Multilevel aging phenomena analysis in complex ultimate CMOS designs

Dolly Natalia Ruiz Amador Ruiz Amador

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Nanoélectronique et Nanotechnologies**

Arrêté ministériel : 7 août 2006

Présentée par

Natalia RUIZ

Thèse dirigée par **Lorena ANGHEL** et
codirigée par **Vincent HUARD**

préparée au sein du **Laboratoire de Techniques de l'informatique et de la Microélectronique pour l'Architecture des systèmes intégrés (TIMA)**.

dans **l'École Doctorale d'Electronique, Electrotechnique, Automatique & Traitement du Signal (EEATS)**

Multilevel aging phenomena analysis in complex ultimate CMOS designs

Thèse soutenue publiquement le **01 février 2012,**
devant le jury composé de :

M. Michael NICOLAIDIS
Chargée de recherche au CNRS-TIMA, Président
M. François MARC
Maitre de Conférences à l'Université de Bordeaux I, Rapporteur
M. Alain BRAVAIX
Professeur à l'ISEN de Toulon, Rapporteur
Mme. Nadine BUARD
Docteur ingénieur à EADS (Suresnes), Examinatrice
M. Vincent HUARD
Docteur ingénieur à STMicroelectronics (Crolles), Invité

This thesis is dedicated to my family for their love and support

Abstract

Integrated circuits evolution is driven by the trend of increasing operating frequencies and downscaling of the device size, while embedding more and more complex functionalities in a single chip. However, the continuation of the device-scaling race generates a number of technology challenges. For instance, the downscaling of transistor channel lengths induce short-channel effects (drain-induced barrier lowering and punch-through phenomena); high electric field in the devices tend to increase Hot electron effect (or Hot Carrier) and Oxide Dielectric Breakdown; higher temperatures in IC products generates an increase of the Negative Bias Temperature Instability (NBTI) effect on pMOS devices. Today, it is considered that the above reliability mechanisms are ones of the main causes of circuit degradation performance in the field.

This dissertation will address the Hot Carrier (HC) and NBTI impacts on CMOS product electrical performances. A CAD bottom-up approach will be proposed and analyzed, based on the Design–in Reliability (DiR) methodology. With this purpose, a detailed analysis of the

NBTI and the HC behaviours and their impact at different abstraction level is provided throughout this thesis.

First, a physical framework presenting the NBTI and the HC mechanisms is given, focusing on electrical parameters weakening of nMOS and pMOS transistors. Moreover, the main analytical HC and NBTI degradation models are treated in details. In the second part, the delay degradation of digital standard cells due to NBTI, HCI is shown; an in-depth electrical CAD analysis illustrates the combined effects of design parameters and HCI/NBTI on the timing performance of standard cells. Additionally, a gate level approach is developed, in which HC and NBTI mechanisms are individually addressed. The consequences of the degradation at system level are presented in the third part of the thesis. With this objective, data extracted from silicon measures are compared against CAD estimations on two complexes IPs fabricated on STCMOS 45nm technologies.

It is expected that the findings of this thesis highly contribute to the understanding of the NBTI and HC reliability wearout mechanisms at the system level.

Acknowledgement

I would like to express my gratitude to Ing. Vincent Huard and Prof. Lorena Anghel, my thesis advisors. They not only gave me important guidance, but also allowed me to explore various approaches and topics freely during various stages of this project.

I am indebted to Prof. François Marc to Prof. Alain Bravaix to Ing. Nadine Buard and Ing. Michael Nicolaidis to spend their valuable time in serving in my thesis committee.

I am grateful to Ing. Xavier Federspiel, Ing. Emmanuel Pion, Ing. Florian Cacho, Ing. Emmanuel Vincent, Ing. Remy Chevallier, Ing. Sylvain Landelle, Ing. Pierre F. Ollagnon, Ing. Sylavin Engels, Ing. Maryline Parra, Ing. Manuel Sellier and Ing. Etienne Maurin of STMicroelectronics. They have provided great help on this project.

TABLE OF CONTENTS

List of Acronyms

AC	Alternating current
AO	AND-OR gate
AOI	AND-OR-INVERTER gate
BFX	Buffer gate
BSIM	Berkeley Short-channel IGFET model for MOS transistor
CAD	Computer-aided design
CCC	Channel Cold Carrier
CHC	Channel Hot Carrier
CHE	Channel Hot Electron
CLK	Clock
CMOS	Complementary Metal-Oxide-Semiconductor
CPU	Central Processing Unit
DAHC	Drain Avalanche Hot Carrier
DC	DC Direct Current
DEE	Direct Electron Excitation
DiR	Design-in Reliability
DUT	Device Under Test
EDM	Energy Driven Model
EEDF	Electron Energy Distribution Function
EES	Electron-Electron Scattering
EFL	Early Fails Failure
EFR	Early Failure Rate
ELA	Equivalent Life Active
EM	Electromigration
ESD	Electrostatic Discharge
F-F	Flip-Flop
FO	Fan-out
HC	Hot Carrier
HCI	Hot Carrier Injection
HL	High-to-Low transition
HTOL	High Temperature Operating Life
IC	Integrated Circuit
IP	Intellectual Property (database)
ITRS	International Technology Roadmap for Semiconductors
IV	Input Inverter
IVX	Inverter gate
JEDEC	Joint Electron Device Engineering Council
LDPC	Low-Density Parity-Check
LH	Low-to-High transition
LEM	Lucky Electron Model
L-R	Limited Reaction
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor

MVE	Multiple Vibrational Excitation
NAND	Not AND gate
nMOSFET	*n*-type MOSFET
NBTI	Negative Bias Temperature Instabilities
NOR	Not OR gate
OLT	Operational Life Test
OTF	On-The-Fly
OV	Output Inverter
pMOSFET	*p*-type MOSFET
PT	Primetime
POT	Power-On-Time
PVT	Process Voltage Temperature
R-D	Reaction-Diffusion
RO	Ring Oscillator
SC	Standard Cell
SCE	Short Channel Effect
SGHC	Secondary Generated Hot Carrier
SHC	Substrate Hot Carrier
SHE	Substrate Hot Electron
SP	Signal Probability
SPICE	Simulation Program with Integrated Circuit Emphasis
SRAM	Static Random Access Memory
SRH	Shockley Read Hall
SVE	Single Vibrational Excitation
STA	Static Timing Analysis
STM	STMicroelectronics
TDDB	Time Dependent Dielectric Breakdown
TTF	Time-To-Failure
UDRM	User Define Reliability Model
XNOR	Exclusive OR gate

List of Symbols

C_{dm}, C_{DIFF}, C_{EXT}, C_{INT}, C_L, C_M	F	Bulk depletion capacitance, diffusion parasitic capacitance, Fan-Out load, total internal cell capacitance, load capacitance and coupling capacitance
C_{ox}	$F.m^2$	Gate oxide capacitance
d	m	Interface thickness
E_a, E_{dm}, E_d	eV	Activation energy, median dissociation and Si-H bond dissociation energies, respectively.
F_{lat}, F_{ox}	$V.m^{-1}$	Lateral electric field (F_{lat}= E_{lat}) and oxide field (F_{ox}≡ E_{ox})
g_m	S	Transconductance
I_{Nmax}, I_{Pmax}	A	Maximum values of the discharging currents for a nMOS and a pMOS devices, respectively.
I_{dlin}, I_{dsat}	A	Linear and saturation currents, respectively.
k_P, k_F, k_R		Surface recombination velocity at oxide/poly, forward dissociation rate constant and annealing rate constant
l, L	m	Pinch-off point-drain length and channel length
m		Body effect coefficient
q	C	Electron charge
R_{ds}, R_L	Ω	Total source and drain series resistances and channel resistance
S_I	s	Input signal ramp
T_{PHY}	m	Oxide thickness
V_{bs}, V_{ds}, V_{gs}	V	Bulk-source, drain-source and gate-source voltages, respectively
V_{dd}	V	Supply voltage
V_{FB}	V	Flatband voltage
V_{th}, $V_{th,ext}$	V	Threshold voltage and extrapolated threshold voltage
W	m	Transistor gate width
μ_0, μ ,μ_{eff}	$A.m^{-2}.V^{-1}$	Surface mobility, channel mobility and effective channel mobility
Φ_{MS}	V	Poly to Silicon workfunction difference
Φ_B, ψ_B	eV	Fermi potential
ψ_S, φS	V	Potential surface
Θ	V^{-1}	Mobility degradation coefficient (Spice parameter)
τ_{IN}	s	Input signal ramp
τ_{OUT}	s	Output signal ramp
σ		Distribution spread

Chapter 1

Introduction

The semiconductor industry has made great progress in the last four decades. Discrete MOS devices have been embedded into integrated circuits. Integrated circuits have gradually incorporated more and more functions on each chip. Today, it is possible to put an entire mainframe computer on a single piece of silicon. This tremendous growth is directly attributable to the scalability of the silicon MOS transistor to VLSI dimensions.

With the continued scaling of the CMOS devices into the sub-65nm, the impact of the process, voltage, temperature and other environmental variations on the performance of the digital integrated circuits has become extremely important. The increased on-chip operating temperature and the non-proportional scaling of the transistor dimensions with respect to the supply voltage have affected the reliability and the lifetime of these circuits. The accurate modeling of these variations and the accounting of their effects in the estimation of timing performance of the circuits has become one of the most critical issues in the field of IC design.

Negative Bias Temperature Instability (NBTI) and Hot Carrier (HC) are considered to be the dominant drift-related aging effects in current technologies. This thesis addresses the impact of these reliability wearout phenomena on the timing performance of the digital circuits.

In order to give a global context about the reliability mechanisms in CMOS technology, the next sections present an overview of both reliability basic concepts and wearout phenomena such as Electromigration (EM), Time dependent dielectric breakdown (TDDB), NBTI and HC.

1.1 Review of Reliability concepts

The reliability is defined as the probability that a system performs its intended function for a specified period of time under stated conditions. The purpose of the reliability is to ensure that the life of the system will be longer than the targeted life and that the failure rate during the normal operating life of the system will be below the target failure rate [Strong06]. The reliability of semiconductor devices is represented by the failure rate curve (called the "bathtub curve"), which is illustrated in the figure 1-1. The curve can be divided into the three following regions: (1) infant failures, which occur within a relatively short time after a device starts to be used, (2) operating life failures, which occur over a long period of time, and (3) wearout failures, which increase as the device approaches the end of its life.

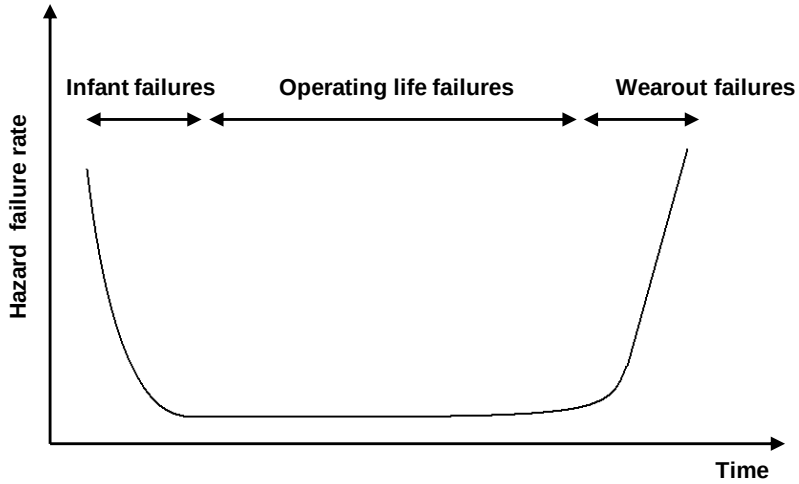


Figure 1-1 Reliability bathtub curve, a hazard failure rate plot in units of fails per time

Infant failures are considered to occur when a latent defect is formed in the course of the manufacturing process. For example, a defect can be formed by tiny particles in a wafer during the lithography process, resulting in a device failure later. Often the infant fails reveal technology problems that still to be addressed. These failures may manifest under the stress of the operation. The failure rate tends to decrease with time because only devices having latent defects will fail.

Operating life failures occur during the operational life of the product. In this period, the instantaneous failure rate should be small and the remaining high-quality devices should operate stably. The failures that occur during this period can usually be attributed to randomly occurred excessive stress, such as power surges, and soft errors.

Wearout failures occur due to many factors such as manufacturing process and aging of devices. The failure rate tends to increase rapidly in this period. Semiconductor devices are therefore designed so that wear-out failures will not occur during their guaranteed lifetime.

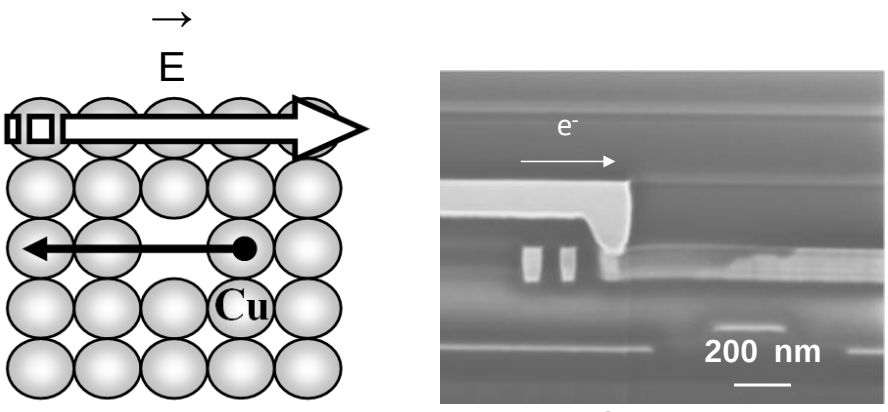
Accordingly, for the production of highly reliable semiconductor devices, it is important to reduce the initial failure rate to ensure the long life, or durability against wear-out failures [Renesas08].

1.2 Failure Mechanisms

This chapter introduces typical failure mechanisms of semiconductor devices that can be encountered in present nanoelectronic technologies.

1.2.1 Electromigration

Most semiconductor integrated circuits use Al, AlCu or Cu metallization wires for routing signals into and out of the chip or from one part of the chip to another, forming networks of devices on the chip. Failures on chip wiring fall into four categories: open circuit, elevated resistance, short circuit and leakage [Sullivan09]. Interconnect-induced failures can be provoked by several causes including corrosion, mechanical fracture (generally generated by differences in thermal expansion between adjacent materials during thermal cycling), etc. These phenomena can occur during chip operation life.



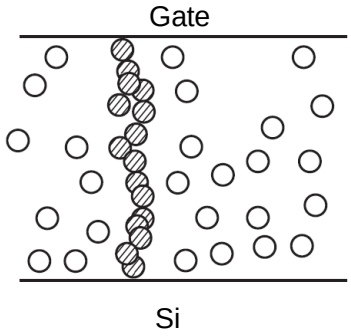
(a) (b)

Figure 1-2 a. EM failure mechanism; schematic illustration. b. Electromigration example in Cu wire

Electromigration is a phenomenon that resulting from the movement of metal atoms due to the current flow in a metallic wire. Cu atoms move in the direction of the electron flow. A void occurs near the negative electrode increasing line resistance, eventually leading to failure of the line (cf., Fig. 1-2). For the case of Cu wires, the failure times usually follow a lognormal distribution with the median lifetime depending on the quality of the interface [Hu99], [Arnaud00].

1.2.2 Time dependent dielectric breakdown

The Dielectric Breakdown is the irreversible change of physical properties of a local spot in the dielectric from low conduction to higher conduction. In the other words, within the dielectric area a tiny spot develops with increased conductivity compared to the rest of the dielectric area that remains nearly unchanged. A breakdown happens after a certain amount of time during which the oxide is subjected to an electrical stress during product operation. Due to the high conductance of the spots, the electric behavior of the area where these spots are localized change drastically. This local change of properties causes the stop operation of the product in most of the cases [Wu06].



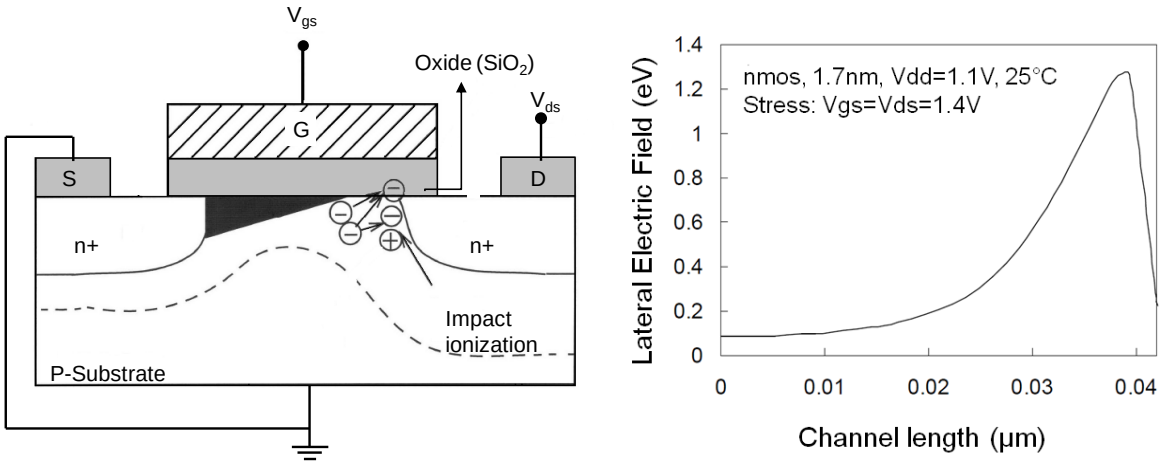
Interface states generation + connection → Oxide break down

Figure 1-3. TDDB failure mechanism

For instance, when an electric field is applied on a transistor gate an injection of the holes in direction to Si bulk is provoked throughout the oxide. As consequence a larger quantity of interface traps are generated. As the number of traps increases, an electric current via the traps is observed due to hopping or tunneling. It has been reported that if the number of traps continues to increase and the traps connect between the gate electrode and the Si substrate, the connection carries a high current that causes the gate oxide film to break down [Renesas08] (cf., Fig. 1-3).

1.2.3 Hot Carrier

As CMOS device scaling continues without reducing the supply voltage proportionally, hot carrier degradation became a major problem for long-term device reliability. When a MOS transistor operates in the saturation mode, a high lateral electrical field is generated near the drain (cf., Fig. 1-4). This large electrical field accelerates the carriers (electrons or holes) in the channel, driven them to become energetic or "hot". The hot carrier effects can overcome the potential barrier at the Si/SiO2 interface and penetrate into the oxide, generating a degradation of the electrical characteristics of devices such as mobility and threshold voltage [Hu85]. In the second chapter, the HC phenomenon as well as its impact on electrical performance of nMOS devices will be explained in detail.



(a) (b)

Figure 1-4 (a) Illustration of hot carrier stress configuration for an nMOSFET. (b) Lateral electric field in the channel [Guerin08]

1.2.4 Negative Bias Temperature Instability

NBTI occurs only in PMOS devices. When the gate of pMOSFET is under a negative bias (at high temperature condition), holes in the channel dissociate the Si-H bonds at the Si/SiO2 interface generating interface states (cf., Fig. 1-5a). The accumulation of these defects (N_{it}) leads to an increase of the PMOS threshold voltage (cf., Fig. 1-5b). Removal of the voltage stress can partially anneal the interface traps (recovery part), but the permanent NBTI component remains, thus the electrical performance of pMOS devices is deteriorated. The

main physic concepts about the NBTI phenomenon as well as some analytical models at transistor level will be presented in chapter two.

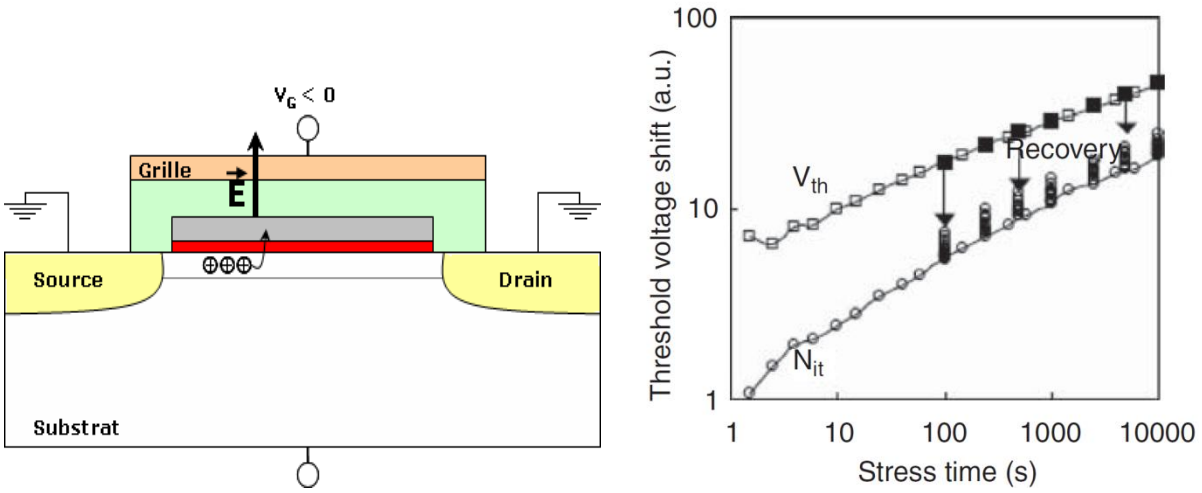


Figure 1-5 (a) NBTI DC stress electric configuration. (b) V_{th} shift (squares) and N_{it} –induced V_{th} shift (circles) during an NBTI DC stress [Huard04]

1.3 Design-in Reliability

As mentioned above, besides the NBTI and HC degradation phenomena there are other mechanisms that take parts in the yield reduction of circuits such as Electromigration, TDDB, Electrostatic discharge (ESD) [Schreier78]. The interaction between the different wearout mechanisms at distinct design levels have augmented the complexity of reliability analysis and in the same time have reduced the IC design margins. As a consequence, accurate reliability simulations are necessary to guarantee realistic circuit performances at a given time. Design-in Reliability (DiR) approach searches to provide an assessment and reliability improvement at all design levels. DiR methodology refers to the ensemble of models that can provide both correct simulation capability and high analysis capacity for any design. This manuscript addresses only NBTI and HCI reliability wearout mechanisms into the context of DiR methodology [Huard07b].

Several approaches have been proposed since 90's [Minehane00] [Zhihong06]. An example [Leblebici93] of design for reliability flow is showed in figure 1-6. As shown in the figure, the design for reliability flow follows a standard implementation with an exception. The main difference between the standard design flow and that for reliability consists in a reliability feed back into the electrical simulation and circuit design stages. An early ageing analysis

allows validating the design characteristics (i.e., topology, size of gates, etc) in order to ensure the life of the circuit.

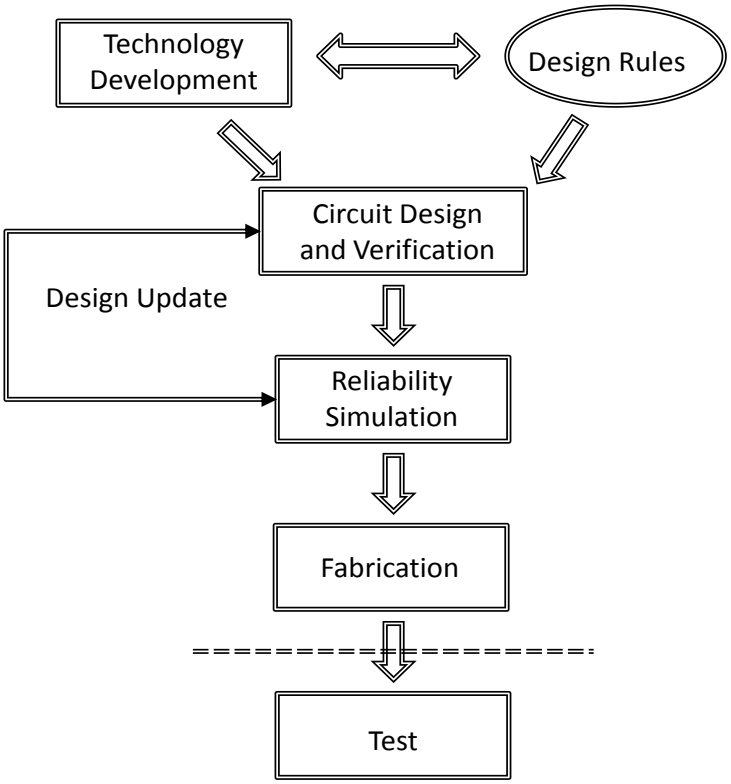


Figure 1-6 VLSI design approach with design for reliability

Various circuit reliability simulation tools have been developed since late 80's [Aur87], [Sheu89], which are built around transistor-level circuit simulation environments such as SPICE. This type of tool allows obtaining the magnitude of one or more circuit parameters, such as current, voltage, making possible the compute of device degradation.

1.4 Motivation

HCI and NBTI effects have increased the complexity of the timing optimization process in advanced nodes at all design abstraction levels (such as transistor, gate and system levels) due to diverse device offerings and overdrive product requirements.

In this context, one of the main current needs is to propagate the timings upwards into the design hierarchy, where reliability issues can be treated as other verification corner [Takayanagi05], using timings originated from the low level characterization of standard cells (§3.1).

Addressing these issues, DiR methodology can be used in order to deploy NBTI and HCI effects from transistor level hierarchically up to gate and system level in advanced nodes. Throughout this manuscript we provided details on how DiR methodology is used to develop a bottom-up reliability modeling, which is addressed to digital circuits.

1.5 Outline of the Thesis

The outline of this dissertation is organized as follows. Chapter 2 reviews the major NBTI and Hot Carrier degradation mechanisms. It also reviews the impact of these phenomena on electrical transistor performance.

Chapter 3 focuses on NBTI and Hot Carrier reliability modeling at gate level. Lasbouygues's propagation delay model is detailed here. It also explains the relationship between the reliability simulations and the design flow. In the second part of this chapter, we propose a new framework to propagate the timing degradations upwards into the design hierarchy while considering the mission profile of design under consideration.

Chapter 4 addresses the NBTI reliability modeling at system level. Basic definitions and terminology about Static Timing Analysis (STA) are presented in the first part of this chapter. In the second part, a silicon-CAD combined aging analysis is presented. This analysis is carried out on two IPs: a Low-Density Parity-check (LDPC) codec circuit and a cortex–A9 ARM processor. CAD results are compared against silicon measurements obtained from experimental experiences.

Finally, it demonstrates that it is possible by a bottom-up approach to build transistor- and gate-level models with sufficient accuracy to allow direct comparison with experimental degradations at system-level.

Chapter 5 summarizes the research results of this thesis and discusses about future research work.

The references to sections are identified by the symbol "§".

Chapter 2

Transistor Reliability Modeling

In order to illustrate the NBTI and HCI phenomena and their impact on transistor performance we present an overview of the physics of these phenomena. We will focus on the correlation between the degradation of physical parameters and their impact on electrical parameters.

The first part of this chapter presents the main physical concepts about Negative Bias Temperature Instability (NBTI). A detailed description of Reaction-Diffusion and Limited Reaction models will be treated to explain the NBTI effects on CMOS transistors.

The second part describes the electrical implications and the physical behavior of the Hot Carrier Injection (HCI) phenomenon at transistor level. In the last part of chapter the reliability simulation models and their application in the design-in reliability context will be shown.

In the last part of the chapter we will introduce the main characteristics of the reliability simulation as well as the transistor level analytical models for NBTI and HC mechanisms.

2.1 NBTI physics overview

Negative Bias temperature Instability is a degradation phenomenon occurring in pMOS transistors at elevated temperatures when a negative voltage is applied to the transistor gate. As a result, interface states (ΔN_{it}) are generated at the Si/SiO_2 interface and build up positively charged bulk defects [Deal67, Frohman71]. NBTI mainly impacts device electrical parameters such as threshold voltage (V_{th}), linear (I_{dlin}) and saturation (I_{dsat}) drain currents, and transconductance (g_m), therefore NBTI is a potential threat in current technologies and designs. Due to the aggressive CMOS scaling trend adopted from 0.13 µm technological node, an enhancement of NBTI degradation has been observed [Alam03]. This fact is mainly explained by the following reasons:

a- The introduction of dual poly-process to replace buried channel (n^+ poly gate) with surface channel (p^+ poly) pMOSFET and to reduce short channel effects with the transistor scaling trend.

b- Slower scaling of the supply voltage compared to more aggressive scaling of oxide thickness produces the gate oxide electric fields increases to values large enough to accelerate the NBTI damage.

c- Power supply reduction increases the circuit sensibility to electrical parameters shifts caused by NBTI (i.e. V_{th}).

d- The introduction of nitrogen at the Si/SiO_2 interface of surface channel pMOSFETs to reduce leakage current and boron penetration in p^+ poly has made NBTI worse.

NBTI degradation is a result of the contribution of two phenomena. First, the hole trapping and second the generation of the interface states. These mechanisms are explained in detail in the next section.

2.1.1 Physical mechanisms contributing to NBTI damage: Interface Traps and Positive Charges

The NBTI-induced physical damage has two independent contributions: interface states generation (ΔN_{it}) and generation or activation of positively charged defects. In turn, this

damage is made up of two parts (§2.2); a permanent component (D_P) "nonrelaxable" made in equal proportion of interface states and positive fixed charges; a recoverable component (D_R), caused by the activation and neutralization of pre-existing holes traps in the oxide. Figure 2-1 shows the relative shift of the interface states generation and the V_{th} increase normalized to 1 when NBTI stress is stopped (2500 sec). During the stress phase *A* (0-2500 sec), the relative N_{it} concentration and the relative V_{th} increase rapidly. In the relaxation phase *B*, a strong reduction is observed for the threshold voltage while the interface states density remains almost unchanged. The different relaxation behaviors of the N_{it} and V_{th} suggest the dual nature of the NBTI damage. Additionally, it can be observed that the threshold voltage does not depend only on the interface states density and that the recovery effect is associated with both neutralization of activated hole traps and repassivation of the generated interface states.

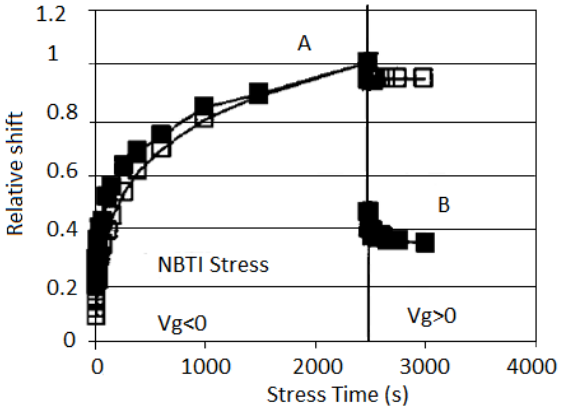


Figure 2-1. Relative shifts of interface states density (open squares) and threshold voltage (filled squares) [Huard03]

2.1.2 Impact of NBTI to pMOSFET electrical parameters

As mentioned early, due to the generation of interface traps and the activation or generation of positively charges defects, the NBTI damage impacts mainly two physical parameters of transistor: the flatband voltage (V_{FB}) and the effective channel mobility (μ_{eff}). Consequently electrical parameters such as I_{dlin}, I_{dsat}, V_{th} and g_m are also affected. In order to describe this relation, we will consider the following expressions [Rosa09, Liang86]:

$$V_{FB} = \left(\frac{-q}{C_{ox}}\right) \cdot \left(\Delta N_{it}^{D} + N_f^+ + N_R^+\right), \qquad (2\text{-}1)$$

where N^{+}_{f} and N^{+}_{R} are fixed and recoverable positively charges, C_{ox} is the gate oxide capacitance , ΔN^{D}_{it} is the donor-type portion of the interface state generated during the NBTI stress and electrically activated at time *t* in the pMOS under consideration.

$$\mu_{eff} = \frac{\mu_0}{\{1+\theta(V_g - V_{th})\}}, \tag{2-2}$$

where θ (V^{-1}) is the mobility degradation coefficient, V_g is the gate voltage and μ_0 is the surface mobility (this equation is only valid in the context of Spice simulations).

During the linear regime ($V_{gs} > V_{th,ext}$), the drain current (I_{dlin}) can be written as [Taur98]:

$$I_{dlin} = \left(\frac{C_{ox} + (V_{gs} + V_{th,ext}) \cdot W \cdot \mu_{eff} \cdot V_{ds}}{L + C_{ox} (V_{gs} - V_{th,ext}) \cdot W \cdot \mu_{eff} \cdot R_{ds}} \right), \tag{2-3}$$

where $V_{th,ext}$ is the extrapolated threshold voltage (assuming a long channel device).

In this relation, R_{sd} is the total source and drain series resistances. The extrapolated V_{th} is directly related to the flatband voltage by:

$$V_{th,ext} = V_{FB} + 2\Phi_B + \frac{\sqrt{4\varepsilon_S \, qN\Phi_B}}{C_{ox}}, \tag{2-4}$$

where $\Phi_B = (\frac{k_B T}{q}) \cdot \ln\left(\frac{N_A}{n_i 0}\right)$ (defined as Fermi potential at time zero). From equation (2-4), it can be seen that the shift in $V_{th,ext}$ relates to the change in V_{FB} through the next expression:

$$\Delta V_{th,ext} = \Delta V_{FB} \tag{2-5}$$

Defining the channel resistance as $R_L \equiv \frac{L}{(C_{ox}(V_{gs} - V_{th,ext}) \cdot W \cdot \mu_{eff}}$ and assuming that $R_{sd} << R_L$, it can be deduced (from the equation (2-3)) that the NBTI-induced shift of I_{dlin} (ΔI_{dlin}) at a given V_{gs} is related to $V_{th,ext}$ and μ_{eff} shift by the equation

$$\frac{\Delta I_{dlin}}{I_{dlin0}} \approx \frac{\Delta V_{th,ext}}{V_{gs} - \Delta V_{th,ext0}} + \frac{\Delta \mu_{eff}}{\mu_{eff0}}, \tag{2-6}$$

with $V_{th,ext0}$, μ_{eff0} and I_{dlin0} representing the initial values of μ_{eff} ,$V_{th,ext}$, I_{dlin} respectively (before stress).

Under the $R_{ds} << R_L$ condition [Rosa09], a linear dependence between the saturation current shift (ΔI_{dsat}) and linear current shift (ΔI_{dlin}) can be quantified, since:

$$\frac{I_{dsat}}{I_{dsato}} \approx \theta \cdot \frac{\Delta V_{th,ext}}{V_{gs} - \Delta V_{th,exto}}, \text{ where } 1 \leq \theta \leq 2. \quad (2\text{-}7)$$

The above equation suggests that both I_{dlin} and I_{dsat} follow the same behavior in the time during the NBTI stress (cf., Fig. 2-2), nevertheless the shift of I_{dlin} is almost twice less than the shift of I_{dsat}.

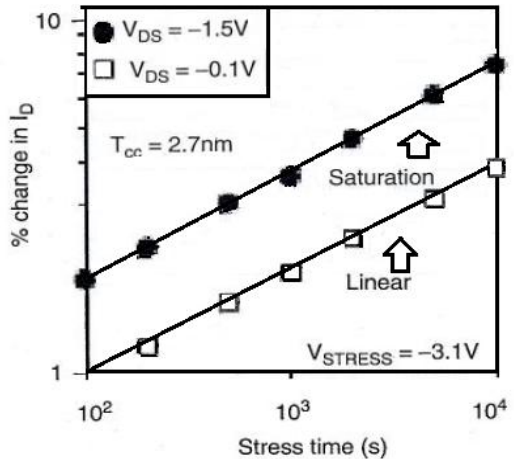


Figure 2-2. I_{dlin} and I_{dsat} evolution as function of time during and NBTI stress [Krishnan03]

Assuming the maximum transconductance (g_m) as $g_m = (\delta I_{dlin} / \delta V_{gs})_{V_{ds}}$, the impact of the NBTI degradation on g_m can be defined as follows:

$$\left|\frac{\Delta g_m}{g_{mo}}\right| \approx \left|\frac{\Delta \mu_{eff}}{\mu_{effo}}\right| \quad (2\text{-}8)$$

In figure 2-3 it can be observed, that both V_{th} and g_m experience the same evolution with the time, showing that in this case the increase in interface states (N_{it}) and positively charged defects affect both parameters [Kimizuka00].

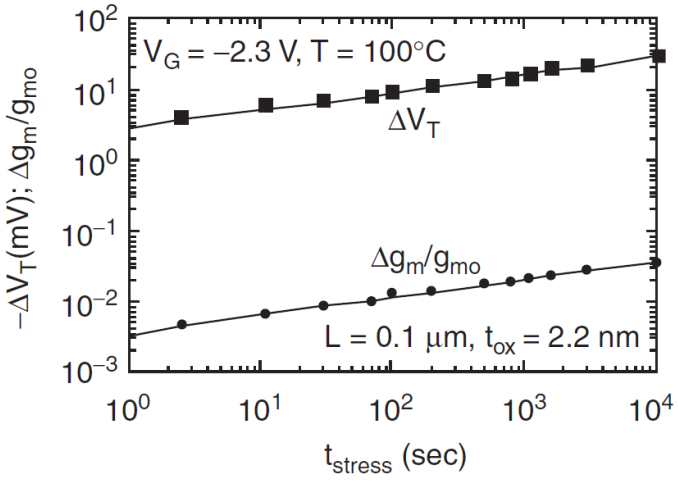


Figure 2-3. Vth shift and Gm degradation induced by NBTI.

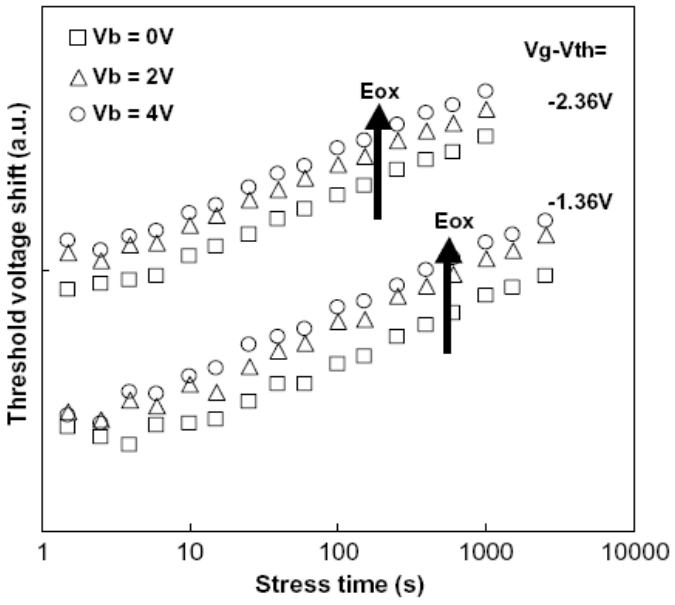
An important requirement of any physical model for NBTI is that it should explain the NBTI physical behavior observed in experimental studies. Before presenting the main NBTI modeling approaches, the following section describes key physical aspects of NBTI damage.

2.1.3 Physical picture of NBTI

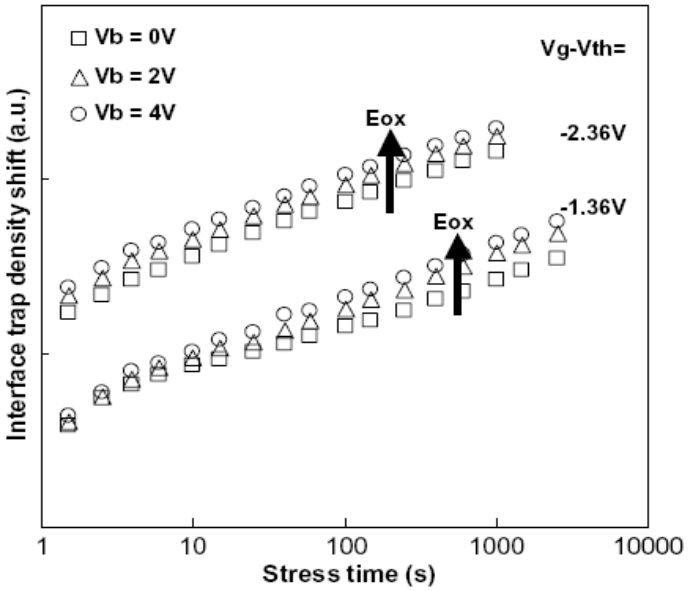
In order to describe the evolution of transistor parameters under NBTI stress conditions, in this section we describe some of the main NBTI physical characteristics.

2.1.3.1 Gate oxide electric field dependence

The NBTI dependence on the gate oxide field (F_{ox}) is clearly illustrated in figure 2-4 [Huard06]. Figures 2-4a and 2-4b show the threshold voltage and interface trap density dynamics with the stress time, where the gate voltage (V_g) is changed in a way to keep (V_g - V_{th}) holding almost constant the channel hole population. It is observed that both ΔV_{th} (Fig. 2-4a) and the ΔN_{it} (Fig 2-4b) variations rise with increase in F_{ox} [Mitani02].



