Reliability and Physics-of-Healthy in Mechatronics: Part I: Entropy and Physics-of-Healthy: Some Concepts to Model Predictive Reliability of Microelectronics for Automotive, Aeronautic and Space Missions, pp.3-138 Part II: Failure and Analysis of Systems Engineering, pp. 139-140

Alain Bensoussan, Joseph Bernstein, Alain Bravaix

Physics-of-Healthy: Predictive Reliability Models and Case Studies of Microelectronics for Aeronautic, Space, Automotive and Transport missions.

Version issue 5

Date: 18 October 2020

Bensoussan Alain
Dr.-Ing, Dr-HDR, IEEE senior member
Thales Alenia Space
Meta Competence Center
Toulouse, France
alain.bensoussan@thalesaleniaspace.com

Bernstein Joseph
Professor, IEEE senior member
Laboratory for Failure Analysis and Reliability of Electronic Systems
Ariel University
Ariel, Israel
josephbe@ariel.ac.il

Bravaix Alain
PhD, HDR, IEEE senior member
REER : Radiation Effects & Electrical Reliability, IM2NP, ISEN
Yncrea Mediterranée, Toulon, France
alain.bravaix@isen.fr

SUMMARY

Electronics components and devices, including equipment to systems, are fabricated from materials and structures that degrade with time under normal operational condition. It is necessary to anticipate and quantify system failure occurrence but the goal of "reliability engineering" is to clarify first the failure paradigm in term of statistics and Physics of Failure or simply Part Count approach. Many statistical approaches and underlying mathematics have been developed in the past decades of the last century to describe failure rates and the well-known bathtub curve showing a schematic of failure rate behavior with time.

From Wikipedia definition, *predictive (or logical) analysis encompasses a variety of techniques derived from statistics, data extraction and game theory that analyze past and present facts to make predictive assumptions about future events.*

In this applied reliability chapter, we would like to share our perception on the experimentally development applied on existing reliability and maintenance paradigms developed since decades by many authors and papers.

This Part I titled "Reliability Basic Tools for SHM protocols" aims to remind a) how applied engineering in predicting failure and monitoring SHM of electronic equipment and systems are implemented and b) to present basic statistic tools defined for reliability modelling implementation and studies related to active microelectronic parts. The second Part II is dedicated to the experimental application titled "Applied Engineering on Physics-of-Healthy and SHM on microelectronic equipment for aeronautic, space, automotive and transport operation" addressing how innovative technologies and Commercial-Off-The-Shelf (COTS) devices are pondered.

How to implement predicting failure and monitoring SHM of electronic equipment and systems?

High Temperature Operational Lifetest (HTOL) Standard is relatively poor to predict Mean-Time-To-Failure (MTTF) associated constant failure rate and reveals mostly ineffective for wearout modelling as no failure are observed. Furthermore, the MTTF50% is not a zero-failure guarantee (rather by definition is a 50% lot failure) and we should define rather a MTTF0.1% for high reliability application. At the same time, attention to long-life cycles on the part of manufacturers has nearly vanished due to the relatively overwhelming demand in the consumer market sector.

Three key questions are reviewed including a) How to apply reliability prediction tools to innovative technologies? b) What are the hypotheses made of and methodology implemented in these tools; and to what extent are they able to cover emerging technologies and applications? c) Prognostic Failure Model (PFM) what could be a proposed guideline to organize all concepts? We describe our approach in what is needed to answer these and other fundamental questions as for example:

The goal of reliability engineering is to anticipate and quantify failure occurrence but the science required is to clarify first the paradigm in term of statistics and Physics of Failure [1], [2] or Part Count failure approach [3] [4], [5]. Many questions are raised, and we try to address as listed here:

In term of Physics and Phenomenological Degradation:

What are the existing Standards, their advantages and drawbacks?

How to manage random and wearout failure rate model to implement predictive reliability and maintenance?

What role the "stressors" are playing?

How "indicators" are characterized and which criteria value to define?

How to deal with multiple stresses?

In very complex new technologies (less than 7 nm FinFET, GaN, Carbone NanoTube, …), what are the new failure mechanisms?

What if multiple failure mechanisms are coexisting? [5]

What is the impact on the accelerating factor (AF) in such complex mixed model (multiple stress, multiple failure mechanism)? [6]

How to apply reliability prediction tools to innovative technologies?

What are the hypotheses made of and methodology implemented in these tools; and to what extent are they able to cover emerging technologies and applications?

How to consider the effect of initial device quality, the influence of use and design options, the effect of mission profiles to model the Robustness [7] and the Reliability of complex systems and new technologies [8], [9], [10]?

In term of Application and experiments:

In very complex new technologies, observed failure mechanisms can often be accelerated simultaneously, causing a dilemma for reliability prediction.

In the domain of telecommunications, automotive, aerospace, satellite, military and the like, the need for accurate reliability prediction is as important as it ever has been. At the same time, attention to long-life cycles on the part of manufacturers has nearly vanished due to the relatively overwhelming demand in the consumer market sector.

How to explore field return and define optimized maintenance survey based on reliability prediction tools?

Assuming a qualification test experiment on 130 samples tested during 2000 hrs of 150°C HTOL with 0 failure observed.

What would be the device operating projected lifetime verified if the operational application is at T=45°C?

What should be the sample size and duration to demonstrate a Cumulative Density Function (CDF) equal of less than 0.1%

after 30 years in mission operation?

For Reliability prediction purpose do we need to consider Time-To-Failure for each failure mechanism either for a constant failure rate based on a Poisson statistic, or for wearout failure mechanism?

What would be the equivalent total λ_{total} and corresponding Time-To-Failure if all failure mechanisms are equally activated?

What the basic mathematics are?

Thermodynamics modeling established by S. Arrhenius (1896) [11], L. Boltzmann (1886) [12], then M. Evans and M. Polanyi in 1938 [13], E. Winner (1938) [14], S. Glasstone (1941) [15], G. Hammond (1955) [16], R. Drenick (1960) [17], E. Snow (1965) [18], N. Sedyakin (1966) [19], D. Cox (1972) [20] and periodically improved by H. Eyring et al. (1980) [21], J. McPherson (1999) [22] [23], or recently by E. Suhir (2013) [24] [25]. Modelling was supported by series of experimental papers from MIT (D. Jin and Del Alamo) [26], RAC Univ. of Maryland, University of Padova and G. Meneghesso's team [27], INTEL and IBM teams [28]. Since the break of 217^{+} standard, FIDES French group guideline is still in place and its next evolution is thought to consider non-constant failure rate model for predictive reliability. Not leaving aside main actors in this paradigm development we have to name most representative experts in the field which have influenced the written of these chapter J. Stathis, V. Huard, A. Bravaix, J. Bernstein, M. Meneghini, H. Zanoni, E. Wu and J. Sune, P.A. Tobias and D.C. Trindade, M. Nikulin and V. Couallier.

