N° Ordre...../Institut/UMBB/2020

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE M'HAMED BOUGARA-BOUMERDES

Institut de **G**énie **E**lectrique et **E**lectronique

Thèse de Doctorat

Présentée par :

CHENOUF Amel

En vue de l'obtention du diplôme de **DOCTORAT** en **SCIENCES**

Filière : Génie électrique

Option : Microélectronique

RELIABILITY AWARE DESIGN OF INTEGRATED CIRCUITS

Soutenue le : 08 Avril 2021

Devant le jury composé de :

M. RAHMOUNE	Fayçal	Professeur	UMBB	Président
M. BENTARZI	Hamid	Professeur	UMBB	Directeur de thèse
M. DJEZZAR	Boualem	Maitre de recherche A	CDTA	Co-directeur de thèse
M. SLIMANE	Abdelhalim	Directeur de recherche	CDTA	Examinateur
M. BENNAAMANE	Kamel	Maitre de conférences A	UMMTO	Examinateur

2020/2021

AKNOWLEDGEMENT

Then all the praises and thanks are to ALLAH, lord of the worlds, for his blessings and mercy! Without ALLAH willing, this PhD thesis would never have been achieved!

This PhD research work has been carried out at Microelectronics & Nanotechnology Division of Centre de Développement des Technologies Avancées; CDTA in collaboration with Institut de Genie Electrique et Electronique; IGEE of Universite M'hamed BOUGARA Boumerdès UMBB.

I would like to express my deep gratitude to a number of people who supported me during my work, in scientific, and in personal form.

First of all, I want to thank Prof. Hamid Bentarzi for giving me the opportunity to work on a PhD theme of my choice, for his guidance, support, and for his empathy.

I want to thank my thesis committee members: Prof. Fayçal RAHMOUNE, Dr. Abdelhalim SLIMANE, and Dr. Kamel BENNAAMANE for devoting time to evaluate this research work and writing feedback on my PhD thesis.

I am thankful to the editors and reviewers of both IEEE TDRM and IET Circuits and Systems journals for their peer reviewing of my submitted papers, and whose constructive remarks and recommendations allowed us to further discuss the obtained results and enrich our papers to get them published.

I cannot thank enough Dr. Boualem Djezzar for his scientific support, proof reading my papers and PhD thesis. I am thankful to him for encouraging me to achieve a goal I thought unreachable! I owe him respect for the high standards, he established in FCS team, for conducting research work!

I am fully grateful to the whole FCS research team: Without their scientific discussions, technical assistance, this work would not have been possible! I would like to thank, in particular, Mr. Abdelmadjid Benabdelmoumene and Mr. Boumediene Zatout who were always there, for helping and encouraging me. We can see that we are in brotherhood!

I want to thank all my colleagues and friends in CDTA who were of a great support for me, who stood by me, and who, when I felt discouraged and losing enthusiasm, came to push me! I would like to thank in particular Karima Ait Saadi and her husband Redha Hamoudi who came always to help me, prayed for me. "How can I help you to get your thesis done" is my favorite encouraging and compassion words I heard from them!

Amel Azzouz with whom I start my journeys, in CDTA, around a good coffee, and leave to go working in my office with her best wishes of luck and success!

Assia Smaali & Slimane Lafane, Farida Hamadi for their help, encouragements, shared concerns, and for the all great journeys we spent together in Algiers!

Fayçal Hadj-Larbi, Hakim Tahi, Mohamed Boubaaya, Abderrazak Smatti, Ahmed Liacha, Rafik Serhane, Noureddine Derguini, Abdelkrim Kamel Oudjida, Messaoud Bengherabi, Khadidija Messaoudi, Khalida Ghanem, Hadjira Haridi, Ratiba Touati, Kahina Mahdi, Yasmina Belaroussi, Majda Kahli, Nadjet Bouhlel, Safia Ben Ahmed, Maouïa Bentlemsan, Amina Ghazli, Samia Fouzar, Souhila Mazighi, Samia Aggoune, Karima, Annou, Amel Bounekta for their encouragements.

I am grateful to my friends and mates who supported me, encouraged me, and shared with me the stressful time I went through along my PhD studies and for the funny time we spent together in particular:

Fouzia Guessab, the person who shared with me joy and sadness over more than 13 years! And who was concerned by my thesis achievement more than anyone else!

Nadjat Hamdaoui, who cared for me like my little sister! Who stood for me, preparing meals, when I was overwhelmed! And who kept awake to encourage me, and to print the whole thesis copies.

Fatima Zohra Bouchibane, who softened my hard moments by making jokes, who is the best person to ask for preparing tea to keep awake working overnight!

Mounira Bitchikh, who kept advising me on the way to get things done, how to manage time to be more efficient at work!

Leila Souadkia, for her support and wise advices! For her believe in me. “You can do it Amel, you deserve it” were always her words that encouraged me a lot!

Nabila Dib, who never stopped supporting me, her availability and that of her family to help me along theses years in Algiers, cannot be thanked enough!

Above all, I want to pay tribute to my family whom I owe them all my achievements! My parents: who raised me with unconditional love! My father, who passed away in 2009 “Allah Yerhamou” was always proud of me. May ALLAH grant him Paradise! My mother, may ALLAH bless her, who brought me through the hard times, and keeps me always in her thoughts and prayers!

My brothers and sisters, who never stopped taking care of me! Caring about my career! And who were always there to support and encourage me to pursue my PhD studies!

The list of people to thank won’t end if I were to mention all of them here. So, I want to thank all people who I have not mentioned, here, for inspiring, helping, and praying for me!

الملخص

مهدت التطورات المذهلة في مجالي تصميم الدارات المدمجة بمساعدة الحاسوب وتقنيات التصنيع الطريق لهندسة وتصنيع دارات معقدة للغاية وعالية الأداء، تدمج أكثر من 100 مليون ترانزستور في رقاقة سيليكون مساحتها بضع مليمترات مربعة . إلا أن هذه الكثافة العالية جلبت معها المزيد من التحديات لمصممي هذه الدارات المدمجة من منظور ضمان موثوقيتها.
في الواقع ، نظرًا للتصغير الشديد للدارات ، وبسبب تأثيرات الإهتراء، تتغير خصائص مكوناتها الشبه الناقلة بمرور الوقت ، مما يتسبب في فشلها في الايفاء بالمواصفات التي صممت من أجلها.
بيد أن الحل القائم على التكنولوجيا ليس دائمًا ممكنًا ، ويرجع ذلك أساسًا إلى أن مهندسي عناصر أشباه النواقل يركزون عادةً على تطوير ترانزستورات أصغر وأسرع وأقل استهلاكًا للطاقة. هذا يجبر المصممين على تخفيف هذا التدهور لضمان الأداء المنوط بداراتهم. وبالتالي ، تصبح محاكاة الشيخوخة ضرورية للتنبؤ بتدهور أداء الدارات المدمجة بسبب التغيرات الزمنية. علاوة على ذلك ، فإن إدخال تقنيات تصميم جديدة و التي تعتبر الموثوقية كشرط لا يقل أهمية عن السرعة والمساحة واستهلاك الطاقة ، أصبحت أكثر من ضرورية لضمان تقديم دارات وأنظمة موثوقة وفقًا لمفهوم التصميم من أجل الموثوقية.
في هذا الصدد ، نقترح ، من ناحية، نقل تحليل الموثوقية من مستوى الترانسستور إلى مستوى أعلى من التجريد. هذا يسمح بتقييم أفضل للتدهور المستحث في أداء الدارات. ومن ناحية أخرى ، نقدم تقنية جديدة لتصميم دارات موثوقة.
بالنسبة لدراسة الحالة لأطروحة الدكتوراه هذه ، إخترنا التعامل مع ،أحد أكثر آليات الإهتراء التي تتسبب في الشيخوخة المبكرة للدارات عالية الدمج، ألا و هو عدم الإستقرار الناجم عن الإسقتطاب السلبي المتزامن مع الحرارة المرتعة في الدارات.
نقدم نتائج تحليل أثره على مستوى الدارة مباشرة ونقدم نموذجا له يمكن إدماجه في جهاز محاكاة تجاري.
ومن ناحية أخرى ، نقدم منهجية لتصميم دارات مدمجة تأخذ بعين الإعتبار هذه المشكل وتستند على الإختيار الأمثل لحجم الترانزستور من أجل إستحداث خلايا ذاكرة للحاسوب ذات مناعة أكبر للإهتراء.

الكلمات المفتاحية:
تصميم الدارات المدمجة، الإهتراء في أنصاف النواقل، الشيخوخة المبكرة للدارات، العمر الإفتراضي للدارات المدمجة
المحاكات الإلكترونية، تصميم الدارات المدمجة من أجل الموثوقية.

RÉSUMÉ

Les progrès impressionnants de la conception assistée par ordinateur et des procédés de fabrication de circuits intégrés ont permis de produire de puces très complexes, hautes performances intégrant plus de 100 millions de transistors dans quelques millimètres carrés de silicium. Cependant, cette densité élevée a entraîné plus de défis pour les concepteurs de circuits intégrés en termes de garantie de fiabilité de leurs circuits.

En fait, en raison de cette forte intégration, et des effets d'usure, les paramètres électriques des dispositifs à semi-conducteurs dérivent plus rapidement avec le temps, ce qui fait que les circuits intégrés échouent à satisfaire les spécifications pour lesquelles ils ont été conçus.

Or une solution basée sur la technologie n'est pas toujours réalisable, du fait que les industriels des semi-conducteurs se concentrent plus sur le développement de transistors plus petits, plus rapides et moins gourmands en énergie. Cela contraint les concepteurs à modérer cette dégradation et à améliorer la durée de vie de leurs circuits durant la phase de conception. Une simulation du vieillissement s'avère alors indispensable pour prédire la dégradation des performances des circuits intégrés due aux variations temporelles. De plus, l'introduction de nouvelles techniques de conception qui considèrent la fiabilité comme une contrainte de conception aussi importante que la vitesse, la surface et la consommation d'énergie, devient plus que nécessaire pour garantir la délivrance de circuits et de systèmes fiables conformément au concept de conception en vue de la fiabilité (DFR). Dans cette perspective, nous proposons, d'une part, de migrer l'analyse de fiabilité du niveau dispositif vers un niveau d'abstraction supérieur. Cela permettrait une meilleure évaluation de la dégradation induite sur les performances des circuits. D'autre part, nous proposons une approche DFR pour concevoir des circuits fiables.

Pour cette de doctorat, nous avons choisi de traiter l`instabilité due à la polarisation négative à haute température dite NBTI. NBTI est le mécanisme d'usure réduisant le plus la durée de vie des circuits intégrés submicroniques profonds. Nous présentons nos résultats de caractérisation au niveau du circuit NBTI, la mise en œuvre d`une interface de simulation intégrant notre modèle NBTI sur un simulateur commercial. D'autre part, nous présentons une approche de mitigation d`NBTI basée sur le dimensionnement des transistors que nous proposons pour la conception de cellules 6T-SRAM robustes.

Mots clés:

Conception de circuits intégrés, fiabilité des CIs. Vieillissement des dispositifs semi-conducteurs, mécanismes d'usure, NBTI, simulation de fiabilité, NBTI. Conception de circuits intégrés en vue de la fiabilité.

ABSTRACT

The striking advances in both computer-aided integrated circuit design and manufacturing technologies have paved the way for designing and manufacturing highly complex, high-performance chips integrating more than 100 million transistors into a few square millimeters of silicon. However, this high density has brought with it more challenges for IC designers in terms of their circuits reliability sign-off.

In fact, due to the aggressive scaling, and to wear out effects, the electrical parameters of semiconductor devices are shifting over time, causing for ICs the failure to meet the specifications for which they were designed. However, a technology-based solution is not always feasible, mainly because semiconductor engineers usually focus on developing smaller, faster, and less energy-intensive transistors. This compels designers to moderate this degradation and to improve the lifetime of their circuits during the design phase. A simulation of aging becomes therefore essential to predict the performance degradation of the ICs due to temporal variations. Moreover, the introduction of new design techniques which consider reliability as a design constraint as important as speed, area, and power consumption, becomes more than necessary to warranty delivering reliable circuits and systems by adopting design for reliability (DFR) concept. In this prospect, we propose, on one hand, to migrate reliability analysis from device-level to a higher level of abstraction. This allows a better assessment of the induced degradation on the circuits' performance. On the other hand, we propose a DFR approach to deign reliable circuits.

For this PhD thesis we choose, to deal with NBTI, which is one of the most wear-out mechanisms shrinking the lifetime of deep submicron ICs. We present our NBTI circuit-level characterisation results, the implementation of our NBTI model on a commercial simulator. On the other hand, we present an NBTI mitigation approach based on transistor sizing we propose for designing robust 6T-SRAM cells.

Keywords:

IC design, Semiconductors Reliability, wear-out mechanisms, NBTI, reliability simulation, NBTI mitigation. IC Design for Reliability

LIST of ACRONYMS

ABB	Adaptive Body Biasing
AHT	As-grown Hole Traps
ANPC	Anti-Neutralization Positive Charge
BEOL	Back-End Of Life
BERT	BErkeley Reliability Tools
BSIM	Berkeley Short-channel IGFET Model
BTI	Bias Temperature Instability
CAD	Computer Aided Design
CHC	Channel Hot Carrier
CMOS	Complementary Metal Oxide Semiconductor
CPC	Cyclic Positive Charge
CPU	Central Processing Unit
CR	Cell Ratio
CUT	Circuit Under Test
DAC	Digital Analog Converter
DFR	Design For Reliability
DIBL	Drain-Induced Barrier Lowering
DLHT	Deep-Level Hole Traps
DRAM	Dynamic Random Access Memory
DVFS	Dynamic Voltage and Frequency Scaling
DVS	Dynamic Voltage Scaling
EDA	Electronic Design Automation
EM	Electro Migration
FBB	Forward Body Bias
FEOL	Front-End Of Line
FF	Fast nMOS and Fast pMOS
FS	Fast nMOS and Slow pMOS
GAA	Gate-All-Around
HCI	Hot Carrier Injection
HKMG	High-K Metal Gate
HSPICE	Hailey Simulated Program with Integrated Circuit Emphasis
IBM	International Business Machines Corporation
IC	Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
IGFET	Insulated-Gate Field-Effect Transistor
IoT	Internet of Things
ISCAS'85	1985 International Symposium on Circuits & Systems
ISCAS'89	1989 International Symposium on Circuits & Systems
ISiT	Institute for Silicon Technology of Fraunhofer
ITC'99	1999 International Test Conference
ITRS	International Technology Roadmap for Semiconductors
LDD	Lightly Doped Drain
LER	Line Edge Roughness

LIST of ACRONYMS

LWR	Line Width Roughness
MCNC'91	1991 Microelectronics Center of North Carolina
MFCP	Multi Frequency Charge Pumping
MNOS	Metal Nitride Oxide Semiconductor
MPE	Multi-Phonon Emission
MPFAT	Multi-Phonon-Field-Assisted Tunneling
MPW	Multi Project Wafer
MOS	Metal Oxide Semiconductor
MOSFET	Metal Oxide Field Effect Transistor
MRAM	Magnetic RAM
MTCMOS	Multi Threshold CMOS
N_{IT}	Interface Traps
NMH	Noise Margin High
NML	Noise Margin Low
nMOS	n-type MOSFET
N(P)BTI	Negative (Positive) BTI
Op-Amp	Operational Amplifier
OPC	Optical Proximity Correction
OTFM	On-The-Fly Measurement
PDN	Pull-Down Network
PCRAM	Phase-Change Random Access Memory
pMOS	p-type MOSFET
PNO	Plasma Nitride Oxide
PPAC	Performance Power Area Cost
PR	Pull-up Ratio
PTM	Predictive Technology Model
PUN	Pull-Up Network
PVT	Process Voltage Temperature
RAM	Random Access Memory
RBB	Reverse Body Bias
R-D	Reaction-Diffusion
RDF	Random Dopant Fluctuation
ReRAM	Resistive RAM
RTN	Random Telegraph Noise
SCE	Short Channel Effect
SF	Slow nMOS and Fast Pmos
SILC	Stress Induced Leakage Current
SINM	Static current Noise Margin
SiON	Silicon Oxynitride
SNM	Static Noise Margin
SRAM	Static Random Access Memory
SS	Slow nMOS and Slow pMOS
STA	Static Timing Analysis
STT-MRAM	Spin-Torque Transfer Magnetic RAM
SVNM	Static Voltage Noise Margin
TCAD	Technology Computer Aided Design

LIST of ACRONYMS

T-D	Trapping -Ditrapping
TDDB	Time-Dependent Dielectric Breakdown
TDDS	Time-Dependent Defect Spectroscopy
TNO	Thermal Nitride Oxide
TSMC	Taiwan Semiconductor Manufacturing Company
TTL	Transistor-Transistor Logic
UDVS	Ultra Dynamic Voltage Scaling
UFM	Ultra-Fast Measurement
ULSI	Ultra Large Scale Integration
URI	Unified Reliability Interface
VFB	Flat Band Voltage
VIH	Input Voltage High
VIL	Input Voltage Low
Vinv	Inverter logic threshold
VOH	Output Voltage High
VOL	Output Voltage Low
VTC	Voltage Transfer Curve
VTH	Threshold Voltage
WNM	Write Noise Margin
WRM	WRite Margin
WTI	Write Trip current
WTV	Write Trip Voltage
WWL	Write WordLine

Figures List

Figures List

CONTENTS

CHAPTER 1 INTRODUCTION

The striking advance in both integrated circuit (IC) design flows and manufacturing processes have led to design very complex chips compacting more than 100 million of transistors on a few square millimeter area running at very high speed (see Fig. 1.1) [1]

Yet, this advent in the era of ultra large scale integration (ULSI) has brought new challenges to IC designers, in addition to the design perspective, validation, test, and reliability aspects have been taken into consideration. In fact, one of the major medium to long-term design issues, which have been reported by the ITRS for IC design, is their increasing variability [2].

IC performance variability can be due to the manufacturing process (lot-to-lot, die-to die, intra-die variations (spatial variation), to the environment, and to the intrinsic nature of the devices themselves and their wear-out (temporal variation) [3].

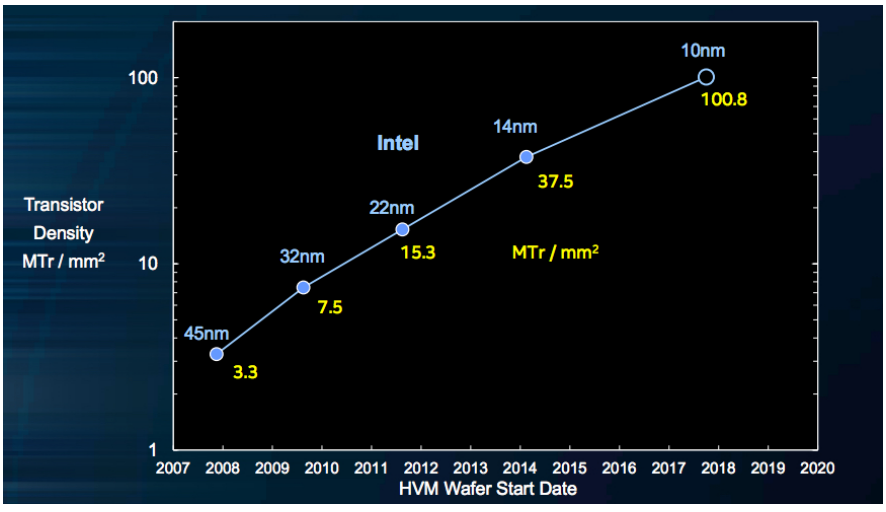


Fig. 1.1: Logic transistor density in Intel chips [1]

Governed by technology scaling, the important shrinking of the gate oxide makes circuits' devices to undergo a high vertical electric field. Moreover, the large scale integration has resulted

in a considerable heating of modern chips operating at high frequencies. Therefore, the high electric field and/or elevated temperature are exacerbating circuits' aging.

Due to wear-out effects, MOS devices' parameters shift, over time, altering thereby the performance and stability of the circuits integrating them, and which causes their failures to meet the specification targets for which they have been designed. Consequently, deep submicron CMOS ICs are prone to accelerate aging and make their lifetime shorter (see Fig. 1.2) [4]. Therefore, aging is among the critical reliability issues which faces present and future deeply downscaled CMOS devices, and the major aging concern is the so-called Bias Temperature Instability (BTI) [5].

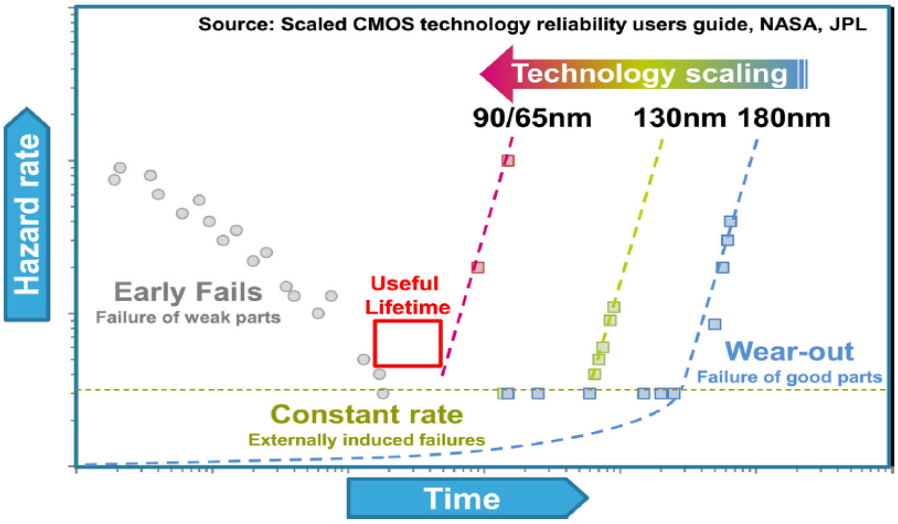


Fig. 1.2: Lifetime shrinking with technology scaling [4]

So far, lifetime reliability analysis has been mainly confined within the community of reliability physics at device level using accelerated stresses of single transistors. Failure criterions have been usually chosen as a shift of 10 % of I_{DS} or a shift of 30-50 mV of the transistor threshold voltage V_{th}. However, analyzing the device aging impact on circuit lifetime has not been well studied. Consequently, IC designers have been obliged to introduce large design margins called guard bands to avoid timing violation due to process and temporal reliability variations. Frequency reduction, supply voltage over-rating and transistors over sizing represent typical guard band concepts that designers adopt during the IC design step to mitigate aging effect appearing once the circuits are fabricated and during their lifetime [6]. However, Guard bands can be very pessimistic as they are based on the worst-case degradation and consequently

result in important overheads in silicon area and power consumption [6] [7]. Henceforth, addressing reliability issues during the early stages of IC design becomes more than an option to design reliable circuits including unreliable devices. Such design approach is known as Design For Reliability « DFR » (and Design-in-reliability) where reliability is considered as a design constraint as important as performance, area, and power consumption.

DFR is twofold research objective. One deals with the analysis/assessment of the aging-induced degradation on the circuit performance reliability while the other with combating/alleviating the predictable degradation and enhances the useful lifetime of the circuit (see Fig. 1.3) [4].

Hence, the research work conducted in this PhD thesis has been motivated by the willing to contribute to the reliability aware IC design. For this aim, we propose, on one hand, to migrate reliability analysis from device- to circuit-level to better assess the reliability degradation of circuits performance due to the aging of their device. On the other hand, we propose a circuit-level aging mitigation technique.

The circuit-level reliability study has been conducted through both experimental characterisation and simulation. In this latter, we have implemented an At^n model of NBTI degradation in an industrial simulator using a unified reliability interface (URI).

The proposed circuit-level design-aware approach is based on transistor sizing technique to ensure SRAM cell reliability during its useful lifetime.

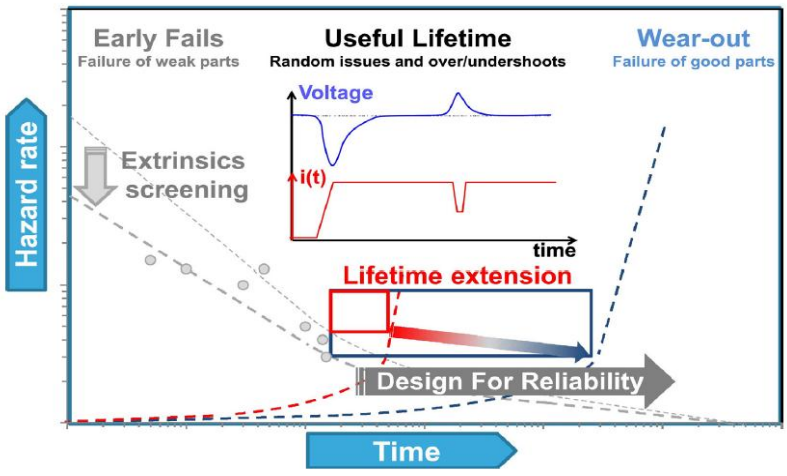


Fig. 1.3: Design for Reliability enhancing the circuits useful lifetime [4]

CHAPTER 1 INTRODUCTION

1.1 Scope

To better understand the PhD project research work, two main points have been taken in consideration:

- Among the multiple variations and reliability issues to which are prone deep submicron ICs, we are interested in studying aging effects due to negative bias temperature instability (NBTI).
- Among the different types of integrated circuits (digital analog and mixed), we are interested in digital circuits and in particular in memory structures.

Such choices have been steered, on one hand, by the objectives of the R&D project: "Bias Temperature Stress" of the FCS team within Microelectronics division of CDTA and, on the other hand, by the following motivations.

1.2 Motivations

Why NBTI?

It has been shown that NBTI is the major reliability challenge when gate dielectric thickness becomes thinner than 2 nm [8] [9] [10], and that for gate oxide thinner than 4 nm, the shift of threshold voltage caused by NBTI for the PMOS is more important than that caused by hot-carrier induced degradation (HCI) of the NMOS transistor [11]. Furthermore, and contrary to HCI that occurs only during dynamic switching, NBTI is caused during static stress on the oxide even without current flow. Besides, compared to positive bias temperature instability (PBTI), it has been revealed that NBTI causes 2.3 times more degradation to the PMOS transistors than PBTI causes to the NMOS transistors [12].

Why SRAM?

CMOS SRAM (Static Random Access Memory) continues to be the technology of the choice for cache memory in high performance processors [13]. Yet, with the advances of process technology, SRAMs are more vulnerable to device parameters variability due to the process, the environment, and device aging. In fact, whatever the stored data, one of the SRAM cell's PMOS transistor is found to be under NBTI stress.

1.3 Contributions

1. By migrating the analysis of NBTI degradation from device to circuit level, we have been able to show experimentally that:

- ✓ NBTI degradation is workload dependent and that the focus on the low-level details does not provide insight about the circuit degradation due to NBTI on a larger scale.
- ✓ The degradation induced by NBTI on the CMOS inverter performances under AC NBTI stress is less pronounced than under DC NBTI stress.
- ✓ The NBTI-induced degradation is voltage, temperature, and time dependent.
- ✓ A correlation between circuit-level and device-level NBTI revealed some aspects about the NBTI degradation.

2. By considering reliability during the design stage, we have been able to:

- ✓ Simulate and analyze circuits' behavior under aging effects and assess the induced degradation on the circuits' performance and reliability.
- ✓ Propose a design technique to mitigate NBTI degradation in 6T-SRAM. The technique is based on transistor sizing. It ensures better cell stability in terms of hold and read noise margins and write-ability to enhance its lifetime.

1.4 Thesis outline

Following this introductory chapter, three chapters come expanding the statement made here about reliability aware IC design circuits. Accordingly, chapter 2 presents the background of the variability problems that have emerged with increasing CMOS technology scaling and the challenges and implication of these new process technologies to IC design. A particular interest is given to NBTI as one of the main reliability issues in deep submicron ICs. We review its analysis and modelling at different design flow abstraction levels. Besides, we present the basics of two CMOS circuits which served as benchmarks for circuit-level aging analysis and NBTI mitigation conducted in this work. Subsequently, chapter 3 presents our experimental analysis of NBTI degradation at circuit level where we report the results of CMOS inverter behavior under DC and AC NBTI stresses, and we draw some conclusion on the correlation between the degradation at device-and circuit-level. Being aware of the NBTI degradation, we present in chapter 4 a DFR approach to mitigate NBTI degradation in the 6T-SRAM. We show how to

optimize its design for a better reliability. Finally, we end-up with concluding remarks and perspectives for future work.

It is worth to note that chapters 3 and 4 are based respectively on the following publications:

1. A. Chenouf, B. Djezzar, A. Benabdelmoumene and H. Tahi, "On the Circuit-Level Reliability Degradation Due to AC NBTI Stress," in IEEE Transactions on Device and Materials Reliability, vol. 16, no. 3, pp. 290-297, Sept. 2016, doi: 10.1109/TDMR.2016.2578040.
2. A. Chenouf, B. Djezzar, H. Bentarzi and A. Benabdelmoumene, "Sizing of the CMOS 6T-SRAM cell for NBTI ageing mitigation," in IET Circuits, Devices & Systems, vol. 14, no. 4, pp. 555-561, 7 2020, doi: 10.1049/iet-cds.2019.0307.

References

[1] R. Courtland. (2017) Intel Now Packs 100 Million Transistors in Each Square Millimeter. *IEEE Spectrum*. Available: https://spectrum.ieee.org/nanoclast/semiconductors/processors /intel-now-packs-100-million-transistors-in-each-square-millimeter

[2] IRDS. (2018). *Update More Moore*. Available: https://irds.ieee.org/images/files/pdf/2018/ 2018IRDS_MM.pdf

[3] S. Ghosh, "Effect of Variations and Variation Tolerance in Logic Circuits," in *Low-Power Variation-Tolerant Design in Nanometer Silicon*, S. Bhunia and S. Mukhopadhyay, Eds., ed Boston, MA: Springer US, 2011, pp. 83-108.

[4] V. Huard, S. Mhira, F. Cacho, and A. Bravaix, "Enabling robust automotive electronic components in advanced CMOS nodes," *Microelectronics Reliability,* vol. 76-77, pp. 13-24, 2017/09/01/ 2017.

[5] T. Nigam and A. Kerber, "Enabling Scaling of Advanced CMOS Technologies: A Reliability Perspective," in *2015 IEEE Computer Society Annual Symposium on VLSI*, 2015, pp. 199-199.

[6] K. Jeong, A. B. Kahng, and K. Samadi, "Impact of Guardband Reduction On Design Outcomes: A Quantitative Approach," *IEEE Transactions on Semiconductor Manufacturing,* vol. 22, pp. 552-565, 2009.

[7] L. Zhang and R. P. Dick, "Scheduled voltage scaling for increasing lifetime in the presence of NBTI," in *2009 Asia and South Pacific Design Automation Conference*, 2009, pp. 492-497.

[8] S. P. Park, K. Kang, and K. Roy, "Reliability Implications of Bias-Temperature Instability in Digital ICs," *IEEE Design & Test of Computers,* vol. 26, pp. 8-17, 2009.

[9] R. Degraeve, M. Aoulaiche, B. Kaczer, P. Roussel, T. Kauerauf, S. Sahhaf, and G. Groeseneken, "Review of reliability issues in high-k/metal gate stacks," in *2008 15th International Symposium on the Physical and Failure Analysis of Integrated Circuits*, 2008, pp. 1-6.

[10] L. L. Lewyn, T. Ytterdal, C. Wulff, and K. Martin, "Analog Circuit Design in Nanoscale CMOS Technologies," *Proceedings of the IEEE,* vol. 97, pp. 1687-1714, 2009.

[11] N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai, and T. Horiuchi, "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," in *1999 Symposium on VLSI Technology. Digest of Technical Papers (IEEE Cat. No.99CH36325)*, 1999, pp. 73-74.

[12] S. Khan, S. Hamdioui, H. Kukner, P. Raghavan, and F. Catthoor, "BTI impact on logical gates in nano-scale CMOS technology," in *2012 IEEE 15th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*, 2012, pp. 348-353.

[13] M. Marinella, "The future of memory," in *2013 IEEE Aerospace Conference*, 2013, pp. 1-11.

CHAPTER 2 BACKGROUND

2. 1 INTRODUCTION

In 1965, Gordon E. Moore predicted that the number of transistors that can be packed into an integrated circuit (IC) will double every year for the next ten years [1]. That prediction called afterwards “Moore law” has moved the innovation in the direction of integrating more devices on a single chip which sets the pace for the modern digital revolution. From careful observation of an emerging trend, Moore extrapolated that computing would dramatically increase in power, and decrease in relative cost, at an exponential pace [2].

The ability to improve performance while decreasing power consumption has made the complementary metal oxide semi-conductor (CMOS) architecture the dominant technology for integrated circuits. As such, CMOS technology downscaling has been needed to meet more requirements of speed, complexity, circuit density, power consumption and ultimately the cost imposed by many advanced applications.

Today’s applications such as high performance computing, mobile computing, and autonomous sensing and computing (Internet of things (IoT)) (see Fig. 2.1) drive the requirements of “More Moore” technologies to bring the PPAC value from node-to-node scaling every 2-3 years[3]:

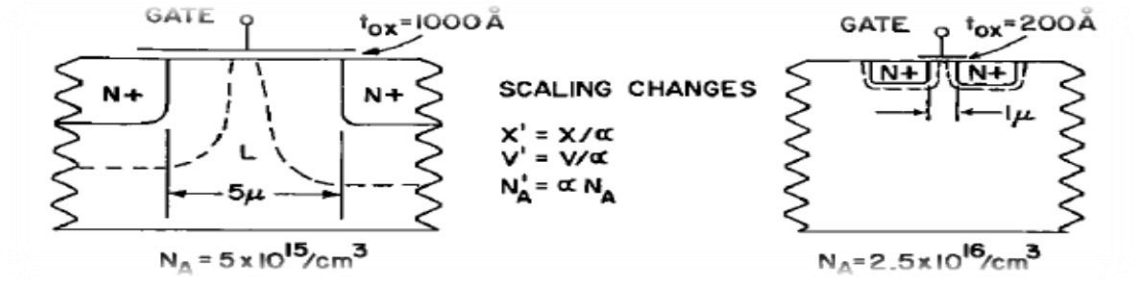
- (P)erformance: >15% more maximum operating frequency at constant energy,
- (P)ower: >30% less energy per switching at a given performance,
- (A)rea: >30% area reduction,
- (C)ost: < 30% wafer cost 30-40% less die cost for scaled die.

Fig. 2.1: Big data and instant data [3]

This impressive progress has been made possible thanks to the technology scaling.

Historically, MOS scaling has been achieved by following two principles: constant field scaling and constant voltage scaling. The story behind constant field technology scaling has its roots in the 1970s when IBM was searching for a technology to fill the cost/performance "file gap" between movable head magnetic disks (which had low cost/bit but high latency time) and random access main memory (RAM) which had high performance but high cost/bit) for transaction based systems [4]. Bob Dennard, among others who were working on alternatives, proposed a preliminary design based on his 1-Transitor- DRAM cell. Yet, many advances were required to achieve the technical and cost goals such as shrinking the dimensions on the chip to about 1µm, improving the yield to allow larger chips and higher resolution lithography, and providing a means of sensing the very small signals on the bit lines [4]. Once the problems related to the lithography, the process and the sensor were resolved, a 5X shrink of the existing technology was needed to achieve 1µm dimensions. At that time, Bob Dennard and Dale Critchlow decided, that rather than designing the 1µm technology from scratch, they would scale from some well-characterized devices which had channel lengths of about 5 µm and could be operated along with voltages up to 20 V. They observed that if the electric fields were kept constant, the reliability of

the scaled devices would not be compromised. The key to that scaling was to scale down all dimensions including wiring and depletion layers and all voltages including threshold voltage. Consequently, Dennard and Fritz Gaensslen derived the constant-field scaling theory and its limitations (see Fig. 2.2) known as the "Scaling paper" [5] and which was presented at IEDM 1972 and published in IEEE Journal of Solid-State Circuits in 1974 [5]. It was this realization of scaling theory and its usage in practice which has made possible the better-known "Moore's Law". In 1975, with more data, Moore revised the estimate of the doubling period of devices to be roughly two years [2].



New depletion thickness = $x'_D = \sqrt{\frac{2\varepsilon_{si}(V/\alpha+\Psi)}{q(\alpha N_A)}} \cong \frac{x_D}{\alpha}$

New threshold voltage = $V'_t = \frac{1}{\varepsilon_{ox}}\left(\frac{t_{ox}}{\alpha}\right)\left[\left(-Q_{eff} + \sqrt{2\epsilon_{si}q(\alpha N_A)\left(\frac{V_{s-sub}}{\alpha}\right)} + \Psi_s\right)\right] + (\Delta w_f + \Psi_s) \cong \frac{V_t}{\alpha}$

New current = $I'_D = \frac{\mu\varepsilon_{ox}}{t_{ox}/\alpha}\left(\frac{W/\alpha}{L/\alpha}\right)\left(\frac{V_g - V_t - V_d/2}{\alpha}\right)\left(\frac{V_d}{\alpha}\right) = \frac{I_D}{\alpha}$

Where: x_D is the depletion thickness, ε_{si} ε_{ox} the dielectric constant for silicon and silicon dioxide, V the supply voltage, α the scaling factor, q the electron charge, N_A the substrate accepter density, t_{ox} the gate oxide thickness, Q_{eff} the effective oxide charge Vs-sub the source voltage relative to substrate, Ψs the band bending in silicon at the onset of strong inversion for zero substrate voltage, Δw_f the work function difference between gate and substrate, μ the mobility, W and L the transistor width and length, Vg, Vd are the gate and drain voltages and Vt the transistor threshold voltage, and I_D the drain current.

Fig. 2.2: Technology scaling theory of Dennard [4]

By applying the scaling theory, the structure of the next-generation process could be known in advance, easing the process development task and the performance of circuits in the next generation could be predicted in a straightforward fashion from that in the present generation.

It is worth to note even though the device density is increased by κ^2, the power density remains constant due to the reduced power dissipation per device by κ^2 (see table 1). Following field constant scaling (in other words: the full scaling), ICs became smaller by a factor of ~ 0.8

corresponding to a density increase of 56%, making the number of the integrated transistors to roughly double every 2 years to exceed, recently, 100 million transistors/mm^2 (see Fig. 2.3) [6].

Table 1 Scaling results for device characteristics

Performance of the device	Symbol	Expression	Scaling factor
Number of device per unit area	N_{tr}	α 1/(W.L)	k^2
Gate oxide capacitance per unit area	C_{ox}	α $1/t_{ox}$	k
Gate oxide capacitance	C_{gate}	C_{ox}.W.L	1/k
Drain saturation current	I_D	$\frac{1}{2}\mu C_{ox} \cdot \frac{W}{L}(V_{gs} - V_{th})^2$	1/k
Intrinsic delay per device	τ	$C_{gate} \cdot V_{DD} \cdot I_D$	1/k
Power dissipation per device	P	$V_{DD} \cdot I_D$	$1/k^2$
Power density	P_{den}	P.N_{tr}	1
Power delay product	PDP	P. τ	$1/k^3$

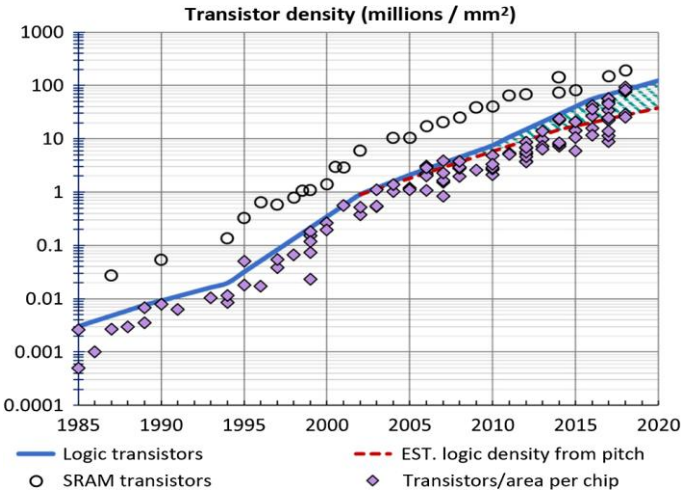


Fig. 2.3: The trend of the density of logic and 6T-SRAM chips over the last 3 decades [6].

Unfortunately, a problem came out with constant field scaling made supply voltage scaling, from a generation to another, undesirable from a system point of view. Besides, interfacing CMOS with 5 V TTL technologies imposed the supply voltage to be maintained at 5V for many years. This has led to a constant voltage scaling where all dimensions but not voltages are scaled (see Fig. 2.4) [7]. The obtained results were quite different. In fact, the circuits became faster by k^2, but the power/unit area increases by k^3. Consequently, constant voltage scaling had brought new challenges and issues [7].

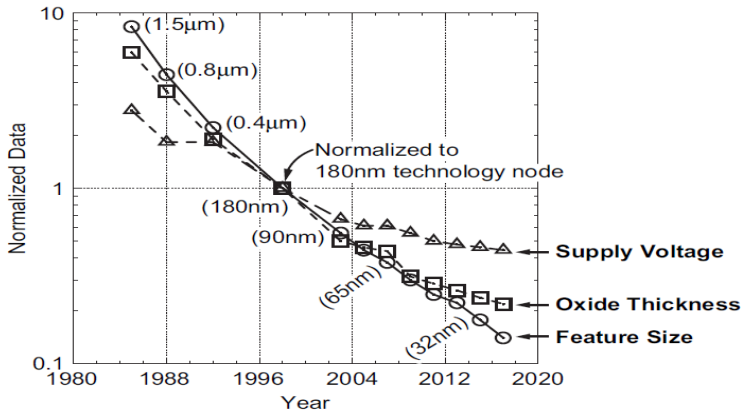


Fig. 2.4: Scaling trends of technology node, oxide thickness, and supply voltage [7]

2. 2 TECHNOLOGY SCALING IMPACT ON IC VARIABILITY

Governed by technology scaling, and in order to achieve high current drive (see table 1) and attain low power operation, the gate oxide thickness has been aggressively reduced. Yet, going to these ultra-scaled CMOS devices also brings some drawbacks. The problem at small gate lengths is that the drain voltage reduces the barrier height at the source, which causes a low source-to-channel barrier height even with the gate voltage off, and hence leads to the decrease of the threshold voltage (V_{th}) and an undesirable large off-state leakage. This phenomenon is referred to as drain-induced barrier lowering (DIBL) (see Fig. 2.5 (a)) [7]. This is due to the phenomenon known as short channel effect (SCE) (see Fig. 2.5)[7]. SCE stands for the fact that V_{th} of the scaled transistor exhibits gate length dependence [7]. SCE makes V_{th} to vary as the channel length and applied voltage change, thereby causing the degradation of circuit performance.

On the other hand, the exponential increase in the gate leakage current when scaling the gate

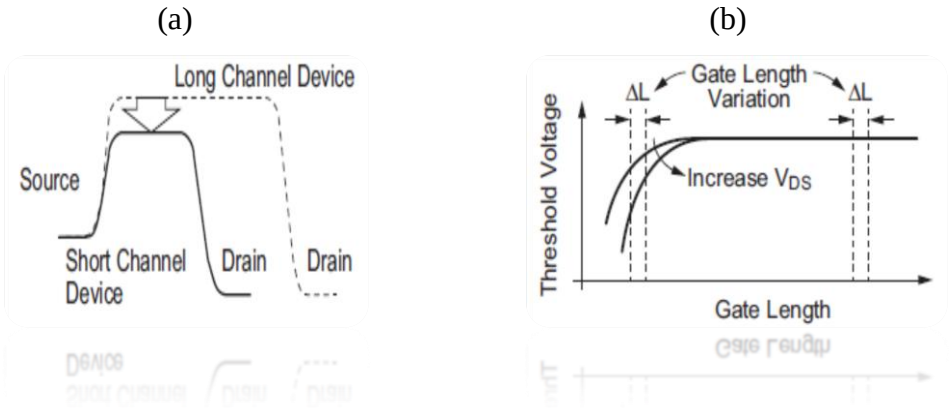


Fig. 2.5: Short channel tranasitors associated effects: DIBL in (a) and SCE in (b) [7]

oxide thickness of CMOS transistors, forced device engineers to introduce new gate materials with a higher dielectric constant compared to traditional SiO_2 or SiON gate dielectrics [8]. This allows further increase of the gate oxide capacitance while keeping the physical gate thickness sufficiently large [8].

Actually, scaling targets have driven the industry toward a number of major technological innovations, including material and process changes such as high-κ gate dielectrics and strain enhancement, new structures such as gate-all-around (GAA); alternate high-mobility channel materials, and new 3D integration schemes allowing heterogeneous stacking/integration [3]. Unfortunately, the introduction of new materials and devices associated with the further reduction of the lateral transistor dimensions reduces circuit reliability [9] .

In the general sense, reliability is the "ability of an item to perform a required function under stated conditions for a stated period of time". As such, reliability is an important requirement for almost all users of integrated circuits. This challenge is exacerbated due to (1) aggressive scaling, (2) the introduction of new materials and devices, (3) more demanding mission profiles (higher temperatures, extreme lifetimes, high currents), and (4) increasing constraints of time and money [3].

Reliability requirements are highly application dependent. Applications in harsh environments that require higher reliability levels, and/or longer lifetimes are more difficult than the mainstream office and mobile applications [10]. In fact, safety-critical applications, such as sensor interfaces embedded to monitor vital parameters in a car or biomedical products such as a pacemaker, require reliable circuit operation with a lifetime of ten to twenty years [10] . In such applications, circuits endure large temperature variations and electromagnetic interferences. Moreover, when they are used in harsh environment conditions such as those on satellites or in nuclear reactors, ICs suffer from radiation effects. Such applications impose more severe reliability requirements on ICs compared to consumer electronics where a short time to market is required to survive in a very competitive market. Despite the applications where they are used, circuits always need to operate correctly over their lifetime (see Fig. 2.6) [10].