(a)



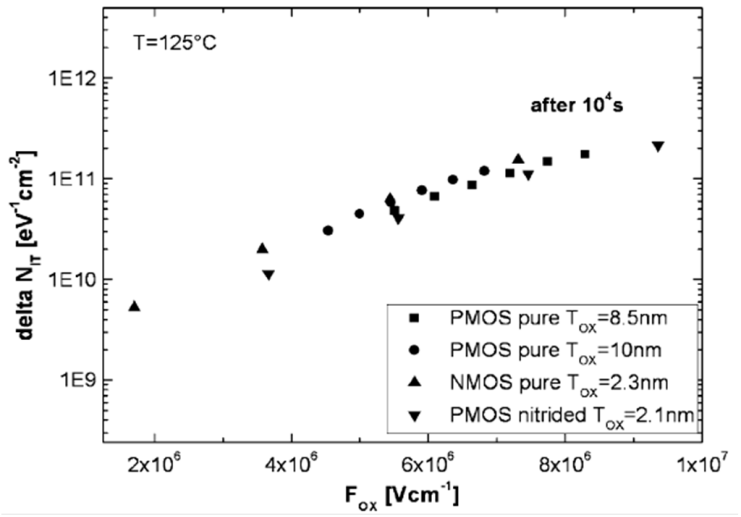
(b)

Figure 2-4. Threshold voltage shifts (a) and interface state generation (b) as a function of stress time for different bulk biases V_b=0V(squares), V_b=2V(triangles) and V_b=4V (circles) at 125°C. ($E_{ox} \equiv F_{ox}$)

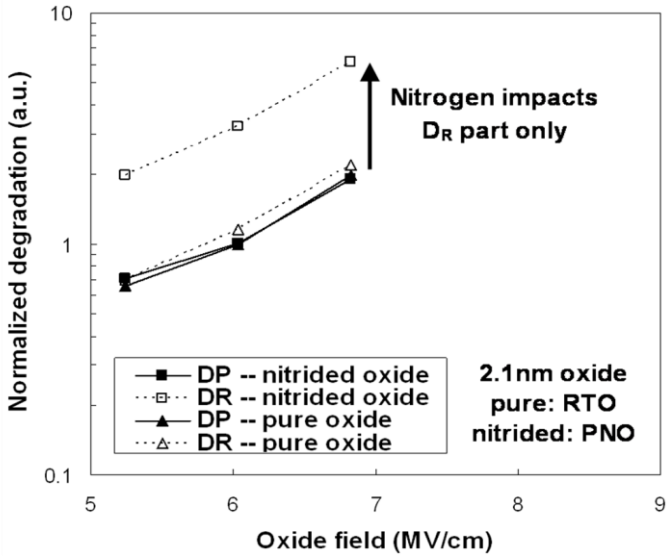
2.1.3.2 Process dependence (Nitrogen)

Despite the fact that multiple publications [Kimizuka00] assert the forceful impact of the incorporation of nitrogen on the NBTI degradation, the role of nitrogen on the NBTI enhancement is still not well explained. In [Hook05, Tan05, Fujieda05] authors suggest that the introduction of the nitrogen at Si/SiO2 interface generates a mechanical stress in the atomic structure. Therefore the nitrogen atoms should modify the interface states creation rate. Nevertheless Huard *et al.* [Huard06][1] show that the interface states generation is identical for pure and nitride oxides (cf., Fig. 2-5a). On the contrary figure 2-5b shows that the incorporation of nitrogen increases the recoverable component D_R of the NBTI degradation (§2.1.1.1). Notwithstanding the nitrogen helps in the reduction of both gate leakage current and boron diffusion, nitrogen related-traps close to the Si/SiO_2 interface have a negative influence on electrical parameters of device such as ΔV_{th}, ΔI_{dsat}.

[1] Experiments were carried out on MOSFETs with oxide thickness ranging from 1.6 nm to 10 nm. The MOSFETs have gate lengths ranging from 0.1 µm to 10 µm and typically a width of 10 µm. NBTI stress was applied with the gate electrode held at a low constant negative bias (ranging from_0.75 V to_3.5 V) under a temperature ranging from 25 _C to 200 _C while the source/drain and n-well electrodes were grounded



(a)



(b)

Figure 2-5. (a) Oxide field dependence of the interface traps creation for p-MOSFETs for pure and nitride oxides with large oxide thickness range [Huard06]. (b) Nitrogen impact on NBTI behavior. Recoverable part is affected by the nitrogen incorporation into gate while permanent part remains unchanged.

2.1.3.3 Temperature dependence

By definition, NBTI degradation is a thermally activated process. Figure 2-6 shows the influence of the temperature on the interface states generation. Two different regimes are observed. First, for low periods of stress time, the N_{it} shift follows a power law behavior. Secondly, for long stress time, the N_{it} reaches the saturation behavior. Figure 2-7 shows that the power law exponent increases linearity with the temperature, this implies that the activation energy E_a (as determined through degradation levels in an Arrhenius [Huard02]) at

a given time is not a constant value, which is in contradiction with the principles of RD model (§ 2.2.1) which assumes that the dissociation energy of Si-H bonds can be defined by a single value.

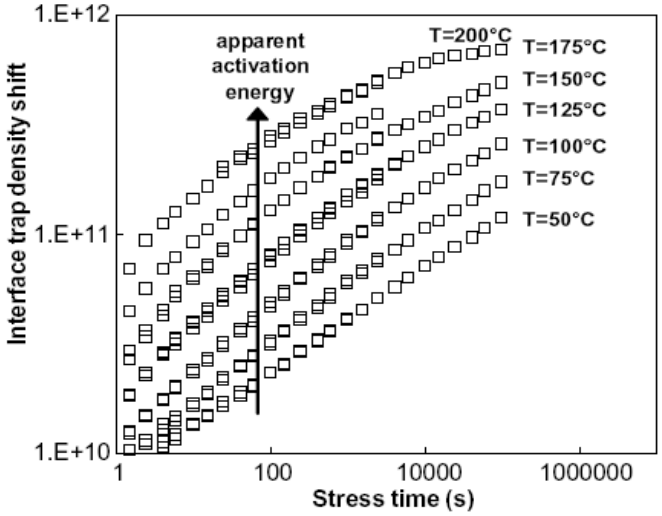


Figure 2-6. Interface traps density shift at a constant gate voltage stress for temperatures ranging from 50°C to 200°C [Huard06]

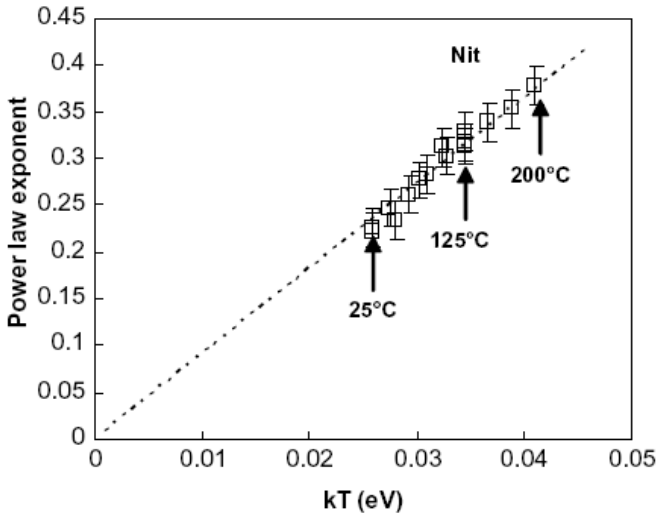


Figure 2-7. Power law exponent describing the dynamics of interface states generation at low times as a function of stress temperature [Huard06]

2.1.3.4 NBTI frequency independence

Frequency independence of the interface traps generation has been observed over a wide range of frequencies [Alam05, Schoder05]. Although the exact range of frequencies over which this phenomenon holds is still not very clear. According to [Alam03, Chen03] the frequency independence is clearly seen in the 1Hz-1MHz range, however more recently it has

been demonstrated [Huard06] that this behavior holds in the range of 1Hz-2GHz. The frequency independence is clearly observed in figure 2-8, in which the AC NBTI voltage acceleration is compared with the DC NBTI stress. Both DC and AC NBTI voltage accelerations are shown to be equivalent.

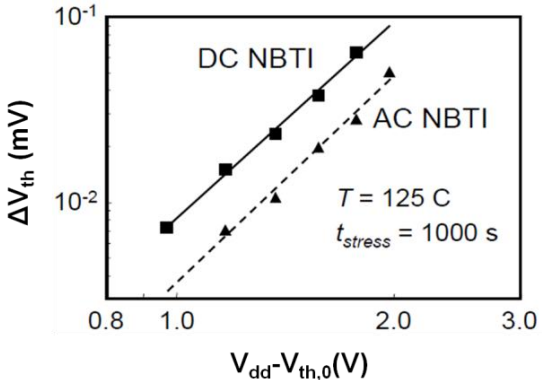


Figure 2-8. Frequency-independent AC NBTI voltage acceleration is equivalent to the DC NBTI voltage acceleration [Fernandez06]

2.2 NBTI modeling

In this section, we present a description of two main analytical models for NBTI. Detailed illustrations will be provided for the Reaction-Diffusion (R-D) and the Limited Reaction (L-R) models. Based on experimental observations we detail the limitations of the RD model and explain why L-R model allows modeling the NBTI degradation in a better way.

2.2.1 R-D Model

In order to explain the mechanism of negative bias temperature Jeppson and Svensson proposed the first R-D model [Jeppson77]. This is based on a structural model of Si-SiO_2 interface defects, which assumes the existence of a large number of hydrogen-terminated trivalent Si bonds $Si_3{\equiv}Si\text{-}H$ at the interface (electrically inactive). When a gate voltage is applied, it provokes a field-dependent reaction at the Si/SiO_2 interface that breaks the passivated Si-H bonds generating a neutral hydrogen atom and an interface trap. An interface trap is defined as an interfacial trivalent silicon atom with a non saturated electron $Si_3{\equiv}Si^{\bullet}$. Several years later, through an extensive mathematical analysis of the charges diffusing phenomenon Ogawa [Ogawa95] proposed that the rate at which the Si/SiO_2 interfaces traps are generated is controlled by the diffusion of hydrogen released from the hydrogen

passivated defect sites previously. Subsequently, Alam [Alam03, Alam04] presented an analytical model including the field, temperature and process dependencies of NBTI. This model provided a method to estimate the number of interface traps for a single stress phase, followed by a relaxation phase under three main assumptions: a-) a monoenergetic dissociation of Si-H bond. b-) the breaking of the passivated Si-H bonds at Si/SiO_2 interface when a gate voltage is applied. c-) the neutrality of diffusing species. The mechanism of trap generation as well as the hydrogen diffusion through the interface is explained by the following equations:

$$\frac{\delta N_{IT}\ (t)}{\delta t} = k_F(N_O - N_{it}) - k_R N_H N_{it} \qquad (x = 0) \qquad (2\text{-}9)$$

$$\frac{\delta N_{IT}\ (t)}{\delta t} = D_H\left(\frac{\delta N_H}{\delta x}\right) + \left(\frac{d}{2}\right)\frac{\delta N_H}{\delta t} \qquad (0 < x < d) \qquad (2\text{-}10)$$

$$D_H\left(\frac{\delta^2 N_H}{\delta x^2}\right) = \frac{\delta N_H}{\delta t} \qquad (d < x < T_{PHY}) \qquad (2\text{-}11)$$

$$D_H\left(\frac{\delta N_H}{\delta x^2}\right) = k_F N_H \qquad (x > T_{PHY}) \qquad (2\text{-}12)$$

Where x denotes the distance between the interface and the gate, N_{it} is the number of interface traps, N_O is the initial number of unbroken Si-H bonds, N_H is the hydrogen concentration, k_F is the oxide field dependent forward dissociation rate constant, k_R is the annealing rate constant, D_H is the hydrogen diffusion coefficient, *d* is the interface thickness, T_{PHY} is the oxide thickness and k_P is the surface recombination velocity at the oxide/poly-silicon interface.

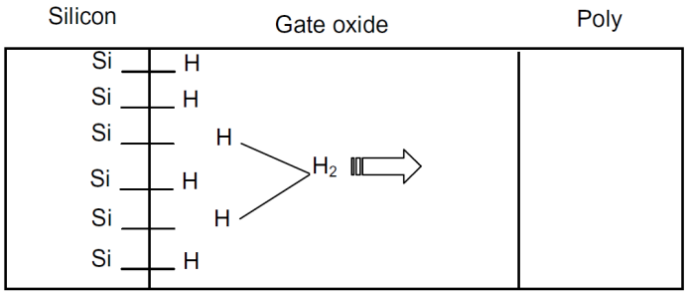


Figure 2-9. Schematic description of the reaction-diffusion model

Figure 2-9 shows a description of the R-D model used to interpret interface state generation during NBTI stress. Broken Si-H bonds at the Si/SiO_2 create Si^+ (interface traps) and H. Initial ΔN_{it} rate depends on Si-H bond dissociation (reaction), while the later rate depends on H removal (diffusion).

2.2.1.1 Description of the R-D model

The general solution of the R-D model in the stress phase anticipates five different regimes of time-dependent interface trap generation. The next paragraphs describe each one of these regimes.

First regime

During the initial stress phase the concentration of hydrogen atoms and interface are both very low, and there is virtually no hydrogen diffusion. Hence the term N_{it} increases with time linearly as:

$$N_{it}(t) \approx k_F N_{0t} \tag{2-13}$$

Second regime

In this phase, the initial density of Si-H bonds is larger than the number of interface traps that are generated, so N_O-$N_{it} \approx N_O$. This allows us to make the following approximation for the reaction equation:

$$k_F N_O \approx k_R N_H(x = 0) N_{it} \tag{2-14}$$

Third regime

During this stage, the hydrogen diffusion begins to control the interface traps generation. In these conditions the equation (2-11) suggests:

$$x = (D_H t)^{1/2} \tag{2-15}$$

In the same way the equation (2-10) expresses:

$$\frac{\delta N_{it}}{\delta t} \sim D_H \left(\frac{\delta N_H}{\delta x}\right) \sim \left(\frac{D_H N_H(x=0)}{D_H t^{\frac{1}{2}}}\right) \tag{2-16}$$

Then the approximate solution is:

$$N_H (x = 0) \sim \left(\frac{t}{D_H}\right)^{1/2} \left(\frac{\delta N_{it}}{\delta t}\right) \tag{2-17}$$

The above relation is substituted in (2-9) assuming that net trap generation is so slow that it is negligible in comparison to the right hand side of (2-9), thus it obtains:

$$N_{it} \sim \left(\frac{k_F N_O}{k_R}\right)^{1/2} (D_H t)^{1/4} \quad (2\text{-}18)$$

Fourth regime

In this phase the hydrogen reaches the oxide/poly-interface. According to equation (2-12) the incoming flux is defined by the following expression:

$$\frac{D_H (N_H(x=0) - N_H(x=T_{PHY}))}{T_{PHY}} = k_P N_H(x = T_{PHY}) \quad (2\text{-}19)$$

From equation (2-10) the hydrogen concentration (N_H) becomes:

$$N_H(x = 0) = \left(\frac{1}{k_P} + \frac{T_{PHY}}{D_H}\right)\left(\frac{\delta N_{it}}{\delta t}\right) \quad (2\text{-}20)$$

Substituting the latter equation in (2-9) and assuming that the traps generation is negligible ($\frac{\delta N_{it}}{\delta t} \approx 0$), it obtains:

$$N_{it} = A \left(\frac{k_F N_o}{k_R}\right)^{1/2} (D_H t)^{1/2}, \quad (2\text{-}21)$$

where A is defined as:

$$A = \left[2\left(\frac{D_H}{k_p} + T_{PHY}\right)\right]^{1/2}$$

Fifth regime

At the end of the stress phase, the general solution of R-D model anticipates quasi-equality between the quantity of Si-H bonds broken and the number of interface traps generated. In addition, this value becomes constant: → $N_{it} \approx N_o$ = *Constant*.

The five stages of R-D model are represented in figure 2-10. Note that each regime has a different value of slope.

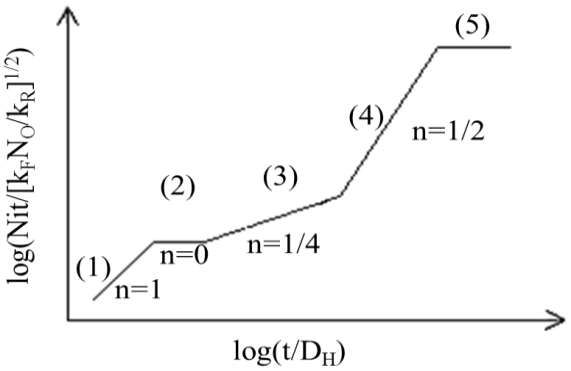


Figure 2-10. Five regimes of time dependent interface-trap generation as obtained from the general solution of the reaction diffusion equations during NBTI stress phase

2.2.2 Limited Reaction model

In this subsection, we will analyze the limitations of R-D model, which was presented in the precedent part.

2.2.2.1 Forward interface traps reaction rate

The R-D model fails to explain the complex behavior of NBTI phenomenon, particularly during the recovery phase as shown in figure 2-11.

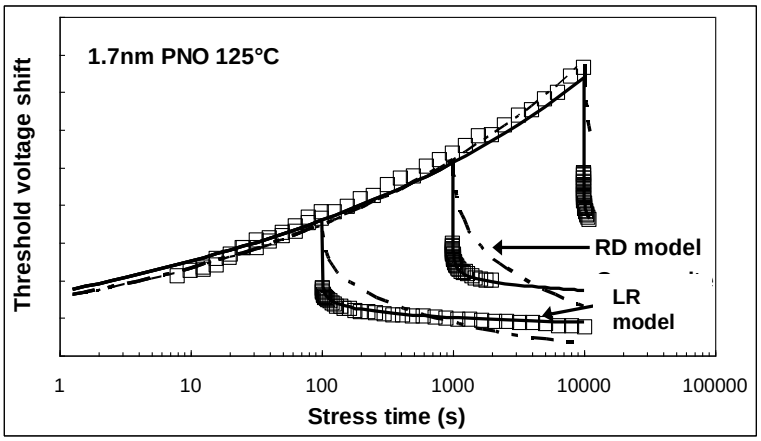


Figure 2-11. Models benchmarked against experiments for both the stress and the recovery phases using four devices under tests (DUTs) per stress/recovery conditions (12 DUTs in this case) [Huard07b]

According to original the diffusion-reaction model proposed by Jeppson and to the related R-D models [Kufluoglu04] the breaking Si-H bonds follows a linear dynamic reaction at low stress times:

$$N_{it}(t) = (1 - e^{-k_F N_{ot}}) \approx k_F N_{ot}, \quad (2\text{-}22)$$

where the dissociation rate k_F is assumed as constant for a given oxide field. However experimental experiences in [Huard06] show than the dissociation rate k_F is not constant (see Fig 2-12) but distributed as Fermi derivative distribution g(E, σ) defining as follows:

$$g(E_{d,\sigma}) = \frac{1}{\sigma} \frac{e^{(\frac{E_{dm} - E_d}{\sigma})}}{(1 + e^{(\frac{E_{dm} - E_d}{\sigma})})^2}, \quad (2\text{-}23)$$

where E_{dm} is the median dissociation energy and σ is the spread of the distribution (about 0.1eV) and E_d represents the Si-H bond dissociation energy for each one of Si-H bonds at the interface.

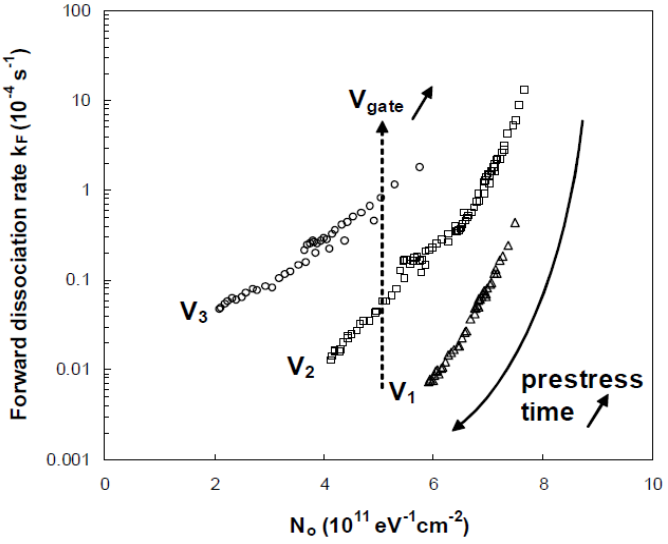


Figure 2-12. Forward dissociation rate k_F as a function of No [Huard06]

Equation 2-24 defines in an analytical way the degradation rate of every defect, taking into account that every single bond is broken according to a first-order equation, where each of them has a specific time constant depending on their own dissociation energy.

$$\frac{\Delta N_{it}}{N_{itmax}}(t) = \int_0^{\infty} g(E_d, \sigma)(1 - e^{-k_F E_d N_o k_f t}) \delta E_d \propto \frac{1}{1 + (\frac{t}{\tau})^{-\alpha}} \quad (2\text{-}24)$$

Where $\tau = \tau_0 \exp(\frac{E_d(E_{ox})}{kT})$, $\alpha = kT/\sigma$ for $\tau_{min} < t < \tau$, T is the bond temperature and τ_{min} is the time constant of the weakest defect.

Figure 2-13 illustrates the theoretical evolution of the interface traps generation as function of stress time, by using equation (2-24). This figure can be broken down into 3 parts. First part, represents the short periods of stress ($t < \tau_{min}$), the degradation is linear, with a slope linked to the defect generation rate of the weakest dissociation energy bonds. Its slope is defined by $\Delta N_{it}(t)/N_0 \approx t/\tau_{min}$. For longer stress time (second part) more and more different bonds (with more elevated dissociation energy) participate to the degradation. Finally, in the last part, the degradation rate reaches a saturation point according to the reduction of the bond numbers to be broken.

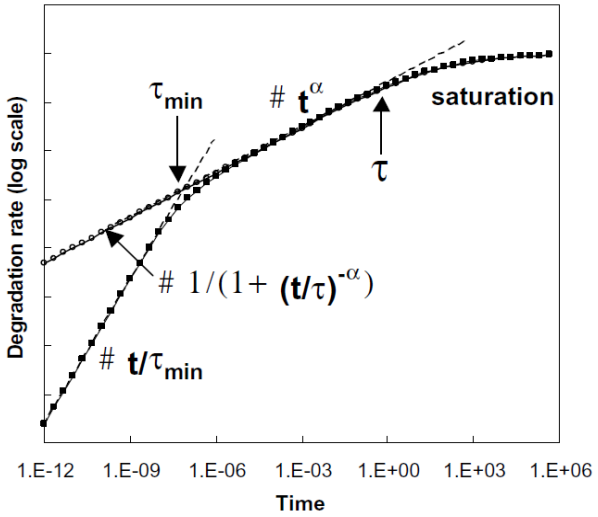


Figure 2-13. Evolution of theoretical degradation rate (filled symbols) with stress time with a disorder-induced variation of dissociation energies for a given temperature, as defined by equation (2-24) [Huard06]

Figure 2-14 show the measured interface traps density shift (symbols). Lines represent the predictions obtained from Eq. (2-24). It can be observed that the experimental dynamics is well reproduced by the analytical model (Eq. 2-24).

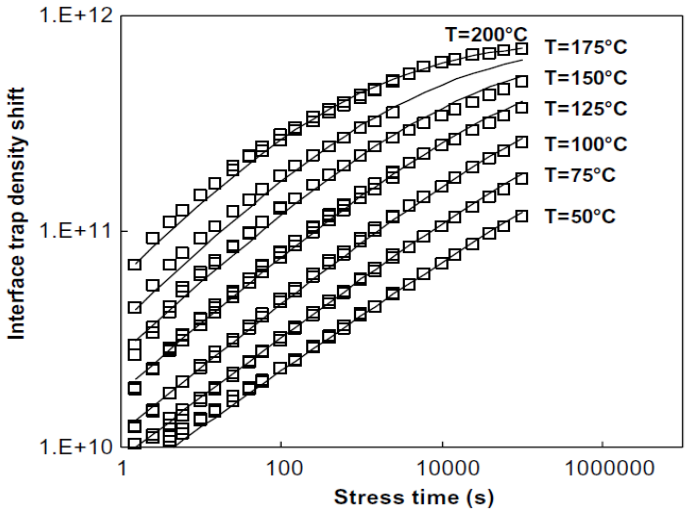


Figure 2-14. Interface states generation density shift for constant gate voltage stress (symbols) and fits according to equation 24 (lines) for a temperature range given [50°C-200°C] [Huard06]

2.2.2.2 Interface generation into permanent part

Another disagreement point with the R-D model is the fact that the R-D model predicts a passivation of the interface traps (created during the stress phase) once the stress is removed. In other words, once the stress is removed the hydrogen released by the dissociation of the Si-H bonds can diffuse back restoring their passive Si-H state. Figure 2-15 shows that though N_{it} level holds constant after one week of recovery, Vth shifts decrease to the constant N_{it} level (after one week). This experience indicates that the interface traps do not get passivated in this timeframe and are the main contributor of the permanent part [Huard07].

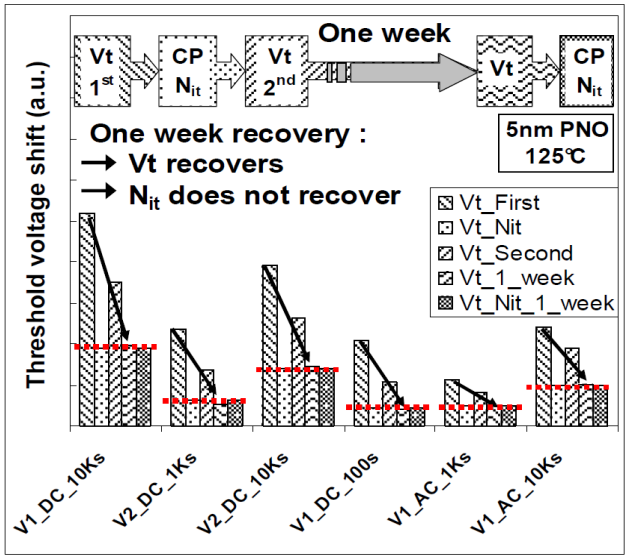


Figure 2-15. ΔVth and ΔNIT during NBTI recovering phase. For various stress conditions (stress voltage, dc/ac, and stress time) [Huard07]

As conclusion, we consider that the NBTI degradation is composed by two components: a permanent part $D_{P\text{-}NBTI}$ and a recoverable part $D_{R\text{-}NBTI}$ illustrate in figure 2-16. The $D_{P\text{-}NBTI}$ and $D_{R\text{-}NBTI}$ are generated by ΔN_{it} and the hole trapping, respectively.

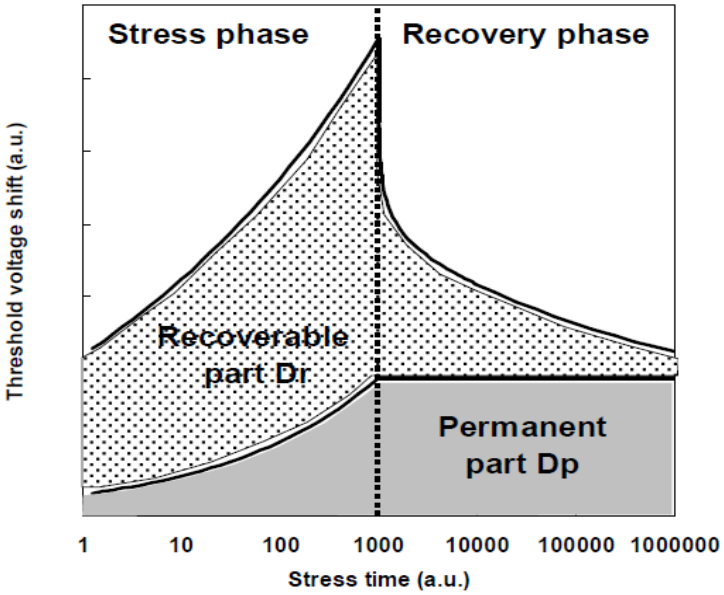


Figure 2-16. General scheme of the NBTI degradation made up of permanent and recoverable parts [Huard07].

According to [Parthasarathy06] the permanent damage can be modeled as a power law:

$$\Delta D_{P-NBTI} = A_P \cdot \exp(\gamma_P \cdot V_{gs}) \cdot \exp\left(\frac{E_{AP}}{KT}\right) \cdot t^n, \quad (2\text{-}25)$$

where V_{gs}, T and t denote the applied gate voltage, temperature and time respectively. γ, E_a and n are the voltage acceleration factor (§3.5.1), activation energy and time exponent respectively. A is a variable factor that depends of the transistor bias (*i.e.*, V_{ds}) and k the Boltzmann's constant.

In the same way the recoverable damage can be modeled as follows:

$$\Delta D_{R-NBTI} = A_R \cdot \exp(\gamma_R \cdot V_{gs}) \cdot \exp\left(\frac{E_{AR}}{KT}\right) \cdot \ln\left(1 + \frac{t}{\tau_d}\right), \quad (2\text{-}26)$$

where, τ_d is a time constant factor.

In the rest of this manuscript the L-R model will be used as base model of NBTI degradation to generate all our results.

2.3 Hot Carrier Overview

Hot carrier is a well known degradation phenomenon in MOSFETs, which is provoked by high kinetic energy carriers (holes or electrons). When passing through the channel, the carriers can gain enough energy and momentum to be injected into the gate and/or sidewalls oxides, which in turn affects the electrical characteristics of transistors [Takeda95].

Figure 2-17 describes the MOSFET HC-induced wearout process [Rauch09]. The three mechanisms involved in this process are:

a- HC damage is generated by carriers flow in the inverted channel ($V_{gs} > V_{th}$), which gain enough energy to pass the channel pinch-off point [Sze81].

b- The energetic carriers may lose their energy via impact ionization [Pagey03], contributing to the substrate current.

c- Hot carriers can acquire sufficient energy to break H bonds at the Si/SiO_2 interface generating interface states (ΔN_{it}). These interface defects lead MOSFET parameter shifts. This instability in the transistor parameters may generate fails in the long term operation of the circuits using these transistors [Rauch09].

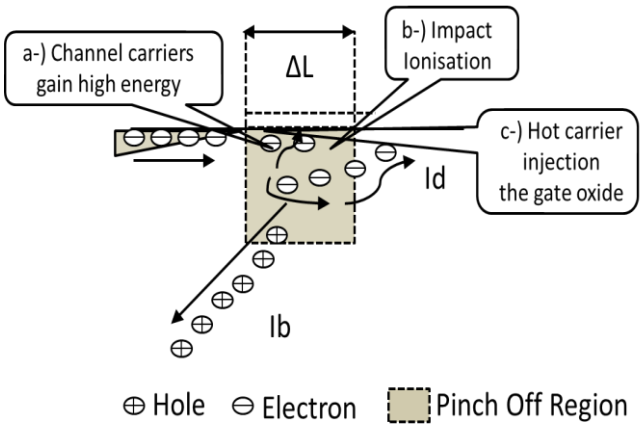


Figure 2-17. Representation of the physical mechanisms contributing to HC degradation for the nMOSFET case.

Taking into account the fact that HC degradation depends on both the number of carriers in the channel and their energy, it is important to understand the evolution of these two parameters with technological nodes. In order to illustrate this evolution we will use next expressions:

- lateral electric field (F_{lat}) is defined as follows

$$F_{lat} \propto \frac{V_{ds}-V_{dsat}}{l}, \tag{2-27}$$

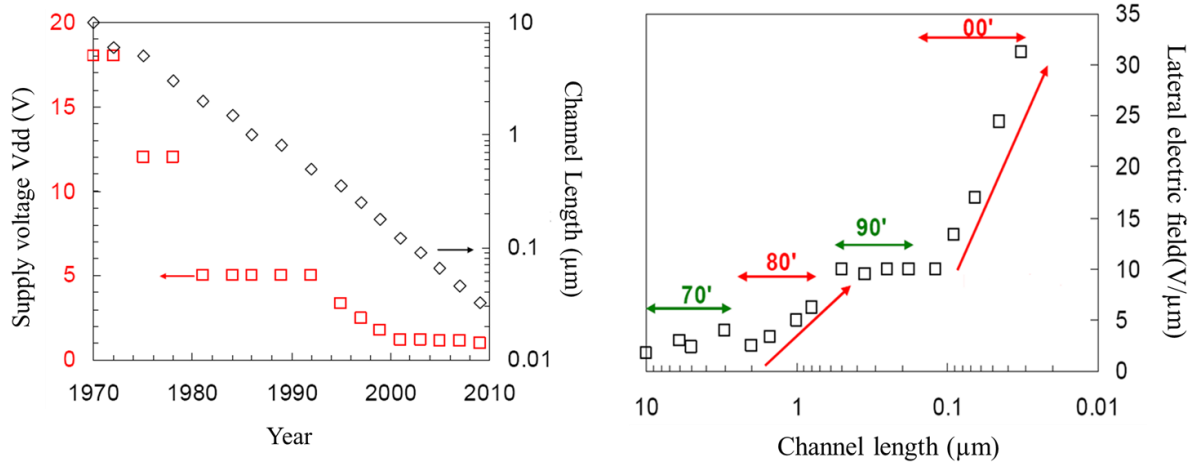
- an approximate value of F_{lat} is given by:

$$F_{lat} \propto \frac{V_{dd}}{L}, \tag{2-28}$$

where l represents the length transistor and V_{dd} is the supply voltage.

Between 0.5µm and 0.13µm nodes, the scaling dimensions as well as the voltage supply reduction were done in such a way that the lateral electric field remained constant. From 90 nm node, HC degradation has regained importance since supply voltage scaling has been slower despite the rapid reduction of the gate length. This happened because the difference between V_{th} and V_{dd} must be kept in a certain ratio in order to avoid the unwanted transistor switching (cf., Fig. 2-18).

On the other hand, a detailed study of the strongly dependence of the hot carrier degradation with bias polarization was performed by Takeda. In the next section we present in detail the four HC injection modes published in [Takeda83].



(a) (b)

Figure 2-18. (a) Vdd scaling and channel length trends (b) Lateral electric field evolution as function of channel length from equation (2-28) These data are those of Intel until 2003 [Moore03], they were supplemented by those of STMicroelectronics until 2009. [Guerin08]

2.3.1 Hot Carrier injection modes

Considering that the nature of HC damage is strongly depending of the bias conditions, four types of HC injection mechanisms can be distinguished:

I. *Channel hot electron (CHE) injection* is generated by high kinetic energy electrons "lucky electrons" [Shockley61], which have sufficient energy to surmount the energy barrier at the Si/SiO_2 interface. This mechanism occurs mainly at $V_{gs} = V_{ds}$ condition.

II. *Drain avalanche hot carrier (DAHC) injection* .This takes place when a high voltage applied at the drain under non-saturated conditions (Vgs < Vds) results in very high electric fields near the drain, which accelerate channel carriers into the drain's depletion region.

III. *Substrate hot electron (SHE) injection*, which dominates at V_{ds}= 0V, $V_{gs} > 0$ and $V_{bs} > 0$. This is due to the homogeneous emission of leakage electrons from the substrate into the interface.

IV. *Secondarily Generated Hot Carrier (SGHC) injection* is induced by secondary impact ionization of substrate carriers [Bravaix01] or photon generation [Mihnea02]. This injection become more important at $V_{bs} < 0$ condition for the nMOSFET case.

2.3.2 Electrical parameters degradation caused by Hot Carrier phenomenon

As mentioned early, HC impacts the MOSFETs electrical performance. This degradation is mainly produced by two mechanisms. First, the creation of N_{it} at the Si/SiO_2 interface, and, second, the defect generation in surround oxide films silicon bulk. The HC-induced physical damage affects channel parameters such as the effective mobility (μ_{eff}) and the flatband voltage (V_{FB}), which in turn modify the parameters such as I_{dlin}, I_{dsat}, V_{th} and g_m.

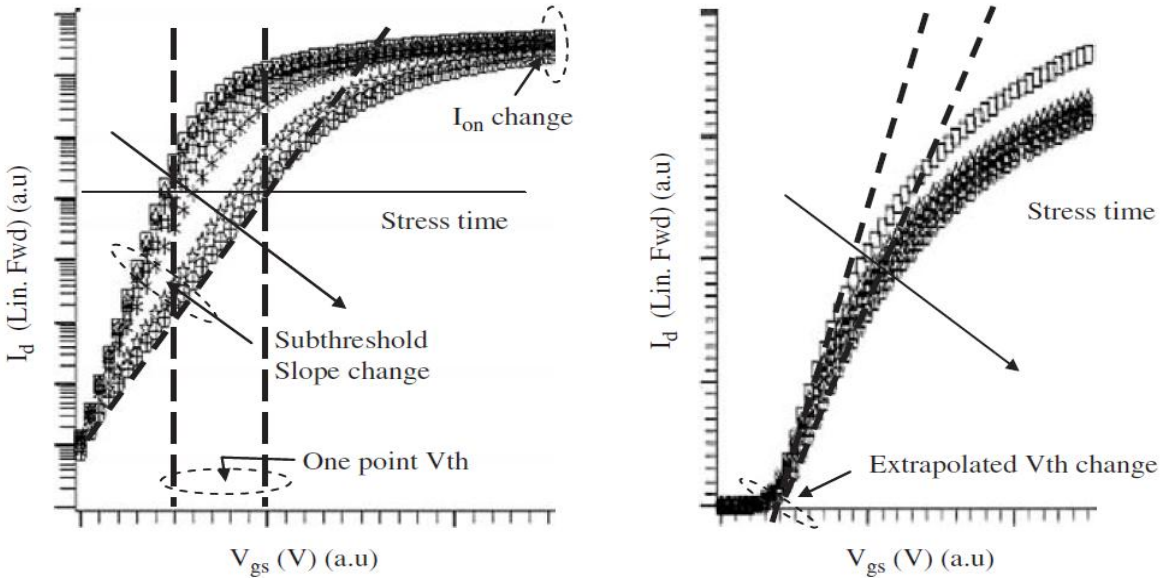


Figure 2-19. Idlin evolution as function of V_{gs} of nMOSFET tested at $V_{ds} < V_{dsat}$ condition, during a peak of substrate current (I_s). The HCI damage decreases the subthreshold slope as well as I_{on} (a) and increases the extrapolated V_{th} (b) [Rauch09]

Figure 2-19[2] describes the I_{dlin} evolution as function of V_{gs} at different stress time. It can be observed that the HC degradation reduces the I_{dlin} subthreshold slope, which in turn increases the V_{th} (see figure 2-19a). In addition, we can be seen that the V_{th} extrapolated ($V_{th,ext}$) experiments a small shift during the stress (cf., 2-19b).

Figure 2-20 illustrates the I_{dlin} degradation in percentage, which follows a power law dependence in stress time ($n \approx 0.5$). This behavior is generally associated with the degradation mobility, which is provoked by the channel ΔNit generation. The parallel evolution of the I_d variation ($\%I_d$) with the time between testing at $V_g \approx V_{th}$ and V_{dd} is a further indication that in this case, the HC-induced damage is controlled by the generation of interface states. In order to explain the correlation between the physical and electrical parameters we will analyze an nMOSFET (in 0.25µm CMOS technology).

[2] For a nMOSFET of the 0.25 µm CMOS technology with source and drain design made up of a single shallow. As drain (8.0E14, 10 Kev) and halo BF2 (2.4E13, 55 Kev, 301).

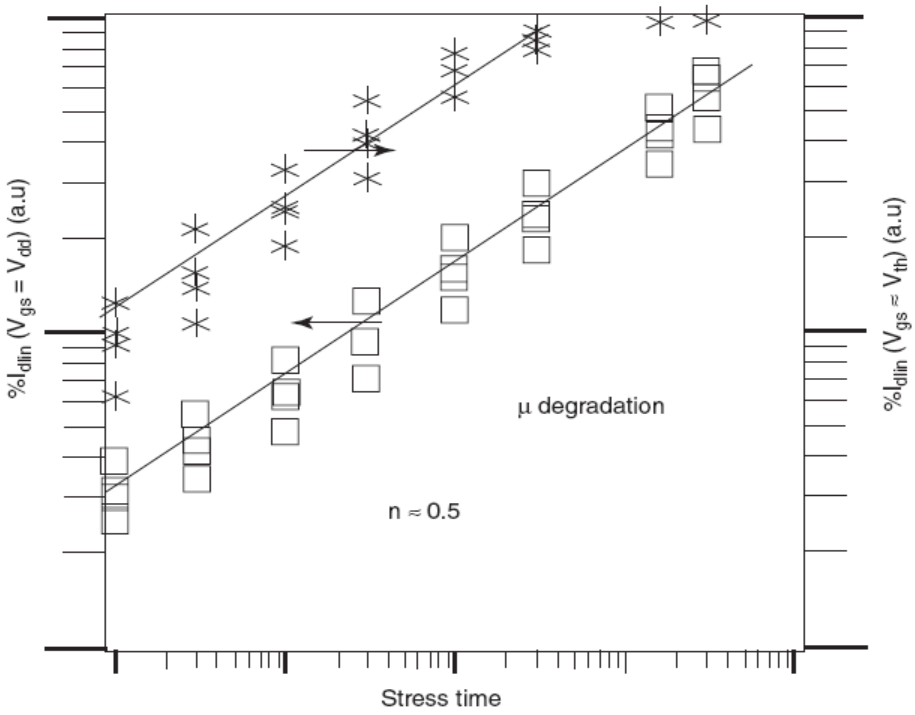


Figure 2-20. I_{dlin} degradation in percentage as function of stress time at both ($V_{gs} \approx V_{th}$) and ($V_{gs} = V_{dd}$) conditions [Rauch09]

Figure 2-21 illustrates the evolution of device parameters for weak and strong inversion assuming a long channel device. In order to explain the shifts of electrical device parameters provoked by HCI, in the next section our explanation will be based in the I_{dlin}–V_{gs} relation.

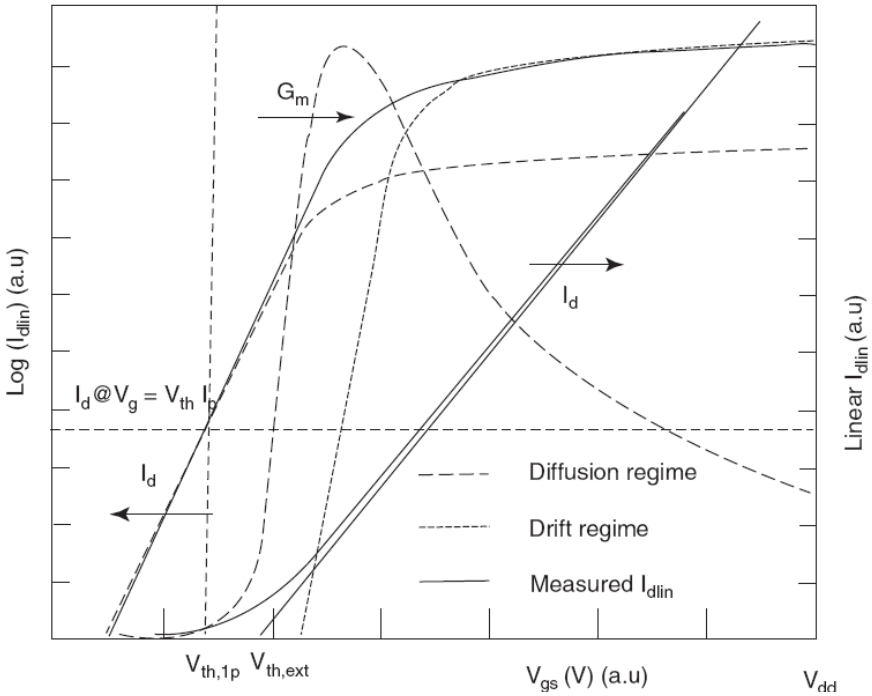


Figure 2-21. Idlin evolution versus V_{gs}. Linear drain current is plotted on both logarithmic and linear scale. [Rauch09]

2.3.2.1 Weak Inversion ($V_{gs} < V_{th,ext}$)

In this regime, the drain current is dominated by charge carrier diffusion through the channel. According to [Taur98], I_{dlin} can be defined as follows:

$$I_{dlin} \approx \mu_{eff} C_{ox} \frac{W}{L} \cdot \left(\frac{kT}{q}\right)^2 \cdot e^{\frac{q(V_g - V_{th,ext})}{mkT}} \left(1 - e^{-\frac{qVds}{kT}}\right), \quad (2\text{-}29)$$

where W and L are the device channel width and length, respectively. C_{ox} is the gate oxide capacitance and μ_{eff} is the effective channel mobility. $V_{th,ext}$ is the extrapolated threshold voltage at the stress time *t*.

