 025705, Jan. 2019, doi: <https://doi.org/10.1063/1.5059381>.
- **X. Ju**, H. Z. Zhang, Z. Liu, L. R. Feng, and D. S. Ang, “All-Organic Low-power Flexible Optoelectronic Neuronic Device toward Bionic Eyes Systems”, *Nano Energy* (In preparation).

- **X. Ju**, and D. S. Ang, “Biological Synaptic Functionality Emulated in Logic Transistor: A Simple Enabler for Scalable Neuromorphic Computing” Advanced Intelligent Systems (In preparation).
- **X. Ju**, C. J. Gu, and D. S. Ang, “A Vacancy-Interstitial Defect Model for Channel Hot-Hole Stressing Induced Hole Switching Transformation in the P-MOSFETs,” Journal of Applied Physics (In preparation).

REFERENCES

- [1] G. E. Moore, "Progress in digital integrated electronics," in *IEDM Tech. Dig.*, 1975, pp. 11-13.
- [2] G. E. Moore, "No exponential is forever, but "forever" can be delayed," in *ISSCC Tech. Dig.*, 2003, pp. 20-23.
- [3] E. J. Nowak, "Maintaining the benefits of CMOS scaling when scaling bogs down," *IBM Journal of Research and Development*, vol. 46, pp. 169-180, 2002.
- [4] E. Mollick, "Establishing Moore's law," *IEEE Annals of the History of Computing*, vol. 28, pp. 62-75, 2006.
- [5] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp. 114 ff," *IEEE Solid-State Circuits Society Newsletter*, vol. 20, pp. 33-35, 2006.
- [6] D. C. Brock and G. E. Moore, Understanding Moore's law: four decades of innovation: Chemical Heritage Foundation, 2006.
- [7] International Technology Roadmap for Semiconductors (ITRS). [Online]. Available: <http://www.itrs2.net/>.
- [8] S. Thompson, et al., "A 90 nm Logic Technology Featuring 50nm Strained Silicon Channel Transistors, 7 layers of Cu Interconnects, Low k ILD, and 1 μ m² SRAM Cell," in *IEDM Tech. Dig.*, 2002, pp. 61-63.
- [9] K. Mistry, et al., "A 45nm Logic Technology with High-k/Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193 nm Dry Patterning, and 100% Pb-free Packaging," in *IEDM Tech. Dig.*, 2007, pp. 247-250.
- [10] C. Auth, et al., "A 22nm High Performance and Low-Power CMOS Technology Featuring Fully-Depleted Tri-Gate Transistors, Self-Aligned Contacts and High Density MIM Capacitors," in *VLSI Symp. Tech. Dig.*, 2012, pp. 131-132.
- [11] S. Narasimha, B. Jagannathan, A. Ogino, D. Jaeger, et al., "A 7 nm CMOS technology platform for mobile and high performance compute application," in *IEDM Tech. Dig.*, 2017, pp. 689-692.
- [12] R. Xie, P. Montanini, K. Akarvardar, N. Tripathi, et al., "A 7 nm FinFET technology featuring EUV patterning and dual strained high mobility channels," in *IEDM Tech. Dig.*, 2016, pp. 47-50.
- [13] C. Auth, A. Aliyarukunju, M. Asoro, et al., "A 10 nm high performance and low-power CMOS technology featuring 3rd generation FinFET transistors, self-aligned quad patterning, contact over active gate and cobalt local interconnects," in *IEDM Tech. Dig.*, 2017, pp. 673-676.
- [14] S.-Y. Wu, C.Y. Lin, M. C. Chiang, J. J. Liaw, J. Y. Cheng, S. H. Yang, et al., "A 7 nm CMOS platform technology featuring 4th generation FinFET transistors with

- a 0.027 μm^2 high density 6-T SRAM cell for mobile SOC applications,” in *IEDM Tech. Dig.*, 2016, pp. 43-46.
- [15] J. H. Stathis, M. Wang, R. G. Southwick, E. Y. Wu, B. P Linder, E. G. Liniger, G. Bonilla, and H. Kothari, “Reliability challenges for the 10 nm node and beyond,” in *IEDM Tech. Dig.*, 2014, pp. 522-525.
 - [16] A. S. Oates, “Will Reliability Limit Moore’s Law? ,” in *IEDM Tech. Dig.*, 2014, pp.526-529.
 - [17] G. Groeseneken, et al., “BTI reliability of advanced gate stacks for Beyond-Silicon devices: challenges and opportunities,” in *IEDM Tech. Dig.*, 2014, pp.828-831.
 - [18] T. J. K. Liu, *VLSI Symp. Short Course*, 2012.
 - [19] G. Groeseneken, *IEDM, Short Course*, 2014.
 - [20] Guido. Groeseneken, R. Degraeve, B. Kaczer, and K. Martens, “Trends and perspectives for electrical characterization and reliability assessment in advanced CMOS technologies,” in *IEEE ESSDERC*, 2010, pp. 64-72.
 - [21] Y. Taur, and T. H. Ning, *Fundamentals of modern VLSI devices*, 2nd ed., Cambridge University Press, London, 2009.
 - [22] V. Huard, F. Cacho, X. Federspiel, W. Arfaoui, M. Saliva, and D. Angot, “Technology scaling and reliability: challenges and opportunities,” in *IEDM Tech. Dig.*, 2015, pp. 551-556.
 - [23] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo, G. A. Sai-Halasz, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind and H.-S. P. Wong, “CMOS scaling into the nanometer regime,” in *Proceeding of the IEEE*, vol. 85, no. 4, pp. 486-504, Apr. 1997.
 - [24] K. Aadithya, S. Venogopalan, A. Demir, and J. Roychowdhury, “Mustard: a coupled, stochastic/deterministic, discrete/continuous technique for predicting the impact of random telegraph noise on SRAMs and DRAMs,” in *ACM/IEEE DAC*, 2011, pp. 292–297.
 - [25] K. Agarwal, and S. Nassif, “Statistical analysis of SRAM cell stability,” in *ACM/IEEE DAC*, 2006, pp. 57–62.
 - [26] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, “Statistical design and optimization of SRAM cell for yield enhancement,” in *IEEE/ACM, ICCAD 2004*, pp. 10–13.
 - [27] K. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari and S. Mudanai, “Process technology variation,” *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2197-2208, Aug. 2011.
 - [28] J. Zou, R. Wang, N. Gong, R. Huang, X. Xu, J. Ou, C. Liu, J. Wang, J. Liu, J. Wu, S. Yu, P. Ren, H. Wu, S. Lee, and Y. Wang, “New insights into AC RTN in scaled high-k/metal gate MOSFETs under digital circuit operation,” in *VLSI Symp. Tech. Dig.*, 2012, pp. 139-140.
 - [29] J. Zou, *et al.*, “Deep understanding of AC RTN in FinFETs though new characterization method and impacts on logic circuits,” in *VLSI Symp. Tech. Dig.*, 2013, pp. T186-T187.