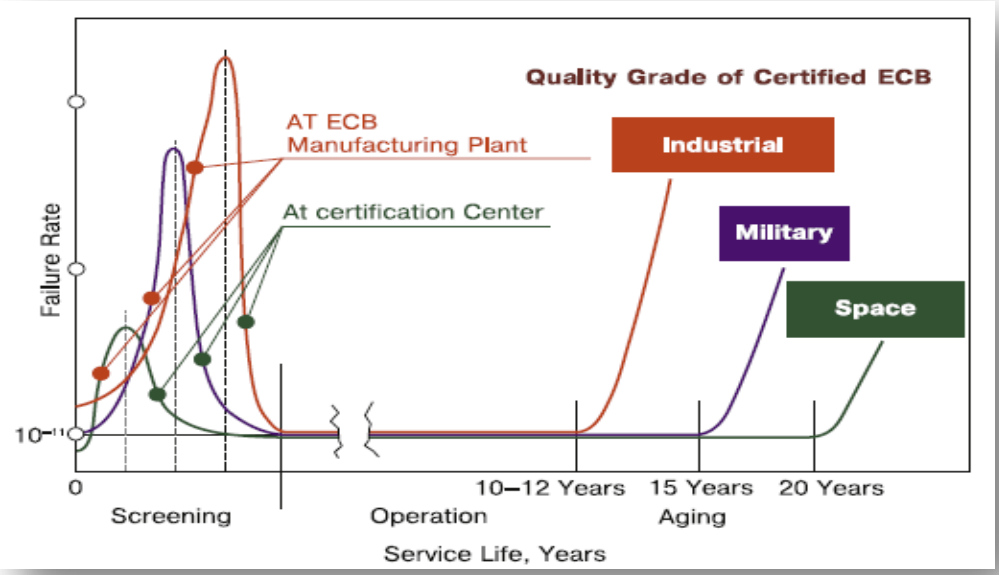


Fig. 2.6: Electronic components failure versus service life [10]

The reliability of ICs can be affected by: 1) process variations, during their manufacturing, 2) temporal variations of the devices they enfold, caused by the fluctuations of the voltage and temperature due to high density of the chips, their switching activities, and noise integrity of their different blocks during their runtime, and to 3) their wear-out mechanisms at the end of their useful lifetime. As such, the device parameters variations can be spatial (t=0), or time dependent (t>0). The taxonomy of variations of semiconductor device parameters is shown on Fig. 2.7[11].

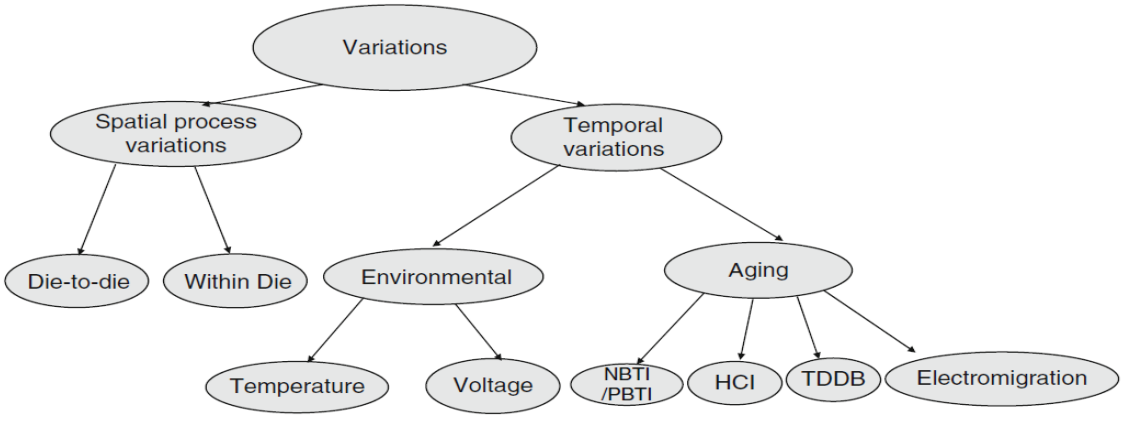


Fig. 2.7: Taxonomy of variations in integrated circuits [11]

2.2.1. Process Voltage Temperature (PVT) variability

Due to the technology process, device parameters can vary globally (inter-die) from: lot-to-lot, wafer-to-wafer, die-to-die, or locally (intra-die) between adjacent devices within the same die. Local variability results in parametric variations of identically designed transistors across a short distance known as mismatch while global variability causes a shift in the mean value of design parameters such as channel length (L), channel width (W), gate oxide thickness (Tox), resistivity or doping density, and body effect [12, 13].

Local variability can be caused by optical proximity correction (OPC), layout-induced strain and well-proximity effects, random dopant fluctuation (RDF) effects, line edge and width roughness (LER and LWR), fixed charges in the gate dielectric and interface roughness (see Fig. 2.8) [13].

While process variations are immediately noticeable just after production and are fixed in time, temporal unreliability effects, on the other hand, they are time varying and change depending on operating conditions such as the operating voltage, temperature, switching action, presence and activity of neighboring circuits. Temporal variations can be due to the environment or to aging effects.

In circuit design, circuits are usually simulated with PVT corners to account for both process and environment induced variability. For the process variability, four worst-case corners are modeled to know: SS, FF, SF, and FS. On one hand, SS stands for slow nMOS and slow pMOS while FF stands for fast nMOS and fast pMOS. On the other hand, FS stands for fast nMOS and

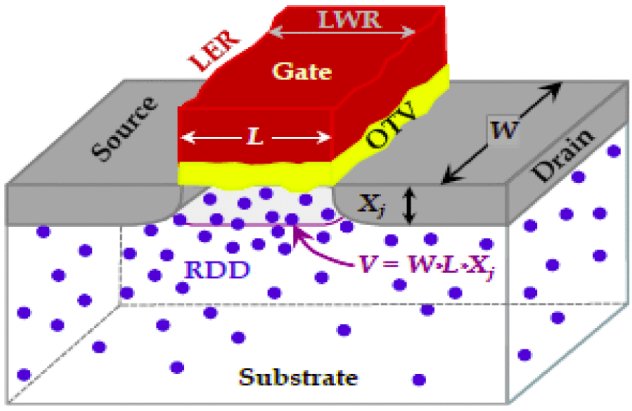


Fig. 2.8: Source of local variations in CMOS devices [13]

slow pMOS while SF stands for slow nMOS and fast pMOS. The first two corners model the worst case for speed and power for analog applications whereas the two other corners model the worst-case `1,' the worst-case `0' for digital applications (see Fig. 2.9) [12].

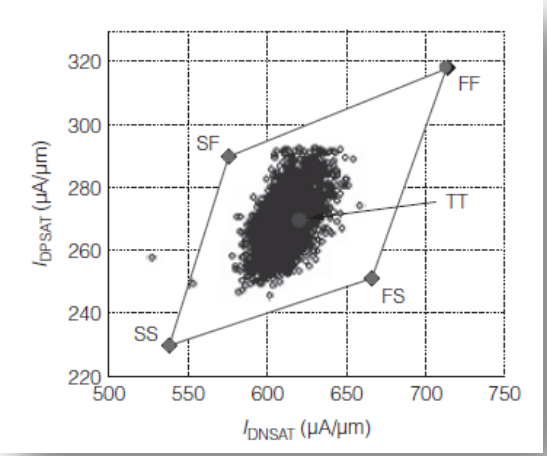


Fig. 2.9: Process corners [12]

For voltage fluctuations, circuits are usually simulated to account for fluctuations of the supply voltage within ±10% of the nominal voltage V_{DD}. For temperature fluctuations, a range of -40°C-125 °C is considered.

Thus, IC designers being aware of device parameters variability, due to process and environment, still need to know about the variability caused by wear out effects and its impact on the useful life of their circuits.

2.2.2. Aging effects

The correct operation of a MOS transistor (see Fig. 2.10) relies on the insulating properties of the dielectric layer below the gate electrode of the transistor (the insulator in bulk CMOS is SiO_2). The stability of the Si/SiO_2 interface is of a great importance.

The two most critical factors that have been limiting the scaling of the thickness of SiO_2 are: the gate leakage current which is increasing the standby power consumption and the intrinsic reliability of the oxide.

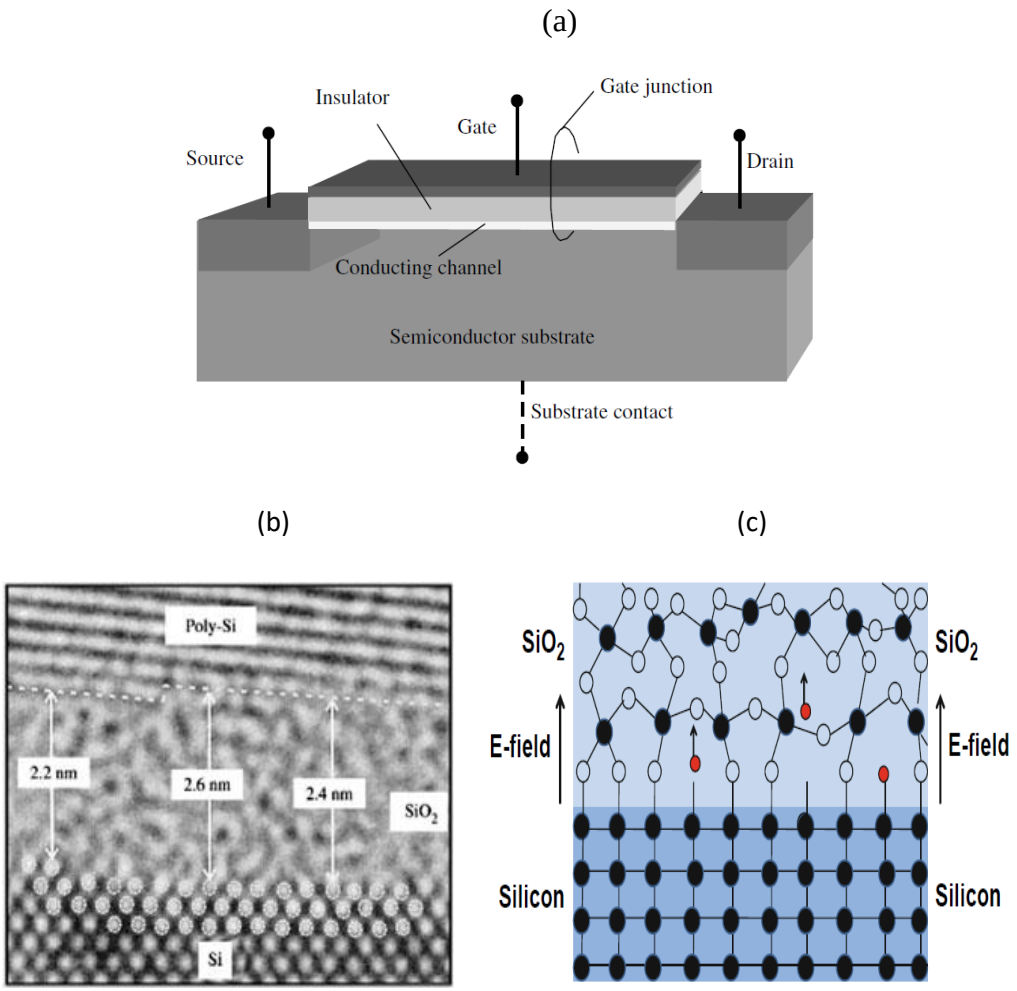


Fig. 2.10: Schematic illustration of a field effect transistor (FET) in (a). pMOS-FET with silicon dioxide (SiO_2) as insulator for CMOS bulk technology in (b source IBM) , interface states roughness (c)

As such, failure mechanisms associated with the MOS transistor also known as front-end of line (FEOL) reliability continue to be one of the main reliability challenges for ICs and systems [3]. FEOL failure mechanisms are influenced by the operating conditions of the circuit including temperature, voltage bias, and current density. The failure could be caused by either breakdown of the gate dielectric or threshold voltage change beyond the acceptable limits. The main IC aging phenomena observed in sub-90nm CMOS technologies are: hot carrier injection (HCI), time-dependent dielectric breakdown (TDDB), and bias temperature instability (BTI) [11] .

2.2.2.1. Time-Dependent Dielectric Breakdown

The correct operation of a MOS transistor relies on the insulating properties of its dielectric material. Each dielectric material has a maximum electric field that can bear, and when this limit is reached, the oxide breaks down and the device fails [11]. This is known as hard breakdown. Before the hard breakdown happens, the device might undergo a soft breakdown. This latter leads to a partial loss of the dielectric properties, resulting in a small increase of the gate current and a significant increase of the gate current noise [14].

It is generally assumed that due to the applied voltage or the resulting tunneling electrons create defects in the volume of the oxide film, these defects accumulate with time and eventually reach a critical density triggering for a sudden loss of dielectric properties. A flow of current produces a large localized rise in temperature leading to permanent structural damage in the silicon-oxide film [15]. The defects generation rate (which strongly depends on the stress conditions) and the critical defect density (which mainly depends on the oxide thickness and area) are the main elements required to relate the generation of defects to the breakdown [15]. The time to a first breakdown event is decreasing with scaling. However, depending on the circuit it may take more than one soft breakdown to produce an IC failure, or the circuit may function for longer time until the initial "soft" breakdown spot has progressed to a "hard" failure [14].

2.2.2.2. Hot Carrier Injection

Hot carriers are particles (electrons and holes) with a high kinetic energy which can be accelerated in a high electric field (in the pinch-off region during saturation mode). Hot carriers can gain enough energy and be injected into the gate oxide, then they cause interfacial damage and introduce instabilities in the electrical characteristics of a MOSFET device [16, 17] . They can lead to shifts in V_{th}, transconductance (g_m), and the drain current I_D. The ΔV_{th} and $\Delta I_D/I_D$ usually follow a power law time dependence with stress (At^n). The time exponent n of HCI degradation is found to be in the range of [0.5-0.7] for long channel bulk CMOS and at $V_G = V_D/2$ when the impact ionization is maximal, and in the range [0.2-0.3] for short channel devices where the hot carriers injection (HCI) condition changes to be $V_G = V_D$ [18].

2.2.2.3. Negative Bias Temperature Instability

Negative bias temperature instability (NBTI), is usually observed in p-channel metal–oxide–semiconductor field-effect transistors (pMOSFETs) when they are stressed with negative gate

voltage at elevated temperature [19] [20] . NBTI has been important when threshold voltages have been scaled down [19] and silicon oxy-nitride has replaced silicon dioxide in order to reduce Boron penetration in the gate insulator [21] [22]. Burn-in options (usually conducted at 125°C with +30 % V_{DD}) to enhance reliability of end-products may be impacted, as it may accelerate negative bias temperature instability (NBTI) shifts [23, 24] . Introduction of high-κ gate dielectric may impact on both the insulator failure modes (e.g., breakdown and instability) as well as the transistor failure modes such as hot carrier effects, positive and negative bias temperature instability. The replacement of polysilicon with metal gates also impacts insulator reliability, raises new thermo-mechanical issues and makes NBT-stress lifetime of copper/low-k interconnects shorter than that of aluminum/$Si0_2$ interconnects. The simultaneous introduction of high-κ and metal gate (HK/MG) makes it even more difficult to determine reliability mechanisms [25] [26]. In fact, the introduction of HK/MG has been associated with the emergence of positive bias temperature instability (PBTI) aging mechanism affecting on nMOS transistors.

NBTI is triggered in negatively biased p-channel MOSFETs within a CMOS inverter as shown in Fig. 2.11 or with (VG < 0 V, VS = VD = 0 V) configuration during accelerated stress. In fact, typical stress temperatures lie within the 100–250 °C range with the oxide electric fields

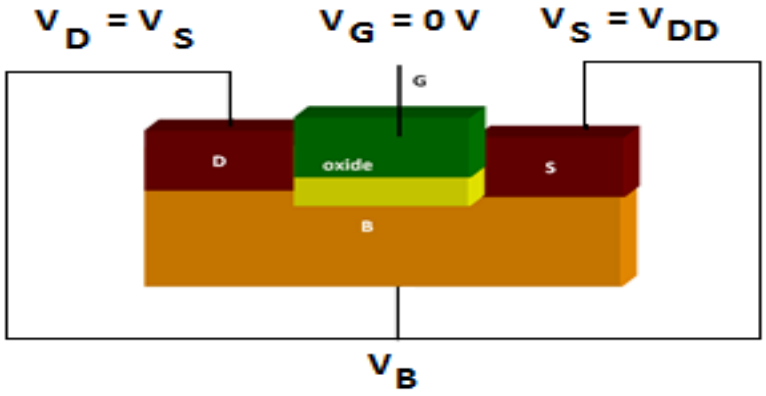


Fig. 2.11: NBTI configuration of the pMOS transistor within CMOS Inverter

typically below 6 MV/cm, i.e., electric fields beyond this value may lead to hot carrier degradation [19].

NBTI has been associated with the Si–H bond breaking along the silicon-oxide interface leading to interface traps and positive charge in the oxide [20]. Under NBTI stress, carriers

(holes) injected from the channel are captured by the traps. In large devices with many traps, this results in a well-defined change in the threshold voltage (ΔV_{th}), which is characterized by a positive shift in the absolute value of the pMOS threshold voltage $|V_{th}|$ [27] leading to a decrease in drain current I_{DS} and consequently to the performances degradation of CMOS integrated circuits [20] [28].

2. 3 NBTI MODELING & SIMULATION

To design an IC, designers usually perform many circuit simulations, make design changes and apply optimizations to evaluate the performance of their circuits before implementation using silicon materials. Therefore, it would be inconceivable to bypass the analysis of the circuit until it is fabricated and rely exclusively on the physical test of manufactured circuit to discover the errors of non-compliance of performance to specifications. Yet, it has been mainly the way in which the reliability of ICs has been treated so far. In fact, the non-deterministic of device parameters with scaling prevent circuits to fulfill their function. Hence, to improve design predictability to properly meet its specification during its lifetime, accurate modeling, and simulation tools for reliability are indispensable for both digital and analog circuits [29]. We review in the following sections NBTI aging models and simulation tools dedicated to digital circuits.

2.3.1.1. Device-level NBTI models

At device level, despite the tremendous effort made by scientists on how to properly measure the effect and to explain the microscopic behavior of NBTI, no consensus has been formed yet on the physical mechanisms governing the kinetics of NBTI (see Fig. 2.12 [30]) during stress and recovery, in large and small area devices [31-35] . Nevertheless, two classes of models have emerged as the strongest competitors: the reaction-diffusion (R-D) model [36] and hole trapping/detrapping (TD) model to explain the phenomena [37].

The first model assumes that it is the diffusion of the released hydrogen that dominates the dynamics while the second model is based on the notion that it is the reactions which essentially limit the dynamics, and that the reaction rates are distributed over a wide range.

In the R-D model, the threshold voltage degradation is modeled as power function of time (t^n) [38] [39], while in the T-D model, it is modeled as a logarithmic function of time (log(t)) [40] [41] [42] [43].

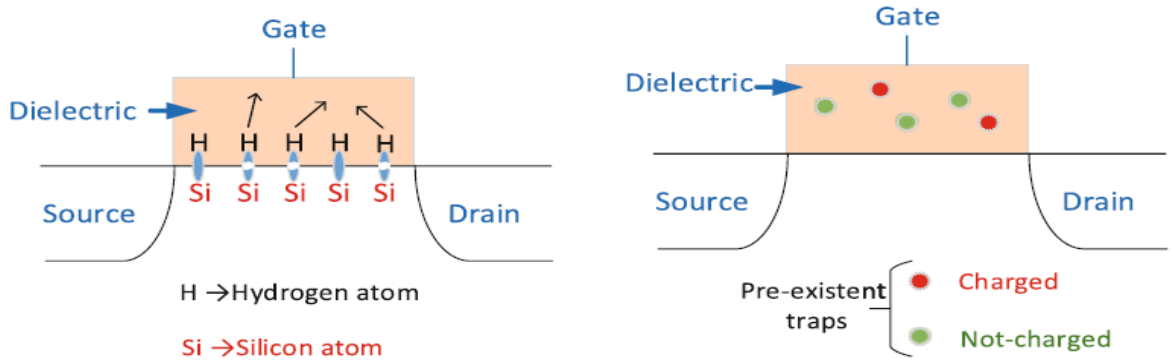


Fig 2.12: NBTI mechanisms: (a) Reaction-Diffusion (R-D) mechanism. (b) Trapping-Detrapping (TD) mechanism [30]

Each of these two models has pros and cons when interpreting experimental results of NBTI characterization on pMOS transistors with different gate oxide dielectrics (pure Silicon dioxide SiO_2, Silicon oxynitride (SiON), High-K Metal Gate (HKMG) based planar and FinFETs, the device size, and under different stress type (DC or AC NBTI) [31]. We briefly review them in the following sections.

A. The R-D model

The R-D model has been originally proposed by Jeppson & Svensson, more than 40 years ago, to explain the surface-trap growth in metal nitride oxide semiconductor (MNOS) memory transistors under negative bias stress [44]. This model has been refined by Alam *et al* [45] to explain NBTI degradation in pMOSFETs. In its old version, R-D model suggests hole-assisted depassivation of Si-H bonds at the Si/SiO_2 interface, and subsequent diffusion of Hydrogen species (as atomic H or molecular H_2,) during stress, and the reverse processes during recovery phase after stress [46] . In fact, the recovery was explained by hydrogen, recombining with the defects upon stress removal soon enough before the hydrogen diffused away forever. Once the hydrogen diffused away a certain distance, the open bond remains open and leaves a permanent damage. This standard R–D model had predicted power-law time dependence and the corresponding exponent to be independent of temperature, time, or oxide thickness. R-D attributed the observed experimental n dependence on temperature, time and thickness to the measurement delay [47] [48] [49].

According to R-D model, NBTI degradation is field-driven, accelerated by the temperature, and related to interface traps (N_{IT}) generation at the Si/SiO_2 interface (see Fig. 2.12) [30].

accordingly, the pMOS transistor threshold voltage is increased and its shift is expressed as [46]:

$$\Delta V_{th} \sim A exp(-nE_D/KT)t^n \quad (2.1)$$

with n ~ 0.25 and E_D ~0.5 eV.

First, the discrepancy between the temperature activation energy E_A ~ 0.12-0.15 and the time exponent = 0.25 is not of a big deal as the specific values of activation energies of forward dissociation (E_F), reverse annealing (E_R), and diffusion coefficient (E_D) were unknown. However, generalized scaling arguments consistently showed that $E_A \sim E_D/n$, with $E_F \sim E_R$ and that measured E_A provided a direct measure of E_D~0.5-0.6 eV which in turn implicated H_2 diffusion [45] [46]. In addition, other experiments showed a saturation of NBTI degradation at long stress time (10^5-10^7s) and independent of recovery effect (due to measurement delay) [50] [51]. In fact, Rangan *et al*. [50] found the recovery exhibiting a universal behavior independent of stress voltage, stress time and temperature (below 25°C). They also observed that activation energy varies from 0.06eV to 0.16eV and the time constant varies from 0.7 to 0.13 when the measurement delay varies from 1ms to 1000s. They attributed various activation energies, reported in the literature, to differences in the interruption times for making measurements. On the other hand, Aono *et al.*[51] showed that the degradation tends to saturate and that the dependence of lifetime on electric field (E_{ox}) is expressed as a power-law of the oxide field. By making measurement step for log time scale and extending the stress time to (10^5-10^7s) compared the measurement delay (30s) and without applying positive voltage to the gate during measurement, they minimized the recovery effect. As a result, the recovery effect becomes smaller and smaller. And so, they argued that the saturation of degradation in a long stress time region is independent of recovery effect. These works were the first indications that pre-2003 R-D analysis of NBTI may not be sufficient [47]. So, to cope with the different observations made by experimental studies reporting degradation saturation, frequency independence observed under AC NBTI stress on different types of gate dielectrics like thick thermal SiON, high-k, etc. yet four aspects were problematic for R-D model to deal with to know: the time exponent, field dependence, hole trapping, and recovery and these issues were addressed in [47] [52]. Therefore, the n ~ 0.25 has been attributed to recovery artifact during measurement, and the modified R-D model (depassivation, dimerization and molecular H_2 diffusion) has been used to explain the "artifact-free" experimental value of n ~ 0.16, as shown in Fig. 2.13 [46].

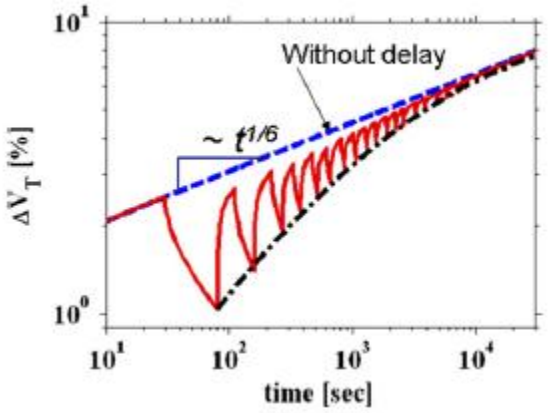


Fig. 2.13: Correction done by RD model to suppress observed saturation in experiments [46]

R-D model has interpreted the field dependence of N_{IT}-specific NBTI in pMOS devices with SiO_2 and plasma SiON as gate dielectric to arise in part from Eox dependence of T_H (tunneling probability of holes that are captured by the Si-H bonds at 1.5 A° away from the interface, and which leads to hole assisted field enhanced thermal generation of interface traps), and another part from E_{ox}-induced thermal barrier lowering (see Fig. 2.14) [47] [33] [53].

To explain the observed NBTI degradation in other dielectrics with highly nitrided dielectric and high-k materials, it has been suggested that in addition to N_{IT} generation, hole trapping in

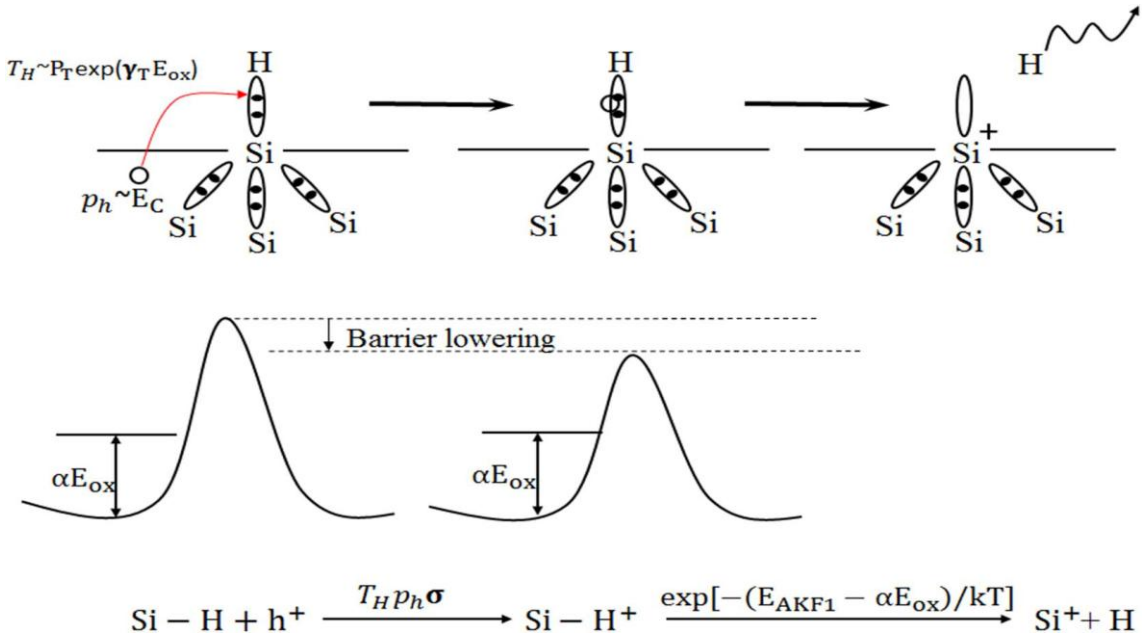


Fig. 2.14: R-D model mechanism principle [33]

preexisting defects plays an important role in determining overall time kinetics of NBTI degradation [47]. By exploring the thickness-dependent time exponents for both plasma nitride oxide (PNO) and thermal nitride oxide (TNO), Islam *et* al. [47] found that hole trapping controls the total degradation at a short stress time. It also results in a decrease in n at an intermediate stress time. Then, the saturation of ΔV_H is shown to be reached within ~1 ms and time exponent of ΔV_{th} tends to be below 1/6 within (within the range of 1−1000 s). Moreover, the contribution of hole trapping to ΔV_{TH} in PNO films has been found to be approximately two orders-of-magnitude smaller than that of TNO films. Consequently, the overall ΔV_{th} time exponent was reduced in TNO films. Further, the short term degradation (1-10s), characterized with time exponent higher than 1/6 experimentally, observed while using the On-the-fly measurement (OTFM), fast and ultrafast measurement (UFM) techniques, and it has been attributed to the gradual atom-molecule H-H_2 dimerization. As such, the NBTI results were generally understood in terms of the R-D model and the hole trapping kinetics in type I (SiO_2, low plasma nitrided SiON) and type II (highly nitrided and High-k films) [47] [49].

Despite, the H_2-R-D model can predict the power law time dependence with exponent n~1/6, a feature of DC NBTI widely accepted in the industry [39], it was not easy for R-D to figure out the fast NBTI relaxation at the early stage and the slow one at the final stage [54] [55] [56] [57] [58] [59] . R-D model anticipated a slower recovery of over 1–2 decades in timescale especially for type I devices shown to be N_{IT} governed NBTI degradation while it was shown to extend to eight-ten decades [54] [60] [61]. This substantial discrepancy in interpreting the recovery experiments has initiated researchers to look for alternate models that involve hole trapping and detrapping [27] [33] [52] [55].

B. Hole trapping-detrapping (T-D) model

A hole trapping- detrapping (T-D) model has been mainly focused on the recovery of NBTI [41] [52]. It is based on physics and supported by microscopic measurements of discrete defects in small-area devices (device area smaller than (W × L = 100 nm × 100 nm)) [32] [62]. It emphasizes on hole capture and emission in oxide traps [63]. In order to account for the asymmetry between stress and relaxation of NBTI and the bias dependence of the recovery, Grasser *et al.* [41, 63] proposed a new model for NBTI, where defect creation proceeds via a two-stage process: first, upon application of stress, the holes can be trapped into oxygen vacancies (E'-center) near the Si/oxide interface via thermally activated multi-phonon emission/multi-

phonon-field-assisted tunneling (MPE/MPFAT) mechanism. Then, a positively charged E'-center can trigger and boost the creation of interface traps (Pb-centers in SiO_2 layers and K_N-centers in oxynitrides) assumed that, via hydrogen exchange with a Pb center at the interface, the defect can transform into a more permanent state (see Fig. 2.15) [41].

Motivated by the stochastic nature of defects is deep scaled devices, and to be able to study NBTI

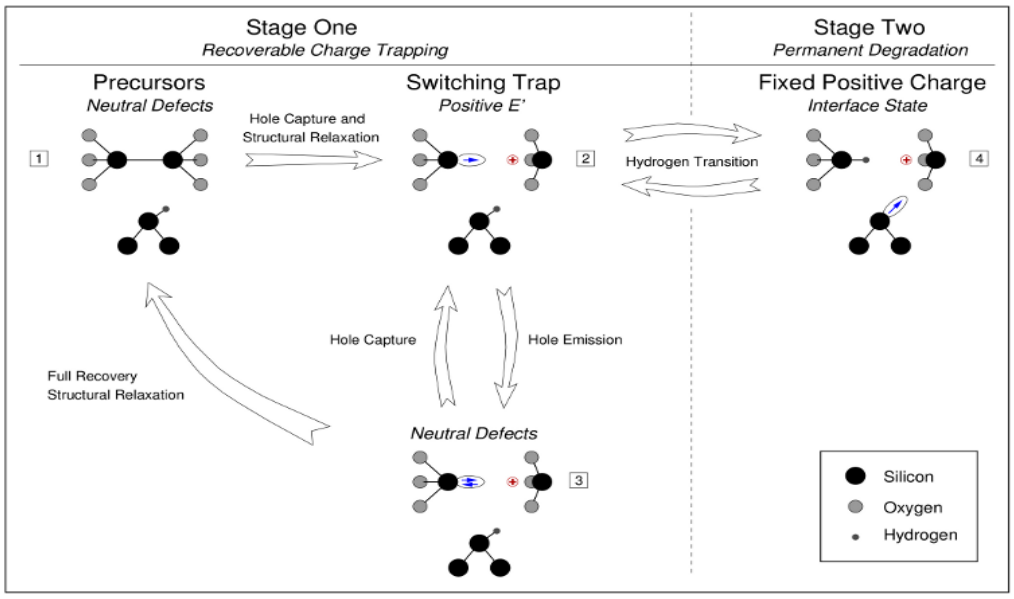


Fig. 2.15: Two-stage NBTI model [41]

in small-area devices at the single-defect level, Grasser *et al.*[62] proposed to use time-dependent defect spectroscopy (TDDS) to bring more insight on the NBTI physical mechanisms. TDDS analyzes steps in the drain current I_D due to an event of capture or emission of charge in traps and it also analyzes a given event repeatedly for many times. According to this model, the defects responsible for the recoverable component of NBTI are identical to those causing random telegraph noise (RTN), and NBTI can be considered the non-equilibrium response of the defects while RTN is a consequence of their quasi-stationary behavior. The model states that NBTI recovery is due to discharging of individual defects with a wide distribution of time scales. It disagrees with species diffusion in the oxide and argues the capture and emission time constants are temperature activated and that they are uncorrelated. No signs of a temperature-independent elastic tunneling process could be found. It has shown that capture time constants to strongly depend on field whereas the emission time constant around Vth may be either weak or strong,

depending on the configuration of the defect. The total number of defect precursors is pre-existing and no signs of newly created defects [62] .

For stress phase, Grasser *et al*. [32] found that the average degradation of the small area devices is larger by 30% of that of large devices, and that the kinetics during both stress and recovery are identical. In particular, during stress a power-law slope of 1/6 was observed in both large and small-area devices (for a measurement delay of 100 μs) while the averaged recovery is roughly logarithmic over the relaxation time. In fact, Grasser *et al.* [41] said that NBTI degradation initially follows a logarithmic time dependence while the long-time data may be approximated by a power-law with the exponent that may differ from 1/6 (n=0.11 in [41]).

C. Composed model

Despite the great efforts done by R-D and T-D models to explain the physical mechanisms responsible of NBTI, none of these two models is capable alone to interpret the large spread of experimental results of NBTI degradation observed in pMOS devices incorporating different dielectrics, built with different technology processes, and characterized by different techniques [31]. Nevertheless, an agreement starts to be established admitting that NBTI degradation could be understood in the prospect of a combination of interface states and hole trapping and detrapping in preexisting traps and new generated defects [31]. These latter can also be divided into two groups: cyclic positive charges (CPCs) and anti-neutralization positive charges (ANPCs) [64]. Accordingly, NBTI can be understood in terms of two components: recoverable and permanent (or quasi-permanent as can be recoverable with bake at high temperature >300°C [65]) and which both present different voltage, temperature and process dependences [66-68] . By eliminating the geometric component from charge pumping (CP) signal, Tahi *et al.* [68] found the permanent component to be composed by interface states N_{IT} and deep-level oxide traps, and to be mainly located in the lightly doped drain (LDD) region while the recoverable hole trapping in the shallow oxide traps takes place in the effective channel. This non-uniformity has been verified by technology computer aided design TCAD simulation [69], and the permanent component as well as the atomic structure of defect precursors behind NBTI-induced traps have been further investigated by Djezzar *et al*. [67] [70]. In fact, by applying the multi frequency charge pumping (MFCP) technique and given that for thick oxides (20 nm in the case of our devices), the pathway tunneling distances of the oxide defects can be averaged by a direct relationship between tunneling distance and frequency, Djezzar *et al.* [67] have been able to scan

the NBTI induced border traps. They have shown that the cumulative border traps ΔN_{bt} increases rapidly with Z (the distance from the interface) and exhibits a strong dependence on both temperature and stress voltage around 6-7Å from the Si/SiO_2 interface, and tends to saturate with high stress voltage and temperature. They have shown that border traps are newly generated traps and look electrically like interface traps based on their apparent activation energy and power law trend with stress time.

To more understand the origin of the NBTI degradation in the effective channel, and in order to identify and separate the contribution of different mechanisms of the permanent component of NBTI, Benabdelmoumene *et al.*[66] proposed to track the degradation in both nMOS transistors and nMOS capacitors. In fact, by characterizing NBTI in n-MOS capacitors under different stress voltages and temperatures and tracking the change in their flat band voltage (V_{FB}), they observed that the trend of ΔV_{FB} in nMOS capacitor mimics that of ΔV_{th} of pMOS transistors by following a power law with stress time. Even thought so, the time exponent is dependent on the oxide field. Moreover, the activation energy presented an unexpected evolution with the oxide field by increasing at low oxide field. These trends of the n and Ea of ΔV_{FB} in the n- MOS capacitors were also observed in ΔV_{th} in nMOS transistors [71]. Thanks to the turn-around phenomenon they observed in nMOS transistors, and based on the amphoteric nature of the interface states, they were able to identify two types of traps with opposite signs, and whose amount and rate will decide of ΔV_{th} shift sign. As such, they attributed the n and Ea increase at the beginning of the stress time to the contribution of oxide traps even during interface traps creation. Based on the observed increase of ΔN_{IT} during recovery phase, they could point out to the contribution of newly generated oxide traps (see Fig. 2.16) [66]. Moreover, they were able to delimit the NBTI degradation from stress induced leakage current (SILC) degradation [66]. These findings corroborate with those reported in. Recently, Djezzar *et al.* [72] have thoroughly investigated NBTI degradation in nMOS transistors in order to identify and separate the generated traps contributing to the permanent and recoverable components. They confirmed, on one hand, the observation of a turnaround phenomenon in nMOS transistors and they showed, on the other hand that the recovery tends to saturate at + 1V. By varying the voltage during the recovery phase, they have been able to disclose the presence of cyclic component that comes from the contribution of positive oxide traps rather than interface traps (see Fig. 2.17) [72]. Hence, they revealed that NBTI-induced ΔV_{th} is due to a permanent component (ΔV_{thp}) caused by interface

traps, a cyclic component (ΔV_{thc}), and a recoverable component (ΔV_{thr}) both caused by oxide traps [70].

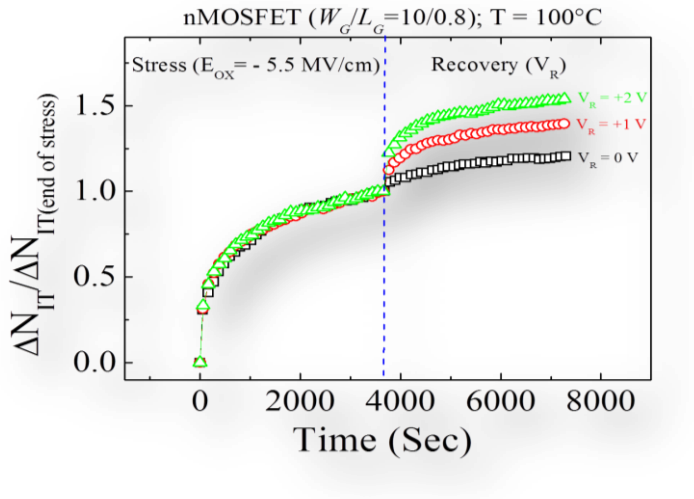


Fig. 2.16: normalized Nit degradation to end of stress for different electric fields during recovery [66].

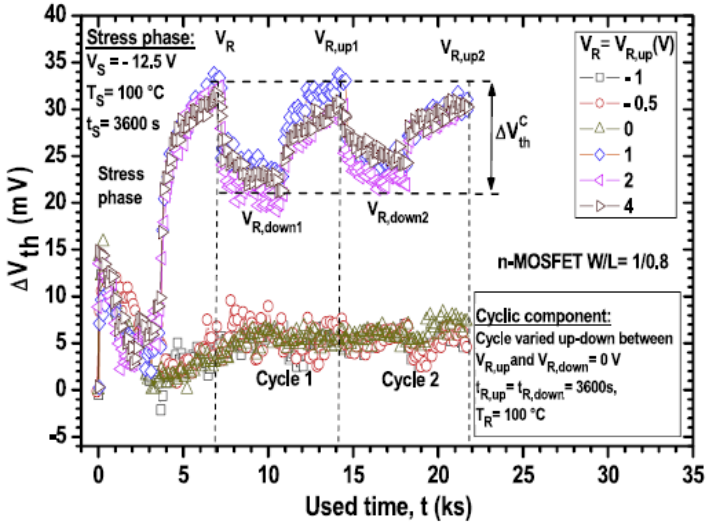


Fig. 2.17: NBTI-induced cyclic component in the gate oxide of n-MOSFET [72].

2.3.1.2. Circuit-level NBTI modeling

Circuit aging analysis which was largely confined to device reliability researchers has been expanded to the design community to foresee the impact of device reliability on system performance, built upon physics-based models.

At circuit-level, the NBTI degradation on pMOS transistor is better understood in terms of the shift of device electrical parameters such as the Vth, drain current I_D, trans-conductance g_m on the circuit performance and reliability.

It is well known that NBTI is characterized by a positive shift in the absolute value of the pMOS V_{th}, when the device is biased in strong inversion. This is interpreted in the perspective of CMOS circuit by a "0" logic applied to its input. For example, when a "0" is applied at the input of the CMOS inverter (see Fig. 2.18 (a)) then, its pMOS is negatively biased as its V_{GS} = -V_{DD}. and when operating for long tie at high temperature it undergoes NBTI, rendering it weak (V_{th} increases, I_{DS} degraded). This makes its low-high transition slower which increases the CMOS inverter delay and alters its performance Fig. 2.18 (b).

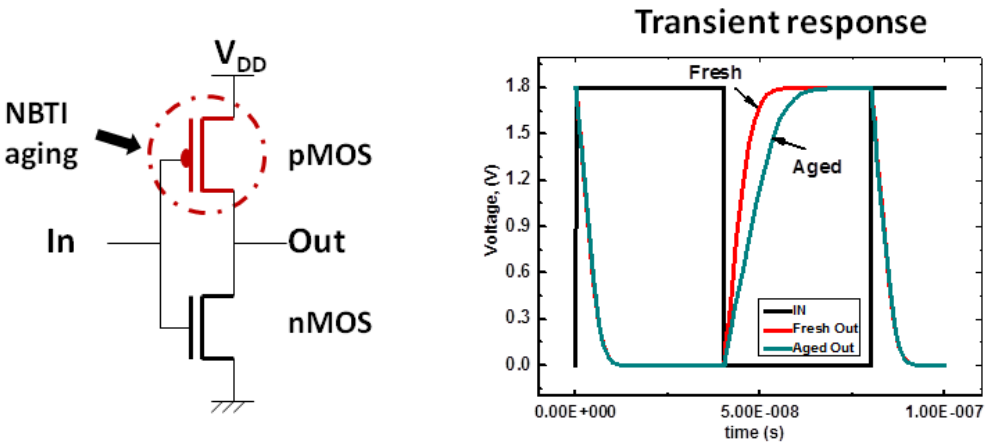


Fig. 2.18: illustration of the Impact of NBTI on the CMOS inverter performance

Based on the alpha power law model proposed in [73], the intrinsic gate delay is approximated **by:**

$$D_0 = \frac{C_L V_{DD}}{I_{ds}} = \frac{\gamma A}{(V_{gs} - V_{th})^{\alpha}} \quad (2.2)$$

where C_L is the load capacitance and V_{DD} the supply voltage, γ is the duty cycle (i.e. the ratio of the input signal that stresses pMOS transistor to the total input cycle), and A is a technology

dependent constant. By differentiating eq.(2.2) with respect to Vth, the contribution of NBTI induced transistor degradation to the gate delay of a circuit can be express as [28]:

$$\frac{\Delta D}{D_0} = \frac{\alpha \Delta V_{th}}{V_{gs} - V_{th}} \tag{2.3}$$

where, α is the velocity saturation index and Do is the gate delay without NBTI. Therefore, by monitoring the threshold voltage degradation, the change in gate delay can be easily estimated with a high degree of accuracy. By adopting the R-D model for Vth shift and accounting for mobility degradation due to NBTI, Paul *et al*.[28] expressed the degraded gate as:

$$\log \frac{\Delta D}{D} = n\log(t) + \log\left(\frac{\alpha A}{V_{gs} - V_{th}}\right) \tag{2.4}$$

where A is expressed as:

$$A = (m+1)\frac{qx\sqrt{E_{ox}e^{\left(\frac{E_{ox}}{E_0}\right)}}}{C_{ox}} \tag{2.5}$$

Where m accounts for excess shift due to mobility degradation, E_0 the field acceleration factor, x represents the field independent terms

Motivated by the fact of the logarithmic dependency, the second term could be treated as constant for a specific range of time (ten years), $log(\Delta D/D)$ slope will linearly change with log(t) with the same slope as the V_{th} degradation. Therefore, Paul *et al.*[28] concluded that by monitoring the threshold voltage degradation, the change in gate delay can be easily estimated with a high degree of accuracy. By considering the delay of a circuit as an accumulation of all individual gate delays in its critical path, the performance degradation of the whole circuit would follow the power law as the pMOS $V_{th.}$ As such, they predicted the worst case degradation of ISCAS benchmark circuits for ten years would not exceed 10% and would be about the ¼ of that of Vth degradation. This circuit NBTI degradation model has been refined by Khan and Hamdioui [74] to account for the NBTI impact of the adjacent gates on the degraded gate delay. Hence, the NBTI-induced delay of gate Y is expressed in terms of its intrinsic delay and the impact of its adjacent gate x as:

$$\Delta D_y = \left(\frac{\alpha \Delta V_{thy}}{V_{gsy} - V_{thy}}\right).A + \left(\frac{\alpha \Delta V_{thx}}{V_{gsx} - V_{thx}}\right).C \tag{2.6}$$

By considering that model, Khan & Hamdaoui found that NBTI can cause up to 19.00 % delay increment due to gate own transistors degradation and up to 4.80 % additional delay due to transistors degradation in the adjacent gates.