And, the body effect coefficient m is defined as:

$$m = 1 + \frac{C_{dm}}{C_{ox}}, \quad (2\text{-}30)$$

where C_{dm} is the bulk depletion capacitance per unit area at $\psi_S = 2\psi_B$ condition.

ψ_S is the potential surface defined as: $\frac{2kT}{q} \cdot \ln\left(\frac{Na}{ni}\right)$,where N_a and ni are the acceptor impurity density and the intrinsic carrier density, respectively.

The C_{dm} depends on depletion (C_{depl}) and the interface states (C_{it}) capacitances as follows:

$$C_{dm} = \frac{1}{\frac{C_{it}+C_{depl}}{C_{depl}C_{it}}}, \quad (2\text{-}31)$$

where C_{it} is given by

$$C_{it}(\psi s) \equiv q \cdot \frac{\delta Nit_{it}^{A}(\psi s)}{\delta \psi s}, \quad (2\text{-}32)$$

with $Nit_{it}^{A}(\psi s)$ representing the acceptor-type N_{it} at the Si/SiO_2 interface.

The increase of interface states number during a channel hot carrier (CHC) stress will increase the $C_{it}(\psi s)$ with a consequent reduction of C_{dm}. Considering that subthreshold slope (*SS*) is: $ss = \frac{1}{\left(\frac{\delta log10(I_{dlin})}{\delta V_{gs}}\right)}$. From equation (2-29), the *SS* can be defined by:

$$ss = 2.3\frac{kT}{q}\left(1 + \frac{Cdm}{Cox}\right), \quad (2\text{-}33)$$

From equations (2-32) and (2-33) it can be observed that the $V_{th,1p}$ and $V_{th,ext}$ shifts are related to each other as next:

$$\Delta V_{th,1p} = \Delta V_{th,ext} \cdot \%ss \tag{2-34}$$

A large *SS* variation (%*ss)* due to an increase in N_{it} during a CHC stress results into a large ($\Delta V_{th,1p}$) for a small ($\Delta V_{th,ext}$), as observed in figure 2-21.

Assuming a ΔN_{it} = 0, C_{it} becomes also equal a zero, and according to equation (2-31), $C_{dm} = C_{depl,}$ $V_{th,ext}$ can be defined as :

$$\Delta V_{th,ext} = \Delta V_{th,1p} \tag{2-35}$$

2.3.2.2 Strong Inversion ($V_{gs} > V_{th,ext}$)

In this regime, the drain current is dominated by charge carrier drift through the channel. According to [Taur08], I_{dlin} is given by:

$$I_{dlin}(t) \cong \frac{WC_{ox}}{L} \cdot \mu_{eff}(t) \cdot \left(V_{gs} - V_{th,ext}(t)\right) V_{ds}{}^{*}(t), \tag{2-36}$$

where V_{ds}* is defined as follows:

$$V_{ds}{}^{*}(t) \equiv V_{ds} - I_{dlin}(t) \cdot R_d(t), \tag{2-37}$$

R_d represents the drain-source series resistance at a given stress time *t*. In this case we assumed that the initial source $R_S(0)$ and drain $R_d(0)$ series resistances (before stress) are negligible and, due to localized damage in the drain region, only R_d is affected during the stress. Combining equations (2-36) and (2-37) we get:

$$I_{dlin}(t) = \frac{C_{ox}(V_{gs} - V_{th,ext}) \cdot W \cdot \mu_{eff} V_{ds}}{L + C_{ox}(V_{gs} - V_{th,ext}) \cdot W \cdot \mu \cdot R_d} \tag{2-38}$$

Assuming a negligible increase in drain series resistance, this equation becomes:

$$I_{dlin}(t) = \frac{C_{ox}(V_{gs} - V_{th,ext}) \cdot W \cdot \mu_{eff} \cdot V_{ds}}{L} \tag{2-39}$$

From the latter equation I_{dlin} can be defined at a given stress time as follows:

$$\frac{I_{dlin}}{I_{dlin0}} \approx \frac{\Delta V_{th,ext}}{V_{gs}-\Delta V_{th,ext0}} + \frac{\Delta \mu_{eff}}{\mu_{eff0}} \approx \frac{\Delta V_{th,ext}}{V_{gs}-\Delta V_{th,ext0}} + \frac{\Delta g_m}{g_{m0}}, \quad (2\text{-}40)$$

The index "0" designates the initial conditions of the different equation terms (before stress).

Figure 2-22 gives an example of CHC impact on I_{ds} for nMOSFET device. The N_{it} generation reduces the drain saturation current I_{ds} and the mobility in the channel [Bravaix09]. The relation between physical damage (ΔN_{it}) and the channel mobility as well as $V_{th,ext}$ in linear conditions is described in the next section.

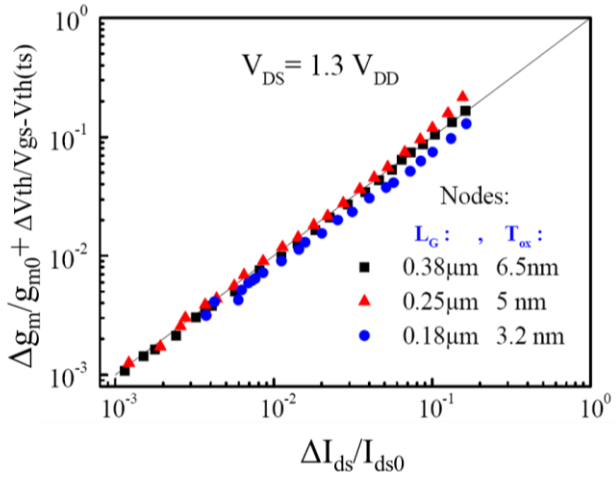


Figure 2-22. Correlation between the degradation of linear parameters as Ids reduction is directly proportional to the mobility reduction and V_{th} shift at the stress bias of maximum I_s [Bravaix09].

2.3.2.3 HC impact on mobility degradation Δµ and threshold voltage shift ΔV_{th}

In the strong inversion regime, the effective channel mobility at ($V_{ds}<V_{dsat}$) condition is given by:

$$\mu_{eff} \cong \frac{\mu_0}{1+\left(\theta(V_g-V_{th})\right)}, \quad (2\text{-}41)$$

where μ_0 is the channel mobility factor, and θ the channel mobility attenuation factor.

The interface states generation ΔNit can be related to channel mobility factor through

$$\frac{\Delta\mu_0(t)}{\mu_0(0)} \approx \frac{1}{1+\Delta Nit(t)}, \quad (2\text{-}42)$$

According to equation (2-4), the flatband voltage (V_{FB}) can be defined as:

$$V_{FB} = \Phi_{MS} - \frac{Qf+Qit}{Cox}, \quad (2\text{-}43)$$

with Φ_{MS} is the poly to Si workfunction difference, Q_f the fixed charge density, and $Q_{it}(\varphi S)$ the interface states induced charge density at the surface potential φS. Taking into account that only the change of Q_f and Q_{it} contribute to the $V_{th,ext}$ shift during the stress, V_{th} becomes:

$$\Delta V_{th,ext}(t) = q\left(\frac{\Delta Nt(t) + \alpha \Delta N_{it}(t)}{Cox}\right), \quad (2\text{-}44)$$

Where α is equal to + or – for acceptor- or donor-type N_{it}.

2.3.3 Hot Carrier modeling

Various hot carrier modeling approaches have been developed in order to represent in an analytical way the HC degradation modes and their implications at transistor level. In this section we will see the framework of the Lucky Electron Model (LEM) and the Energy Driven Model (EDM).

2.3.3.1 Lucky Electron model (LEM)

The injection of carriers into the SiO_2 layer has been typically described by the *lucky electron* approach [Shockley61]. In the LEM model the probability that a hot electron at a distance *d* reach the Si/SiO_2 interface without suffering any energy-robbing collision is given by:

$$P(d) = e^{(-d/\lambda)}, \quad (2\text{-}45)$$

where λ is the scattering mean free path of the hot electron in silicon[Tam84]. Figure 2-23 illustrate the injection hot carrier process as well as the relation between the parameter *d*, the electron kinetic energy and the hot electron emission energy barrier. An electron at x = *d* has just enough kinetic energy qV(x) to overcome energy barrier for emission, if it can travel from x = *d* to the interface at x = 0 without encountering energy-losing collision. That is, qV(*d*) represents the effective energy barrier for emission [Taur98].

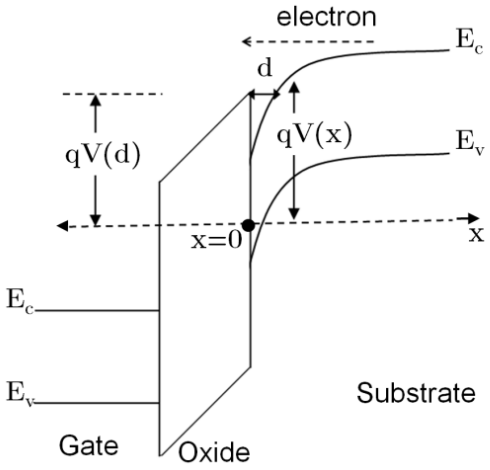


Figure 2-23. Illustration of the hot electron traveling towards the Si/SiO_2 interface

2.3.3.1.1 Bulk current in the LEM model

Considering that the bulk current is a good monitor for hot electron effects, since HC degradation is driven by the channel electric field (E_{lat}), authors in [Hu85] proposed a model based on drain and bulk current, which expresses the interface generation states probability as:

$$I_{bs} = \underbrace{C1}_{1} \cdot I_{ds} \underbrace{\cdot \exp\left(-\frac{\varphi i}{q\lambda E_{lat}}\right)}_{2}, \tag{2-46}$$

where term 1 is a constant and term 2 represents the probability that a channel electron travels a sufficient distance to gain energy φi without suffering collision. λ is the scattering free path of channel electron. The relation between physical damage (ΔN_{it}) and the channel mobility as well as $V_{th,ext}$ in linear conditions is described in the next section. Recently, newer energy-driven theories have been introduced to generalize the ideas of the LEM, which are presented in the next section.

2.3.3.2 Energy Driven mechanisms

Experimental evidence shows that the LEM allows describing the HC degradation at elevated values of biases $V_{ds} \gtrsim 4.0V$. However this model requires the additional energy gain mechanisms [Rauch01] above the simple model "pure" of heating caused by a high electrical field in order to explain the HC degradation that has been observed in experimental experiences for $V_{ds} \lesssim 3.7V$(cf., 2-24) [Tam83]. An attempt to extent the LEM to modern

technological nodes (beyond the 0.13µm) was carried out by authors in [Rauch05], which is known as Energy Driven Model (EDM). In this model Rauch included electron-electron scatterings effects and defines the hot carrier rate as follows:

$$Rate = \int f(E)S(E)dE, \tag{2-47}$$

where f(E) is the electron energy distribution function (EEDF) and S(E) is the scattering rate function, this term can be given by S_{ii} (effective impact ionization) or S_{it} (for interface traps generation). The integrand of Eq. (2-47) generally has one or several peak points or "knee", which are points of high curvature associated with dominant energies (*Edom*).

There are three major mechanisms that affect the hot electron rate. The first of them was evidenced by authors in [Bude95], which showed that the EDDF has a significant knee near the maximum energy available (V_{EFF}) from the steep potential drop at the drain (*i.e.* at the pinch off point). A second knee is also observed at about $2V_{EFF}$, which is produced by the electron-electron scattering (EES).The second mechanism is the single vibrational excitation which is linked to high energy carriers, which adding energy at the system through ionization by impact of phonons with the lattice.

The third mechanism was evidenced by Hess *et al.*, in [Hess99]. They showed that the bonds may be broken by channel cold carriers (CCC) effects, through a fourth mechanism corresponding to multiple vibrational excitations (MVE). This corresponds to direct excitation of the vibrational modes of the bond by multiple carrier impacts, each of which individually have a low energy, but which can cumulatively break the bond.

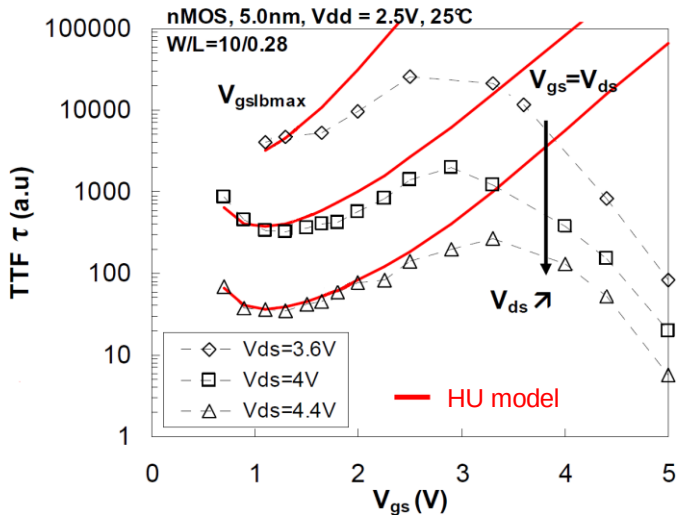


Figure 2-24 Comparison between experimental lifetime (symbols) and LEM (lines). The measured TTF have been determinate for 10% I_{dsat} [Guerin07b].

Figure 2-24 shows the Time-To-Failure (TTF) against V_g. The classical LEM picture [Hu85] predicts monotonous increase of TTF when the gate voltage is increased above $V_g = V_d/4$, which contradicts experimental observations. Note that LEM model not considers the MVE mechanisms.

2.3.3.3 Multi-mode model

Based on the above mechanisms, a HC degradation multi-mode model for CMOS device was introduced in [Guerin08]. This uses quantum-mechanical models to explain the process of carriers gaining energy, through three different mechanisms:

I. High-energy channel hot carriers based on direct electron excitation (DEE). This regime is characterized by a low drain current, *i.e.* low V_g. The generated hot holes are attracted to the interface due to negative oxide field generated at ($\sim V_g$-V_d) condition, with sufficient energy to break interface Si-H bonds.

II. Medium energy electrons which is based on electron-electron scattering. In this regime the value of drain current is moderated *i.e*. moderate energy V_d and V_g. Most of the carriers induced by impact ionization have not enough energy to break the Si-H bonds. Nevertheless, through carrier-carrier interactions, some of them can access higher energies and thus being able to break the bonds.

III. Channel cold carriers based on multiple vibrational excitation (MVE). In this regime distinguish by a high drain current *i.e.* high V_g, the carriers do not have enough energy to break a Si-H bond in a single direct excitation. Besides, their density increases rapidly with V_g close to the interface. Due to confinement, the electron density close to the interface strongly increases, which favors the low-energy interactions between electrons and bonds by MVE.

Therefore this model defines the hot carrier rate as follows:

Rate = $\frac{1}{\tau}$, where τ is the device lifetime.

$$\text{Rate} = C_1 \left(\frac{I_{ds}}{W}\right)^{a1} \left(\frac{I_{bs}}{I_{ds}}\right)^{m} + C_2 \left(\frac{I_{ds}}{W}\right)^{a2} \left(\frac{I_{bs}}{I_{ds}}\right)^{m} + C_3 \, {V_{ds}}^{\frac{a3}{2}} \left(\frac{I_{ds}}{W}\right)^{a3} exp\left(\frac{-E_{emi}}{K_B T}\right), \quad (2\text{-}48)$$

where the three terms represent the HC degradation in the high-energy mode, the medium-energy mode and channel cold carriers, respectively. C_1, C_2 and C_3 are constants.

Figure 2-25 shows the HCI degradation calculated by the equation (2-48). As we can observe, the experimental data have a good agreement even in transition regions between the different degradation modes. Note the strongly dependence of degradation with V_d.

Despite the fact that the published works on HC in domain of electronic design automation have been based on LEM, we consider that it is obsolete. Therefore, in the rest of this manuscript the multi-mode model will be used as base model of HCI degradation to generate all our results.

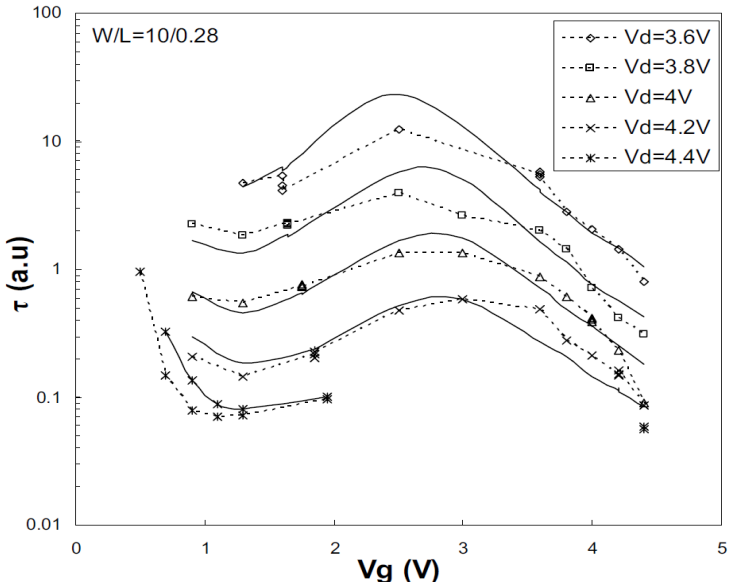


Figure 2-25 Comparison of multi-mode model (lines) with different regimes to experimental lifetimes (symbols) at five V_d different conditions [Guerin07b]

2.4 Electric Simulation

The behavior of an electric circuit depends on the individual behaviors of each one of the circuit elements. The electrical simulators use analytical models or *SPICE models* to represent the elements of a circuit in order to calculate their operation conditions. Based on electrical laws such as Ohm's law, Kirchhoff's law [Dorf06] simulators can carry out a complete electric analysis. Solving linear and non-linear equations, these simulators calculate currents and voltages of each circuit elements. Subsequently, this information can be observed to be observed in a graphical interface. In order to launch an electrical simulation or SPICE simulation, it is necessary to specify three input files: the circuit initial conditions, the circuit elements (Netlist) and the SPICE models.

2.4.1 Reliability Simulation

Figure 2-27, shows a typical reliability flow. First, a "fresh simulation" (before stress) is done with the fresh and original SPICE models using an analog circuit simulator such as ELDO (Mentor Graphics™), or HSPICE (Synopsys™). In the fresh simulation the circuit activities of devices are evaluated in order to calculate their impact on circuit lifetime. In the second stage flow an "age simulation" (which is divided in two parts: stress phase and characterization phase) is carried out with a new device model "a degraded SPICE model", which is calculated from NBTI/HCI transistor models

An example of NBTI degradation on an inverter (reliability electrical simulation) is shown in figure 2-28. Due to the complementary nature of n- and p MOSFETs only one transistor is active at each stable state of the input [0/1]. When *v1* (source voltage) is set to zero, the gate-to-source potential difference (V_{ds}) of the pMOSFET is equal at V_{dd} (one logic), which turns pMOS on. At the same time, the gate-to-source potential difference (V_{ds}) of the nMOSFET is equal at V_{ss} (zero logic), thus nMOSFET is off.

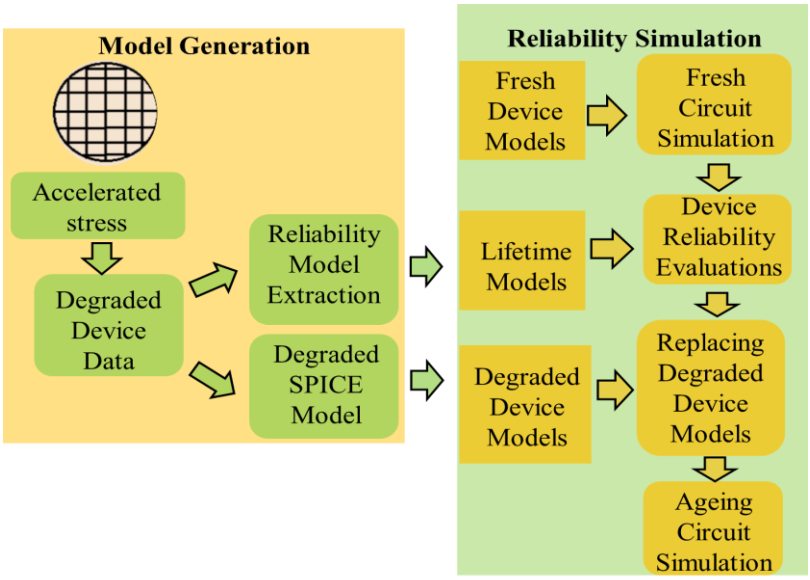
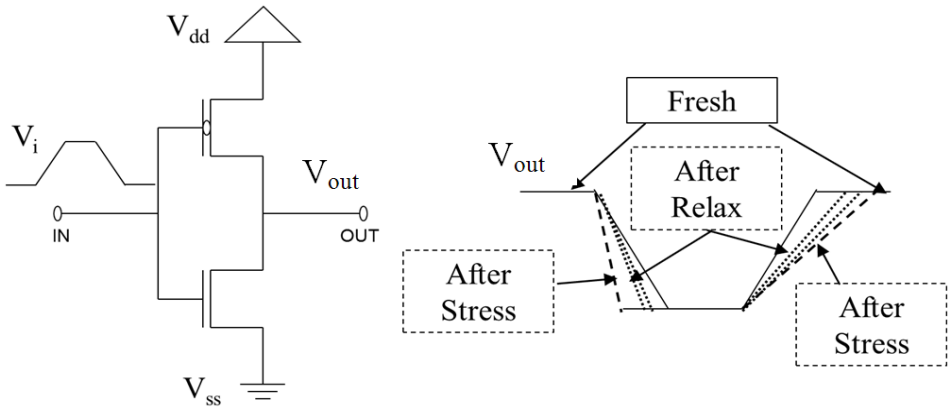


Figure 2-27 Reliability simulation flow and reliability model generation

Comparing the outputs before and after stress, it can be observed that the electrical response at input switching 1 to 0 (pMOSFET conducting) is faster than the electrical response at input switching 0 to 1 (nMOSFET conducting). This is provoked by the electrical degradation of

pMOSFET (§2.1.2), see figure 2-28b. Due to the NBTI recovery component the inverter electrical response will become at a state closer to its initial state.



(a) (b)

Figure 2-28 (a) Inverter schematic. (b) NBTI impact on inverter electrical response [Parthasarathy06]

In next section we will introduce the main characteristics of STMicroelectronics (STM) reliability simulation flow, which is used to carry out all reliability simulations presented in this manuscript.

2.4.2 STM Reliability simulation flow

As mentioned earlier in order to execute the reliability simulation is necessary to generate a degraded SPICE model. In the STM reliability flow this work is carried out by User Defined Reliability Module (UDRM). The UDRM is an interface application, which interacts with ELDO simulator to calculate the degradation of each circuit devices. In the first stage of the flow, ELDO computes the operating points of each circuit transistors. Subsequent this information is used by the UDRM to evaluate the amount of damage of each transistor in the circuit (ΔD). This evaluation is carried out with the same step times than ELDO simulations. This evaluation is updated at the end of every time interval and a new simulation is carried out. This process is repeated from T_0 (initial time) to T_{age} (final age), where the characteristics of the circuit are evaluated. These continual evaluations consume larger resources of CPU time. As consequence, a short transient simulation is chosen to be run. It represents a sample of the operating conditions, which is assumed to repeat through the circuit's lifetime. Finally, the value of the degradation is obtained by an extrapolation of these results at the total of circuit's lifetime in order to generate the degraded SPICE models.

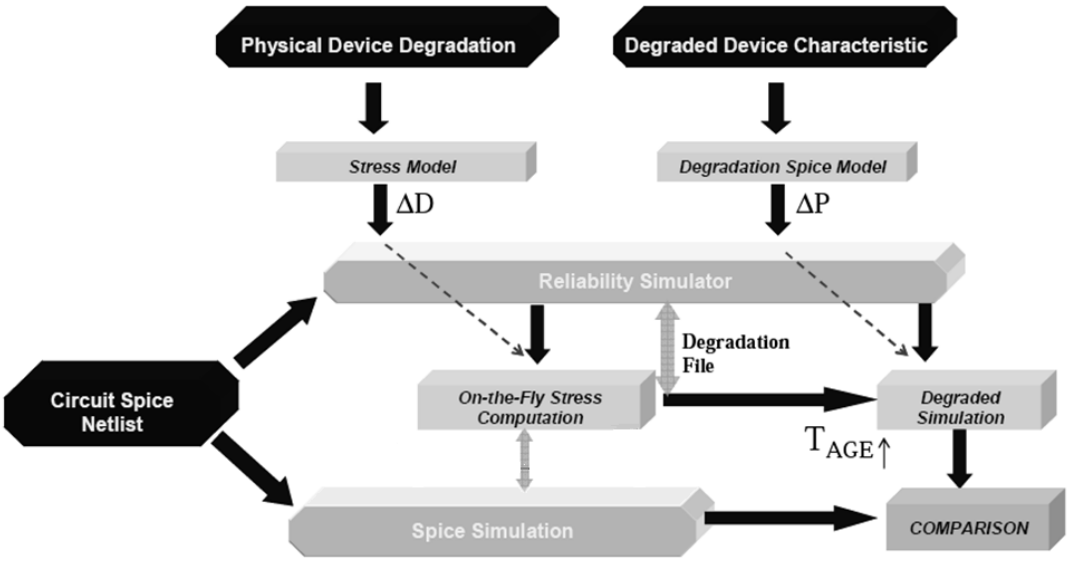


Figure 2-29 NBTI impact on inverter electrical response [Parthasarathy06]

Figure 2-29 illustrates the STM simulation reliability flow. ΔP represents the degradation of the Spice transistor model parameters and it is defined as follows:

$$\Delta P = f(\Delta D), \tag{2-49}$$

where ΔD defines the NBTI/HCI damage for each transistor in the circuit. The ΔD is defined as a function of several parameters such as (*operating voltages*, *operating currents*, *temperature*, *stress time*).

2.4.2.1 Age function

As mentioned above, a reliability simulation is performed from a degraded device SPICE model, which is calculated considering the dynamic operations of the circuit (such as stimuli, temperature, supply voltage, etc). However this practice involves the utilization of larger quantity of CPU and memory resources. In order to optimize this simulation process, only the transistor parameters affected by ageing phenomena are considered in the Age function (cf., Fig 2-30). The *Age* function describes the ageing of each circuit transistor i.e. degradation of V_{th} or μ. From this information the simulator calculates the degradation of the whole circuit.

The *Age* is a linear function with respect to the time, which is computed from the ΔD.

$$f(Age) = (\Delta D)^{1/n} \tag{2-50}$$

From equation 2-50, the dynamics degradation of Spice parameters can be defined as follows:

$$\Delta P = f(\Delta Age) = (\Delta D)^{1/n} \quad (2\text{-}51)$$

As we can see in figure 2-30, compact models have a larger quantity of parameters; this has increased the complexity level of electrical simulations as well as the CPU time. As consequence, the Age function solely considers the electrical parameters affected by NBTI/HCI.

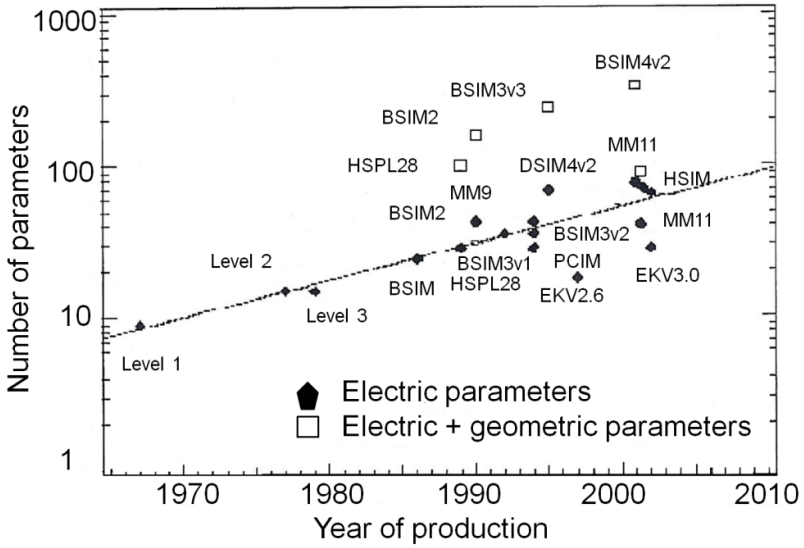


Figure 2-30 Compact models evolution

2.4.2.2 HCI degradation into reliability simulation

The model of HC degradation incorporated in the UDRM module is defined as follows [Guerin08] (§2.3.3.2):

$$\text{Age} = \text{Age}_1 + \text{Age}_2 + \text{Age}_3 \quad (2\text{-}52)$$

where Age1, Age2, Age3 describe the degradation for the high energy mode, medium energy mode and channel cold carriers mode respectively. Equation 2-52 can be also expressed as follows:

$$\text{Age} = \frac{1}{\tau} = \text{t} \cdot \left[C_1 \cdot \left(\frac{I_{ds}}{W}\right)^{a1} \left(\frac{I_{bs}}{I_{ds}}\right)^{m} + C_2 \cdot \left(\frac{I_{ds}}{W}\right)^{a2} \left(\frac{I_{bs}}{I_{ds}}\right)^{m} + C_3 \cdot V_{ds}^{\frac{a3}{2}} \left(\frac{I_{ds}}{W}\right)^{a3} exp\left(\frac{-E_{emi}}{K_B T}\right)\right. \quad (2\text{-}53)$$

Note that equation 2-52 is used to compute the HC degradation for each transistor of the circuit.

2.4.2.3 NBTI degradation into reliability simulation

According to the limited reaction model, the permanent component of NBTI degradation can be modeled for a power law in stress time and voltage which are translated in our modeling as:

$$\Delta D_P \approx K_P V_g^a t_s^b \tag{2-54}$$

Considering that the recovery component of NBTI follows a logarithmic law [Tewskbury92], this has been modeled as follows:

$$\Delta D_R \approx K_R V_g^c \cdot \ln\left(1 + \frac{t_S \tau_e}{t_R \tau_c}\right) \tag{2-55}$$

where t_S represents the stress time and t_R defines the recovery time.

The maximum degradation after a stress time is expressed (when no recovery is considered) as follows:

$$\Delta D_{MAX}(t_s) = \Delta D_P\,(t_s) \tag{2-56}$$

Therefore the total degradation after a stress time and a recovery time is defined as follows:

$$\Delta D_{NET}(t_s, t_R) = \Delta D_P\,(t_s) + \Delta D_R(t_s) - \Delta R(t_s, t_R), \tag{2-57}$$

where ΔR represents the recovery, which depends on the recovery time.

2.5 Silicon Validation of NBTI and HC models

In the first part of this section we present some examples of ageing analysis (simulation data), which are compared with experimental data allowing us to validate both NBTI and HCI reliability models at transistor level. The figures presented in the subsections 2.5.1 and 2.5.2 are taken from the manuscript referenced in [Parthasarathy06].

In order to validate delay degradation predictions of the HCI and NBTI models we must increase the level of abstraction from transistor to gate level. In the second part of this chapter

various ageing analysis are carried out on ring oscillator circuits and SRAM memories, the respective CAD-Silicon delay comparisons are illustrated.

2.5.1 Transistor characteristics: HC mechanism

2.5.1.1 CHC substrate current

Figure 2-31 illustrates the comparison between the simulation results (ELDO simulations) and the experimental measures. As we can observe, the correspondence between the CAD data and silicon data is good for this study case (2.1nm gate oxide nMOS (10mm/0.1mm) in 90nm technology).

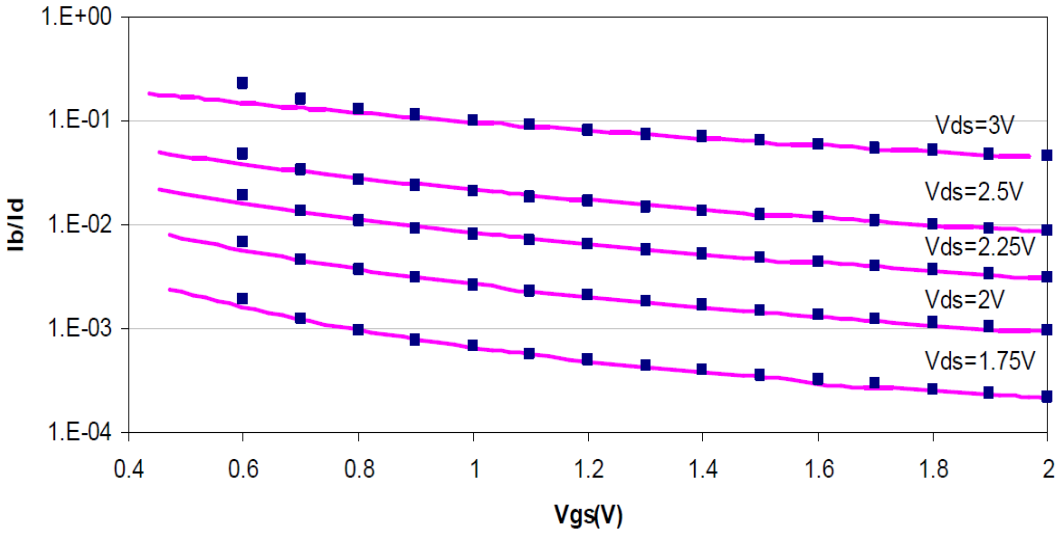


Figure 2-31 I_b/I_d matching between SPICE simulations and silicon data. Symbols represent silicon measures and the lines correspond to simulation data

2.5.1.2 CHC I-V Characteristics

For the same device as above, the figure 2-32 shows the results before and after degradation at a stress conditions of V_{gs}= V_{ds}= 2.1V compared to silicon results after 12000s of stress. Note that there can be some mismatch between simulation and silicon measures even on fresh device, but the amount of degradation is seen to be the same in both silicon experiments as well as ELDO simulation.

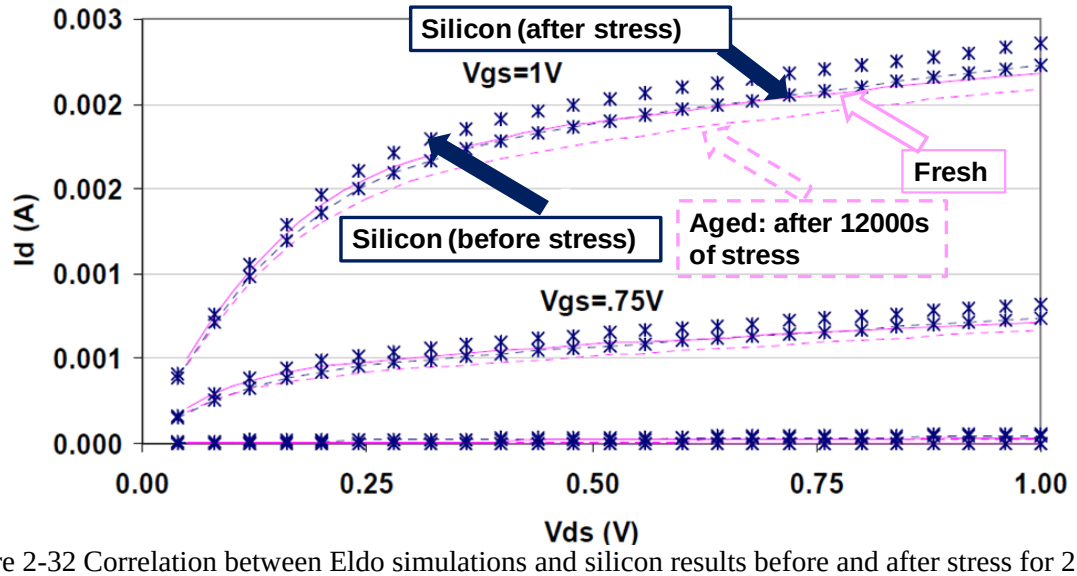


Figure 2-32 Correlation between Eldo simulations and silicon results before and after stress for 2.1nm gate oxide 10/.1 mm NMOS: I_d-V_{ds} for large V_{gs}

2.5.2 Transistor characteristics: NBTI mechanism

Figure 2-33 shows the comparison between silicon measures data and the model. In order to validate NBTI model over a large range of frequencies, a DC NBTI stress (at a constant V_{gs} =-1.5V, 125C) and an AC NBTI stress (0 to -1.5V, 100KHz, 50%SP[3], 125C) were carried out. The model previsions (lines) are shown to coincide with the silicon data (circular patterns).

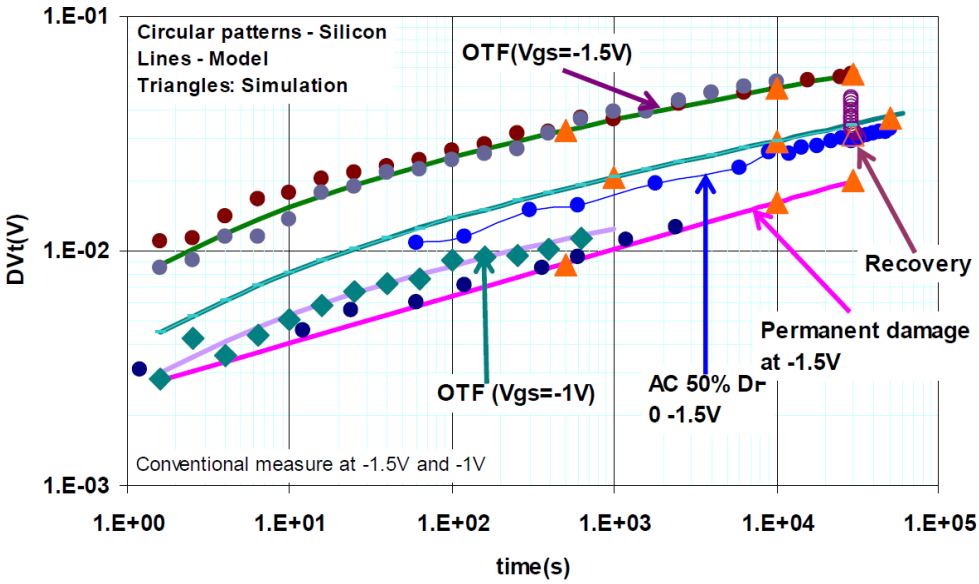


Figure 2-33 Correlation between simulation results (triangles) and silicon data (data points) [Parthasarathy06].

[3] Signal Probability (SP) (§3.2.1.3)

2.5.3 NBTI validation: Ring Oscillators

Several standard cell-based ring oscillators fabricated in 45 nm, were tested at different stress conditions, including different voltages, temperatures, etc. (cf., Fig. 2-34). The resulting drifts (open symbols) are compared to reliability simulations run on the netlists including all parasitic capacitances, resistances, etc.

Two main conclusions are arising from this analysis. First, the L-R model (full black lines) is very accurate in reproducing experimental frequency drifts which allows us to consider the transistor-level model results as silicon results in the rest of the manuscript. The second conclusion is that the R-D model (dashed lines) is pessimistic by about 40%, due to its inaccurate recovery dynamics modeling.

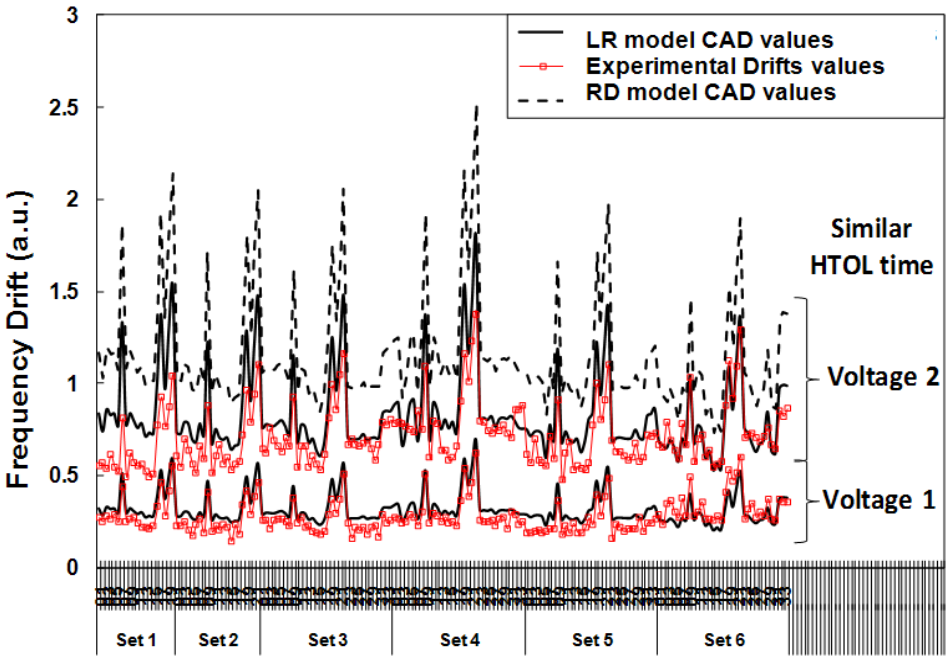


Figure 2-34. Frequency drifts from 200+ standard cells-based ring oscillators separated in various sets measured on silicon

2.5.4 NBTI and HCI validation: Static Random Access Memory (SRAM)

Figure 2-35 shows a typically SRAM memory, which is mainly divided in 4 parts. Control block leads the control signals i.e. clock, enable, reset, etc. Decoder block selects the address inside memory array for both read and write cycles. Memory array block is composed by all

memory points. Finally Input/Output block is the communicator channel of the SRAM. The operation in the clock cycle is calculated by the control block. When read or write cycle is performed, the address is chosen by the decoder block, subsequently the right word is chosen inside the memory array. In parallel, the input/output block (I/O) collects the data from the memory array for a read cycle or from outside (environment memory, i.e., processor bus) the new data to be saved in the memory array (write cycle).