The paradigm of the Transition State Theory (TST) developed by E. Wigner in 1934 [14] and by M. Evans, M. Polanyi in 1938 [13] is an approach we can benefit too by adapting the concept of a unified semiconductor reliability model and multiple failure mechanisms related to Physics of degradation. Indeed in the early last century, the TST was applied to chemistry transformations by H. Eyring [21], and S. Glasstone et al. [15] in 1941. The TST was developed in chemistry based on the Hammond's postulate [16] published in 1955 applied to physical organic chemistry.

Multiple failure mechanisms and Physics of degradation in semiconductors may occur in a single set of time-to-failure data but without obvious points of inflection to help separate the mechanisms. J. McPherson in his book 3rd edition of Reliability Physics and Engineering provides the basics of reliability modelling [23], [22] recalled *generally, materials/devices exist in metastable states. These states are referred to as being metastable because they are only apparently stable. Metastable states will change/degrade with time. The rate of degradation of the materials (and eventual time-to-failure for the device) can be accelerated by an elevated stress (e.g., mechanical stress, electrical stress, electrochemical stress, etc.) and/or elevated temperature*.

Gibbs free energy diagram recalled in 3rd edition of J. McPherson book has provided the main inputs to describe multiple stressors environment effect including entangled accelerating factor picture as fully detailed in Part I. Numerical application for a study case on a FinFET technology assuming three major failure mechanisms defined by the following Arrhenius reliability parameters are also detailed.

These mathematics and physic approaches show how the activation energy is an Eyring model related to the stress and temperature applied and can no-longer be considered as a constant to extrapolate some experiment under high stress to nominal mission operation profile. That's the reason why several end-user Industries and Institutions are very cautious to perform lifetest conditions as close as the nominal conditions because of the change of activation energy attributed sometime wrongly to new (or different) failure mechanism while it is simply explained by the interaction of stress and temperature effect on the measured activation energy or to a best extend to Eyring law.

From these equations we can observe the **equivalent activation energy is dependent of the temperature T_0 and is increasing with temperature**. So, under the stress conditions, **the failure mechanism model is a non-uniform acceleration mechanism**.

The goal of Part I is to present basic statistic tools defined for reliability modelling implementation and studies related to active microelectronic parts (Integrated circuits, power transistors, etc) when exploited in operational environment for long term high reliability application.

Basics mathematics on serie-parallel systems reliability are presented with some approximation considerations for distribution queues statistics.

Conclusion and perspective open the door for the next Part II related to Predictive Reliability supported by experimental and Physics of Failure (PoF) or what we call Physics of Healthy (PoH).

Physics of degradation in semiconductors may occur in a single set of time-to-failure data but without obvious points of inflection to help separating the mechanisms. In his book, J. McPherson titled "Reliability Physics and Engineering provides the basics of reliability modelling" [23], [22] stated "*generally, materials/devices exist in metastable states. These states are referred to as being metastable because they are only apparently stable. Metastable states will change/degrade with time. The rate of degradation of the materials (and eventual time-to-failure for the device) can be accelerated by an elevated stress (e.g., mechanical stress, electrical stress, electrochemical stress, etc.) and/or elevated temperature"*.

Part II goal is to concentrate on experiment supported by models and field return showing the "true life" observing simultaneous stress environment (various mission profiles) and multiple failure mechanisms.

Keywords— Probabilistic Design-for-Reliability, SHM, Reliability, PoF, PoH, Transition State Theory, Statistics, Semiconductors, Wide Band Gap, Deep Sub-micron, FinFET.

NOTATION

BAZ	Boltzmann-Arrhenius-Zhurkov
BEOL	Back End Off Line
BTI	Bias Instability (NBTI or PBTI)
CA	Constructional Analysis
CCC (or CC)	Channel Cold Carrier
CDF CFET	Cumulative Distribution Complementary Transistors Function
CHC or HC)	Channel Hot Carrier
DIBL	Drain-Induced Barrier Lowering
DoE	Design of Experiment
DSM	Deep-Submicron technology
$c(x_i)$	Energy factor related to stress parameter x_i
E_a	Activation energy
$E_{a\_effective}$ or E_{aa}	effective activation energy related to multiple stress reliability test sequence
EES	Electron-electron scattering
ECSS	European Cooperation on Space Standardization
EMC	Electromagnetic compatibility
EM	Electromigration
ESD	ElectroStatic Discharge
EOS	Electrical Overstress
EOT	Equivalent Oxid Thickness
EVD	Extreme Value Distribution
$\in$ or $\in(t)$	Gibbs Free Energy for a device
FEOL	Front End Off Line
FF	FinFET
FI	Fan In
FO	Fan Out
FRAME	Failure Risk Analysis Methodology
GaAs	Gallium Arsenide
GaN	Gallium Nitride
G	Gibbs Free energy
GIDL	Gate-Induced Drain Leakage
γ_i	Stress factor as percentage of burnout limits for label i as current *I*, voltage *V*, P_{DC} power, P_{in} signal input power
HBD	Hard breakdown
HCD	Hot carrier degradation
HCI	Hot carrier injection mechanism
I_{BO}	Current burnout limit value
I_{G_leak}	Gate leakage current between gate and source at V_{GS} and V_{DS}
I_{G_LBO}	Gate leakage current between gate and source at high V_{GS} and V_{DS} close to burnout or breakdown.
I_g	peak gate current during stressing
I_{HCG}	Current due to hot carrier generation
ILD	Inter-layer dielectric
I_{sub}	peak substrate current during stressing
$\lambda(t)$ or *IFR*	Instantaneous Failure Rate or Hazard function
k	Boltzmann's constant ($1.3807 \cdot 10^{-23}$ J/°K or $8.6174\ 10^{-5}$ eV/°K).
MESFET	Metal-semiconductor field effect transistor
MR	Maximum rating
M-STORM	*M*ulti-phy*S*ics mul*T*i-stress*O*rs predictive *R*eliability *M*odel
M-TOL	Multiple Temperature Overstress life-test
MVE	Multi-Vibrational Excitation
NAF	NMOS Acceleration time Factor
NBTI	Negative bias temperature instability
NTF	DC NMOS Acceleration Time Factor
PAF	PMOS Acceleration Factor
PBTI	Positive bias temperature instability
P_{BO}	Power burnout limit value
PDF	Probability Density Function
PDfR	Probabilistic Design-for-Reliability
PHEMT	Pseudomorphic High Electron Mobility Transistor
PMOS	P type Metal Oxide Semiconductor
PoF	Physics of Failure
PoS	Physics of Stress
P_q	Poisson statistic function
q	Number of failures in Poison statistic function
QBD	Charge at breakdown
SBD	Soft breakdown
S_e	electrical indicator or signature of the failure mechanism
S_i	Stress parameter for i as current *I*, voltage *V* or P_{DC} power consumption or P_{in} signal input power
S_{iBO}	Stress parameter at burnout limit value
SiC	Silicon Carbide
SM/SV	Stress migration/voiding
SOA	Safe Operating Area
SVE	Single Vibrational Excitation
TDDB	Time Dependent Dielectric Breakdown
TST	Transition State Theory
TSRM	Transition State Reliability Model