To predict the circuit performance degradation due to NBTI aging, other researchers have considered the atomistic trap-based model for Vth shift, where each device is characterized by the number of defects, their capture and emission times τc and τe, and their impact on the device when charged. Thus, the device BSIM model was enhanced by adding the trap kinetics equations implemented in Verilog-A component. The parameters of the enhanced model are passed to the electric simulator to generate the circuit instances to conduct statistical circuit simulation. At each time step of the transient simulation, the probability of the trap to change its state is computed. Based on this information the trap's next state is randomly defined. The impact of each populated trap on V_{th} is added to the threshold voltage parameter of the transistor model [75] [76].

A comparison of the device-level models of NBTI degradation at circuit-level has been conducted by kukner *et al.*[77] to analyze and assess their predictability of the degradation of many basic digital circuits. They found, on one hand, both models are in consistency in terms of gate delay with respect to its type, its drive strength, frequency, and duty factor. On the other hand, the analysis showed the atomistic trap-based model is superior from the point of non-periodicity and instant degradation, while the R-D model gets advantageous in case of long-term aging [77].

To consider NBTI at a higher level of abstraction Amrouch *et al*. [78] proposed to provide the synthesis tool with a degradation-aware cell library to allow addressing aging concerns and to design reliable digital circuits (see Fig. 2.19) [78]. The degradation-aware cell library was built by considering a duty cycle λ increasing with a step of 0.1 in the range of [0, 1] and 49 operating conditions (OPCs) (7 input signal slews (S) and 7 output load capacitances (C)) for each gate. The required boundaries were determined as follows: Smin, the minimum slew rate was set based on the fastest gate in the library under no fanout while Smax, the maximum slew rate was set based on the slowest gate connected to the maximum fanout. Cmin and Cmax are determined based on the smallest (i.e. no fanout) and largest (i.e. maximum allowed fanout) [78]. The as-built library is plugged to a timing analysis tool will enable accurate analysis of the timing behavior of the entire circuit in the scope of static and/or dynamic aging stress. Hence, the required guard band can be obtained. Additionally, synthesizing the circuit based on the degradation-aware cell library enables optimization algorithms to consider aging effects (in every gate/cell) [78].

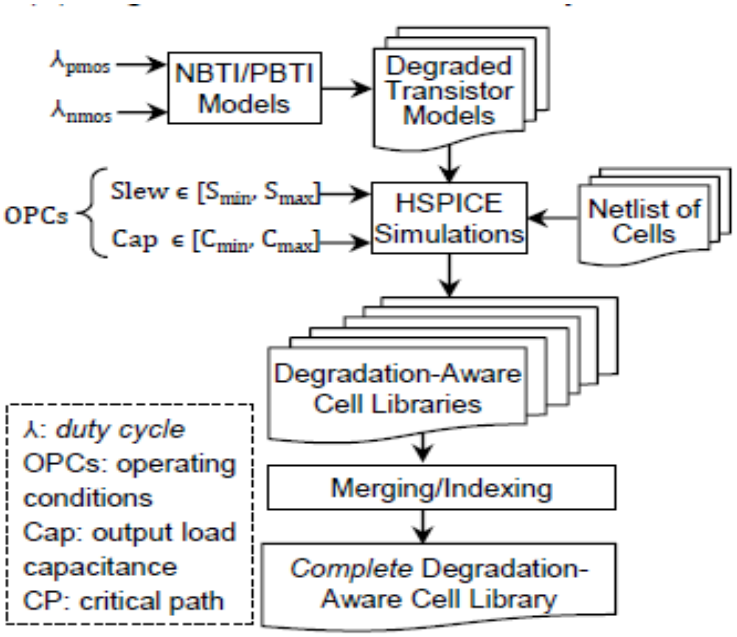


Fig. 2.19: Degradation-aware cell library creation [78]

2. 4 DESIGN BENCHMARKS

CMOS Inverter (the basic cell in logic design) and the 6T-SRAM cell (the dominant architecture used for cache memories) are the benchmarks we have chosen in this thesis for circuit-level NBTI analysis and reliability aware IC design approach. The two design benchmarks function and features are presented hereafter.

2.4.1 The CMOS inverter

The CMOS inverter provides one of the basic functions in digital logic that is the "NOT" operation. The inverter is the simplest cell built in CMOS technology using a pMOS, nMOS pair that shares a common gate as input and the drain as output while their source are connected to the supply voltage and to the ground respectively (see Fig. 2.20).

The inverter operation can be understood by examining the relationship between the input voltage and the gate-source voltages of the two transistors as follows:

$$V_{gsn} = V_{in} \quad \& \quad V_{gsp} = V_{in} - V_{DD} \tag{2.7}$$

$$V_{dsn} = V_{out} \quad \& \quad V_{dsp} = V_{out} - V_{DD} \tag{2.8}$$

When a high input voltage is applied to the inverter input (Vin = V_{DD}). Then,

$$V_{gsn} = V_{DD} \ \& \ V_{gsp} = 0 . \quad (2.9)$$

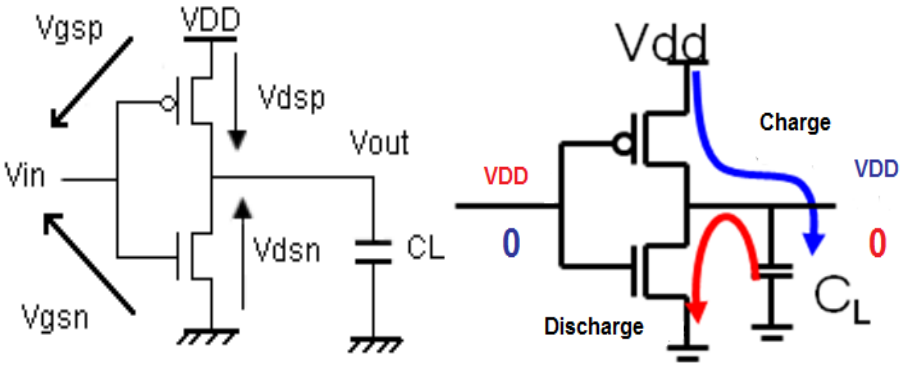


Fig. 2.20: CMOS inverter input and output voltages

In this case, the pMOS transistor is cutoff while the nMOS is conducting in the non-saturated mode. Then, nMOS provides a current path to ground, resulting in an output voltage of $min(V_{out}) = V_{OL} = 0\,V$. Where V_{OL} is called the output low voltage, and represents the smallest voltage available at the output. V_{OL} the maximum voltage at the output considered as logic "0" . In the other case, when a low input voltage is applied to inverter (Vin = V_{DD}) then:

$$V_{gsn} = 0 \ \& \ V_{gsp} = -V_{DD}. \quad (2.10)$$

In such biasing condition, the nMOS is in cutoff state while pMOS conducts to the non-saturated mode. Then, pMOS provides a conductive path to the power supply and gives $max(V_{out}) = V_{OH} = V_{DD}$, which defines the output high voltage V_{OH} of the circuit; which is the largest value of Vout.

The CMOS inverter features that are characterized under both DC and AC NBTI (as it will be discussed in chapter 3) are its DC characteristics and transient response. DC features include inverter logic threshold (Vinv), the acceptable input and output logic voltages V_{IH}, V_{IL}, V_{OL} and V_{OH} (see Fig. 2.21), and the noise immunity which characterize the stability of the inverter, and expressed in terms of two factors noise margin high (NMH) and noise margin low (NML) (see Fig. 2.22).

A. The inverter logic threshold voltage (Vinv)

The Inverter logic threshold also known as the switching voltage delimits logic 1 from logic 0 and is defined where Vin = Vout on its VTC

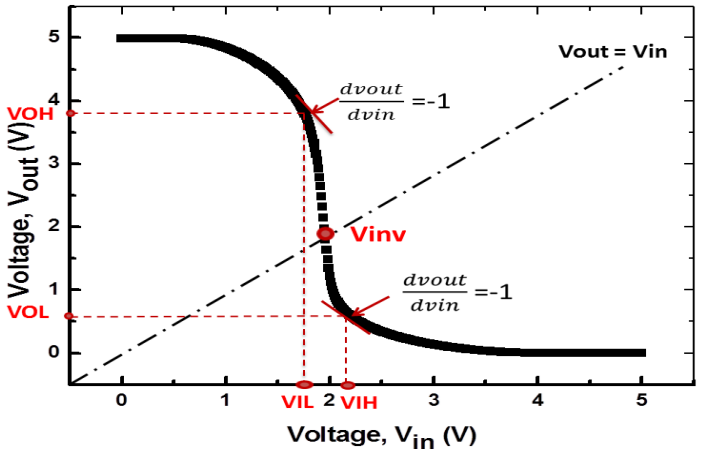


Fig. 2.21: CMOS Inverter DC characteristics

At this point, pMOS and nMOS transistors are both operating in the saturation mode. By equaling their drain currents (Eq. 11), Vinv can be resolved and expressed by Eq.2.12.

$$\frac{1}{2}\beta_n\left(v_{inv}-v_{thn}\right)^2=\frac{1}{2}\beta_p\left(v_{DD}-v_{inv}+v_{thp}\right)^2 \tag{2.11}$$

$$v_{inv}=\frac{r\left(v_{DD}+v_{thp}\right)+v_{thn}}{r+1} \tag{2.12}$$

where

$$r=\sqrt{\frac{\beta_p}{\beta_n}}=\sqrt{\frac{\mu_p W_p}{\mu_n W_n}} \tag{2.13}$$

By differentiating Eq. (2.12) with respect to $V_{thp,}$ of the logic threshold shift due to NBTI, ΔV_{inv} can be expressed as:

$$\Delta v_{inv}=\frac{r}{r+1}\Delta v_{thp} \tag{2.14}$$

Then, it is expected that Vinv will follow the same trend of V_{thp} in terms of time dependence but with a lower magnitude. This will be thoroughly examined in chapter 3.

B. The Inverter logic levels

the critical voltages V_{IL}, V_{IH}, V_{OL} and V_{OH} of the CMOS inverter VTC are defined by the slope of the VTC that has a value of -1 as shown in Fig. 2.21. The input low voltage V_{IL} represents the largest value of Vin that can be interpreted by logic “0”, whereas the input high voltage V_{IH} represents the smallest value of Vin that can be interpreted as logic “1”. The output

low voltage V_{OL} defines the largest value of Vout's logic "0", while the output high voltage V_{OH} represents the smallest value of Vout logic "1". All these voltages are defined by:

$$\frac{d v_{out}}{d v_{in}} = -1 \tag{2.15}$$

C. The Inverter noise immunity

The word "noise" in the context of digital circuits means unwanted variations of voltages and currents at the logic nodes [79]. This noise can be due to a capacitive, an inductive cross talk, or it can be internally-generated power supply. Whatever its source, for a gate to be robust and insensitive to noise, it is essential that the "0" and "1" intervals be as large as possible. A measure of the inverter sensitivity to noise is given by noise margin low (NML) and noise margin high (NMH), which quantize the size of the legal "0" and "1", respectively, and set a fixed maximum threshold on the noise value (see Fig. 2.22). So, we define the voltage noise margin for logic 1 (high) as:

$$NMH = V_{OH} - V_{IH} \tag{2.16}$$

Similarly, the voltage noise margin for logic 0 (low) as:

$$NML = V_{IL} - V_{OL} \tag{2.17}$$

A CMOS inverter is said to be robust if its noise immunity is as large as possible. In chapter 3, the NBTI effect on the noise margins of the inverter and noise immunity will be presented.

2.4.2 The 6T-SRAM

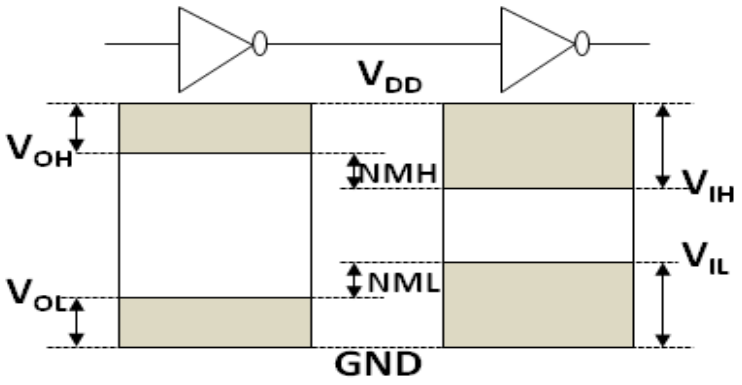


Fig. 2.22 : CMOS Inverter logic Levels and Noise margins

It is well known that without a memory no information can be stored or retrieved in a computer. A computer memory is organized in a hierarchy and which can be characterized by a

number of parameters. Among these parameters are the access type, capacity, cycle time, latency, bandwidth, and cost. The term access refers to the action that physically takes place during a read or write operation. The capacity of a memory level is usually measured in bytes. The cycle time is defined as the time elapsed from the start of a read operation to the start of a subsequent read. The latency is defined as the time interval between the request for information and the access to the first bit of that information. The bandwidth provides a measure of the number of bits per second that can be accessed. The cost of a memory level is usually specified as dollars per megabytes [80].

A typical memory hierarchy starts with a small, expensive, and relatively fast unit, called the cache, followed by a larger, less expensive, and relatively slow main memory unit (see Fig. 2.23). Cache and main memory are built using solid-state semiconductor material. They are followed in the hierarchy by far larger, less expensive, and much slower magnetic memories that consist typically of the (hard) disk and the tape [80].

The idea behind using a cache as the first level of the memory hierarchy is to keep the information expected to be used more frequently by the CPU close to it.

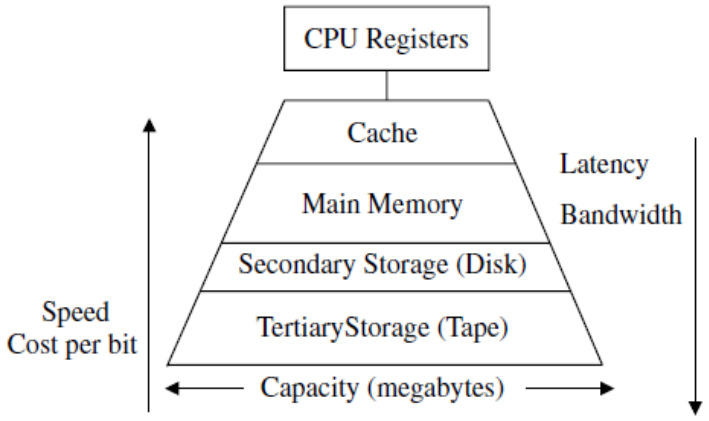


Fig. 2.23: Typical memory hierarchy [80]

The speed of a cache is typically in inverse proportion to its size. As such, modern computers have more than one cache level, frequently up to 3 levels. Additionally, it is common to have a separate Level 1 (and even Level 2) cache for instructions and another for data as shown on Fig. 2.24 [81].

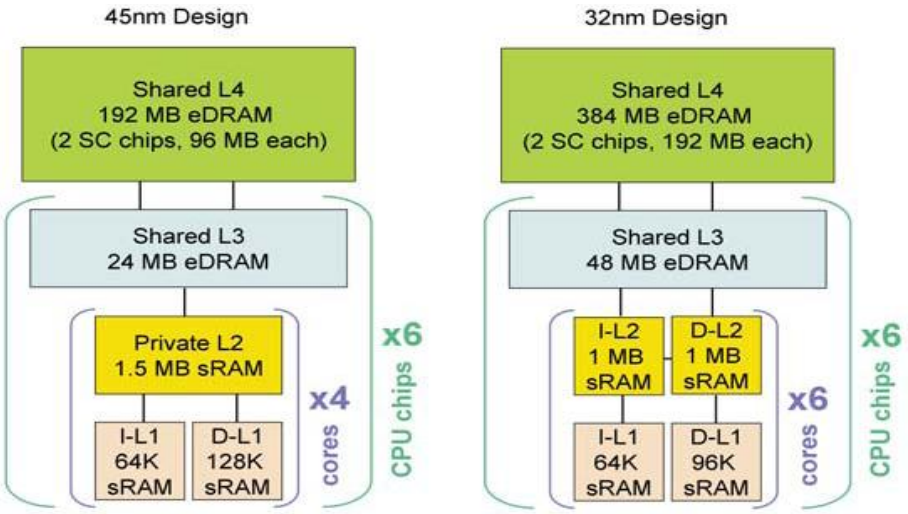


Fig. 2.24: Processor with 3 level cache [81]

Despite the progressive introduction of new emerging memory technologies such as Phase-Change Random Access Memory (PCRAM), Resistive RAM (ReRAM), Magnetic RAM (MRAM), Spin-Torque Transfer Magnetic RAM (STT-MRAM), the CMOS SRAM continues to be the technology of choice for cache memory [82].

6T-SRAM cell is made up of two cross coupled inverters (Q1, Q3 and Q2, Q4), and two access transistors Q5 and Q6 that provide the read and write operations (see Fig. 2.25).

Upon the activation of the word line, the access transistors connect the two internal nodes of the cell to the true (BL) and the complementary (BLB) bit lines.

An SRAM cell is a circuit that has three main operations.

- Write - A data bit is stored in the circuit
- Hold - The value of the data bit is maintained in the cell
- Read - The value of the data bit is transferred to an external circuit

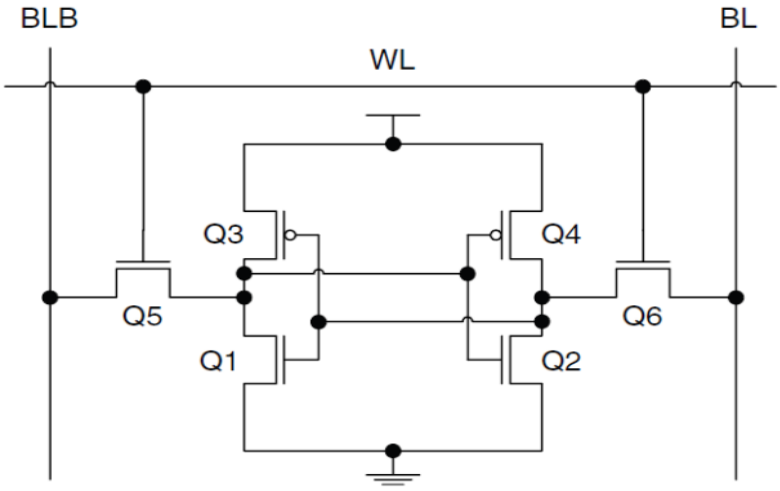


Fig. 2.25: 6T-SRAM Cell schematic

A. Hold operation

The hold state of the cell is achieved by bringing the word line signal to WL = 0. This places both access transistors in cutoff, and isolates the storage cell from the BL and BLB. An SRAM cell is capable of holding a data bit so long as the power is applied to the circuit.

B. Read operation

Prior to initiating a read operation, bit lines are pre-charged to V_{DD}. Then, the operation is initiated by enabling the word line (WL) and the access transistors Q5 and Q6 connect the pre-charged bit lines, BL and BLB, to the internal nodes of the cell. Upon read access the bit line voltage V_{BL} remains at the pre-charged level. The complementary bit line voltage V_{BLB} is discharged through transistors Q1 and Q5 connected in series. Effectively, transistors Q1 and Q5 form a voltage divider whose output is now no longer at zero volt and is connected to the input of inverter Q2−Q4 (Fig. 2.26). Sizing of Q1 and Q5 should ensure that inverter Q2−Q4 does not switch causing a destructive read. In other words, 0 +ΔV should be less than the switching threshold of inverter Q2−Q4 [83].

Ignoring the short-channel and body effects, the maximum allowed value 0+ΔV of the node storing logic “0” during read access can be obtained by:

$$\Delta V = \frac{V_{DSATn}+CR(V_{DD}-V_{thn})\pm\sqrt{V_{DSATn}^2(1+CR)+CR^2(V_{DD}-V_{thn})^2}}{CR} \quad (2.18)$$

where *CR* is the cell ratio defined as:

$$CR = \frac{W_1/L_1}{w_5/l_5} \tag{2.19}$$

Typically, and in order to ensure a non-destructive read and an adequate noise margin, *CR* must be greater than one and can be varied depending on the target application of the cell from approximately 1 to 2.5 [83].

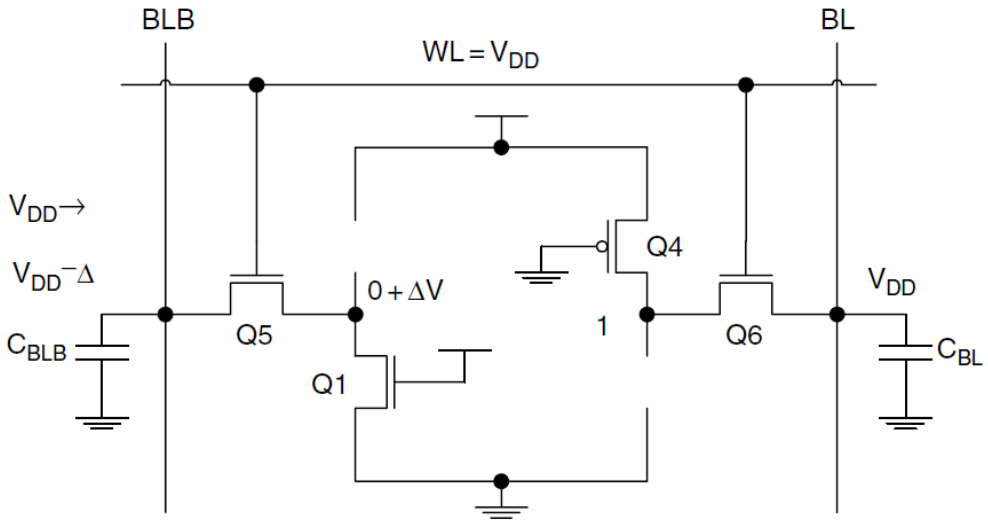


Fig. 2.26: 6T-SRAM during read operation

C. Write operation

Write operation of a 6T-SRAM cell is nothing but a flipping of the bit storing "1".Thus, one of the bit lines, BL in Fig. 2.27, should be driven from pre-charged value (V_{DD}) to the ground potential by a write driver through transistor Q6. If transistors Q4 and Q6 are properly sized, then the cell is flipped and its data is effectively overwritten.

$$V_{"1"} = V_{DD} - V_{thn} - \sqrt{(V_{DD} - V_{thn})^2 - 2\frac{\mu_p}{\mu_n} PR\left((V_{DD} - |V_{thp}|)V_{DSATp} - \frac{V_{DSATp}^2}{2}\right)} \tag{2.20}$$

where PR is the Pull-up ratio defined as:

$$PR = \frac{w_4/L_4}{w_6/l_6} \tag{2.21}$$

D. 6T-SRAM stability

The static noise margin (SNM) is the measurement of the cell's stability and in maintaining its stored data against noise. SNM is the maximum amount of noise voltage V_n that can be tolerated at the both inputs of the cross-coupled inverters in different directions while inverters

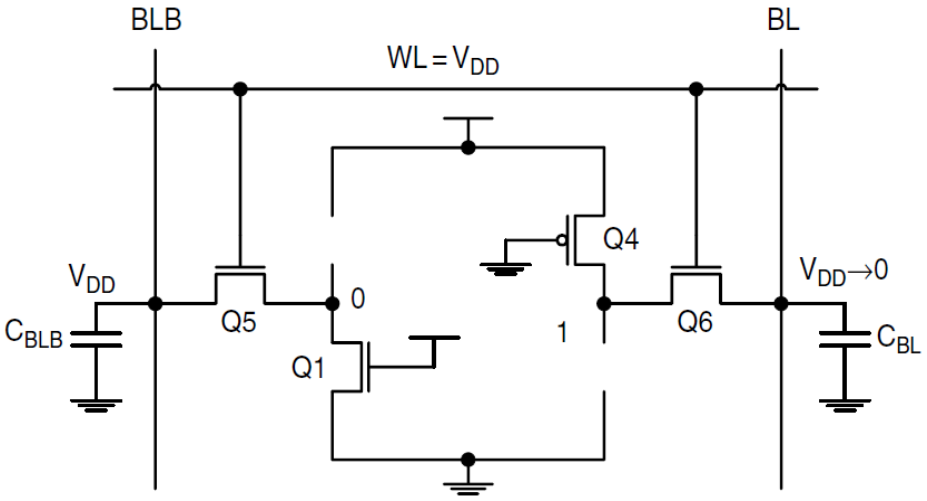


Fig. 2.27: 6T-SRAM during write operation

still maintain bi-stable operating points and cell retains its data. In other words, SNM quantifies the amount of noise voltage V_n required at the storage nodes of SRAM to flip the cell data. The SNM can be calculated by the min edge of the largest square which can be embodied by the two voltage transfer function curves VTCs of the SRAM (see Fig. 2.28).

The static noise margin can be divided into two types: the Hold SNM and the Read SNM. The first is the SNM of the cell in hold mode (WL = 0) and when in access mode (WL = 1) for read operation.

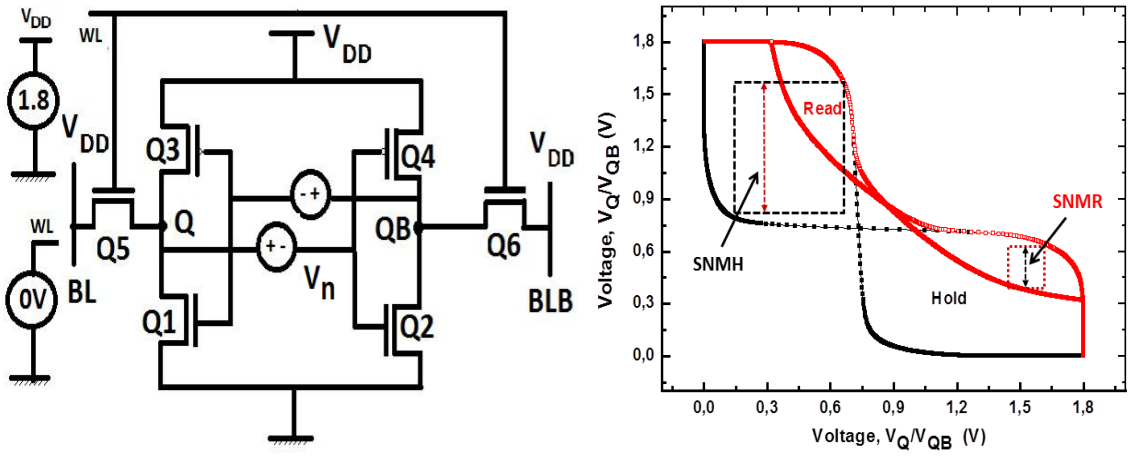


Fig. 2.28: 6T-SRAM SNM setup and butterfly

The Read SNM is more critical than Hold SNM because the cell is fragile during read operation as shown in Fig. 2.28 and Fig. 2.29 for the 6T-SRAM designed in TSMC 0.18µm CMOS technology. It is clear that when the cell is hold mode the noise amount to disturb it should be higher than the 1/3 V_{DD} (> 680mV) while the amount 320mV is sufficient to disturb it and flip the bit cell higher during read operation.

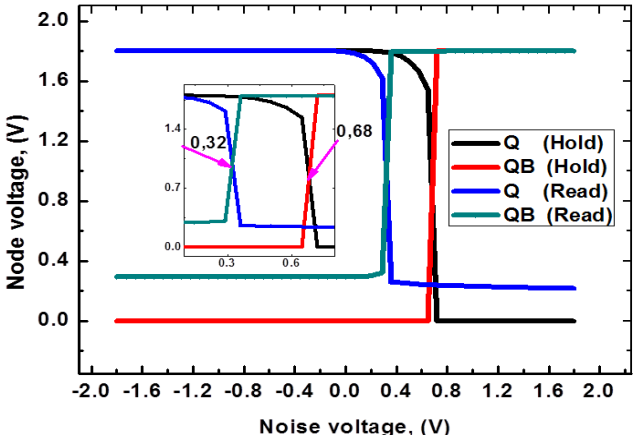


Fig. 2.29: Example of hold and read SNM of a 0.18µm 6T-SRAM cell

A statistical measure of SRAM cell write-ability is defined as write margin. The cell data is written by forcing the BL pair to the differential levels of "1" and "0" while WL is asserted to allow the access transistor to be connected to the BL. The potential of the corresponding storage node is pulled down to the critical level that is dependent on the ratio of transistor strengths between the pull-up and the access transistors. This ratio is referred to as PR. In order to ensure robust write operation, the critical level has to be lowered than the trip point of connected inverter before the level of "0"-written BL is reached to the end-point (e.g., GND). The write margin (WRM) is defined as the rest of potential difference between the BLB level at which the data is flipped and the end-point (e.g., GND) [84].

If the cell data is flipped when the BLB comes at X mV, where X mV is allowed to reach the GND level, WRM is defined as X mV. The lower value is, the harder is to write the cell, implying a smaller write margin. Therefore, a careful design of PR ratio is needed to maintain the required WRM. An example of the write margin of the 6T-SRAM designed in TSMC 0.18µm CMOS technology is shown in Fig. 2.30. It outcomes that the BLB should be lowered by 1.1 to get the cell flipped and the data written. In this case, the write margin equals 0.7V.

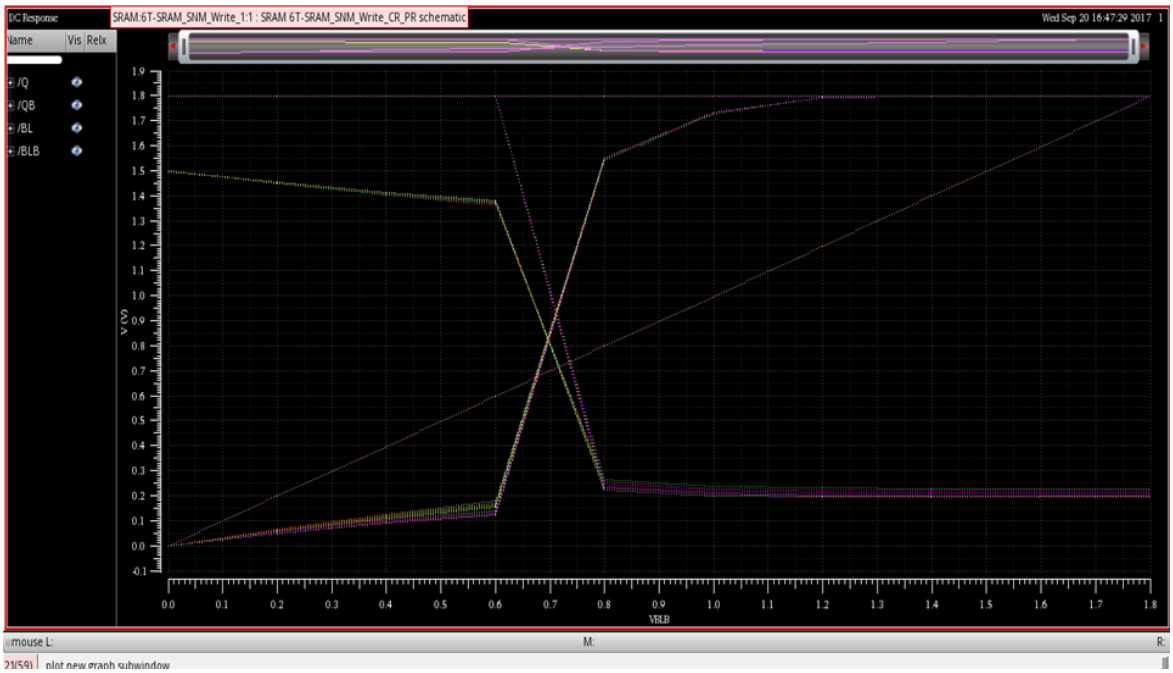


Fig. 2. 30: Example write margin of a 0.18µm 6T-SRAM cell

Both good noise margin during read operation and a good write-ability are required for 6T-SRAM correct operation. However, these operations are entailing conflicting constraints on access transistors sizing. Besides, under NBTI, these constraints become challenging when the pMOS transistors parameters will shift as function of time. Then, an appropriate 6T-SRAM design would take into consideration devices temporal variability to insure data integrity in the memory as further examined in chapter 4.

References

[1] 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. 11, pp. 33-35, 2006.

[2] P. J. Denning and T. G. Lewis, "Exponential laws of computing growth," *Commun. ACM,* vol. 60, pp. 54–65, 2016.

[3] IRDS. (2018, 12/12/2020). *update More Moore*. Available: https://irds.ieee.org/images/files/pdf/2018/2018IRDS_MM.pdf

[4] D. L. Critchlow, "Mosfet Scaling-the Driver of VLSI Technology," *Proceedings of the IEEE,* vol. 87, pp. 659-667, 1999.

[5] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits,* vol. 9, pp. 256-268, 1974.

[6] M. Rieger, "Retrospective on VLSI value scaling and lithography," *Journal of Micro/Nanolithography, MEMS, and MOEMS,* vol. 18, p. 040902, 2019.

[7] T. Iizuka, "CMOS technology scaling and its implications," in *Digitally-Assisted Analog and Analog-Assisted Digital IC Design*, X. Jiang, Ed., ed Cambridge: Cambridge University Press, 2015, pp. 1-20.

[8] M. M. Frank, "High-k / metal gate innovations enabling continued CMOS scaling," in *2011 Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, 2011, pp. 25-33.

[9] R. Degraeve, M. Aoulaiche, B. Kaczer, P. Roussel, T. Kauerauf, S. Sahhaf, and G. Groeseneken, "Review of reliability issues in high-k/metal gate stacks," in *2008 15th International Symposium on the Physical and Failure Analysis of Integrated Circuits*, 2008, pp. 1-6.

[10] V. A. Kharchenko, "Problems of reliability of electronic components," *Modern Electronic Materials,* vol. 1, pp. 88-92, 2015/09/01/ 2015.

[11] S. Ghosh, "Effect of Variations and Variation Tolerance in Logic Circuits," in *Low-Power Variation-Tolerant Design in Nanometer Silicon*, S. Bhunia and S. Mukhopadhyay, Eds., ed Boston, MA: Springer US, 2011, pp. 83-108.

[12] S. K. Saha, "Modeling Process Variability in Scaled CMOS Technology," *IEEE Design & Test of Computers,* vol. 27, pp. 8-16, 2010.

[13] S. K. Saha, "Compact MOSFET Modeling for Process Variability-Aware VLSI Circuit Design," *IEEE Access,* vol. 2, pp. 104-115, 2014.

[14] H. J. Lee and K. K. Kim, "Analysis of time dependent dielectric breakdown in nanoscale CMOS circuits," in *2011 International SoC Design Conference*, 2011, pp. 440-443.

[15] "Dielectric Breakdown of Gate Oxides: Physics and Experiments," in *Reliability Wearout Mechanisms in Advanced CMOS Technologies*, ed, 2009, pp. 209-329.

[16] "Hot Carriers," in *Reliability Wearout Mechanisms in Advanced CMOS Technologies*, ed, 2009, pp. 441-516.

[17] S. Mahapatra, C. D. Parikh, V. R. Rao, C. R. Viswanathan, and J. Vasi, "Device scaling effects on hot-carrier induced interface and oxide-trapped charge distributions in MOSFETs," *IEEE Transactions on Electron Devices,* vol. 47, pp. 789-796, 2000.

[18] D. Vuillaume, A. Bravaix, and D. Goguenheim, "Hot-carrier injections in SiO2," *Microelectronics Reliability,* vol. 38, pp. 7-22, 1998/02/20/ 1998.

[19] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *Journal of Applied Physics,* vol. 94, pp. 1-18, 2003.

[20] D. K. Schroder, "Negative bias temperature instability: What do we understand?," *Microelectronics Reliability,* vol. 47, pp. 841-852, 2007/06/01/ 2007.

[21] L. K. Han, D. Wristers, J. Yan, M. Bhat, and D. L. Kwong, "Highly suppressed boron penetration in NO-nitrided SiO/sub 2/ for p/sup +/-polysilicon gated MOS device applications," *IEEE Electron Device Letters,* vol. 16, pp. 319-321, 1995.

[22] S. S. Tan, T. P. Chen, C. H. Ang, and L. Chan, "Mechanism of nitrogen-enhanced negative bias temperature instability in pMOSFET," *Microelectronics Reliability,* vol. 45, pp. 19-30, 2005/01/01/ 2005.

[23] A. Chakraborty and D. Z. Pan, "Controlling NBTI degradation during static burn-in testing," in *16th Asia and South Pacific Design Automation Conference (ASP-DAC 2011)*, 2011, pp. 597-602.

[24] 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 Transactions on Device and Materials Reliability,* vol. 11, pp. 219-226, 2011.

[25] J. S. Suehle, "Ultrathin gate oxide reliability: physical models, statistics, and characterization," *IEEE Transactions on Electron Devices,* vol. 49, pp. 958-971, 2002.

[26] A. Suzuki, K. Tabuchi, H. Kimura, T. Hasegawa, and S. Kadomura, "A strategy using a copper/low-k BEOL process to prevent negative-bias temperature instability (NBTI) in p-MOSFETs with ultra-thin gate oxide," in *2002 Symposium on VLSI Technology. Digest of Technical Papers (Cat. No.01CH37303)*, 2002, pp. 216-217.

[27] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability,* vol. 46, pp. 1-23, 2006/01/01/ 2006.

[28] B. C. Paul, K. Kunhyuk, H. Kufluoglu, M. A. Alam, and K. Roy, "Impact of NBTI on the temporal performance degradation of digital circuits," *IEEE Electron Device Letters,* vol. 26, pp. 560-562, 2005.

[29] Y. Cao, J. Velamala, K. Sutaria, M. S. Chen, J. Ahlbin, I. S. Esqueda, M. Bajura, and M. Fritze, "Cross-Layer Modeling and Simulation of Circuit Reliability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 33, pp. 8-23, 2014.

[30] S. Tan, M. Tahoori, T. Kim, S. Wang, Z. Sun, and S. Kiamehr, "Introduction," in *Long-Term Reliability of Nanometer VLSI Systems: Modeling, Analysis and Optimization*, S. Tan, M. Tahoori, T. Kim, S. Wang, Z. Sun, and S. Kiamehr, Eds., ed Cham: Springer International Publishing, 2019, pp. 3-12.

[31] J. H. Stathis, S. Mahapatra, and T. Grasser, "Controversial issues in negative bias temperature instability," *Microelectronics Reliability,* vol. 81, pp. 244-251, 2018/02/01/ 2018.

[32] T. Grasser, K. Rott, H. Reisinger, M. Waltl, F. Schanovsky, and B. Kaczer, "NBTI in Nanoscale MOSFETs—The Ultimate Modeling Benchmark," *IEEE Transactions on Electron Devices,* vol. 61, pp. 3586-3593, 2014.

[33] S. Mahapatra and N. Parihar, "A review of NBTI mechanisms and models," *Microelectronics Reliability,* vol. 81, pp. 127-135, 2018/02/01/ 2018.

[34] J. H. Stathis, "The physics of NBTI: What do we really know?," in *2018 IEEE International Reliability Physics Symposium (IRPS)*, 2018, pp. 2A.1-1-2A.1-4.

[35] J. H. Stathis and S. Zafar, "The negative bias temperature instability in MOS devices: A review," *Microelectronics Reliability,* vol. 46, pp. 270-286, 2006/02/01/ 2006.

[36] N. Parihar, N. Goel, A. Chaudhary, and S. Mahapatra, "A Modeling Framework for NBTI Degradation Under Dynamic Voltage and Frequency Scaling," *IEEE Transactions on Electron Devices,* vol. 63, pp. 946-953, 2016.

[37] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "Understanding negative bias temperature instability in the context of hole trapping (Invited Paper)," *Microelectronic Engineering,* vol. 86, pp. 1876-1882, 2009/07/01/ 2009.

[38] S. Chakravarthi, A. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability," in *2004 IEEE International Reliability Physics Symposium. Proceedings*, 2004, pp. 273-282.

[39] S. Mahapatra, V. Huard, A. Kerber, V. Reddy, S. Kalpat, and A. Haggag, "Universality of NBTI - From devices to circuits and products," in *2014 IEEE International Reliability Physics Symposium*, 2014, pp. 3B.1.1-3B.1.8.

[40] M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, M. Bidaud, G. Ribes, D. Barge, L. Vishnubhotla, B. Tavel, Y. Rey-Tauriac, F. Perrier, N. Revil, F. Arnaud, and P. Stolk, "New hole trapping characterization during NBTI in 65 nm node technology with distinct nitridation processing [MOSFETs]," in *IEEE International Integrated Reliability Workshop Final Report, 2004*, 2004, pp. 121-124.

[41] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "A two-stage model for negative bias temperature instability," in *2009 IEEE International Reliability Physics Symposium*, 2009, pp. 33-44.

[42] J. B. Velamala, K. Sutaria, T. Sato, and Y. Cao, "Physics matters: Statistical aging prediction under trapping/detrapping," in *DAC Design Automation Conference 2012*, 2012, pp. 139-144.

[43] M. Denais, C. Parthasarathy, G. Ribes, Y. Rey-Tauriac, N. Revil, A. Bravaix, V. Huard, and F. Perrier, "On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET's," in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.*, 2004, pp. 109-112.

[44] K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," *Journal of Applied Physics,* vol. 48, pp. 2004-2014, 1977.

[45] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability,* vol. 45, pp. 71-81, 2005/01/01/ 2005.

[46] M. A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, "A comprehensive model for PMOS NBTI degradation: Recent progress," *Microelectronics Reliability,* vol. 47, pp. 853-862, 2007/06/01/ 2007.

[47] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative-Bias Temperature Instability: Initial Degradation, Field Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *IEEE Transactions on Electron Devices,* vol. 54, pp. 2143-2154, 2007.

[48] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam, "On the dispersive versus arrhenius temperature activation of nbti time evolution in plasma nitrided gate oxides: measurements, theory, and implications," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, 2005, pp. 684-687.

[49] S. Mukhopadhyay and S. Mahapatra, "An Experimental Perspective of Trap Generation Under BTI Stress," *IEEE Transactions on Electron Devices,* vol. 62, pp. 2092-2097, 2015.

[50] S. Rangan, N. Mielke, and E. C. C. Yeh, "Universal recovery behavior of negative bias temperature instability [PMOSFETs]," in *IEEE International Electron Devices Meeting 2003*, 2003, pp. 14.3.1-14.3.4.

[51] H. Aono, E. Murakami, K. Okuyama, A. Nishida, M. Minami, Y. Ooji, and K. Kubota, "Modeling of NBTI saturation effect and its impact on electric field dependence of the lifetime," *Microelectronics Reliability,* vol. 45, pp. 1109-1114, 2005/07/01/ 2005.

[52] V. Huard and M. Denais, "Hole trapping effect on methodology for DC and AC negative bias temperature instability measurements in PMOS transistors," in *2004 IEEE International Reliability Physics Symposium. Proceedings*, 2004, pp. 40-45.

[53] A. E. Islam, G. Gupta, S. Mahapatra, A. T. Krishnan, K. Ahmed, F. Nouri, A. Oates, and M. A. Alam, "Gate Leakage vs. NBTI in Plasma Nitrided Oxides: Characterization, Physical Principles, and Optimization," in *2006 International Electron Devices Meeting*, 2006, pp. 1-4.

[54] T. Grasser, W. Gos, V. Sverdlov, and B. Kaczer, "The Universality of NBTI Relaxation and its Implications for Modeling and Characterization," in *2007 IEEE International Reliability Physics Symposium Proceedings. 45th Annual*, 2007, pp. 268-280.

[55] 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 Transactions on Device and Materials Reliability,* vol. 11, pp. 19-34, 2011.

[56] 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 *2006 IEEE International Reliability Physics Symposium Proceedings*, 2006, pp. 448-453.

[57] Z. Q. Teo, D. S. Ang, and C. M. Ng, "“Non-Hydrogen-Transport” Characteristics of Dynamic Negative-Bias Temperature Instability," *IEEE Electron Device Letters,* vol. 31, pp. 269-271, 2010.

[58] P. Hehenberger, P. J. Wagner, H. Reisinger, and T. Grasser, "On the temperature and voltage dependence of short-term negative bias temperature stress," *Microelectronics Reliability,* vol. 49, pp. 1013-1017, 2009/09/01/ 2009.

[59] C. Shen, M. Li, C. E. Foo, T. Yang, D. M. Huang, A. Yap, G. S. Samudra, and Y. Yeo, "Characterization and Physical Origin of Fast Vth Transient in NBTI of pMOSFETs with SiON Dielectric," in *2006 International Electron Devices Meeting*, 2006, pp. 1-4.

[60] T. Grasser and B. Kaczer, "Negative bias temperature instability: Recoverable versus permanent degradation," in *ESSDERC 2007 - 37th European Solid State Device Research Conference*, 2007, pp. 127-130.

[61] H. Reisinger, R. P. Vollertsen, P. J. Wagner, T. Huttner, A. Martin, S. Aresu, W. Gustin, T. Grasser, and C. Schlunder, "The effect of recovery on NBTI characterization of thick non-nitrided oxides," in *2008 IEEE International Integrated Reliability Workshop Final Report*, 2008, pp. 1-6.

[62] T. Grasser, 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 *2010 IEEE International Reliability Physics Symposium*, 2010, pp. 16-25.

[63] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. T. T. Luque, and M. Nelhiebel, "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction–Diffusion to Switching Oxide Traps," *IEEE Transactions on Electron Devices,* vol. 58, pp. 3652-3666, 2011.

[64] J. F. Zhang, Z. Ji, and W. Zhang, "As-grown-generation (AG) model of NBTI: A shift from fitting test data to prediction," *Microelectronics Reliability,* vol. 80, pp. 109-123, 2018/01/01/ 2018.

[65] A. A. Katsetos, "Negative bias temperature instability (NBTI) recovery with bake," *Microelectronics Reliability,* vol. 48, pp. 1655-1659, 2008/10/01/ 2008.

[66] A. Benabdelmoumene, B. Djezzar, A. Chenouf, B. Zatout, and M. Kechouane, "Analysis of NBTI Degradation in nMOS-Capacitors and nMOSFETs," *IEEE Transactions on Device and Materials Reliability,* vol. 18, pp. 583-591, 2018.