A reliability analysis was performed in the SRAM with the purpose to identify the most sensitive SRAM part to degradation using our in-house reliability modeling solution (§2.4.2.2), (§2.4.2.3) [Huard10a]. In parallel, an ebeam analysis was performed on silicon in order to cross-check simulations results with silicon measures. The ebeam analysis has been done on dedicated pads connected to inner nets at the interfaces of main blocks to obtain timing correlations (before and after stress) at block level. Electrical stress on silicon was performed at-speed in order to simulate real operating conditions. In order to speed-up the simulations[4] only the bottom and top row memory array are considered. We kept the worst case critical paths (slowest). Note that accuracy of the simulations is kept because the not used memory points are modeled by the equivalent static loads.

Two sets of electrical simulations were performed. First, only NBTI phenomenon was enabled. And, second, NBTI and HCI were considered. Table 2-1 shows CAD[5]-Silicon timing correlation calculated on the SRAM critical paths. All values presented in this table are expressed in picoseconds. Note that the precision of ebeam measure is +/- 50ps, which is equivalent to a relative error of 5.01%. Timing paths are divided in seven parts. Each part has an initial point and an end point. These points were chosen in order to represent both the write and the read cycles. In figure 2-35 these points are illustrate by the numbers [(1)-(8)]. For instance, the operation read “0” at bottom row cycle starts in the point (1), which represents the CLK node (clock) in the control block. The clock signal enables seven sub blocks (designated by the points 2, 3, 4, 5, 6, 7, 8), which allow collecting in the right memory point the data asked.

[4] For instance in a SRAM (512 words x 32 bits) the memory array is composed by 16384 memory points. As a consequence the simulations have become slow since the larger number of transistors.
[5] The CAD data used in this correlation were calculated, taking into account both reliability phenomena (NBTI and HCI).

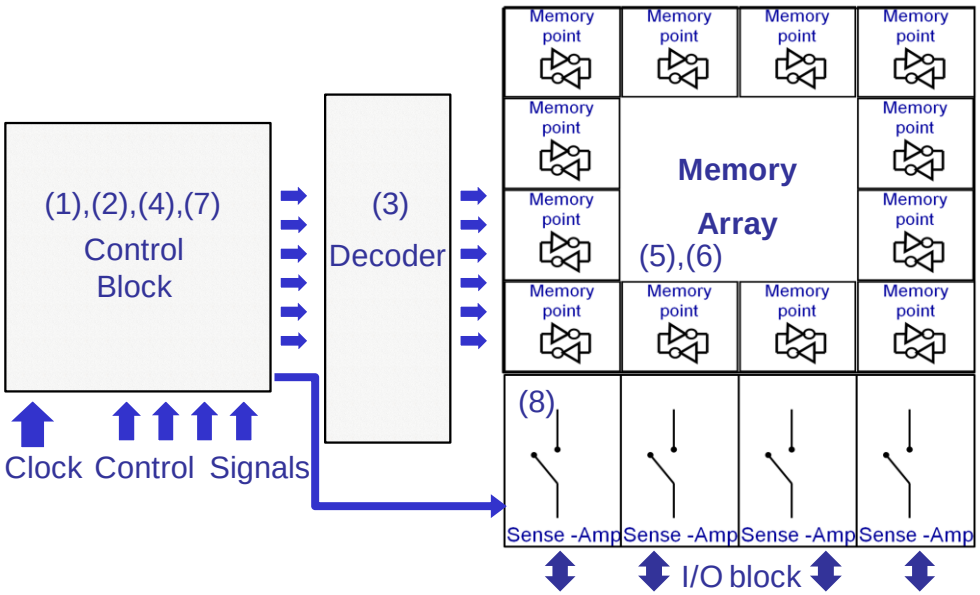


Figure 2-35 SRAM schematic including wrapped control logic

On the other hand, the write cycle starts in the point (1) and ends in the point (6), which represents a point memory. In the table 2-1 the first column illustrates the operation under consideration. Second column shows both the initial and the end points of the timing path. The columns three and four show the comparison between the CAD and Silicon data at two points, before and after the stress, respectively. Finally the columns five and six show the relative error of Eldo simulations with respect to silicon measures (fresh and ageing, respectively). Let us note that this table is a simplified representation of the real SRAM data table, which is shown in the Appendix B.

Three main conclusions are arising from this analysis. The first conclusion is that the degradation was generated largely by NBTI; however the HCI contribution is not negligible (in the experimental experiences). Note that the electrical stress was performed with an external clock at high frequency in order to reproduce a real activity, thus the HC mechanism was activated.

In spite of some cases the relative error reaches a maximum of 16%, we considered that the timings monitored by ebeam technique on the SRAM are well reproduced by ageing simulations. This error can be explained by the fact that the process variability of the silicon wafer in which the measures were done was not properly center with respect to the CAD corners defined for this technology node. Figure 2-36 illustrates a typical case of the difference between CAD and Silicon corners. As we can clearly observe the definition of the CAD corners depends on several factors such as the die variability, wafer variability, etc.

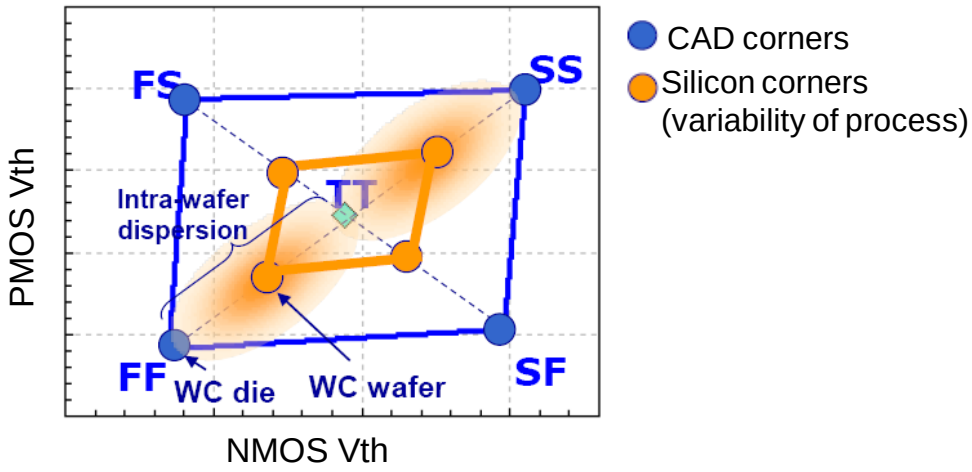


Figure 2-36 Comparison between the CAD and Silicon Corners (defined with respect to threshold voltage of pMOS and nMOS devices) of a given technological node

The third conclusion is that the control logic block was the more impacted by the degradation than the memory points.

OPERATION	TIMING PATH	(Silicon/CAD)		Relative Error(%)	
		Fresh(t=0)	Ageing	Fresh(t=0)	Ageing
WRITE 0 AT BOTTOM ROW	(1) to (4)	970/835	936/880	-13,9	-6,0
	(3) to (6)	1410/1330	1269/1370	-5,7	8,0
READ 0 AT BOTTOM ROW	(1) to (4)	630/601	774/641	-4,6	-17,2
	(3) to (8)	1450/1459	1341/1510	0,6	12,6
WRITE 1 AT TOP ROW	(1) to (4)	885/822	873/862	-7,1	-1,3
	(3) to (6)	1230/1314	1188/1353	6,8	13,9
READ 1 AT TOP ROW	(1) to (4)	630/604	747/644	-4,1	-13,8
	(3) to (8)	1475/1457	1305/1509	-1,2	15,6

Table 2-1 Silicon measures versus CAD data (Eldo simulations) for a SRAM

2.5.5 HC validation: RO

Figure 2-37 shows frequency evolution of 27 RO, which are composed by different cells such as Inverter, NOR and NAND. These RO were carried out in 40nm technology and were stressed at several conditions (including different stress voltages and stress times) [Huard12]. In order to favor the HCI degradation all stresses were performed at 25°C. As it can be

observed in the next figure, the silicon results (dash line) are in agreement with CAD results (full line). On the other hand, we observed that each standard-cell based RO presents a different response to reliability stimuli. This can be explained by the fact that the standard cells (*i.e.*, NOR, NAND) have different design (defined as physical configuration of transistors in the design), which react differently at ONE logic or at ZERO logic (§3.2.1.3). For instance is widely known that the NOR gates are more sensitive to NBTI degradation (because they have pMOS stacks). Therefore they are more sensitive to a stimuli low (ZERO logic) than the inverters.

It is worth to notice that the RO do not have the same number of elements, for example the smallest RO has 8 stages and the bigger has 134 stages.

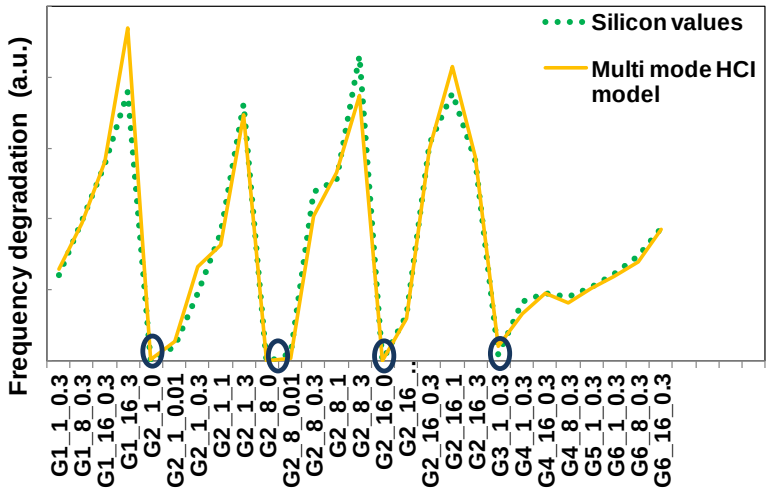


Figure 2-37 Frequency drifts from specific delay chains stressed at various frequencies ranging from 10Mhz to 3GHz. Some chains (blue circle) are blocked (i.e. no oscillation) during life tests [Huard12]

2.6 Summary

In this chapter, we have reviewed the physics of the NBTI and HCI degradation. We have illustrated the main analytical models for NBTI and HCI wearout mechanisms throughout of the last decades. In particular, we have explained the main characteristics of NBTI L-R model and the HC multimode model (in-house models). In the last part of this chapter we presented the silicon validation of the NBTI L-R and HC multimode models at transistor level. Additionally, a CAD-Silicon timing checking section was presented in order to validate the delay degradation predictions of these two models. In the next chapter we will describe the NBTI and HCI impact at gate level.

Chapter 3

Gate Level Reliability Modeling

In the first part of this chapter we address NBTI and HC reliability effects at gate-level. The influence of design parameters on the standard cell electrical performances is explained in detail in the sections 3.1 and 3.2.

Section 3.3 presents the characterization flow of standard cells focusing in the timing characterization aspects. In sections 3.5 and 3.7 we present a novel methodology to model NBTI and HC phenomena at gate level.

3.1 Overview of gate level reliability

The layouts of both analog and digital basic cells are designed in a compact manner in order to obtain a best optimization in terms of area, timing and power consumption. Each layout is packed into an abstract representation called a *standard cell* (SC) and a group of these cells form a *standard cell library*.

In order to understand how the NBTI and the HC mechanisms affect the electrical performance of the digital standard cells, it is necessary to take into account in the reliability analysis all design parameters of the circuit. Some examples of design parameters are: input signal waveform transition time (fall/rise edge), load capacitance, input signal frequency, etc.

The delay is the main characteristic analyzed during the reliability characterization, since NBTI and HC solely affect the timing performance of standard cells. Multiple designs factors, as well as the individual transistor ageing degradation impact the delay propagation of the signals through the combinational and sequential gates. Before introducing the impact provoked by the NBTI and the HC phenomena on the gate delay, we present an overview of the delay modeling of gates.

It is widely known that the propagation delay follows a nonlinear behavior with respect to the capacitance and the operating conditions. During the last two decades several delay analytic models have been proposed in [Wang00], [Jeppson94], [Maurine02]. In order to explain the propagation delay concepts in detail, we present the delay analytical model for a CMOS inverter proposed by Lasbouygues in [Lasbouygues06]. Despite the fact that the illustration of this approach is only based on the CMOS inverter, most of the formulations and performance factors can be used to other cells (i.e. NAND and NOR). In the following paragraphs, delay model general concepts and its parameters are described.

3.1.1 Propagation delay model at gate level

According to [Mead80], the basic switching process of a CMOS inverter cell can be considered as an exchange of charges between the intrinsic cell and its output loading capacitance. The output transition time (defined as the transition time of the ideal inverter

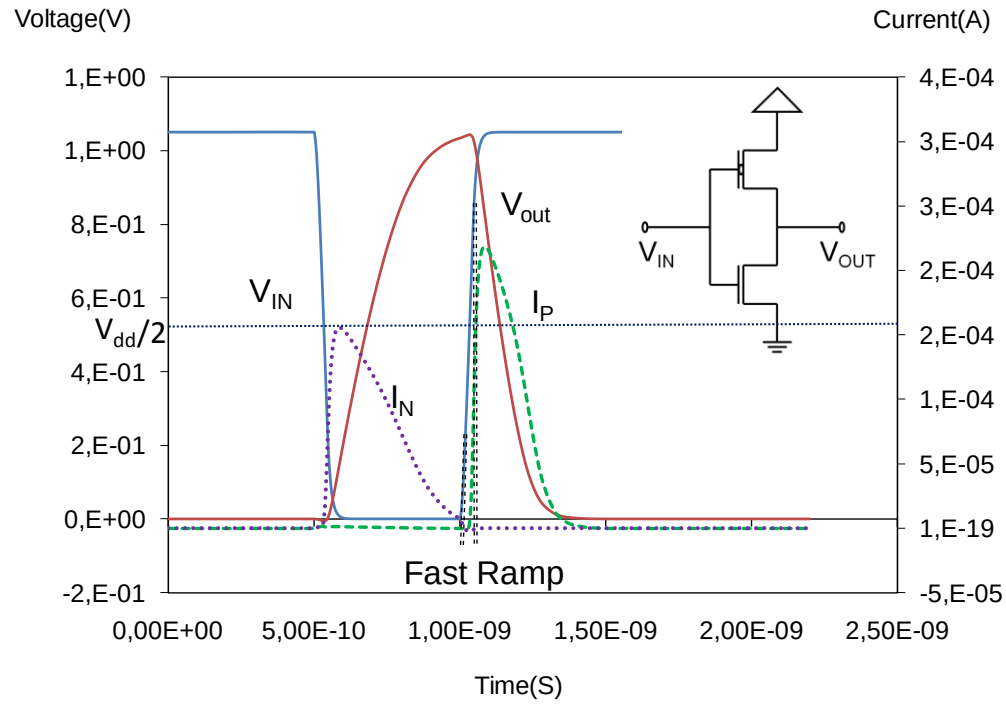
output signal voltage (V_{OUT}) can then be directly obtained from the model of charging/discharging current process that occurs during the switching of the cell between two stables states [High-to-Low (HL) or Low-to-High (LH)] and from the amount to charge (C_{load_total}*V_{dd}) to be exchanged with the output node such as

$$\tau_{OUTHL} = \frac{C_{load\_total} \cdot V_{dd}}{I_{Nmax}},$$

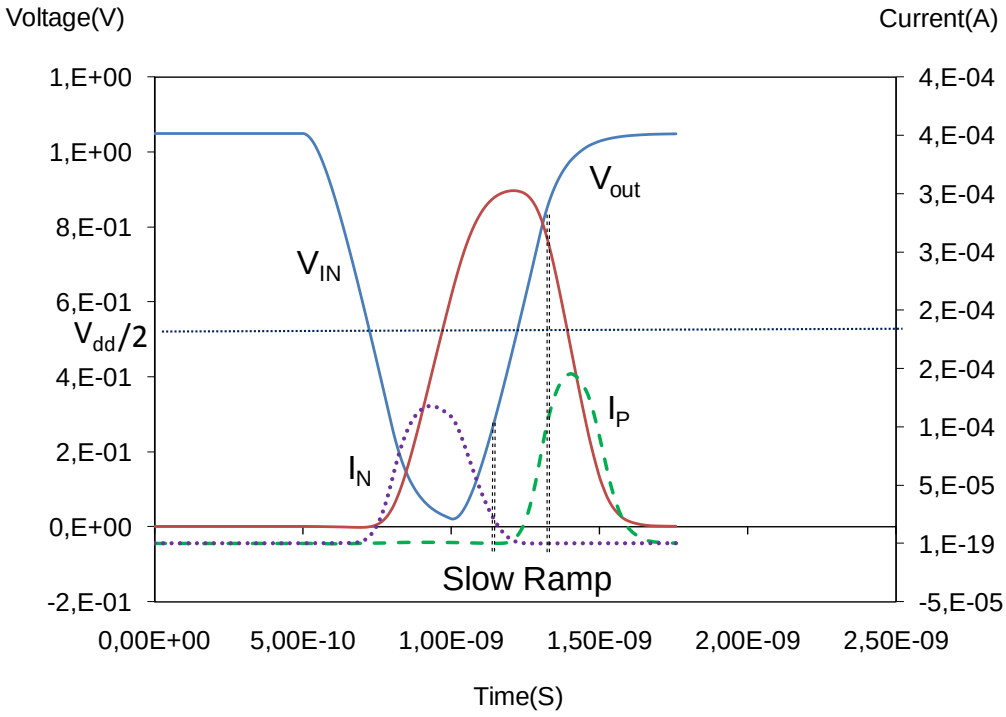
$$\tau_{OUTLH} = \frac{C_{load\_total} \cdot V_{dd}}{I_{Pmax}}, \quad (3\text{-}1)$$

In the equation 3-1 C_{load_total} is the cell output total capacitance, which is given by the sum of internal load, C_{INT}, and fan out load, C_{EXT}. The first term C_{INT} is the sum of the coupling capacitance, C_M, between the inputs and output nodes [Tsividis03] and of the transistor diffusion parasitic capacitance C_{DIFF} [Taur98]. Finally I_{Pmax} and I_{Nmax} are the maximum values of the discharging currents in the inverter output node generated by the pMOSFET and nMOSFET devices, respectively.

In order to extend the ideal output transition time metric to the real domain (delay), it is necessary to consider the dynamic of the inverter electrical currents. Following [Sakurai90], the current flow supplied by an inverter during the switching process depends of the slope input signals (slow or fast). In figure 3-1, we can observe that the switching current is not constant. In fact, its dynamic evolves over time according to both the charge and the operation conditions. Figures 3-1a and 3-1b show the waveform of the inverter input voltage υ_{IN}, for a fast transition time *(fast ramp)* and for a slow transition time (*slow ramp*), respectively.



(a)



(b)

Figure 3-1 Voltage and Current evolution as a function the time for a CMOS inverter, showing an example of a slow ramp and fast ramp

For the charging up-case (high-to-low switching), we can observe that for the fast input signal ramp (cf., Fig 3-1a), V_{OUT} reaches the V_{dd} level very fast, since the short circuit period is shorter [Daga99]. The short circuit inverter mode has a stronger dependence on the input

ramp signal; the faster the input ramp, the sooner the N transistor will be OFF and the P transistor will be able to charge the output capacitance with a maximum current. For the slow slope input signal ramps, the short circuit period is much longer than for the fast input ramps. Thus the discharge current is much lower (cf., Fig. 3-1b). Taking into account all these considerations, the delay of an inverter cell can be defined as follows [Lasbouygues06]:

$$d_{HL,LH} = \frac{\tau_{IN}}{\tau_{out\,HL,LH}^{Fast}} \cdot \frac{v_{thN,P}}{2} + \frac{2\alpha_{HL,LH}}{\left(\alpha_{HL,LH} + \frac{C_{DIFF}}{C_{IN}}\right) + \frac{C_{load}}{C_{IN}}} \cdot \left(1 - A_{HL,LH}\frac{\tau_{IN}}{\tau_{out\,HL,LH}^{Fast}}\right) + \frac{\tau_{out\,HL,LH}}{2\tau_{out\,HL,LH}^{Fast}} \cdot \left(1 - A_{HL,LH}\frac{\tau_{IN}}{\tau_{out\,HL,LH}^{Fast}}\right), \quad (3\text{-}2)$$

where $\alpha_{HL,LH}$ is the Meyer coefficient [Tsividis03], τ_{IN} is the input signal ramp, $A_{HL,LH}$ is a fitting parameter, V_{th} is the threshold voltage (N or P transistors), C_{IN} is the input capacitance, C_{load} is the load capacitance, C_{DIFF} is the diffusion parasitic capacitance and $\tau_{out\,HL,LH}^{Fast}, \tau_{out\,HL,LH}$ represent the transition time for the fast and slow input signal ramps, respectively.

3.2 Impact of ageing phenomena on standard cells delay

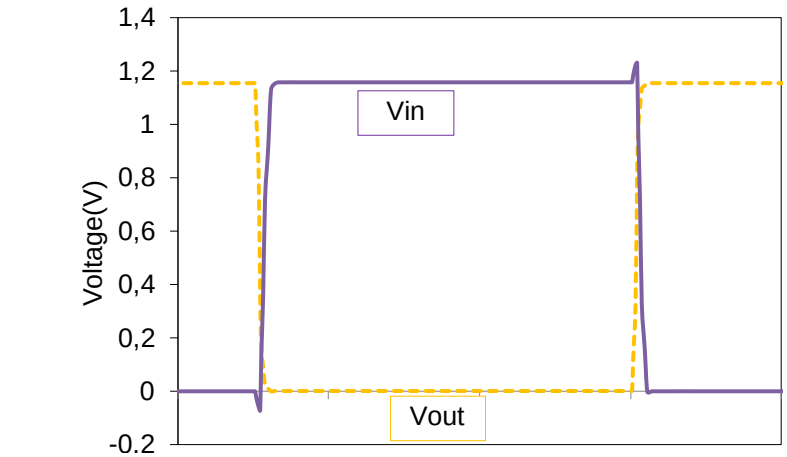
It was already shown in the chapter 2 that induced-NBTI/HC damage in CMOS devices is caused by the degradation of physical parameters such as mobility (μ), flatband voltage (V_{FB}), etc, which in turn affect the electrical performance of transistors. Therefore the electrical performance at standard cell level is also impacted; the performance can be degraded or improved depending of the standard cell environmental factors (operation conditions and design parameters). In order to illustrate how these factors modify the standard cell degradation behavior, an ageing analysis was carried out, where both values of the interface states shift (ΔN_{it}) and its rate of generation during the stress phase were monitored.

3.2.1 NBTI and HC degradation dependence on Signal Probability, Input Ramp and Load Capacitance.

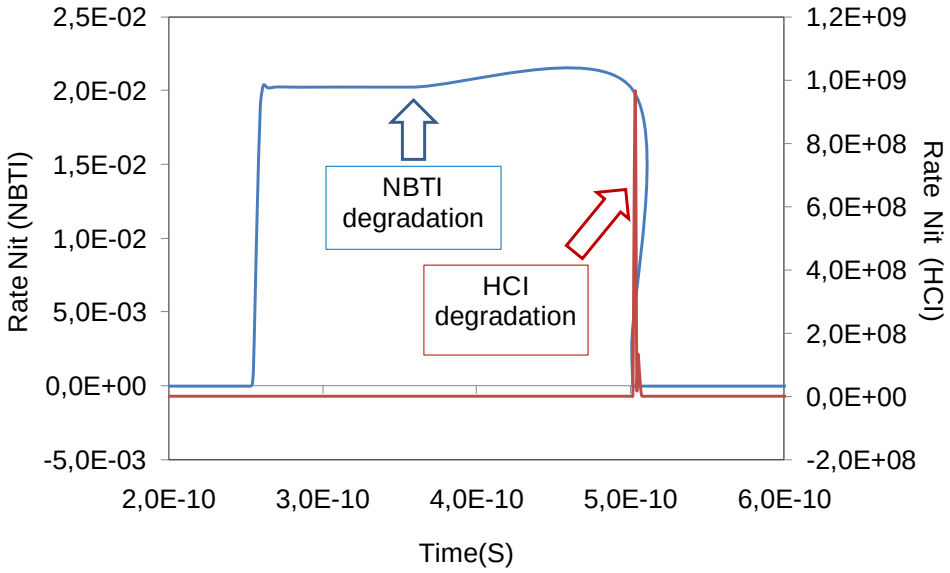
In the first part of this subsection we will use the *stress monitor* function [Parthasaraty06], to describe the differences between the NBTI and HC behaviors during the stress phase.

Subsequently we will present the ageing evolution with both design parameters and operation conditions.

Figure 3-2b illustrates the N_{it} generation rate for both NBTI and HCI stress calculated on an inverter (45nm technology) by the transistor models (§2.4.2.2), (§2.4.2.3) using the *stress monitor* function of the *Eldo* simulator. As it can be observed, the degradation HC is produced during the toggling input signal, in contrast to the NBTI degradation that appears when the input signal *Vin* is stable (cf., Fig. 3-2a), more specifically during the high input state. It is worth to note that in order to generate HC damage it is necessary to have a flow of electrons in other words a current, on the contrary NBTI can be caused solely by a voltage.



(a)



(b)

Figure 3-2 (a) Input signal form applied on the inverter under consideration during the phase stress. (b) HC and NBTI rate ΔNit during the stress phase of the reliability simulation

Considering the difference between the HC and NBTI behaviors, the next section will address the influences of input ramp for both phenomena. Signal probability (§3.2.1.3) impact will be analyzed only for NBTI. And, finally the Fan-Out effect on the delay will be consider.

3.2.1.1 Input slope signal on delay degradation

Let us take an example of combined HC and NBTI degradation for an inverter (pMOS transistor is impacted by NBTI degradation and the nMOS device is affected by HC) with a ratio of Wp/Wn=2, in 45nm CMOS technology. Ageing simulations results are shown in figure 3-3.

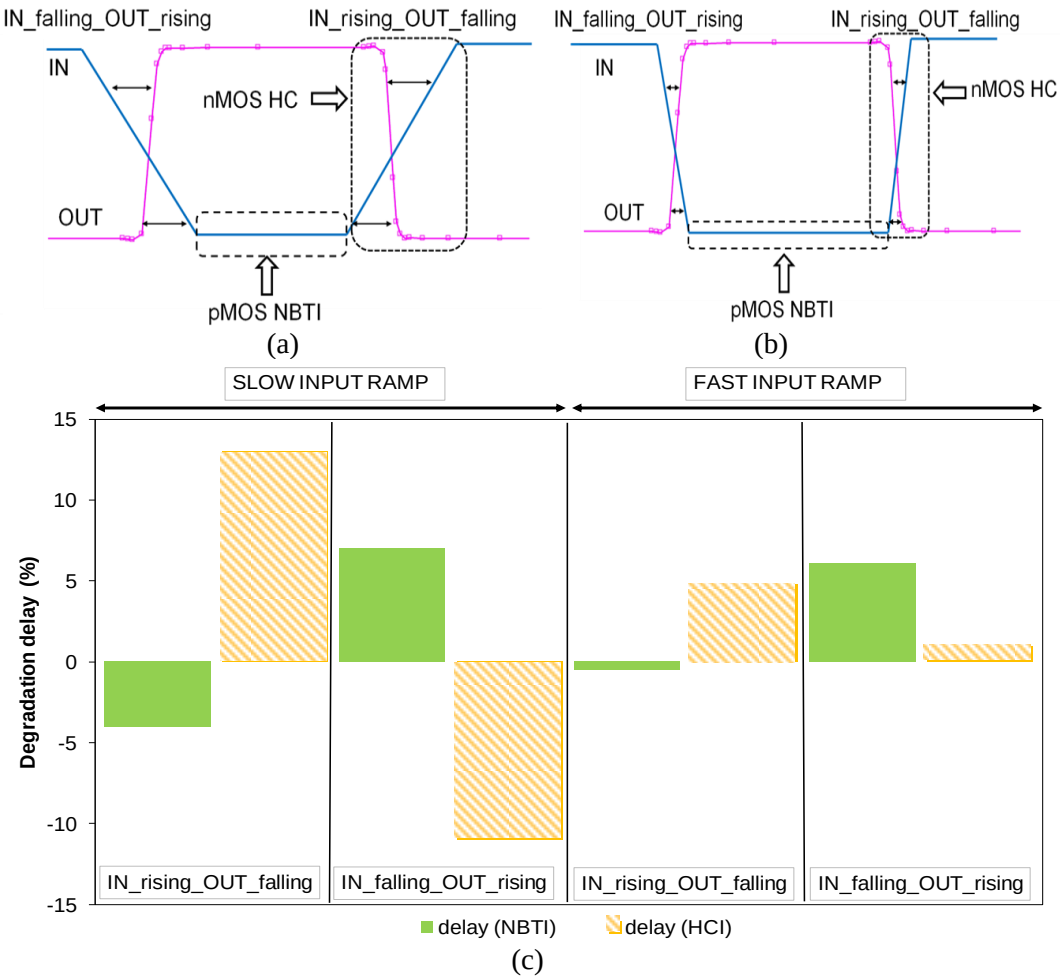


Figure 3-3 Delay shifts provoked by HC on nMOSFET and NBTI on pMOSFET showing that the degradation is strongly dependent of the circuit operation conditions. (a) Slow input ramps. (b) Fast inputs ramps. (c) Summary of total degradation (NBTI+HCI) observed on an inverter [Parthasaraty06]

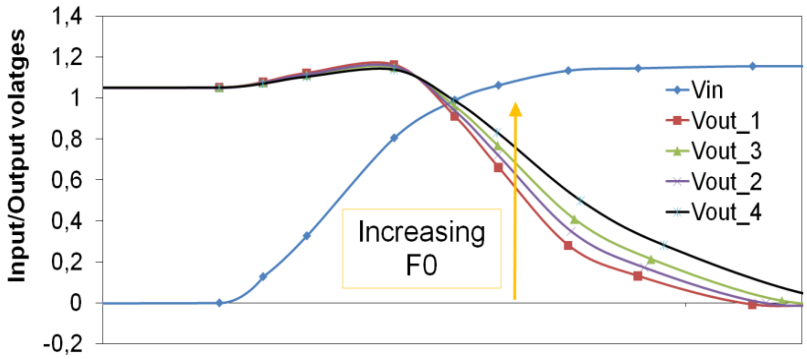
It can be observed that the total degradation depends on the input signal ramp. If the ramp is slow (cf., Fig. 3-3a) NBTI and HC effects on delay are opposite, since the two devices (P and

N transistors) are in saturation zone (ON state). In this case the two transistors are affected by the degradation. In order to analyze in detail the effect of each phenomenon in the simulation tool we enabled individually each one of the degradation modes. First, let us consider only the NBTI mode (i.e., pMOS device is degraded). The output fall transitions become faster than before (negative change) since nMOS can discharge faster and the output rise transitions are slower, because the PMOS is weaker. On the other hand, if HC mode is considered only, the situation is inverted. Thus the output rise transitions become faster.

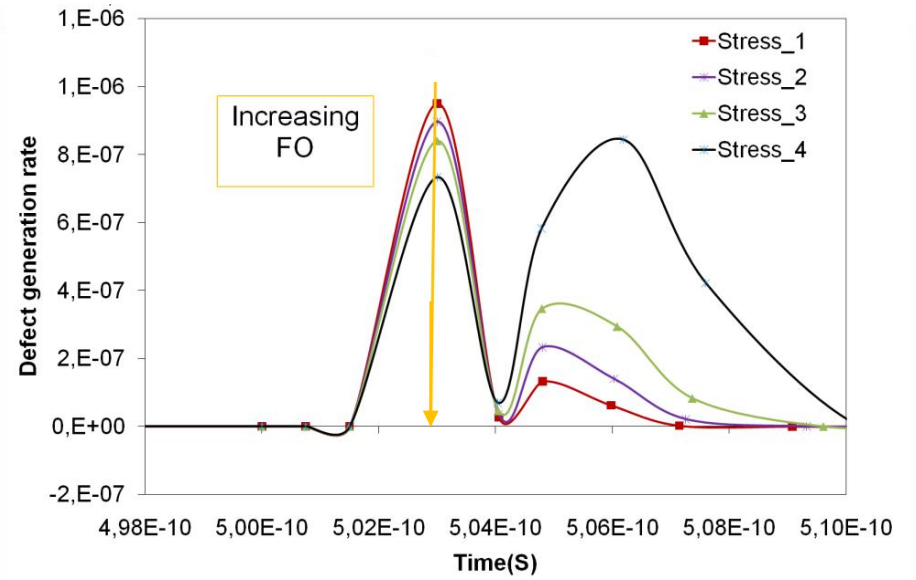
For the fast ramp case of input signal (cf., Fig. 3-3b), the ageing impact on delay is only determined by one transistor because the fast input ramps turn ON a transistor and OFF the other one. In this case, the total degradation in delay is dominated by the damage generated on the stressed transistor, the nMOS in the case of the input toggling down and the pMOS in the case of input toggling up. Figure 3-3c illustrates the difference in the levels of NBTI and HCI degradation for fast and slow input ramps. As we observe in this figure, the HC damage dominates on the NBTI damage for the slow input ramps. On the contrary for the fast input ramps, the HC and NBTI damages have almost the same level. Note that the NBTI degradation increase is due to the augment of duration of the input signal low state.

3.2.1.2 Fan–Out impact in the degradation

Figure 3-4 shows the impact of the Fan-Out (defined as the ratio of load capacitance to input capacitance, FO= C_{load}/C_{in}) on the HC degradation for an inverter carried out in 32 nm CMOS technology. As discussed by [Leblebici93], [Huard09] the HC physical damage has a strong dependence on the transition shape, which in turn depends of FO and input ramp. As we can see in figure 3-4b, as the FO increases, the defect generation rate augments too.



(a)



(b)

Figure 3-4 FO impact on HC physical damage. Simulation was carried out on an Inverter cell

Note that in this figure two peaks appear. This is explained by the fact that the hot carriers can be generated in three different manners. The first peak is produced by the hot carriers generated by the mode I (low value of drain current) and mode II (moderate values of V_d and V_g) while the second peak is caused by the mode III (§2.3.3.4).

Figure 3-5 shows the delay degradation caused by the NBTI versus the FO for a group of five inverter cells in 45nm CMOS technology. As we can observe in this figure the NBTI degradation has a weak dependence of FO. This is explained by the fact that NBTI damage is generated only during the input signal stable states (the low state) and considering that the value of the FO slightly affects the duration of the low state signal, we can assume that NBTI is not dependent of FO. Note that the delay degradation increase is due to the driver of the inverter under consideration. As the inverter driver (defined as physical size of the cell) decreases, the electrical capacity to charge a large load capacitance diminishes.

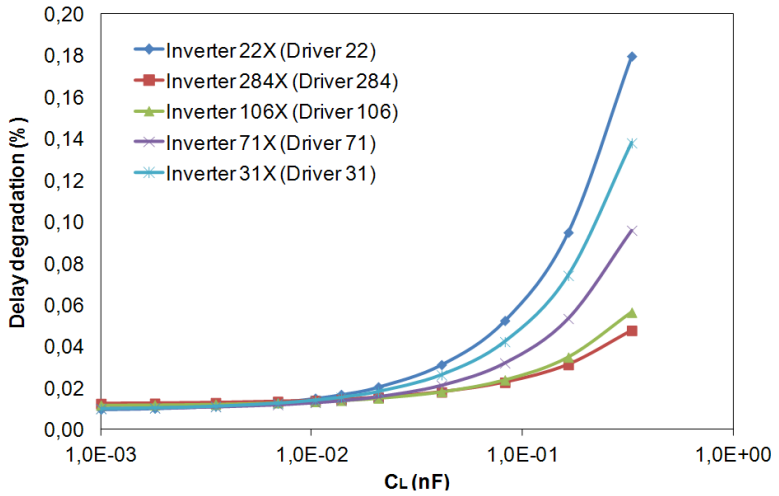


Figure 3-5 Representation of delay degradation versus CL. This result allows assuming non-NBTI degradation dependence with FO

Figure 3-6 summarizes the effects of the input signal ramp (S_I) and the load capacitance (C_L) on the delay degradation provoked by HC phenomenon on an inverter (45 nm CMOS technology). As we can see, the HC damage on nMOS device is stronger for the small values of the C_L, and the large values of S_I.

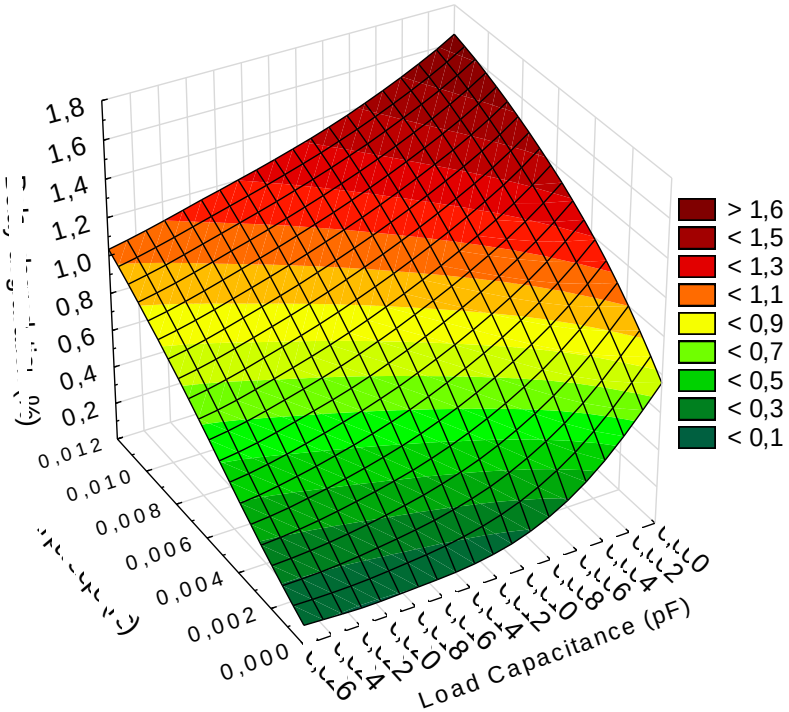


Figure 3-6 Degradation of nMOSFET in a CMOS inverter cell (45 nm) as a function of input ramp and capacitance load. For this case the HC impact on delay was calculated for a time *t* =10 years.

3.2.1.3 NBTI delay degradation dependence on Signal Probability

As it is widely known, the NBTI delay degradation is strongly dependent on input cell signal probability [Sanjay07]. On the contrary, HC degradation is not influenced by signal probability. Figure 3-7 represents an example of inverter delay degradation dependence with the signal probability (defined as the probability that the input signal of a cell is high and denoted by SP) at given S_I and C_L values. The SP provokes a maximum damage at SP100% and a null damage at SP0% (for the inverter cells). Overall, the curve shows an upward trend of degradation with SP.

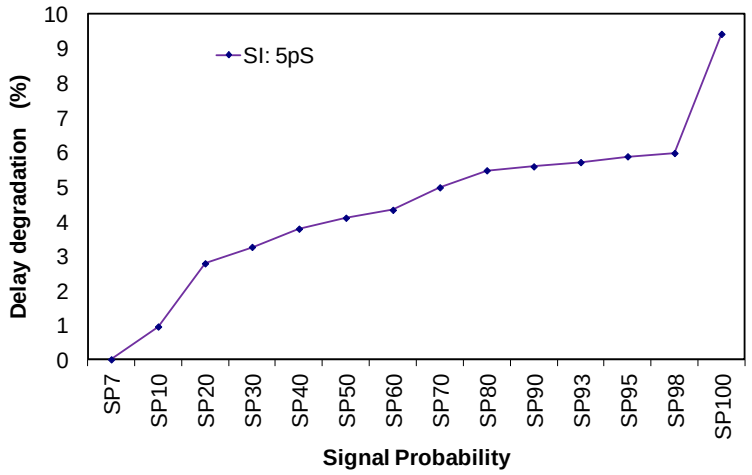


Figure 3-7 Delay degradation (input vector: IN_falling_OUT_rising) as a function of SP. For a 45 nm CMOS technology inverter

This can be explained by the fact that a high value of SP involves a larger period of stress and a lower period of recovery. Let us compare the extreme points of the figure (SP10 and SP100). The difference between the degradations of the two points is larger, approximately 8%.

Figure 3-8 shows the delay degradation dynamic for several S_I values at C_L given. It can observe that the upward trend is conserved. Here, the degradation dependence on S_I is clearly observed.

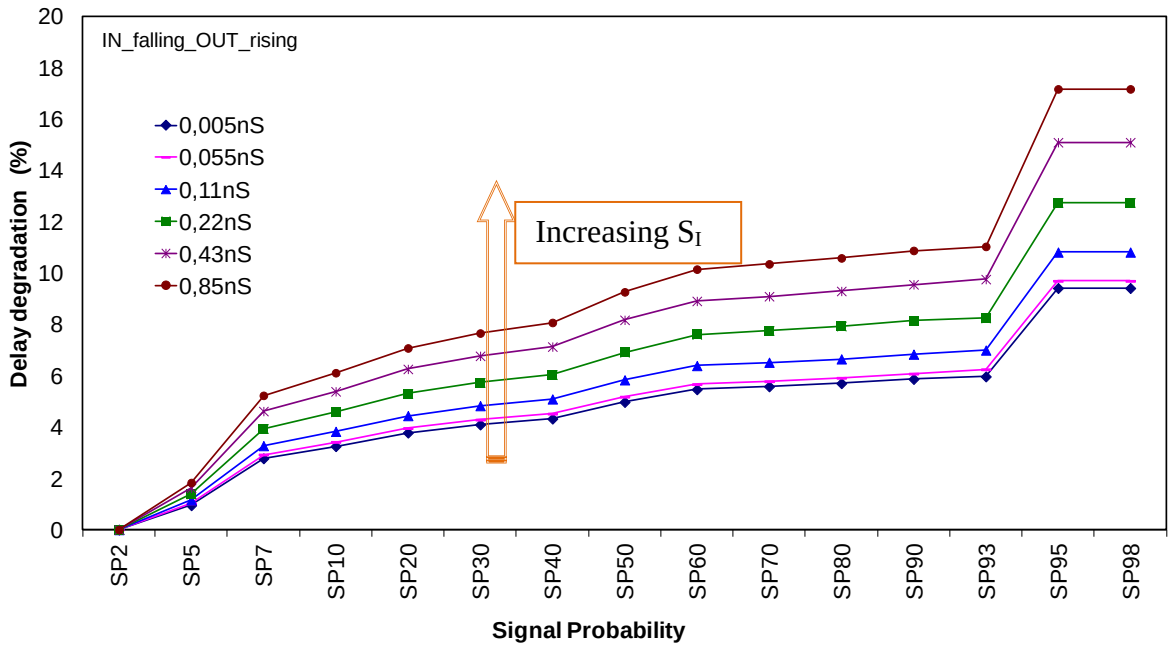


Figure 3-8 Rise delay of an inverter as a function of SP, including the effects of S_I (45 nm CMOS technology)

3.3 Standard Cell Library Characterization Flow

Standard cell library characterization flow generates a complete description of standard cells electrical behavior. Timing (delay performance) and power consumption are characterized for a given technological node specifications, SPICE models and operating conditions. In this manuscript we focus on timing characterization. The standard cells are characterized by applying a ramp voltage source at the input of the cell and an explicit load capacitance at the output of the cell. Then several circuit simulations (*Eldo in our case*) are executed varying the transition time of the ramp and the value of the load capacitance in order to calculate both the propagation delay and the output transition time (or output signal ramp) for each pair of (S_I, C_L). These results are stored in a table, which is used in the subsequently design steps such as the synthesis, place and route, etc (§4.1).