CONTENT

Part I: Basic Reliability Tools for SHM protocols

Bensoussan Alain
Dr. Ing., Dr.- HDR, IEEE senior member
Thales Alenia Space
Meta Competence Center
Toulouse, France
alain.bensoussan@thalesaleniaspace.com

Bernstein Joseph
Professor, IEEE senior member
Laboratory for Failure Analysis and Reliability of Electronic Systems
Ariel University
Ariel, Israel
josephbe@ariel.ac.il

Bravaix Alain
PhD, HDR, IEEE senior member
REER : Radiation Effects & Electrical Reliability, IM2NP, ISEN
Yncrea Mediterranée, Toulon, France
alain.bravaix@isen.fr

"As long as mathematical laws refer to reality, they are not absolute, and as long as they are absolute, they do not refer to reality"

"Tant que les lois mathématiques renvoient à la réalité, elles ne sont pas absolues, et tant qu'elles sont absolues, elles ne renvoient pas à la réalité."

Discours à l'Académie Scientifique de Prusse, Janvier 1921. Albert Einstein

I. Introduction

Trending articles and news related to the semiconductor industry are published in various web sites. In a paper signed by Khaveen Jeyaratnam, a specialist in financial analysis, valuation and investment research presented the Semiconductor Industry Value Chain and discussed the Top 3 Semiconductor Foundries In The World in March 21, 2019 (https://seekingalpha.com/article/4250209-top-3-semiconductor-foundries-world). He highlights "*the semiconductor industry is experiencing heavy growth in demand from the onset of the fourth wave of technology which includes robotics, artificial intelligence, nanotechnology, quantum computing, biotechnology, the Internet of Things, fifth-generation wireless technologies (5G), augmented/virtual reality, 3D printing, and autonomous vehicles*."

The top 4 worldwide semiconductor foundries by revenue for 1H18 are TSMC (56.1% market share), GlobalFoundries (9.0%), UMC (8.9%), and Samsung (7.4%). Currently TSMC ($32B+) is the number one foundry by a very large margin with GF ($6B+), UMC (>$5B+), Samsung (>$5B+), and SMIC $3B+) barely visible in the rearview mirror. In Semiwiki web site it was mentioned the full year 2018 revenues of TSMC, by technologies represent 9% for the 7-nanometer node size (25% in 2019), 11% for the 10-nanometer, 23% for the 16/20-nanometer, and 63% for the 28-nanometer and below. On the other side, GlobalFoundries offer CMOS, FinFET and FD-SOI technologies.

Electronics components are made from devices, including equipment to complete systems, are fabricated from materials and structures that degrade with time under normal operational condition. This is a fundamental consequence of the second law of thermodynamics which cannot be avoided in our physical world. We cannot afford uncertainty of systems based on various domains including Aeronautics, More Electrical Aircraft, Space, Automotive and autonomous vehicles relying on hundreds, thousands and millions of electronic devices per vehicle and employing fleets of thousands. Degradation has been observed also, to depend on many environment parameters, both intrinsic and extrinsic, as follow due to:

- Packaging and component manufacturing and process distributions
- Original material and geometry arrangement into devices
- Internal and external environment conditions named "***stressors***" so as biasing (voltage, current, power dissipation, transient, signal level), temperature (low, high, constant, variable), irradiation levels (particles, X-ray, γ-ray), humidity, other pollution gases, etc [23], [5],

We can model failure rates and degradation, assuming independently thermal activation processes and at random, until some critical device parameter can no longer meet the required specification for proper device or system functionality. Some parameters (considered as electrical signatures) can be identified as **precursor** of failure occurrence. So, the failure is really defined when a specific functional parameter is "enough" affected to create a malfunction of the system. The threshold, for an indicator degradation leading to a final system parametric failure or a catastrophic failure depends on the stressor intensities (level of stresses with respect to the intrinsic limits supported by a device) and time.

Many statistics and mathematics have been developed over the centuries [11], [12], [13], [21], [23], [29] to describe failure rates and the well-known bathtub curve showing a schematic of failure rate behavior with time. In this approach three domains were traditionally established: Infant Mortality, Random failure and Wearout Failure [2], [30], [31]. The common "Time-to-Failure" parameter (TTF) e.g. the time for a given population to survive, to have its number of "good" elements be ½ the original population [23].

This simplified viewpoint which has drawbacks and suffers from some vagueness in the definition of failure when considering innovative commercial-off-the-shelf (COTS) devices use (Deep Sub-Micron technology node size lower than 20 nm [32] or Wide Band Gap semiconductor based devices [33]):

- Infant mortality, random failures and wearout are generally based on catastrophic failures observations in the field. When considering parametric failures this may induce different interpretation and results [6] [34], [35].
- Temperature definition is a key parameter which can modify the interpretation and the reading of results [3], [36]: for example temperature can be considered as ambient or case or junction.
- Failure criteria definition is observed to be a sensitive parameter impacting the TTF [37], [38], [39].

Saying this, broad application of disruptive technologies in embedded Hi-Rel systems requires all reliability aspects to be modeled and quantified in order to predict TTF and wearout Remaining Useful Life (RUL) both impacted by mission profile of electronic systems.

This chapter will first detail the existing models, tools and methodologies existing for COTS new emerging technologies. We will see why and how the concept of activation energy is an Eyring model related to the stress and temperature applied and the use of an approximated constant activation energy can no-longer be considered as a constant to extrapolate some experiment under high stress to nominal mission operation profile.

We will present hypotheses, baselines and definitions of reliability parameters through the prism of Transition State Theory (TST) in order to show how the multi-dimensional stress environment can be simply modelled. Reliability prediction is then supported by the extrapolation of lifetime models based on what we call, Physics of Healthy (PoH). PoH in term of reliability prediction is the calculation of the statistical failure probability described by physics-based rate processes. The methodology presented here is a direct application of the well-known, standard, proportional sum-of-failure-rates where the rates are

calculated by thermodynamically determined statistical processes.