[67] B. Djezzar, H. Tahi, A. Benabdelmoumene, A. Chenouf, M. Goudjil, and Y. Kribes, "On the permanent component profiling of the negative bias temperature instability in p-MOSFET devices," *Solid-State Electronics,* vol. 106, pp. 54-62, 2015/04/01/ 2015.

[68] H. Tahi, B. Djezzar, A. Benabdelmoumene, A. Chenouf, and M. Goudjil, "Investigation of interface, shallow and deep oxide traps under NBTI stress using charge pumping technique," *Microelectronics Reliability,* vol. 54, pp. 882-888, 2014/05/01/ 2014.

[69] A. Benabdelmoumene, B. Djezzar, H. Tahi, A. Chenouf, M. Goudjil, R. Serhane, F. H. Larbi, and M. Kechouane, "Does NBTI effect in MOS transistors depend on channel length?," in *2014 26th International Conference on Microelectronics (ICM)*, 2014, pp. 52-55.

[70] B. Djezzar, H. Tahi, A. Benabdelmoumene, and A. Chenouf, "A propagation concept of negative bias temperature instability along the channel length in p-type metal oxide field effect transistor," *Solid-State Electronics,* vol. 82, pp. 46-53, 2013/04/01/ 2013.

[71] A. Benabdelmoumene, B. Djezzar, A. Chenouf, H. Tahi, B. Zatout, and M. Kechouane, "On the turn-around phenomenon in n-MOS transistors under NBTI conditions," *Solid-State Electronics,* vol. 121, pp. 34-40, 2016/07/01/ 2016.

[72] B. Djezzar, A. Benabdelmoumene, B. Zatout, D. Messaoud, A. Chenouf, H. Tahi, M. Boubaaya, and H. Timlelt, "Recovery investigation of NBTI-induced traps in n-MOSFET devices," *Microelectronics Reliability,* vol. 110, p. 113703, 2020/07/01/ 2020.

[73] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journal of Solid-State Circuits,* vol. 25, pp. 584-594, 1990.

[74] S. Khan and S. Hamdioui, "Modeling and mitigating NBTI in nanoscale circuits," in *2011 IEEE 17th International On-Line Testing Symposium*, 2011, pp. 1-6.

[75] B. Kaczer, S. Mahato, V. V. d. A. Camargo, M. Toledano-Luque, P. J. Roussel, T. Grasser, F. Catthoor, P. Dobrovolny, P. Zuber, G. Wirth, and G. Groeseneken, "Atomistic approach to

variability of bias-temperature instability in circuit simulations," in *2011 International Reliability Physics Symposium*, 2011, pp. XT.3.1-XT.3.5.

[76] V. V. A. Camargo, B. Kaczer, T. Grasser, and G. Wirth, "Circuit simulation of workload-dependent RTN and BTI based on trap kinetics," *Microelectronics Reliability,* vol. 54, pp. 2364-2370, 2014/11/01/ 2014.

[77] H. Kükner, S. Khan, P. Weckx, P. Raghavan, S. Hamdioui, B. Kaczer, F. Catthoor, L. V. d. Perre, R. Lauwereins, and G. Groeseneken, "Comparison of Reaction-Diffusion and Atomistic Trap-Based BTI Models for Logic Gates," *IEEE Transactions on Device and Materials Reliability,* vol. 14, pp. 182-193, 2014.

[78] H. Amrouch, B. Khaleghi, A. Gerstlauer, and J. Henkel, "Reliability-aware design to suppress aging," in *2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2016, pp. 1-6.

[79] J. P. Uyemura, "The CMOS Inverter," in *Circuit Design for CMOS VLSI*, J. P. Uyemura, Ed., ed Boston, MA: Springer US, 1992, pp. 79-113.

[80] "Memory System Design I," in *Fundamentals of Computer Organization and Architecture*, ed, 2004, pp. 107-133.

[81] J. Warnock, Y. H. Chan, H. Harrer, D. Rude, R. Puri, S. Carey, G. Salem, G. Mayer, Y. Chan, M. Mayo, A. Jatkowski, G. Strevig, L. Sigal, A. Datta, A. Gattiker, A. Bansal, D. Malone, T. Strach, H. Wen, P. Mak, C. Shum, D. Plass, and C. Webb, "5.5GHz system z microprocessor and multi-chip module," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2013, pp. 46-47.

[82] T. Endoh, H. Koike, S. Ikeda, T. Hanyu, and H. Ohno, "An Overview of Nonvolatile Emerging Memories— Spintronics for Working Memories," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* vol. 6, pp. 109-119, 2016.

[83] "SRAM Circuit Design and Operation," in *CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies: Process-Aware SRAM Design and Test*, A. Pavlov and M. Sachdev, Eds., ed Dordrecht: Springer Netherlands, 2008, pp. 13-38.

[84] J. Wang, S. Nalam, and B. H. Calhoun, "Analyzing static and dynamic write margin for nanometer SRAMs," in *Proceeding of the 13th international symposium on Low power electronics and design (ISLPED '08)*, 2008, pp. 129-134.

CHAPTER 3 CIRCUIT-LEVEL NBTI CHARACTERIZATION

3. 1 Introduction

At device level, many research works have been conducted to study the fundamentals behind the NBTI effect on pMOS transistor, developing fast measurement techniques to capture and separate the degradation components under both static and dynamic stresses [1-7]. This has made modeling theories for the NBTI degradation to be still under debate [8-10]. However, the focus on the low-level details does not provide insight into the NBTI-induced degradation on a larger scale. This is due to the fact that NBTI degradation on circuit performance is mainly dependent on the type of the circuit, its workload and its application rather than the absolute value of the degradation at the device level [11-15]. Jha *et al*. [11] analyzed NBTI impact on many analog circuits, and found that, the cascade current mirror undergoes the maximum drift (12.5%) in the output current while the constant g_m biasing circuit experiences the least degradation (1%) for the same amount of NBTI-induced Vth shift. In addition, they revealed the two-stage CMOS operational amplifier (op-amp) to be slightly affected by △Vth while it undergoes a considerable offset voltage variation when it is used as a comparator [11]. Moreover, variations in biasing currents due to NBTI have resulted in a gain error for the digital analog converter (DAC). Subsequently, they concluded that NBTI can present a serious reliability concern for such circuits, even a small variation in bias currents can cause significant gain errors in circuits like DACs [11].

NBTI-induced delay degradation of random logic circuits analysis using static timing analysis (STA) of ISCAS'85 benchmark circuits synthesized with 65nm PTM, Kang *et al.* [12] showed that NBTI can result in the degradation of the operating frequency f_{max} up to 8.8% in 3 years at 125°C. They also showed that it can be reduced to 3.3% if the operating temperature was lowered to 25°C [12]. They demonstrated that unmatched signal probability between the two pMOS transistors in a 6T-SRAM might result in nearly 30% reduction of its static noise margin (SNM) [12].

Khan *et al.* [13] analyzed the BTI degradation in both basic and complex gates by considering the impact of the inputs' duty cycle and the frequency at which they switch as well as location of the stressed transistor within the gate/circuit. They found that BTI is gate dependent and more pronounced in complex gates [13]. They also found that static NBTI (PBTI) impact is more important in NOR (NAND) gates than in NAND (NOR) gates. Yet, the delay of NAND gate has been more affected by low duty-cycle NBTI than by PBTI. In fact, the delay shift induced by NBTI, with 20% duty-cycle in 10 years, is about 10.75% while it is induced by PBTI is just of 4.13%. [13]. In addition, they showed that BTI is independent of the operating frequency even in complex logic gates. The delay degradation due to NBTI is 2-3 times higher than that due to PBTI even at 45 nm [13].

Eghbalkhah *et al.*[14] analyzed the impact of both die temperature and workload variation on NBTI aging and lifetime prediction of ISCAS'89 and ITC'99 benchmark suits. They showed that the thermal and activity profiles within a circuit can significantly affect the degradation rate caused by NBTI [14]. In fact, their simulation results revealed that the predicted timing degradation have errors ranging from -135% to +98% if the dynamic operating conditions are not considered.

Mintarno *et al.*[15] analyzed the impact of aging in an industrial processor core using two sub-45nm aging models, reaction-diffusion (RD) and trapping-detrapping (TD). They found that aging effects depend significantly on cell-topology and library parameters. By applying the same input slew, output load, and worst-case RD threshold voltage degradation, they found that the shift of the delay and the output transition time for all 700 standard-cells and >11K timing-arcs (logic condition, driving pin, output pin, rising-falling) varied widely from 1.2% to 20.2% [15]. Similarly, paths ranking was significantly influenced and the timing degradations among them

were non-uniform, as for a sample of > 660K critical paths, the largest degradation was found to be 17.6% while the smallest one to be 4.9%, with an average of 9.5% [15].

Nigam [16] studied the relaxation of NBTI stress on pMOS device subjected to pulse stress, from a circuit perspective, using a ring oscillator (RO) as CUT (circuit under test). He found that frequency degradation in the RO is 5 times less than in DC NBTI stress devices, and the time exponent of the frequency degradation differs from 0.16. In fact, he found that it is 0.12 for 100-MHz RO and 0.18 for 3 GHz RO [16].

Tsai *et al.* [17] found that the AC to DC NBTI lifetime ratio extracted from a RO is more than 10000 (for frequency near 1GHz). Hence, due to the dynamic nature of digital circuits operation, the prediction of digital circuit lifetime based only on DC NBTI would overestimate the degradation of circuit performance [16, 17]. As such, an extra effort has to be made to bridge the gap between circuit and device reliability analysis. A few works have dealt with experimental analysis of NBTI at circuit level and correlate it with that of device level [18] [19-22]. In fact, Fernandez *et al*. [18] found that the shift of the Inverter logic threshold to be governed by the shift of pMOS threshold voltage. Moreover, it is independent of the frequency in the entire 1 Hz - 2 GHz range [18, 19]. They found that the Vth shift due to AC NBTI was following the same trend of that due DC NBTI, that is a power law, but with time exponent of 0.17-0.20 [19]. However, we found that pMOS V_{th} does not wholly capture the degradation of the Inverter logic threshold voltage under certain conditions [22].

Pong-Fei Lua et *al.* [21] analyzed long-term NBTI degradation under real-use conditions in IBM microprocessors, and found that the frequency shift of a ring oscillator over 700 days of work is consistent with models based on accelerated testing of pMOS transistor and that the degradation showed a power law with different time exponents of 0.174, 0.211 and 0.186 [21]. In the same prospect, we show with experimental evidence how NBTI degradation at circuit-level is dependent on the circuit configuration and its operation mode. We particularly show how the degradation time dependence varies with voltage and temperature rather than showing a constant value due to the switching nature of the circuit. Moreover, when correlated with device-level degradation, the analysis of the shift of CMOS Inverter DC response and performance due to AC NBTI revealed a signature of the co-existence of more than one physical mechanism behind NBTI degradation where the dominance of one mechanism over another depends on stress time, voltage, and temperature conditions.

The remaining sections of the chapter is organized as follows. In section 2, the circuit under test, the experimental DC and AC NBTI setup details, and the Measure/Stress/Measure (MSM) protocol are described. The obtained results for NBTI characterization on the Inverter DC response and performance are analyzed in section 3. These results are further discussed in section 4 and 5 to highlight the correlation between the device-circuit NBTI degradation while section 6 concludes the chapter.

3. 2 Experimental Setup for Reliability Analysis of the CMOS Inverter

A. The circuit under test (CUT)

CUTs consist of CMOS Inverters made with nMOS and pMOS transistors with aspect ratios of 1µm/0.8µm and 2µm/0.8µm, respectively. These transistors have 20 nm SiO_2 gate oxide grown with dry O_2 using a conventional CMOS process at ISiT (Institute for Silicon Technology) of Fraunhofer, Germany (the CUT shown in Fig. 3.1).

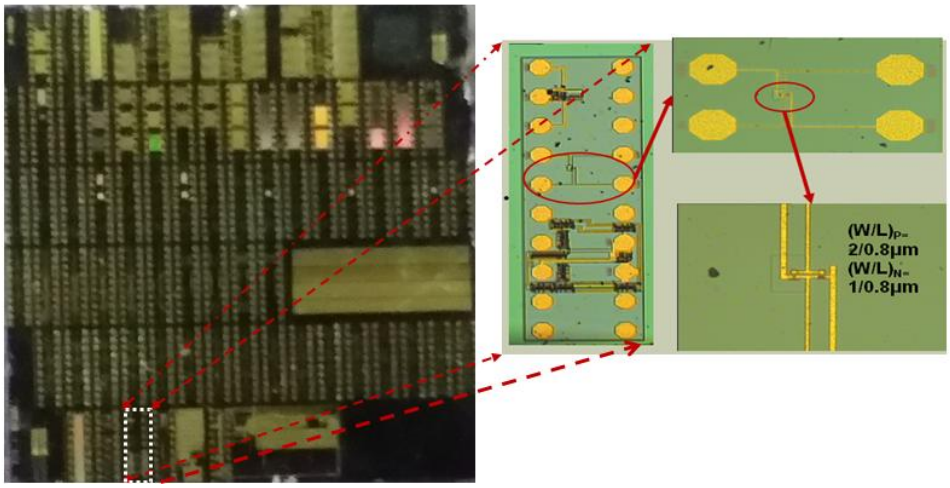


Fig. 3.1: Circuit under test: CMOS Inverter.

It is worth to note that the thick oxide of our test structures does not significantly affect the study as NBTI degradation depends on the electric field but not on the oxide thickness. In fact, Gregor Pobegen *et al.*[23] have conducted a study on devices with different oxide thicknesses (5 to 30 nm) and they have shown experimentally that the basic mechanisms behind the NBTI are the same in thin and thick oxide technologies. However, the devices of different gate oxide thicknesses T_{ox} show the same parameter degradation at the same oxide field E_{ox}.

B. The test protocol for NBTI characterization of the Inverter DC features

The experimental setup is based on measure/stress/measure (MSM) technique, where the sequences have been performed using fully automated bench. The bench included an Agilent HP 4156C Semiconductor Parameter Analyzer for both DC polarization and stress of CUTs, an Agilent 16440A SMU/Pulse Generator Selector to ensure automatic switch of the output between stress and measurement, and a Tektronix Digital Oscilloscope TDS 3054 to track the Inverter temporal response, a Karl Suss PA300 micro-manipulator probe station, and a PID-controlled hotplate to vary the temperature of the chip during the experiments. The test circuit chip (bare-die) within the probe station is isolated from vibration and enclosed in a grounded faraday cage to avoid both RF and light effects as shown in Fig. 3.2.

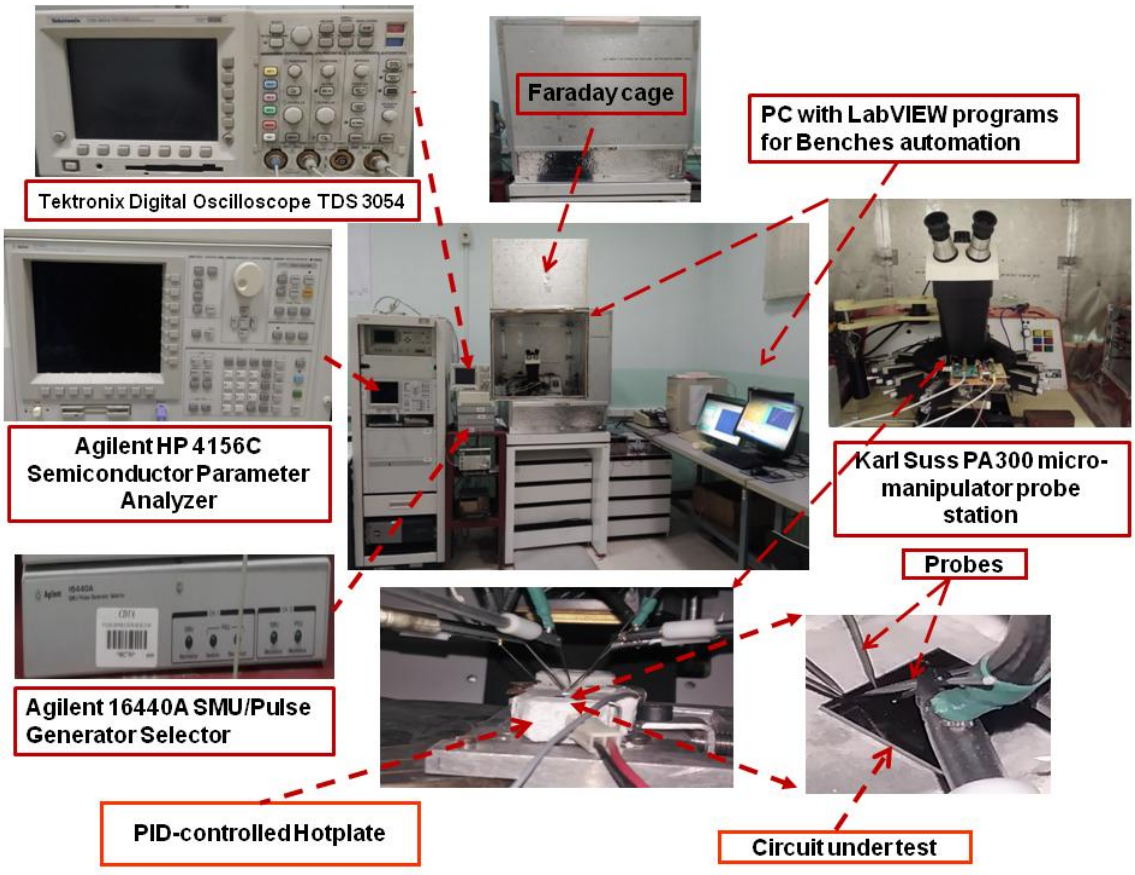


Fig. 3.2: NBTI Characterization test bench for the CMOS Inverter.

The MSM was conducted by making a measure at room temperature of the virgin (unstressed) Inverter's voltage transfer curve (VTC). Then, the Inverter undergoes a DC (AC pulse signal of 0.5 duty-cycle and 10 kHz) NBTI stress for a period of 3600s. Under NBTI DC stress, negative voltages of -8V,-10V,-13 and -14V were applied to the Inverter's input (inferring negative fields

ranging from -4MV/cm to -7MV/cm) whereas a pulse signal with the same magnitude of -8 to -14V was applied to the Inverter input during NBTI AC stress while the of CUT pads were grounded as shown in Fig. 3.3. After the stress period, another measure of the Inverter's VTC is done and their data has been saved to be compared with that before stress. This test protocol has been repeated at different temperatures in the range 80°C-140°C with a 20°C step using a PID-controlled hotplate in order to combine large field with elevated temperature. The obtained results are discussed in next section 3.3.

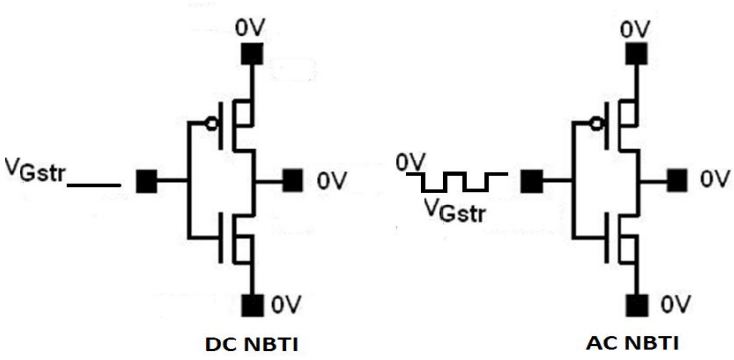


Fig. 3.3: CMOS Inverter configuration under DC & AC NBTI stresses.

C. The test protocol for NBTI characterization of the Inverter transient response

A stress framework of 3 hours and half (12600 s) was used to characterize the CMOS Inverter temporal response. During the stress period, an AC pulse signal of 10 kHz and 0.5 duty-cycle with a V_{Gstr} ranging between -8 and -14 V was applied to the input of the Inverter while the remaining pins of the circuit were grounded (see left-side of Fig.3.4).

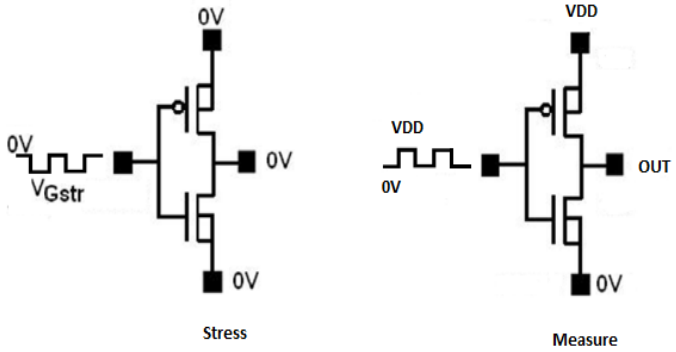


Fig. 3.4: Stress & measure configurations used to characterize the CMOS Inverter temporal performance under AC NBTI stress.

During the measurement using a positive voltage, a square pulse with the same frequency and duty cycle of AC NBTI stress was applied to the Inverter input (see right-side of Fig. 3.4) before and after the stress. A measure of the Inverter temporal response was done and saved after each 1800 s of the stress time which lasted for a total period of 3 hours and half. The switch between stress and measure configurations was automated via a LabVIEW program. The obtained results are given in section 3.4

3.3 NBTI Impact on CMOS Inverter DC Features

Figure 3.5 shows the shift of the Inverter DC response (VTC) due to both DC and AC NBTI stresses at room temperature in (a) and at elevated temperature in (b). Solid symbols correspond to DC NBTI, whereas open symbols correspond to the AC NBTI. Virgin curves denote VTCs of

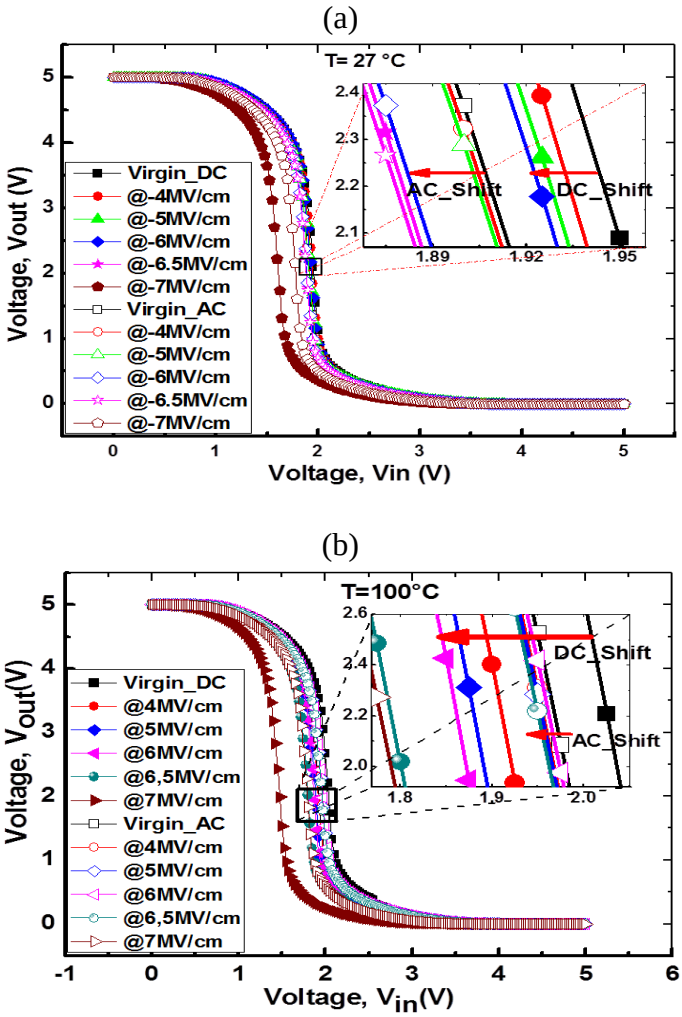


Fig. 3.5: CMOS Inverter VTC shift subjected to DC & AC NBTI stresses at room temperature in (a) and at elevated temperature in (b). [24]

unstressed Inverters and the other curves correspond to the VTCs of the stressed ones with a series of electric field ranging from -4 to -7MV/cm [24].

It is clear that a VTC shift (with respect to unstressed VTC) caused by DC NBTI stress is larger than caused by AC NBTI stress at both room temperature and elevated temperature but with a higher magnitude at the later. This is due to the fact that while the Inverter undergoes a field stress at room temperature, the stress is exacerbated by the high temperature (100°C), resulting in a negative bias temperature instability of the inerter response..

It is worth to note that DC NBTI-induced shift of the Inverter`s VTC increases with negative electric field as shown in (a) and with temperature as shown (b) in Fig. 3.5. However, AC NBTI-induced VTC`s shift does increase with field in (a) but not, too much, with the temperature (b). This is due to the fact that temperature could accelerate the recovery during the AC stress. This observation will be more analyzed later.

To understand more about the difference between AC and DC NBTI stresses impact on the CMOS Inverter VTC, we have tracked their impact on its main critical voltages in [24]. These voltages correspond to the Inverter`s logic (switching) threshold voltage, valid input and output logic levels and noise margins which characterize the Inverter circuit robustness. These voltages are extracted from the VTC as shown in Fig. 3.6 [25]. The detailed comparison between DC and AC NBTI impacts on the Inverter's robustness can be found in [24]. Hence, we review here the shift of one of the main Inverter DC features that is ΔVinv and its trend under AC NBTI stress in the prospect to correlate it with PMOS threshold voltage Vth shift (ΔVthp) [26].

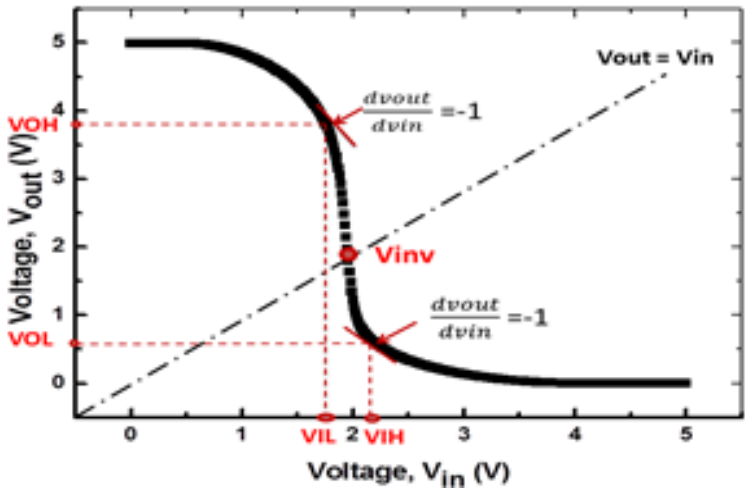


Fig. 3.6: The critical voltages of the CMOS Inverter VTC [25].

The shift of the CMOS Inverter logic (switching) threshold voltage (Vinv) under both DC- & AC-type NBTI in semi-log scale, with respect to the energy in (a) and to the negative field in (b), is depicted in Fig.3.7. The solid symbols represent experimental data of the shift due to DC NBTI, while open symbols represent those due to AC-type stress. In the same way, the solid and dashed lines are the linear fit of DC and AC data, respectively.

The obtained results show that Vinv undergoes under AC NBTI, a shift quasi-parallel of that under DC stress, but with a lower degradation magnitude. Actually, the AC slopes are slightly lower, within a range of 1 %, than DC slopes.

It can be noted that from the slopes of the fitted data of the Vinv shift in (a) and (b) of Fig. 3.7,

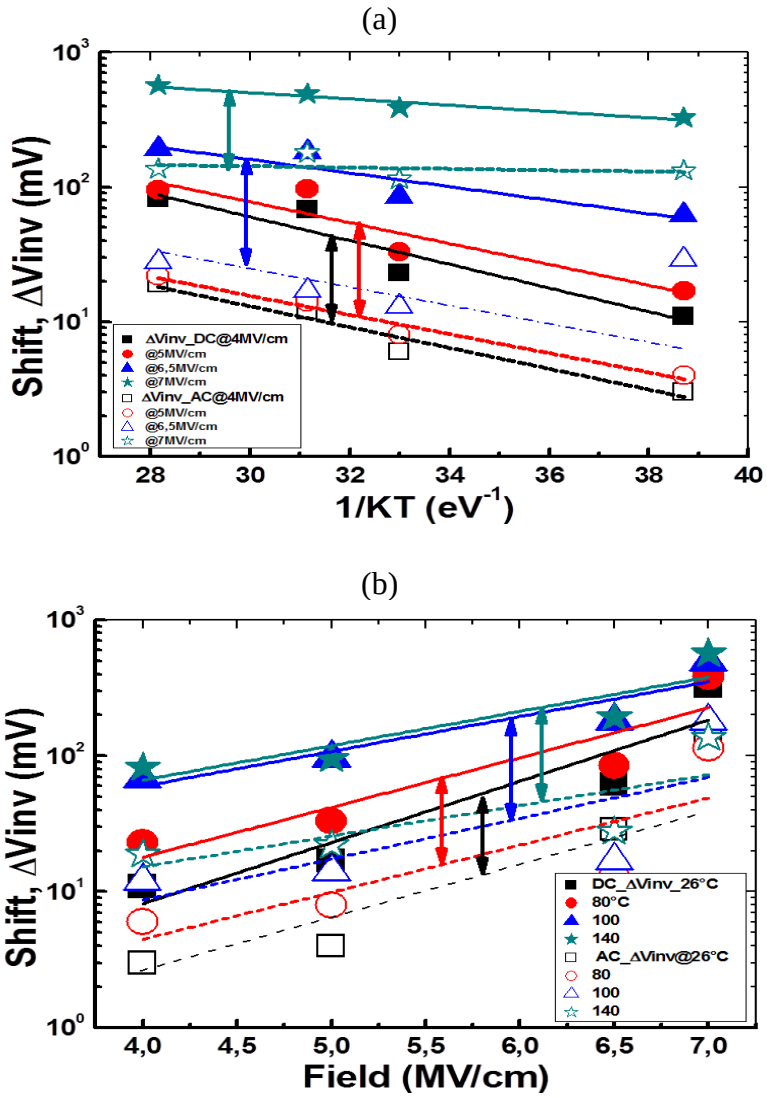


Fig.3.7 Inverter switching threshold shift under DC & AC-type NBTI stresses with respect to the energy (a) and the applied field (b) [25].

the influence of the electric field (s = 0.2-0.4) is much more important than the temperature (0.1-0.04). In fact, the shift slopes of the fitted data due to the applied negative field (Fig.3.7. (b)) are almost 5 times superior to those due to the temperature (Fig.3.7.(a)). Moreover, the shift becomes temperature independent under both DC and AC stresses when the field exceeds 6.5 MV/cm as the ΔVinv slope corresponding to 7 MV/cm (see Fig.3.7 (a)) is about 0.04 for DC and 0.01 for AC stress.

Furthermore, the Inverter undergoes under a 0.5 duty-cycle AC NBTI, the half stress time of that under DC NBTI, and it is supposed that the former induces a shift equals to the half of that induced by the later. However, this has not been observed here (in these experiments) as it is found much less than a half. In fact, the magnitude of the switching logic shift ΔVinv caused by AC stress represents a 0.25 to 0.46 of the shift caused by DC stress (see Fig.3.8).

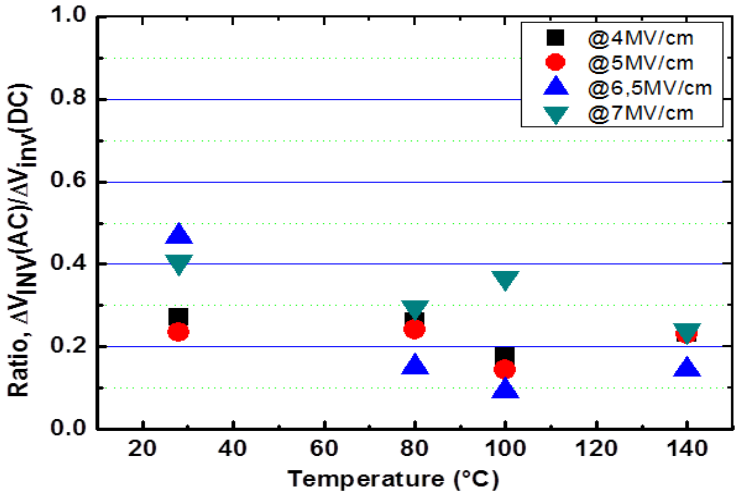


Fig. 3.8: AC/DC ratio of the Inverter switching threshold voltage shift [25].

This ratio is much less than that expected in the literature and reported elsewhere for a 0.5 duty-cycle AC stress [27]. This ratio tends to decrease with both temperature and negative field E_{ox} < 7 MV/cm. This ratio`s decrease could be attributed to both the recovery through bias switches, during the half period of the AC stress, and temperature acceleration on recovery [27]. This holds true except for the field of 7MV/cm for which the activation energy drastically drops to 0.01 while it is around 0.16 for lower field values (as shown in Fig.3.9). This implies that another physical process is coming out and which could be related to the creation of fixed traps in the oxide and these charges do not switch or recover after releasing (DC) or alerting (AC) NBTI stress. This leads to think this creation does not have to be attributed to the pure NBTI

degradation and it has only an additional effect in a highly accelerated field stress. So, if we take this in consideration, the AC NBTI-induced shift of ΔVinv in Fig. 3.7(b) can be replotted as shown in Fig.3.10.

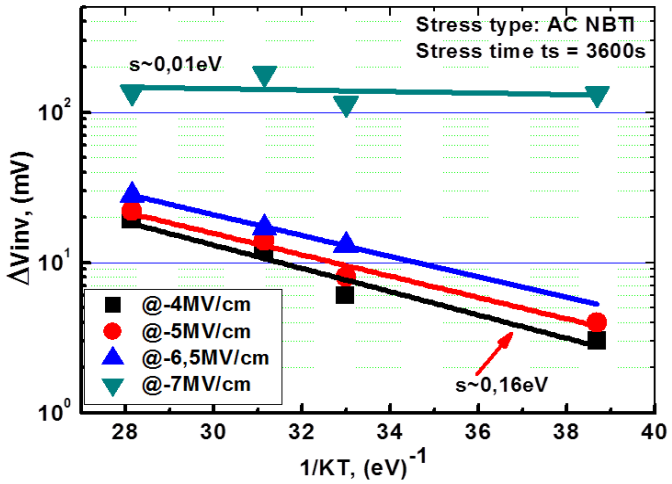


Fig. 3.9 Inverter switching threshold shift due to AC NBTI with respect to the energy [26].

Based on results depicted in Fig. 3.9 and Fig. 3.10, it is clear that ΔVinv shows two tendencies: degradation with low field acceleration (0.15 cm/MV) associated with high activation energy (0.16 eV) then degradation with high field acceleration (4 cm/MV) and low activation energy (0.01 eV).

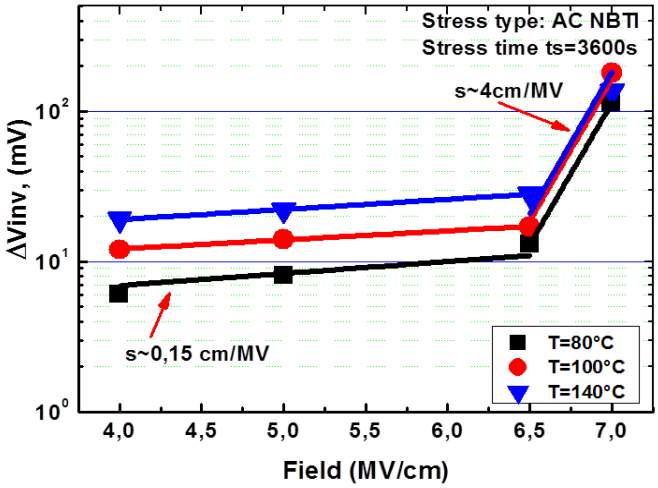


Fig. 3.10: Inverter logic threshold shift due to AC NBTI with respect to the applied field. [26]

To give more insight about this assumption, we examine, in the following, ΔVinv against ΔVthp. The shift of the logic threshold under NBTI can be expressed in terms of the shift of the pMOS threshold voltage as (ref to chapter 2, section 2.4.1):

$$\Delta V_{inv} = \frac{r}{r+1}\Delta V_{thp} \tag{3.1}$$

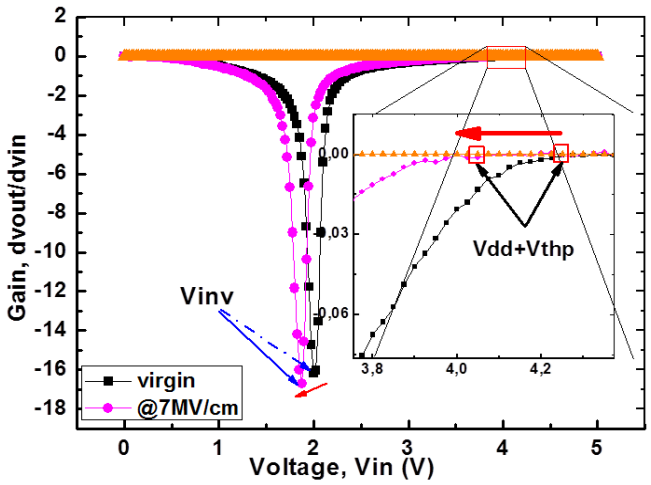


Fig. 3.11: Method of extracting PMOS threshold (Vthp) from the Inverter DC gain [26].

Based on Eq. (3.1) and given r is constant, then one can expect, if governed by the same physical mechanisms, ΔVinv and ΔVthp under NBTI to show similar temperature dependencies. For this aim, we have extracted the pMOS threshold voltage from the Inverter VTC before and after stress by differentiating this latter and resolving where the slope drops down to zero as shown in Fig. 3.11. The corresponding shift with respect to the energy is depicted in Fig. 3.12

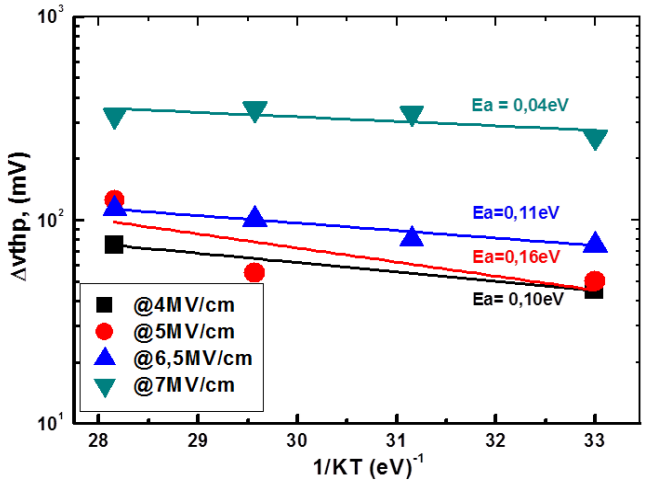


Fig. 3.12: The shift of the extracted PMOS threshold voltage under AC NBTI with respect to the energy under different electric field [26].

It is clear that ΔVthp is larger in magnitude than ΔVinv, but they are with slightly different activation energies. In fact, ΔVthp shows activation energy (Ea) of 0.11 eV against 0.16 eV of ΔVinv under AC NBTI with $E_{ox} \leq 6$ MV/cm. Then, it shows an Ea of 0.04 eV against 0.01 eV of ΔVinv with Eox >6.5MV/cm. This slight difference can be attributed to the influence of the shift of nMOS threshold (V_{thn}) and that of the transistor's transconductances on the measured ΔVinv. We can conclude that the Inverter threshold shift is governed by the same physical mechanisms behind the pMOS threshold shift. The range of the activation energy (Ea) found here for ΔVthp (0.11 eV) and 0.04 eV) is consistent with values of (0.063 eV) and (0.115 eV) reported by Huard *et al*. [28] and Ang *et al.* [5], respectively.

Given the discrepancy between Ea of ΔVthp and that associated with the interface states, Huard and Ang made evidence of the co-existence of two mechanisms of the degradation. The first one is related to hole trapping with a low Ea (0.02 eV) [28] and (0.06 eV) [5] and the second one is related to the interface states generation with a higher Ea (0.16 eV) [28], and (0.144 eV) [5]. Following the same analysis, we can say that the discrepancy of activation energies shown by both ΔVinv and ΔVthp (see Figs 3.9 and 3.12 under Eox $\leq$ 6.5 MV/cm (~0.16 eV) and under a higher oxide field; Eox > 6.5 MV/cm (~0.01 eV), would advise on the signature of more than one mechanism behind the degradation which is dominated by interface states under low oxide field and by hole trapping in new generated traps when the oxide field goes higher.

On the other side, the first Ea found here for ΔV_{thp} under Eox $\leq$ 6.5 MV/cm that is 0.11-0.16 eV matches with the Ea measured by Huard *et al.* [28] and Denais *et al.* [29] for ΔVthp, at the end of the recovery phase of dynamic NBTI, and by Ang et al.[5], after a dynamic NBTI of 30 cycles of stress. They found this energy to be consistent when interface traps are extracted using charge pumping method. So, they concluded that the permanent component of dynamic NBTI is attributed to interface traps. However, if the permanent component NBTI is only due to interface traps, then how would we explain the large ΔVthp with low activation energy undergone when the oxide field goes higher than 6.5 MV/cm in our experiment? There might be another component which slowly recovers like what interface states do but it has low temperature dependence. Relying on the resemblance between this latter and that reported for hole trapping [5] [28], we can say that this part may be related to hole traps, with large recovery times. Accordingly, we can say the characterization of the Inverter DC response subjected to dynamic NBTI stress has advised us on the existence of more than one physical mechanism behind the NBTI degradation.

3.4 NBTI impact on CMOS Inverter temporal performance

After analyzing voltage and temperature dependencies of the NBTI-induced degradation, its dependence on stress time is also analyzed in this section. For this aim, one Inverter was used for each (voltage, temperature) pair. Each Inverter was stressed for 3 hours and half by a 10 kHz, 0.5 duty-cycle AC NBTI pulse.

Figure 3.13 shows the temporal response of the CMOS Inverter subjected to an AC NBTI stress. The out_virgin curve denotes the Inverter's temporal response before applying any stress, while the other curves correspond to its response after 600, 1800, 3000, and 3600 s of the stress time.

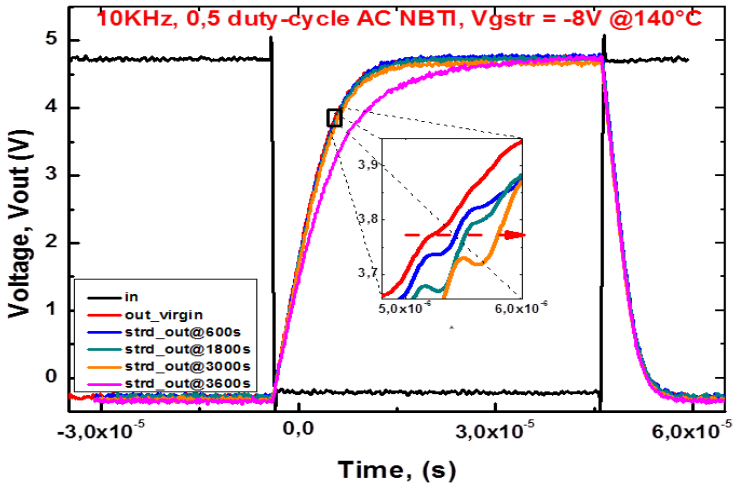


Fig. 3.13 CMOS Inverter temporal response subjected to AC NBTI stress. [26]

It can be observed that the low-to-high switching feature of the Inverter response experiences an increasing shift towards the right-side. This shift becomes more appearent when the stress time increases. Subsequently, two temporal features of the Inverter might be influenced by this shift: the risetime (Tr) and the low-to-high delay (D_{LH}). The former corresponds to the time that takes the output to rise from 10% to 90% of its steady state value, while the latter corresponds to the time that takes the output to reach 50% of its steady state value after the input has crossed it. In this section, we focus on the analysis of the risetime shift under AC NBTI.

To obtain a first-order estimation of the switching time, a simple RC model can be used. During the risetime interval, the Inverter's nMOS transistor is cutoff, while its pMOS transistor is conducting from the power supply. In such model, the nonlinear resistance of pMOS is approximated by its best-case value where the transistor is assumed to be saturated. Then, the Inverter risetime can be expressed as [30]:

$$T_r = \frac{c_{Load} \times 2V_{DD}}{\beta_p (v_{DD} - |v_{thp}|)^2} = \frac{K}{(v_{DD} - |v_{thp}|)^2} \quad (3.2)$$

where V_{DD} is the supply voltage, C_{load} is the Inverter load capacitance, and βp is the pMOS gain factor which is given by:

$$\beta_p = \mu_p c_{ox} \left(\frac{w}{l}\right)_p \quad (3.3)$$

By differentiating the risetime in Eq.(3.2) with respect to Vthp, the relative risetime shift can be expressed in terms of $\Delta Vthp$ as:

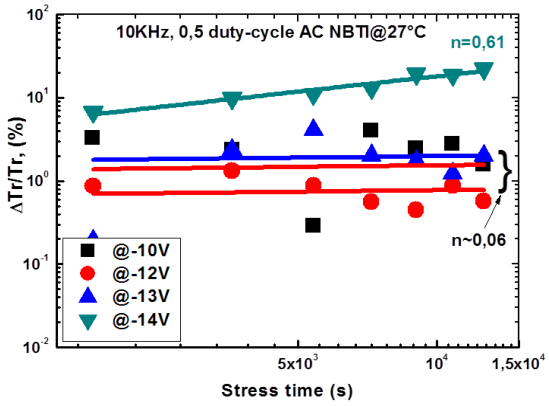
$$\frac{\Delta T_r}{T_r} = 2 \frac{\Delta v_{thp}}{(v_{DD} - |v_{thp0}|)} \quad (3.4)$$

According to Eq. (3.4), the relative shift of the Inverter's risetime follows a similar time dependence as that of ΔVthp under NBTI stress. Hence, if the shift of Vthp follows a power law with stress time, then the shift of the relative risetime would do too. To verify such statement, we have ploted, in log-log scale, the shift of extrated risetime with respect to the stress time at different tempratures and under different voltages as shown in Fig. 3.14.