The PVT (Process, Voltage, Temperature) corner approach is widely used in timing library characterization. It is widely known that the timing performance dynamic depends strongly on this triplet. This method evaluates the performance of cells in three configurations: worst case, nominal case and best case in order to validate the timing performance for all possible combinations of the PVT triplet. For instance, a high voltage, low temperature and a fast

process will favor the timing performance (best case), than a low voltage, high temperature and slow process. This assertion is based on three facts. First, as the voltage increases, the available carrier numbers for the switching process increases as well. Second, the low temperature reduces the carrier-phonon scattering, and finally, the fast process ensures a device fast timing response.

Due to the large amount of possible PVT combinations per cell, the characterization of a standard cell library takes time (probably more than a six months) supposing the availability of a huge memory space. As an example a typical cell library can be constituted by at least 500 cells.

The corner method has two main disadvantages. First it consumes a lot of CPU resources. And, second, this method is not very accurate. Due to the impossibility to simulate all possible combinations of PVT and design parameters present in an IC, in most of the cases an interpolation is done to calculate the delay for certain conditions. This interpolation may induce a significant error in the timing analysis. There are several models used to calculate the delay by interpolation method such as Non-Linear Delay Model (NLDM), Composite Current Source model (CCS), etc [Synopsys11] (cf., Fig.3-9).

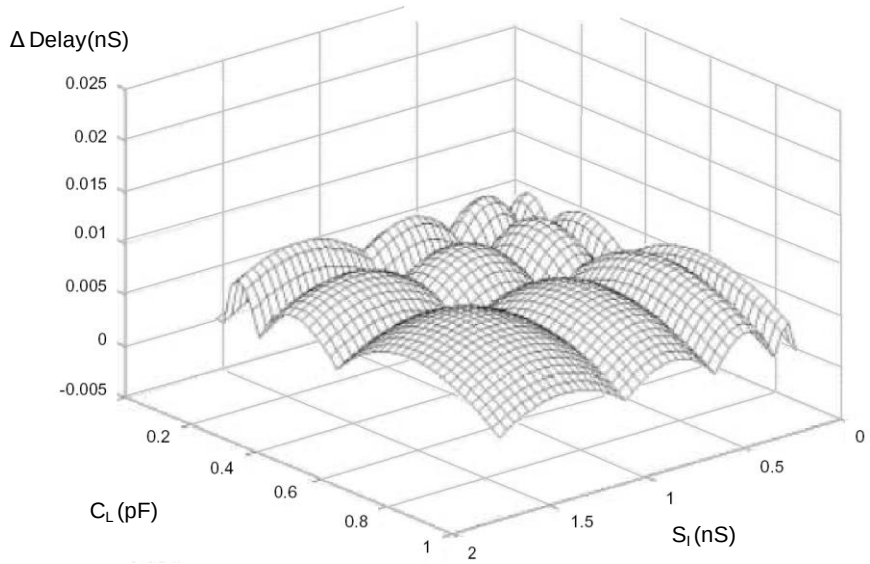


Figure 3-9 Interpolation absolute error considering a non-linear model for a table of [5X5]; five load capacitances and five input ramps.

For instance, the NLDM uses lookup tables with C_L and S_I as indexes. Due to the fact that its accuracy is proportional to the table size, this approach needs larger tables to provide a correct accuracy delay calculations. This can lead to memory blow-out. Moreover, with this approach voltage and temperature are linearly de-rated. Figure 3-9 shows an example of the

absolute error induced by the NLDM method [STA00] for a lookup table of [5X5]. For instance, the error induced by this method is 5ps for S_I= 1nS and C_L= 0.2pF.

A second example of the use of the NLDM is illustrated in the figure 3-10. In this case we need to calculate the propagation delay of a certain cell for a pair (S_I= 0.32, C_L= 0.05) at a given PVT, assuming that this pair was not characterized by simulation. It is therefore necessary to do an interpolation from the nearest neighbor data. The dots in this figure represent index values defined in output file of the library characterization. The delay calculation is carried out as follows. The input ramp, S_I= 0.32, falls between the index values 0.098 and 0.587, and the load capacitance C_L= 0.05 falls between index values 0.03 and 0.06. These index values correspond to the next x and y coordinates of the four neighboring table points as shown in figure 3-10 (circles). Using the interpolation method the value of z coordinate (delay) is 0.275nS (dashed circles) [Synopsys11].

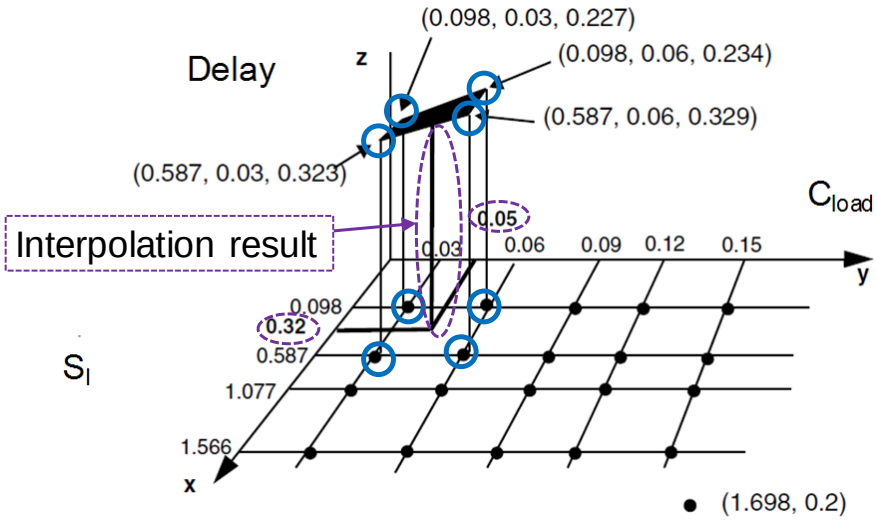


Figure 3-10 Result of delay calculation. Delay value is calculated by two-dimensional interpolation method

3.3.1 Limitations of STM standard cell library characterization flow: Fresh and Ageing PVTs

Long-term functionality of circuits is based on the reliable operation of each component (cells). In order to assure the correct operation of the circuit, variations introduced by the interaction between environmental parameters (PVT, C_L, S_I, etc.) and ageing phenomena have to be taken into account in the characterization flow. However testing a large number of configurations is impractical in current industrial practice. As such, in the current STM

standard cell characterization flow, the effects of ageing phenomena are considered for several combinations of PVT, C_L and S_I, however only one ageing condition is taken into account. For this condition the value of SP is worth 50%.

Each cell has a specific margin of yield (after stress) imposed by the technical specifications of circuit (mission profile). The mission profile defines the operation mode of a circuit i.e. frequency, voltage, power consumption, etc. A circuit can have different mission profiles depending on the applications in which it will embedded (i.e., smart phone, computers, etc). For instance, the table 3-1 shows some application profiles for the ARM Cortex-A9 processor (§4.2.3).

Next-Generation Devices	Typical Cortex-A9 Configuration
Mobile Handsets **Connected Mobile Computers** 	**High-end mobile devices (1500-3000DMIPS)** 2-3 core processor advanced power management 32K Instruction and Data caches, 256-512K shared L2 cache using PL310, partitioned AXI NEON technology-based Media Processing Engine
	Mid-range, cost reduction, (900-1500DMIPS) Single core processor with NEON or FPU 16K or 32K instruction and data caches 128-256K L2 cache using PL310, single AMBA AXI bus
	Feature-rich mass market (600-900DMIPS) Single core processor with FPU 16K instruction and data caches, single AXI
Consumer and Auto-infotainment 	**Consumer: user interactions (800-3000DMIPS)** 1-4 core processors giving design scalability across family of devices 32K instruction and data caches with 0-512K L2 cache NEON technology for advanced media and DSP processing Advanced bus interface unit for high-speed memory transfers between on-chip 3D engines and network interface MACs AMP configurations using separate CPU for real-time RTOS
Networking / Home Gateways 	**Enterprise market (4000-8000DMIPS)** 3-4 core performance optimized implementation 32K+64K instruction and data cache 512K-2MB L2 cache, dual 64 bit AMBA AXI interfaces
	Consumer devices (800-1500DMIPS) 1x or 2x multicore utilizing coherent accelerators 32+32K instruction and data, with 256-512K shared L2 cache NEON or VFP when offering media gateway or services
Embedded 	**Embedded media and imaging (800-2000DMIPS)** 2x multicore utilizing coherent accelerators 32+32K instruction and data with 256K shared L2 cache FPU for postscript and image manipulation and enhancement Code migration through selective AMP/SMP deployments

Table 3-1 Cortex A9 processor example application profiles

Table 3-2 shows an example of a smart phone mission profile. As it can be observed, each mode has a different stress profile, which in turn may generate different degradation levels. In order to illustrate this concept we give an example in the next paragraph.

Consider the figure 3-11, which shows a critical path of a small circuit existing in a smart phone. Let us assumes that the operation mode of the smart phone (during 50% of the time) is

the number five (from the table 3-2). In this mode, both CLK and DATA signals are set to "zero", thus the input signals of cells *a*, *b* and *e* are "zero" for 100% of the time (represented in the figure 3-11 whit the dark color). Note that D node of the flip-flop (*g* cell) is also set to zero. As consequence, cells *a*, *b*, *e* and *g* are stressed under a NBTI DC stress condition (100% of SP).

Now let us consider that the smart phone is activated in mode number 2 at a time= *t1*, the CLK and DATA signals change to "one" (represented in the figure with light color). In this new scenario, the timing performance is affected by the electrical weakness of pMOS transistors of *a*, *b*, *e* and *g* cells, thus the delay in the data path becomes negative (it has become faster than previous case). On the contrary, the clock path delay remains stable, this situation augments the probability of *setup violations* (§4.1.3)

Mode number	Mode name	Lifetime % active	Voltage mode [Vdd]	Temperature in transistor junction [°C]
1	Cellular Voice	10	0.85	75
2	Music playback	5	0.85	75
3	Messaging (SMS, email)	15	1.00	86
4	Internet navigation	5	1.15	100
5	Cellular standby	55	0.6	30

Table 3-2 Smart phone example mission profile

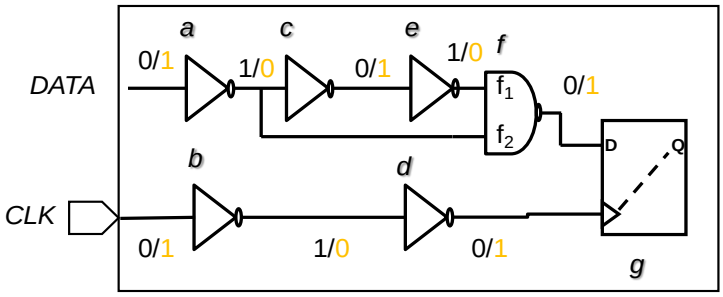


Figure 3-11 Example of a critical path (smart phone circuit)

As mentioned above one of main purpose of standard cell library characterization is abstracting the detailed characteristics of the SC into cell level timing in order to be used in synthesis phase, place and route, parasitic extraction [Keating07]. With the objective to compute accurately the gate timing performances at a given time, circuit operating conditions (mission profile) should be included in the standard cell characterization flow. However this practice is impracticable because of the large numbers of possible scenarios.

Consequently, we have developed an ageing gate model, which allows estimating the timing degradation considering every single product mission profile from a fresh PVT and an aged PVT. The background of the delay model at gate level as well as our approach will be detailed in the following sections.

3.4 Overview of ageing models at gate level

In order to quantify the impact of NBTI and HCI on the temporal delay of standard cells, several CAD frameworks for managing CMOS wearout mechanisms have been developed in the last years.

3.4.1 State of art of NBTI gate models

Wang *et al*., in [Wang07] presented an analytical model based on the Reaction-Diffusion (R-D) framework. For a given set of environmental conditions (i.e., supply voltage, temperature, etc) the authors utilize a NBTI transistor model defined by equation (3-3) to compute the threshold degradation (ΔV_{th}) of a given transistor in a certain gate.

$$\Delta V_{th,t} = \left(\frac{\sqrt{K_v^2 \, T_{clk} \cdot \alpha / \min(\alpha, 1-\alpha)}}{1-\beta_t^{1/2n}} \right)^{2n}, \tag{3-3}$$

where β_t is the fraction parameter of the recovery, T_{clk} represents the time period of one stress-recovery cycle, α defines the SP, ξ_1 and ξ_2 are constants, n is the time exponent parameter, which is worth 1/6 (in this case) and K_V is a complex parameter, which depends on electrical field and temperature (among others).

With the degraded threshold voltage (ΔV_{th}), the authors employ SPICE simulations to extract the propagation gate delay (t_P) under discrete working conditions (i.e., discrete values of ΔV_{th}). The discrete values of t_P (calculated by SPICE simulations) are then connected by fitting Chebyshev polynomial series in order to obtain a gate delay model depending on ΔV_{th}. Subsequently authors use an algorithm of circuit timing analysis (considering NBTI effects), to evaluate the timing degradation due to NBTI for each gate (i) in a circuit. The inputs of this algorithm are: input signal probability, the slew rate of the input signal and the load

capacitance of the gate under consideration. Once the timing degradation for the gate under consideration is computed (from the NBTI –aware library), this timing is added to the intrinsic delay of the gate (delay without NBTI effects) and a final gate delay is obtained.

Authors in [Kang07] focus their approach in the effect of random dopant fluctuation (RDF) on NBTI degradation. Using a RD-based ΔV_{th} model, the authors developed a statistical NBTI model to predict the impact of the degradation on lifetime circuit performance which is defined as follows:

$$\sigma_{Nit}^2 = \mu N_{it} = \frac{C_{ox}\mu\Delta V_{th}}{q} = \frac{\epsilon_{ox} A_G \mu\Delta V_{th}}{q t_{ox}}, \quad (3\text{-}4)$$

where σ_{Nit} and μ_{Nit} represent the mean and the standard deviation of Nit generation, which are modeled as Poisson random variables. $\mu\Delta V_{th}$ is the nominal (mean) V_{th} degradation due to NBTI, where ΔV_{th} with respect to the time is defined as: $\Delta V_{\text{th}}(t) = \frac{-qN_{it}(t)}{C_{ox}}$

Finally, A_G represents the effective channel area.

Approaches in [Lorentz09], [Paul06] are based on static Timing Analyzer methodology to calculate degradation impact at circuit level. The two frameworks provide the degraded gate performance for a given operating condition sets over lifetime circuit.

Unlike what is proposed in the last four approaches by the authors we consider that the R-D model is not the correct physical base to reproduce the NBTI behavior and we will use the model presented in (§2.2.2).

3.4.2 State of art of HCI gate models

In order to predict the HC damage on a circuit several macromodels have been developed based on transistor level simulation [Weber91], [Tu91], [Kuo88]. However the reliability simulation is inertly slow, since the transient electrical simulations incorporate the ageing of device parameters during the computation, which results in a larger computational CPU time. The macromodel methodologies offer a solution to this problem, increasing thus the computational efficiency of reliability analysis.

In this section, two such approaches based on algorithms and look-up tables will be examined.

3.4.2.1 iRULE: Fast Hot Carrier reliability diagnosis method

iRULE method [Teng94] avoids simulation during the reliability analysis. It employs netlist- and geometry–based rules to determine the potentially critical transistors in a circuit. It uses a macromodel (look-up) table, which contains hot carrier degradation values as a function of circuit operation conditions. These tables are generated from a library characterization framework (§3.3). The physical basis of this approach is the average bond-breaking substrate current method (I_{BB}*)[1].The authors consider that I_{BB}* over one period of time (T) provides an accurate measure of the dynamic device degradation.

This method transforms series transistors in multi-input logic into an equivalent transistor for simplification. Despite the fact that this framework can handle larger circuits than transistor-level ones, it is still difficult to consider the behavior of the multi-input logic and the complex logic correctly.

3.4.2.2 Timing simulation methodology

Kawakami *et al.*, propose a gate aged timing simulation methodology, which is made-up from a ratio based model explained in [Yonezawa98], [Yonezawa00] and an unique algorithm, in which the hot carrier degradation timing library is automatically created during the library characterization step (§3.3) by running the transistor-level simulator. According to the authors, this approach unlike the above one, can easily handle a full chip with multi-input logics. The developed algorithm consists of three parts. First, a fresh simulation is performed by using transistor-level models. Second, all circuits gates are decomposed into primitives. For instance, AND gate is composed by two primitives a NAND gate and an inverter (primitive-level). In this state each primitive is analyzed in order to calculate their respectively HC–induced damage. Then this stage provides primitive delays and ratios (α). Finally, in the last stage after recombining primitive data to be gate data, aged timing (T_{aged}) is obtained by scaling fresh timing (T_{fresh}) with above ratio according to the next expression:

[1] $I_{BB}{}^{*} = \frac{1}{T}\int_0^T I_{BB}(t)\delta t = (S_I, W_n/C_L)$, where I_{BB} represents the substrate current, which is calculated using the model devised by Kuo *et al.*, [Kuo88]. Following Leblebici *et al.*, [Leblebici90] the degradation of I-V characteristics nMOS devices is attributed primarily to generation of interface states (ΔNit), which in turn is defined in terms on I_{BB}.

$$T_{aged} = \sum_{j=1}^{m} \left(\frac{t_{aged\,j}}{t_{fresh\,i}} \cdot t_{fresh\,j} \right) = \sum_{j=1}^{m} (\alpha_j \cdot t_{fresh\,j}), \quad (3\text{-}5)$$

where *m*, t_{fresh} and t_{aged} are the number of primitives in the target gate, fresh and aged input-to-output delays at primitive-level respectively.

We consider that the physical model in which these last two approaches are based fails to explain the complex features of HCI degradation, which were presented in (§2.3.3.2). Therefore we will use the multi mode model (§2.3.3.3) as base of our HCI gate model.

3.5. Gate level NBTI analytical model

In this subsection we will propose a new gate level NBTI model, which will be described in detail. First we will analyze the NBTI effect on the simplest standard cell "the Inverter". An extension of the model is presented in section 3.5.2, which is validated for more complex cells such as XNOR, AOI, NAND, etc.

3.5.1 First stage: Inverter

Various ageing simulations were performed at different stress ages considering a larger range of C_L and S_I in order to verify if the time power law observed at transistor level remains true at the gate level. Let us consider three stress time, say t_1 =0.02years, t_2 =2years and t_3=20years. The NBTI delay degradation was calculated for an inverter for these three points using the L-R model (§2.4.2.3). Figure 3-12 shows the evolution of the delay degradation against stress time for an inverter, carried out in 45nm CMOS technology (driver 9X). As we can observe, the delay shift (Δd) grows up linearly with the stress time. As consequence, it can be asserted that the degradation delay power law with stress time remains valid for the inverter cells.

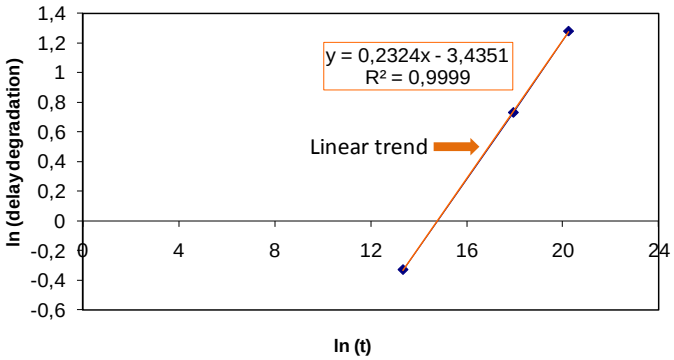


Figure 3-12 Power law delay degradation (provoked by NBTI) as a function of stress time. A linear trend is clearly observed between delay degradation (Δd) and stress time variables

An ageing analysis was performed on a family of 10 inverters (IVX) in 45nm CMOS technology. Table 3-3 summarizes the main characteristics of this inverter group. Several ELDO simulations were performed at various scenarios (including several stress voltage, SP, temperatures, supply voltage for a wide range of [C_L, S_I] pairs). Subsequently, least squares method was used to estimate the linearity correlation between the delay degradation values and the stress time values.

Inverter	Pmos number	Pmos W (nm)
IVX2	1	0.22
IVX4	1	0.41
IVX7	2	0.61
IVX9	2	0.79
IVX22	3	0.68
IVX31	4	0.71
IVX40	5	0.73
IVX71	8	0.81
IVX106	12	0.81
IVX284	32	0.81

Table 3-3 Standard cell general characteristics (Inverter family)

Figure 3-13 shows a histogram of the linearity coefficients calculated on the data base. The data base is composed of 594 elements (represented in the figure as N 594). Note that only the LOW-to-HIGH (L-to-H) input vector was considered in this case. It can be observed that the linearity coefficient histogram confirmes accurately the results shown in the figure 3-12. In the others words, the power law with the stress time is kept despite the complexity of certains inverters under consideration (such as IVX106 and IVX284, which are composed by various parallel stage of pMOS and nMOS devices).

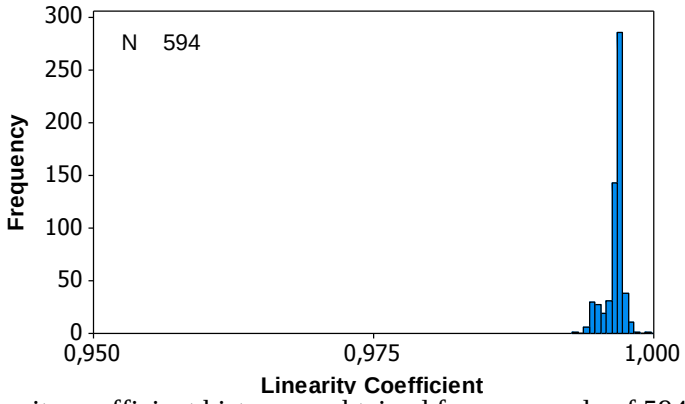


Figure 3-13 Linearity coefficient histogram obtained from a sample of 594 *L-to-H* vectors

From these results we propose a 3-parameters simplified ageing model for the inverter cells, which is defined as follows:

$$\frac{\delta t}{t} = S \cdot (POT)^n, \tag{3-6}$$

where *S* represents the sensitivity of the inverter to ageing, which mainly depends on S_I and C_L. The second parameter is the power-on time (*POT*) which is the equivalent stress time with respect to the product mission profile and the nominal voltage. The third parameter is the time exponent *n*, assuming a power law dependence for the delay degradation.

Two elements form the basis of this methodology. The first element is a well-established ageing dynamics, which is calculated from two PVT corners (a PVT fresh and a PVT aged), since it is impossible to calculate the ageing effects (with enough accuracy) on a standard cell from a single PVT corner (§2.1.2 and §2.3.2). The second element is a robust procedure to determine the delay of the standard cell for given PVT corners. We use physics-based models to calculate the delay for fresh and ageing PVT (§2.4.2.2 and §2.4.2.3), since uncorrelated reliability transistor models with silicon physics means a larger reduction of performance at system level [Ruiz11].

A library characterization-based methodology is the heart of this approach. The parameters *S* and *n* are obtained from the ageing dynamics (cf., Fig.3-14). On the other hand, the *POT* parameter is calculated from the generic mission profile (cf., Table 3-4).

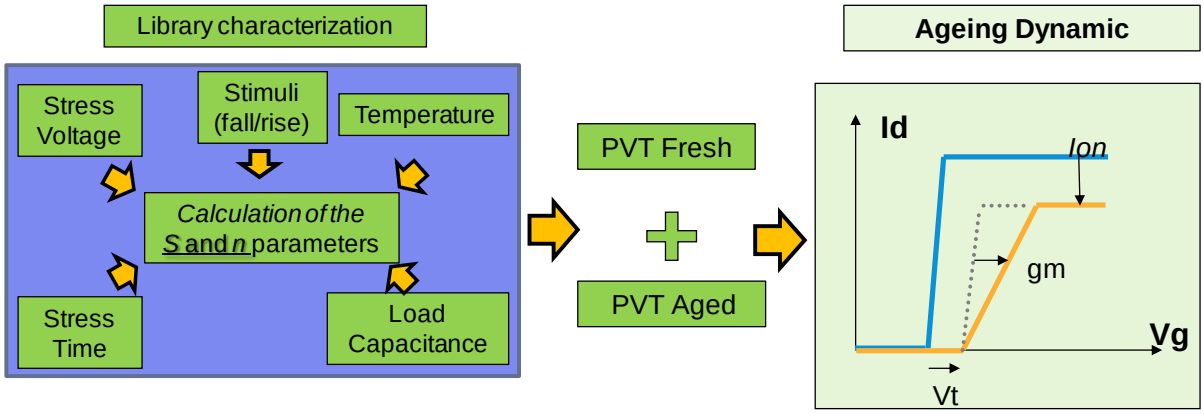


Figure 3-14 Library characterization schematic, showing the computation method of model parameters: sensitivity to ageing *(S)* and time exponent (*n*)

The generic mission profile is calculated from the Early Life Failure Rate (ELFR) method defined in [JEDEC74], which is used to test the early life failure rate of a product (silicon experiences). The acceleration models with respect to voltage and temperature parameters are defined by equations 3-7 and 3-8 respectively. These models are used to calculate the acceleration factor (*A*) for the Early Fails Failure (ELF).

$$A_V = \exp\left[\left(\frac{K}{X}\right) \cdot (V_A - V_U)\right] = \exp[(\gamma_v \ \cdot (V_A - V_U)], \tag{3-7}$$

where A_V defines the voltage acceleration factor, K represents the electric field constant (which is determined experimentally), X defines the thickness of stressed dielectric, γ_V is defined as the ratio (K/X) [V^{-1}], V_A is the stress voltage in accelerated ELF test and V_U represent the voltage at normal use conditions.

$$A_T = \exp\left[\left(\frac{E_{aa}}{K}\right) \cdot \left({}^{1}/_{T_u} - {}^{1}/_{T_A}\right)\right], \tag{3-8}$$

where A_T defines the temperature acceleration factor, E_{aa} represents the apparent activation energy (eV), K is the Boltzmann's constant (eV/°K), T_U defines the junction temperatures at normal use conditions of the product (°K) and T_A represents the junction temperatures at accelerated conditions (°K).

Precisely, the acceleration factor *A* is defined as the product of voltage acceleration and temperature factors,

$$A = A_T \cdot A_V \tag{3-9}$$

Finally, the Power-On-Time parameter is defined as the sum of the result of the multiplications of the *A* factor and of the Equivalent Life Active (ELA) [seconds] for a certain number of modes.

$$POT = \sum A \cdot ELA \tag{3-10}$$

Note that the acceleration factors are extracted from the High-Temperature Operating Lifetests (HTOL).

	I	II	III	IV	V	VI	VII	VIII	IX
	Mode number	**Mode name**	**Lifetime (%)active**	**Equivalent Life active (seconds)**	**Voltage mode [Vdd]**	**Temperature in transistor junction [°C]**	**A_T**	**A_V**	**Normalized Stress time at a condition (V and T°) VII*VIII*IV**
	1	Cellular Voice	10	$18x10^6$	0.85	75	0.116	1.325	$2.77x10^5$
	2	Music playback	5	$9x10^6$	0.85	75	0.075	1.54	$1.49x10^6$
	3	Messaging (SMS, email)	15	$27x10^6$	1.00	86	0.00153	0.0009	$3.72x10^6$
	4	Internet navigation	5	$9x10^6$	1.15	100	0.1284	1.55	$1.79x10^6$
	5	Cellular standby	55	$99x10^6$	0.6	30	0.0002	0.00	0.000
Total	Generic Mission Profile		100	Equivalent stress time at given condition (voltage and temperature)					$35.56x10^6$

POT

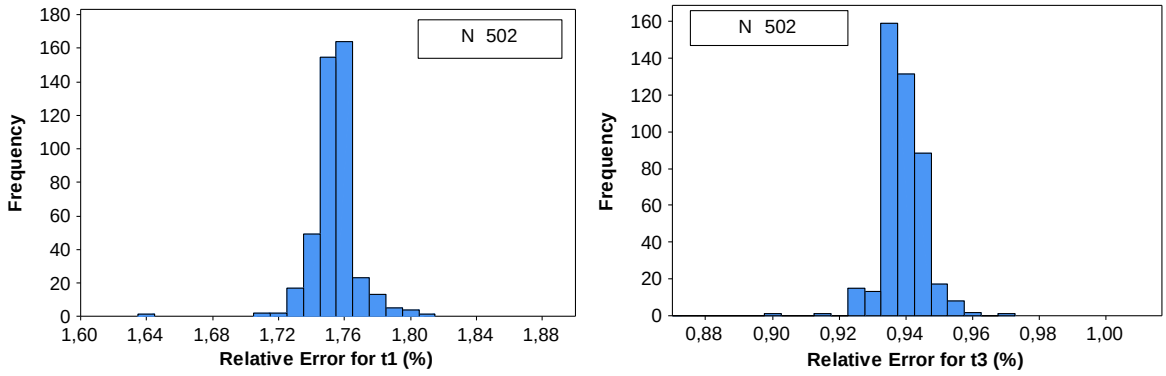
Table 3-4 Generic mission profile equivalent computation from a smart phone mission profile (Note that this example is not a real mission profile)

In order to check the validity of the proposed gate model, a comparison with ELDO-gate model was performed. With this purpose, two extra ageing PVT corners (t_1, t_3) were characterized for the inverter family (note that these extra corners are not part of the STM flow). Thousands of ELDO simulations were performed at two values of the stress time (t_1, t_3), including several stress voltages and a wide range of operating conditions [temperature, supply voltage, C_L, S_I]. The value of SP chosen for this study was 50%[2] . Parallel, the equation (3-6) was used to calculate the delay degradation on the inverter family at t_1 and t_3 corners based on the available PVT corners in STM characterization flow.

Figure 3-15 shows the gate model relative error with respect to Eldo simulations. This analysis shows that the gate-level model is accurate to reproduce delay degradations with a

[2] This is the SP value considered in STM characterization flow (§3.3.1)

maximum margin of error of 1.809% of the total degradation (cf., Fig. 3-15a). For example, if the mean delay degradation is 10% gate-level model is accurate within 0.1809%.



(a) (b)

Figure 3-15 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 (a) and t_3 (b) PVT corners. Gate-level model is slightly pessimistic with transistor-level model by about 1.81% of the degradation

An important point to relieve is the difference between the behaviors of parameters *S* and *n*. Figure 3-16 shows the distribution of *S* parameter calculated from equation (3-6) for the inverter family. As we can observe this distribution presents a wide range of variation. This fact can be explained by the strong dependence of *S* with electrical parameters (C_L, S_I).

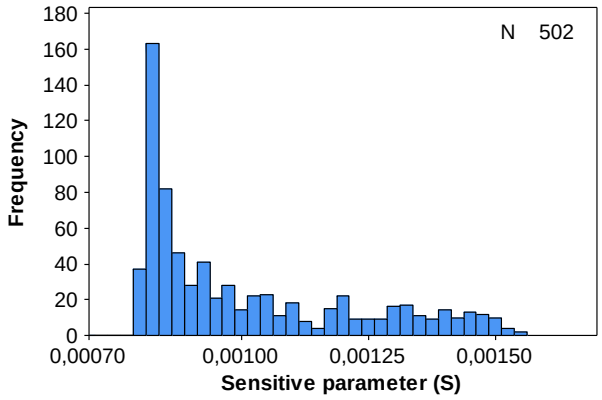


Figure 3-16 Sensitivity *S* histogram, showing the strong dependence of *S* with timing arcs

On the other hand, the distribution of *n* parameter has a small sigma (σ) (cf., Fig. 3-17). From this result it can be seen that the time exponent *n* has a weak dependence with the design inputs.

In order to simplify the calculation, a single value of *n* was assumed. Using (3-6), the degradation delay was evaluated taking into account this approximation. As observed in

figure 3-18, the relative error between the gate predictions (with fixed n) and Eldo simulations is smaller.

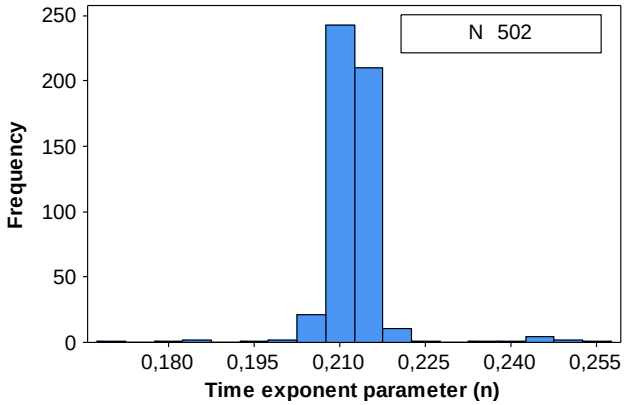
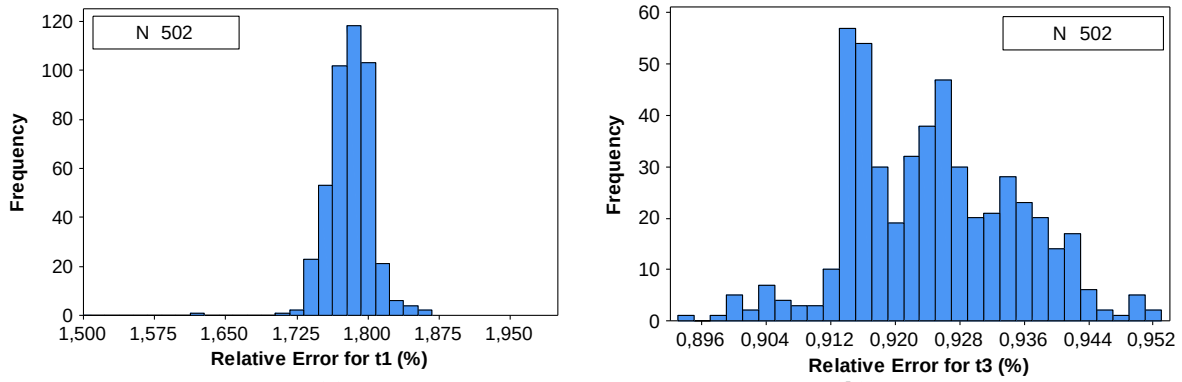


Figure 3-17 Time exponent n histogram, showing the weak dependence of parameter n with timing

The new delay degradation results (calculated on the inverter family) agree very well with the original values (cf., Fig. 3-15). The maximum error obtained for t_3 and t_1 points is the 1.872%. Therefore the total error generated by the model after *n* approximation is ~ 1.9%. Retaking our last example, for 10% of mean delay degradation, gate model is accurate within 0.19%.



(a) (b)

Figure 3-18 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 (a) and t_3 (b) time points assuming a unique value of *n* parameter

3.5.2 Extension of the gate analytical model to more complex standard cells

Following the approach presented in the previous section, a data base was built considering the NAND, AND, NOR and AO standard cells. The conditions of the ageing simulations remain identical to section 3.5.1. Linearity analysis was applied on the new data base in order to corroborate the observed results on the inverter family. The data base is composed by 4470

elements (or N 4470). As we can see in figure 3-19, the coefficient of linearity allows extending the time law power of the degradation NBTI at gate level.

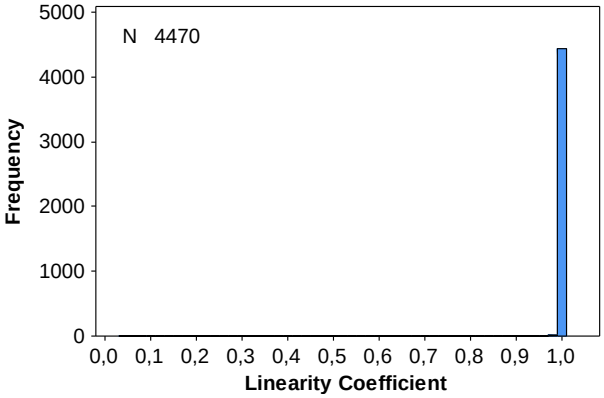


Figure 3-19 Histogram of the linearity coefficients for a simple of 4470 timing arcs

Figure 3-20 shows the histogram of time exponent parameter *n* (Eq. 3-6) for several standard cells. The limits of the *n* parameter range remain almost stable [0.16 to 0.29].

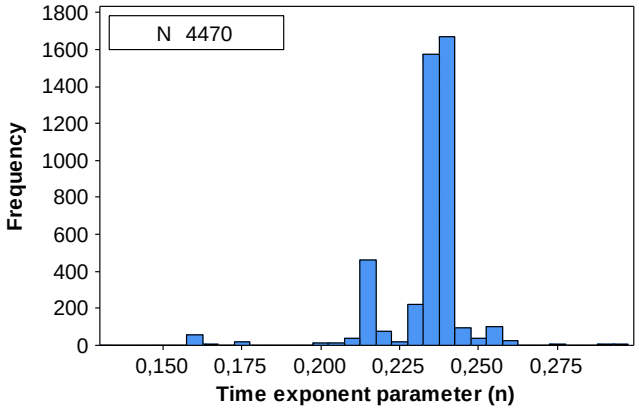


Figure 3-20 Time exponent *n* histogram

As we observe in figure 3-21, the sigma of the S parameter population is larger. This is due to the strong dependence of *S* parameter with both standard cell architecture and timing arcs. Consequently, we consider that all delay degradations of standard-cells library must be characterized by simulators type SPICE in order to ensure an adequate accuracy level in the computes of timing estimations.

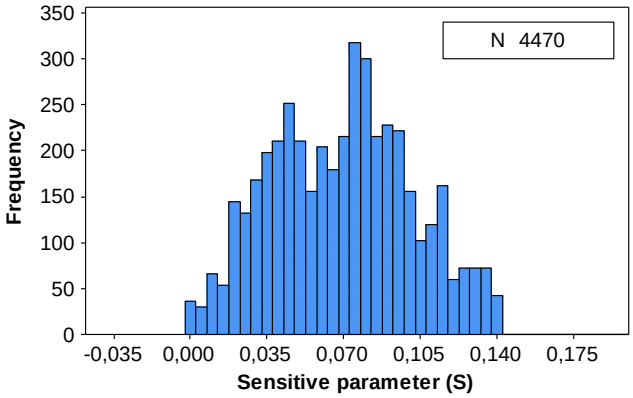
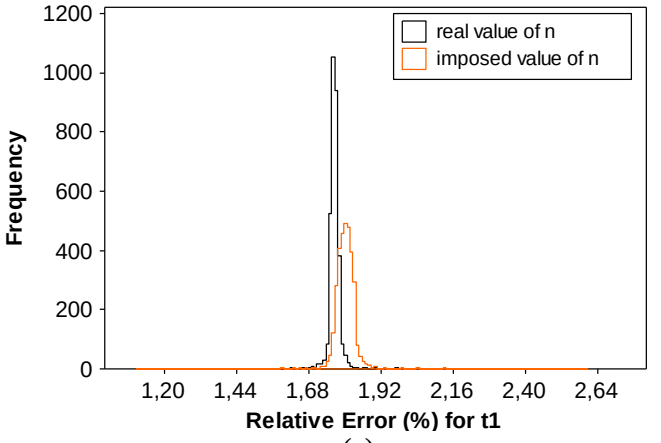
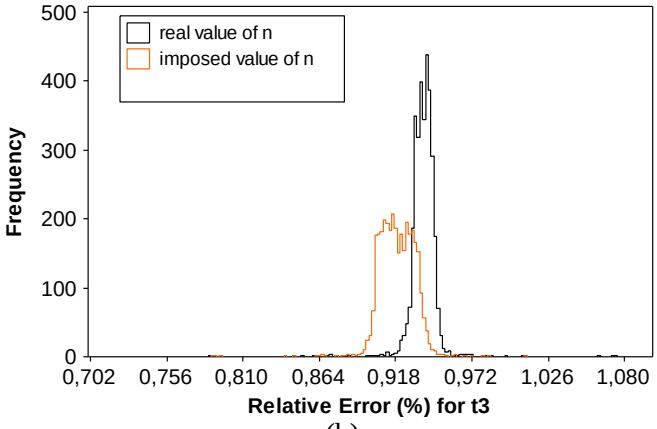


Figure 3-21 Sensitivity *S* histogram

Figure 3-22 shows the delay degradation relative error between the gate level and the transistor level models for the t_1 (cf., Fig. 3.22a) and t_3 (cf., Fig. 3-22b). As we can see, the trend observed in the above section is conserved. The maximum error in the calculation is made for *t1* point (at *n*-imposed condition). The largest error of model for this population is 2,636% on the delay degradation.



(a)



(b)

Figure 3-22 NBTI Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 (a) and t_3 (b) time points, assuming a unique value of *n* parameter

Therefore if we retake the example of 10% of mean delay degradation, the gate level model is accurate within 0.267%. Despite the extra error introduced by the hypothesis made on *n* parameter (*n* fixed), we consider this level of accuracy satisfactory to yield accurate timing analysis at gate level.

3.6 NBTI analysis on 32 nm CMOS technology

In this section, we present the application of NBTI gate model on CMOS 32nm standard cells. The NBTI gate model can be applied to any technology node since the gate model ageing dynamics is based on physical transistor models. Table 3-5 summarizes the standard cells chosen for this study and their corresponding timing arcs.