A section will address the Transition State Reliability Model and reliability Physics of Healthy defined as the probability of a product performing it intended function for a given amount of time and mission profile. Basic approaches are recalled so that engineering designs and materials are detailed to construct reliability tools as defined by conventional reliability standards.

A last section will focus on system reliability (series and parallel) mathematic recalls and proposed approximations through four Lemmas.

L1: A series system constituted of *n* identical and independent elements, each described by a POISSON distribution (λ_0) reliability model can be approximated by a general equivalent POISSON distribution with parameter $n \cdot \lambda_0$.

L2: A parallel system constituted of *n* identical and independent elements, each described by a POISSON distribution (λ) reliability model can be approximated by a general equivalent WEIBULL distribution with parameters λ and $\beta = n$ with an error lower than 1% for time operation lower than 30%·MTTF for the example shown.

L3: A parallel system constituted of *n* identical and independent elements, each described by a reliability WEIBULL distribution (α, β) model can be approximated by a general equivalent WEIBULL distribution with parameter (α, $\beta \cdot n$) with an error lower than 1% for time operation lower than 75%·MTTF.

L4: In a series-parallel system constituted of *n* identical and independent elements, each described by a reliability POISSON distribution (λ), its reliability model can be approximated by a general equivalent WEIBULL distribution with parameter (θ, $\beta = m$) with an error lower than 3% for time operation lower than 50%·MTTF.

The fact that microelectronic devices are manufactured in such large quantities with established variability, deviations (fluctuations) and uncertainties make them the quintessential framework through which all reliability prediction and evaluation should be understood. Existing reliability models may sometimes be less accurate for emerging and non-mature technologies because of lack of customer and user's feedbacks due to the short product life cycle. The consumer market relies on short life cycles. Consequently, the very nature of the electronics industry, where so many parts having so many interacting physical phenomena, electrical performances, geometrical and material parameters, and application design conditions makes them the ideal example for reliability evaluation that can be applied to any other product or industry.

This chapter Part I will prepare the foundation to answer three central questions:

1. How to apply reliability prediction tools for innovative technologies?
2. What are the hypotheses made of and methodology implemented in these tools; and to what extent are they able to cover emerging technologies and applications?
3. Prognostic Failure Model (PFM): what could be a proposed guideline to organize all concepts described?

A. *COTS and Emerging technologies in Deep-Sub-Micron technologies: short overview*

In a recent synthetic paper published on SemiWiki on LithoVision 2019 – Semiconductor Technology Trends and their impact on Lithography, Scotten Jones [40] presented highlights on NAND scaling with layers, DRAM peripheral scaling and Logic high performance and IoT. Evidence is shown on the complexity of market push technology in the 3D NAND mask counts and bit density trends. "The transition from 2D NAND to 3D is enabling the continuation in bit density scaling by using the third dimension (see **Figure II-1**). DRAM scaling is limited by capacitor size and facing physical limits. Logic is continuing to scale but fundamental limits on 2D shrinks are looming. Leading edge logic has evolved from planar transistors to a split roadmap with FinFET for high performance and things like FDSOI for IOT. Longer term gate-all-around is on the horizon."

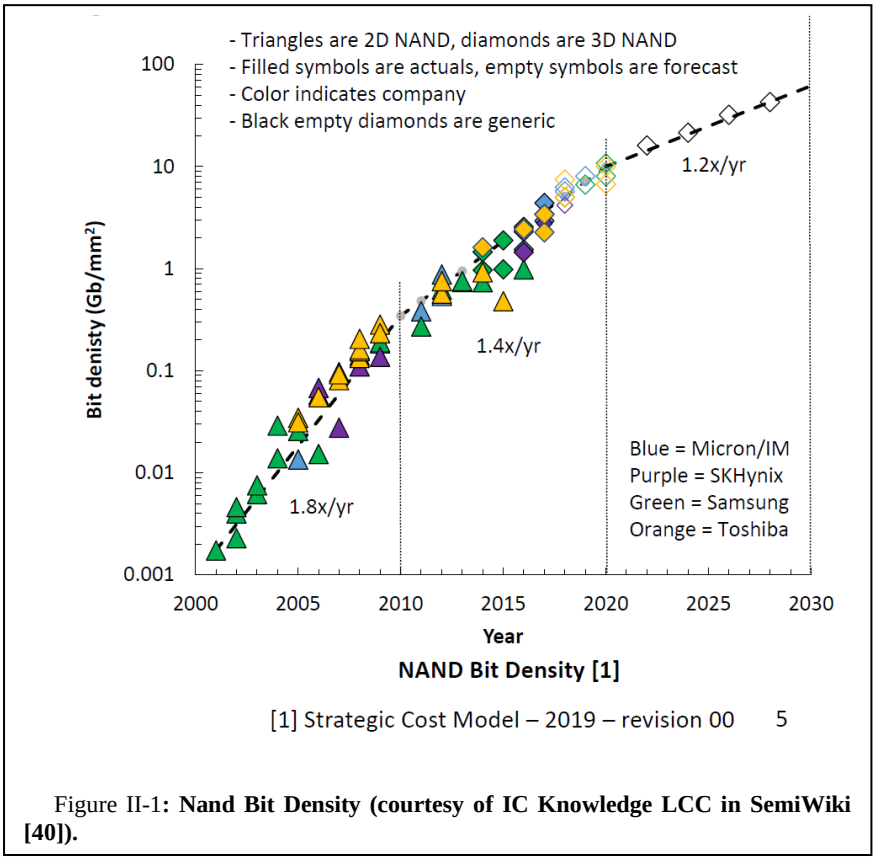
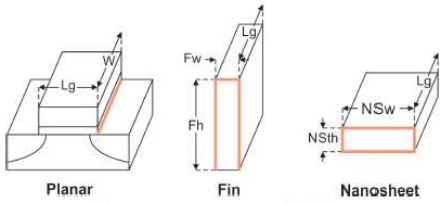


Figure II-1**: Nand Bit Density (courtesy of IC Knowledge LCC in SemiWiki [40]).**

From **Figure II-2** extracted from ref [40] we can see that for planar transistors the effective gate width (Weff) was defined by the transistor gate width. For FINFETs the Weff depends on the width and 2x the height of the fin and therefore Weff can only be varied in discrete increments. With GAA the Weff is 2x thickness and 2X width.

Logic Gate All Around (GAA)

- At 3nm recent imec work suggests that FinFETs are viable but every scaling booster option is required and nanosheets offer more margin.
- Nanowires provide the best electrostatics, FinFETs provide the best drive current, nanosheet width can tune the trade-off.



Effective channel width (Weff).