- [30] H. M. Tsai, *et al.*, “The Understanding of the Trap Induced Variation in Bulk Tri-gate Devices by a Novel Random Trap Profiling (RTP) Technique,” in *VLSI Symp. Tech. Dig.*, 2012, pp. 189-190.
- [31] C. Liu, *et al.*, “Towards the systematic study of aging induced dynamic variability in nano-MOSFETs: adding the missing cycle-to-cycle variation effects into device-to-device variation,” in *IEDM Tech. Dig.*, 2011, pp. 571-574.
- [32] C. Liu, *et al.*, “New Observations on AC NBTI Induced Dynamic Variability in Scaled High-k/Metal-gate MOSFETs: Characterization, Origin of Frequency Dependence, and Impacts on Circuits,” in *IEDM Tech. Dig.*, 2012, pp. 466-469.
- [33] B. Kaczer, *et al.*, “Origin of NBTI variability in deeply scaled pFETs,” in *Proc. Int. Rel. Phys. Symp.*, 2010, pp. 26-32.
- [34] G. Ghibaudo, C. Mezzomo, A. Bajolet, A. Cathignol and R. DiFrenza, “Characterization and modeling of device variability in advanced CMOS technologies,” *SISPAD Workshop*, 2010.
- [35] M. Yamaoka, H. Miki, A. Bansal, S. Wu, D. J. Frank, E. Leobandung, and K. Torii, “Evaluation methodology for random telegraph noise effects in SRAM array,” in *IEDM Tech. Dig.*, 2011, pp. 745-748.
- [36] H. S. Momose, S. I. Nakamura, T. Ohguro, T. Yoshitomi, E. Morifuji, T. Morimoto, Y. Katsumata, and H. Iwai, “Study of the manufacturing feasibility of 1.5-nm direct-tunneling gate oxide MOSFETs: uniformity, reliability, and dopant penetration of the gate oxide,” *IEEE Trans. Electron Devices*, vol. 45, no. 3. pp. 691–700, 1998.
- [37] B. Kaczer, *et al.*, “Statistics of Multiple Trapped Charges in the Gate Oxide of Deeply Scaled MOSFET Devices—Application to NBTI,” *IEEE Electron Device Lett.*, vol. 31, no. 5, pp. 411-413, 2010.
- [38] L. Gerrer, *et al.*, “Interplay Between Statistical Reliability and Variability: A Comprehensive Transistor-to-Circuit Simulation Technology,” in *Proc. Int. Rel. Phys. Symp.*, 2013, pp. 3A.2.1-3A.2.5.
- [39] A. Ghetti, *et al.*, “Comprehensive Analysis of Random Telegraph Noise Instability and Its Scaling in Deca–Nanometer Flash Memories,” *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1746-1752, 2009.
- [40] K.P. Cheung, *et al.*, “On the Magnitude of Random Telegraph Noise in Ultra-Scaled MOSFETs,” in *ICICDT*, 2011.
- [41] M. Toledano-Luque, *et al.*, “Degradation of time dependent variability due to interface state generation,” in *VLSI Symp. Tech. Dig.*, 2013, pp. 190-191.
- [42] S. Kim, Y. Lee, H. D. Kim, and S. J. Choi, “Parallel weight update protocol for a carbon nanotube synaptic transistor array for accelerating neuromorphic computing,” *Nanoscale*, 12(3), pp. 2040-2046, 2020.
- [43] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, 521(7553), pp. 436-444, 2015.
- [44] J. Hasler, and H. B. Marr, “Finding a roadmap to achieve large neuromorphic hardware systems,” *Frontiers in neuroscience*, 7, 118, 2013.

- [45] D. S. Jeong, K. M. Kim, S. Kim, B. J. Choi, and C. S. Hwang, "Memristors for energy-efficient new computing paradigms," *Adv. Electron. Mater.*, 2(9), 1600090, 2016.
- [46] W. Xu, S. Y. Min, H. Hwang, and T. W. Lee, "Organic core-sheath nanowire artificial synapses with femtojoule energy consumption," *Science advances*, 2(6), e1501326, 2016.
- [47] L. F. Abbott, and W. G. Regehr, "Synaptic computation," *Nature*, 431(7010), pp. 796-803, 2004.
- [48] M. Prezioso, F. Merrih-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, "Training and Operation of an Integrated Neuromorphic Network based on Metal-Oxide Memristors," *Nature*, 521, pp. 61-64, 2015.
- [49] J. Park, M. Kwak, K. Moon, J. Woo, D. Lee, and H. Hwang, "TiO_x-Based RRAM Synapse With 64-Levels of Conductance and Symmetric Conductance Change by Adopting a Hybrid Pulse Scheme for Neuromorphic Computing," *IEEE Electron Device Lett.*, vol. 37, no. 12, pp. 1559–1562, 2016.
- [50] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale Memristor Device as Synapse In Neuromorphic Systems," *Nano Lett.*, 10 (4), pp. 1297–1301, 2010.
- [51] D. Kuzum, R. G. D. Jeyasingh, B. Lee, and H. S. P. Wong, "Nanoelectronic Programmable Synapses Based on Phase Change Materials for Brain-Inspired Computing," *Nano Lett.*, 12 (5), pp. 2179–2186, 2012.
- [52] M. Suri, O. Bichler, D. Querlioz, O. Cueto, L. Perniola, V. Sousa, D.; Gamrat, C. Vuillaume, and B. DeSalvo, "Phase-Change Memory as Synapse for Ultra-dense Neuromorphic Systems: Application to Complex Visual Pattern Extraction," in *IEDM Tech. Dig.*, 2011, pp. 4.4.1–4.4.4.
- [53] S. B. Eryilmaz, D. Kuzum, R. Jeyasingh, S. Kim, M. BrightSky, C. Lam, and H.-S. P. Wong, "Brain-Like Associative Learning Using a Nanoscale Non-volatile Phase Change Synaptic Device Array," *Front. Neurosci.*, 8, 205, 2014.
- [54] M. Wu et al., "Extremely Compact Integrate-and-Fire STT-MRAM Neuron: A Pathway toward All-Spin Artificial Deep Neural Network," in *VLSI Symp. Tech. Dig.*, 2019, pp. T34-T35.
- [55] M. Bavandpour et al., "Mixed-Signal Neuromorphic Inference Accelerators: Recent Results and Future Prospects," in *IEDM Tech. Dig.*, 2018, pp. 20.4.1-20.4.4.
- [56] H. Mulaosmanovic, J. Ocker, S. Müller, M. Noack, J. Müller, P. Polakowski, T. Mikolajick, and S. Slesazeck, "Novel Ferroelectric FET based Synapse for Neuromorphic Systems," in *VLSI Symp. Tech. Dig.*, 2017, pp. T176–T177.
- [57] S. Oh, T. Kim, M. Kwak, J. Song, J. Woo, S. Jeon, I. K. Yoo, and H. Hwang, "HfZrO_x-Based Ferroelectric Synapse Device With 32 Levels of Conductance States for Neuromorphic Applications," *IEEE Electron Device Lett.* Vol. 38, no. 6, pp. 732–735, 2017.
- [58] M. Jerry, P.-Y. Chen, J. Zhang, P. Sharma, K. Ni, S. Yu, and S. Datta, "Ferroelectric FET Analog Synapse for Acceleration of Deep Neural Network Training," in *IEDM Tech. Dig.*, 2017, pp. 6.2.1–6.2.4.