Gate (Xxx denotes the cell driver)	Timing Arc
NAND2X12, NAND2X13, NAND2X27, NAND2X50	AFZR1, BFZR1
NOR2X27, NOR2X41, NOR2X65, NOR2X84	AFZR0, BFZR0
BFX134, BFX25, BFX42, BFX67	AFZF
IVX134, IVX 25, IVX 58	AFZF
NAND3X12, NAND3X 24, NAND3X6	AFZR11, BFZR11, CFZR11
XNOR2X6, XNOR2X4, XNOR2X17, XNOR2X25	AFZF1,AFZFR0, BFZF1, BFZR0
XNOR3X4, XNOR3X16, XNOR3X25, XNOR3X18	AFZF01, AFZR00, AFZR11, AFZF10, BFZF01, BFZR00, BFZR11, BFZF10, CFZF01, CFZR00, CFZR11, CFZF10

Table 3-5 Timing arcs of several standard cells

In this table A, B, C, D denotes the input pins of the cell under consideration and Z denotes the output pin. *F* represents falling edge and *R* defines a rising edge. Finally 1 and 0 defines the values of the other pins of the cell. For example a vector AFZR1 for a gate of two inputs means that the pin A is set to zero (AF), the pin B is set to one (1) and the output predicted is one (ZR). Due to the complexity of certain gates such as XNOR3X4, which allows a large number of possible combinations of pin values, only the NBTI worst cases are chosen. In the other words the timing arcs presented in this table were chosen in such a way to generate the maximum NBTI delay degradation of the gate.

NBTI gate model results are compared with reliability simulation data. In figure 3-23 it can be seen that there is a very good agreement between the gate model predictions and

simulations. The maximum error introduces by the model is 2.45%. This result ascertains that the model accuracy was not affected by the change of technological node.

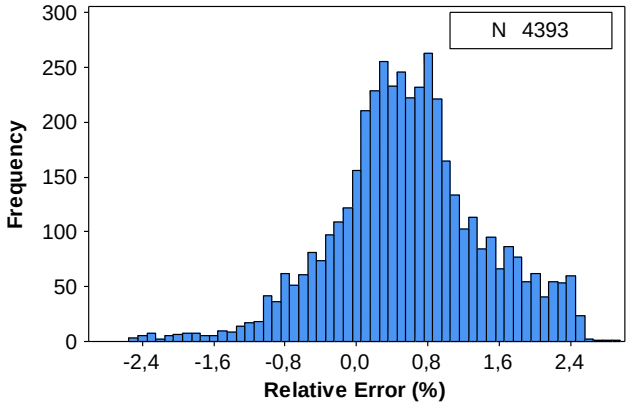


Figure 3-23 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 PVT assuming a unique value of *n* parameter

3.7 HC gate level analytical model

As HC degradation follows a time power law with the stress time as we have seen for the NBTI degradation, we will apply the same method used to address NBTI degradation to model the HC degradation. It is worth to note that although the origins (i.e., bias polarization) of NBTI and HCI degradations are different and that these phenomena act on different devices (NBTI → pMOSFET and HCI → nMOSFET), their physical consequences on device are the same (§2.1.2) (§2.3.2).

We did a depth HC analysis on an inverter in 45nm CMOS technology. Several simulations were performed at different conditions including various stress voltages [1.4, 1.5, 1.6, 1.7] volts, stress times [10, 48,168, 500, 1000] hours, two stress temperatures [25, 80] °C and a larger set of (C_L, S_I). First we calculate the linearity coefficient of the sample, which is shown in figure 3-24a. Figure 3.24b illustrates the distribution of values of *n* parameter. It can be observed that the dispersion of the data is very wide. In the figure 3.24c we observe that the values of sensitivity parameter sample are very scattered. The causes of this behavior (*n* and *S* parameter large dispersions) can be explained by two reasons. First, the physical difference (*i.e.*, t_{ox}) between the transitors under consideration. The data presented in these figures were generated from simulations launched on five different libraries. Each library was designed for special purposes, thus the design of the elementary devices (transistors) is different in each one of these five libraries.

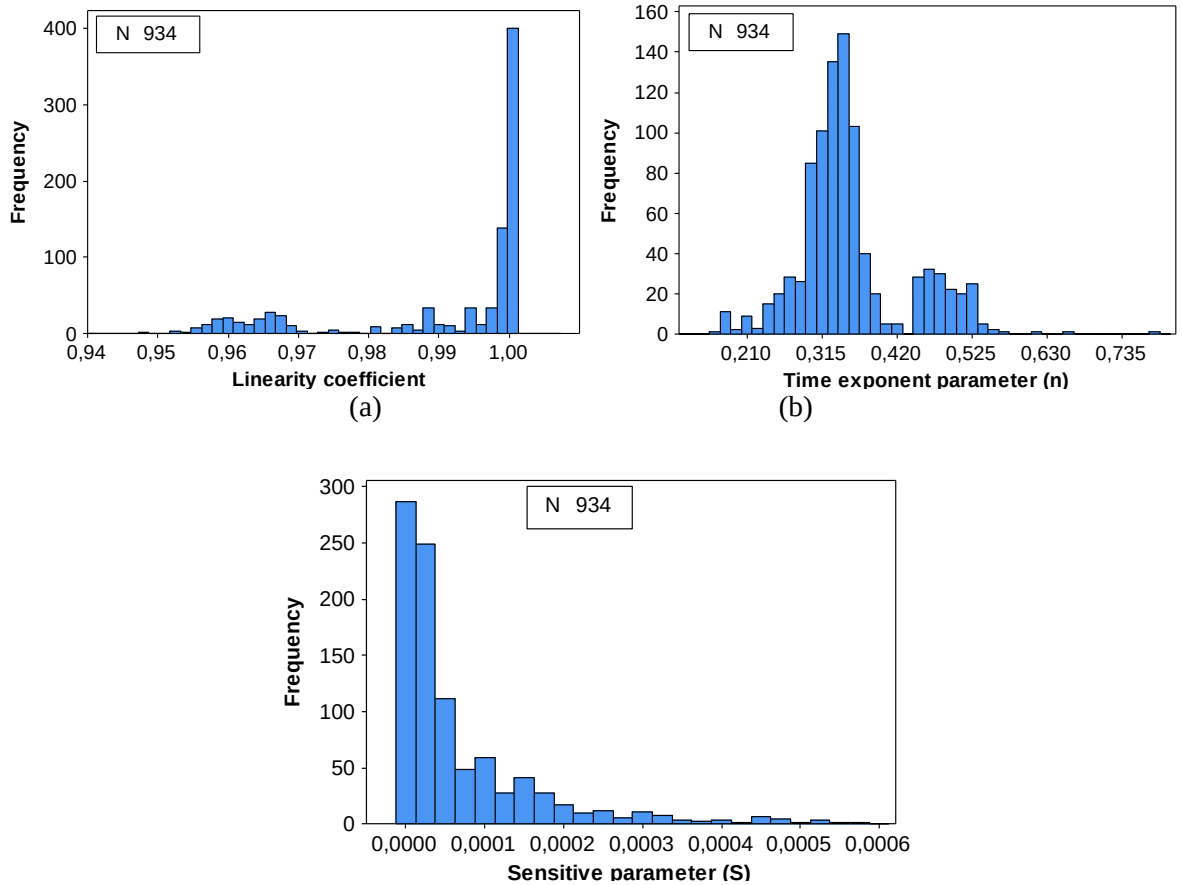


Figure 3-24 (a) Histogram of the linearity coefficients for a simple of 4470 timing arcs, showing the non-linearity of HC degradation with stress time for larger values of C_L and smaller values of S_I. (b) Time exponent n histogram. (c) Sensitive parameter histogram

As a consequence, the acceleration characteristic factors of each library are represented in this sample (§2.3.3.4) (cf., 3-24c).

The second reason is of this behavior is the strong dependence of HC degradation with respect to the design parameters (i.e., C_L and S_I). From this figure it can be observed that the time power law is not conserved at gate level for all timing arcs cases. This is due to the strong electrical parameters influence on delay degradation (especially for the larger values of C_L and smaller values of S_I). Then a linear model is not the best representation of the HC degradation. Nevertheless the maximum margin of the relative error (with the reference *Eldo* simulations) introduced by the equation (3-6) is 2.15% (cf., Fig. 3-25). Note that in this case the order of magnitude of the data sample is picoseconds. In the others terms, the maximum absolute error made by the model is 0.442ps at a nominal degradation of 17.7ps.

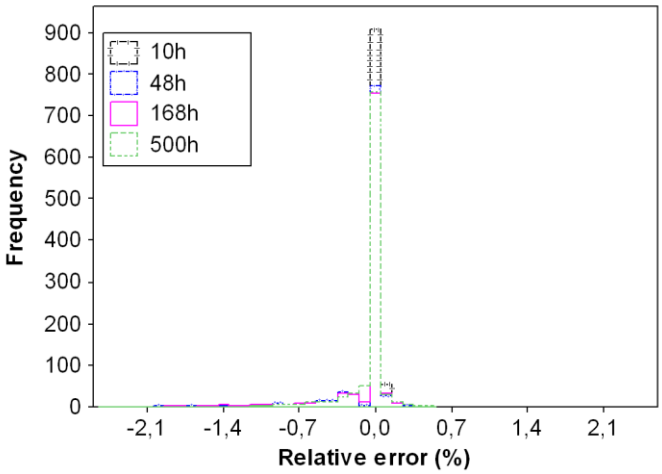


Figure 3-25 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for different values of stress time [10h, 48h, 168h, 500h]

3.7.1 Inverter and Buffer standard cells HC characterization

As mentioned early, the HC damage is generated during the periods of switching signals, therefore this phenomenon is highly dependent of the operation frequency of the circuit. As a consequence, we focus our work on Inverters and Buffers, which are very used in circuit clock paths.

3.7.1.1 Inverter case

We see in figure 3-26 that the gate model predictions are optimistic in most of the cases (in this example) with respect to the transistor model. This can be explained by the fact that HCI degradation decreases at longer stress times, indicating a tendency towards the saturation. Therefore the HCI degradation previsions of the model can be pessimistic or optimistic depending on the characteristics of the PVTs (*i.e.*, age) used to feed the gate model (Eq. 3-6).

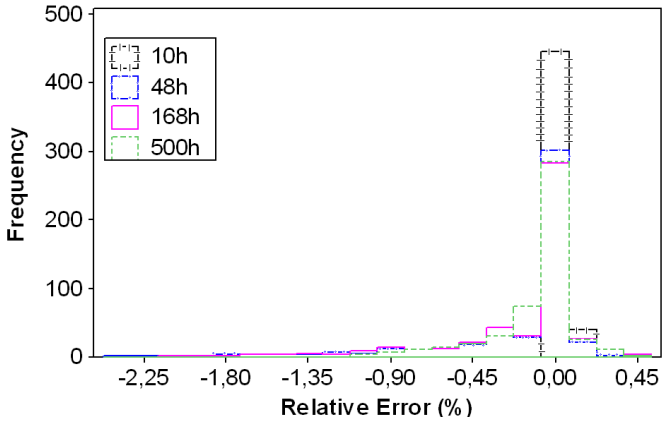


Figure 3-26 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 time point assuming a unique value of *n* parameter. The negative values of relative error show the optimistic trend of gate model with respect to transistor model (or Eldo simulations)

3.7.1.2 HC characterization for Buffer gates

This section presents the relative error of the gate model for a buffer family (several drivers) carried out in 32nm CMOS technology. In this case we considered various libraries of standard cells for low power applications. Each library was designed with a different type of MOSFET, which changes the speed conduction of the transistor.

Figure 3-27 shows the relative error of the gate model compared with the transistor model (ELDO simulations). As we can observe in the figure, the maximum error margin of the model is 0.55% on the mean delay degradation. For instance if the standard cell degradation calculated by Eldo is the 2.5% (which is already quite large with current product usage) gate model prediction will be 3.05%.

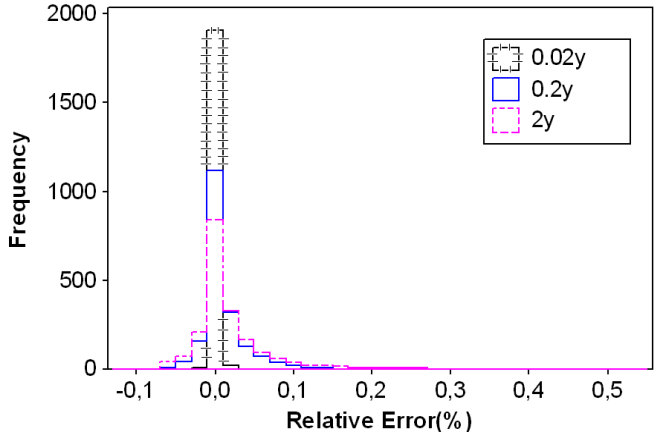


Figure 3-27 Relative error of the gate model delay degradation predictions with transistor model (Eldo simulations) for t_1 time point assuming a unique value of *n* parameter

3.7.2 Slope/Capacitance approach

3.7.2.1 Buffer cells

In order to model the HC degradation, a second analytical approach was studied, which is based on the degradation power trend with the ratio of S_I/C_L.

This approach is also valid for inverters cells. As it can be seen in the figure 3-28, a buffer is composed by certain X number of inverters. Therefore this cell can be analyzed as a single structure or as group of structures. In our case we decided to analyze it as a group of inverters. This type of methodology is widely used because it reduces the complexity of the delay degradation analysis.

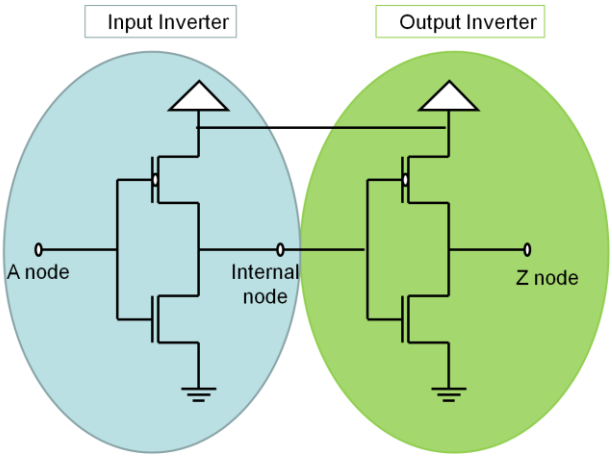


Figure 3-28 Schematic representation of a buffer

Figure 3-29 shows the delay degradation as a function of the ratio input slope/capacitance load for a buffer driver 4 (X4) carried out in 45nm CMOS technology. Two populations are easily identifiable.

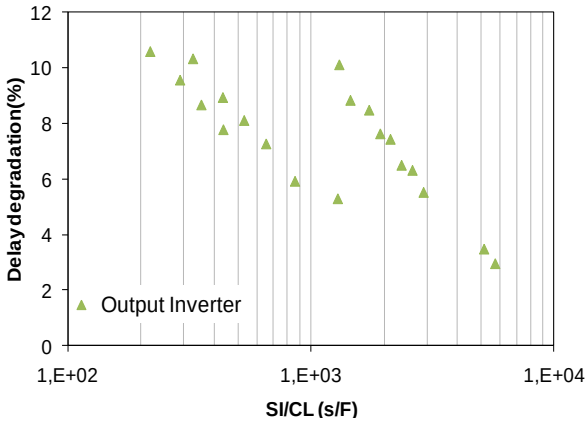


Figure 3-29 HCI Delay degradation with the ratio (S_I/C_L) for a buffer driver 4 carried out in 45nm

In the first part of this figure, we can observe that the first inverter (input inverter) contribution to the delay and in the second part the contribution of the second inverter (output inverter). Considering the output slope of *input inverter* (*IV*) as the input slope of the *output inverter* (*OV*) and the capacitance of the output node of the *IV* as the load capacitance of the *OV,* we observe a new behavior presented in the figure 3-30. As we can see in this figure, a trend appears confirming the dependence of the delay degradation with S_I/C_L ratio. It should be noted that the dispersion of the output inverters delays is highest (represented in the figure by the triangles) than the input inverter data (represented in the figure by *X* symbols) this fact is explained by the magnitude difference between: the C_L of the internal node (C_{LI}) and the C_L of the Z node (C_{LZ}), and, the slope of the node A (S_{IA}) and the slope of the internal node (S_{IN}).

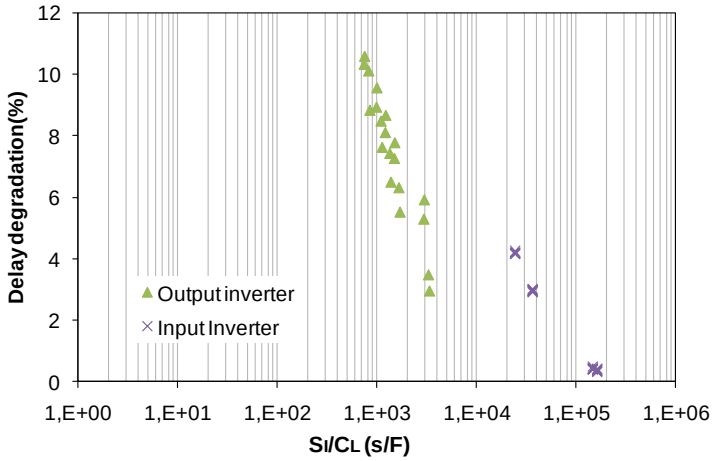


Figure 3-30 HCI delay degradation in percentage as a function of the ratio (S_I/C_L) for input inverter (delay on internal node) and output inverter (delay on Z node)

The C_{LI} values are ranged between 12.6fF and 13.4fF while the C_{LZ} values are ranged between 3.50fF and 13.8fF. The S_{IA} range is defined between 3ps and 18ps and for the S_{IN} between 7.61ps and 9.53ps.

On the other hand, it should not be forgotten that the switching process is strongly dependent of the slope (§3.1.1). Figure 3-31 focuses on the *output inverter* data. In this figure it is clearly observed two populations of ramps, fast ramps (light circle) and slow ramps (dark circle). The degradation generated by the slow ramps tends to be lower than the fast ramps; this can be explained by the fact that the electric current level, available during the switching process for the slow ramps is not the maximum current level as it is the case for the fast ramps.

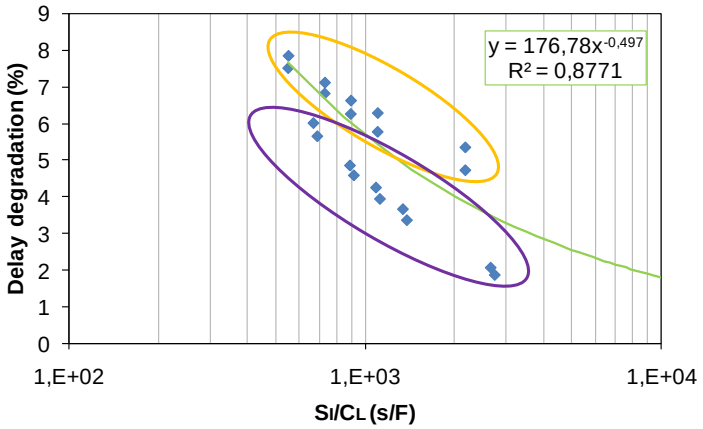


Figure 3-31 HCI delay degradation in percentage as a function of the ratio (S_I/C_L) output inverter (delay on Z node)

This approach is proved impracticable as S_{IA}, C_{LI} and the switching electrical current are impossible to obtain from the STM library characterization flow. Therefore this approach was not completely developed; this point cannot be explained more in details because we should need to discuss technical aspects, which are protected by an agreement *Confidentiality*.

3.8 Summary

In the first part of this chapter we explained the influence of the design parameters on delay degradation. A detailed description of the standard library characterization flow was presented in the section 3.3.

In the sections 3.5 and 3.7 we proposed a framework to model the timing degradations provoked by NBTI and HCI phenomena, respectively.

In the next chapter, we will address the validation of the gate model at system level.

Chapter 4

Reliability Modeling at SoC level

This chapter describes the reliability estimation at system level using the NBTI gate model presented in chapter 3. It in this context, we present two reliability analyses carried out on a Low-Density Parity-check (LDPC) codec circuit and a Cortex–A9 ARM IPs (§4.2) in order to verify the accuracy predictions of the gate model at system level. HCI phenomenon is not considered at this level, this point is explained in detail in appendix A.

In order to carry out the verification of gate model predictions at system level, several timing analyses were performed using the Static Timing Analysis (STA) tools. Therefore, in the first part of this chapter, a Static Timing Analysis framework is presented, in which the main basic concepts of STA are explained. The basic definitions of this subsection were taken out of the book referenced in [Dharchoudhury06].

4.1 Static Timing Analysis

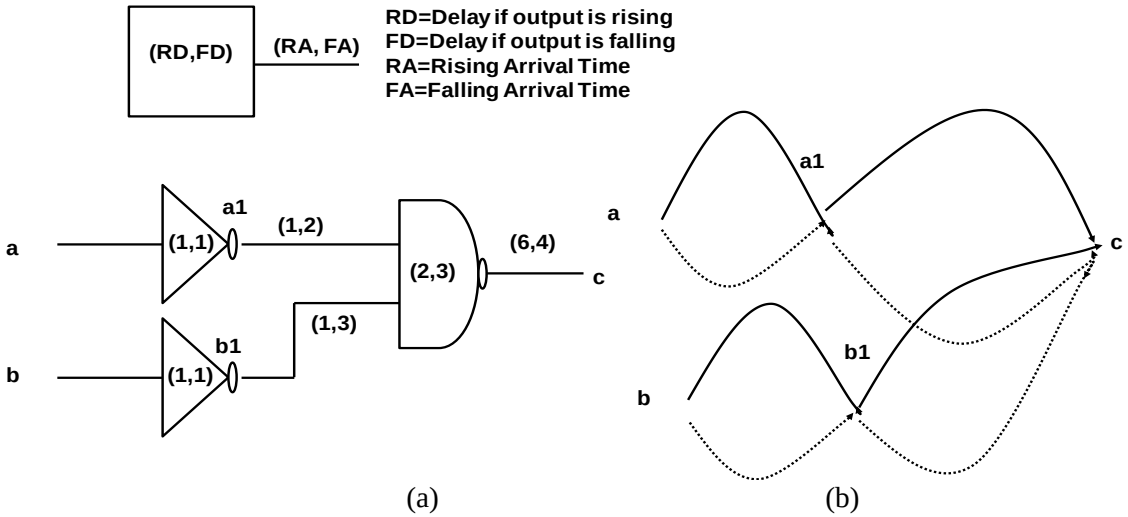
Timing analysis is one of the most important verification steps in IC design since it checks the speed specifications of the design, prior to the fabrication [Sakallah90], [Jouppi87], [Burks 95]. Static Timing Analysis (STA) is a method for computing the expected timing of a design without requiring simulation by checking all possible paths for timing violations (§4.1.3). This technique is pattern independent, which means that it allows the verification of all signal propagations throughout the design, and is applicable to very large designs.

The aim of STA is to find the *critical paths* in the design. Critical paths are those signal propagation paths that determine the maximum operating frequency of the design. For instance, in a microprocessor, the critical paths are mainly located between the cache memory and the CPU logic. In order to determine the critical paths, STA tool follows the next steps. Starting at the inputs of the circuit, the latest time at which signals arrive at a node in the circuit is determinated from the arrival times (§4.1.2) at a certain fan-in (defines as the capacitance of inputs nodes) nodes.
This latest arrival time is then propagated toward the primary outputs. At each primary output, we obtain the latest possible arrival time of signals and the corresponding longest path. If the longest path does not meet the timing constraints imposed by the designer, then a violation is observed. Alternatively, if the longest path meets the timing constraints, then all other paths in the circuit will also satisfy the timing constraints.

4.1.1 Timing Graph

The elementary structure in STA is the timing diagram. The timing diagram is a graphical representation of the design, where each vertex in the diagram corresponds to an input or an output node of the design. Each timing arc has a direction defined by the type of transition at the input and output nodes. For example, there are two timings arcs from the input to the output of an inverter: one corresponds to the IN_rising_OUT_falling and the other to IN_falling_OUT_rising. Each timing arc in the diagram is annotated with the propagation

delay of the signal from the input to the output. Figure 4-1 shows the timing diagram for a small combinational circuit.



(a) (b)

Figure 4-1 Basic diagram with delays and arrival times (a), showing the signal timing arcs (b)

Assuming that *a* and *b* are 0, the dotted line arcs correspond to rising input transitions at nodes a1, b1 and falling output transitions at node c. The solid line time arcs represent the falling input transitions at nodes a1, b1 and the rising output transitions at node c when *a* and *b* are set 1.

4.1.2 Arrival Times

Given the times at which the signals at the primary inputs of the circuit are stable (or storage elements to primary output), the minimum (earliest) and maximum (latest) arrival times of signals at all the nodes in the circuit can be calculated with a single breadth-first pass through the circuit in topological order. The early arrival time $t_{Ae}(c)$ is the smallest time by which signals arrive at node *c* from a certain node (i.e., *a* node) (cf., Fig. 4-1) and is given by:

$$t_{Ae,a}(c) = min_{a \in FI(c)}[t_{Ae}(a) + d_{ac}], \quad (4\text{-}1)$$

Similarly, the late arrival time $t_{Al}(c)$ is the latest time by which signals arrive at node *c* and is given by

$$t_{Al,a}(c) = max_{a \in FI(c)}[t_{Al}(a) + d_{ac}], \quad (4\text{-}2)$$

In these equations *FI(c)* represents the set of all fan-in nodes of *c*, i.e., all nodes that have and edge to *c*, d_{ac} is the delay of an edge from *a* to *c*.
These equations can be used to calculate both the arrival times and the delays at node *c* from its fan-in nodes. A breadth-first pass through the timing graph using Eqs. 4-1 and 4-2 will yield the arrival times at all nodes in the circuit.

In order to illustrates the concept of arrival time let us consider the timing graph of the figure 4-1a, and assuming that the arrival times at the primary inputs *a* and *b* are 0. Let's define the maximum arrival time for a rising signal at node *a1* as $t_{Al,r}$= 1, and the maximum arrival time for a falling signal as $t_{Al,f}$= 2. The timing conditions for the node *b1* are $t_{Al,r}$= *1* and $t_{Al,f}$= *3*. The delays of cells are defined as $d_{aa1,f}$= $d_{bb1,f}$= $d_{aa1,r}$= $d_{bb1,r}$= 1 and $d_{a1c,r}$= $d_{b1c,r}$= 2 and $d_{a1c,f}$ = 2 and $d_{b1c,f}$ =3, where the subscripts *r* and *f* denote the polarity of the signal (rising or falling).

For instance if the node *a* is set 0 and the node *b* is set 1 we can calculate the total latest arrival time at node *c* using Eq. 4-2. For this condition, the t_{Al} at node *c* is given by the *b* path delay (d_{bc}), which is defined as $t_{Al,r,b}(c)$= 1+ 3+ 2= 6 (longest path) while that the *a* path delay is given by $t_{Al,r,a}(c)$= *1+ 1+ 2= 4*. Note that in this case we considered that nodes a and b are the firsts nodes of the circuit and we assume the input delays as $t_{Al,f,r}$ (a,b)= 0.

4.1.3 Required Times and Slacks

Usually the constraints are placed on the arrival time of signals at the primary output nodes of a circuit based on performance or speed requirements. Additionally, the timing constraints are imposed on the clocked elements inside the design such as flip-flops, latches, etc. These timing constraints allow checking the correct functionality of the circuit in terms of speed. The nodes in the circuit where the timings checks are done are called *sink nodes*.

Given the earliest and the latest signal arrival times at the sink nodes, the required times at all other nodes in the circuit can be calculated by processing the circuit in reverse topological order considering each node only once.

The late required time $R_l(c)$ at the node *c* (or the time before which a signal must arrive at a node *c)* is given by:

$$R_l(c) = max_{a \in FO(c)}[R_l(a) - d_{ac}]. \quad (4\text{-}3)$$

Similarly, the early required time $R_e(c)$ is the required time on the early arriving signals (or the time after which the signal must arrive) and it is given by:

$$R_e(c) = min_{a \in FO(c)}[R_e(a) - d_{ac}], \quad (4\text{-}4)$$

In the above equations, *FO(c)* defines the set of fan-out nodes of *c* (i.e., the nodes to which there is a timing arc from node *c*), $R_l(a)$ is the late required time at the node *a*, d_{ac} is the delay of an edge from node *a* to node *c*, $R_e(a)$ defines the early required time at the node *a*.

The difference between the late required time and the late arrival time at the node *c* is defined as the *late slack* at that node and is given by:

$$S_l(c) = R_l(c) - t_{Al}(c) \quad (4\text{-}5)$$

Similarly, the *early slack* at node c is defined by:

$$S_e(c) = R_e(c) - t_{Ae}(c) \quad (4\text{-}6)$$

Figure 4-2 illustrates the concepts of *arrival time* (i.e., data arrival), *required time* (i.e., data required) and *slack* in the context of the STA. In this figure we can observe an example of static timing verification between the signals of the data node (represented in the figure as *D*)

and the clock node (represented in the figure as *clk*) of the flip-flop *b* (F-Fb). The clock signal arrives at the circuit at the node CLK and subsequently this signal is divided in two in order to feed the *arrival time path* and the *required time path*. In the figure 4-2b we can observe that the signal of data arrival is controlled by the cells U1, U2, flip-flop *a* (F-Fa) and U3 while the clock signal is controlled by the cells U1 and U4.

Let us evoke that a flip flop is a *register*, in which the output (Q node) follows the previously input stored data when the clock signal is high[1] (*positive–edge-triggered)*.

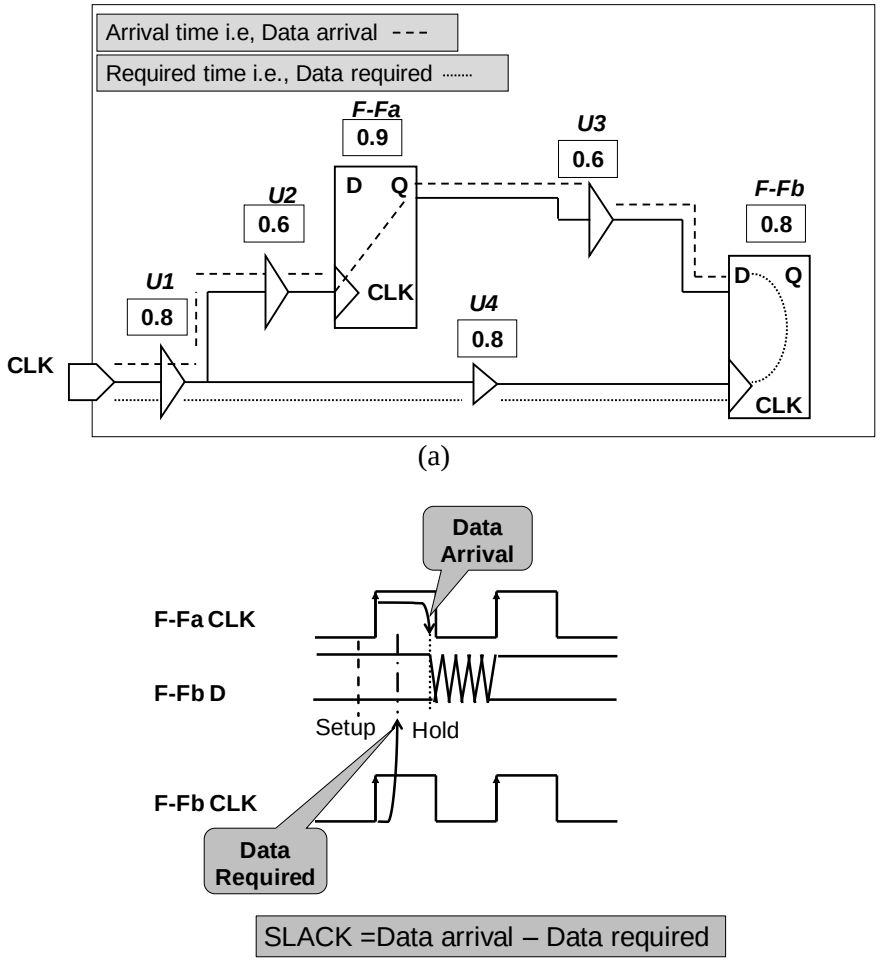


Figure 4-2 Static timing verification of F-Fb. a) Circuit schematic of .b) Diagram of timing: data and clock signals

[1] In typical flip-flops, the input data is stored either on the rising edge (low-to-high transition) or on the falling edge (high-to-low transition) of the clock signal. These flip-flops are known as *positive–edge-triggered* and *negative-edge-triggered*, respectively.

In order to assure that the F-Fb stores a correct new data, this data should arrive and hold to a stable level (data arrival) for a certain period of time before (*Setup*) and after (*Hold*) the rising edge of CLK (data required). In the more recent technologies (32nm and 28nm) the value of this time period is very small (around some picoseconds).

For instance, when a F-F has been affected by NBTI the timing balance between data arrival and data required can be broken provoking an error in the logic function of the circuit. In the section 4.1.6 it will present a practical example of this timing balance on a typical critical path.

4.1.4 Interconnection Parasitic and STA

The effect of interconnects on circuit performance have to be considered in order to carry out an accurate STA. The interconnect (or interconnect parasitic) represents the effective load seen by the driving gate in terms of resistances and capacitances allowing to calculate the delay of a given path. Figure 4-3 shows the representation RC of a wire of metal between two cells.

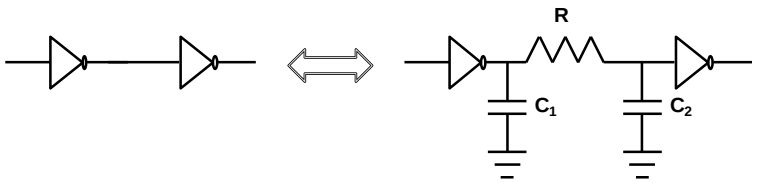


Figure 4-3 Illustration of the interconnect load for static timing analysis

Let us evoke that a STA is done after the place and route step [Kilts06] in order take into account in the circuit delay compute the delay contribution of the interconnect parasitic.

4.1.5 Process and Environmental Variations

Unavoidable variations in the environmental parameters (such as temperature, supply voltage, etc), provoke variations in circuit performances. Moreover, the disturbances present in IC

manufacturing processes also cause variations in device parameters and circuit performances [Zhang95]. As a result, STA have to consider the effect of process and environmental variations. Typically, statistical process and environmental variations are considered by performing analysis using the PVT approach (§3.3). Usually, STA is performed for three cases: worst case, nominal case and best case in order to check the circuit performance in whole range of conditions.

4.1.6 Timing analysis

At standard cell level the analysis is carried out by using the lookup table generated during the library characterization step, in which the timing arcs from the inputs to the outputs of all library gates are characterized (cf., Fig. 4-4). Using these data, STA commercial tools (such as Synopsys Primetime®) calculate the path delays of a given design. In order to calculate the reliability impact on STM designs two STA are done. First, a fresh STA (at time t= 0), which does not considers the ageing phenomena. Secondly, an aged STA, which includes the ageing impacts on the timing. The results of the STA are presented by Primetime (PT) as a timing report[2]. Note that STA are performed considering signal probability (SP)[3] at 50%. Figure 4-4 shows an example of a timing report generated from of a fresh STA and an aged STA (at time $t= t_{age}$) [PT09]. A timing report provides detailed timing information about any number of requested paths. Note that in this example we analyzed only one path. The report includes the *start* and *end* pin names of the path under consideration, the delays of the gate path (*Incr*. column) and the sum of all gate delays (*Path.* column). This report is divided in three parts. In the first part, the data path information is presented and the total delay of the data path is represented by the line *data arrival time*. Second part describes the clock path. The total delay of the clock path is defined by the line *data required.* Finally, in order to verify if the path under consideration has a timing violation, the *slack* is calculated (third part). In this case, the

[2] PT is the STA tool used in this manuscript.

[3] In the STM characterization flow the value of the SP is set to 50%.

data arrival time is smaller than the data required, thus the slack is positive after stress (non timing violation).

In the third column of the timing report (cf., Fig.4-4) we can observe the intrinsic delay (time propagation) of the different path cells (at time $t= 0$) and in the fifth column we can see these same delays at time $t= t_{age}$. We can observe that the cell delays have increased on an average of 4% after stress, which does not provoke a timing violation since the slack is positive; however this reduces the speed performance of this circuit path. It is worth to notice that in this example we only illustrate a small portion of a circuit block, however a real product is composed by thousands of blocks, thus although the NBTI does not cause an error in the functionality of the block, the progressive lost of performance in chain of block can induce an error in the logic function of the product.

		STA without NBTI at time *t=0*		STA with NBTI at time *t= tage*	
Header	Startpoint : CLK1				
	Endpoint: FFb/CLK				
	Path Group: CLK				
	Point	**Incr**	**Path**	**Incr**	**Path**

Data Arrival	clock Clk (rise edge)	0.00	0.00	0.00	0.00
	clock network delay(propagated)	1.10	1.10	1.10	1.10
	U1/Y	0.81	1.91	0.82	1.92
	U2/Y	0.60	2.51	0.64	2.56
	F-Fa/CLK	0.00	2.51	0.00	2.56
	F-Fa/Q	0.90	3.41	0.95	3.51
	U3/Y	0.61	4.02	0.66	4.17
	F-Fb/D	0.00	4.00	0.00	4.17
	F-Fb/Q	0.80	4.82	0.85	5.02
	data arrival time		*4.82*		*5.02*

Data Required	clock Clk(rise edge)	4.00	4.00	4.00	4.00
	clock network delay (propagated)	1.00	5.00	1.00	5.00
	U1/Y	0.81	5.81	0.83	5.83
	U4/Y	0.79	6.60	0.84	6.67
	F-Fb/CLK	0.0	6.60	0.0	6.67
	data required		*6.60*		*6.67*

Slack	*data required*		*6.60*		*6.67*
	data arrival time		*-4.82*		*-5.02*
	slack		*1.78*		*1.65*

Figure 4-4 Primetime report timing generated from STA fresh (without NBTI effects) and a STA age (with NBTI effects, showing the timing library characterization values (orange circles). This report timing is based on the circuit described in figure 4.2. Both STA are carried out at 50% of SP

As mentioned earlier, due to the technical difficulties in terms of CPU and engineer time, in the industrial ambit is impractical to execute the reliability STA for each one of the mission profile scenarios. Therefore, we use the gate model in order to estimate analytically the reliability impact on the design (or reliability STA) taking into account the mission profile. Next section illustrates the application of NBTI gate level model on complex designs.

4.2 Evaluation of NBTI degradation at system level

In order to validate the NBTI gate modeling approach at system level, a silicon-CAD combined ageing analysis was performed. We used the accelerated stress results from qualification reports (experimental experiences) of two IPs carried out in 45nm STM technology: a Low-Density Parity-check (LDPC) codec circuit and a Cortex–A9 ARM, with the objective to compare the F_{max} degradation measures (defined as the shift in the maximum IPs operating frequency (ΔF) provoked by the NBTI) against CAD data estimations obtained through of the analytical reliability STA. Each IP qualification report contains the Operating Lifetests (HTOL) experiments including various stress voltages, stress times and several stress temperatures.

On the other hand, in order to obtain the delay degradation dynamic of critical paths a fresh STA (defined as a STA at time $t= 0$) and an ageing STA (at t_{age}) are executed on the two databases. These STA were executed using the transistor model presented in chapter 2 (§2.4.2.3). Subsequently the different IP HTOL stress conditions were translated into their equivalents Power-On Time (POT) values (§3.5.1). The gate model was fed with both the POT values and the delay degradation dynamic values of the gates in order to generate the analytical reliability STA.

4.2.1 NBTI Analysis on LDPC

The LDPC circuit was carried out in a 45nm Low-Power bulk technology. This LDPC codec supports the 12 modes required by the 802.11n standard (Wifi). This circuit is commonly used to decode noisy transmission channel. The top circuit and the LDPC decoder micro

architecture are presented in figure 4-5. This IP was stressed at 10 different conditions including several stress time (up to 1000hrs), stress voltage for a total of 1200 parts tested.

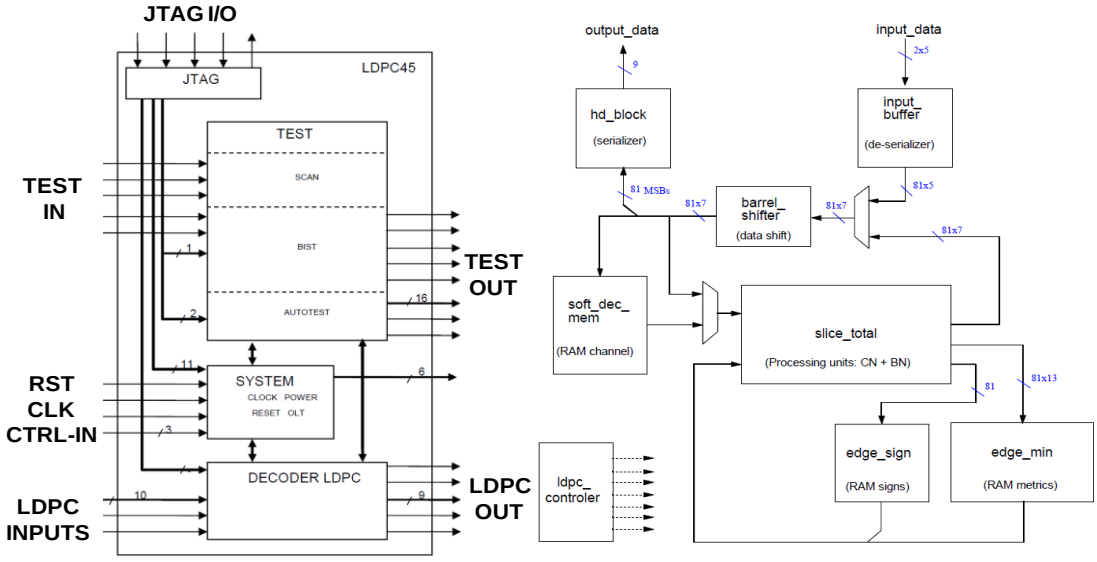


Figure 4-5 (a) Top circuit view and LDPC. (b) Decoder micro architecture

Figure 4-6, shows the histogram of the delay degradation at five different values of POTs. These histograms were calculated using the gate model for 60.000 more critical DATA paths (focusing on paths with *setup* violations). For the sake of comparison, we also include the experimental F_{max} degradation data (vertical lines).