Weff

- Planar (1 side gate) Weff = W
- FinFET Weff (3 side gate) = 2Fh + Fw
- Nanosheet (4 side gate) Weff = 2NSth + 2NSw

Device	nanowire	FinFET	nanosheet	nanosheet
Dimensions	5nm/5nm	7nm/40nm	18nm/5nm	5nm/50nm
Relative Weff	0.69	1.00	1.06	1.26
Electrostatics	Best	Good	Better	Better

Nanowire, FinFET and nanosheet relative performance

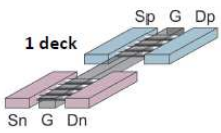
IC KNOWLEDGE LLC

11

Figure II-2 : Logic Gate All Around (GAA) size definition according to ref **[40]**).

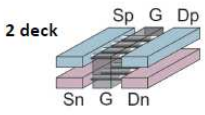
Leading Edge 2D to 3D Logic Roadmap

	N28	N20	N14	N10	N7	N5 [1]	N3.5 [1]	N2.5 [1]	N1.75	N1.25
Designation [2]	28.1	20.1	14.1	10.1	7.1	5.1	3.1	2.1		
Device	Planar	Planar	FF	FF	FF	FF	HNS/FF	HNS		
CPP	128	90	78	66	57	57/50	45 [4]/48	37 [4]		
M2P	90	64	64	44	40	36/28	32/28	26 [5]		
Tracks	7.00	9.00	9.00	8.25	6.00	6.00/6.00	5.00/5.00	5.00		
SDB/DDB	SDB	SDB	SDB	DDB	DDB	SDB/SDB	SDB/SDB	SDB		
Density (MTx/mm2)	19.72	30.05	34.68	55.10	96.49	126.53/ 185.46	216.37/ 231.83	323.89		
Density improvement [6]		1.52	1.15	1.59	1.75	1.31/1.92	1.71/1.25	1.50/1.40		



Designation [2]	14.7		3.2	3.3	3.4
Device	CFET		CFET	CFET	CFET
CPP	78		45 [4]	37 [4]	37 [4]
M2P	64		28	26 [5]	26 [5]
Tracks	6.00		5.00	5.00	5.00
SDB/DDB	SDB		SDB	SDB	SDB
Density (MTx/mm2)	129.63		386.38	454.57	534.79
Density improvement	1.34		1.79/1.67	1.18	1.18

[1] Values in red are projected values for Samsung/TSMC.
[2] Values are "node" linewidths followed by number of decks.
[3] Both companies claim 54nm but 57nm seen in actual standard cells.
[4] CPP limit as derived based on Lg, t_{spacer} and $W_{contact}$ porjected limits.
[5] 1D limit for EUV.
[6] Intel transistor density metric.



IC KNOWLEDGE LLC

12

Figure II-3: A roadmap from 2D planar transistors at the 28nm and 20nm nodes to FinFETs and then HNS and eventually stacked 3D CFETs (ref **[40]**).

From **Figure II-3**, for Node 5 (N5) and N3.5 we have specific projections for Samsung and TSMC. At N2.5 we have a generic forecast with both companies converged on HNS.

For 3D we have options beginning with a relaxed 14nm design rule CFET with 7 layers as well as more lithographically aggressive 3.2, 3.3 and 3.4 Complementary Transistors (CFETs) with 3nm lithography and 2, 3 and 4 layers. The pictures show an nFET and pFET side by side for a single layer device and then a pFET over and nFET for a 2 layer CFET.

According to Scotten Jones it is shown how the EUV helps to mitigate mask count increases and about 60 to 70 mask sets. CFET at a 1.75nm node also helps to control the lithographic difficulty by being highly self-aligned.

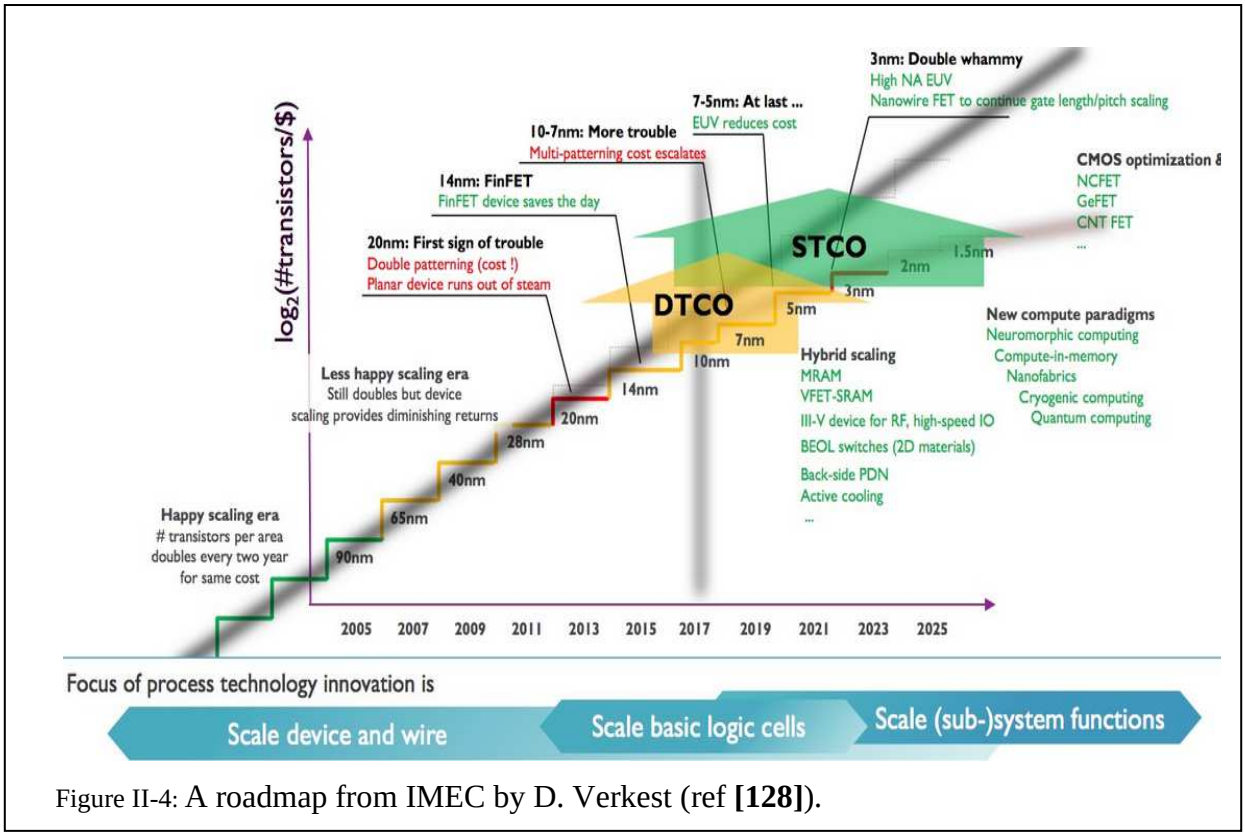


Figure II-4: A roadmap from IMEC by D. Verkest (ref **[128]**).