- [59] S. Kim, J. Yoon, H. D. Kim, and S. J. Choi, "Carbon nanotube synaptic transistor network for pattern recognition," *ACS Appl. Mater. Interfaces*, vol. 7, no. 45, pp. 25479-25486, 2015.
- [60] L. Zhu, C. Wan, L. Guo, Y. Shi, and Q. Wan, "Artificial synapse network on inorganic proton conductor for neuromorphic systems," *Nat. Commun.*, vol. 5, no. 1, pp. 1-7, 2014.
- [61] Y. Zang, H. Shen, D. Huang, C. A. Di, and D. Zhu, "A dual-organic-transistor-based tactile-perception system with signal-processing functionality," *Adv. Mater.*, vol. 29, 1606088, 2017.
- [62] J. Yu, L. Liang, L. Hu, H. Duan, W. Wu, H. Zhang, J. Gao, F. Zhuge, T. Chang, and H. Cao, "Optoelectronic Neuromorphic Thin-Film Transistors Capable of Selective Attention and with Ultra-Low Power Dissipation," *Nano Energy*, vol. 62, pp. 772-780, 2019.
- [63] D. K. Schroder, "Negative bias temperature instability: what do we understand?" *Microelectron. Reliab.*, vol. 47, no. 6, pp. 841-852, 2007.
- [64] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative-bias temperature instability," *IEEE Trans. Device & Mat. Reliab.*, vol. 8, no. 1, pp. 22-34, 2008.
- [65] M. Duan, J. F. Zhang, Z. Ji, W. D. Zhang, B. Kaczer, S. De Gendt, and G. Groeseneken, "New insights into defect loss, slowdown, and device lifetime enhancement," *IEEE Trans. Electron Devices*, vol. 60, no. 1, pp. 413-419, 2013.
- [66] D. S. Ang, Z. Q. Teo, T. J. J. Ho, and C. M. Ng, "Reassessing the mechanisms of negative-bias temperature instability by repetitive stress/relaxation experiments," *IEEE Trans. Device & Mat. Reliab.*, vol. 11, no. 1, pp. 19-34, 2011.
- [67] N. Kimizuka, K. Yamaguchi, K. Imai, T. Iizuka, C. T. Liu, R. C. Keller, and T. Horiuchi, "NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10- μ m gate CMOS generation," in *Tech. Dig. VLSI Symp.*, 2000, pp. 92-93.
- [68] K. Onishi, R. N. Choi, C. S. Kang, H. J. Cho, Y. H. Kim, R. E. Nigh, J. Han, S. A. Krishnan, M. S. Akbar, and J. C. Lee, "Bias-temperature instabilities of polysilicon gate HfO₂ MOSFETs," *IEEE Trans. Electron Devices*, vol. 50, no. 6, pp. 1517-1524, 2003.
- [69] K. Zhao, J. H. Stathis, A. Kerber, and E. Cartier, "PBTI relaxation dynamic after AC vs. DC stress in high-k/metal gate stacks", in *Proc. Int. Rel. Phys. Symp.*, 2010, pp. 50-54.
- [70] M. Cho, J.-D. Lee, M. Aoulaiche, B. Kaczer, P. Roussel, T. Kauerauf, R. Degraeve, J. Franco, L.-A. Ragnarsson, and G. Groeseneken "Insight Into N/PBTI Mechanisms in Sub-1-nm-EOT devices", *IEEE Trans. Electron Devices*, vol. 59, no. 8, pp. 2042-2048, 2012.
- [71] D. K. Schroder, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing", *Journal of Applied Physics*, vol. 94, no. 1, pp. 1-18, 2003.
- [72] J. H. Stathis, and S. Zafar, "The negative bias temperature instability in MOS devices: A review", *Microelectron. Reliab.*, vol. 46, no. 2/4, pp. 270-286, 2006.

- [73] T. Grassler, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, "The paradigm shift in understanding the bias temperature instability: from reaction-diffusion to switching oxide Traps", *IEEE Trans. Electron Devices*, vol. 58, no. 11, pp. 3652-3666, 2011.
- [74] K. Sutaria, A. Mohanty, R. Wang, R. Huang, and Y. Cao, "Aging statistics based on trapping/detrapping: compact modeling and silicon validation", *IEEE Trans. Dev. Mater. Reliab.*, vol.14, no.2, pp.607-615, 2014.
- [75] J. Bhaskarr Velamala, K. B. Sutaria, H. Shimizu, H. Awano, T. Sato, G. Wirth, and Y. Cao, "Compact modeling of statistical BTI under trapping/detrapping," *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3645-3654, 2013.
- [76] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI Degradation and Recovery-Behavior Based on Ultra-Fast VT-Measurements," in *Proc. Int. Rel. Phys. Symp.*, 2006, pp. 448-453.
- [77] G. A. Du, D. S. Ang, Z. Q. Teo, and Y. Z. Hu, "Ultrafast Measurement on NBTI," *IEEE Electron Device Lett.*, vol. 30, no. 3, pp. 275-277, 2009.
- [78] K. Zhao, J. H. Stathis, A. Kerber, and E. Cartier, "PBTI relaxation dynamic after AC vs. DC stress in high-k/metal gate stacks", in *Proc. Int. Rel. Phys. Symp.*, 2010, pp.50-54.
- [79] Z. Y. Tung, D. S. Ang, and Y. Gao, "On electron-trap transformation and its unexpected frequency dependence under dynamic positive-bias temperature stressing," in *Proc. Int. Rel. Phys. Symp.*, 2014, pp. XT.10.1-XT.10.6.
- [80] C. L. Chen, Y. M. Lin, C. Wang, and K. Wu, "A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2nm ultra-thin oxide," in *Proc. Int. Rel. Phys. Symp.*, 2005, pp. 704-705.
- [81] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability," in *Proc. Int. Rel. Phys. Symp.*, 2002, pp. 248-254.
- [82] M. Toledano-Luque, et al., "From mean values to distributions of BTI lifetime of deeply scaled FETs through atomistic understanding of the degradation", in *VLSI Symp. Tech. Dig.*, 2011, pp. 152-153.
- [83] J. Franco, B. Kaczer, M. Toledano-Luque, P. J. Roussel, J. Mitard, L.-A. Ragnarsson, L. Witters, T. Chiarella, M. Togo, N. Horiguchi, G. Groeseneken, M. F. Bukhori, T. Grassler, and A. Asenov, "Impact of single charged gate oxide defects on the performance and scaling of nanoscaled FETs," in *Proc. Int. Rel. Phys. Symp.*, 2012, pp. 5A.4.1-5A.4.6.
- [84] J. P. Chiu, et al., "A Comparative Study of NBTI and RTN Amplitude Distributions in High-κ Gate Dielectric pMOSFETs", *IEEE Electron Device Lett.*, vol. 33, no. 2, pp. 176-178, 2012.
- [85] M. Cho, et al., "Interface Trap Characterization of a 5.8-Å EOT p-MOSFET Using High-Frequency On-Chip Ring Oscillator Charge Pumping Technique," *IEEE Trans. Electron Devices*, vol. 58, no. 10, pp. 3342-3349, 2011.