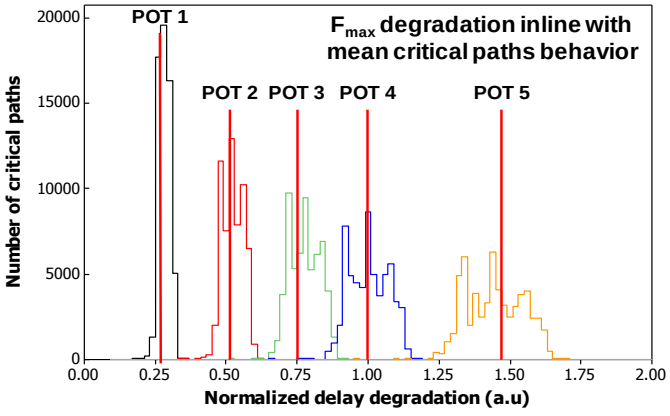


Figure 4-6 Histograms of delay degradations based on gate model for the most 60000 critical paths of LDPC IP with respect to the setup violations for various stress conditions (POTs); POT1<POT2<POT3<POT4<POT5

We can observe that the F_{max} degradation is very well aligned with the mean delay degradation of the whole critical path population under consideration. Thus, we consider that this approach provides an accurate way to predict product macroscopic feature (such as F_{max}) degradation.

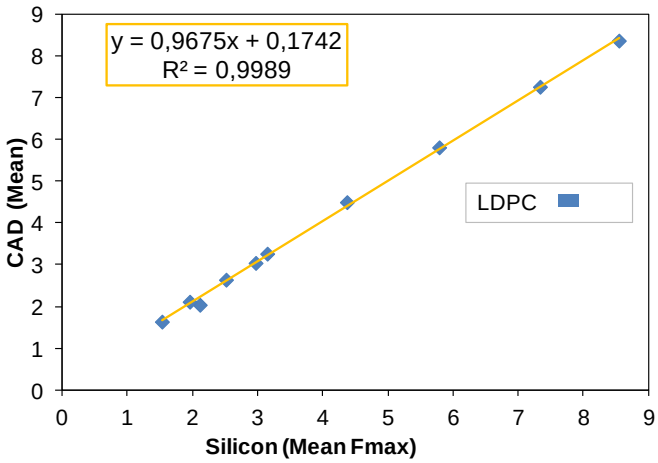


Figure 4-7 Correlation plot between the mean delay degradation for the LDPC set of critical paths under consideration and the mean CAD

Figure 4-7 shows the correlation between mean delay degradation (calculated from gate model) as a function of the experimental mean F_{max} degradation for the complete set of POTs (10 values in total) under consideration in our testcase. As we can observe the predictions of the gate model have a correlation very close to 1:1 with the experimental observations. The gate model reproduces the delay degradations with a margin of error of 2% of the total degradation. Therefore if the mean delay degradation is worth 10%, the error made by the gate model is 0.2%.

4.2.2 NBTI Analysis on Cortex A9 processor

The ARM Cortex™-A9 processor provides high levels of performance and power efficiency. It is usually used in applications such as: cellular phones, PDAs and games consoles. The A9 processor supports the configuration of 16, 32 or 64KB four way associative L1 caches, with up to 8MB of L2 cache through the optional L2 cache controller. On a 45 nm (STM standard

General Purpose technology) it reaches a clock frequency of 800 MHz [Arm09] (cf., Fig 4-8). This IP was stressed at 3 different conditions including several stress time, which represent a real mission profile (up to 1000hrs) for a total of 160 parts tested.

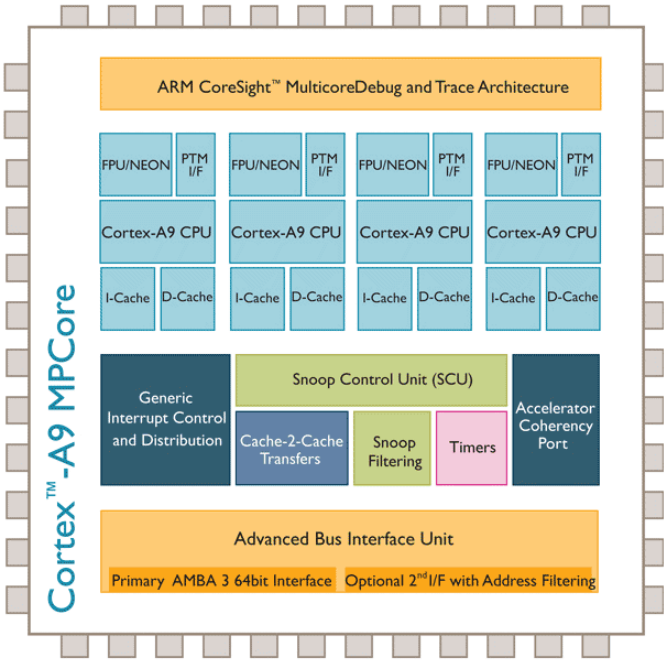
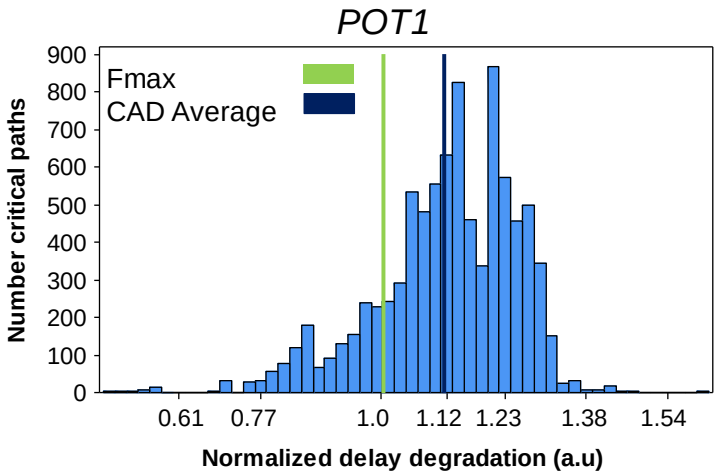


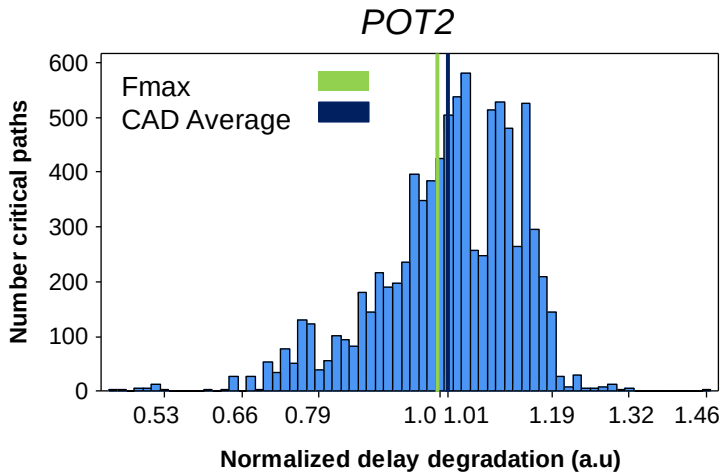
Figure 4-8 The ARM Cortex™-A9 processor core

Figure 4-9 shows the histograms of delay degradations predicted by the gate model for three stress different conditions including several stress time (up to 1000hrs). For the comparison, the experimental F_{max} degradation for each one of the three stress conditions is represented by the vertical light line.

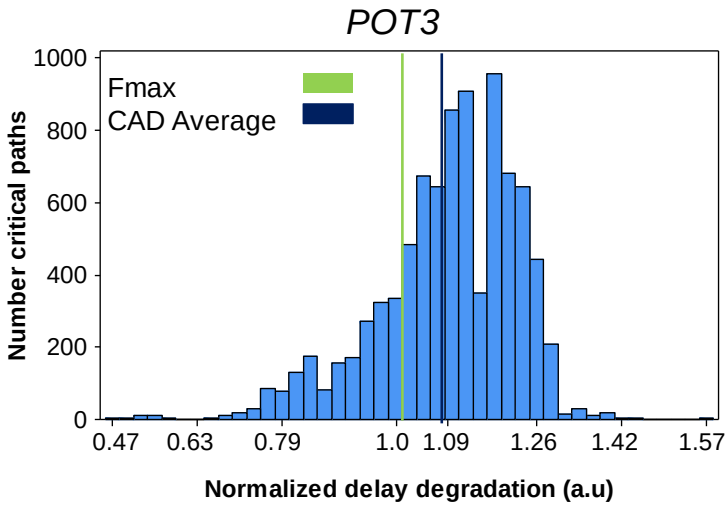
It is worth noticing the good agreement between the CAD averages (the estimated F_{max} degradations) versus the experimental F_{max} degradations.



(a)



(b)



(c)

Figure 4-9 Delay degradations normalized with respect to average of measured F_{max} for the slowest 8825 critical paths of A9 IP for three different stress conditions (POTs). The experimental F_{max} degradation for similar stress conditions is well reproduced.

Figure 4-10 shows the correlation between the mean of the F_{max} degradations with respect to the CAD mean for the A9 and LDPC critical path populations. This shows that our modeling approach can be used successfully to analyze complex databases.

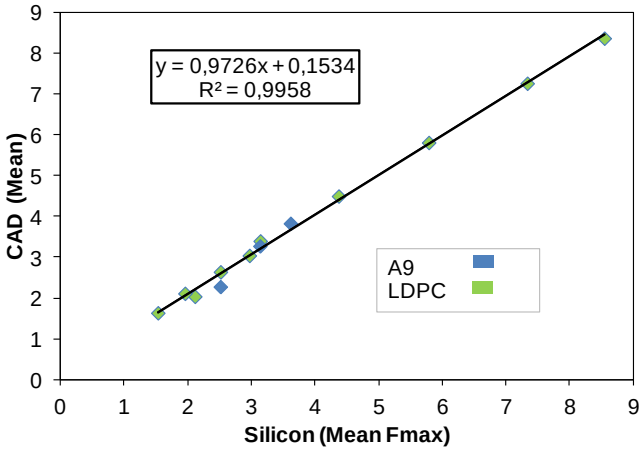


Figure 4-10 Correlation plot between the mean delay degradation for both LDPC and A9 set of critical paths and the mean CAD, which was calculated from NBTI gate level

4.3 Conclusion

In this chapter, we have presented the application of the NBTI model methodology on complex digital applications. We have demonstrated that it is possible by a bottom-up approach to build a gate-level model with sufficient accuracy to allow direct comparison with experimental degradations at system-level.

Chapter 5

Conclusions and Future Work

5.1 Conclusion

This thesis presents a reliability study on CMOS technologies addressing digital applications. Two major topics are considered in this manuscript, the NBTI and HCI reliability effects at gate level. The NBTI gate analytical model was analyzed in details and used in order to propagate the timing degradation upwards into the design hierarchy.

In the first part of the thesis a deep-analysis is carried out on a large data base of standard cells with the purpose of understanding the influences of the electrical and design parameters in the delay degradation at gate level. From this study, a three parameter analytical approach was proposed to model NBTI effects at gate level. In the same way, a HC dedicated database was also analyzed in order to provide the performance delay degradation caused by HC phenomena, this data base is based on 45 nm and 32 nm standard cells. From of these results, a HC analytical model was proposed in the chapter 3 (section 3.7).

The second part of the manuscript discussed the NBTI impact at system level. The gate level NBTI model was used to compute the effects of NBTI degradation on two IPs: a Low Density Parity Check (LDPC) codec circuit and a cortex A9 processor. Several analytical reliability Static Timing Analysis were calculated using the gate model and subsequently were compared with the experimental data measured on silicon. The NBTI gate level approach was validated with respect to silicon and was showed that the model is accurate to reproduce the NBTI degradation behaviour at system level.

5.2 Future Work

This study contains broad reliability analysis at three levels: transistor level, gate level and system level for the NBTI wearout mechanism. However, the characterization framework of the HC phenomena at system level was not carried out. A substantial amount of efforts is still needed between different teams (design and electrical characterization) to achieve the reliability analysis.

Suggestions on the future work are listed below:

The study of the interaction of HC and NBTI phenomena should continue at different abstraction levels, but especially at system level.

In the near future, STM CMOS bulk technology will be replaced by the FDSOI. In this perspective, the influences of the wearout mechanisms on IC performance will probably change. It is mandatory that the new behavior of the wearout phenomena be carefully analyzed. However, the reliability modeling framework presented in this thesis should be able to be integrated with a little effort the new FDSOI reliability analysis.

Finally, some research efforts should be done with the purpose of integrating the Time-Dependent Dielectric Breakdown (TDDB) impacts in a three-part gate model (HC, NBTI and TDDB).

Appendix A

Technical limitations to include HC phenomenon on library characterization flow

In current STM library characterization flow only the NBTI mechanism is considered. Due to the stronger dependence of HC with C_L and S_I (§3.2.1.2), this phenomenon is not analyzed in the characterization flow. Let us briefly look back on main characteristics of reliability analysis. First, it is divided in three parts: a.) fresh simulation, b.) reliability simulation (§2.4.2), c.) characterization simulation[1].

For the NBTI analysis the fresh and the reliability simulations can be executed for the worst case pair C_L, S_I (in terms of timing degradation) and only the characterization simulations are performed for the whole range of (C_L, S_I) previously defined in the technical specifications. This consideration is based on weak dependence of NBTI with C_L, S_I (cf., Fig. A-1). In this way all library is analyzed and its performance is ensured. On the other hand in HC analysis these three types of simulations must to be executed for the same pair (C_L, S_I). This requires larger CPU resources and involves larger execution time.

[1] The simulation of characterization uses the degraded SPICE models.

It is worth noting that the conditions of characterization simulation are exactly the same as the conditions of the fresh simulation.

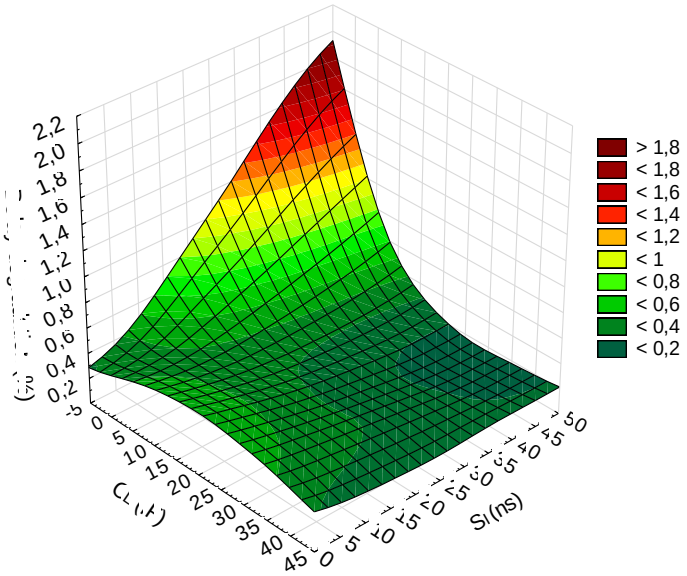


Figure A-1 NBTI impact on inverter electrical response

Figure A-1 shows the dependence of timing performance degradation with the input signal ramp S_I, and the load capacitance C_L. It can be observed that the delay degradation has a weak dependence with respect to S_I and C_L. Let us note that for the small values of load capacitance and larger values of input signal slope, the increase in the delay degradation is given by the electrical conduction aspects of the inverter cell (§3.1.1) (§3.2.1.1). This degradation in the delay is not defined by the NBTI phenomenon.

Appendix B

SRAM Silicon/CAD comparison table

In the table B-1 the first column defines operation (*i.e.*, write or read) under consideration. The second column shows both the initial and the end points of the timing path. Columns three and four show the data measured on silicon, before and after the stress, respectively. Finally the columns five and six describe the data obtained by Eldo simulations (fresh and ageing, respectively).

OPERATION	***TIMING PATHS***	***SILICON***		***ELDO***	
		Fresh (t=0)	***Ageing***	***Fresh (t=0)***	***Ageing***
WRITE 0 AT BOTTOM ROW	(1) to (2)	120	117	43	43
	(2) to (3)	280	333	228	260
	(3) to (4)	570	486	564	577
	(3) to (5)	590	504	580	593
	(3) to (6)	840	783	766	793
READ 0 AT BOTTOM ROW	(1) to (2)	100	90	42	43
	(2) to (3)	250	306	220	247
	(3) to (4)	280	378	339	351
	(3) to (5)	320	270	354	366
	(3) to (6)	570	540	542	568
	(3) to (7)	400	378	384	396
	(7) to (8)	160	153	179	180
WRITE 1 AT TOP ROW	(1) to (2)	105	99	42	43
	(2) to (3)	300	324	224	251
	(3) to (4)	480	450	556	568
	(3) to (5)	500	468	573	585
	(3) to (6)	750	738	758	785
READ 1 AT TOP ROW	(1) to (2)	110	126	42	43
	(2) to (3)	230	288	222	249
	(3) to (4)	290	333	340	352
	(3) to (5)	320	270	355	367
	(3) to (6)	580	513	543	569
	(3) to (7)	400	360	385	397
	(7) to (8)	175	162	174	176

Table B-1 Silicon measures versus CAD data (Eldo simulations) for a SRAM

References

[Alam03] M.A. Alam, “A critical examination of the mechanics of dynamic NBTI for PMOSFETs”, Electron Devices Meeting. Technical Digest. IEEE International , vol., no., pp. 14.4.1- 14.4.4, 8-10, 2003.

[Alam04] M. A. Alam. “A comprehensive model of Pmos NBTI degradation”, Microelectronics Reliability, v 45, no. 1, pp 71-81, 2005

[Alam06] M.A. Alam, “On the Reliability of Micro-electronic Devices: An Introductory Lecture on Negative Bias Temperature Instability”. http://nanohub.org/resources/193. in Nanotechnology 501 Series, 2006.

[Ang05] D.S. Ang, S. Wang, C.H. Ling, “Evidence of two distinct degradation mechanisms from temperature dependence of negative bias stressing of the ultrathin gate p-MOSFET”, Electron Device Letters, IEEE , vol.26, no.12, pp. 906- 908, 2005.

[Arm09] “The ARM Cortex-A9 processors”, White paper, http://www.arm.com/files/pdf/ARMCortexA-9Processors.pdf, 2009.

[Arnaud00] L. Arnaud, G. Tartavel, T. Berger, D. Mariolle, Y. Gobil, I. Touet, Microelectronics Reliability 40, 77-86, 2000.

[Aur87] S. Aur, D. E. Hocevar, Y. Ping, “Circuit hot electron effect simulation”, Electron Devices Meeting, 1987 International , vol.33, no., pp. 498- 501, 1987.

[Bravaix01] A. Bravaix, D. Goguenheim, V. Huard, M. Denais, C. Partharasaty et al. , “Hot carrier reliability study of second and first impact ionization degradation in 0.15mm channel length N-MOSFETS”, Micro. Eng., 2001.

[Bravaix09] A. Bravaix, C. Guerin, V. Huard, D. Roy, D, J. M. Roux, E. Vincent, “Hot-Carrier acceleration factors for low power management in DC-AC stressed 40nm NMOS node at high temperature”, Reliability Physics Symposium, 2009 IEEE International , vol., no., pp.531-548, 26-30, 2009.

[Bude95] J. D. Bude, M. Mastrapasqua, “Impact ionization and distribution functions in sub-micron nMOSFET technologies”, Electron Device Letters, IEEE , vol.16, no.10, pp.439-441, Oct 1995.

[Burks 95] T. Burks, K. A. Sakallah, T.N. Mudge, “Critical paths in circuits with level-sensitive latches”, IEEE Trans. Very Large Scale Integration Systems, IEEE Transactions on , vol.3, no.2, pp.273-291, 1995.

[Chen03] G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng, D. L. Kwong, “Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling”, Electron Device Letters, IEEE , vol.23, no.12, pp. 734- 736, 2002.

[Chen07] Chen Wai-Kai, “The VLSI Handbook”, pp. 28-1, 32-6. Ed. CRC Press.

[Daga99]. A comprehensive delay macro modeling for submicrometer CMOS logics

[Dharchoudhury06] A. Dharchoudhury, D. Blaauw, S. Ganguly, “The VLSI handbook - Timing and Signal Integrity Analysis” pp. 63-3, 63-34, Ed. CRC Press, 2006.

[Deal67] B.E. Deal, M. Sklar, A.S. Grove, E.H. Snow, “Surface-State Charge of Thermally Oxidized Silicon”, J. Electrochem. Soc., 114, p 266, 1967.

[Devine97] R. A. B. Devine et al. “Structural nature of the Si/SiO [sub 2] interface through infrared spectroscopy, Appl. Phys. Lett. 68, 3108 ,1996.

[DiMaria95] D. J. DiMaria, D. A. Buchanan, J.H. Stathis, R. E. Stanlbush, “Interface states induced by the presence of trapped holes near the silicon–silicon-dioxide interface”, Journal of Applied Physics , vol.77, no.5, pp.2032-2040, 1995.

[Dorf06] R. C. Dorf and J.A. Svoboda, “Introduction to electric circuits”, pp 1-200, Ed. John Wiley and Sons, 2006.

[Eldo07] Eldo user manual UDRM version 2010

[Frohman71] Frohman-Bentchkowsky D, “A fully decoded 2048-bit electrically programmable FAMOS read-only memory”, Solid-State Circuits, IEEE Journal of , vol.6, no.5, pp.301-306, 1971.

[Fujieda05] S. Fujieda S, Y. Miura, M. Saitoh, Y. Teraoka, A. Yoshigoe, “Characterization of interface defects related to negative-bias temperature instability in ultrathin plasma-nitrided SiON/Si⟨100⟩ systems” Microelectronics Reliability, v 45, n 1, p 57-64, 2005.

[Guerin07] C. Guerin, V. Huard, A. Bravaix, “The Energy-Driven Hot-Carrier Degradation Modes of nMOSFETs”, Device and Materials Reliability, IEEE Transactions on , vol.7, no.2, pp.225-235, 2007.

[Guerin07b] C. Guerin, V. Huard, A. Bravaix, “Hot-carrier damage from high to low voltage using the energy-driven framework”, Device and Materials Reliability, IEEE Transactions on , vol.7, no.2, pp.225-235, 2007.

[Guerin08] C. Guerin, PhD dissertation, University of Provence d'Aix-Marseille I, 2008.

[Guerin09] C. Guerin, V. Huard, A. Bravaix, “General framework about defect creation at the Si/SiO_2 interface”, Journal of Applied Physics , vol.105, no.11, pp.114513-114513-12, 2009.

[Haggag06] A. Haggag, M. Kuffler, D. Zhang, M. Sadaka, P. Grudowski, M. Moosa, “Novel Model for HCI Degradation and Impact of Conventional and Non-Conventional Scaling”, Reliability Physics Symposium Proceedings, 2006. 44th Annual., IEEE International , vol., no., pp.737-738, 26-30, 2006.

[Hess99] K. Hess, L. F. Register, W. McMahon , B. Tuttle, O. Aktas, U. Ravaioli, J.W Lyding , and I.C. Kizilyalli. Theory of channel hot carrier degradation in MOSFETs, Physical B, 1999.

[Hook05] Hook TB et al, R. Bolam, W. Clark, J. Burnham, N. Rovedo,. L. Schutz, “Negative bias temperature instability on three oxide thicknesses (1.4/2.2/5.2 nm) with nitridation variations and deuteration” Microelectronics Reliability, v 45, n 1, p 47-56, 2005.

[Huard02] V. Huard et al., “Interface traps and oxide charges during NBTI stress in pMOSFETs”, Integrated Reliability Workshop Final Report, 2002. IEEE International , vol., no., pp. 135- 138, 21-24, 2002.

[Huard03] V. Huard, F. Monsieus, G. Ribes, S. Bruyere, “Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs”, Reliability Physics Symposium Proceedings, 2003. 41st Annual. 2003 IEEE International , vol., no., pp. 178- 182, 2003.

[Huard04] V. Huard, M. Denais, “Hole trapping effect on methodology for DC and AC negative bias temperature instability measurements in PMOS transistors”, Reliability Physics Symposium Proceedings, 2004. 42nd Annual. 2004 IEEE International , vol., no., pp. 40- 45, 25-29,2004.

[Huard06] V. Huard, M. Denais, C. Parthasarathy “NBTI degradation: From physical mechanisms to modeling”, Microelectronics Reliability, v 46, n 1, p 1-23, 2006.

[Huard07a] V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammase, D. Barge, C. Ouvrard , “New characterization and modeling approach for NBTI degradation from transistor to product level”, Electron Devices Meeting, 2007. IEDM 2007. IEEE International , vol., no., pp.797-800, 10-12, 2007.

[Huard07b] V. Huard, C. R. Parthasarathy, A. Bravaix, T. Hugel, C. Guerin, E. Vincent, “Design-in-Reliability Approach for NBTI and Hot-Carrier Degradations in Advanced Nodes”, Device and Materials Reliability, IEEE Transactions on , vol.7, no.4, pp.558-570, 2007.

[Huard09] V. Huard, C. R. Parthasarathy, A. Bravaix, C. Guerin, E. Pion, “CMOS device design-in reliability approach in advanced nodes,” Reliability Physics Symposium, 2009 IEEE International , vol., no., pp.624-633, 26-30 April 2009.

[Huard10] V. Huard et al. “Two independent components modeling for Negative Bias Temperature Instability”, Reliability Physics Symposium (IRPS), 2010 IEEE International, vol., no., pp.33-42, 2010.

[Huard10a] V. Huard, R. Chevallier, C. Parthasarathy, A. Mishra, N. Ruiz-Amador, F. Persin, V. Robert, A. Chimeno, E. Pion, N. Planes, D. Ney, F. Cacho, N. Kapoor, V. Kulshrestha, S. Chopra, N. Vialle, “Managing SRAM reliability from bitcell to library level”, Reliability Physics Symposium (IRPS), 2010 IEEE International , vol., no., pp.655-664, 2-6, 2010

[Huard12] V. Huard, N. Ruiz Amador, E. Pion, F. Cacho, D. Croain, V. Robert, S. Engels, P. Flatresse. “A predictive hierarchical approach to digital system reliability”, Reliability Physics Symposium, 2012 IEEE International. (This article will be published in the march 2012).

[Hu85]. C. Hu, C. Simon, Tam, H. Fu-Chieh, K. Ping-Keung, C. Tung-Yi, K.W. Terrill, "Hot-Electron-Induced MOSFET Degradation -- Model, Monitor, and Improvement", Solid-State Circuits, IEEE Journal of , vol.20, no.1, pp. 295- 305, 1985.

[Hu99] C. –K Hu, R. Rosenberg, H.S. Rathore, D. B. Nguyen, B. Agarwala, "Scaling effect on electromigration in on-chip Cu wiring" Interconnect Technology, IEEE International Conference , vol., no., pp.267-269, 1999.

[JEDEC74] JEDEC Solid State Tech. Association. "Early Life Failure Rate Calculation Procedure for Semiconductor Components" (JESD74), 2007.

[Jeppson77] K.O. Jeppson, C.M. Svensson. "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices". J. Appl. Phys., 48, pp 2004-2014, 1977.

[Jeppson94] K. Jeppson, "Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay", Solid-State Circuits, IEEE Journal of , vol.29, no.6, pp.646-654, Jun 1994.

[Jouppi87] N.P. Jouppi, "Timing analysis and performance improvement of MOS VLSI designs", IEEE Transactions on , vol.6, no.4, pp. 650- 665, 1987.

[Kaczer08] B. Kaczer, T. Grasser, P. J Roussel, J. Martin-Martinez, R. O'Connor, B.J. O'Sullivan, G. Groeseneken, "Ubiquitous relaxation in BTI stressing—New evaluation and insights", Reliability Physics Symposium, 2008. IRPS 2008. IEEE International , vol., no., pp.20-27, 2008.

[Kang07] K. Kang, Park S. P., Roy K. and Alam M. "Estimation of statistical variation in temporal NBTI degradation and its impact on lifetime circuit performance", Computer-Aided Design, 2007. ICCAD 2007. IEEE/ACM International Conference on , vol., no., pp.730-734, 4-8. 2007

[Kawakami00] Y. Kawakami, F. Jingkun, H. Yonezawa, N. Iwanishi, W. Lifeng, C. I-Hsien, N. Koike, C. Ping, Y. Chune-Sin, L. Zhihong, "Gate-level aged timing simulation methodology for hot-carrier reliability assurance", Design Automation Conference Proceedings of the ASP-DAC 2000. Asia and South Pacific, vol., no., pp.289-294, 9-9, 2000.

[Keating07] M. Keating, D. Flynn, R. Aitken, A. Gibbons, K. Shi." Low Power Methodology Manual for System-on-Chip Design ", pp 184,190. Ed. Springer, 2007.

[Kilts06] S. Kilts, “Advanced FPGA Design: Architecture, Implementation and Optimization”, pp 241-255, Ed. Wiley, 2006.

[Kimizuka00] N. Kimizuka, K. Yamaguchi, K, Imai, T. Iizuka, C.T. Liu; R.C. Keller, T. Horiuchi, “NBTI enhancement by nitrogen incorporation into ultrathin gate oxide for 0.10-μm gate CMOS generation”, VLSI Technology, 2000. Digest of Technical Papers. 2000 Symposium on , vol., no., pp.92-93, 2000.

[Krishnan03] A.T Krishnan, V. Reddy, S. Chakravarthi, J. Rodriguez, S. John, S. Krishnan, “NBTI impact on transistor and circuit: models, mechanisms and scaling effects MOSFETs”, Electron Devices Meeting, 2003. IEDM '03 Technical Digest. IEEE International , vol., no., pp. 14.5.1- 14.5.4, 8-10, 2003.

[Kufluoglu04] H. Kufluoglu and M. A. Alam, “A generalized reaction diffusion model with explicit H-”, *IEEE Trans. Electron Devices*, vol. 54, p.1101 , 2007.

[Kuo88] M. Kuo, K. Seki, P. Lee, J. Choi, P. Ko, C. Hu , “Simulation of MOSFET lifetime under AC hot-electron stress”, Electron Devices, IEEE Transactions on , vol.35, no.7, pp.1004-1011, 1988.

[Lasbouygues06] B. Lasbouygues, S. Engels, R. Wilson, P. Maurine, N. Azémard and D. Auvergne, “Logical effort model extension to propagation delay representation”, Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on , vol.25, no.9, pp.1677-1684, Sept. 2006

[Lasbouygues06a] B. Lasbouygues, “Analyse statique temporelle des performances en presence de variations de tension d’alimentation et de temperature.” PhD dissertation, Grenoble University I, 2006.

[Lasbouygues07] B. Lasbouygues, R. Wilson, N. Azémard, and P. Maurine, “Temperature and Voltage Aware Timing Analysis: Application to Voltage Drops”, Design, Automation & Test in Europe Conference & Exhibition, 2007. DATE '07 , vol., no., pp.1-6, 16-20 April 2007.

[Leblebici90] Y. Leblebici, S. M Kang, “An integrated hot-carrier degradation simulator for VLSI reliability analysis”, IEEE Int’l Conf. on Computer-Aided Design pp 400-403, 1990.

[Leblebici93] Y. Leblebici, S. M Kang, “Hot Carrier reliability of MOS VLSI circuits”, pp 19- 179 Ed. Kluwer Academic publishers, 1993.

[Liang86] M.S. Liang et al. “Inversion level capacitance and mobility of very thin gate oxide MOSFETs”. IEEE Trans. Electron Devices, ED-, 33 p 409, 1986.

[Lorentz09] D. Lorenz, G. Georgakos, U. Schlichtmann. “Aging analysis of circuit timing considering NBTI and HCI”, On-Line Testing Symposium. IOLTS 2009. 15th IEEE International, vol., no., pp.3-8, 24-26. 2009.

[Maurine02] P. Maurine, M Rezzoug, N. Azémard and D. Auvergne, "Transition time modeling in deep submicron CMOS," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on , vol.21, no.11, pp. 1352- 1363, 2002.

[Mead80] C. Mead, L. Conway, “Introduction to VLSI systems”, Ed. Addison-Wesley, 1980.

[Mihnea02] A. Mihnea, P. J. Rudeck, C. Chun, K. D. Prall, R. Ghodsi, “Direct observation of secondary ionization current in n-channel MOSFETs”, Electron Devices, IEEE Transactions on , vol.49, no.12, pp. 2301- 2307, 2002.

[Minehane00] S. Minehane, R. Duane, P. O’Sullivan, K.G McCarthy and A Mathewson, “Design for reliability”, Microelectronics Reliability, v 40, n 8-10, p 1285-94, 2000.

[Mitani02] Y. Mitani, M. Nagamine, H. Satake, A. Toriumi, “NBTI mechanism in ultra-thin gate dielectric - nitrogen-originated mechanism in SiON”, Electron Devices Meeting, 2002. IEDM '02. Digest. International , vol., no., pp. 509- 512, 2002.

[Moore03] Moore, G.E “ No exponential is forever: but "Forever" can be delayed! [semiconductor industry]”, Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC. 2003 IEEE International, vol., no., pp. 20- 23 vol.1, 2003.

[Ogawa95] S. Ogawa and N. Shiono, “Generalized Diffusion-Reaction Model for the Low-Field Charge Build Up Instability at the Si/SiO2 Interface,” Phys. Rev. B, vol. 51, no. 7, pp. 4218–4230, 1995.

[Pagey03] M. P. Pagey, “Hot carrier reliability simulation in aggressively scaled MOS transistors”, PhD dissertation, 2003.

[Parthasarathy06], C. R. Parthasarathy, PhD dissertation, University Aix-Marseille I, 2006.

[Paul06] B. Paul, K. Kunhyuk, H. Kufluoglu, A. Ashraful, K. Roy, “Temporal Performance Degradation under NBTI: Estimation and Design for Improved Reliability of Nanoscale Circuits”, Design, Automation and Test in Europe, 2006. DATE '06. Proceedings , vol.1, no., pp.1-6, 6-10, 2006.

[PT09] Primetime User Manual, version D.2009-12, 2009.

[Rauch09] Stewart E. Rauch. “Reliability Wearout Mechanisms in Advance CMOS Technologies”, pp 441-514, Ed. Wiley, 2009.

[Renesas08] “Semiconductor reliability handbook” http://documentation.renesas.com/eng/products/others/rej27l0001_reliabilityhb.pdf. 2008.

[Rosa09]. G. La Rosa, “Reliability Wearout Mechanisms in Advance CMOS Technologies”, pp 331-434, Ed. Wiley, 2009.

[Ruiz11] N. Ruiz Amador, V. Huard, E. Pion, F. Cacho, D. Croain, V. Robert, S. Engels, P. Flatresse, L. Anghel, “Bottom-up digital system-level reliability modeling”, Custom Integrated Circuits Conference (CICC), IEEE , vol., no., pp.1-4, 2011.

[Sanjay07] S. V. Kumar, C.H. Kim, S.S. Sapatnekar, “NBTI-Aware Synthesis of Digital Circuits”, Design Automation Conference, 2007. DAC '07. 44th ACM/IEEE , vol., no., pp.370-375, 4-8 June 2007.

[Sakallah90] K. A. Sakallah, T.N. Mudge, and O.A. Olukotun, “Timing verification and optimal clocking of synchronous digital circuits”, Proc. IEEE Intl.Conf. ICCAD-90. Digest of Technical Papers., 1990 IEEE International Conference on, vol., no., pp.552-555, 11-15, 1990.

[Sakurai90] T. Sakurai, R. Newton, “Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas”, Solid-State Circuits, IEEE Journal of , vol.25, no.2, pp.584-594, Apr 1990.

[Schreier78] L. A. Schreier, “Electrostatic Damage Susceptibility of Semiconductor Devices”, Reliability Physics Symposium, 1978. 16th Annual , vol., no., pp.151-153, 1978.

[Schoder05]. D. K. Schroder, “Negative Bias Temperature Instability: Physics, Materials, Process, and Circuit Issues”, IEEE Solid-State Circuit Society, 2005.

[Shen06] C. Shen, M-F Li; C.E. Foo, T. Yang, D.M. Huang, A. Yap, G. S. Samudra, Y.C. Yeo, “Characterization and Physical Origin of Fast Vth Transient in NBTI of pMOSFETs with SiON Dielectric”, Electron Devices Meeting, 2006. IEDM '06. International , vol., no., pp.1-4, 11-13, 2006.

[Sheu89] B.J. Sheu, W. J. Hsu, B.W. Lee, “An integrated-circuit reliability simulator-RELY”, Solid-State Circuits, IEEE Journal of, vol.24, no.2, pp.473-477, 1989.

[Shockley61] W. Shockley, “Problems related to p-n junctions in silicon.” Solid-State Electronics, Volume 2, pp. 35–67, 1961.

[Stesmans00] A. Stesmans, V. V. Afanas, “Paramagnetic defects at the interface of ultrathin oxides grown under vacuum ultraviolet photon excitation on (111) and (100) Si”, Appl. Phys. Lett. 77, 1469, 2000.

[STA00] S.E. Start, E. N. Comparini, “Constructing Better Non-linear delay models (NLDM) to improve timing closure”. SNUG Boston 2000.

[Strong06]. A. W. Strong, “Reliability Wearout Mechanisms in Advance CMOS Technologies”, pp 1-69, Ed. Wiley, 2009.

[Sullivan09]. T. D. Sullivan, “Reliability Wearout Mechanisms in Advance CMOS Technologies”, pp 517-564, Ed. Wiley, 2009.

[Synopsys11] Library Compiler: Modeling Timing, Signal Integrity and Power in Technology Libraries User Guide, version F-2011.09, pp 3-14 - 3-20, Synopsys, 2011.

[Takayanagi05] T. Takayanagi, J. L. Shin, B. Petrick, J. Y. Su, H. Levy, Pham Ha, J. Son, N. Moon, D. Bistry, U. Nair, M. Singh, V. Mathur, A. S. Leon, A.S, “A dual-core 64-bit ultraSPARC microprocessor for dense server applications” Solid-State Circuits, IEEE Journal of , vol.40, no.1, pp. 7- 18, 2005.

[Takeda83] E. Takeda, N. Suzuki, T. Hagiwara, “Device performance degradation to hot-carrier injection at energies below the Si-SiO2energy barrier”, Electron Devices Meeting, 1983 International, vol.29, no., pp. 396- 399, 1983.

[Takeda95] E. Takeda, C.Y. Yang, and A Miura-Hamada. “Hot carriers effects in MOS Devices”, Ed. Academic Press, 1995.

[Tam83] S. Tam, F.C. Hsu, R.S Muller, P.K. Ko, “Hot-Electron current in very short channel MOSFETs ”, IEEE Electron Device Letters, v EDL-4, n 7, 249-51, 1983.

[Tam84] S. Tam, P.K Ko, and C. Hu, “Lucky-electron model of channel hot-electron injection in MoSFETs”, IEEE Transactions on Electron Devices, v ED-31, n 9, p 1116-25, 1984.

[Tan05] S. S. Tan, T. P. Chen, C. H. Ang, L. Chan, “Mechanism of nitrogen-enhanced negative bias temperature instability in pMOSFET”, Microelectronics Reliability, Volume 45, Issue 1, January 2005, Pages 19-30.

[Taur98] Y Taur, T. H. Nin, “Fundamental of Modern VLSI Devices”, p70, pp 103-243, p258, Ed. Cambridge University Press, 1998”.

[Taur98a] Y Taur, T. H. Nin, “Fundamental of Modern VLSI Devices”. pp 134-135, Ed. Cambridge University Press, 1998.

[Teng94] C-C Teng, W. Sun, S. M. Kang, P. Fang, “iRULE: fast hot-carrier reliability diagnosis using macro-models”, Custom Integrated Circuits Conference, Proceedings of the IEEE 1994 , vol., no., pp.421-424, 1994.

[Tewskbury92] Tewskbury, PhD dissertation, MIT, 1992.

[Tsividis03] Y. Tsividis, “Operation and Modeling of the MOS Transistors”, Ed. Oxford University Press, 2003.

[Tu91] R. Tu, et al., “BERT-Berkeley Reliability Tolls”, Memorandum N.UCB/ERL M91/107, University of California, Berkeley, 1991.

[Wang00] F. Wang and S. Chang, “Scalable delay model for logic and physical synthesis”, The 16th IFIP World Computer Congress, ICDA, August 2000.

[Wang07] W. Wang, S. Yang, S. Bhardwaj, S. Vrudhula, F. Liu, Y. Cao, “The Impact of NBTI Effect on Combinational Circuit: Modeling, Simulation, and Analysis”, Very Large Scale Integration (VLSI) Systems, IEEE Transactions on , vol.18, no.2, pp.173-183, 2010.

[Weber91] W. Weber, M. Bronx, T. Kuenemund, H. Muehlhoff, and D. Schmitt-Landsiedel, “Dynamic degradation in MOSFETs-Part II: application in the circuit environment”, IEEE Trans. Electron Devices, pp.1859 -1867, 1991.

[Wu06]. E. Y. WU, R. –P. Vollertsen, J. Suñé, Reliability Wearout Mechanisms in Advance CMOS Technologies", pp 71-208, Ed. Wiley, 2009.

[Yang05] T. Yang, M. F. Li, C. Shen, C.H. Ang, Zhu Chunxiang, Y.C. Yeo, G. Samudra, S.C. Rustagi, M. B. Yu, "Fast and slow dynamic NBTI components in p-MOSFET with SiON dielectric and their impact on device life-time and circuit application", VLSI Technology, 2005. Digest of Technical Papers. 2005 Symposium on, vol., no., pp. 92- 93, 14-16, 2005.

[Yonezawa98] H. Yonezawa, F. Jingjun, Y. Kawakami, N. Iwanishi, W. Lifeng, A.I.-H Chen, N. Koike, C. Ping, Y. Chune-Sin, L. Zhihong, "Ratio based hot-carrier degradation modeling for aged timing simulation of millions of transistors digital circuits", Electron Devices Meeting. IEDM '98 Technical Digest., International, vol., no., pp.93-96, 6-9, 1998.

[Yonezawa00] Y. Kawakami, F. Jingkun, H. Yonezawa, N. Iwanishi, W. Lifeng, I-Hsien Chen, N. Koike, C. Ping, Y. Chune-Sin, L. Zhihong, "Gate-level aged timing simulation methodology for hot-carrier reliability assurance", Design Automation Conference. Proceedings of the ASP-DAC 2000. Asia and South Pacific , vol., no., pp.289-294, 9-9, 2000.