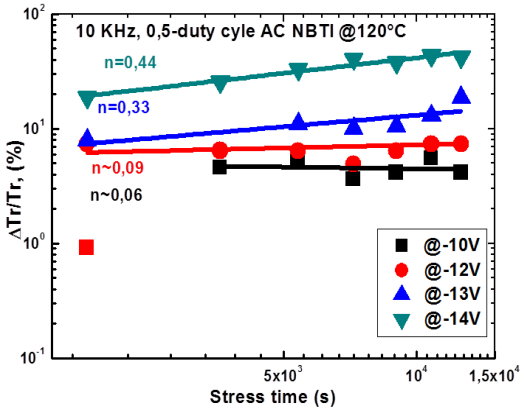
Two different behaviors can be observed. The first one is at room temperature and under low AC NBTI stress voltage. The second is revealed under higher voltage and at elevated temperature. In the first case (see Fig. 3.14(a)), under an AC NBTI stress at room temperature, the risetime shift is trivial and remains quasi constant with time stress under different stress voltages, except under -14 V. The same trend can be observed at elevated temperature when the stress voltage is set at -10 V (-5 MV/cm) (see the fitted data in black line (b) and (c) of Fig. 3.14). The second behavior is observed when the AC NBTI pulse voltage is set at a higher value than 10 V and the temperature is elevated above 100 °C. In such case, the amount of the risetime shift becomes more important and starts to increase with stress time increasing. In fact, under an AC NBTI of -12 V (-6 MV/cm) the degradation does not show a time exponent (n) higher than 0.1 until the temperature is elevated to 140°C (Fig. 3.14 (c)). However, for an AC NBTI stress voltage of -13 V (-6.5 MV/cm), a temperature of 120 °C was large enough to induce a degradation with a similar time exponent (Fig. 3.14. (b)). So, unlike other works conducted at circuit level which found that the degradation showing a power law with time exponent n of 0,16 [31], the time exponent found in our experiment framework does not show a constant value.

Acoordingly, we can wonder on the voltage and temperature dependence of the time exponent. We further analyse the observed trends in the following section.

(a)



(b)



(c)

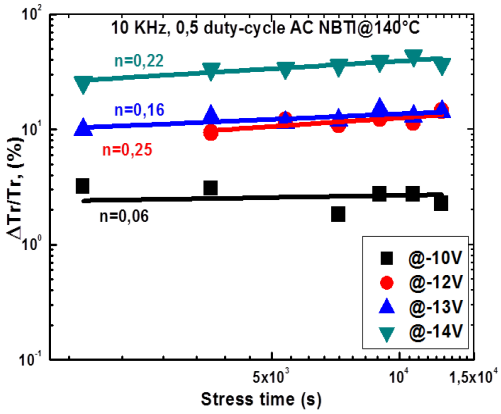


Fig. 3.14: The shift CMOS Inverter risetime with respect to stress time under a series of electric field at different temperatures. [26]

3.5 Correlation between device-level and circuit-level NBTI

The relatively low degradation with marginaly stress time dependence (n = 0.06),shown by the Inverter risetime at room temperature and under low AC NBTI stress voltage (-10 V), seems to mimic the trend of hole trapping in preexisting traps, while that shown at higher voltages and at elevated temperature seems to include two parts which could be attributed to both hole trapping in preexisting oxide traps and the generation of interface traps. This can be justified as follows :

At room temperature and under AC NBTI stress with negative voltages lower than 14 V, the insignificant shift, and which remains quasi constant with stress time, means no creation of traps has happened. Therefore, the observed shift could be attributed to hole trapping in preexisting deep level traps. This thought is consolidated by the fact that our circuits are made with dry thermaly grown SiO_2 and an experimental evidence of hole trapping in such oxides at room temperature has been reported in [32]. It was also found that NBTI degradation of PNO based pMOS transistors below 290 K is dominated by hole trapping process and that it is unnecessarily related to nitrogen but the incorporation of nitrogen in the gate dielectric increases the probability of hole trapping in the NBTI process [33].

When the temperature is elevated, the degradation becomes more important and the time exponent starts to increase obove 0.1. This increase of the time exponent is induced by the combination of voltage and temperature. On one hand, if the stress voltage of the AC NBTI is low (Vgstr < 13 V), a higher temperature (T >120 °C) is needed to induce an n > 0.1. On the other hand,when the stress voltage is increased (Vgstr ≥ 13 V), less temperature is needed to observe an n > 0.1. This trend of the time exponent n of the Inverter risetime shift risetime with voltage and temperature looks like that observed on its logic threshold as shown in Figs 3.10 and 3.12. It imitates high temperature acceleration coupled with low field acceleration for E_{ox} < 6.5 MV/cm, and high field acceleration coupled with low temperature acceleration for E_{ox} ≥ 6.5 MV/cm. Hence, we can assume that the degradation of Inverter risetime would be dominated by the interface states creation in the first case and by hole trapping generation in new generated oxide traps in the second case. Figure 3.14 can give more insight on the possible origins of the degradation. Indeed, in the case of an AC NBTI stress under low voltage, the activation energy of the risetime is around 0.06 eV, while it equals 0.25, 0.17, and 0.09 eV under -12, -13, and -14 V, respectively.

This trend consolidates the idea that the NBTI degradation can be due to trapping in pre-existing traps (with E_a ~0.06 eV), or to new generated traps (interface traps with higher activation energy (0.17-0.25 eV) or to both interface states and oxide traps under certain conditions with dominance of this latter as the E_a decreases to 0.11 eV then to 0.04 eV with stress time

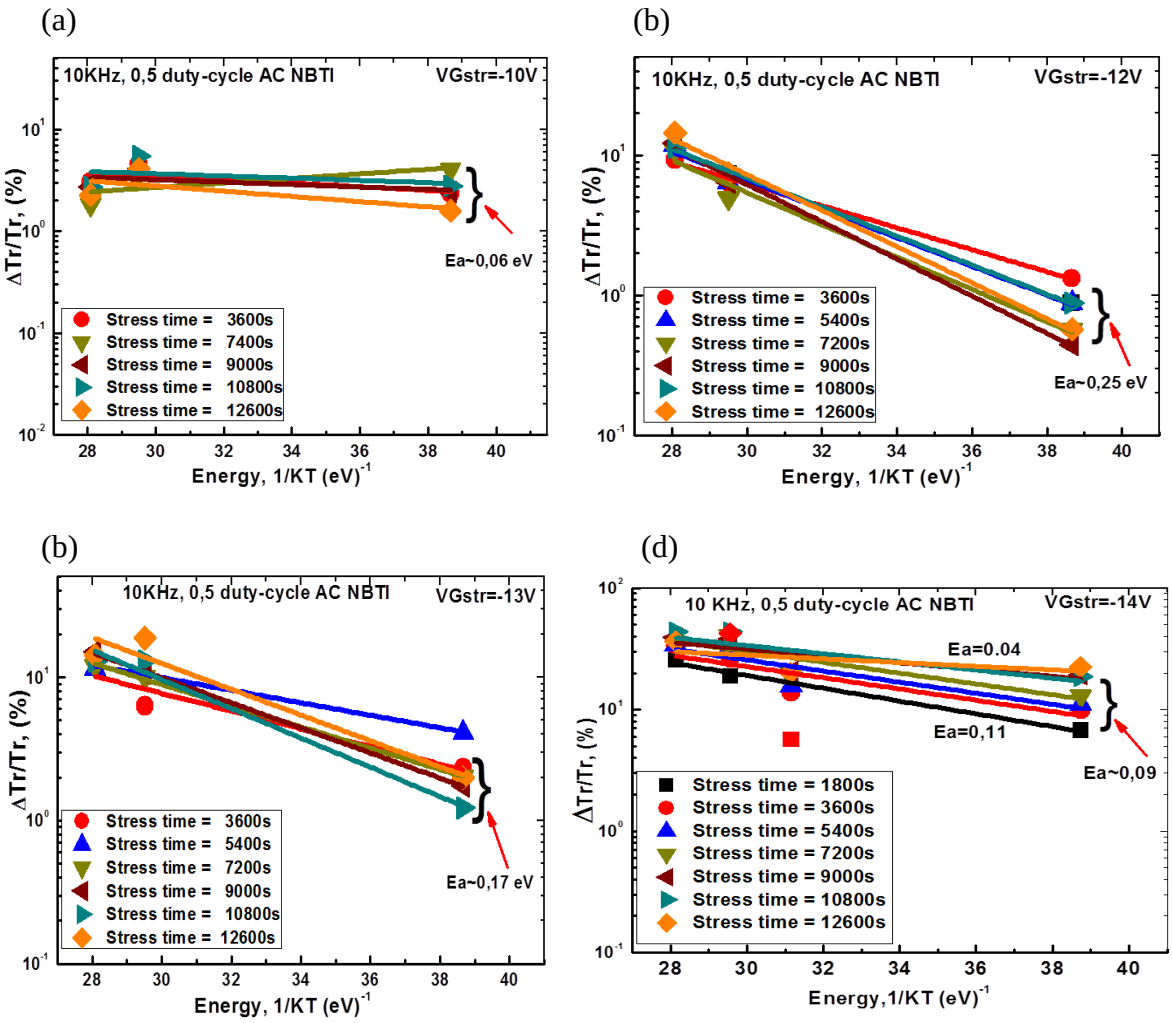


Fig. 3.15 CMOS Inverter risetime shift with respect to the energy under different AC NBTI stress voltages. [26]

increasing. Therefore, the signature of the NBTI-induced Inverter risetime shift when correlated with the pMOS threshold shift also reveals the co-existence of more than one physical mechanism behind the degradation. Moreover, higher the electric field and temperature are, longer the stress time is. It is more likely to observe a power law dependence with time stress of the circuit performance degradation.

Besides, when the AC NBTI stress voltage is set at -14 V, the shift of the risetime at room temperature becomes significant and shows an increase with stress time inceasing with an important time exponent (n = 0.6). The shift continues to increase at higher temeprature but with a decreasing time exponent. In fact, at the beginning of the stress, the shift becomes more important when the temperature is elevated from room temperature to 120 and 140 °C, because more traps are generated. Then, it tends to increase slowly with both temperature and stress time increasing. This could be due, on one hand, to the temperature accelerated recovery [34] during the half period of the AC stress when temperature is elevated from 120 to 140 °C, and on the other hand, to decrease of availible defect precursors to break when stress time goes longer. Such trend of time exponent decrease with higher temperature was interpreted by Pobegen *et al*. [7] by the time-temperature concept merging from the logarithmic nature of NBTI recovery.

As point out, the NBTI degradation under AC NBTI stress is much less important than that under DC NBTI stress. Indeed, the configuration of the circuit and its operation mode (switching) largely influence the way the circuit is degraded. The results of the characterization of the Inverter DC response and temporal performance presented here have confirmed many aspects and revealed others.

- **The confirmed aspects are*:***

The circuit-level NBTI degradation is voltage, temperature and stress time dependent. In fact, NBTI circuit-level degradation follows that of device-level in terms of voltage and temperature dependence.

- **The revealed aspects are**:

An experimental evidence of the signature of more than one physical mechanism is behind the degradation. Moreover, the dominance of one mechanism over the other depends on voltage, temperature, and stress time conditions.

The time exponent found here is not a fixed value but it rather depends on both voltage and temperature unlike what is reported elsewhere about the fixed value of the time exponent to 0.16 for digital circuit NBTI degradation [31]. In our experiments, it can be observed that Time exponent increases with stress voltage increasing at low temperature or with temperature increasing under low stress voltage. This finding of voltage and temperature dependence of the Inverter switching time agrees other works conducted at circuit-level [21], where the authors found that time exponent depends on the circuit workload. In fact, *Lu et al.*[21] found system 1`s

frequency degradation to show a time exponent higher than that shown by system 2 within the same chip due to the fact that system1 is operated at a voltage which is about 80 mV higher than that of system2. In addition, the average temperature of sytem1 is about 13 degrees higher than system 2 due to the difference in the cooling. On the other hand, the time exponent decreases at higher temperature under high voltages leading to saturation with stress time [21].

In the light of these findings, the degradation of digital circuits can be minimized if one could maintain time exponent under a certain threshold value. This can be possible by keeping the stress voltage and temperature within a certain limit. For the Inverter, it can noticed that if the oxide field does not reach 6.5 MV/cm and the temperature does not exceed 120 °C, the power law time exponent n would not exceed 0.1. Hence, in the latter case, it would require about 5 times longer, to reach the same degradation, than the case where the time constant is assumed to be 0.16 as in [31] for the power law exponent of digital circuits degradation due to NBTI. Thereby, the circuit lifetime can be enhanced in the first case by lowering V_{DD}. On the other side, if lowering the supply voltage is not possible, then managing temperature profile is the only way to reduce the induced degradation. Either by cooling down the circuit to reduce the induced degradation during its on state (stress phase) or elevating its temperature during its off state (recovery phase) to accelerate its recovery and thereby compensate the induced degradation during stress phase.

Finally, if the circuit is kept operating under certain conditions, the NBTI-induced degradation of its temporal performance could be minimized.

3.6 Conclusion

In this chapter, we have presented an experimental analysis of NBTI degradation at circuit-level. The CMOS Inverter has been characterized under both DC and AC NBTI stresses. The impact of NBTI has been examined on both the DC features and temporal performance of the Inverter. The obtained results revealed that the degradation of the circuit DC response and transient under AC NBTI stress is much less than that undergone under DC NBTI stress. When examined against device-level NBTI degradation, circuit-level NBTI degradation has also shown both voltage and temperature dependences with more than one acceleration factor for each. However, it does not unconditionally present stress time dependence. Actually, the time exponent is found to be governed by both voltage and temperature. Moreover, the induced degradation tends to saturate with both temperature and stress time.

Furthermore, circuit-level NBTI characterization has revealed a signature of the co-existence of more than one physical mechanism that is behind the degradation where the dominance of one mechanism over another depends on voltage, temperature and the stress time.

By analyzing NBTI degradation at circuit-level, designers are more aware about its impact, better guided to develop techniques for alleviating NBTI and assure their circuits' reliability lifetime that will be presented in chapter 4.

References

[1] M. Jo, M. Chang, S. Kim, H. Jung, R. Choi, and H. Hwang, "Contribution of Interface States and Oxide Traps to the Negative Bias Temperature Instability of High-k pMOSFETs," *IEEE Electron Device Letters,* vol. 30, pp. 291-293, 2009.

[2] S. Mahapatra, V. D. Maheta, A. E. Islam, and M. A. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs," *IEEE Transactions on Electron Devices,* vol. 56, pp. 236-242, 2009.

[3] Z. Q. Teo, D. S. Ang, and C. M. Ng, "Separation of Hole Trapping and Interface-State Generation by Ultrafast Measurement on Dynamic Negative-Bias Temperature Instability," *IEEE Electron Device Letters,* vol. 31, pp. 656-658, 2010.

[4] B. Djezzar, H. Tahi, A. Benabdelmoumene, A. Chenouf, and Y. Kribes, "A New Method for Negative Bias Temperature Instability Assessment in P-Channel Metal Oxide Semiconductor Transistors," *Japanese Journal of Applied Physics,* vol. 51, p. 116602, 2012/11/05 2012.

[5] 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 Transactions on Device and Materials Reliability,* vol. 11, pp. 19-34, 2011.

[6] S. Mahapatra, N. Goel, S. Desai, S. Gupta, B. Jose, S. Mukhopadhyay, K. Joshi, A. Jain, A. E. Islam, and M. A. Alam, "A Comparative Study of Different Physics-Based NBTI Models," *IEEE Transactions on Electron Devices,* vol. 60, pp. 901-916, 2013.

[7] G. Pobegen and T. Grasser, "On the Distribution of NBTI Time Constants on a Long, Temperature-Accelerated Time Scale," *IEEE Transactions on Electron Devices,* vol. 60, pp. 2148-2155, 2013.

[8] T. Grasser, K. Rott, H. Reisinger, M. Waltl, F. Schanovsky, and B. Kaczer, "NBTI in Nanoscale MOSFETs—The Ultimate Modeling Benchmark," *IEEE Transactions on Electron Devices,* vol. 61, pp. 3586-3593, 2014.

[9] S. Mahapatra and N. Parihar, "A review of NBTI mechanisms and models," *Microelectronics Reliability,* vol. 81, pp. 127-135, 2018.

[10] J. H. Stathis, S. Mahapatra, and T. Grasser, "Controversial issues in negative bias temperature instability," *Microelectronics Reliability,* vol. 81, pp. 244-251, 2018/02/01/ 2018.

[11] N. K. Jha, P. S. Reddy, D. K. Sharma, and V. R. Rao, "NBTI degradation and its impact for analog circuit reliability," *IEEE Transactions on Electron Devices,* vol. 52, pp. 2609-2615, 2005.

[12] K. Kang, S. Gangwal, S. P. Park, and R. Kaushik, "NBTI induced performance degradation in logic and memory circuits: how effectively can we approach a reliability solution?," in *2008 Asia and South Pacific Design Automation Conference*, 2008, pp. 726-731.

[13] S. Khan, S. Hamdioui, H. Kukner, P. Raghavan, and F. Catthoor, "BTI impact on logical gates in nano-scale CMOS technology," in *2012 IEEE 15th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*, 2012, pp. 348-353.

[14] B. Eghbalkhah, M. Kamal, H. Afzali-Kusha, A. Afzali-Kusha, M. B. Ghaznavi-Ghoushchi, and M. Pedram, "Workload and temperature dependent evaluation of BTI-induced lifetime degradation in digital circuits," *Microelectronics Reliability,* vol. 55, pp. 1152-1162, 2015/07/01/ 2015.

[15] E. Mintarno, V. Chandra, D. Pietromonaco, R. Aitken, and R. W. Dutton, "Workload dependent NBTI and PBTI analysis for a sub-45nm commercial microprocessor," in *2013 IEEE International Reliability Physics Symposium (IRPS)*, 2013, pp. 3A.1.1-3A.1.6.

[16] T. Nigam, "Pulse-Stress Dependence of NBTI Degradation and Its Impact on Circuits," *IEEE Transactions on Device and Materials Reliability,* vol. 8, pp. 72-78, 2008.

[17] Y. S. Tsai, N. K. Jha, Y. Lee, R. Ranjan, W. Wang, J. R. Shih, M. J. Chen, J. H. Lee, and K. Wu, "Prediction of NBTI degradation for circuit under AC operation," in *2010 IEEE International Reliability Physics Symposium*, 2010, pp. 665-669.

[18] R. Fernandez, B. Kaczer, J. Gago, R. Rodriguez, and M. Nafria, "Experimental characterization of NBTI effect on pMOSFET and CMOS Inverter," in *2009 Spanish Conference on Electron Devices*, 2009, pp. 231-233.

[19] R. Fernandez, B. Kaczer, A. Nackaerts, S. Demuynck, R. Rodriguez, M. Nafria, and G. Groeseneken, "AC NBTI studied in the 1 Hz -- 2 GHz range on dedicated on-chip CMOS circuits," in *2006 International Electron Devices Meeting*, 2006, pp. 1-4.

[20] N. Berbel, R. Fernandez, and I. Gil, "Modelling and experimental verification of the impact of negative bias temperature instability on CMOS Inverter," *Microelectronics Reliability,* vol. 49, pp. 1048-1051, 2009/09/01/ 2009.

[21] P.-F. Lu, K. A. Jenkins, T. Webel, O. Marquardt, and B. Schubert, "Long-term NBTI degradation under real-use conditions in IBM microprocessors," *Microelectronics Reliability,* vol. 54, pp. 2371-2377, 2014/11/01/ 2014.

[22] A. Chenouf, B. Djezzar, A. Benabedelmoumene, and H. Tahi, "Does PMOS Vth shift wholly capture the degradation of CMOS Inverter circuit under DC NBTI?," in *2012 IEEE International Integrated Reliability Workshop Final Report*, 2012, pp. 191-194.

[23] G. Pobegen, T. Aichinger, M. Nelhiebel, and T. Grasser, "Dependence of the negative bias temperature instability on the gate oxide thickness," in *2010 IEEE International Reliability Physics Symposium*, 2010, pp. 1073-1077.

[24] A. Chenouf, B. Djezzar, A. Benadelmoumene, H. Tahi, and M. Goudjil, "Reliability analysis of CMOS Inverter subjected to AC & DC NBTI stresses," in *2014 9th International Design and Test Symposium (IDT)*, 2014, pp. 142-146.

[25] A. Chenouf, B. Djezzar, A. Benadelmoumene, and H. Tahi, "Deep experimental investigation of NBTI impact on CMOS Inverter reliability," in *2012 24th International Conference on Microelectronics (ICM)*, 2012, pp. 1-4.

[26] A. Chenouf, B. Djezzar, A. Benabdelmoumene, and H. Tahi, "On the Circuit-Level Reliability Degradation Due to AC NBTI Stress," *IEEE Transactions on Device and Materials Reliability,* vol. 16, pp. 290-297, 2016.

[27] G. Pobegen, T. Aichinger, M. Nelhiebel, and T. Grasser, "Understanding temperature acceleration for NBTI," in *2011 International Electron Devices Meeting*, 2011, pp. 27.3.1-27.3.4.

[28] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability,* vol. 46, pp. 1-23, 2006/01/01/ 2006.

[29] M. Denais, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, N. Revil, and A. Bravaix, "Interface trap generation and hole trapping under NBTI and PBTI in advanced CMOS technology with a 2-nm gate oxide," *IEEE Transactions on Device and Materials Reliability,* vol. 4, pp. 715-722, 2004.

[30] "The CMOS Inverter: Analysis and Design," in *CMOS Logic Circuit Design*, J. P. Uyemura, Ed., ed Boston, MA: Springer US, 2001, pp. 103-154.

[31] S. Mahapatra, V. Huard, A. Kerber, V. Reddy, S. Kalpat, and A. Haggag, "Universality of NBTI - From devices to circuits and products," in *2014 IEEE International Reliability Physics Symposium*, 2014, pp. 3B.1.1-3B.1.8.

[32] R. F. DeKeersmaecker and D. J. DiMaria, "Hole trapping in the bulk of SiO2 layers at room temperature," *Journal of Applied Physics,* vol. 51, pp. 532-539, 1980.

[33] X.-L. Ji, Y.-M. Liao, F. Yan, Y. Shi, G. Zhang, and Q. Guo, "Direct Experimental Evidence of Hole Trapping in Negative Bias Temperature Instability," *Chinese Physics Letters,* vol. 28, p. 107302, 2011/10 2011.

[34] T. Aichinger, M. Nelhiebel, and T. Grasser, "Unambiguous identification of the NBTI recovery mechanism using ultra-fast temperature changes," in *2009 IEEE International Reliability Physics Symposium*, 2009, pp. 2-7.

CHAPTER 4 AGING-AWARE CIRCUIT DESIGN

4.1 INTRODUCTION

As semiconductor manufacturing moves to more advanced technology nodes, process variability and temporal reliability issues can no longer be ignored during the early stages of the design cycle. Among the temporal reliability issues, NBTI continues to present an important concern of IC design in deep submicron technologies [1-5]. However, a technology based solution is not always possible, because the focus of device engineers is mainly on developing smaller and faster transistors with lower power consumption. This compels the designers to deal with reliability issues during the design phase. Hence, they are in need of means to assess NBTI degradation on their circuits' performance. Likewise, they need to introduce techniques to moderate the predicted degradation in order to improve the circuit's lifetime. As a result, aging simulation becomes indispensable to predict the performance degradation of the ICs due to temporal variations. The introduction of new design techniques which consider reliability as a design constraint as important as speed, area, and power consumption, becomes also necessary in order to warranty delivering reliable circuits built on unreliable devices. Such a design approach is known as design for reliability (DFR) [6-8].

In the view of this work, DFR is twofold objective. One concern circuit-level analysis and assessment of NBTI degradation, while the other consists in combating/alleviating the induced degradation.

The remaining of this chapter is organized as follows: in sections 2 and 3, the state of the art of reliability simulation and aging mitigation techniques are reviewed respectively. The proposed NBTI-aware design techniques for 6T-SRAM is presented in section 4, and further discuss in section 5.

4. 2 AGING RELIABILITY SIMULATION

Guaranteeing circuit reliability in the presence of transistor aging has been a problem since the early 1980s [9, 10]. So, circuit reliability simulators have been proposed in literature and integrated into the industrial design flows of the main EDA tools companies such as Cadence [11], Synopsys [12], and Mentor Graphics [13].

Current reliability simulations are based on physics of failure models, empirical data and statistical models to investigate new manifestations of existing failure mechanisms like EM, TDDB, HCI, and NBTI on circuit performance and stability [10].

Reliability analysis can be conducted at different abstraction levels in the IC design flow (see Fig. 4.1) [14]. Depending on the targeted abstraction level, an appropriate aging model ought to be provided to monitor the device parameter's shift on the circuit performance.

The overall modeling for reliability simulation at circuit-level is shown on Fig.4.2 [15]. It is twofold process. First, an aging model is generated from accelerated stresses; then the extracted data is used to generate the degraded SPICE model that defines the electrical behavior of the device over aging. This latter together with the lifetime definition criteria are used by the simulator to extrapolate the aged device parameters to the simulation time and generate the corresponding waveforms of the circuit response with the aged devices. At the end of the simulation, results of the circuit performance with both fresh and aged devices are used by the designer to assess his (her) circuit performance over time and redesign it if necessary to meet its target specifications during its useful lifetime.

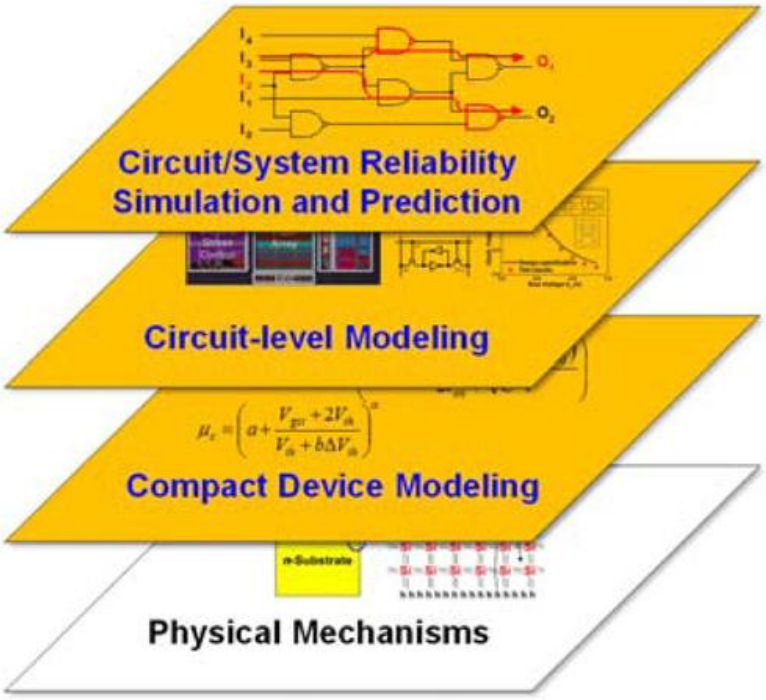


Fig. 4.1: Cross-layer approach of circuit reliability analysis [14].

Circuit-level reliability simulators have been inspired by BERT "Berkeley Reliability Tools" modeling and simulation concepts [16, 17]. BERT basically simulates changes in the circuit performance due to hot carriers' effect, and calculates the probability of the circuit to fail due to

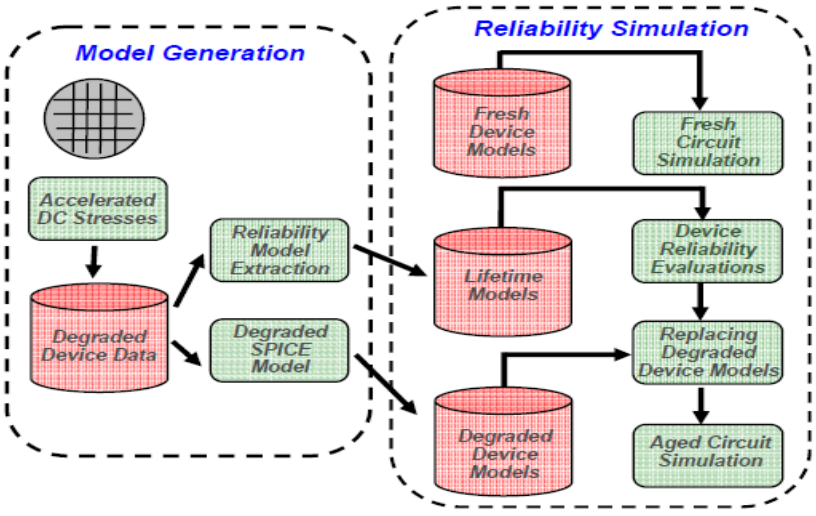


Fig. 4.2: Typical reliability modeling and simulation framework [7].

its oxide breakdown, or electro-migration. BERT is designed in a modular way around 4 tools:

- CAS (Circuit Aging Simulation) to simulate hot carriers injection (HCI) in MOSFETs.
- BiCAS (Bipolar Circuit Aging Simulation) to simulate aging in bipolar transistors.
- CORS (Circuit Oxide Reliability Simulation) to simulate time dependent dielectric breakdown (TDDB).
- EM (Electo-migration) to simulate the electro-migration.

The overall simulation flow of BERT is depicted in Fig 4.3. We briefly review how CAS module works. BERT CAS module requires the circuit simulator to provide it with waveforms of terminal voltages of all MOS transistors in the circuit. The post-processor uses these waveforms to compute the degradation rate of each transistor based on bias conditions and time through a parameter called AGE. The amount of degradation induced to transistor model parameters, such as threshold voltage, mobility, transconductance, is then related to this AGE parameter. The post-processor calculates the AGE parameter for each transistor and stores the result in a table. Then, it uses this Age table to generate new transistor model parameters for each stressed device. This is done by interpolation between the process files provided by the user (see Fig.4.4). These files are based on stress measures on real transistors and contain the drifts of the device's parameters based on the values of the corresponding AGE parameters. Finally, the user can simulate the so-modified netlist in order to analyze the impact of the degradation for example HCI on the circuit to be tested.

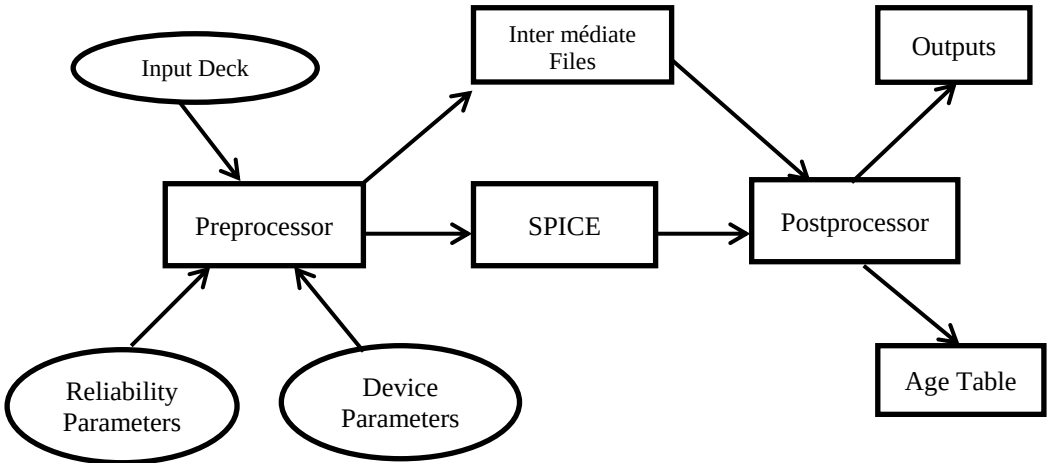


Fig. 4.3 : BERT structure [7]

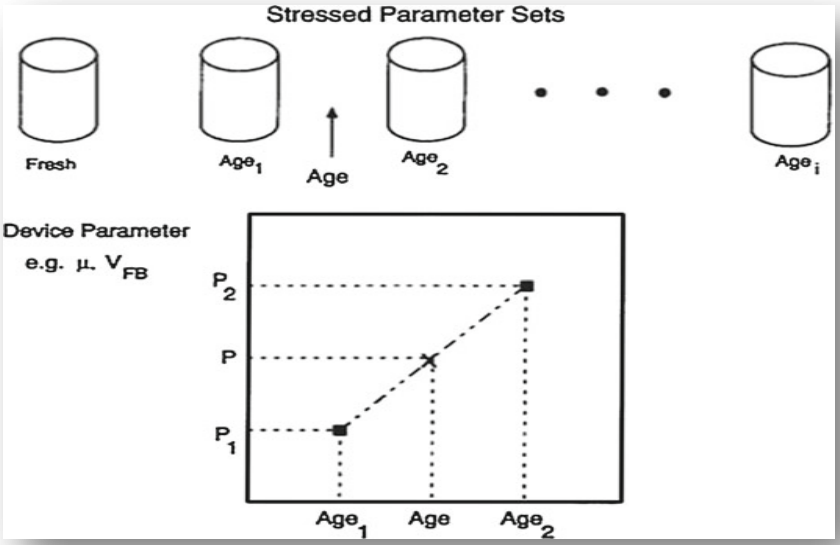


Fig. 4.4: Calculation of the aged parameter from pre-stressed model parameter sets. [7].

BERT CAS module was later integrated into Cadence reliability simulation framework invoked through "Virtuoso Ultrasim" or in the "ADE" Analog Design Environment.

To perform aging simulation, Cadence supports two methods for modeling aging effects for a specific technology: Aged Model and AgeMOS.

- **Aged Model** is a table of transistor models where the parameters of the SPICE model of the aged transistor are extracted from a fresh device at a number of stress intervals. These model parameters form a set of aged model files as seen on fig 4.4. During reliability simulation, the aging for each transistor in the circuit is calculated on the basis of interpolation or regression of values in these files. This approach is inspired by BERT.
- **AgeMOS** is an analytical aging model describing each aging effect. IC manufacturers can provide a universal model to all of their IC design customers without SPICE model compatibility issues. AgeMOS is more accurate than Aged model as the device parameters at any age value can be calculated, so no interpolation or regression is needed. Simulation with the AgeMOS model is easier to perform and faster for the simulator.

Therefore, we adopt this methodology to conduct circuit-level aging simulation using our own NBTI model. The interaction of the developed model with the simulator is ensured through Virtuoso unified reliability interface (URI) shown in Fig. 4.5. The implemented NBTI model is described hereafter

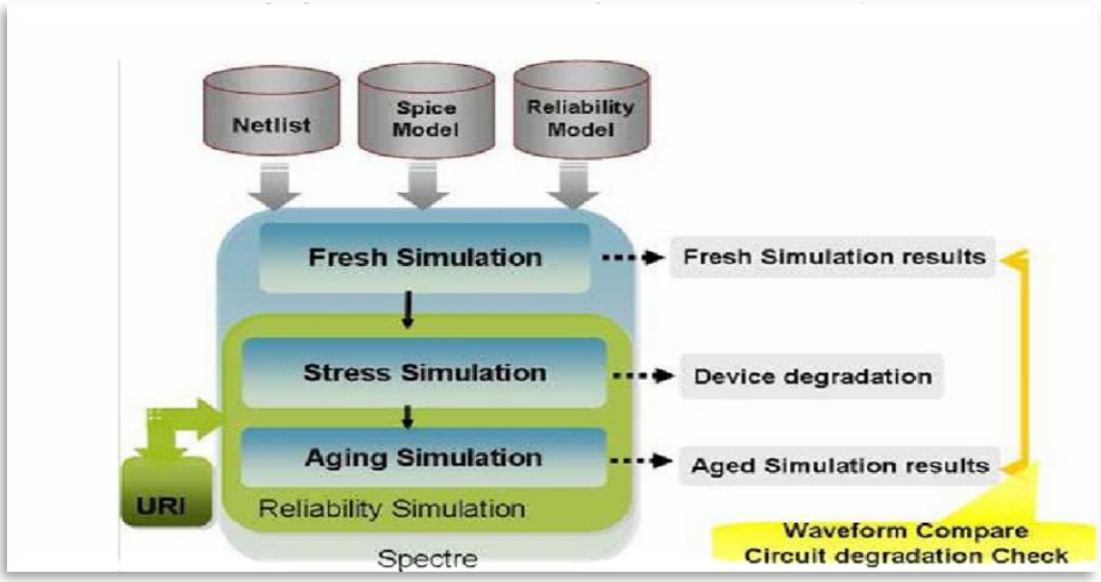


Fig. 4.5: Cadence Ultrasim reliability URI

4.2.1 Our developed NBTI simulation interface on Cadence

Despite leading EDA companies provide with reliability simulators aging models, it is not possible to use these built-in models to design reliable circuits. In fact these tools and model are usually proprietary and customized to a specific technology. Nevertheless, they provide IC designers with a means, usually an interface, to incorporate their own reliability models for aging simulation. As such, we have added our own NBTI model to Cadence Relxpert simulator through its unified reliability interface (URI). The NBTI model is implemented as a shared library (.so). This library advises the simulator, on run time, the way to compute the device model parameters for the considered age, how to evaluate the degradation, and how to calculate the lifetime associated with a given amount of the NBTI induced degradation. For this aim, Cadence provides users through its Virtuoso Unified Reliability Interface Reference manual with guidelines to implement their models and building the shared library [18].

In order to integrate the reliability model into the Relxpert simulator, it is essential to know the data (data structures) that are exchanged between the simulator and the interface, to understand the sequence of calls from the simulator to these structures, and to define the equations implementing the reliability model in order to be able to update the values of the SPICE parameter chosen to monitor the degradation. In view of that, the simulator uses the routine of the URI, as illustrated in the chart depicted in Fig. 4.6 [18].

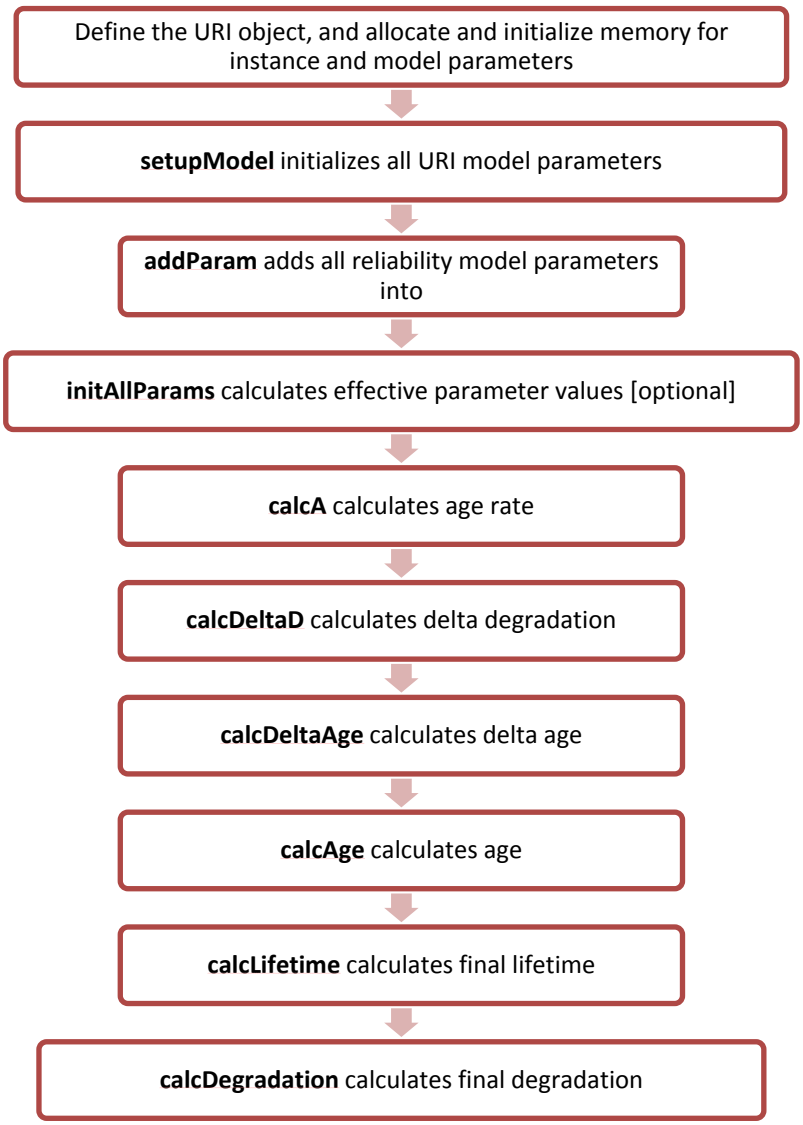


Fig. 4.6: Unified Reliability Interface Call Sequence [18].

Actually, each degradation model uses a quantity called **Age** to represent the amount of damage caused by its mechanism. Using Age, the degradation of any physical characteristic of the device can be derived. The model must provide the derivative of Age or **age rate** (A), which is used by the simulator to calculate the device age.

Age can be a function of device geometry width (w), length (l), device bias (V_{gs}, V_{ds}, or V_{bs}), SPICE models, or SPICE model quantities (such as I_{ds} or V_{th}) as:

$$A = f\ (V_{gs}, V_{DS}, V_{bs}, I_{DS}, V_{th}, w, l, \dots\dots) \tag{4.1}$$

Age is then calculated by integrating A on the total interval of the transient simulation t composed of N entry signal period. Consequently, Age is given by:

$$age(t) = N \int_0^T A\ dt \tag{4.2}$$

The Age calculation is provided with the reliability interface through the **calcA** function.

The degradation monitor is chosen by the user, in terms of transconductance ($\Delta g_m/g_m$), drain current ($\Delta I_{ds}/I_{ds}$), threshold voltage shift (ΔV_{th}), and can be predicted at any time t using the user-defined D degradation equation based on AGE as follows:

$$D = f(A, t_{stress}, t) \ \text{ or } \ D(t) = f(AGE) \tag{4.3}$$

where **Age** corresponds to the amount of degradation at the t sterss moment and t is the time at which degradation should be predicted. The D calculation is provided with the reliability interface through the **calcDegradation** function.

The calculation of the lifetime is also defined by the user, and it is solved by the time when the degradation reaches a certain value for example 10% of the I_{DS} degradation, or 30-50 mV of V_{th}. It is defined as:

$$t_{life} = f(D_{life}, A, t_{stress}) \tag{4.4}$$

where t_{life} is the lifetime, D_{life} is the amount of degradation that defines the lifetime and A is the age after a stress time given by t_{stress}. The lifetime is provided with the reliability interface through **calcLifetime** function.

According to the call sequence of the interface by the simulator shown in Fig. 4.6 the following three intermediate equations are to be implemented: **calcDeltaD**, **calcDeltaAge**, and **calcAge**.

- **calcDeltaD:** calculates and returns degradation up to the simulation time in the transient time period. This function is called at every simulation time point.
- **calcDeltaAge:** calculates and returns the age value up to the simulation time in the transient time period. This function is called at every simulation time point.
- **calcAge:** calculates the total age at the specified future time. This function is called once to calculate the final device age value.

The user reliability models are passed to Relxpert (Ultrasim) through a shared library. The creation of this latter follows the chart illustrated in Fig. 4.7. Once the library built and compiled, as shown in Fig. 4.9, it can be and called by the simulator at the runtime.

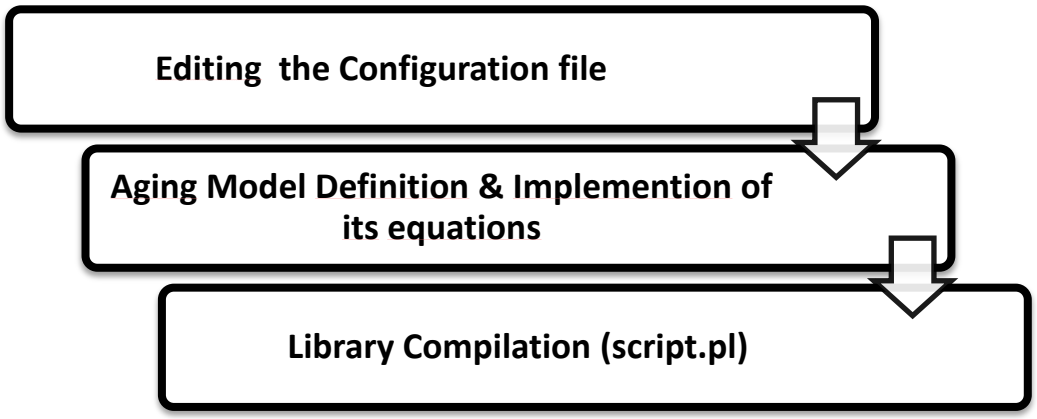


Fig. 4.7: Shared library creation steps

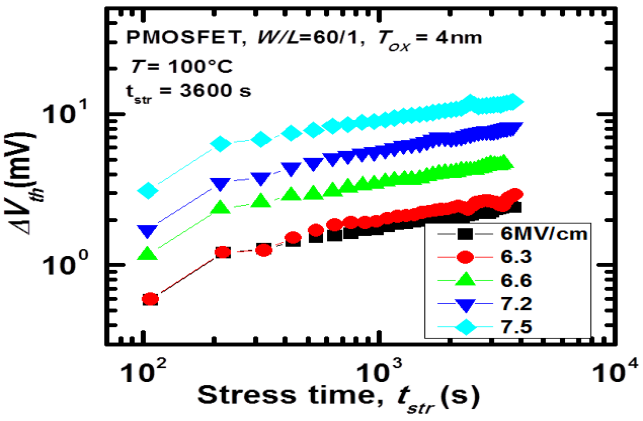
4.2.2 Our NBTI degradation model

The implemented NBTI model for aging simulation is a power law model At^n expressed as:

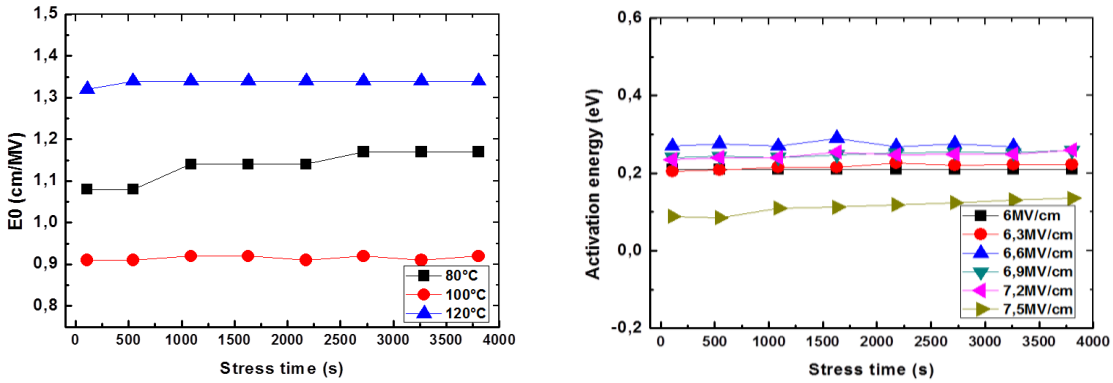
$$\Delta v_{th} = A_0 . exp\left(\frac{E_{ox}}{E_0}\right) exp\left(-\frac{E_a}{KT}\right)\ t^n \tag{4.5}$$

where the fit parameters E0, Ea and n can be extracted from the OTFOT-based NBTI characterization [19] of 0.18µm pMOS transistors as shown in Fig. 4.8.