- [86] T. Grassler, H. Reisinger, P. -. Wagner, F. Schanovsky, W. Goes and B. Kaczer, "The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability," in *Proc. Int. Rel. Phys. Symp.*, 2010, pp. 16-25.
- [87] T. Grassler, K. Rott, H. Reisinger, M. Walzl, J. Franco and B. Kaczer, "A unified perspective of RTN and BTI," in *Proc. Int. Rel. Phys. Symp.*, 2015, pp. 4A.5.1-4A.5.7.
- [88] T. Grassler, M. Walzl, W. Goes, Y. Wimmer, A.-M. El-Sayed, A. L. Shluger, and B. Kaczer, "On the volatility of oxide defects: Activation, deactivation, and transformation," in *Proc. Int. Rel. Phys. Symp.*, 2015, pp. 5A.3.1-5A.3.8.
- [89] T. Grassler, B. Kaczer, H. Reisinger, P.-J. Wagner, and M. Toledano-Luque, "On the frequency dependence of the bias temperature instability," in *Proc. Int. Rel. Phys. Symp.*, 2012, p. XT.8.1-XT.8.7.
- [90] T. Grassler, H. Reisinger, W. Goes, T. Aichinger, P. Hehenberger, P.-J. Wagner, M. Nelhiebel, J. Franco, and B. Kaczer, "Switching oxide traps as the missing link between negative bias temperature instability and random telegraph noise," in *IEDM Tech. Dig.*, 2009, pp. 1-4.
- [91] T. Grassler, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, "The paradigm shift in understanding the bias temperature instability: from Reaction-Diffusion to switching oxide traps," *IEEE Trans. Electron Devices*, vol. 58, no. 11, pp. 3652-3666, 2011.
- [92] P. M. Lenahan, J. P. Campbell, A. T. Krishnan, and S. Krishnan, "A Model for NBTI in Nitrided Oxide MOSFETs Which Does Not Involve Hydrogen or Diffusion," *IEEE Trans. Electron Devices Reliab.*, vol. 11, no. 2, pp. 219-226, 2011.
- [93] T. Grassler, W. Goes, Y. Wimmer, F. Schanovsky, G. Rzepa, M. Walzl, K. Rott, H. Reisinger, V. V Afanas'ev, A. Stesmans, A.-M. El-Sayed, and A. L. Shluger, "On the microscopic structure of hole traps in pMOSFETs," in *IEDM Tech. Dig.*, 2014, p. 21.1.1-21.1.4.
- [94] P. E. Blöchl, and J.H. Stathis, "Aspects of defects in silica related to dielectric breakdown of gate oxides in MOSFETs," *Phys. Condens. Matter*, no. 273-274, pp. 1022-1026, 1999.
- [95] Z. Y. Tung and D. S. Ang, "Transient to temporarily permanent and permanent hole trapping transformation in the small area SiON P-MOSFET subjected to negative-bias temperature stress," in *IPFA*, 2014, pp. 254-257.
- [96] A. M. El-Sayed, M. B. Watkins, V. V. Afanas'ev, and A. L. Shluger, "Nature of intrinsic and extrinsic electron trapping in SiO₂," *Phys. Rev. B*, no. 89, 2014.
- [97] M. Walzl, "Ultra-Low Noise Defect Probing Instrument for Defect Spectroscopy of MOS Transistors," *IEEE Trans. Electron Devices Reliab.*, vol. 20, no. 2, pp. 242-250, 2020.
- [98] T. Grassler, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectron. Reliab.*, vol. 52, no. 1, pp. 39-70, 2012.

- [99] M. Wautl, "Reliability of Miniaturized Transistors from the Perspective of Single-Defects," *Micromachines*, vol. 11, no. 736, 2020.
- [100] S. Ramey, Y. Lu, I. Meric, S. Mudanai, S. Novak, C. Prasad, and J. Hicks, "Aging model challenges in deeply scaled tri-gate technologies," in *IIRW*, 2015, pp. 56-62.
- [101] J. J. Ellis-Monaghan, R. B. Hulfachor, K. W. Kim, and M. A. Littlejohn, "Ensemble Monte Carlo study of interface-state generation in low-voltage scaled silicon MOS devices," *IEEE Trans. Electron. Devices*, vol. 43, no. 7, pp. 1123–1132, 1996.
- [102] A. Ghetti, L. Selmi, and R. Bez, "Low-voltage hot electrons and soft programming lifetime prediction in nonvolatile memory cells," *IEEE Trans. Electron. Devices*, vol. 46, no. 4, pp. 696–702, 1999.
- [103] A. Bravaix, V. Huard, D. Goguenheim, and E. Vincent, "Hotcarrier to cold-carrier device lifetime modeling with temperature for low power 40 nm Si-bulk NMOS and PMOS FETs," in *IEDM Tech. Dig.*, 2011, pp. 27.5.1–27.5.4.
- [104] A. Bravaix, C. Guerin, D. Goguenheim, V. Huard, D. Roy, C. Besset, S. Renard, Y.M. Randriamihaja, and E. Vincent, "Off State Incorporation into the 3 Energy Mode Device Lifetime Modeling for Advanced 40 nm CMOS Node," in *Proc. Int. Rel. Phys. Symp.*, 2010, pp. 55–64.
- [105] S. Tyaginov and T. Grasser, "Modeling of hot-carrier degradation: Physics and controversial issues," in *IIRW*, 2012, pp. 206-215.
- [106] B.S. Doyle, and K. R. Mistry, "The characterization of hot carrier damage in p-channel transistors," *IEEE Trans. Electron. Devices*, vol. 40, pp. 152–156, 1993.
- [107] E. Bury, A. Chasin, M. Vandemaele, S. Van Beek, J. Franco, B. Kaczer, and D. Linten, "Array-Based Statistical Characterization of CMOS Degradation Modes and Modeling of the Time-Dependent Variability Induced by Different Stress Patterns in the {VG, VD} bias space," in *Proc. Int. Rel. Phys. Symp.*, 2019, pp. 1–6.
- [108] N. L. Cohen, R. E. Paulsen, and M. H. White, "Observation and characterization of near-interface oxide traps with C-V techniques," *IEEE Trans. Electron. Devices*, vol. 42, pp. 2004–2009, 1995.
- [109] S. Tyaginov, I. Starkov, H. Enichlmair, J. Park, C. Jungemann, and T. Grasser, "Physics-Based Hot-Carrier Degradation Models," *ECS Trans.*, pp. 321–352, 2011.
- [110] S. Tyaginov, M. Bina, and J. Franco, et al. "Physical modeling of hot-carrier degradation for short-and long-channel MOSFETs", in *Proc. Int. Rel. Phys. Symp.*, 2014, pp. XT. 16.1-XT. 16.8.
- [111] A. Bravaix, C. Guerin, and V. Huard, et al. "Hot-carrier acceleration factors for low power management in DC-AC stressed 40 nm NMOS node at high temperature", in *Proc. Int. Rel. Phys. Symp.*, 2009, pp.531-548.
- [112] S. Tyaginov, M. Bina, J. Franco, Y. Wimmer, B. Kaczer, and T. Grasser, "On the importance of electron–electron scattering for hot-carrier degradation", *Jpn. J. Appl. Phys.*, vol. 54, no. 4S, pp. 04DC18-1-04DC18-6, 2015.