[Zhang95] J.C Zhang and M.A. Styblinski, "Yield and variability optimization of integrated circuits", Ed. Kluwer Academic, 1995.

[Zhihong06] L. Zhihong, B. W. McGaughy, J.Z. Ma, "Design tools for reliability analysis", Design Automation Conference, 43rd ACM/IEEE , vol., no., pp.182-187, 2006.

Résume

Introduction

L'industrie des semi-conducteurs a fait de grands progrès au cours des quatre dernières décennies. Des dispositifs MOSFET (*Metal Oxide Semiconductor Field Effect Transistors)* ont été intégrés par milliards dans les circuits a fin d'incorporer de plus en plus de fonctions sur chaque puce. Cette croissance phénoménale est directement due à la grande capacité d'intégration des transistors MOS (Very Large Scale Integration **VLSI**).

Avec la réduction continue des échelles des transistors MOS, l'impact des différents paramètres de fabrication, de design (tels que la qualité de processus de fabrication, la valeur de la tension d'alimentation et de la température) et des autres variations environnementales sur la performance temporelle des circuits intégrés est devenu extrêmement important notamment pour les technologies sub-65nm.

L'augmentation accrue de la température de fonctionnement des IC (Integrated Circuits) (par exemple dans un processeur) et la-non proportionnalité de la réduction des dimensions du transistor par rapport à la réduction de la tension d'alimentation a impacté de façon significative la fiabilité et la durée de vie de ces circuits. La modélisation précise de ces variations et de leurs effets dans la dégradation ou l'amélioration de la performance temporelle des circuits est devenue un des problèmes le plus critique dans le domaine de la conception des IC.

L'instabilité en température ou NBTI par ses sigles en anglais (Negative Bias Temperature lnstability) et la dégradation par Porteurs Chauds ou HC (Hot Carrier) sont considérés comme les phénomènes dominants de vieillissement dans les technologies actuelles. Cette thèse porte sur l'impact de ces deux phénomènes sur la performance des circuits numériques.

Les grandes lignes de la thèse sont précisées par la suite. La première section examine les principaux mécanismes de la dégradation NBTI et Porteurs Chauds. Elle présente aussi l'impact de ces phénomènes sur le rendement électrique du transistor.

La section 2 met l'accent sur la modélisation des phénomènes NBTI et Porteurs Chauds au niveau «porte logique». Dans cette partie, nous proposons un nouveau modèle de propagation de la dégradation de délai des portes logiques vers le haut dans la hiérarchie de conception (niveau système) tout en tenant compte le profil de mission de la porte considérée.

La section 3 porte sur la fiabilité NBTI au niveau système. Dans cette section nous présentons une analyse combinée de vieillissement (Silicium-CAO). Cette analyse est effectuée sur deux IPs: un détecteur/correcteur d'erreur de type LDPC et un processeur ARM cortex–A9. Des résultats CAO obtenus par simulations sont comparés à des mesures, qui ont été obtenues à partir des expériences expérimentales sur Silicium.

Finalement la section 4 présente les conclusions de cette thèse.

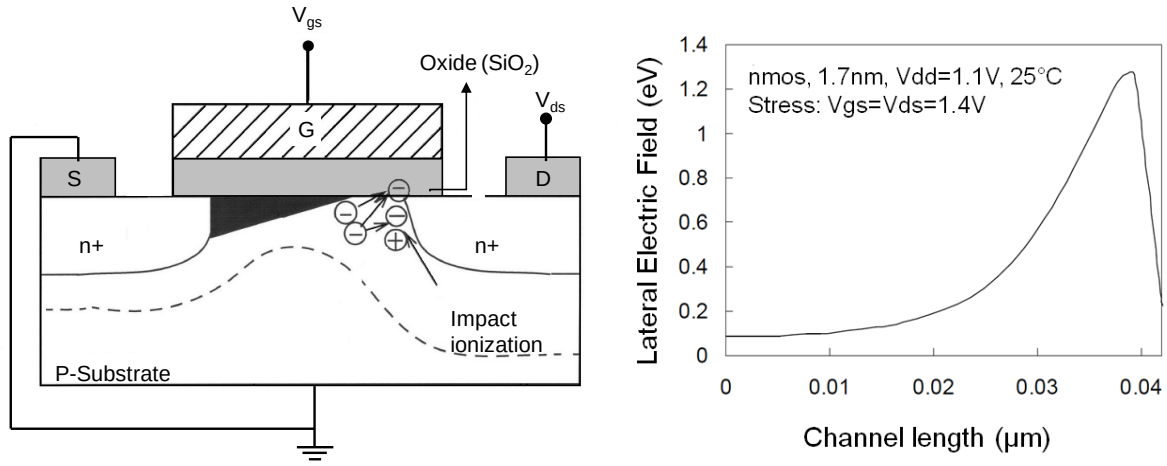
Les références aux sections sont identifiées par le symbole « § ».

1. Modélisation de la fiabilité de transistor

L'objective de cette section est d'introduire les notions générales à propos des phénomènes NBTI et HCI. Ensuite nous présenterons leur impact sur les performances du transistor. Ces deux phénomènes seront décrits ici de manière succincte, pour plus de détails je vous invite à consulter le chapitre 2 du manuscrit qui présente ces mécanismes de manière plus exhaustive.

1.1. Porteurs Chauds

La réduction de la taille du transistor a permis de réaliser des circuits de plus en plus complexes et rapides. Malheureusement, cette réduction dans taille des transistors comporte aussi des conséquences négatives. Une des plus graves est l'apparition du mécanisme Hot Carrier, qui est devenu un problème majeur pour la fiabilité à long terme du dispositif. Lorsqu'un transistor MOS fonctionne dans le mode saturation, un champ électrique latéral très élevé est généré près du drain (cf., Fig. 1-4). Ce vaste champ accélère les porteurs (électrons ou trous) dans le canal en les conduisant à devenir énergiques. L'injection de porteurs chauds dans l'oxyde modifie les caractéristiques électriques du transistor MOS puisqu'elle produit des défauts dans l'interface Si/SiO_2.



(a) (b)

Figure 1-1 (a) Configuration de stress «Porteurs Chauds» pour un diapositif nMOS. (b) Champ électrique latéral dans le canal d'inversion dans un transistor nMOS [Guerin08]

La figure 1-2 décrit le processus de dégradation produit par le mécanisme HC dans un nMOSFET [Rauch09]. Les trois mécanismes impliqués dans ce processus sont :

a – La dégradation HC est générée par les porteurs qui traversent le canal d'inversion (V_{gs}> V_{th}), lesquels peuvent acquérir assez d'énergie pour traverser le point de pincement du canal [Sze81].

b - Les porteurs énergétiques peuvent perdre leur énergie par ionisation par impact [Pagey03], contribuant ainsi à augmenter la valeur du courant du substrat.

c – Les porteurs chauds peuvent acquérir suffisamment d'énergie pour casser les liaisons d'hydrogène à l'interface Si/SiO_2 générant des états d'interface (ΔN_{it}). Ces défauts entraînent un changement des paramètres électriques du MOSFET. Cette instabilité dans le transistor peut générer des défaillances à long terme dans les circuits ou ces transistors sont localisés [Rauch09].

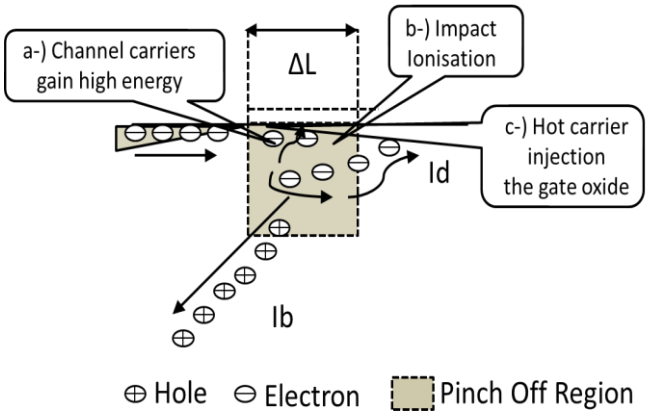


Figure 1-2 Représentation des mécanismes physiques qui contribuent à la dégradation HC pour le cas un transistor nMOS

1.1.1 Les modes d'injections des porteurs chauds

Etant donné que la dégradation provoquée par les porteurs chauds est fortement dépendante des conditions de polarisation, on peut distinguer quatre types de mécanismes d'injection HC:

I. *L'injection de porteurs chauds du canal (Channel Hot Carrier **CHC**)* est généré par des électrons avec une haute énergie cinétique «électrons chanceux» [Shockley61], qui ont suffisamment d'énergie pour surmonter la barrière d'énergie de l'interface de Si/SiO_2. Ce mécanisme se produit principalement lorsque la condition $V_{gs} = V_{ds}$ devient vraie.

II. *L'injection de porteurs chauds au drain en mode d'avalanche (Drain Avalanche Hot Carrier **DAHC**).* Elle a lieu quand une tension haute est appliquée au drain sous conditions de non saturation (V_{gs}< V_{ds}). Comme résultat on peut observer la

génération des forts champs électriques très près du drain qui accélèrent les porteurs dans la région de déplétion du drain.

III. *L'injection des porteurs chauds du substrat (Substrate Hot Carrier **SHC**)*, qui est dominante à V_{ds}= 0V, V_{gs} > 0V et V_{bs} > 0V. Elle est due à l'émission homogène d'électrons (courants des fuites) du substrat à l'interface.

IV. *L'injection des porteurs chauds secondaires* (*Secondarily generated Hot Carrier **SGHC)***. Elle est induite par l'ionisation par impact secondaire des porteurs dans le substrat [Bravaix01]. Elle peut être assiste aussi par des photons [Mihnea02]. Cette injection devient plus importante lorsque la condition V_{bs} < 0 est vraie.

1.2. NBTI

Le mécanisme de dégradation NBTI se produit seulement dans les dispositifs pMOS. Lorsque le transistor est soumis à une polarisation négative (à haute température), les trous dans le canal de conduction dissocient les liaisons Si-H à l'interface Si/SiO_2 [Deal67, Frohman71] générant des états d'interface (Fig. 1-3a). L'accumulation de ces défauts (N_{it}) entraîne une augmentation de la tension de seuil du transistor (Fig. 1-3b). Avec la suppression de la contrainte de tension une relaxation de la dégradation NBTI est observée. Cela est due à l'auto guérison partielle des pièges générées dans l'interface (composante recouvrable du NBTI), mais la composante permanente du NBTI reste, ainsi le rendement électrique des dispositifs pMOS est détérioré.

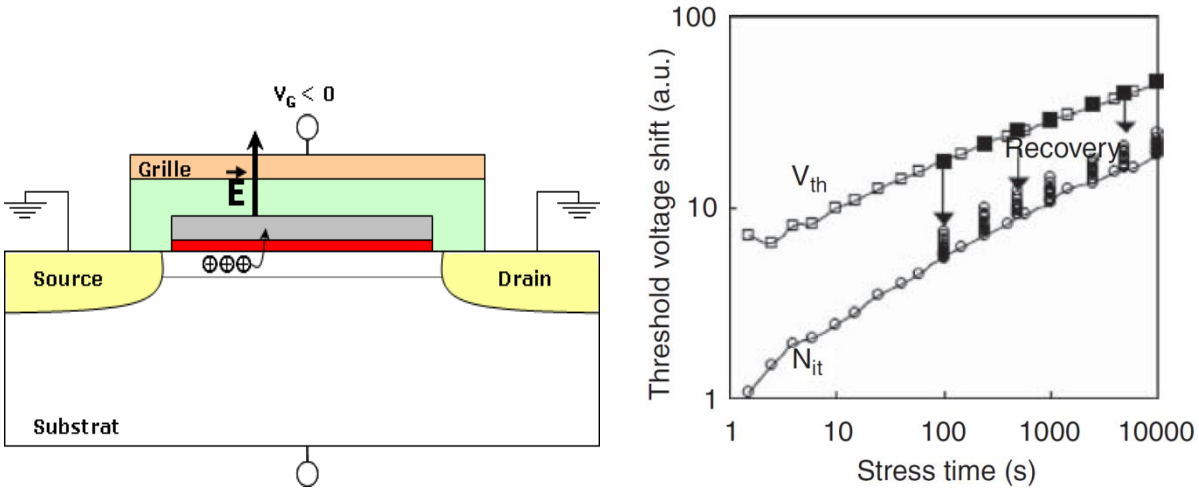


Figure 1-3 (a) Configuration électrique de un stress NBTI de type DC. (b) Evolution du V_{th} (carres) et du N_{it} (circles) pendant un stress NBTI DC [Huard04]

La dégradation NBTI affecte principalement la tension de seuil (V_{th}), le courant du drain linéaire (I_{dlin}), de saturation (I_{dsat}) et la transconductance (g_m). Le NBTI est donc une menace potentielle dans les technologies actuelles.

En raison de l'agressive tendance à la réduction des échelles adoptée à partir de la technologique 0,13 µm, une augmentation dans la dégradation NBTI a été observée [Alam03]. Cela est expliqué principalement par les raisons suivantes:

a -L'introduction du processus double du poli-silicium pour remplacer l'ancien procède de fabrication à canal enterré (n^+ poly) avec un canal surfacique (p^+ poly), afin de réduire les effets des canaux courts.

b -La lente réduction de la tension d'alimentation par rapport à la réduction de l'épaisseur d'oxyde (plus agressive) a accru les valeurs des champs électriques verticaux ainsi que la dégradation NBTI.

c -La réduction de la valeur de la tension d'alimentation a augmenté la sensibilité du circuit à la dégradation des paramètres électriques causés par l'NBTI (par exemple ΔV_{th}).

d -L'introduction de l'azote à l'interface Si/SiO_2 du pMOSFET afin de réduire la pénétration du bore et les courants de fuite a favorisé la dégradation NBTI.

1.3 Design-in Reliability

L'interaction entre les mécanismes de vieillissement est différente dans chaque niveau d'abstraction du design (ou conception). Cela augmente la complexité de l'analyse de fiabilité et en même temps réduit les marges de conception des IC. Par conséquent, les simulations de fiabilité doivent être précises pour pouvoir prévoir des performances «réalistes» du circuit à un temps donné. Design in-Reliability est une méthodologie qui se propose à améliorer l'évaluation de la fiabilité à tous les niveaux de conception des IC. Cette méthodologie renvoie à l'ensemble des modèles qui peuvent fournir une simulation et une capacité d'analyse élevées et précises pour n'importe quel produit. Cette thèse porte sur la modélisation des effets vieillissement NBTI et HCI dans le cadre de la méthodologie DiR présenté en [Huard07b].

Plusieurs approches ont été proposées depuis les années 90 [Minehane00] [Zhihong06]. Un exemple du flot de conception de fiabilité est montré dans la figure 1-4 [Leblebici93]. Ce flot

suit une implémentation standard avec une exception. Cela consiste en une bouclé de feedback entre l'étape «simulation de fiabilité» et l'étape de «conception du circuit». Une première analyse du vieillissement permet de valider les caractéristiques de conception (telles que: la topologie, la taille des transistors à l'intérieur des portes, etc.) afin d'assurer une durée de vie utile du circuit. Si le circuit est conforme aux spécifications de fiabilité, il peut être envoyé à fabrication.

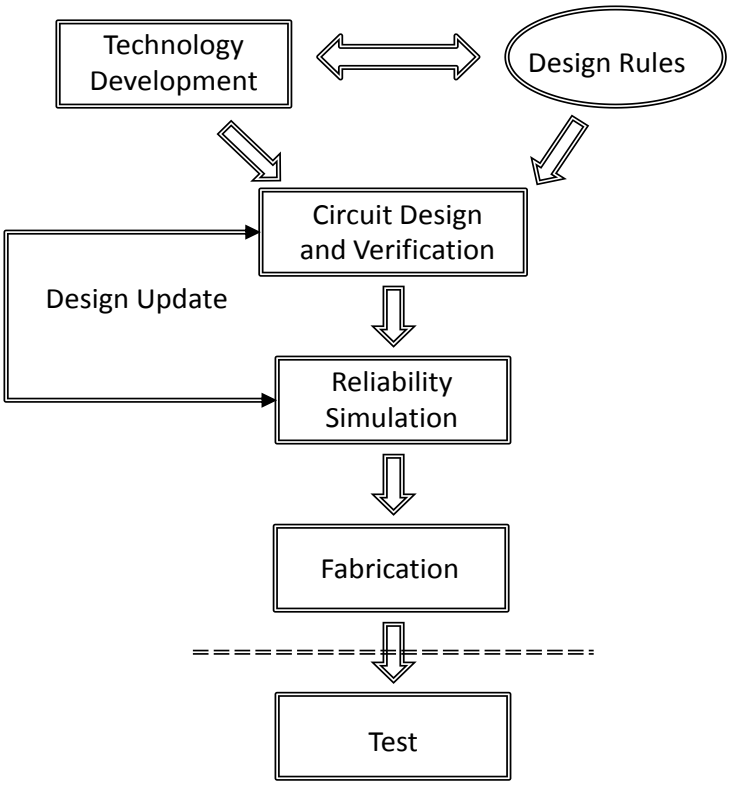


Figure 1-4 Méthodologie de conception pour des circuits du type VLSI en tenant en compte la notion de fiabilité

2 Modèle du NBTI et du HC à niveau Porte

A cause de la vaste gamme des applications et des sur exigences demandés aux circuits, les phénomènes HC et NBTI ont augmenté la complexité du processus d'optimisation de la conception des circuits dans les nœuds avancés. Dans ce contexte, un des principaux besoins actuels est de propager les délais du transistor vers le haut dans la hiérarchie de conception [Takayanagi05], afin de prévoir l'impact de la dégradation au niveau porte (§3.1) et système.

Nous utiliserons la méthodologie DiR afin de déployer les effets du NBTI et du HCI (en commençant au niveau transistor) hiérarchiquement au niveau porte et système. Tout au long du manuscrit nous fournissons des détails sur comment la méthodologie DiR est utilisée pour développer une modélisation des effets de la dégradation du type «bottom-up». Cette méthodologie s'adresse aux circuits numériques.

2.1. Modèle analytique de NBTI niveau porte

Dans cette section, nous allons proposer un nouveau modèle de la dégradation NBTI au niveau porte. Tout d'abord, nous analyserons l'effet du NBTI sur la porte la plus simple «l'inverseur». Une extension du modèle est présentée dans la section 3.5.2 du manuscrit, qui est validée pour des cellules plus complexes tels que XNOR, AOI, NAND, etc. De la même façon, cette méthodologie est valide pour modéliser les effets provoqués par la dégradation HC. Pour avoir plus de détails, le lecteur peut s'adresser au chapitre 3 du manuscrit (§ 3.7)

2.1.1 Première étape: L'Inverseur

Diverses simulations de vieillissement ont été effectuées à différents conditions de stress (en tenant en compte une grande gamme de valeurs de C_L et S_I). Nous avons cherché à vérifier si la Loi de Puissance de la dégradation NBTI par rapport au temps de stress observée au niveau transistor reste vraie aussi au niveau porte. Nous avons défini trois temps de stress, t_1 = 0.02 années, t_2 = 2 années et t_3= 20 années[1]. Figure 1-5 montre l'évolution de la dégradation du délai par rapport au temps de stress. Comme nous pouvons l'observer, le délai (Δd) augment linéairement avec le temps de stress. Conséquemment, on peut affirmer que la Loi de

[1] Afin de calculer l'impact du NBTI dans ces trois points nous avons utilisé le modèle L-R (§2.4.2.3).

Puissance de la dégradation du délai par rapport au temps de stress reste valable pour les Inverseurs.

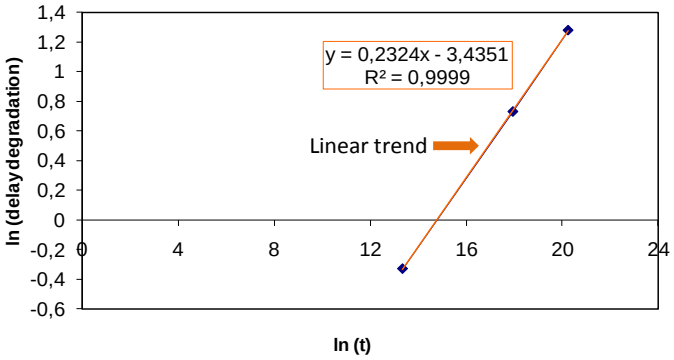


Figure 1-5 Loi de puissance de la dégradation du délai (provoquée par le NBTI) en fonction du temps de stress.

L'analyse du vieillissement a été élargie sur une famille de 10 inverseurs (ou IVX) fabriqués dans la technologie CMOS 45nm. Le tableau 1-1 résume les principales caractéristiques de ce groupe des inverseurs. Plusieurs simulations ELDO ont été effectuées en différents scénarios (y compris plusieurs températures, valeurs de tension, SP, C_L et S_I). Par la suite, la méthode des moindres carrés a été utilisée pour estimer la corrélation de linéarité entre les valeurs de la dégradation du délai et les valeurs de temps de stress.

Inverseurs	Numéro de pMOS	W pMOS (nm)
IVX2	1	0.22
IVX4	1	0.41
IVX7	2	0.61
IVX9	2	0.79
IVX22	3	0.68
IVX31	4	0.71
IVX40	5	0.73
IVX71	8	0.81
IVX106	12	0.81
IVX284	32	0.81

Tableau 1-1 Caractéristiques générales du groupe d'inverseurs analysés

La figure 1-6 montre l'histogramme de coefficients de linéarité calculés sur une base de données composée de 594 éléments (représentées dans la figure comme N 594). Notez que seul le vecteur d'entrée (Low-High) a été examiné. On observe que la Loi de Puissance de la dégradation par rapport au temps de stress est conservé malgré la complexité de certains inverseurs (comme l'IVX106 et l'IVX284, qui sont composés par divers dispositifs MOS en parallèle).

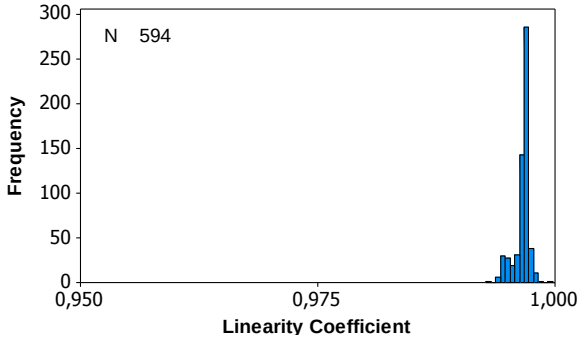


Figure 1-6 Histogramme de coefficient de linéarité obtenu à partir d'un échantillon de 594 vecteurs (*L-H)*

D'après ces résultats, nous proposons un modèle de vieillissement, qui est définie par 3 paramètres:

$$\frac{\delta t}{t} = S \cdot (POT)^n, \tag{1-1}$$

où *S* représente la sensibilité de la porte au vieillissement, cela dépend principalement de la pente du signal d'entrée S_I et de capacité de charge C_L. Le deuxième paramètre «Power-On Time ***POT***» est défini par l'équivalence en temps de stress du profil de la mission du produit (à la tension nominale d'opération). Le troisième paramètre est l'exposant *n*. Nous supposons une dépendance définie par la loi de puissance de la dégradation du délai par rapport au temps de stress.

Deux éléments constituent la base de cette méthode. Le premier élément est une dynamique de vieillissement bien établie, qui est calculée à partir de deux PVTs (un PVT âgé et un PVT frais) (§3.3.1). Puisqu'il est impossible de calculer les effets du vieillissement (avec suffisamment de précision) sur une porte en se basant sur un seul PVT (§2.1.2 et §2.3.2). Le deuxième élément est une méthode robuste de calcul de l'impact du NBTI sur le transistor à un PVT donné. Nous utilisons des modèles basés sur la physique (au niveau transistor) pour calculer la dégradation du délai (§2.4.2.2 et §2.4.2.3). Si les modèles de vieillissement au niveau transistor ne sont pas corrélées avec le silicium, on observe une grande réduction de la performance des IC au niveau système [Ruiz11].

Les paramètres *S* et *n* sont calculés à partir de la dynamique du vieillissement (Fig. 1-7). D'autre part, le paramètre *POT* est calculé à partir du profil de mission générique du produit sous analyse (tableau 1-2).

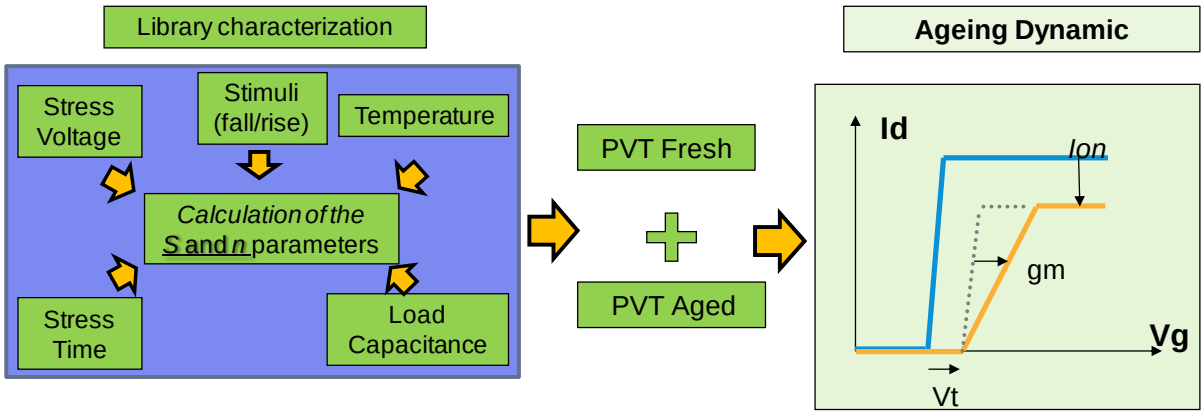


Figure 1-7 Schématique de la caractérisation de bibliothèque de portes (ou Standard Library) montrant la méthode de calcul des paramètres du modèle: sensibilité au vieillissement *(S)* et le exposant (*n*)

Le profil de mission générique est calculé par la méthode du taux de défaillance (au début de la vie utile du produit Early Life Failure Rate **ELFR**) définie dans [JEDEC74], qui est utilisé pour tester le taux d'échec précoce d'un produit (expériences sur silicium). Les modèles d'accélération par rapport aux paramètres de tension et de température sont définies respectivement par les équations 1-2 et 1-3. Ces modèles sont utilisés pour calculer le facteur d'accélération (*A*) de la dégradation.

$$Av = \exp\left[\left(\frac{K}{X}\right) \cdot (V_A - V_U)\right] = \exp[(\gamma_v \ \cdot (V_A - V_U)], \tag{1-2}$$

où A_v définit le facteur d'accélération en tension, K représente la constante de champ électrique (qui est déterminée expérimentalement), X définit l'épaisseur du diélectrique stressé, γ_V est défini comme le ratio (K/X) [V^{-1}], V_A est la tension de stress dans le test accéléré «Early Fails Failure **ELF**» et V_u représentent la tension d'alimentation dans les conditions normales d'utilisation.

$$A_T = \exp\left[\left(\frac{E_{aa}}{K}\right) \cdot \left({}^{1}/_{T_u} - {}^{1}/_{T_A}\right)\right], \tag{1-3}$$

où A_T définit le facteur d'accélération en température, E_{aa} représente l'énergie d'activation apparente (eV), K est la constante de Boltzmann (eV/°K), T_U définit la température de jonction à des conditions normales d'utilisation du produit (°K) et T_A représente la température de jonction à des conditions d'accélération (°K).

Précisément, le facteur d'accélération *A* est défini comme le produit du facteur d'accélération en tension et le facteur d'accélération en température,

$$A = A_T \cdot A_V \tag{1-4}$$

Enfin, le paramètre POT est défini comme la somme des résultats des multiplications du facteur *A* et de *l'équivalent de vie active* (Equivalent Life Active **ELA**) [secondes] pour un certain nombre de modes.

$$POT = \sum A \cdot ELA \qquad (1\text{-}5)$$

Notez que les facteurs d'accélération sont extraits des tests de fonctionnement en haute température et tension (High-Temperature Operating Lifetests HTOL).

	I	II	III	IV	V	VI	VII	VIII	IX
	Mode number	**Mode name**	**Lifetime (%)active**	**Equivalent Life active (seconds)**	**Voltage mode [Vdd]**	**Temperature in transistor junction [°C]**	**A_T**	**A_V**	**Normalized Stress time at a condition (V and T°) VII*VIII*IV**
	1	Cellular Voice	10	$18x10^6$	0.85	75	0.116	1.325	$2.77x10^5$
	2	Music playback	5	$9x10^6$	0.85	75	0.075	1.54	$1.49x10^6$
	3	Messaging (SMS, email)	15	$27x10^6$	1.00	86	0.00153	0.0009	$3.72x10^6$
	4	Internet navigation	5	$9x10^6$	1.15	100	0.1284	1.55	$1.79x10^6$
	5	Cellular standby	55	$99x10^6$	0.6	30	0.0002	0.00	0.000
Total	Generic Mission Profile		100	Equivalent stress time at given condition (voltage and temperature)					$35.56x10^6$

POT

Tableau 1-2 Profil de mission générique pour un smart phone
(Notez que cet exemple n'est pas un véritable profil de mission)

3 Modélisation NBTI au niveau Système

Ce chapitre décrit l'estimation de la fiabilité au niveau système en utilisant le modèle porte, qui a été présenté en la section antérieure. C'est dans ce contexte que nous présentons les analyses de fiabilité effectués sur deux IPs: un détecteur/correcteur d'erreur de type LDPC (*Low Density Parity Check)* et un Cortex–A9 (§4.2), afin de vérifier la précision des prédictions du «modèle porte» au niveau système. Le phénomène HC n'est pas analysé à ce niveau, ce point est expliqué en détail dans l'annexe A.

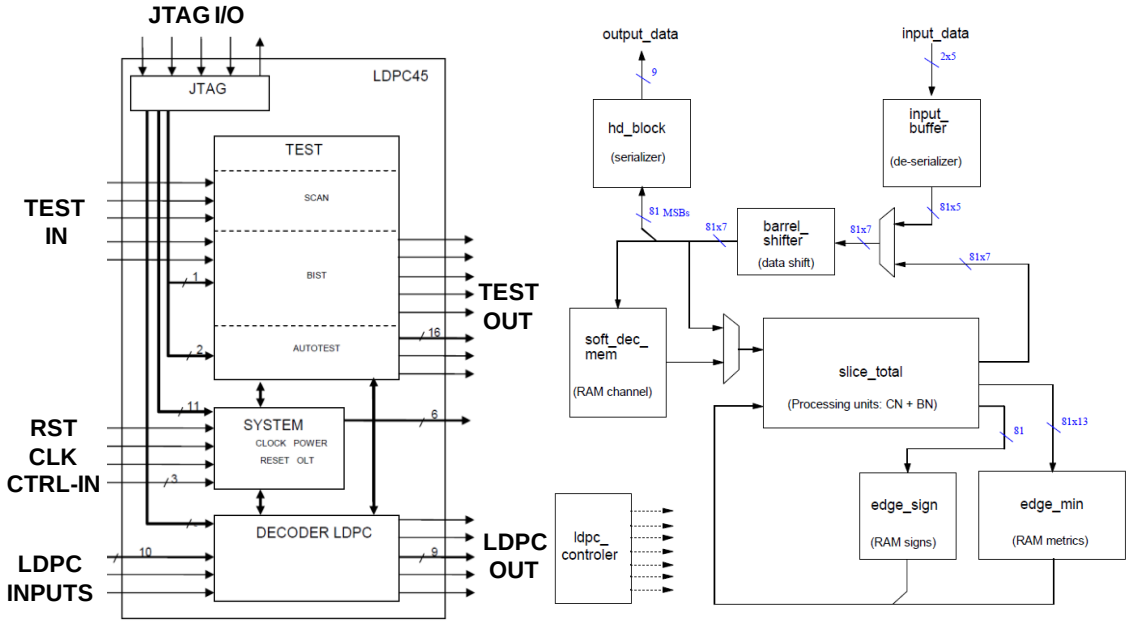
Afin de valider le «modèle NBTI au niveau porte» au niveau système, une analyse combinée de vieillissement Silicium-CAO a été effectuée. Nous avons utilisé les résultats des rapports de qualification (tests HTOL) des deux IPs (le LDPC et le processeur Cortex–A9) afin de comparer les valeurs de la F_{max} (définit comme la dégradation de la fréquence maximale d'opération des IPs (ΔF) provoquée par le NBTI) mesurent sur Silicium avec les estimations CAO obtenues à partir du modèle NBTI au niveau porte (§4.2). Chaque rapport de qualification des IPs contient les mesures HTOL (lesquels ont été réalisés à diverses tensions de stress, temps de stress et températures).

Afin d'obtenir la dynamique de dégradation du délai des chemins critiques un STA «frais» (définie comme une STA à un temps $t = 0$) et un STA «âgé» (à $t=t_1$) ont été exécutées sur les deux bases de données. Ces STA ont été calculés en utilisant le modèle au niveau transistor présenté dans le chapitre 2 (§2.4.2.3). Par la suite les différentes conditions de stress des tests HTOL effectués sur les deux IPs ont été converties en ses équivalents Power-On Time (POT) (§3.5.1). En suit le modèle au niveau porte a été nourri avec les valeurs POT et les valeurs des paramètres S et n (obtenues à partir des STA) afin de calculer l'impact de la dégradation NBTI sur les chemins critiques.

3.1 Analyse du NBTI sur un LDPC

Le circuit LDPC a été fabriqué en technologie CMOS 45nm. Ce circuit est couramment utilisé pour décoder les signaux qui voyagent dans des canaux de transmissions bruyants. La vue du haut niveau du circuit et l'architecture du décodeur du LDPC sont présentés dans la figure 1-8.

Cette IP a été stressée à 10 différentes conditions, dont plusieurs temps de stress (jusqu'à 1000hrs) pour un total de 1200 pièces testées.



(a) (b)

Figure 1-8 (a) Vue du haut niveau du circuit LDPC. (b) Architecture micro décodeur du LDPC

La figure 1-9, montre l'histogramme de la dégradation du délai à cinq valeurs des POTs (on montre seulement cinq de dix valeurs dans la figure pour des raisons de visibilité). Ces histogrammes ont été calculés en utilisant le modèle NBTI au niveau porte pour 60.000 chemins de données (en mettant l'accent sur les chemins avec violations de *temps de setup*). Nous incluons également les valeurs expérimentales de la dégradation de la F_{max} du LDPC (lignes verticales).

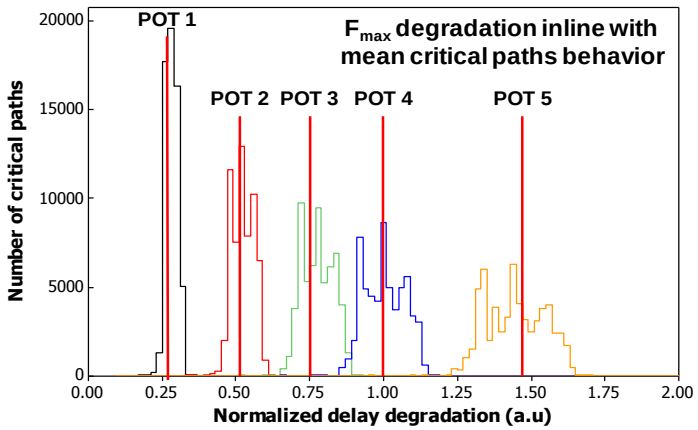


Figure 1-9 Histogrammes des dégradations du délai calculés à partir du modèle porte pour 60.000 chemins critiques du LDPC pour différentes conditions de stress (POTs).
POT1< POT2< POT3< POT4< POT5

Nous pouvons observer que la moyenne de la dégradation de la F_{max} est très bien alignée avec la moyenne CAO des dégradations du délai de la population des chemins critiques. Ainsi, nous considérons que cette approche offre un moyen précis pour prédire la dégradation de la $F_{max,}$ qui est une des principales caractéristiques macroscopiques analysées au niveau système et produit.

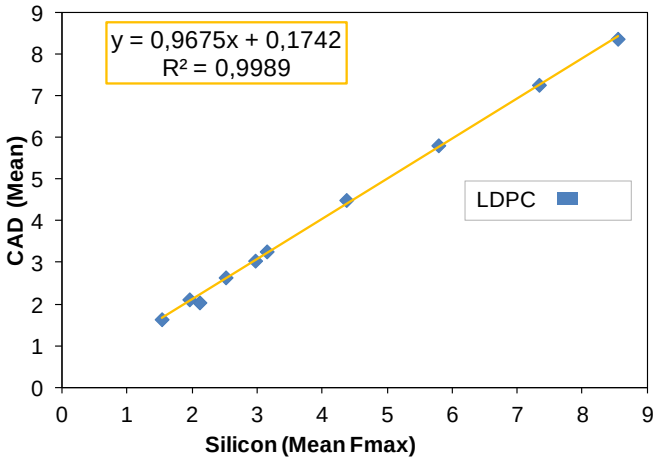


Figure 1-10 Corrélation entre les moyennes de la dégradation de la F_{max} et les moyennes CAO pour 10 conditions de stress (LDPC)

La figure 1-10 montre la corrélation entre les moyennes de la dégradation du délai (calculée à partir du modèle NBTI au niveau porte) en fonction des moyennes de la dégradation expérimentale F_{max} pour l'ensemble des conditions de stress appliquées au LDPC (10 POTs). Comme nous pouvons observer les prédictions du modèle au niveau porte ont une corrélation très proche de 1:1 avec les observations expérimentales. Le modèle NBTI au niveau porte reproduit les dégradations du délai avec une marge d'erreur de 2% de la dégradation totale. Donc si la moyenne de la dégradation du délai vaut 10%, l'erreur commise par le modèle est de 0,2 %.

3.2. Analyse de l'impact de la dégradation NBTI sur un processeur Cortex-A9

Le processeur ARM Cortex-A9 ™ est généralement utilisé en applications telles que: smartphones, consoles de jeux et télévision digitale. Ce processeur a été fabriqué en technologie 45 nm (STM). Il atteint une fréquence d'horloge de 800 MHz [Arm09] (Fig. 1-11). Cette IP a été stressée à 3 conditions différentes, dont plusieurs temps de stress (jusqu'à 1000hrs), qui représentent un vrai profil mission pour un total de 160 pièces testées.

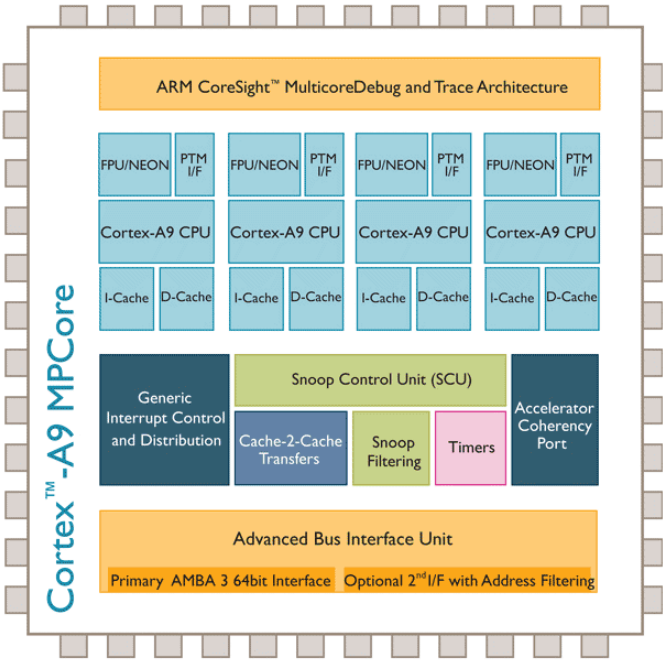


Figure 1-11 Processeur l'ARM Cortex-A9 ™

La figure 1-12 montre la corrélation entre les moyennes des dégradations de la F_{max} et les moyennes CAO pour les populations de chemins critiques de l'A9 et du LDPC. Cela montre que notre approche de modélisation peut être utilisé avec succès pour analyser les bases de données complexes.

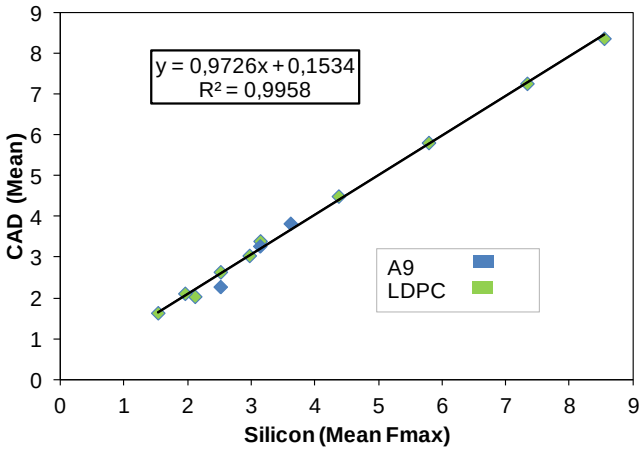


Figure 1-12 Corrélation entre les moyennes de la dégradation de la F_{max} et les moyennes calculées à partir du modèle NBTI au niveau porte (CAO) (LDPC et A9).

4 Conclusion

Cette thèse présente une étude de fiabilité sur les technologies CMOS nanométriques utilisées dans les applications numériques. Deux thèmes majeurs sont considérés dans ce manuscrit, les phénomènes de vieillissement NBTI et HCI au niveau porte.

Dans la première partie de la thèse une analyse détaillée est effectuée sur une base importante de données des portes dans le but de comprendre les influences des paramètres de conception dans la dégradation du délai au niveau porte. De cette étude, une approche analytique définit par trois paramètres a été proposée afin de modéliser les effets de la dégradation NBTI et HCI au niveau porte (section 3.7).

Dans la deuxième partie du manuscrit nous avons discuté de l'impact du mécanisme NBTI au niveau système. Le modèle NBTI au niveau porte a été utilisé pour calculer les effets de la dégradation NBTI sur deux IPs: un LDPC et un processeur Cortex-A9. Plusieurs analyses analytiques de fiabilité (Static Timing Analysis) ont été effectuées en utilisant le modèle au niveau porte et par la suite les résultats obtenus ont été comparés avec les données expérimentales (mesures silicium). Le modèle NBTI au niveau porte a été validé avec le silicium. Nous avons montré que le modèle peut reproduire d’une manière précise le comportement de la dégradation NBTI au niveau système.