Traditional device scaling continues but dimensional scaling impacts performance too as shown in Figure III.4 an extract from Cadence blog Paul McLellan published in June 2018. “Fundamental new devices such as carbon nanotubes or 3D materials are still far from maturity. So there have been three eras. In the first, the focus was to scale the devices and wires. In the second, the focus was reducing the size of basic logic cells (and the SRAM cell). In the future, the focus has to move to scaling entire sub-system functions.”

In another internet link title “Big Trouble at 3nm” published by Mark Lapedus, Executive Editor for manufacturing at Semiconductor Engineering, in June 21st, 2018 [41], reports Dan Mocuta (director of logic integration and devices at Imec) words : “the key element is variable widths. You can control it better than the variable height of a fin,”.

“In a finFET technology, the width of the device is quantized. You can have one fin, two fins, three fins or whatever. In nanosheets (see Figure II-4 and Figure II-5), you have a fixed number of nanosheets on top of each other. But you can play with the width. Now, you have access to a continuum of device widths, which you didn’t have for the finFET,” Mocuta said. “For example, you want to have an area that drives a lot of current. That could be a buffer. Then, you want to have an SRAM with a very small footprint. There are different needs on the chip that can be met.”

Nanosheets are promising, but that isn’t the only option. With a breakthrough, finFETs could extend beyond 5nm. Another option is to wait until the industry develops a better transistor. Still another way is to get the benefits of scaling by putting multiple devices in an advanced package.

For now, gate-all-around technology appears to be the most practical technology after finFETs.

Modeling such novel transistors to establishing reliability figures for risk assessment is tremendously challenging. Market will be driven by system application and Deep Sub Micron technology as a “COTS integration plug and play” will require to implement risk assessment methodology to satisfy market related to Automotive, Aeronautic, and Space where system failures in normal operation and harsh mission profile condition are strictly forbidden for some period of time. Not clearly yet addressing the next question related to MTTF or MTBF for system repairs.

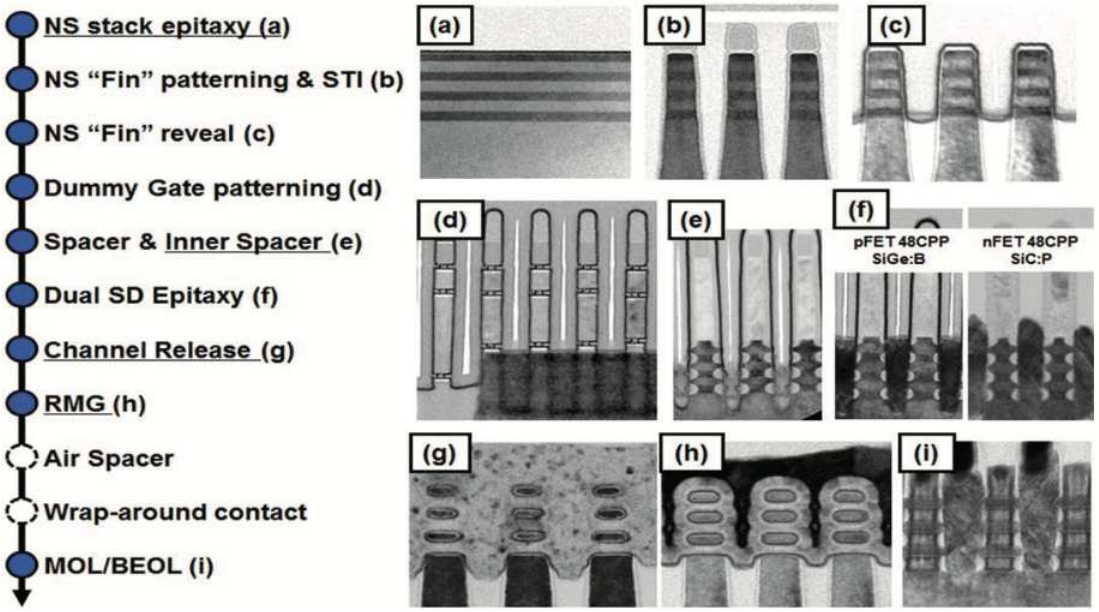


Figure II-5: Stacked nanosheet process sequence and TEM. IBM, Samsung Electronics, Global Foundries, (ref **[129]**).

Having established the hardware is so challenging, we know any system and component in custom or niche application will tend to degrade with time when operating under nominal or even sometimes extreme conditions. Yet the best reliable technology will continue to be affected by change in their intrinsic characteristic and performances or due to extrinsic high stress applied.

B. General overview in GaN device failure mechanisms

Normally-Off high power switching AlGaN/GaN-on-Si heterojunction transistors are promising for high reliability space application. Nevertheless, such and similar GaN technologies suffers well-known failure mode issues as studied since the last decade by several authors on gate current leakage increase or/and current collapse [42], [43], dynamic on-resistance with recovery effects [26], strain relaxations and trap charging effects [44], [45], TDDB mainly related to GaN MIS-HEMTS structures [46], PBTI of GaN MOSFETs [47]. From book chapter on Power GaN Devices M. Meneghini, G. Meneghesso and E. Zanoni [48] (Springer Ed. 2017) in chap. 13, J. Wuerfl (FBH, Berlin) describes the most relevant drift effects that limit the performance of GaN-based power transistors. Degradation and drift effects on semiconductor devices can affect their performances either reversibly or definitively. Many failure mechanisms have been identified and described showing they are imperative for predicting device performance in real system environment.

The effect of temperature on electronic devices is often assessed by extrapolating from accelerated tests at extremely high temperatures based on the Arrhenius law. This method is known to be not necessarily accurate for prediction, particularly when stress induced failures are driven by non-thermal dynamic electrical stresses. Theoretical work in kinetic theory, thermodynamics, and statistical mechanics have developed forms that contain exponential forms similar to Arrhenius. It is observed that the well-known Arrhenius law usually does apply, albeit with some modification, within existing models describing Physics of Failure. This is known, by example, in Black's law, Coffin-Manson or any application of Eyring's law. These include the effect of humidity or the hydrogen poisoning or other effects in semiconductors [5], [29], [21], [49], [50], [51], [52]. When these effects are simultaneously activated under multiple stress conditions inducing series of different failure modes and mechanisms, the standard reliability predictive models are questionable.