- [113] M. Duan et al., "Interaction between hot carrier aging and PBTI degradation in nMOSFETs: Characterization modelling and lifetime prediction", in *Proc. IEEE Int. Rel. Phys. Symp.*, pp. XT-5.1-XT-5.7, 2017.
- [114] M. Duan, J. F. Zhang, Z. Ji, W. D. Zhang, D. Vigar, A. Asenov, L. Gerrier, V. Chandra, R. Aitken, and B. Kaczer, "Insight Into Electron Traps and Their Energy Distribution Under Positive Bias Temperature Stress and Hot Carrier Aging", *IEEE Trans. Electron Devices*, vol. 63, no. 9, 2016.
- [115] M. Duan et al., "Key issues and solutions for characterizing hot carrier aging of nanometer scale nMOSFETs", *IEEE Trans. Electron Devices*, vol. 64, no. 6, pp. 2478-2484, 2017.
- [116] A. Bravaix et al., "Impact of the gate-stack change from 40 nm node SiON to 28 nm high-K metal gate on the hot-carrier and bias temperature damage," in *Proc. IEEE Int. Rel. Phys. Symp.*, 2013, pp. 2D.6.1–2D.6.9.
- [117] B. Ullmann, M. Jech, K. Puschkarsky, G. A. Rott, M. Walzl, Y. Illarionov, H. Reisinger, and T. Grasser, "Impact of Mixed Negative Bias Temperature Instability and Hot Carrier Stress on MOSFET Characteristics-Part I: Experimental", *IEEE Trans. Electron Devices*, vol. 66, no. 1, pp. 232-240, 2018.
- [118] "Keithley's Model 4200-SCS Semiconductor Characterization System Reference Manual," 2013.
- [119] "Keithley's Automated Characterization Suite (ACS) Reference Manual," 2015.
- [120] T. Grasser, K. Rott, H. Reisinger, P. -J. Wagner, W. Goes, F. Schanovsky, M. Walzl, M. Toledano-Luque, and B. Kaczer, "Advanced characterization of oxide traps: The dynamic time-dependent defect spectroscopy," in *Proc. IEEE Int. Reliab. Phys. Symp.*, 2013, pp. 2D.2.1-2D.2.7.
- [121] M. Walzl, W. Goes, K. Rott, H. Reisinger, and T. Grasser, "A single-trap study of PBTI in SiON nMOS transistors: Similarities and differences to the NBTI/pMOS case," in *Proc. Int. Reliab. Phys. Symp.*, 2014, pp. XT-18.1-XT-18.5.
- [122] S. F. Guo, P. P. Ren, R. S. Wang, Z. Q. Yu, M. L. Luo, X. Zhang, and R. Huang, "A new efficient method for characterizing time constants of switching oxide traps," in *Proc. Int. Reliab. Phys. Symp.*, 2014, pp. XT.14. 1-XT.14.4.
- [123] L. Gerrier, J. Ding, S. M. Amoroso, F. Adamu-Lema, R. Hussin, D. Reid, C. Millar, and A. Asenov, "Modelling RTN and BTI in nanoscale MOSFETs from device to circuit: A review," *Microelectron. Reliab.*, vol. 54, no. 4, pp. 682–697, 2014.
- [124] D. Y. Mao, S. F. Guo, R. S. Wang, M. L. Luo, and R. Huang, "Deep understanding of random telegraph noise (RTN) effects on SRAM stability," in *Proc. Int. Symp. VLSI Technol., Sys. & Appl.*, 2016.
- [125] R. Huang, X. B. Jiang, S. F. Guo, P. P. Ren, P. Hao, Z. Q. Yu, Z. Zhang, Y. Y. Wang, and R. S. Wang, "Variability-and-reliability-aware design for 16/14nm and beyond technology," in *IEDM Tech. Dig.*, 2017, pp. 12.4.1-12.4.4.
- [126] Z. Y. Tung and D. S. Ang, "Impact of voltage-accelerated stress on hole trapping at operating condition," *IEEE Electron Device Lett.*, vol. 37, no. 5, pp. 644-647, 2016.

- [127] Z. Y. Tung and D. S. Ang, "Alteration of oxide-trap switching activity at operating condition by voltage-accelerated stressing," in *Proc. Int. Reliab. Phys. Symp.*, 2016, pp. XT-01-1 - XT-01-3.
- [128] E. Amat, T. Kauerauf, R. Degraeve, R. Rodriguez, M. Nafria, X. Aymerich, and G. Groeseneken, "Competing degradation mechanisms in short-channel transistors under channel hot-carrier stress at elevated temperatures," *IEEE Trans. Device & Mat. Reliab.*, vol. 9, no. 3, pp. 454-458, 2009.
- [129] Z. Yu, J. Zhang, R. Wang, S. Guo, C. Liu, and R. Huang, "New insights into the hot carrier degradation (HCD) in FinFET: New observations unified compact model and impacts on circuit reliability," in *Tech. Dig. Int. Electron Devices Meet.*, 2017, pp. 7.2.1–7.2.4.
- [130] M. Wang, R. G. Southwick, K. Cheng, and J. H. Stathis, "Lateral profiling of HCI induced damage in ultra-scaled FinFET devices with I_d - V_d characteristics," in *Proc. Int. Reliab. Phys. Symp.*, 2018, pp. 6E.1-1-6E.1-6.
- [131] X. Ju and D. S. Ang, "Response of switching hole traps in the small-area P-MOSFET under channel hot-hole effect," in *Proc. Int. Reliab. Phys. Symp.*, 2019, pp. 1-4.
- [132] B. Ullmann, M. Jech, S. Tyaginov, M. Walzl, Y. Illarionov, A. Grill, K. Puschkarsky, H. Reisinger, and T. Grasser, "The impact of mixed negative bias temperature instability and hot carrier stress on single oxide defects," in *Proc. Int. Reliab. Phys. Symp.*, 2017, pp. XT-10.1-XT-10.6.
- [133] X. Ju and D. S. Ang, "Alteration of gate-oxide trap capture/emission time constants by channel hot-carrier effect in the metal-oxide-semiconductor field-effect transistor," *IEEE Access*, vol. 8, pp. 14048-14053, 2020.
- [134] P. E. Blochl, "First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen," *Phys. Rev. B*, vol. 62, no. 10, pp. 6158-6179, 2000.
- [135] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors," *Phys. Rev. B Condens. Matter*, vol. 82, no. 24, pp. 245318, 2010.
- [136] T. Grasser, M. Walzl, G. Rzepa, W. Goes, Y. Wimmer, A. M. El-Sayed, A. L. Shluger, H. Reisinger, and B. Kaczer, "The "permanent" component of NBTI revisited: Saturation, degradation-reversal, and annealing," in *Proc. Int. Reliab. Phys. Symp.*, 2016, pp. 5A.2.1-5A.2.8.
- [137] T. J. J. Ho, D. S. Ang, and K. C. Leong, "Increased deep-level hole trapping by combined negative-bias temperature and channel hot-hole stress," *IEEE Electron Device Lett.*, vol. 32, no. 10, pp. 1337-1339, 2011.
- [138] Y. P. Feng, A. T. L. Lim, and M. F. Li, "Negative-U property of oxygen vacancy in cubic HfO_2 ," *Appl. Phys. Lett.*, vol. 87, no. 6, art. no. 062105, 2005.
- [139] T. Grasser, K. Rott, H. Reisinger, M. Walzl, P. Wagner, F. Schanovsky, W. Goes, G. Pobegen, and B. Kaczer, "Hydrogen-related volatile defects as the possible cause for the recoverable component of NBTI," in *IEDM Tech. Dig.*, pp. 15.5.1-15.5.4, 2013.
- [140] D. J. DiMaria, "Defect generation in field-effect transistors under channel-hot-electron stress," *Journal of Appl. Phys.*, vol. 87, no. 12, pp. 8707-8715, 2000.