(a)



(b) (c)



(d)

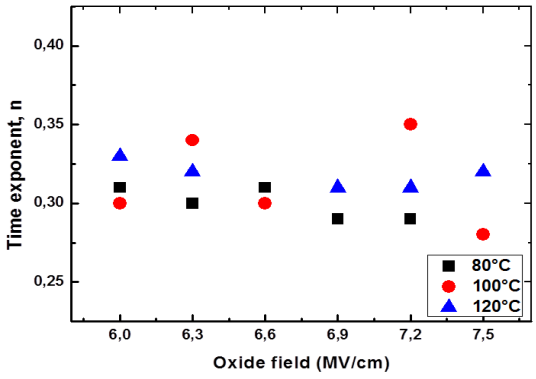


Fig. 4.8: OTFOT-based NBTI characterization of a 0.18µm pMOS V_{th} in (a), and the extracted NBTI Model parameters E_0 in (b), E_a in (c), and the n exponent in (d)

These pMOS devices belong to a test structures chip (see Fig. 4.9) are designed at CDTA with TSMC 0.18 µm CMOS technology, and fabricated via multi project wafer (MPW) of Europractice IC service [20]. Two nMOS and pMOS devices in terms of output and the transfer characteristics are illustrated in Fig 4.10.

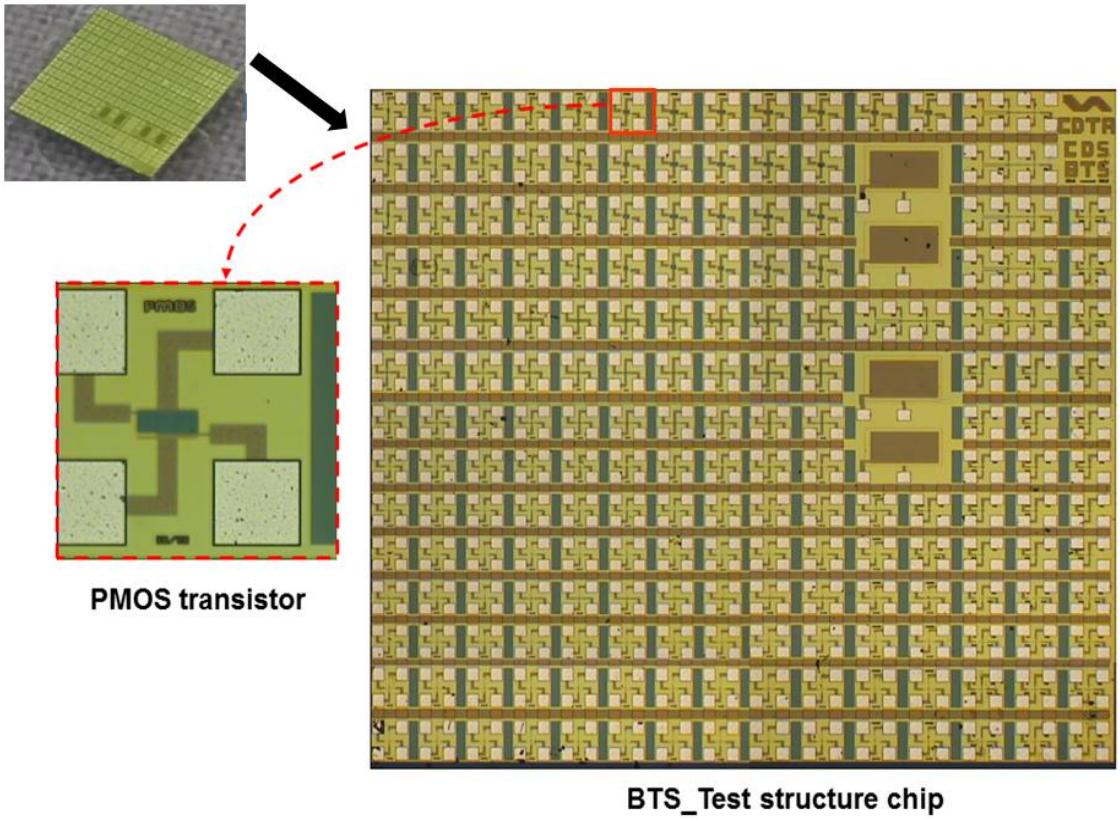


Fig. 4.9: BTS_Test structures chip implemented in TSMC 0.18µm CMOS technology.

4.2.3 NBTI aging simulation examples

Once we have written our NBTI model equations in C-code, compiled and built the shared library and checked it with scripts provided by Cadence (see appendix 1 for implementation details. and further simulation results of the Inverter), it was possible for us to conduct aging simulation at circuit-level. Some examples of aging simulation, using our NBTI model, of the CMOS Inverter, ring oscillator (RO) and the 6T-SRAM are given hereafter.

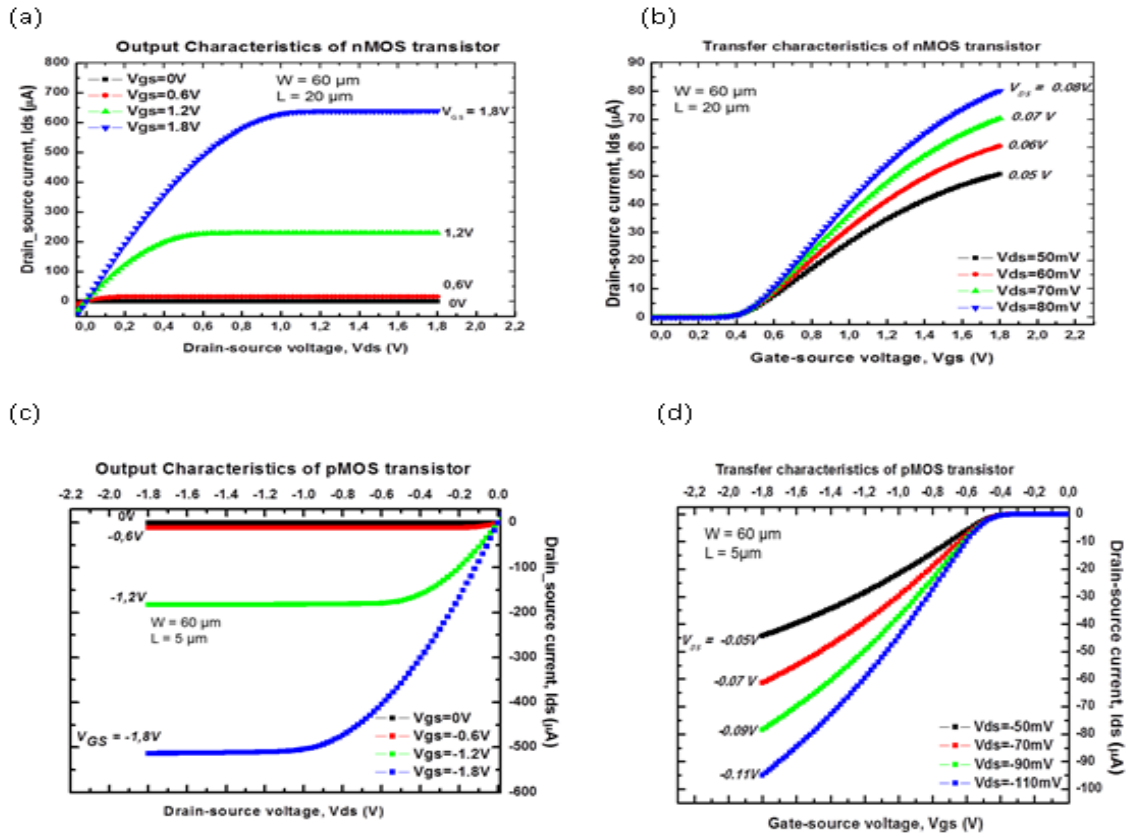


Fig. 4.10: Characterisation results of two nMOS and pMOS devices belonging to BTS_test structures.

The simulation results of the DC and the transient responses of the CMOS Inverter under NBTI are shown in Fig. 4.11. As expected, the inverter VTC has shifted to the left-side with age increasing (see Fig. 4.11(a)) and the shift is exacerbated at high temperature (see Fig. 4.11 (b)). This has led to an increase of the CMOS inverter risetime as shown on Fig. 4.11 (c) and (d). Consequently, the circuit performance has been affected. This can be better assessed when examining the inverter in the context of the ring oscillator (RO) as shown in Fig 4.12. In fact, the RO frequency has been degraded over time as the inverter delay has been increased due to NBTI aging. Moreover, by simulating the aging of the pMOS due to NBTI in the context of the 6T-SRAM cell, as shown in Fig. 4.13, it reveals that data integrity may be negatively affected.

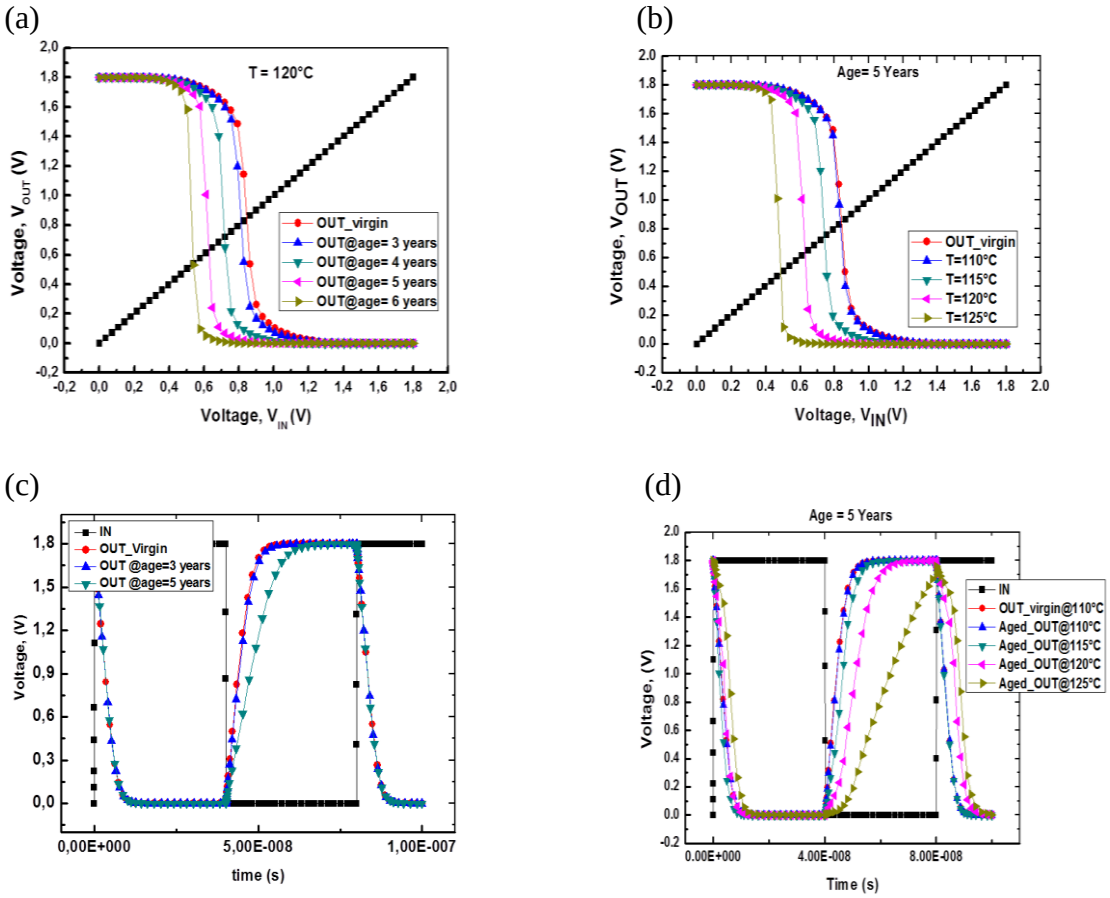


Fig. 4.11: Simulated CMOS Inverter DC and transient response subjected to NBTI (a) for different ages (b) at different temperatures

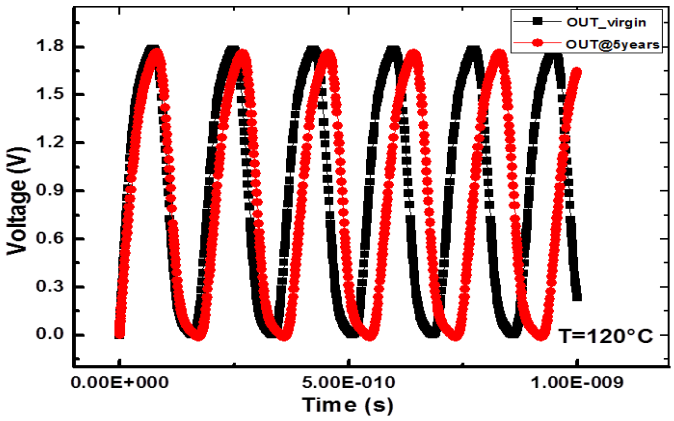


Fig. 4.12: Aging simulation of 15-Stages Ring Oscillator.

Actually, due to NBTI, it is not possible for the aged cell to have a good "1" upon a write operation compared to the fresh cell (see Fig. 4.13 (a)). Upon an operation read is invoked, the bit storing "0" is of the 6T-SRAM get it voltage increased by an amount of ΔV_0. This latter does not affect the cell stability as long as it does not reach the trip of other inverter of the cell. Unfortunately, this voltage is increasing in the aged cell, while it is constant in the fresh one, (see Fig. 4.13 (b)). When, it reaches the trip voltage (logic threshold) of the inverter, this latter change and the data at its output got flipped (see Fig. 4.13(c). Hence, the whole cell bit is inverted while being read and the cell become instable during read operation resulting in read failure. Further aging simulation results of the three circuits are demonstrated.

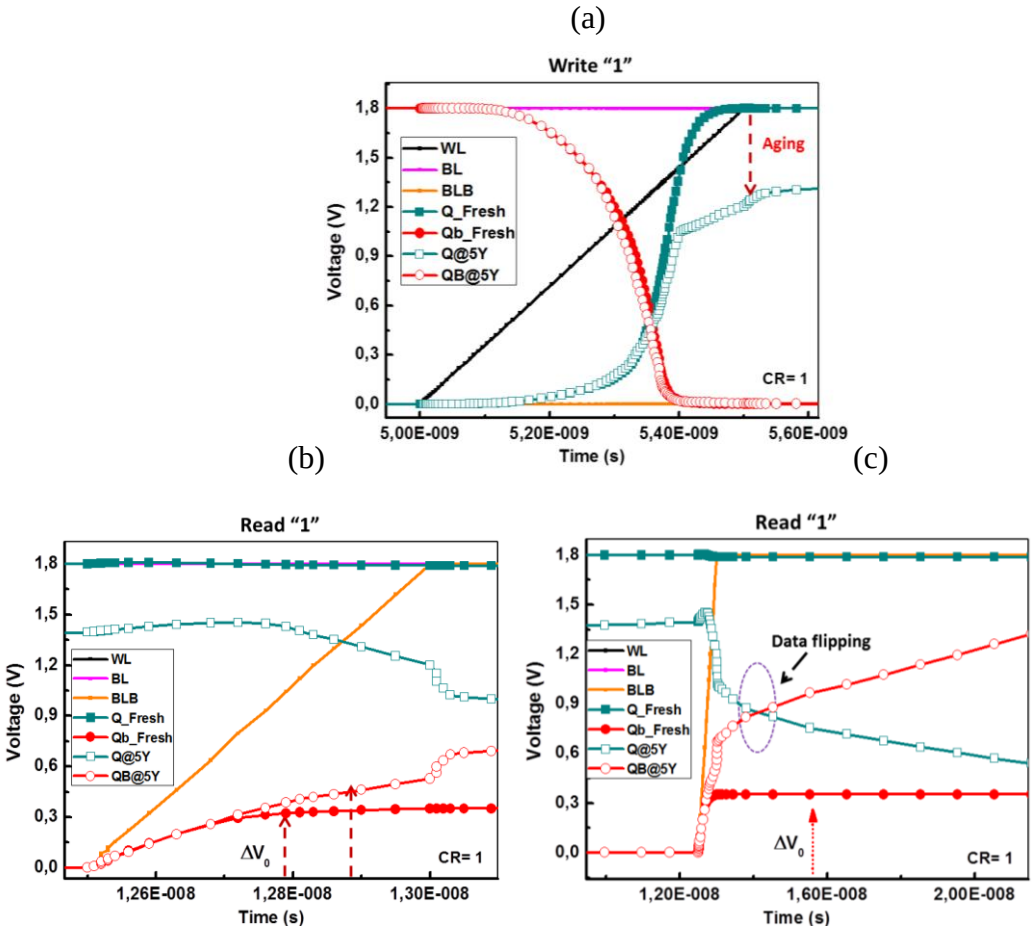


Fig. 4.13: 6T-SRAM aging simulation results

"Bad" 1 in the aged cell in (a) Read operation perturbed in the aged cell in (b) and data flipping and a read failure in the aged cell in (c)

4. 3 NBTI AGING MITIGATION TECHNIQUES

Aging-aware design techniques have been proposed in the literature [6, 7]. These techniques can be introduced in the design flow at different abstraction levels ranging from circuit[8] [21] to system level [22, 23], or through a cross-layer approach [24] [25]. In brief, these techniques aim to enhance the circuit performance at design time, and/or adjust parameters that directly affect the amount of NBTI-induced $\Delta V_{th,}$ and boost its recovery during the circuit runtime. Most of these techniques have been adapted for SRAM design to compensate/mitigate NBTI impact since both the cell memory stability and the array performance can be affected [26].

NBTI mitigation techniques are based on dynamic supply voltage scaling [27], power gating techniques [28-30], pulse width adjusting [31], world line voltage controlling [32], adaptive body biasing [33-36], duty cycle balancing [37, 38], recovery boosting [39], and micro architectural redundancy [40]. Among these techniques, transistor sizing remains the less used technique for NBTI mitigation in SRAMs [41] [42]. In the following, we review some of the NBTI mitigation techniques, and how they were adapted to SRAMs.

4. 3.1 Guard-Banding

For aging, guard-banding, a conventional design technique to deal with uncertainties consists in reducing the clock frequency and/or increasing the supply voltage to account for degradation over the lifetime of a design such that there are no timing violations due to aging during the circuit lifetime. Usually, 10-20% of the cycle time is reserved as a guard-band for the logic structures [43]. Similarly, V_{min}, defined as the minimum voltage at which the SRAM array is functional [44], is increased by 10% as a guard-band to handle 10% increase in threshold voltage (V_{th}) for the storage structures [43]. However, increasing V_{min} over time to compensate for the degraded SRAM stability due to BTI aging is highly undesirable in low power systems in which energy efficiency is often the utmost design consideration particularly in an ultra-dynamic voltage scaling (UDVS) designs [31] [32]. Yet, in such designs, the sub-threshold operation would provide the minimum energy solution with relaxed constraint on performance. However, in the sub-threshold region, the SNM is a limiting factor which is significantly reduced by NBTI [31]. To improve V_{min} in such designs, Kim *et al.* [31] proposed an NBTI/PBTI-aware wordline control technique trough a careful control of the wordline pulse width for 6T-SRAM.

Lee et *al.* [32] extended the approach with wordline pulse voltage controlling to improve the stability of half-selected cell (8T-SRAM) against BTI aging. Despite read and write operations are decoupled in 8T-SRAMs, they might experience unwanted 6T-readlike disturbance sine their bit lines pairs are pre-charged to V_{DD}. In fact, Lee et al. [32] found that the SNM and the time to data flip becomes faster with BTI stress. By lowering write wordline (WWL) voltage, the characteristics of the access transistors are weakened. Consequently, the disturbing currents flowing from the bitlines were reduced and the SNM was consequently improved [32]. Furthermore, the time to data flipping is enlarged using WWL lowering. However, the BTI-improved write margin would be scarified with lower WWL. Using 10000-point Monte Carlo simulation results, they showed that, by WWL lowering, the failure error probability is reduced from 50 mV forward and achieves as high as 96% error probability reduction at 100-mV degradation [32].

To evaluate architecture-level optimization techniques impact on NBTI mitigation, Chan *et al.* [23] developed a numerical simulation engine for NBTI-induced aging, based on the reaction-diffusion model. They performed a study on a commercial 65 nm technology and adapted their model to support voltage scaling, power gating, and activity management which are employed by architectural techniques. To estimate the system-level performance degradation, they measured the maximum delay among the top ten critical paths of OpenSPARC T1 processor and assumed all pMOS transistors in a circuit degrade by the same amount. They found that dynamic voltage scaling (DVS) strategy has benefits just during early lifetime. Indeed, they found that for a total of 10 years lifetime the energy savings of DVS was just 7% with respect to guard-banding. In addition, by observing degradation reduction with respect to power gating factors, they found that high power gating factors may be feasible in designs where activity is naturally low. However, such power gating factors are typically accompanied by significant throughput reduction, limiting the feasibility of power gating as an NBTI mitigation technique [23]. Regarding activity management, they showed that adapting the processor configuration can reduce activity by up to 61% to achieve additional aging reduction and allow more power gating. However, this incurs significant performance degradation (up to 60% reduction in throughput) to achieve frequency benefit (reduced delay degradation) with respect to the baseline processor of only 4%. Relying on

these observations, Chan *et al* [23] concluded that guard-banding may still be the most efficient way to deal with aging.

Zhang & Dick [27] proposed scheduled voltage scaling, where the operating voltage of the IC is gradually increased to compensate for NBTI-related performance degradation. As many ICs have their own voltage regulators for use in dynamic voltage and frequency scaling (DVFS), they faced one challenge that is to integrate the NBTI-aware voltage levels into the existing set of voltages supported by the regulator. Hence, a voltage regulator providing non-uniform distributed voltage levels was required. They solved it by using voltage regulator with higher resolution digital to analog converter (DAC). Another challenge arises is when and how to switch from a lower voltage to a higher one. To tackle this problem, they proposed software and hardware approaches for ICs with and without DVS schemes. By adopting such technique, they found that the lifetime can be increased by 46% for ICs fabricated using a 45 nm process compared to guard-banding [27].

4.3.2 Power gating

Power gating, also called multi threshold CMOS (MTCMOS), is a well-known design technique used to reduce leakage power in conjunction with multi-V_{DD} design [45]. Power gating uses sleep transistors to disconnect the power supply from the rest of the circuit during standby periods [45]. The main drawbacks of power gating are due to the series sleep transistor that reduces the speed during normal operation and increases the circuit area [46]. Calimera *et al*.[47] observed that disconnecting a logic block from the ground voltage using a sleep transistor provides a natural way of reducing the NBTI effects. This stands true for footer-based power gated circuits In fact, when the sleep transistor is off, the ground terminal of the block is attached to a floating node (the virtual ground) which quickly reaches the logic “1” value, thus causing all nodes in the gated circuit to also reach the logic “1”, that is, the recovery state for NBTI. But, the application of power gating to a logic block increases its nominal delay (i.e., at time zero), as a consequence of the nonzero resistance of the sleep transistor when the circuit is active. Therefore, even if power gating reduces the NBTI-induced delay degradation over time, a larger time-zero delay may offset this benefit. To resolve this issue, Calimera *et al.* [47] proposed a clustered power-gating strategy, which allows a better control of the initial delay degradation than a solution in which the entire logic block is power gated (full power gating). Performance penalty

is reduced by power gating the critical cells with a larger sleep transistor (i.e., with a smaller penalty on active delay), and the noncritical cells with a smaller transistor (thus tolerating a higher time-zero delay penalty on them) [47]. To highlight the effectiveness of the proposed approach, they compared it with non-power gating and full power gated schemes on benchmark circuits mapped onto an industrial 45 nm technology, and for which the lifetime was defined as the time at which the circuit degrades its performance by 15% beyond its nominal value. The obtained results revealed that clustered architecture achieved 5X lifetime extension with respect to the non-gated version, with a negligible loss in leakage reduction (about 0.5% penalty with respect to full poser gating) [47].

Likewise, Ricketts *et al*. [28] investigated NBTI impact on different power saving cache strategies employing symmetric and asymmetric (with dual-Vth) 6-transistor (6T) and 8T SRAM cells. Such investigation aimed to figure out the best power saving strategy to be applied in cache design, if lifetime operation is a prime concern [28]. For that aim, they simulated NBTI in SRAM cells, designed with predictive technology model (PTM) for 45 and 32nm technologies [48], with the shifted value of pMOS threshold voltage due to NBTI after 5 years of time for different duty cycles. They found that both read SNM and WNM shifted with about 10% for both symmetric and asymmetric SRAM while, surprisingly, there was a 28.3% reduction in read SNM for the asymmetric SRAM cell as compared to symmetric SRAM cell [28]. Moreover, the WNM of the initial (t=0) SRAM cell is lower than the stressed (t=5years) SRAM cell implying higher voltage would be needed to write the aged cell. They found recovery in read SNM for different SRAM cache configurations varies from 38% to 66%. In addition, an increased rate of recovery of WNM and SNM in different SRAM caches has been seen for lower duty cycles. As such, they were able to identify the best configuration for low power cache memories, and which combines the sleep mode with the disabled mode using a timer that if it expires switches the cache block-off instead of sleeping. [28].

4.3.3 Duty-cycle tuning

One way to alleviate NBTI degradation consists in reducing the stress time and increasing NBTI recovery time through controlling signal probability. Many techniques have been proposed to mitigate NBTI at architecture-level based on balancing the stress and putting the timing-critical gates of the critical path into a relaxed mode [49]. However, it is not possible to put all critical

path nodes into a fully relaxed state (a signal probability of zero, SP(0) = 0%) or a fully stressed state (SP(0) = 100%). So, the object should be to reverse these signal probabilities to obtain signal probabilities that are as balanced as possible (around 50%), rather than reducing one signal probability to avoid NBTI stress. Kumar *et al*. [37] found that SNM worsens by about 8-9% after 10^8 seconds (~3years) for SRAM designed with BPTM 100 nm and 70nm devices [48]. By adopting R-D model, and performing simulations using V_{thp} shifted model files, they showed that about 30% of read stability (measured in terms of SNM) can be restored through cell flipping [37]. The implementation of SRAM cell flipping scheme have two main aspects: the ability to, periodically, flip the contents of all cells, and the ability to, correctly, read and write data during normal course of operation. The cell flipping could be performed either through software or hardware [37].

Jin and Wang [50] proposed an aging-aware design to combat the NBTI induced degradation in the instruction cache by duty-cycle balancing. Their analysis of the lifetime behavior on the instruction cache in the simulated microprocessor EV7 revealed that only 33.3% of the cachelines in the instruction cache are valid state (in use). Therefore, they proposed a periodic bit-flipping for cachelines in the invalid state., For cachelines in valid state, they proposed to first invalidate them, after the cacheline in the instruction cache remains idle for certain predefined interval, and then do the bit-flipping similar to the invalid cachelines [50]. By applying suh idle-time-based invalidation scheme for valid cachelines and bit-flipping/complementing scheme for invalid cachelines, the duty cycle ratio of the entire instruction cache could be balanced to ∼50% with minimized overheads (reduced duty cycle ratio to 56.2% for the valid cachelines with performance loss under 0.9%) [50].

4.3.4 Vth tunning

Adaptive body biasing (ABB) is a dynamic design technique based on tuning the devices threshold voltage at runtime to improve timing performance and instability. As such, an ABB circuit consists of a threshold voltage sensing circuit and an on-chip analog controller [35]. To compensate BTI aging effects and improve performance of SRAMs, Faraji *et al*.[35] proposed a low area overhead (ABB) circuit. The proposed circuit uses a control circuit and word line voltage to adjust the voltage applied to the body of 6T SRAM cell's transistors, in a way that their threshold voltage decreases in active mode (FBB) and increases in standby mode (RBB), so

the BTI effect dependence of threshold voltage is reduced [35]. The investigation has been conducted by simulation based on BSIM 32 nm predictive high-k metal-gate model [40] at 125°C where both static and alternating (with 0.5 duty-cycle) BTI stresses were considered. Under the worst case condition, the proposed ABB reduced the HOLD SNM degradation by 6.85%, the READ SNM degradation by 12.24%, the WRITE margin degradation by 2.16 %, the READ delay by 28.68%, and the WRITE delay by 32.61% compared to the conventional SRAM cell at 3 years [35].

4.3.5 Transistor sizing

Transistor sizing is a well know design technique used for a long time to optimize delay, area, power, or power-delay product. Since NBTI degrades the device drain current and thereby the circuit delay, the use of transistor sizing has been extended to mitigate NBTI degradation. The basic idea of transistor sizing for NBTI mitigation is to pre-calibrate the V_{th} degradation in each

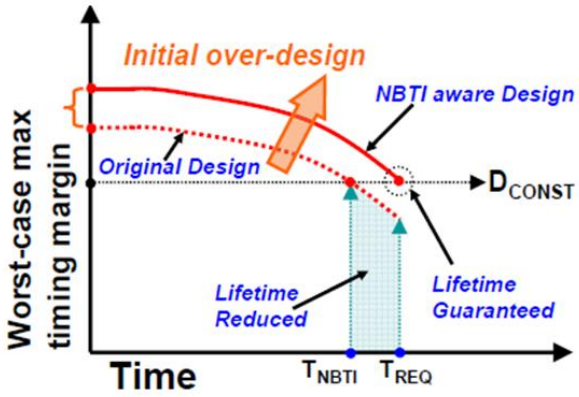


Fig. 4.14: NBTI aware transistor sizing. [51]

pMOS transistor due to NBTI considering the signal probability at the gate input. Then, sizing is applied while assuming that the circuit has degraded over the lifetime (see Fig 4.15 [51]. Many NBTI-aware transistor sizing approaches have been proposed in the literature. Kang *et al.*[51] proposed a transistor-level sizing algorithm based on a modified Lagrangian Relaxation (LR) technique to account for the temporal degradation of circuit and guarantee lifetime reliability of circuit under NBTI. They applied different sizing factors for PUN and PDN contrary to the single sizing ratio for each logic gate employed by the original LR. The simulation results on several ISCAS'85 benchmarks have shown that their transistor sizing approach can reduce the area

overhead of compensating the 3 year NBTI by an average of 43% compared to the conventional cell-level sizing method.

Similarly, Paul et *al.* [52] proposed a gate sizing approach, also using Lagrangian relaxation algorithm for global optimization of gate sizes subjected to NBTI degradation. Their algorithm estimated the Vth degradation of all individual PMOS transistors in a circuit based on their activities and sizes the circuit for a desired performance [52]. They applied their sizing approach on several ISCAS'85 benchmark circuits implemented in 70 nm BPTM [48]. The simulation results showed that sizing these circuits taking into account NBTI for a period of 10 years would entail only an average area overhead of 8.7% [52].

Kaffashian *et al* [53] proposed a genetic algorithm (GA) based transistor sizing approach to combat NBTI induced delay degradation in high fan-in domino OR circuits. Their optimization algorithm targeted to minimize degraded delay, during a lifetime of 3 years, subjected to area, power consumption, and unity noise gain (UNG) constraints. By considering a duty cycle of 50% for the clock signal and an activity factor of 0.5 for the inputs a dynamic OR circuit simulated using 65nm technology, the threshold voltage of pMOS transistors was estimated for 3 years lifetime using a power law NBTI prediction model. The simulation results of their GA-based sizing algorithm showed an improvement of more than 21.6% in delay during the circuit lifetime with a marginal power (1.5-2.4%) and UNG (0.05%) overheads compared to typical design [53].

Khan and Hamdioui [54] proposed an NBTI mitigation based on transistor sizing. Their technique consists in an optimization solution, carried out using the large scale gate sizing (LSGS) technique [55] to minimize the area of transistors in all the gates and ensures that the delay values of the current and adjacent gates are below the targeted value. Their NBTI gate delay model considers both the degradation due to the gate own transistors as well as that due to transistors of its adjacent gates [54]. The simulation results showed that their transistor sizing technique for NBTI mitigation realizes a delay reduction of about 45% as compared to the original circuit applied to ISCAS'85 benchmarks; and adopting 45nm PTM [48] for pMOS V_{th} model. They showed that with an average of 12% area overhead, the circuit delay will not exceed 15% after 10 years operation [54].

Yang et *al.*[56] proposed an NBTI-aware gate sizing approach where the circuit is first mapped onto a technology library (0.13µm technology from MOSIS), a logic simulation is performed to obtain the probability that the pMOS transistors in the circuit. NBTI degradation simulation is conducted to assess each gate delay increase. Then, path sensitization is performed (monitoring the pMOS transistors in the top 10% of the longest paths in the circuit) to reduce a number of constraints. The sizing was fixed using an off-the-shelf mathematical solver. The circuit delay shift due to NBTI degradation over 10 years of circuit operation is calculated. For most MCNC'91 benchmark circuits, their NBTI-aware sizing technique resulted in less than 1% area increase over baseline sizing [56].

For mitigating NBTI in large logic circuits, the main idea of transistor sizing consists in optimizing the library cells to balance their rise and fall delays at the expected lifetime rather than the design time. Library Cells are therefore optimized for NBTI aging considering both the Vth shift to signal probability and the distribution of SP of internal nodes in typical circuits. After obtaining the new aging-aware cells, each cell is characterized to obtain the delay look-up tables (LUTs for different load capacitances, transition times, and ΔVth. In the mapping phase, only the critical gates in the gate-level netlist are replaced with the new cells [57]. Kiamehr et *al*. [57] considered four different scenarios for the discretization of the SP range and showed that the lifetime improvement saturates when the number of SP samples exceeds a particular limit. When compared to the normal standard cell library design in terms of lifetime, their simulation results showed that the lifetime can be improved by approximately 150% with negligible area/power overheads [57].

Amrouch et *al.* [58] proposed degradation-aware cell libraries to suppress aging in cell-based IC design. They proposed to extend the cell library to account for aging by re-characterizing cells under varied aging scenarios (Operation Conditions (OPCs): based on the slew of the gate input signals and the load capacitance of its output. They showed that optimizing aging is not only a matter of balancing the duty cycle of transistors but it is rather about selecting the most suitable gate with respect to aging, based on the existing OPCs. They created 121 degradation-aware cell libraries which can be used by a timing analysis tool enable accurate analysis of the timing behavior of the entire circuit in the scope of static and/or dynamic aging stress. Hence, the required guard-band can be obtained [58].

In the light of the techniques we reviewed above, we can say that no single mitigation technique suits all the circuits undergoing NBTI degradation. In fact, the choice of the NBTI mitigation technique is guided and/or governed by the circuit type, its topology framework application, and the abstraction level at which the technique is introduced. In the next section we present our NBTI-aware 6T-SRAM design approach, and which is based on transistor sizing technique.

4. 4 AGING-AWARE 6T-SRAM DESIGN APPROACH

Conventionally, the 6T-SRAM cell design or sizing has been seen as a trade-off between area, power, performance, and yield [59]. Subsequently, 6T-SRAM cell could be sized for high density, high performance, or low voltage memory. Yet, the cell cannot maintain its stability when it goes old as the devices parameters would shift from the origin values due to aging. This would appeal to re-examine the 6T-SRAM sizing in the prospect of its aging too.

Recall that in the literature of the 6T-SRAM cell design; only the cell ratio (CR) and pull-up ratio (PR) are resolved for read stability and write-ability constraints. But, no clue about sizing the access transistors exists. Actually, CR & PR do resolve the sizing of drive transistors and pull-up transistors against access transistors, but not the access transistor themselves. This renders the challenge in 6T-SRAM design to be the access transistors sizing as their w/l needs to be as large as possible to improve access time and write-ability and on the other hand it needs to be as small as possible for a better read stability of the cell. Thus, no rule other than; “weak pull-up transistors, medium access transistors, and strong drive transistors”, usually found in the literature, exists for 6T-SRAM sizing. On the other hand, NBTI mitigation based on transistor sizing is understood as upsizing the pMOS transistor. In the light of what is said, and on the contrary to what is expected, we propose to size nMOS transistors instead of pMOS transistors to mitigate NBTI degradation undergone by these latter! Our approach consists in sizing the access transistors of the 6T-SRAM cell for a better hold stability, and the rest of the cell transistors based on CR & PR for a better compromise between its read stability and write-ability of the cell till its end of life.

4.4.1 Simulation setup & NBTI model

The aging-aware 6T-SRAM approach we propose is introduced at circuit-level where the NBTI model, expressed by equation (4.5), has been integrated to Cadence Relxpert simulator through its URI as mentioned in section 4.2.1.

The NBTI impact on 6T-SRAM stability has been analysed in terms of hold, read, and write margins using the simulation setups shown on Fig. 4.15.

For hold operation (Fig.4.15 (a)), the access transistors are turned off (WL=0). The noise voltage is swept from -1.8 V to +1.8 V, and the voltage at the cell bit (Q) and its complement (QB) is tracked. For read and write operations (Fig.4.15 (b)), the access transistors are turned on (WL = V_{DD}) and a noise source is connected to the node at level 0 V and swept from 0 to 1.8V, according to n-curve method [60]. The drain current at that node is tracked.

Note that Age parameter used in aging simulation was set to 5 years. The simulation was conducted for typical process, voltage; temperature (PVT) corner and the obtained results are reported hereafter.

(a) (b)

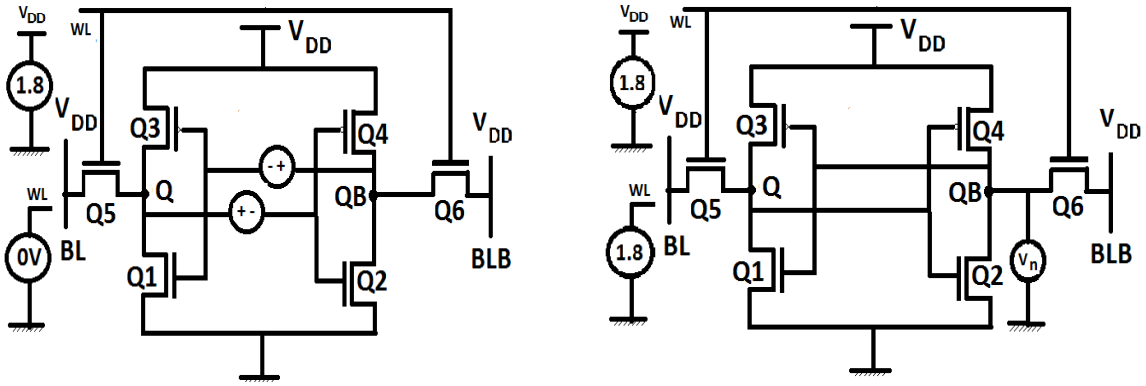


Fig. 4.15: 6T-SRAM simulation setup for hold (a,) read & write operations using n-curve method (b).

4.4.2 Simulation results

Transistor sizing impact on NBTI mitigation in 6T-SRAM cell is analysed in terms of its static stability in hold, read, and write operations.

4.4.2.1 Hold operation

Data retention or the static stability of a 6T-SRAM cell memory is analyzed in terms of its static noise margin (SNM). The SNM characterizes the maximum noise voltage that can tolerate the cell without flipping its data bit. Therefore, higher the SNM is, more stable is the cell.

Recall that SNM can be determined as the side of the largest square that fits in the SRAM's butterfly. This latter can be obtained by overlapping the voltage transfer curve (VTC) of one inverter of the SRAM as shown in Fig. 4.16. When the two lobes are not symmetrical, the SNM is then extracted from the smallest lobe enclosed by the two VTCs.

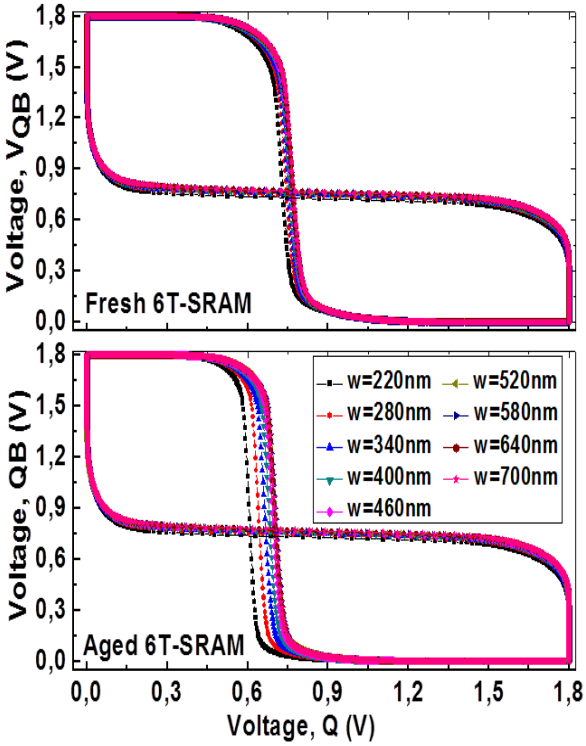


Fig. 4.16: The trend of hold SNM of the 6T-SRAM fresh and aged cell for different transistor sizes

It is clear, from Fig .4.16, that aging does move the butterfly to the left-side narrowing thereby the enclosed lobe which results in the decrease of the SNM for the aged cell. When analysing the cell stability as a function of cell size, it reveals that SNM increases with transistors' width increasing for both the fresh and the aged cell as shown on Fig. 4.17(a).

It is also worth to note for a high-density cell, where all transistors are of the minimum size allowed by the technology (in this case w_n=220 nm) the cell experiences the highest degradation of its SNM with aging.

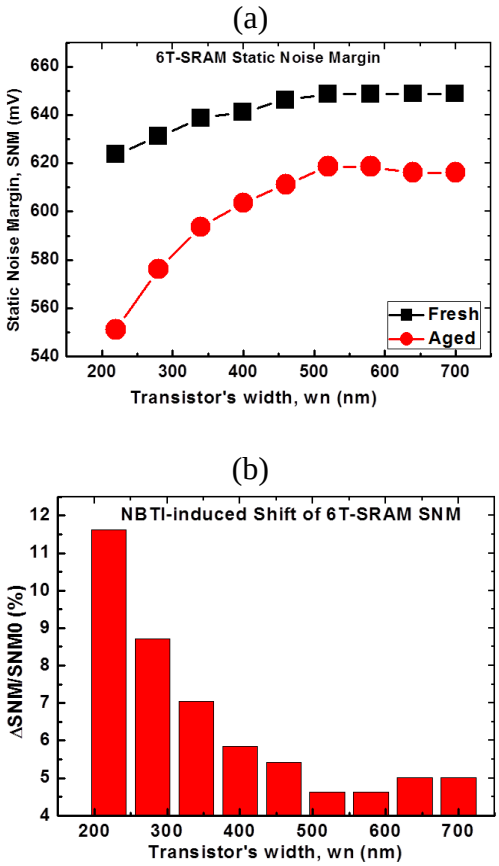


Fig. 4.17: The shift of Hold SNM of a symmetric 6T-SRAM cell subjected to NBTI for different transistor widths

So, enlarging the cell's size would increase its SNM when aged and shorten its shift due to NBTI, as shown on Fig. 4.17(b). For example, when the transistors width is enlarged from 220 nm to 400 nm, the SNM shift would shrink to the half (as it drops from 11.62% to 5.84%). Yet, enlarging the transistors of the SRAM cell beyond a certain value (540 nm) won't boost anymore its SNM. In fact, this latter tends to saturate for both fresh and aged cells. Afterwards, a further enlargement of the size cell beyond that value does not come with any significant improvement of its robustness, and becomes thereby useless. As a result, we can set the width of the nMOS access transistors to a higher than the minimum allowed width by the technology to ensure a

better hold stability. Then, proceed by sizing of the rest of the cell transistors for a better cell read stability and write-ability. It is well known that read stability and write-ability are antagonist targets in 6T-SRAM cell design as they adversely depend on the access transistors width. Hence, we study the behaviour of the aged cell in terms of read stability and write-ability as a function of its sizing. For that aim, we adopted the n-curve metrics [60] to analyse the two operations within the same simulation setup. Fig. 4.18 shows the trend of the n-curve of both fresh and aged 6T-

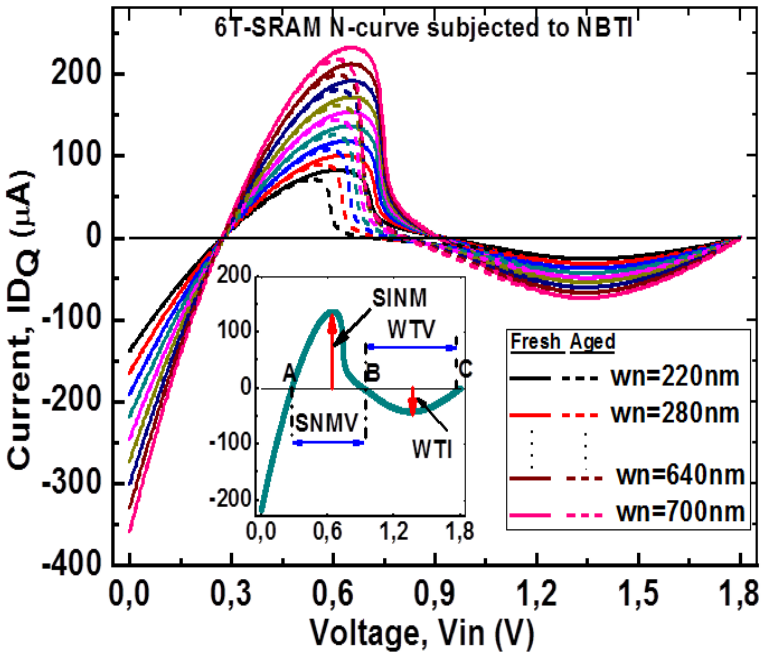


Fig. 4.18: N-curves of fresh and aged 6T-SRAM for different transistors' width and the corresponding extracted read & write features.