Hence regarding the Normally-Off eGaN-on-Si heterojunction transistors, a critical concern with such new technology is reliability. It is mandatory to study the role of temperature and biasing stress on the high-power degradation of this type of technology.

Quite a multitude of drift effects are influencing the performance of power switching devices, reduce switching efficiency and can compromise reliability. **Figure II-6: Schematic cross section of an AlGaN/GaN HEMT at different bias and trapping conditions (from ref).**extracted from ref [53] as well as Figure II-7 and Figure II-8 summarizes the most important reversible drift mechanisms observed in GaN transistors. Often the effects described above are correlated to each other resulting in a great variety of device parameter changes.

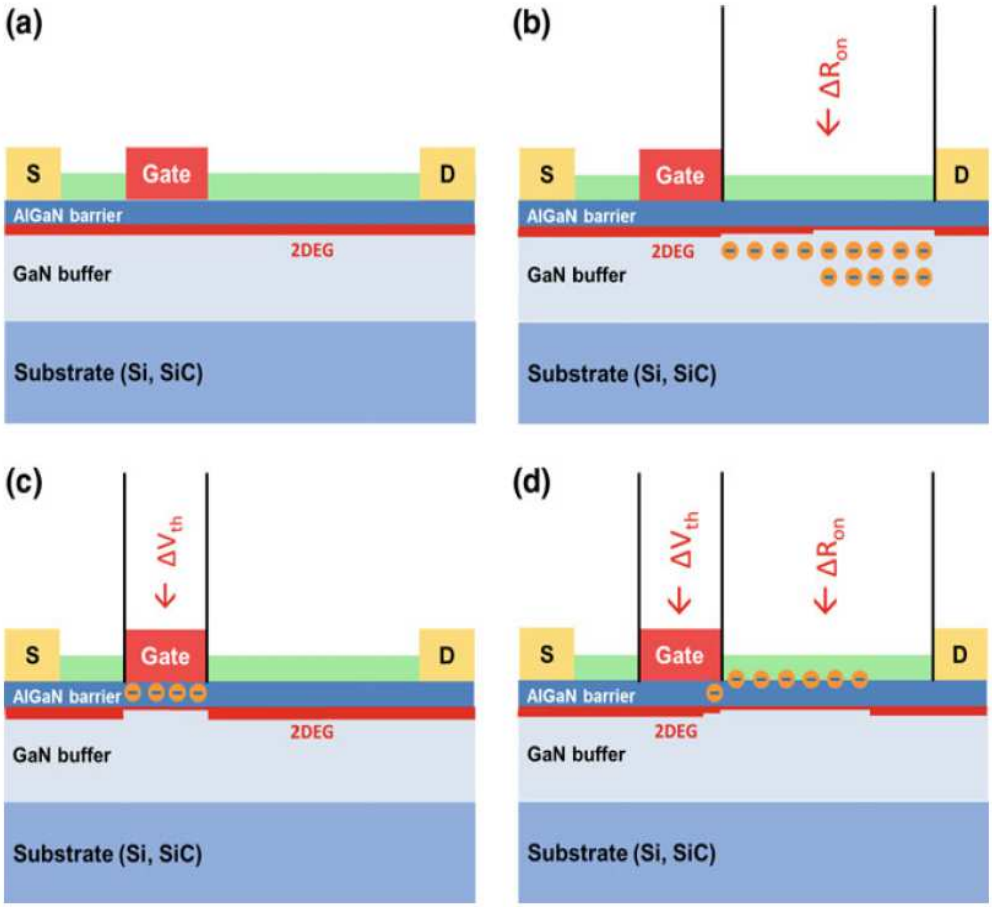


Figure II-6: Schematic cross section of an AlGaN/GaN HEMT at different bias and trapping conditions (from ref **[53]**).

a) Ideal device without any trapping at on-state: The 2DEG is fully populated with electrons.
b) Electron trapping in the vicinity of the drain access region: This situation, for example, happens immediately after switching from off-state to on-state at high drain bias. Trapped electrons **Source spécifiée non valide.** (or depleted deep acceptors **Source spécifiée non valide.**) completely or partially deplete the 2DEG and hence impede current flow. If the source or the drain access regions are influenced by trapping, the on-state resistance changes. This gives rise to the so-called dynamic on-state resistance increase ΔR_{on_dyn}.
c) Electron trapping in the vicinity of the gate (underneath the gate) results in a partial depletion of the channel just underneath the gate: As these charges aid depleting the channel, a lower gate voltage is necessary to fully turn off the device—the threshold voltage shifts to more positive values.
d) Electron trapping at the interface to passivation layer and in the barrier layer: This leads to both a threshold voltage shift and to the formation of a virtual gate at the interface oriented towards the drain ⟶ reduction of 2DEG electron concentration, R_{on_dyn} increase.

Negative charges trapped in the source or drain access region, will reduce the 2DEG electron density resulting in an increase of on-state resistance. Traps confine to regions underneath the gate and in presence of negative charges the equilibrium must be compensated by a positive threshold voltage shift; collaterally threshold voltage can also shift in negative direction if positive charges are introduced (for example, if electrons are emitted from donor traps).

For GaN RF devices G. Meneghesso [54] degradation can still be observed on chosen parameters named "***indicators***" identified as key parameters showing pre-evolution of the future performance of a device.

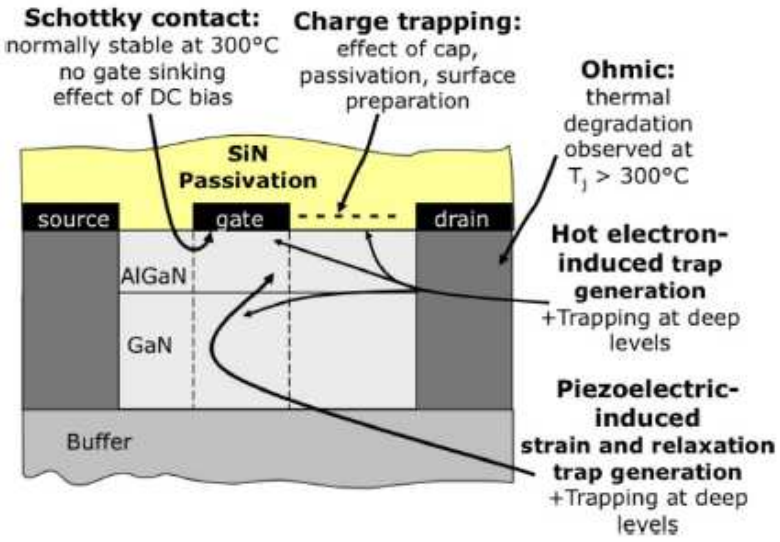


Figure II-7: Schematic cross section of an AlGaN/GaN HEMT RF transistor, identifying critical areas which can be subjected to degradation (as per ref. **[54]**)