- [141] T. Grasser, K. Rott, H. Reisinger, M. Walzl, F. Schanovsky and B. Kaczer, “NBTI in Nanoscale MOSFETs—The Ultimate Modeling Benchmark,” in *IEEE Trans. Electron Devices*, vol. 61, no. 11, pp. 3586-3593, 2014.
- [142] J. Li, K. Kang, and K. Roy, “Novel variation-aware circuit design of scaled LTPS TFT for ultra-low power, low-cost applications,” in *Proc. Int. Conf. IC Design and Technol.*, 2007, pp. 1-4.
- [143] C. Schlünder, J. Berthold, F. Proebster, A. Martin, W. Gustin and H. Reisinger, “On the influence of BTI and HCI on parameter variability,” in *Proc. Int. Reliab. Phys. Symp.*, 2017, pp. 2E-4.1-2E-4.8.
- [144] C. Z. Zhao, M. B. Zahid, J. F. Zhang, G. Groeseneken, R. Degraeve, and S. De Gendt, “Threshold voltage instability of p-channel metal-oxide-semiconductor field effect transistors with hafnium based dielectrics,” *Appl. Phys. Lett.*, vol. 90, no. 14, 143502, 2007.
- [145] Y. Illarionov, M. Bina, S. Tyaginov, K. Rott, B. Kaczer, H. Reisinger, and T. Grasser, “Extraction of the Lateral Position of Border Traps in Nanoscale MOSFETs,” *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2730–2737, 2015.
- [146] K. Hess, I. C. Kizilyalli, and J. W. Lyding, “Giant isotope effect in hot electron degradation of metal oxide silicon devices,” *IEEE Trans. Electron Devices*, vol. 45, no. 2, pp. 406-416, 1998.
- [147] G. La Rosa, and S. E. Rauch III, “Channel hot carrier effects in n-MOSFET devices of advanced submicron CMOS technologies,” *Microelectron. Reliab.*, vol. 47, no. 4-5, pp. 552-558, 2007.
- [148] Z.-Y. Lu, C. J. Nicklaw, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, “Structure, properties, and dynamics of oxygen vacancies in amorphous SiO₂,” *Phys. Rev. Lett.*, vol. 89, no. 28, art. no. 285505, 2002.
- [149] T. Uchino and T. Yoko, “Density functional theory of structural transformations of oxygen-deficient centers in amorphous silica during hole trapping: Structure and formation mechanism of the E'_{γ} center,” *Phys. Rev. B*, vol. 74, no. 12, art. no. 125203, 2006.
- [150] G. Buscarino, S. Agnello, and F. M. Gelardi, “Characterization of E'_{δ} and triplet point defects in oxygen-deficient amorphous silicon dioxide,” *Phys. Rev. B*, vol. 73, no. 4, art. no. 045208, 2006.
- [151] Y. Wimmer, A. El-Sayed, W. Göss, T. Grasser, and A. L. Shluger, “Role of hydrogen in volatile behaviour of defects in SiO₂-based electronic devices,” *Proc. Royal Soc. A*, vol. 472, no. 2180, art. no. 20160009, 2016.
- [152] A. Yokozawa and Y. Miyamoto, “First-principles calculations for charged states of hydrogen atoms in SiO₂,” *Phys. Rev. B*, vol. 55, no. 20, pp. 13783-13788, 1997.
- [153] P. E. Blochl and J. H. Stathis, “Hydrogen electrochemistry and stress-induced leakage current in silica,” *Phys. Rev. Lett.*, vol. 83, no. 2, pp. 372-375, 1999.
- [154] W. Goes, Y. Wimmer, A. M. El-Sayed, G. Rzepa, M. Jech, A. L. Shluger, and T. Grasser, “Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence,” *Microelectron. Reliab.*, vol. 87, pp. 286-320, 2018.

- [155] Y. Mitani, M. Nagamine, H. Satake, and A. Toriumi, "NBTI mechanism in ultra-thin gate dielectric-nitrogen-originated mechanism in SiON," in *Tech. Dig. Int. Electron Devices Meet.*, 2002, pp. 509–512.
- [156] P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, and S. Krishnan, "Atomic-scale defects involved in the negative-bias temperature instability," *IEEE Trans. Mat. Dev. Reliab.*, vol. 7, no. 4, pp. 540–557, 2007.
- [157] C. J. Gu, D. S. Ang, and Z. Q. Teo, "Impact of twofold coordinated nitrogen on the generation of deep-level hole traps under negative-bias temperature stressing," *ECS Trans.*, vol. 35, no. 4, pp. 125–143, 2011.
- [158] J. D. Bude, "Gate current by impact ionization feedback in sub-micron MOSFET technologies," in *Tech. Dig. Symp. VLSI*, 1995, pp. 101–102.
- [159] D. S. Ang, H. Liao, and C. H. Ling, "Nonlocal hot-electron injection as the mechanism for the predominant source-side gate oxide degradation in CHE-stressed deep submicrometer n-MOSFETs," in *IEEE Electron Device Lett.*, vol. 25, no. 6, pp. 417–419, 2004.
- [160] C. Bonnelle, "Multiphonon-induced charge trapping-detrapping and damage in insulators," *Phys. Rev. B*, vol. 81, no. 5, art. no. 054307, 2010.
- [161] X. Liu, X. Xu, C. Gu, R. Gu, W. Wang, W. Liu, and T. Duan, "Investigating the impact of the defect dynamic characteristics on the PBTI in the high- κ gate device," *Microelectron. Reliab.*, vol. 80, pp. 24–28, 2018.
- [162] L. Vandelli, L. Larcher, D. Veksler, A. Padovani, G. Bersuker, and K. Matthews, "A charge-trapping model for the fast component of positive bias temperature instability (PBTI) in high- κ gate-stacks," in *IEEE Trans. Electron Devices*, vol. 61, no. 7, pp. 2287–2293, 2014.
- [163] G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," *J. Chem. Phys.*, vol. 113, no. 22, pp. 9901–9904, 2000.
- [164] C. Liu, K. T. Lee, S. Pae, and J. Park, "New observations on hot carrier induced dynamic variation in nano-scaled SiON/Poly HK/MG and FinFET devices based on on-the-fly HCI technique: The role of single trap induced degradation," in *IEDM Tech. Dig.*, 2014, pp. 34.6.1–34.6.4.
- [165] S. Tyaginov, M. Jech, J. Franco, P. Sharma, B. Kaczer, and T. Grassler, "Understanding and modeling the temperature behavior of hot-carrier degradation in SiON nMOSFETs," *IEEE Electron Device Lett.*, vol. 37, no. 1, pp. 84–87, 2016.
- [166] G. Pobegen and T. Grassler, "On the distribution of NBTI time constants on a long, temperature-accelerated time scale," *IEEE Trans. on Electron Devices*, vol. 60, no. 7, pp. 2248–2155, 2013.
- [167] C. J. Delerue, and M. Lannoo, "Nanostructures: theory and modeling," *Springer Science & Business Media*, 2013.
- [168] N. Clément, K. Nishiguchi, A. Fujiwara, and D. Vuillaume, "One-by-one trap activation in silicon nanowire transistors," *Nature communications*, vol. 1, no. 1, pp. 1–8, 2010.
- [169] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, "Discrete resistance switching in submicrometer silicon