SRAM cells with transistor sizing. Different metrics can be extracted to analyse the read stability and the write-ability of the 6T-SRAM in terms of voltage and current. For read operation, both the static voltage noise margin (SVNM) and static current noise margin (SINM) are examined while for write operation, both the write trip voltage (WTV) and write trip current (WTI) are considered. These features could be extracted from the n-curve, as shown in Fig. 4.18 (inset graph). Note that the n-curve is characterized by three voltages where no current flows from or to the sweep voltage source (Vn) and which are delimited by points A, B, and C in Fig. 4 .18 (inset graph). Voltage at point A is determined by the cell ratio (CR) defined by the ratio between the drive and the access transistor sizes. Voltage at point C is defined by the pull-up (PR), the

ratio between the pull-up to and the access transistor size. Voltage at point B depends on to both the drive to pull-up transistors size ratio and the access transistor width.

4.4.2.2 Read & Write operations

Referring to Fig. 4.18, it seems that only point B is shifting with respect to aging and width variation while points A and C are not. Then, any increase of SVNN will entail the decrease of WTV and vice versa. As the shift is towards to left-side, then it is supposed to decrease the read stability and increase the cell write-ability. In fact, as shown on Fig. 4.19 (a), SVNM decreases with aging. This means, that aging deteriorates the cell read stability. When analyzed with respect to transistor sizing, SVNM tends to slightly decrease in the fresh cell with transistor enlarging, while it increases in the aged cell. This minor decrease of SVNM in fresh cell cannot be interpreted as read stability decreasing with sizing as it is associated with a larger SINM compared to that of the aged cell (see Fig.4.19 (b)). Subsequently, the required noise charge to disturb the fresh cell is found to be larger than that required for disturbing the aged one. Similarly, the cell write-ability ought to be analyzed in terms of both WTV and WTI. WTV is the minimum voltage drop needed to change the internal node "1" with both the bit-lines cell. For the write-ability of the cell it is preferable to have small WTV and WTI. Yet, no shift has been induced in WTI by NBTI (see Fig.4.19 (b)). This means no change in the amount of current is needed to write the aged cell when both bit-lines are kept at V_{DD}. However, WTI should be large enough to cope with read stability requirements [60]. On the other hand, lower is the WTI; higher is the write-trip point of the cell. This justifies the high WTV required to write the aged cell with small transistors. Consequently, the aged cell is becoming slower to write under NBTI. Then, if analyzed in terms of the NBTI-induced shift, SVNM, SIMN, and WTV (see Fig .4.19 (c) and (d)) seem to mimic the hold stability trend with transistors' width (see Fig. 4.17). Therefore, larger the transistors width is, better the cell's read stability and write-ability are when the cell gets old.

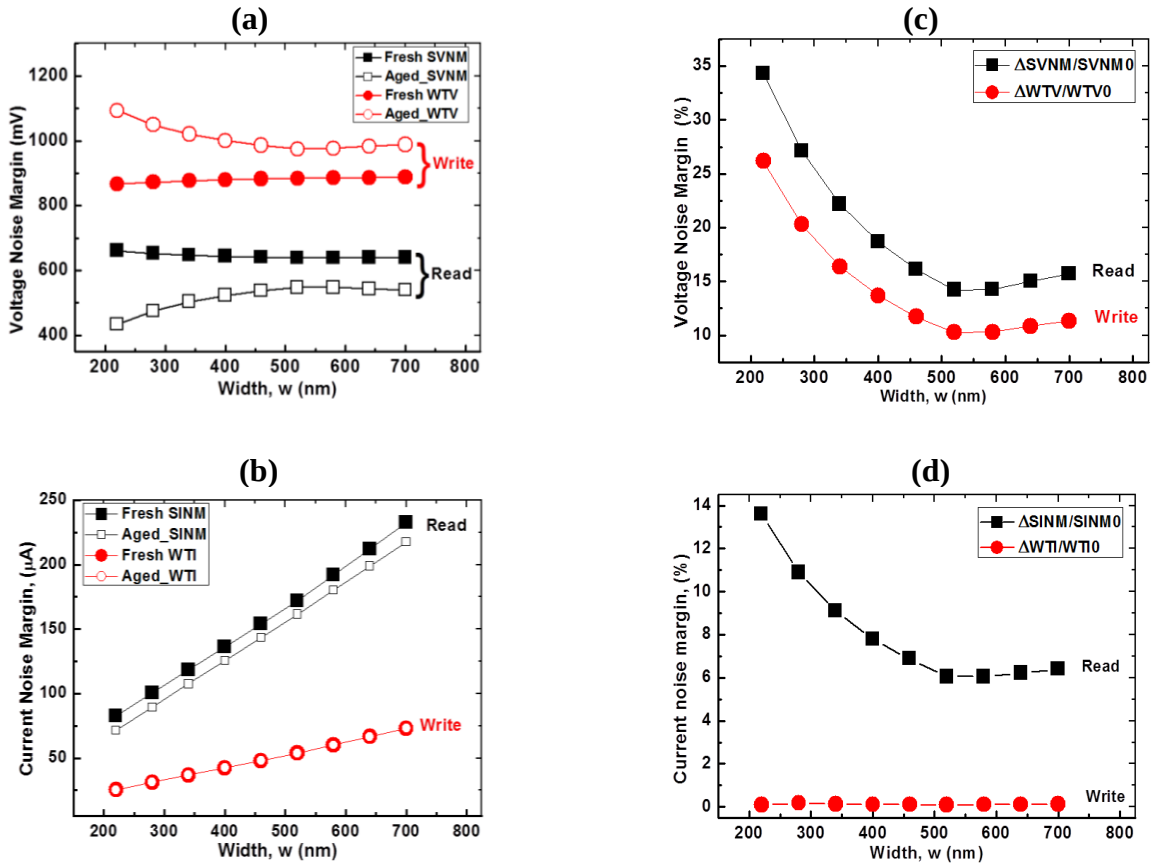


Fig. 4.19: 6T-SRAM read stability and write-ability features subjected to NBTI for different transistor widths.

In the light of the results shown on Figs. 4.17 and 4.19, we can say that a high-density 6T-SRAM (with minimum-sized transistors) is the most vulnerable to NBTI degradation and becomes more prone to read instability when it ages. However, when it comes to other memory applications, this could be a good start point to sizing the 6T-SRAM access transistors to mitigate NBTI aging elucidated hereafter.

4.4.2.3 NBTI-aware Cell Sizing

On the contrary of high-density memory, high-performance 6T-SRAM is always designed with transistors larger than the minimum width allowed by the technology. Its design responds to two antagonist constraints: a nondestructive read and write-ability of the cell. These two features

are governed by the cell ratio (CR) and the pull-up ratio (PR). CR and PR can be simply expressed in terms of widths as:

$$CR = \frac{\left(\frac{w}{l}\right)_1}{\left(\frac{w}{l}\right)_5} = \frac{\left(\frac{w}{l}\right)_2}{\left(\frac{w}{l}\right)_6} = \frac{w_1}{w_5} = \frac{w_2}{w_6} \tag{4.6}$$

$$PR = \frac{\left(\frac{w}{l}\right)_3}{\left(\frac{w}{l}\right)_5} = \frac{\left(\frac{w}{l}\right)_4}{\left(\frac{w}{l}\right)_6} = \frac{w_3}{w_5} = \frac{w_4}{w_6} \tag{4.7}$$

These two ratios are chosen to resolve the read and write ability constraints. Yet, no analytical formula exists to size the nMOS access transistors. In view of that, we recommend to size the access transistor to minimize the NBTI effect in a unit-ratio CMOS 6T-SRAM (CR=PR=1) for an optimum hold stability, then size the drive and pull-up transistors with CR and PR higher than 1 and tune their range according to the stability of the aged cell. In fact, results depicted in fig 4.17 reveal that setting the width of cell transistors to the double of the min allowed width would reduce the NBTI-induced shift in hold operation by 2 times . Thus, we choose to set the access transistors width to 440 nm and vary CR and PR from 1 to 2.5 to assess their impact on the NBTI mitigation in read and write operations.

Figure 4.20 shows the n-curves for 6T-SRAM for PR=1.5. At the first glance, we note the n-curve differs from that obtained for the 6T-SRAM with same sized transistors (CR=PR=1) (see Fig. 4.18). In fact, the starting point for the n-curves point at n = 0V is fixe under different CR. Moreover, voltage point A does shift to the left with CR increasing and also do the maximum currents for SINM and WTI. The extracted 6T-SRAM cell stability features are illustrated in Fig. 4.21. It is clear that SVNM is decreased by NBTI (see Fig. 4.21 (a)). However, the induced shift in SVNM tends to decrease with CR increasing (see Fig. 4.21 (b). On the other hand, PR increasing does not actually come with any gain to mitigate NBTI. Therefore it outcomes that pMOS upsizing does not really alleviate NBTI impact on the read stability of the 6T-SRAM cell. To better understand the trend of read stability with aging and as a function of cell ratio, let us first analyze the n-curve of the fresh cell.

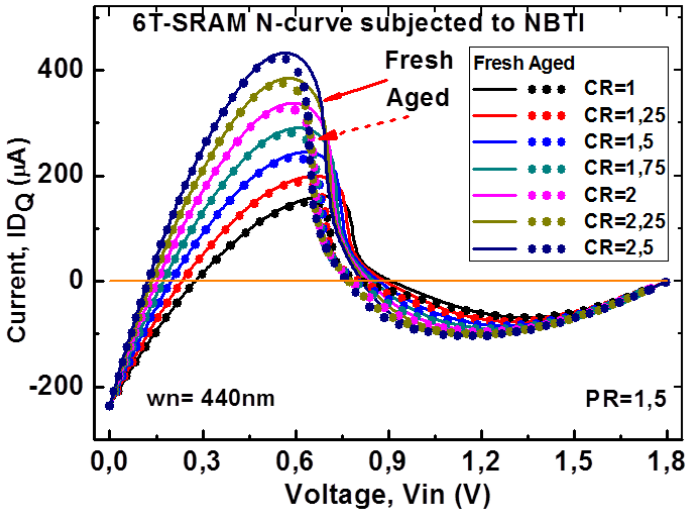


Fig. 4.20: The 6T-SRAM n-curve subjected to NBTI for different cell ratios.

At the beginning Vn = 0, the access and the drive transistors are operating in saturation mode and linear regime respectively, while the pull-up transistor is off. This makes drain current of the former higher than that of the latter. This current difference flows into the sweep voltage source which justifies the negative sign of the current on Fig. 4.20. As Vn increases to reach voltage at point A, the current trough the drive transistor increases until it equals that of the drive transistor and no current flows into the sweep voltage source. When Vn goes higher than voltage at point A, ID_Q continues to increase to reach it maximum noted by SINM and the drive transistor moves to its saturation mode. Then, the pull-up transistor turns on which leads to the current decrease until no current flows from the sweep voltage source when the voltage reaches point B. Then, from where is defined SINM to that of WTI, the three transistors of the right half-cell of 6T-SRAM are in their saturation mode. Then, voltage B is related to drive to pull-up ratio and the access transistor. This justifies the relatively small left-shift of point B compared to that of point A. Yet, the voltage difference between points A and B that is SVNM tends to increase with CR increasing. This means higher the CR, better read stability the 6T-SRAM cell.

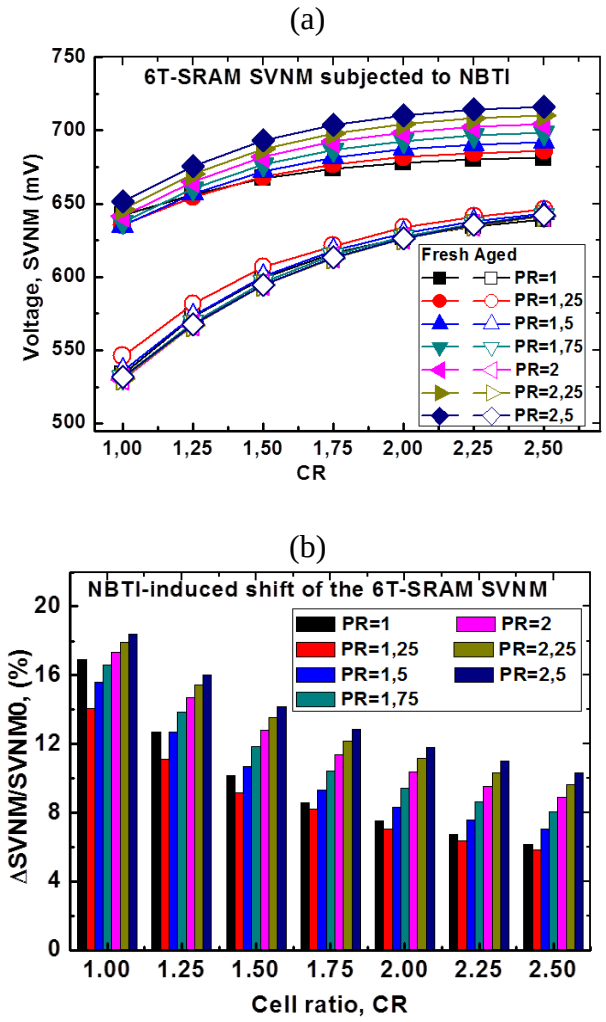


Fig. 4.21: The trend of 6T-SRAM cell read stability (the static voltage noise margin in (a) and its corresponding NBTI- induced shift ratio in (b)) with transistor sizing

Now let us analyze the read stability under NBTI. As the sweep voltage source is connected to cell node stocking 0 (QB in Fig.4.15 (b)), then it is the pull-up transistor of the left inverter that would undergo NBTI stress. Consequently, it would be difficult to maintain a logic ‘‘1” at the input of the right inverter. This would turn on its pull-up transistor earlier. This happens when the three transistors at node QB are conducting, which means transition from SINM to WTI points accelerated are on the n-curve and voltage at point B is reached faster. Therefore, SVNM of the aged cell is decreased.

Therefore, increasing CR would decrease the NBTI-induced shift of SVNM, as shown in Fig. 4.21, making the 6T-SRAM cell more stable during read operation when it ages. On the other side, PR increasing does not significantly impact the read stability. Actually PR slightly improves the SVNM, for a given CR, in the fresh cell but not in the aged one. PR increasing makes the current of the pull-up transistor to increase, while NBTI decreases. Consequently, PR increasing would not come improving the read stability of the cell when it ages.

6T-SRAM write-ability subjected to NBTI is depicted in Fig. 4.22. It can be seen that WTV increases with CR increasing in both the fresh and the aged cells but with a slower rate in the latter. Hence, it becomes more difficult to write the cell with a higher CR. Therefore, increasing CR degrades the cell write-ability. However, increasing PR would decrease WTV of the fresh 6T-SRAM cell improving thereby its write-ability. But, it marginally affects that of the aged cell. In fact, increasing CR would make the drive transistor stronger than the access transistor preventing voltage at node storing “0” from increasing. This makes the cell more stable during read as it renders it more difficult to turn on the left drive transistor to flip its state. However, at WTI the drive transistor moves to the cut-off region then PR increasing would strengthen the pull-up transistors which will increase voltage at node 0 making it possible to turn-on the drive transistor of the left inverter and flip its output. This would then flip the other inerter state making the cell writable. But, PR increasing slightly improves the write-ability of the aged cell because of the compensating effect with NBTI-induced current decrease in the pull-up of the left inverter. Likewise, results of Fig. 4.22 let us saying that lower PR is, better the cell write-ability is when it ages. Therefore, enlarging pMOS pull-up transistors in this case has not come to improve the 6T-SRAM cell write-ability subjected to NBTI. To be more confident about the aforementioned findings, the analysis of the cell stability has been also expressed in terms of both the static power noise margin (SPNM) and the write-trip power (WTP). These two metrics enfold both voltage and current information then can be used analyze the read stability and write-ability trend with transistor sizing instead of the four metrics SVNM, SINM ,VTP, and WTI, when these latter present controversy trends between voltage and current trends [60]. Therefore, NBTI impact on the read and write operations of 6T-SRAM, expressed in terms of SPNM and WTP features, is illustrated in Fig. 4.23.

Recall that higher and lower SPNM and WTP are, better are the read stability and the write-ability of the 6T-SRAM cell. From results depicted in Fig. 4.23, it can be observed that NBTI has degraded both read stability and write-ability. Increasing CR would, on one hand, increase the read stability in both fresh (Fig. 4.23(a)) and aged cell (Fig.4.23 (b)) and decreases its NBTI-induced degradation (Fig.4.23 (c)). On the other hand, it degrades write-ability in both fresh (Fig.

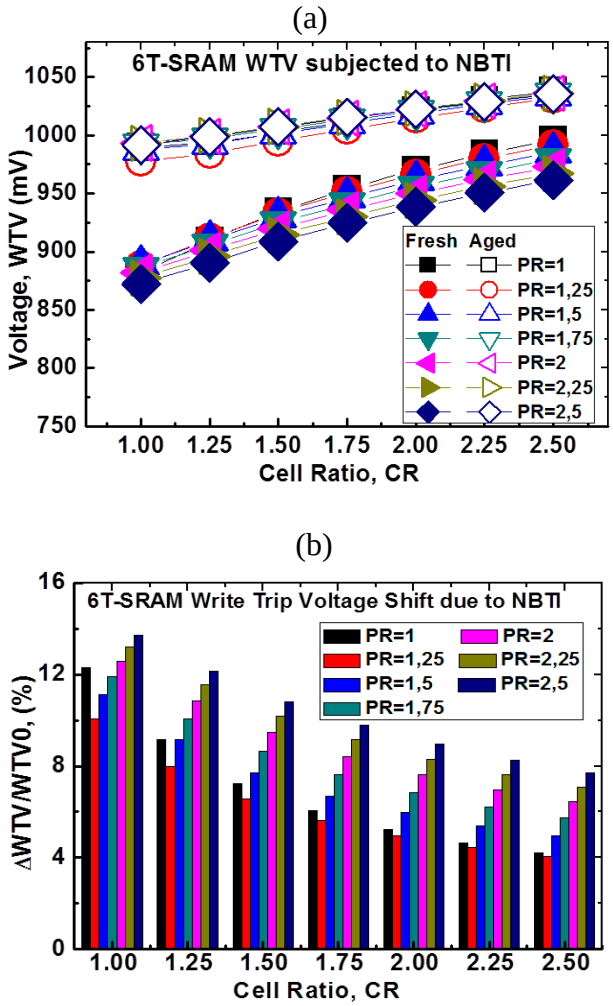


Fig. 4.22: The trend of 6T-SRAM cell write-ability (the write trip voltage in (a) and its corresponding NBTI- induced shift ratio in (b)) with transistor sizing.

4.23 (a)) and aged cell (Fig.4.23 (b)). Hence, CR increasing would not significantly reduce the NBTI-induced write-ability degradation ((Fig.4.23 (d)). Moreover, enlarging PR does not really come improving either the read stability or the write-ability of the cell when its ages. Therefore,

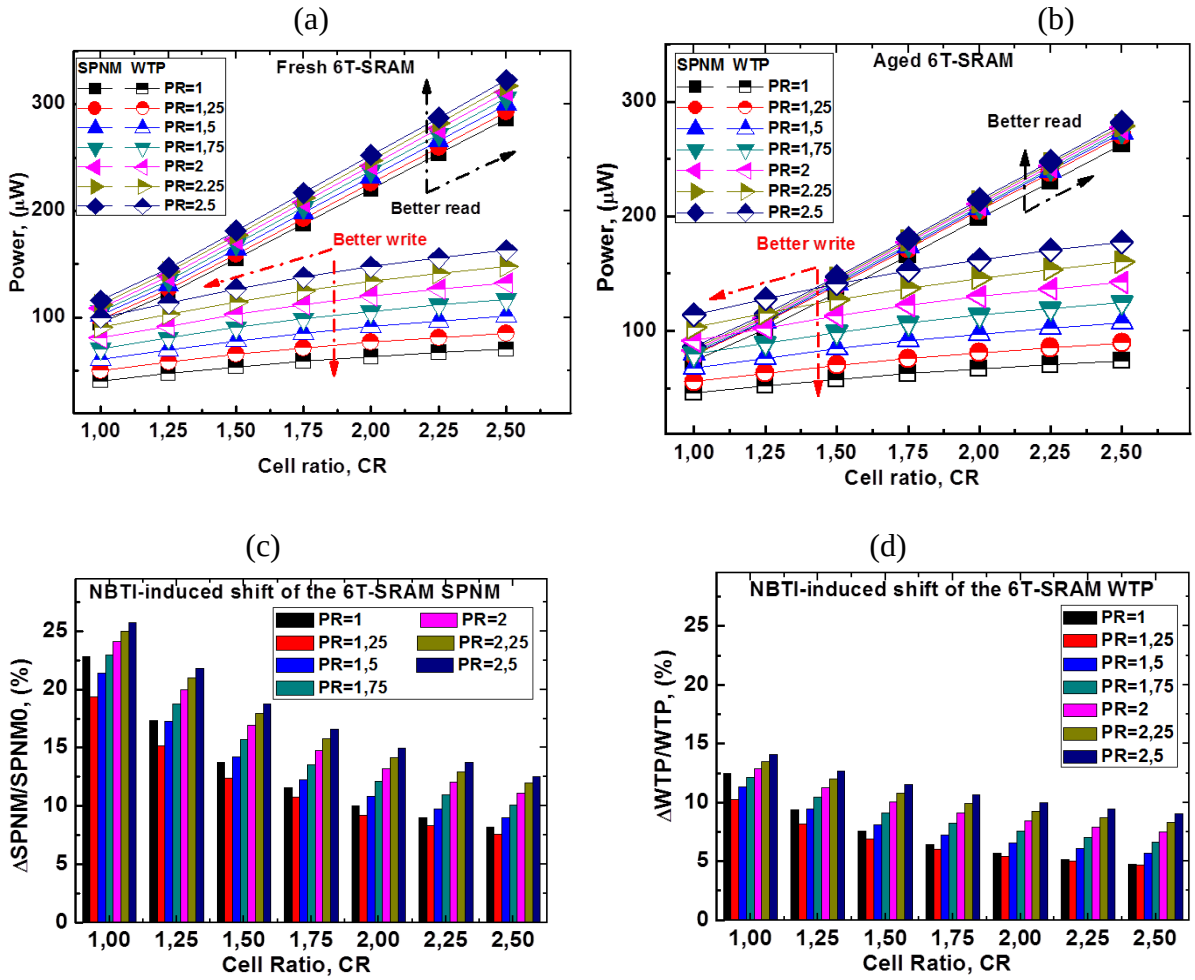


Fig. 4.23: 6T-SRAM power metrics for read stability and write-ability degradation due NBTI as a function of cell sizing (CR and PR ratios)

PR larger than 1.25 ((Fig.4.23 (c) & (d)) would increase the NBTI-induced degradation of both read stability and the write-ability of the cell.

4.5 DISCUSSION

A 6T-SRAM cell is usually sized for a high density, or high performance, or a low voltage memory. Then, cell design consists in optimizing read stability and write-ability at time zero, However, with temporal variation; the cell devices cannot preserve their electrical parameters when they get old. This would entail to re-examine its sizing in the prospect of its aging too.

In fact, we found that for a high-density memory, where the 6T-SRAM cell is designed with minimum-sized transistors (l_{min} = 180 nm & w_{min} = 220 nm for TSMC 0.18µm), the cell undergoes the largest degradation of its SNM exceeding 10% after 5 years (see Fig. 4.17). This shift can be brought below 10% by enlarging the cell transistors by 60 nm. Yet, the generally accepted analytical model of SNM driven in [61] cannot foresee this finding as according to such model the SNM of the 6T-SRAM is supposed to vary as a function of CR and PR rather than the cell transistors size. This is due to the fact that the model uses constant threshold voltages (equal for n and p-channel) and neglects second-order effects such as mobility reduction and velocity saturation. So, the analytical SNM model [61] would resolve sizing the 6T-SRAM drive and pull-up transistors with regard to access transistors but does not advice on the sizing of these latter. Henceforth, we advise sizing access transistors to mitigate NBTI aging in the 6T-SRAM cell. Actually, the simulation results depicted in Fig. 4.17 and 4.19 revealed the existence of an optimal access transistor width, in the range of 2-3 times the minimum width allowed by the technology, that would drop by 2 times the NBTI-induced degradation of 6T-SRAM cell stability features in hold, read, and write operations over 5 years. Once the optimal width of the access transistors is set, the rest of the 6T-SRAM transistors would be sized by choosing CR and PR that reduce NBTI degradation of read stability and write-ability using n-curve metrics. Moreover, when the four n-curve metrics are combined in the two power metrics SPNM and WTP for, a better NBTI-induced shift of the 6T-SRAM cell robustness can be estimated and appropriate CR and PR could be obtained (see Fig. 4.23). As reported earlier, a PR value of 1.25 would be the optimal ratio for pull-up pMOS transistors sizing and a CR value higher than 1.5 for drive nMOS transistors sizing to ensure a hardened cell against NBTI in hold, read, and write operations. These findings make the nMOS transistor sizing rather than pMOS a good means to mitigate NBTI in SRAMs not from the prospect of transistor granularity level but from the whole cell granularity.

4.6 CONCLUSION

In this chapter, we proposed a review of transistor sizing for NBTI mitigation in the 6T-SRAM cell. We showed that contrary to what is expected, pMOS upsizing alone does not significantly alleviate NBTI induced degradation of SRAM performance. In fact, we found upsizing pMOS beyond a certain threshold would be useless in terms of both reliability and area. However, properly sizing the nMOS access transistors would improve significantly the cell stability subjected to NBTI aging. Therefore, the 6T-SRAM design would be re-thought as access transistor sizing for NBTI mitigation. Consequently, we think that 6T-SRAM sizing needs to be re-examined to find, if not an optimal sizing for NBTI aging, the best tradeoff between reliability, area and other cell design constraints. Likewise, if other design constrains other than reliability could not be met by our proposed sizing, then other NBTI mitigation techniques such as footer-based low power technique could be coupled with the optimal sizing to bring the degradation at a lower level as possible.

REFERENCES

[1] 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," *Microelectronics Reliability,* vol. 45, pp. 31-38, 2005/01/01/ 2005.

[2] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability,* vol. 46, pp. 1-23, 2006/01/01/ 2006.

[3] D. K. Schroder, "Negative bias temperature instability: What do we understand?," *Microelectronics Reliability,* vol. 47, pp. 841-852, 2007/06/01/ 2007.

[4] S. Khan, S. Hamdioui, H. Kukner, P. Raghavan, and F. Catthoor, "BTI impact on logical gates in nano-scale CMOS technology," in *2012 IEEE 15th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS)*, 2012, pp. 348-353.

[5] T. Chan, W. J. Chan, and A. B. Kahng, "On Aging-Aware Signoff for Circuits With Adaptive Voltage Scaling," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 61, pp. 2920-2930, 2014.

[6] H. Hong, J. Lim, H. Lim, and S. Kang, "Lifetime Reliability Enhancement of Microprocessors: Mitigating the Impact of Negative Bias Temperature Instability," *ACM Comput. Surv.,* vol. 48, p. Article 9, 2015.

[7] N. Khoshavi, R. A. Ashraf, R. F. DeMara, S. Kiamehr, F. Oboril, and M. B. Tahoori, "Contemporary CMOS aging mitigation techniques: Survey, taxonomy, and methods," *Integration,* vol. 59, pp. 10-22, 2017/09/01/ 2017.

[8] X. Chen, Y. Wang, Y. Cao, Y. Xie, and H. Yang, "Assessment of Circuit Optimization Techniques Under NBTI," *IEEE Design & Test,* vol. 30, pp. 40-49, 2013.

[9] E. Maricau and G. Gielen, "Background on IC Reliability Simulation," in *Analog IC Reliability in Nanometer CMOS*, E. Maricau and G. Gielen, Eds., ed New York, NY: Springer New York, 2013, pp. 79-91.

[10] J. B. Bernstein, M. Gurfinkel, X. Li, J. Walters, Y. Shapira, and M. Talmor, "Electronic circuit reliability modeling," *Microelectronics Reliability,* vol. 46, pp. 1957-1979, 2006/12/01/ 2006.

[11] C. d. systems. (12/12). Available: https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/custom-ic-analog-rf-flows/legato-reliability-solution.html

[12] Synopsys. *Reliability Analysis*. Available: https://www.synopsys.com/verification/ams-verification/reliability-analysis.html

[13] M. graphics. *Addressing IC Reliability Issues Using Eldo*. Available: https://www.mentor.com/products/ic_nanometer_design/resources/overview/addressing-ic-reliability-issues-using-eldo-396d66e1-037c-4c03-b6d1-324f87205f88

[14] Y. Cao, J. Velamala, K. Sutaria, M. S. Chen, J. Ahlbin, I. S. Esqueda, M. Bajura, and M. Fritze, "Cross-Layer Modeling and Simulation of Circuit Reliability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 33, pp. 8-23, 2014.

[15] L. Zhihong, B. W. McGaughy, and J. Z. Ma, "Design tools for reliability analysis," in *2006 43rd ACM/IEEE Design Automation Conference*, 2006, pp. 182-187.

[16] C. Hu, "The Berkeley reliability simulator BERT: an IC reliability simulator," *Microelectronics Journal,* vol. 23, pp. 97-102, 1992/04/01/ 1992.

[17] R. H. Tu, E. Rosenbaum, W. Y. Chan, C. C. Li, E. Minami, K. Quader, P. K. Ko, and C. Hu, "Berkeley reliability tools-BERT," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 12, pp. 1524-1534, 1993.

[18] C. D. Systems, "Virtuoso Unified Reliability Interface

Reference," ed, October 2015.

[19] B. Djezzar, H. Tahi, A. Benabdelmoumene, A. Chenouf, and Y. Kribes, "A New Method for Negative Bias Temperature Instability Assessment in P-Channel Metal Oxide Semiconductor Transistors," *Japanese Journal of Applied Physics,* vol. 51, p. 116602, 2012/11/05 2012.

[20] E. I. service. (12/12/2020). *MPW prototyping*. Available: https://europractice-ic.com/mpw-prototyping/asics/tsmc/

[21] R. Vattikonda, W. Wenping, and C. Yu, "Modeling and minimization of PMOS NBTI effect for robust nanometer design," in *2006 43rd ACM/IEEE Design Automation Conference*, 2006, pp. 1047-1052.

[22] H. Lee, M. Shafique, and M. A. A. Faruque, "Aging-Aware Workload Management on Embedded GPU Under Process Variation," *IEEE Transactions on Computers,* vol. 67, pp. 920-933, 2018.

[23] T. Chan, J. Sartori, P. Gupta, and R. Kumar, "On the efficacy of NBTI mitigation techniques," in *2011 Design, Automation & Test in Europe*, 2011, pp. 1-6.

[24] X. Guo and M. R. Stan, "Implications of accelerated self-healing as a key design knob for cross-layer resilience," *Integration,* vol. 56, pp. 167-180, 2017/01/01/ 2017.

[25] F. Oboril, M. Ebrahimi, S. Kiamehr, and M. B. Tahoori, "Cross-layer resilient system design flow," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2015, pp. 2457-2460.

[26] K. Kang, H. Kufluoglu, K. Roy, and M. A. Alam, "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 26, pp. 1770-1781, 2007.

[27] L. Zhang and R. P. Dick, "Scheduled voltage scaling for increasing lifetime in the presence of NBTI," in *2009 Asia and South Pacific Design Automation Conference*, 2009, pp. 492-497.

[28] A. Ricketts, J. Singh, K. Ramakrishnan, N. Vijaykrishnan, and D. K. Pradhan, "Investigating the impact of NBTI on different power saving cache strategies," in *2010 Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, 2010, pp. 592-597.

[29] A. Calimera, E. Macii, and M. Poncino, "NBTI-aware power gating for concurrent leakage and aging optimization," in *ISLPED*, 2009.

[30] A. Calimera, M. Loghi, E. Macii, and M. Poncino, "Dynamic Indexing: Leakage-Aging Co-Optimization for Caches," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* vol. 33, pp. 251-264, 2014.

[31] T. T. Kim and Z. H. Kong, "Impacts of NBTI/PBTI on SRAM VMIN and design techniques for SRAM VMIN improvement," in *2011 International SoC Design Conference*, 2011, pp. 163-166.

[32] Z. C. Lee, K. C. Leong, Z. H. Kong, and T. T. Kim, "NBTI/PBTI-Aware WWL Voltage Control for Half-Selected Cell Stability Improvement," *IEEE Transactions on Circuits and Systems II: Express Briefs,* vol. 60, pp. 602-606, 2013.

[33] H. Mostafa, M. Anis, and M. Elmasry, "Adaptive Body Bias for Reducing the Impacts of NBTI and Process Variations on 6T SRAM Cells," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 58, pp. 2859-2871, 2011.

[34] H. Mostafa, M. Anis, and M. Elmasry, "NBTI and Process Variations Compensation Circuits Using Adaptive Body Bias," *IEEE Transactions on Semiconductor Manufacturing,* vol. 25, pp. 460-467, 2012.

[35] R. Faraji and H. R. Naji, "Adaptive Technique for Overcoming Performance Degradation Due to Aging on 6T SRAM Cells," *IEEE Transactions on Device and Materials Reliability,* vol. 14, pp. 1031-1040, 2014.

[36] A. P. Shah, N. Yadav, A. Beohar, and S. K. Vishvakarma, "On-Chip Adaptive Body Bias for Reducing the Impact of NBTI on 6T SRAM Cells," *IEEE Transactions on Semiconductor Manufacturing,* vol. 31, pp. 242-249, 2018.

[37] S. V. Kumar, K. H. Kim, and S. S. Sapatnekar, "Impact of NBTI on SRAM read stability and design for reliability," in *7th International Symposium on Quality Electronic Design (ISQED'06)*, 2006, pp. 6 pp.-218.

[38] Y. Kunitake, T. Sato, and H. Yasuura, "Signal probability control for relieving NBTI in SRAM cells," in *2010 11th International Symposium on Quality Electronic Design (ISQED)*, 2010, pp. 660-666.

[39] T. Siddiqua and S. Gurumurthi, "Enhancing NBTI Recovery in SRAM Arrays Through Recovery Boosting," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 20, pp. 616-629, 2012.

[40] J. Shin, V. Zyuban, P. Bose, and T. M. Pinkston, "A Proactive Wearout Recovery Approach for Exploiting Microarchitectural Redundancy to Extend Cache SRAM Lifetime," in *2008 International Symposium on Computer Architecture*, 2008, pp. 353-362.

[41] S. Kothawade and K. Chakraborty, "Analysis and mitigation of BTI aging in register file: An application driven approach," *Microelectronics Reliability,* vol. 53, pp. 105-113, 2013/01/01/ 2013.

[42] Z. Xuan and S. K. Gupta, "Asymmetric sizing: An effective design approach for SRAM cells against BTI aging," in *2017 IEEE 35th VLSI Test Symposium (VTS)*, 2017, pp. 1-6.

[43] J. Abella, X. Vera, and A. Gonzalez, "Penelope: The NBTI-Aware Processor," in *40th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO 2007)*, 2007, pp. 85-96.

[44] M. Agostinelli, J. Hicks, J. Xu, B. Woolery, K. Mistry, K. Zhang, S. Jacobs, J. Jopling, W. Yang, B. Lee, T. Raz, M. Mehalel, P. Kolar, Y. Wang, J. Sandford, D. Pivin, C. Peterson, M. DiBattista, S. Pae, M. Jones, S. Johnson, and G. Subramanian, "Erratic fluctuations of sram cache vmin at the 90nm process technology node," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, 2005, pp. 655-658.

[45] J. T. Kao and A. P. Chandrakasan, "Dual-threshold voltage techniques for low-power digital circuits," *IEEE Journal of Solid-State Circuits,* vol. 35, pp. 1009-1018, 2000.

[46] P.-E. Gaillardon, E. Beigne, S. Lesecq, and G. D. Micheli, "A Survey on Low-Power Techniques with Emerging Technologies: From Devices to Systems," *J. Emerg. Technol. Comput. Syst.,* vol. 12, p. Article 12, 2015.

[47] A. Calimera, A. Macii, E. Macii, and M. Poncino, "Power-Gating for Leakage Control and Beyond," in *Circuit Design for Reliability*, R. Reis, Y. Cao, and G. Wirth, Eds., ed New York, NY: Springer New York, 2015, pp. 175-205.

[48] A. S. University. *Predictive Technology Model*. Available: http://ptm.asu.edu/

[49] H. M. Abbas, B. Halak, and M. Zwolinski, "BTI mitigation by anti-ageing software patterns," *Microelectronics Reliability,* vol. 79, pp. 79-90, 2017/12/01/ 2017.

[50] T. Jin and S. Wang, "Aging-Aware Instruction Cache Design by Duty Cycle Balancing," in *2012 IEEE Computer Society Annual Symposium on VLSI*, 2012, pp. 195-200.

[51] K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy, "Efficient Transistor-Level Sizing Technique under Temporal Performance Degradation due to NBTI," in *2006 International Conference on Computer Design*, 2006, pp. 216-221.

[52] B. C. Paul, K. Kunhyuk, H. Kufluoglu, M. A. Alam, and K. Roy, "Temporal Performance Degradation under NBTI: Estimation and Design for Improved Reliability of Nanoscale Circuits," in *Proceedings of the Design Automation & Test in Europe Conference*, 2006, pp. 1-6.

[53] M. H. Kaffashian, R. Lotfi, K. Mafinezhadand, and H. Mahmoodi, "An optimization method for NBTI-aware design of domino logic circuits in nano-scale CMOS," *IEICE Electronics Express,* vol. 8, pp. 1406-1411, 2011.

[54] S. Khan and S. Hamdioui, "Modeling and mitigating NBTI in nanoscale circuits," in *2011 IEEE 17th International On-Line Testing Symposium*, 2011, pp. 1-6.

[55] S. Joshi and S. Boyd, "An Efficient Method for Large-Scale Gate Sizing," *IEEE Transactions on Circuits and Systems I: Regular Papers,* vol. 55, pp. 2760-2773, 2008.

[56] X. Yang and K. Saluja, "Combating NBTI Degradation via Gate Sizing," in *8th International Symposium on Quality Electronic Design (ISQED'07)*, 2007, pp. 47-52.

[57] S. Kiamehr, F. Firouzi, M. Ebrahimi, and M. B. Tahoori, "Aging-aware standard cell library design," in *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2014, pp. 1-4.

[58] H. Amrouch, B. Khaleghi, A. Gerstlauer, and J. Henkel, "Reliability-aware design to suppress aging," in *2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2016, pp. 1-6.

[59] "SRAM Circuit Design and Operation," in *CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies: Process-Aware SRAM Design and Test*, A. Pavlov and M. Sachdev, Eds., ed Dordrecht: Springer Netherlands, 2008, pp. 13-38.

[60] E. Grossar, M. Stucchi, K. Maex, and W. Dehaene, "Read Stability and Write-Ability Analysis of SRAM Cells for Nanometer Technologies," *IEEE Journal of Solid-State Circuits,* vol. 41, pp. 2577-2588, 2006.

[61] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," *IEEE Journal of Solid-State Circuits,* vol. 22, pp. 748-754, 1987.

CHAPTER 5 CONCLUSION & PERSPECTIVES

In this work, we have presented a methodology for designing reliable integrated circuits on unreliable devices. The methodology consists of tackling the reliability issues at the early stages of the IC design flow.

As among CMOS ICs wear-out mechanisms, Negative Bias Temperature Instability has re-emerged, as the most dominant reliability issue in deep submicron technologies, we have focused our work on NBTI aware design of ICs. Accordingly, we have proposed an IC design for reliability approach that migrate and mitigate NBTI at circuit-level.

Our proposed DFR approach is twofold research objective: on one hand, it deals with migrating the analysis and understanding of NBTI from device- to circuit-level to better asses its impact on the circuit performance and reliability. On the other hands, it proposed a design technique to moderate that NBTI degradation and improve the circuit lifetime.

To consolidate our understanding of NBTI degradation, we have conducted a circuit-level NBTI analysis by characterizing the CMOS inverter under both DC and AC stresses. The results revealed that the degradation depends on the circuit workload rather than a simple prediction from analysis done on a single transistor. In fact, the shift of the inverter threshold voltage is found to be one order less than that of the pMOS threshold under DC NBTI stress. In addition, the shift of the inverter DC characteristics under AC NBTI stress is found to be much less than under DC NBTI, and more precisely much less than expected for a 0.5 duty-cycle NBTI stress.

Correlated to device-level NBTI, the obtained results of the circuit-level NBTI characterization have confirmed many aspects and revealed others. Actually, circuit-level NBTI analysis has confirmed voltage, temperature, and time dependence of the NBTI degradation. However, the time exponent has been shown to be both voltage and temperature dependent advising on the signature of the co-existence of more than one physical mechanism behind the degradation, and where the dominance of one mechanism over another depends on voltage, temperature and the stress time.

To provide more insight into the circuit degradation due to NBTI in a larger scale, we have, on one side, developed a reliability interface to simulate NBTI aging at circuit-level integrating our own NBTI model. On the other side, we have proposed a DFR approach to mitigate NBTI degradation in SRAMs. The proposed DFR approach is based on transistor sizing technique.

We showed that, contrary to what is thought, pMOS upsizing does not significantly alleviate NBTI degradation in the 6T-SRAM cell. As a matter of fact, we found upsizing pMOS beyond a certain value would be useless in terms of both reliability and area. However, properly sizing the nMOS access transistors would improve significantly the cell stability subjected to NBTI aging. Therefore, the 6T-SRAM design for NBTI aging would be re-thought as access transistor sizing for NBTI mitigation.

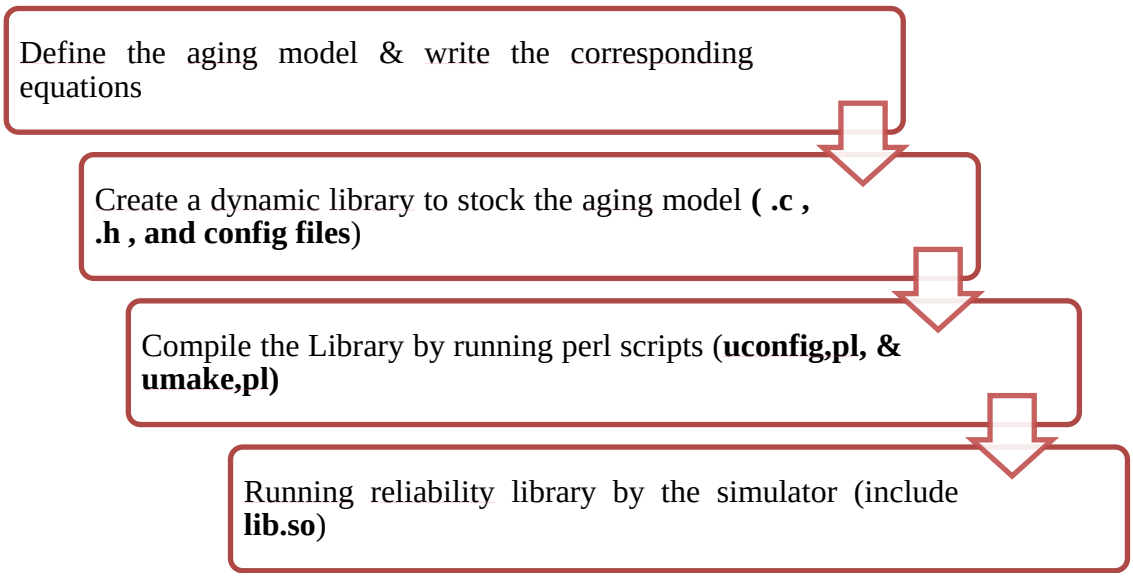
For future work, we intend to implement other NBTI models than power law At^n model, to figure out the implementation of NBTI recovery once the model and fit parameters would be available from experimental fast measurements. We also think about considering process variability alongside with NBTI to better deal with the stochastic nature of device parameters in ultra large scaled CMOS technologies. Moreover, we intend exploring the possibility of combining transistor sizing with other mitigation techniques to extend the proposed 6T-SRAM designs to the memory array.