- inversion layers: Individual interface traps and low-frequency ($1/f$) noise,” *Phys. Rev. Lett.*, vol. 52, no. 3, pp. 228-231, 1984.
- [170] G. Ghibaudo and T. Boutchacha, “Electrical noise and RTS fluctuations in advanced CMOS devices,” *Microelectron. Reliab.*, vol. 42, no. 4-5, pp. 573-582, 2002.
 - [171] A. P. van der Wel, E. A. Klumperink, E. Hoekstra, and B. Nauta, “Relating random telegraph signal noise in metal-oxide-semiconductor transistors to interface trap energy distribution,” *Appl. Phys. Lett.*, vol. 87, no. 18, 183507, 2005.
 - [172] S. Zafar, A. Callegari, E. Gusev, and M. V. Fischetti, “Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks,” *Journal of Appl. Phys.*, vol. 93, no. 11, pp. 9298-9303, 2003.
 - [173] W. H. Lo, T. C. Chang, J. Y. Tsai, C. H. Dai, C. E. Chen, S. H. Ho, H. M. Chen, O. Cheng, and C. T. Huang, “Charge trapping induced drain-induced-barrier-lowering in HfO_2/TiN p-channel metal-oxide-semiconductor-field-effect-transistors under hot carrier stress,” *Appl. Phys. Lett.*, vol. 100, no. 15, 152102, 2012.
 - [174] C. H. Dai, T. C. Chang, A. K. Chu, Y. J. Kuo, W. H. Lo, S. H. Ho, C. E. Chen, J. M. Shih, H. M. Chen, B. S. Dai, G. Xia, O. Cheng, and C. T. Huang, “Impact of static and dynamic stress on threshold voltage instability in high-k/metal gate n-channel metal-oxide-semiconductor field-effect transistors,” *Appl. Phys. Lett.*, vol. 98, no. 9, 092112, 2011.
 - [175] A. A. Boo, and D. S. Ang, “Evolution of hole trapping in the oxynitride gate p-MOSFET subjected to negative-bias temperature stressing,” *IEEE Trans. Electron Devices*, vol. 59, no. 11, pp. 3133-3136, 2012.
 - [176] M. Wang, R. Muralidhar, J. H. Stathis, B. P. Linder, H. Jagannathan, and J. Faltermeier, “Superior PBTI reliability for SOI FinFET technologies and its physical understanding,” *IEEE Electron Device Lett.*, vol. 34, no. 7, pp. 837-839, 2013.
 - [177] M. Yamaoka, H. Miki, A. Bansal, S. Wu, D. J. Frank, E. Leobandung, and K. Torii, “Evaluation methodology for random telegraph noise effects in SRAM arrays,” in *Tech. Dig. Int. Electron Devices Meet.*, pp. 32.2.1-32.2.4, 2011.
 - [178] M. Luo, R. Wang, S. Guo, J. Wang, J. Zou, and R. Huang, “Impacts of random telegraph noise (RTN) on digital circuits,” *IEEE Trans. Electron Devices*, vol. 62, no. 6, pp. 1725-1732, 2015.
 - [179] R. Brederlow, R. Prakash, C. Paulus, and R. Thewes, “A low-power true random number generator using random telegraph noise of single oxide-traps,” *IEEE ISSCC Dig. of Tech.*, pp. 1666-1675, 2006.
 - [180] J. Brown, R. Gao, Z. Ji, J. Chen, J. Wu, J. Zhang, B. Zhou, S. Qi, J. Crawford, and W. Zhang, “A low-power and high-speed True Random Number Generator using generated RTN,” *Symp. on VLSI Tech.*, pp. 95-96, 2018.
 - [181] P. S. Yeh, C. A. Yang, Y. H. Chang, Y. D. Chih, C. J. Lin, and Y. C. King, “Self-Convergent Trimming SRAM True Random Number Generation with In-Cell Storage,” *IEEE J. Solid-State Circuits*, vol. 54, no. 9, pp. 2614-2621, 2019.

- [182] F. Alibart, S. Pleutin, D. Guérin, C. Novembre, S. Lenfant, K. Lmimouni, C. Gamrat, and D. Vuillaume, "An organic nanoparticle transistor behaving as a biological spiking synapse," *Adv. Funct. Mater.*, vol. 20, no. 2, pp. 330-337, 2010.
- [183] H. Tian, Q. Guo, Y. Xie, H. Zhao, C. Li, J. J. Cha, F. Xia, and H. Wang, "Anisotropic black phosphorus synaptic device for neuromorphic applications," *Adv. Mater.*, vol. 28, no. 25, pp. 4991-4997, 2016.
- [184] D. Sarkar, J. Tao, W. Wang, Q. Lin, M. Yeung, C. Ren, and R. Kapadia, "Mimicking biological synaptic functionality with an indium phosphide synaptic device on silicon for scalable neuromorphic computing," *ACS nano*, vol. 12, no. 2, pp. 1656-1663, 2018.
- [185] I. Sanchez Esqueda, X. Yan, C. Rutherglen, A. Kane, T. Cain, P. Marsh, Q. Liu, K. Galatsis, H. Wang, and C. Zhou, "Aligned carbon nanotube synaptic transistors for large-scale neuromorphic computing," *ACS nano*, vol. 12, no. 7, pp. 7352-7361, 2018.
- [186] Y. Xiang, P. Huang, R. Han, Z. Zhou, Q. Shu, Z. Su, H. Hu, L. Liu, Y. Liu, X. Liu, and J. Kang, "Hardware Implementation of Energy Efficient Deep Learning Neural Network Based on Nanoscale Flash Computing Array," *Adv. Mater. Tech.*, vol. 4, no. 5, 1800720, 2019.
- [187] S. Nakanishi, "Molecular diversity of glutamate receptors and implications for brain function," *Science*, vol. 258, no. 5082, pp. 597-603, 1992.
- [188] G. Q. Bi, and M. M. Poo, "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type," *Journal of neuroscience*, vol. 18, no. 24, pp. 10464-10472, 1998.
- [189] R. S. Zucker, and W. G. Regehr, "Short-term synaptic plasticity," *Annual review of physiology*, vol. 64, no. 1, pp. 355-405, 2002.
- [190] G. Voglis, and N. Tavernarakis, "The role of synaptic ion channels in synaptic plasticity," *EMBO reports*, vol. 7, no. 11, pp. 1104-1110, 2006.
- [191] M. R. Azghadi, N. Iannella, S. F. Al-Sarawi, G. Indiveri, and D. Abbott, "Spike-based synaptic plasticity in silicon: design, implementation, application, and challenges," *Proc. of the IEEE*, vol. 102, no. 5, pp. 717-737, 2014.
- [192] E. Farquhar, and P. Hasler, "A bio-physically inspired silicon neuron," *Int. Symp. on Circuits & Systems*, 2004, pp. 309-312.
- [193] G. Indiveri, Linares-Barranco, Bernabé, T. J. Hamilton, Schaik, André van, R. Etienne-Cummings, and T. Delbruck, et al, "Neuromorphic silicon neuron circuits," *Frontiers in Neuroscience*, vol. 5, no. 73, 2011
- [194] S. Kim, C. Du, P. Sheridan, W. Ma, S. H. Choi, and W. Lu, "Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity," *Nano Lett.*, vol. 15, no. 3, pp. 2203-2211, 2015.
- [195] Z. Wang, S. Joshi, Sergey E. Savel'ev, H. Jiang, R. Midya, and P. Lin, et al., "Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing," *Nat. Mater.*, 2016.
- [196] J. Yin, F. Zeng, Q. Wan, F. Li, Y. Sun, Y. Hu, J. Liu, G. Li, and F. Pan, "Adaptive Crystallite Kinetics in Homogenous Bilayer Oxide Memristor for Emulating Diverse Synaptic Plasticity," *Adv. Funct. Mater.*, vol. 28, 1706927, 2018.

- [197] W. Xiong, L. Zhu, C. Ye, F. Yu, Z. Ren, and Z. Ge, "Bilayered Oxide-Based Cognitive Memristor with Brain-Inspired Learning Activities," *Adv. Electron. Mater.*, 1900439, 2019.
- [198] C. D. Wright, P. Hosseini, and J. A. V. Diodado, "Beyond von Neumann computing with nanoscale phase-change memory devices," *Adv. Funct. Mater.*, vol. 23, no. 18, pp. 2248-2254, 2013.
- [199] A. Sebastian, T. Tuma, N. Papandreou, M. Le Gallo, L. Kull, and T. Parnell, et al., "Temporal correlation detection using computational phase-change memory," *Nat. Commun.*, 1115, 2017.
- [200] Z. Wang, W. Zhao, W. Kang, Y. Zhang, and C. Chappert, "Ferroelectric Tunnel Memristor-based Neuromorphic Network with 1T1R Crossbar Architecture," *IEEE IJCNN*, pp. 6-11, 2014.
- [201] R. Guo, Y. Zhou, L. Wu, Z. R. Wang, Z. Lim, and X. Yan, et al., "Control of synaptic plasticity learning of ferroelectric tunnel memristor by nano-scale interface engineering," *ACS Appl. Mater. Interfaces*, vol. 10, no. 15, pp. 12862-12869, 2018.
- [202] M. Ziegler, and H. Kohlstedt, "Mimic synaptic behavior with a single floating gate transistor: A MemFlash synapse," *Journal of Applied Physics*, vol. 114, no. 19, pp. 194506, 2013.
- [203] L. Danial, E. Pikhay, E. Herbelin, N. Wainstein, V. Gupta, N. Wald, Y. Roizin, R. Daniel, and S. Kvatinsky, "Two-terminal floating-gate transistors with a low-power memristive operation mode for analogue neuromorphic computing," *Nature Electronics*, vol. 2, no. 12, pp. 596-605, 2019.
- [204] X. Gu and S. S. Iyer, "Unsupervised Learning Using Charge-Trap Transistors," *IEEE Electron Device Lett.*, vol. 38, no. 9, pp. 1204-1207, 2017.
- [205] M. C. Lemme, J. K. Efavi, T. Mollenhauer, M. Schmidt, H. D. B. Gottlob, T. Wahlbrink, and H. Kurz, "Nanoscale TiN metal gate technology for CMOS integration," *Microelectronic Engineering*, vol. 83, no. 4-9, pp. 1551-1554, 2006.
- [206] D. G. Seo, Y. Lee, G. T. Go, M. Pei, S. Jung, Y. H. Jeong, and T. W. Lee, "Versatile neuromorphic electronics by modulating synaptic decay of single organic synaptic transistor: From artificial neural networks to neuro-prosthetics," *Nano Energy*, vol. 65, 104035, 2019.
- [207] A. J. Arnold, A. Razavieh, J. R. Nasr, D. S. Schulman, C. M. Eichfeld, and S. Das, "Mimicking neurotransmitter release in chemical synapses via hysteresis engineering in MoS₂ transistors," *ACS nano*, vol. 11, no. 3, pp. 3110-3118, 2017.
- [208] P.-Y. Chen, X. Peng, and S. Yu, "NeuroSim+: An integrated device-to-algorithm framework for benchmarking synaptic devices and array architectures," in *Tech. Dig. Int. Electron Devices Meet.*, pp. 6-1, 2017.
- [209] S. Wu, K. Y. Wong, and M. Tsodyks, "Neural information processing with dynamical synapses," *Frontiers in computational neuroscience*, vol. 7, pp. 188, 2013.
- [210] J. T. Yang, G. Chen, J. Y. Du, H. Y. Huang, H. Meng, H. B. Liu, G. Z. Yang, and K. J. Jin, "Artificial synapses emulated by an electrolyte-gated tungsten-oxide transistor," *Adv. Mater.*, vol. 30, 1801548, 2018.

- [211] G. F. Striedter, “Neurobiology: A Functional Approach,” *Oxford University Press*, 2015.
- [212] W. C. Abraham, “Metaplasticity: tuning synapses and networks for plasticity,” *Nature Reviews Neuroscience*, vol. 9, no. 5, pp. 387-387, 2008.
- [213] D. O. Hebb, “The Organization of Behavior: A Neuropsychological Approach,” *John Wiley & Sons: Hoboken*, 1949.
- [214] R. C. Froemke, and Y. Dan, “Spike-timing-dependent synaptic modification induced by natural spike trains,” *Nature*, vol. 416, no. 6879, pp. 433-438, 2002.
- [215] S. Song, K. D. Miller, and L. F. Abbott, “Competitive Hebbian learning through spike-timing-dependent synaptic plasticity,” *Nature neuroscience*, vol. 3, no. 9, pp. 919-926, 2000.
- [216] Y. Li, Y. Zhong, L. Xu, J. Zhang, X. Xu, H. Sun, and X. Miao, “Ultrafast synaptic events in a chalcogenide memristor,” *Scientific reports*, vol. 3, 1619, 2013.
- [217] D. Kuzum, S. Yu, and H. P. Wong, “Synaptic electronics: materials, devices and applications,” *Nanotechnology*, vol. 24, no. 38, 382001, 2013.
- [218] P. Rainville, “Distributed representation of nociception underlying the experience of pain,” *Curr. Opin. Neurobiol.*, vol. 12, pp. 195-204, 2002.
- [219] C. Villemure, and C. M. Bushnell, “Cognitive modulation of pain: how do attention and emotion influence pain processing?” *Pain*, vol. 95, no. 3, pp. 195-199, 2002.