Appendix 1

APPENDIX1: IMLPLEMENTATION OF NBTI MODEL ON CADENE ULTRASIM URI

The NBTI model is implemented in URI through a shared library **".so".** The creation of the library involves three essential steps:

1) Model equations Implementation **(.c** and **.h** files writing)
2) Configuration file Editing
3) Building and compiling the library



1 THE IMPLEMENTED NBTI MODEL

$$\Delta v_{th} = A_0 . exp\left(\frac{E_{ox}}{E_0}\right) exp\left(-\frac{E_a}{KT}\right)\ t^n \quad \text{.....(1)}$$

A_0, E_0, Ea, et n have been extracted from OTFOT characterization of pMOS devices belonging to BTS_Structures chip

$$0.92 \leq E_0 \leq 1.34\ \text{MV/cm}\quad 0.21 \leq E_a \leq 0.25eV \text{ et} \quad 0.20 \leq n \leq 0.3$$

MODEL
Written files: **pchVthDef.c** and **pchVthDef.h** were witten to declare `PCHVTH. Data strutures and the model` equation.l
In addition of config file

Appendix 1

pchVthDef.c

```
//
************************************************************************
******
// Copyright 2014-2015 CDTA/DMN/FCS/BTS_Project/A.Chenouf
// All rights reserved.
//
// This is version 1 of Unified Reliability Interface for NBTI Aging
Analysis
// It is based on the template provided by Cadence Design Systems for HCI
//
//
//
************************************************************************
******

/*****************************
 NBTI Model implementation on URI:
 *****************************/
#include "uri.h"
#include "pchVthDef.h"
#include <string.h>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>

/*********************************************************************

This example illustrates the interface for NBTI reliability model
implementation.

In this example, it is assumed that the Degradation of Vth follows a power
law of stress time:
dvth = A.exp(Eox/E0).exp(-Ea/KT).t^n
As Virtuoso RelXpert reliability simulator calculates the time dependence
of degradation using Degradation = (age.t)^n, then dvth can be rewritten as
:

Degradation= (A0 exp(-Eox/E').exp (Ea'/KT).t)^n
lifetime = (dD)^(1/n) A^-1= dD^1/n. (A^-1. exp(-Eox/E').exp (Ea'/KT))
where
      Eox = Oxyde field =(Vgs-Vth)/Tox
      Vgs = bias Voltage
      Ea = Energy Activation
      t =  time for which degradation should be predicted
      dD = degradation criterion
      T = Temperature at which the degardation is estimated
      K = Boltzman Constant
      A = Process constant
      Ea' = Ea/n
      E'= nE0
      A0 =A^(1/n)
Reliability model parameters: A, E0, Ea, n, T

*********************************************************************/
```

Appendix 1

```
/*
 * this function sets up the reliability parameters needed and the default
 * values for the reliability parameters
 * returns 0 if successful
 */
int pchVthSetupModel( URI_Model *model, const char *const modelType,
                      const int isDebug )
{

    PCHVTH *mymodel = ( PCHVTH * )model;
    mymodel->Ea = 0.21; /* Ea */
    mymodel->n = 0.3; /* n */
    mymodel->E0 = 1.13;
    mymodel->A = 1.8e-1;
    mymodel->nd1_vth0 = 1.0;
    mymodel->nn1_vth0 = 1.0;
    if ( isDebug )
        printf( "pchVthSetupModel done\n" );
    return 0;
}

/*
 * this function assigns the model parameter values from the netlist
 * returns 0 if successful
 */
int pchVthAddParam( URI_Model *model, const char *const name,
                    const double value, const char *strvalue, const int
isDebug )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    if ( isDebug )
        printf( "pchVthAddParam ..." );
    if ( strcasecmp( name, "Ea" ) == 0 ) {
        mymodel->Ea = value;
        if ( isDebug )
            printf( " parameter Ea=%f assigned\n", mymodel->Ea );
        return 0;
    } else if ( strcasecmp( name, "N" ) == 0 ) {
        mymodel->n = value;
        if ( isDebug )
            printf( " parameter N=%f assigned\n", mymodel->n );
        return 0;
    } else if ( strcasecmp( name, "E0" ) == 0 ) {
        mymodel->E0 = value;
        if ( isDebug )
            printf( " parameter E0=%f assigned\n", mymodel->E0 );
        return 0;
    } else if ( strcasecmp( name, "A" ) == 0 ) {
        mymodel->A = value;
        if ( isDebug )
            printf( " parameter A=%f assigned\n", mymodel->A);
        return 0;

    } else if ( strcasecmp( name, "ND1_VTH0" ) == 0 ) {
        mymodel->nd1_vth0 = value;
        if ( isDebug )
            printf( " parameter ND1_VTH0=%f assigned\n", mymodel->nd1_vth0
);
        return 0;
```

Appendix 1

```
    } else if ( strcasecmp( name, "NN1_VTH0" ) == 0 ) {
        mymodel->nn1_vth0 = value;
        if ( isDebug )
            printf( " parameter NN1_VTH0=%f assigned\n", mymodel->nn1_vth0
);
        return 0;
    } else if ( strcasecmp( name, "STRING_ARG" ) == 0 ) {
        if ( strlen( strvalue ) < sizeof( mymodel->str_value ) ) {
            strcpy( mymodel->str_value, strvalue );
            if ( isDebug )
                printf( " parameter STRING_ARG=%s assigned\n", mymodel-
>str_value );
            return 0;
        } else {
            printf( "pchVthAddParam: STRING_ARG length (%d) exceeds the
maximum (%d)\n",
                    strlen( strvalue ), sizeof( mymodel->str_value ) );
            return 1;
        }
    } else if ( strcasecmp( name, "STRING_ARG1" ) == 0 ) {
        if ( strlen( strvalue ) < sizeof( mymodel->str_value1 ) ) {
            strcpy( mymodel->str_value1, strvalue );
            if ( isDebug )
                printf( "parameter STRING_ARG1=%s assigned\n", mymodel-
>str_value1 );
            return 0;
        } else {
            printf( "pchVthAddParam: STRING_ARG1 length (%l) exceeds the
maximum (%d)\n",
                    strlen( strvalue ), sizeof( mymodel->str_value1 ) );
            return 1;
        }
    }

    if ( isDebug )
        printf( " failed\n" );
    return 1; // return 1 if failed
}


/*This function is used to add vector model parameters into URI model.
*/
int pchVthAddVectorParam( URI_Model *model, const char *const name, const
double *value,
                          const int size, const int isDebug )
{
    int i = 0;
    PCHVTH *mymodel = ( PCHVTH * )model;
    if ( isDebug )
        printf( "pchVthAddVectorParam ... \n" );
    if ( !strcasecmp( name, "vth0_table" ) && size > 0 ) {
        mymodel->vth0_table = ( double * )malloc( sizeof( double ) * size
);
        if ( isDebug )
            printf( "vector parameter %s = ", name );
        for ( i = 0; i < size; i++ ) {
            mymodel->vth0_table[i] = value[i];
            if ( isDebug )
                printf( "%f, ", mymodel->vth0_table[i] );
```

Appendix 1

```
        }
        if ( isDebug )
            printf( "assgined\n" );
        return 0;
    }
    if ( isDebug )
        printf( "failed\n" );
    return 1;

}

/*
 * This function is used to free the memory allocated to save vector
parameters.
 * */
int pchVthDeleteModel( URI_Model *model,  URI_sharedInstState *st1,
URI_InstState *st2,
                       const int isDebug )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    if ( isDebug )
        printf( "pchVthDeleteModel ...\n" );
    if ( mymodel->vth0_table ) {
        free( mymodel->vth0_table );
        mymodel->vth0_table = 0;
    }
    return 0;

}

/*
 * this function calculates the age rate value
 * returns age rate value if successful
 */
double pchVthCalcA( URI_Model *model, const URI_Var *const uriVar,
                    const URI_Env *const env, const int isDebug,
                    URI_sharedInstState *st1, URI_InstState *st2 )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    URI_MosVar *uriMosVar = ( URI_MosVar * )uriVar;
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    double ageRate = 0.0;
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;
    double tox, Eox, vt, vgs, T, temp, E0, Ea, E01, Ea1, n, A, k1 ,k2, k3,
invn ;
  if ( uriMosVar->vgs == 0.0 )
        ageRate = 0;

    temp = env->temp;
    vgs = n  = mymodel->n;
    Ea = mymodel->Ea;
    E0 = mymodel->E0;
    A = mymodel->A;
    E01 = E0 * n;
    Ea1 = Ea/n;
    vgs = uriMosVar->vgs;
    vt = uriMosVar->von;
    tox  = uriMosVar->tox;
    Eox = -(vgs + vt)*1e-8/tox;
```

Appendix 1

```
        if (vgs == 0.0)
       k2 = 0;
   else
      k2 = exp(Eox/E01);

   T = temp + 273.15;                        /*  C to K temperature
conversion */
   k3 = exp(-Ea1 / (K * T));
   invn = 1.0/n;
   k1= pow(A, invn);

   ageRate = k1 * k2 * k3;
   /* printf ( "\n ageRate =%e \n", ageRate); */

   s2->agerate = ageRate;

    if ( isDebug )
     printf ( "\n Temperature T = %f \n", T);
     printf ( "\n Vgs =%e \n", uriMosVar->vgs);
     printf ( "\n Oxide thikness Tox =%e \n", uriMosVar->tox);
     printf ( "\n threshold voltage vth =%e \n", uriMosVar->von);
     printf ( "\n Oxide field Eox=%e \n", Eox);
     /*printf ( "\n n =%f \n ", n);
     printf ( "\n E01 =%f \n", E01);
     printf ( "\n Ea1 =%f \n ", Ea1);
     printf ( "\n A =%e \n ", A);
     printf ( "\n invn =%e \n", invn);*/
     printf ( "\n 1st factor =%e \n", k1 );
     printf ( "\n 2nd factor =%e \n", k2);
     printf ( "\n 3rd factor =%e \n", k3);
     printf ( "\n ageRate =%e \n", ageRate);
     printf ( "\n pchVthCalcA done\n" );

    return ageRate;

}


/*
 * this function calculates the age value
 * returns age value if successful
 */
double pchVthCalcDeltaD( URI_Model *model, const URI_Var *const uriVar,
                         const int isDebug, URI_sharedInstState *st1,
URI_InstState *st2,
                         const URI_Env *const env )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    URI_MosVar *uriMosVar = ( URI_MosVar * )uriVar;

    double deltaD = s2->deltaD;
    double n = mymodel->n;
    double tmpDeg = s2->agerate * uriMosVar->dt;
    if ( deltaD > 0.0 )
        deltaD = pow( pow( deltaD, 1.0 / n ) + tmpDeg, n );
    else
        deltaD = pow( tmpDeg, n );
```

Appendix 1

```
    s2->deltaD = deltaD;
    if ( isDebug )
        printf ("\n temporary Deg =%e \n", tmpDeg);
        printf ("\n deltaD =%e \n", deltaD);

        printf( "\n pchVthCalcDeltaD done\n" );
    return deltaD;
}

/*
 * this function calculates the total age value up to the simulation time
 * returns the total age value
 */
double pchVthCalcDeltaAge( URI_Model *model, const URI_Var *const uriVar,
                           const int isDebug, URI_sharedInstState *st1,
URI_InstState *st2,
                           const URI_Env *const env )
{
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    URI_MosVar *uriMosVar = ( URI_MosVar * )uriVar;

    double dage = 0.5 * ( s2->prevage + s2->agerate ) * uriMosVar->dt;
    s2->prevage = s2->agerate;
    s2->totage += dage;
    if ( isDebug )

    printf("\n DeltaAge =%e \n", dage);
    printf("\n totage =%e \n", s2->totage);

    printf( "\n chVthCalcDeltaAge done\n" );
return s2->totage;
}

/*
 * this function calculates the final age value for the times specified
 * in .age.
 * returns the final total age value
 */
double pchVthCalcAge( URI_Model *model, const double futuretime,
                      const int isDebug, URI_sharedInstState *st1,
URI_InstState *st2,
                      const URI_Env *const env )
{
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;
    PCHVTH *mymodel = ( PCHVTH * )model;

    double deg;

    if ( isDebug )
        printf( "Starting CalcAge(): ( dev: %s; mod: %s )\n", s1->devName,
s1->modName );

    if ( env == 0 )
        return 0;

  double ftime = 0.0;
```

Appendix 1

```
    if ( env->futureTime )
        ftime = *env->futureTime;
    else
        ftime = futuretime;

    deg = s2->totage * ftime / ( env->tStop - env->tStart );

    if ( isDebug )

        printf("\n Final age =%e  \n", deg);
        printf( "pchVthCalcAge done\n" );
    return deg;

}

/*
 * this function calculates the final degradation value for the times
 * specified in .age.
 * returns the final degradation value
 */
double pchVthCalcDegradation( URI_Model *model, const URI_Var *const
uriVar,
                              const double futuretime, const int isDebug,
                              URI_sharedInstState *st1, URI_InstState
*st2,
                              const URI_Env *const env )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;
    double deg;

    if ( isDebug )
        printf( "Starting CalcDegradation(): ( dev: %s; mod: %s )\n", s1-
>devName, s1->modName );

    deg = s2->deltaD * pow( futuretime / ( env->tStop - env->tStart ),
mymodel->n );
    if ( isDebug )

        printf("\n Final degraddation =%e  \n", deg);
        printf( "pchVthCalcDegradation done\n" );
    return deg;

}

/*
 * this function calculates the final lifetime value for degradation
 * criterion specified in .degrad statement
 * returns the final lifetime value
 */
double pchVthCalcLifetime( URI_Model *model, const double age,
                           const double degAtLifetime, const int isDebug,
                           URI_sharedInstState *st1, URI_InstState *st2,
                           const URI_Env *const env )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;
```

Appendix 1

```
    double lt;

    if ( isDebug )
        printf( "Starting CalcLifeTime(): ( dev: %s; mod: %s )\n", s1-
>devName, s1->modName );

    lt = ( env->tStop - env->tStart ) * pow( degAtLifetime / s2->deltaD,
1.0 / mymodel->n );
    if ( isDebug )
    printf ("\n degAtLifetime =%e \n", degAtLifetime);
    printf ("\n lifetime =%e \n", lt);
     printf( "pchVthCalcLifetime done\n" );

    return lt;
}

/*
 * this function returns the degraded spice parameter value
 *
 */
double pchVthScaleParam( URI_Model *model,  double age,
                         const char *const name, const double value, const
int isDebug,
                         const URI_Var *const uriVar, const
URI_sharedInstState *const st1,
                         const URI_Env *const env, const URI_InstState
*const st2 )
{
    PCHVTH *mymodel = ( PCHVTH * )model;
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;

    if ( isDebug )
        printf( "Starting ScaleParam(): %s ( dev: %s; mod: %s )\n", name,
s1->devName, s1->modName );

    if ( !strcasecmp( name, "VTH0" ) )
     return ( value * pow( ( age * mymodel->nd1_vth0 ), mymodel->nn1_vth0 )
);
    else
        return value;
}

/*
 * this function adds all instance parameters into URI states, such as l,
w, m.
 * NOTE: All MOS device parameters can be added into URI states with this
func.
 *       but for external users, the argu "name" must be the same as the
param
 *       name in device definition, such as ad, as, w, l ...
 */
int pchVthAddInstParam( URI_Model *model, URI_sharedInstState *st1,
URI_InstState *st2, const char *const name, const double value, const char
*strvalue )
{
    PCHVTHSHAREDINSTSTATE *s1 = ( PCHVTHSHAREDINSTSTATE * )st1;
    // PCHVTHINSTSTATE *s2 = ( PCHVTHINSTSTATE * )st2;
    if ( s1 != 0 ) {
        if ( strcmp( name, "l" ) == 0 ) {
```

Appendix 1

```
            s1->l = value;
            return 0;
        } else if ( strcmp( name, "w" ) == 0 ) {
            s1->w = value;
            return 0;
        } else if ( strcmp( name, "mult" ) == 0 ) {
            s1->mult = value;
            return 0;
        } else if ( strcmp( name, "device_name" ) == 0 ) {
            strcpy( s1->devName, strvalue );
        } else if ( strcmp( name, "model_name" ) == 0 ) {
            strcpy( s1->modName, strvalue );

        }
// TODO: add other instance parameters, such as ad, as ...
    }
    return 1;
}
int pchVthgetNewModelName(URI_Model *model, const char *name, char
*newname)
{

printf(newname, "%  aged", name);
return 0;

}
static PCHVTH pchVthMod;

static URI_ModelDef pchVthModDef = {
    ( URI_Model * ) &pchVthMod,
    103,
    sizeof( PCHVTH ),
    uriModelParams,
    sizeof( uriModelParams ) / sizeof( uriModelParams[0] ),
    pchVthSetupModel,
    pchVthAddParam,
    NULL,  // pchVthInitAllParams
    pchVthCalcA,
    pchVthCalcDegradation,
    pchVthCalcLifetime,
    pchVthScaleParam,
    pchVthCalcAge,
    pchVthCalcDeltaD,
    pchVthCalcDeltaAge,
    pchVthAddVectorParam,
    pchVthDeleteModel,
    pchVthAddInstParam,
    NULL,  // pchVthcopyInstState,
    NULL,  // pchVthGetModelParam,
    NULL,  // pchVthGetModelVectorParam,
    NULL,  // pchVthGetInstParam,
    NULL,  // pchVthGetInstVectorParam,
    sizeof( PCHVTHSHAREDINSTSTATE ),
    sizeof( PCHVTHINSTSTATE ),

};

URI_ModelDef *pchVthDefPtr = &pchVthModDef;
```

Appendix 1

pchVthDef.h

```
//
*****************************************************************************
// Copyright 2011-2011 Cadence Design Systems, Inc.
// All rights reserved.
//
// This work is protected by copyright laws and contains Cadence
proprietary
// and confidential information.  No part of this file may be reproduced,
// modified, re-published, used, disclosed or distributed in any way, in
any
// medium, whether in whole or in part, without prior written permission
from
// Cadence Design Systems, Inc. (Cadence).  If you are a customer of
Cadence,
// such permission may be found in a previously executed license and
// maintenance agreement between authorized representatives of you and
Cadence.
//
// Notwithstanding any previous agreement, written or otherwise, this API
is
// provided to you AS IS with no warranties of any kind, and you agree to
bear
// the entire risk of any customizations you make using this API.
//
*****************************************************************************


#ifndef _PCHVTHDEF_H_
#define _PCHVTHDEF_H_
#define  K 8.6173324e-5               /* eV/K */

typedef struct pchVth {
    /* user defined model goes here */
    double A;
    double E0;
    double Ea;
    double n;
    double T;
    double nd1_vth0;
    double nn1_vth0;
    double *vth0_table;
    char str_value[1024];
    char str_value1[1024];
} PCHVTH;


/* Defined the Spectre native reliability
 * analysis data type. This data structure
 * is only used by Spectre. */
static URI_Param uriModelParams[] = {
    {"A", uriRealType},
    {"E0", uriRealType},
    {"Ea", uriRealType},
```

Appendix 1

```
    {"n", uriRealType},
    {"nd1_vth0", uriRealType},
    {"nn1_vth0", uriRealType},
    {"vth0_table", uriVectorType},
    {"str_value", uriStringType},
    {"str_value1", uriStringType}
};


/* variables passed among uri routines */

typedef struct pchVthSharedInstState {
    /* instance parameters related to geometry */
    double mult;
    double l;
    double w;
    char   devName[1024];
    char   modName[1024];
} PCHVTHSHAREDINSTSTATE;

typedef struct pchVthInstState {
    /* instance parameters related to bias */
    double prevage;
    double agerate;
    double totage;
    double deltaD;
} PCHVTHINSTSTATE;

#endif
```

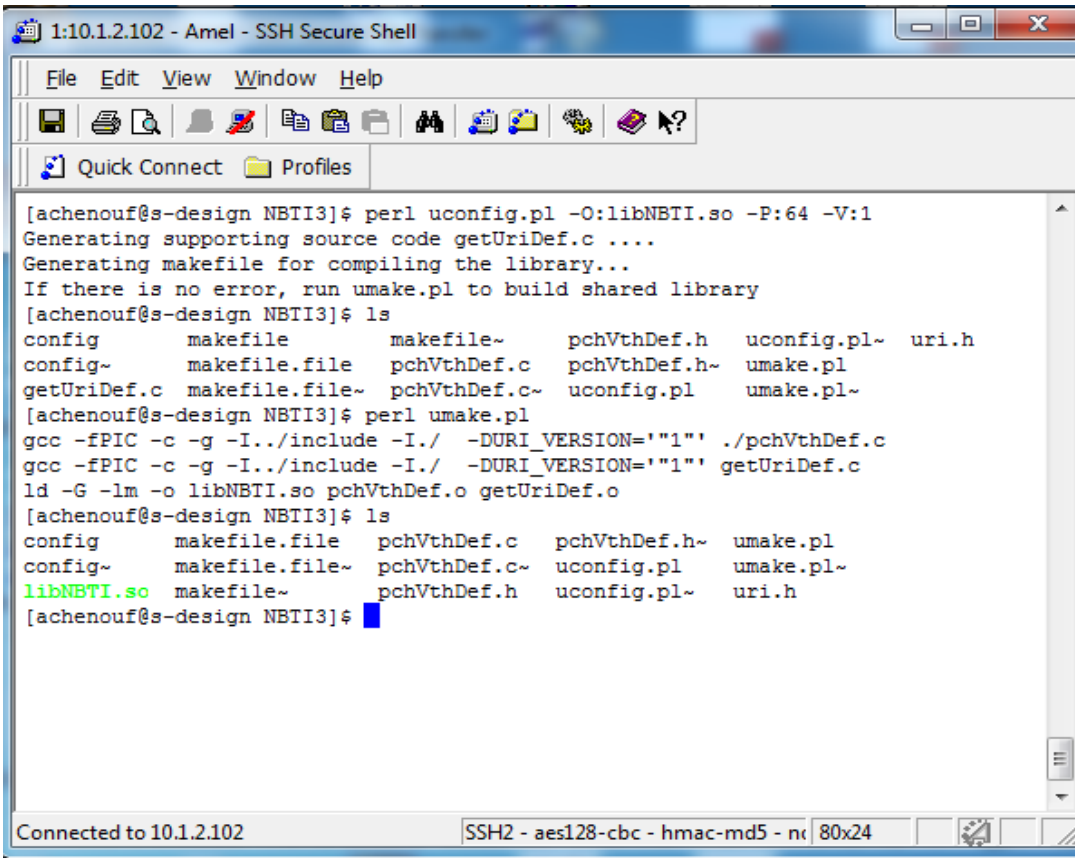
The config file includes these two lines

```
*type level name
  age 103 pchVth
```

The level should be set to a value **above 100,** as those <100 are reserved for Cadence

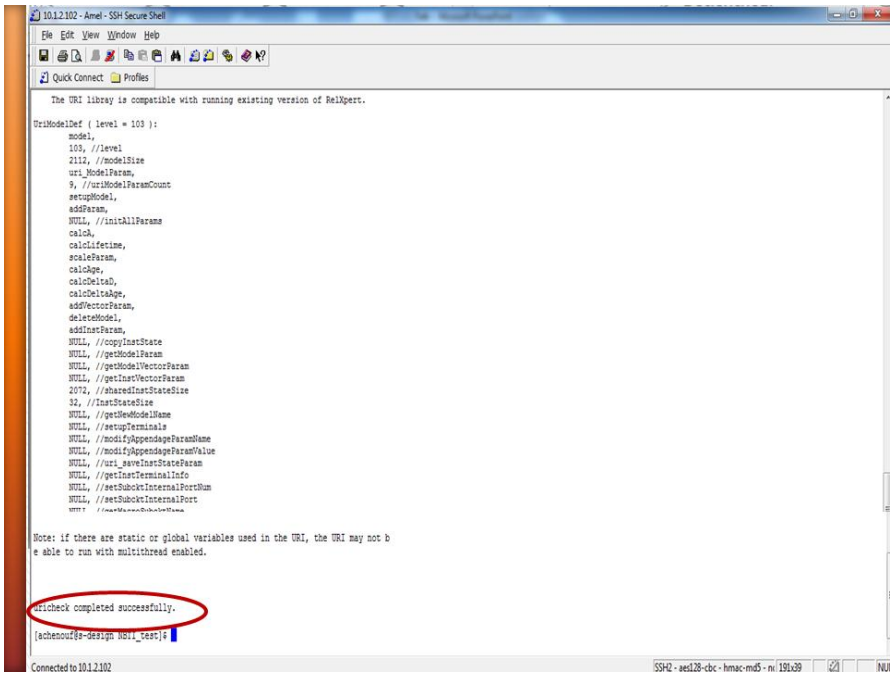
Cadence provides Perl scripts: **uonfig.pl** and **umake.pl** compile and build the shared library libNBTI.so as shown bellow. Then we checked the built library

Appendix 1

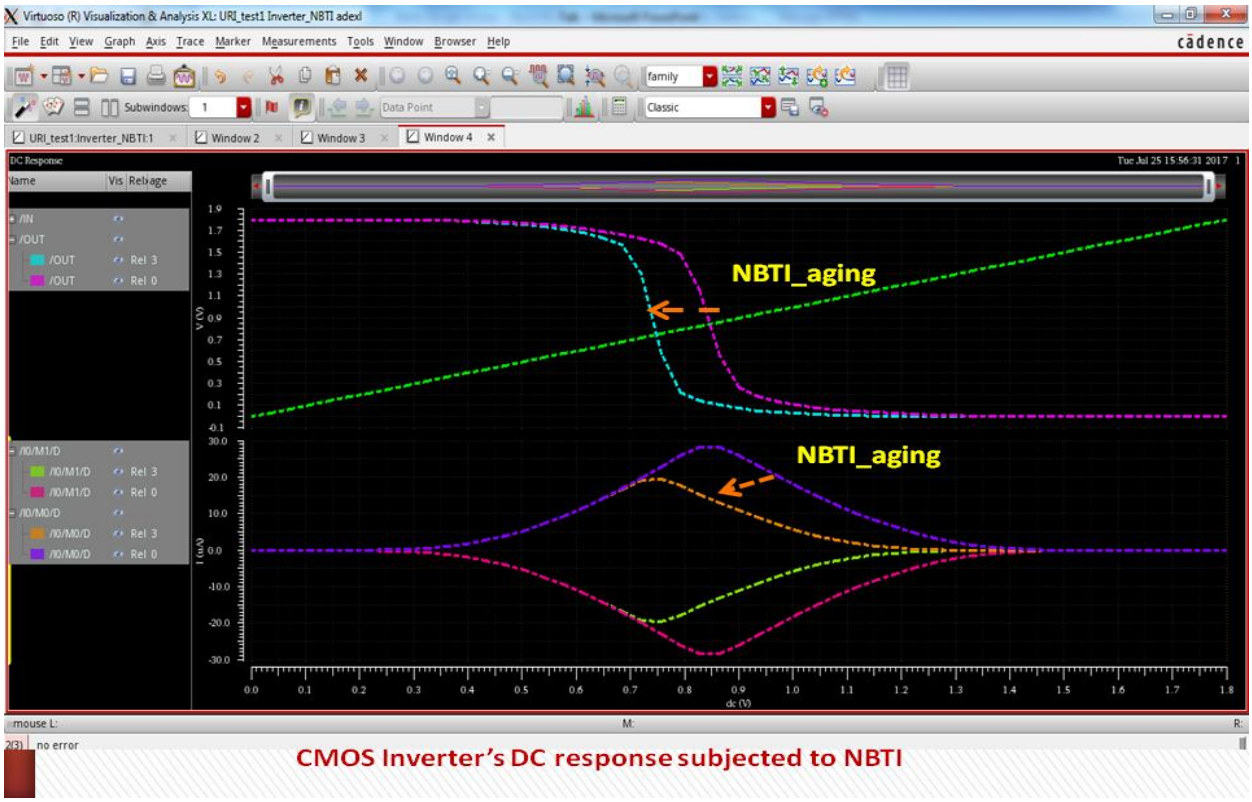


Shared library compilation and building

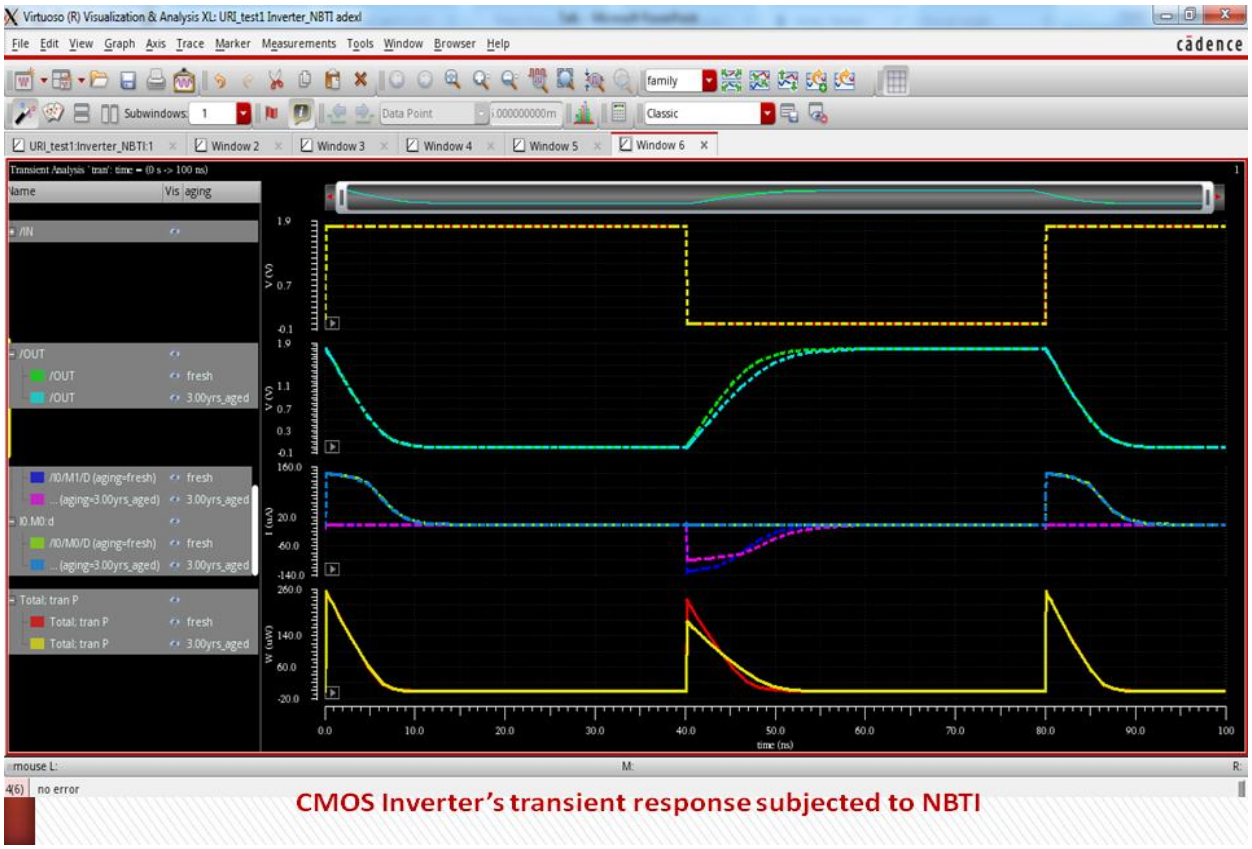
Runnig uricheck libNBTI to verify compatibility of the model with reliability simulator



Appendix 2



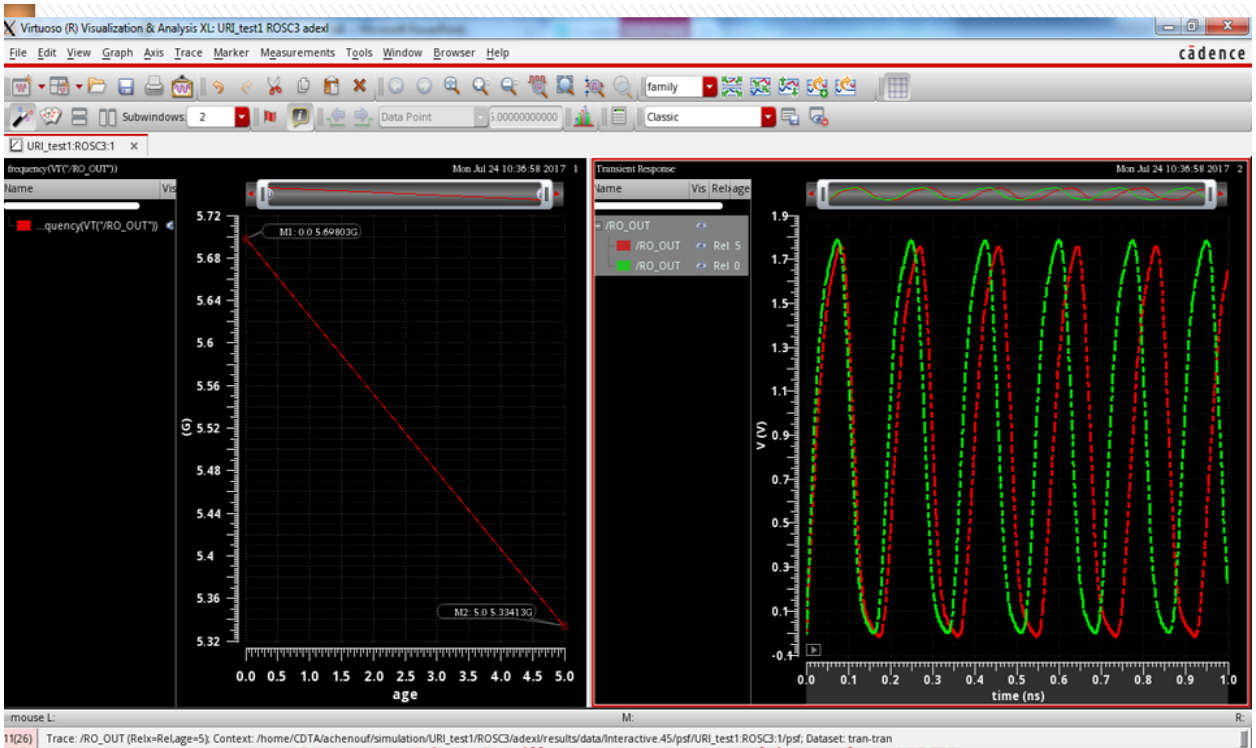
CMOS Inverter's DC response subjected to NBTI



CMOS Inverter's transient response subjected to NBTI

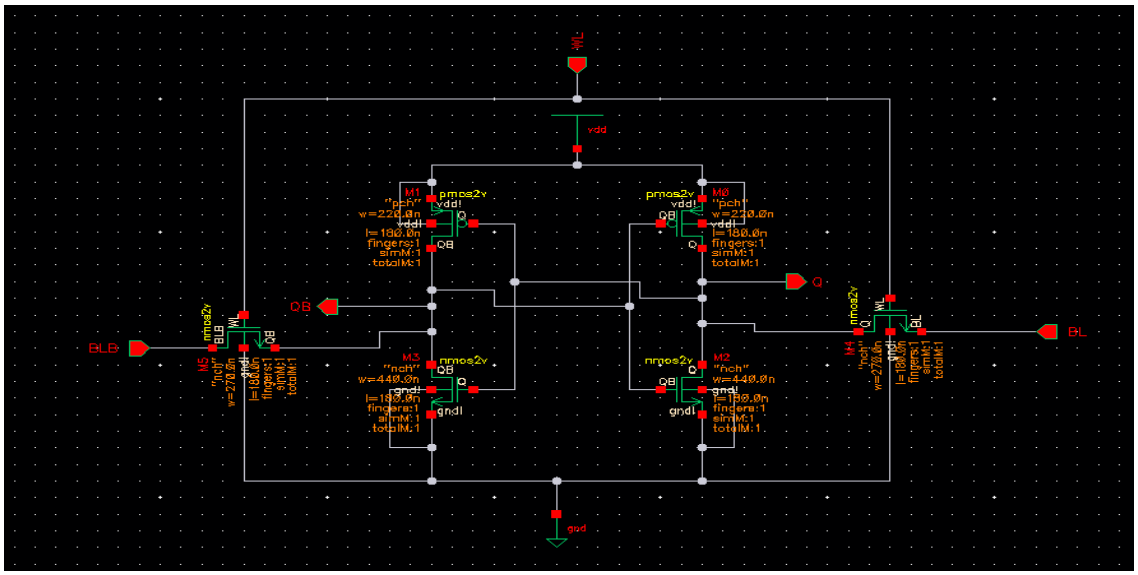
Appendix 2

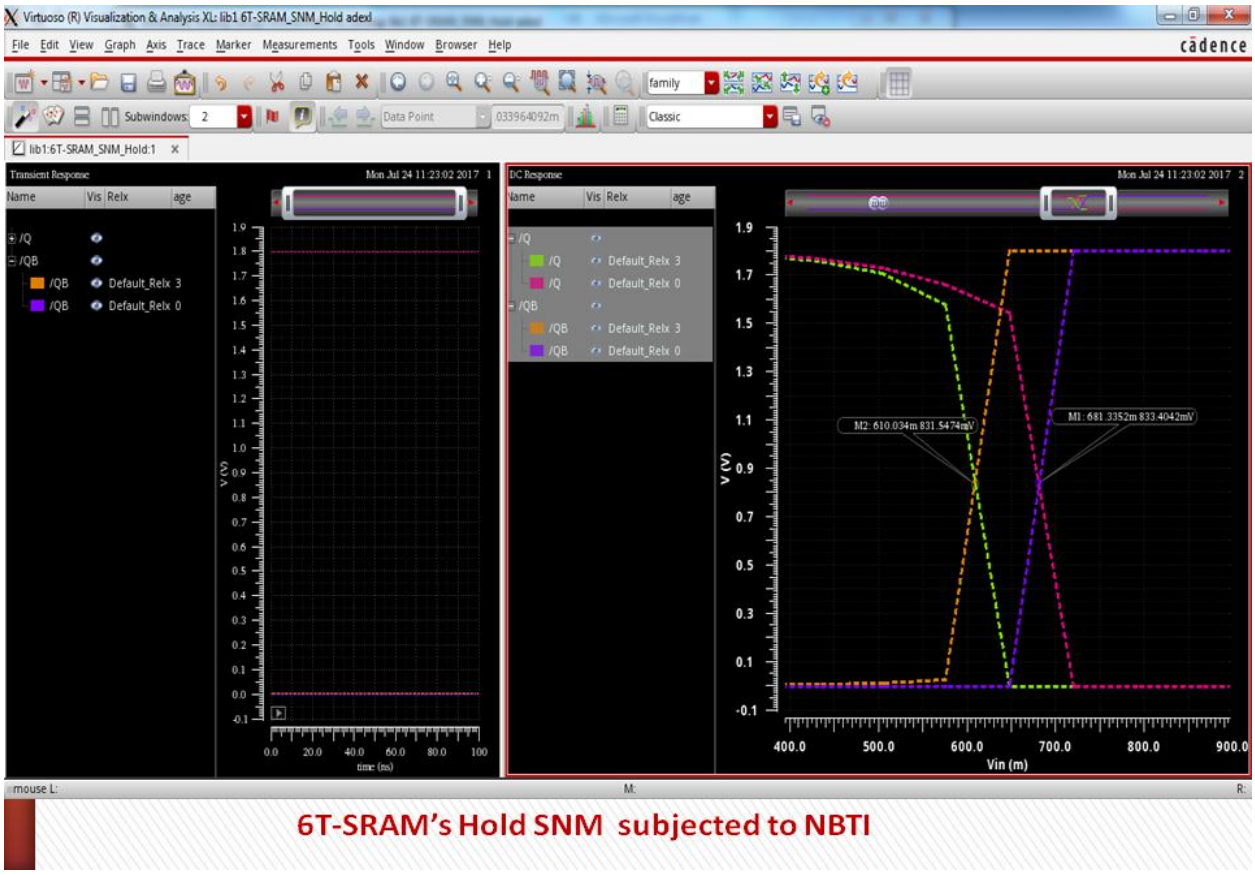
Inverter response under NBTI stress with different duty-cycle



3-stages-Ring Oscillator response subjected to NBTI

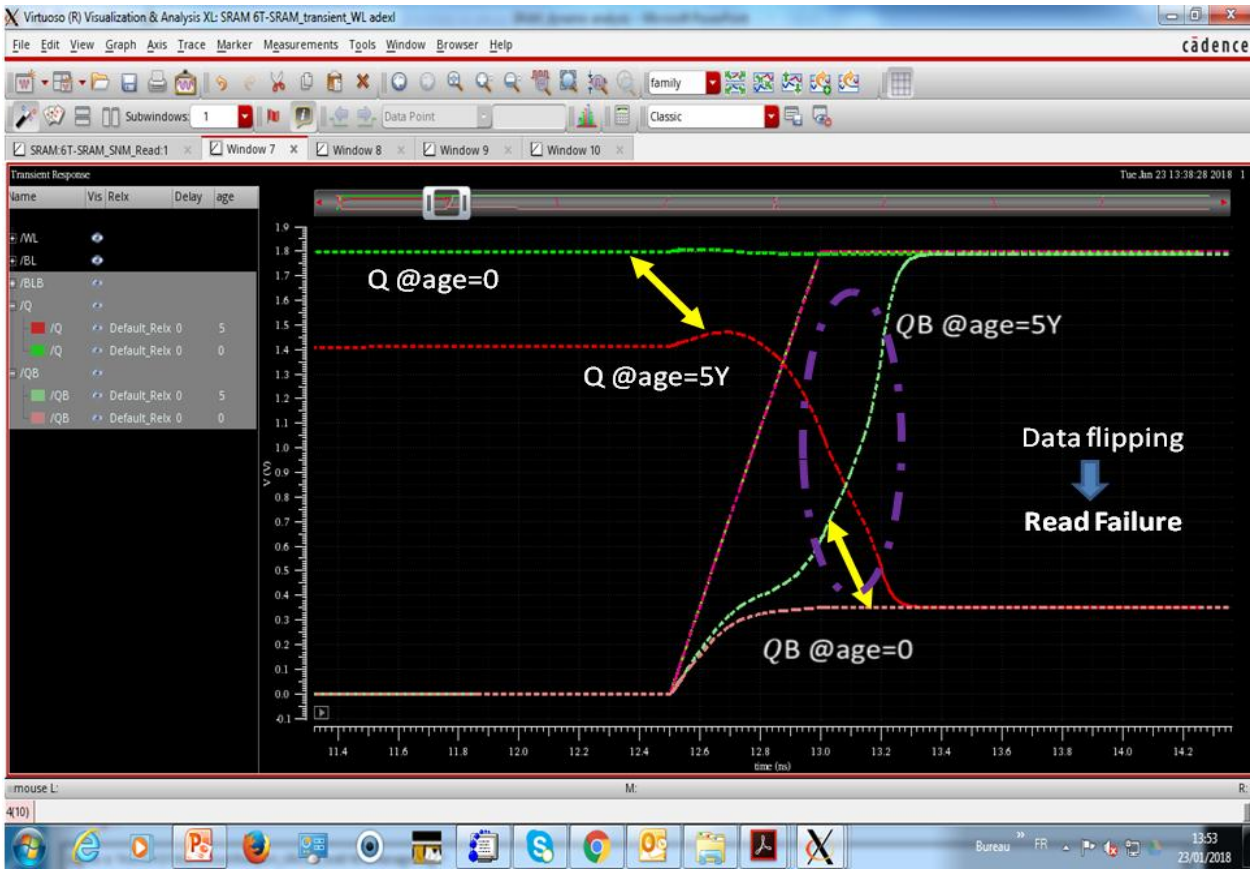
Appendix 2

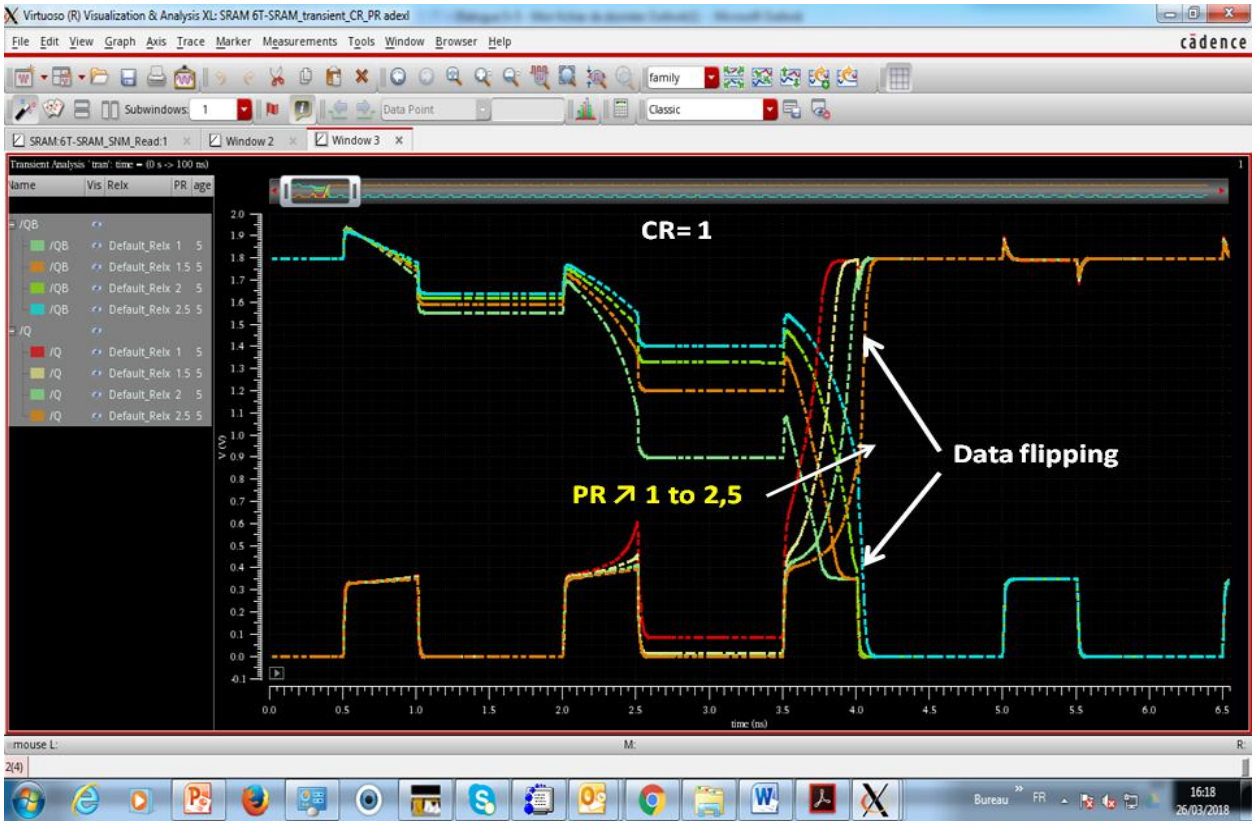




6T-SRAM's Hold SNM subjected to NBTI

Appendix 2


Virtuoso (R) Visualization & Analysis XL: SRAM 6T-SRAM_transient_WL adexl
Q @age=0
Q @age=5Y
QB @age=5Y
QB @age=0
Data flipping
Read Failure


Virtuoso (R) Visualization & Analysis XL: SRAM 6T-SRAM_transient_CR_PR adexl
CR= 1
PR ↗ 1 to 2,5
Data flipping