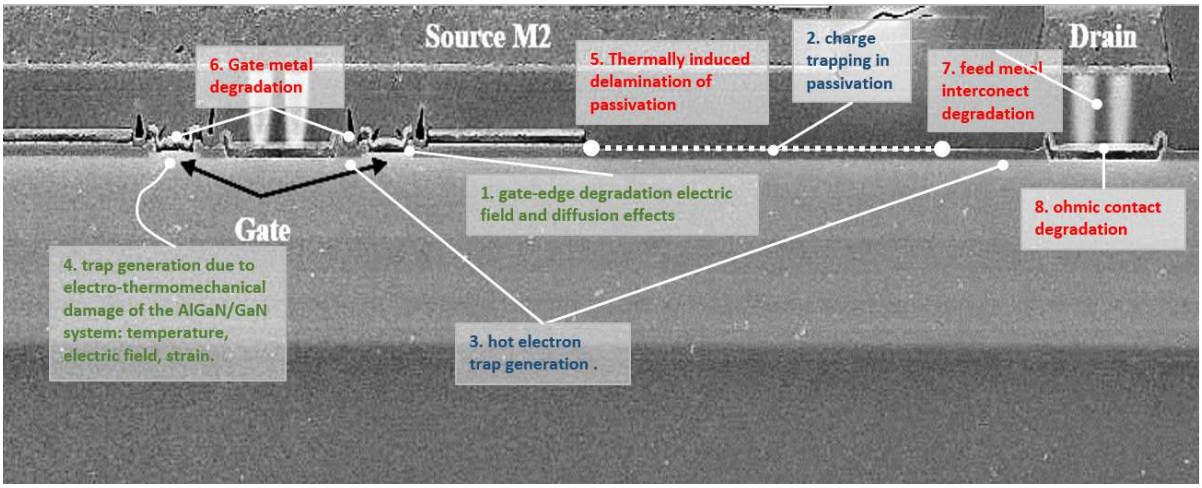


Figure II-8: Failure mechanisms recently identified on GaN HEMTs. Mechanisms identified in red (5, 6, 7, 8) are thermally activated mechanisms. Mechanisms 2 and 3 are related to the presence of hot electrons present at high bias conditions. Mechanisms 1 and 4 (green) are peculiar to GaN devices due to the polar and piezoelectric nature of this semiconductor. (freely arranged from ref. **[27]**)

C. Physical Reliability models applied to DSM technology

In a recent paper published in RAMS 2018 conference titled "Entropic Approach to Measure Damage with Applications to Fatigue Failure and Structural Reliability" [55], H. Yunm et al. suggest several entropies as candidates of proper damage measurements in fatigue process. The entropic approaches are expected to develop in order to apply them in practical reliability analyses and prognosis. The paradigm of the Transition State Theory (TST) developed by E. Wigner in 1934 [14] and by M. Evans, M. Polanyi in 1938 [13] is an approach we can benefit if we adapted to the concept of a unified semiconductor reliability model. Indeed in the early last century, the TST was applied to chemistry transformations by H. Eyring [21], and S. Glasstone et al. [15] in 1941. The TST was developed in chemistry based on the Hammond's postulate [16] published in 1955 applied to physical organic chemistry.

From the 60's and during the following 40 years, several authors report complementary reliability models centered around the mobile-ion (Na+) drift within dielectric [18], to intrinsic mechanism as TDDB [56] for example, while in mid-80's, J. W. McPherson and D.A. Baglee [22], [57] report of stress dependent activation energy and develop a generalized Eyring model in order to better understand thermally activated failure mechanisms (J.W. McPherson, Reliability Physics and Engineering – Time-to-Failure Modeling, 3rd ed., Springer, chapter 9 [23]).

Similarly in 2013, the stress dependent activation energy model based on Boltzmann-Arrhenius-Zhurkov model (BAZ) for multiple stress proposed by E. Suhir [24], [25], [58] presented a generalization of the principle proposed by S.N. Zhurkov [59] in 1965 and also D.R. Cox [60], [20] in 1972. Suhir suggests to consider any type of stress (not only mechanical but also,

electrical or external) as key parameters to modify the equivalent activation energy of a given failure mechanism. In other word, the activation cannot be considered as a constant parameter with respect to any type of stress (temperature range, biasing and signal stress range). We can view this approach as an equivalent TST concept applied to reliability paradigm [25].

Multiple failure mechanisms and Physics of degradation in semiconductors may occur in a single set of time-to-failure data but without obvious points of inflection to help separate the mechanisms. J. McPherson in his book 3rd edition of Reliability Physics and Engineering provides the basics of reliability modelling [23], [22] recalled *generally, materials/devices exist in metastable states. These states are referred to as being metastable because they are only apparently stable. Metastable states will change/degrade with time. The rate of degradation of the materials (and eventual time-to-failure for the device) can be accelerated by an elevated stress (e.g., mechanical stress, electrical stress, electrochemical stress, etc.) and/or elevated temperature*.

The Gibbs free energy description of material/device degradation is illustrated relating from ref [23] in Figure II-9, and Figure II-10. Considering the initial state to be a sound device before aging and the final state of a degraded device (either catastrophically failed or degraded and not compliant to the acceptable performance limit of the device). It is important to note at this point the reliability model is described by parameter drift degradation as a function of time and not as random failure paradigm. In this diagram the net reaction rate is a dynamic equilibrium between forward and reverse reaction meaning the degradation could be reversible. The net reaction process is written:

$$k_{Net} = k_{Forward} - k_{Reverse} \quad \text{Eq. II-1}$$

With $k_{forward}$ and $k_{reverse}$ as function of temperature and stress S applied:

$$k_{Forward} = exp\left(-\frac{\Delta G_F^{\ddagger}(S,T)}{k \cdot T}\right) \quad \text{Eq. II-2}$$

and

$$k_{Reverse} = exp\left(-\frac{\Delta H_R^{\ddagger}(S,T)}{k \cdot T}\right) \quad \text{Eq. II-3}$$

Where T is the temperature in Kelvin, k the Boltzmann constant, and $\Delta G^{\neq}_{F}(S, T)$ represents the free energy of activation associated with the reaction process for the forward reaction. $\Delta H^{\circ}_{R}(S, T)$ represents the change in enthalpy required the final state reached for the reverse reaction. ΔS° is the entropy change, e.g. the driving force for a device degradation.

Figure II-9 represents the simplified Free Gibbs Energy diagram for a reaction activated only by the temperature for a sound device to a degraded device and vice-versa. Let's assume a device placed in an electronic system designed to implement an electronic function for a given application. The electronic function is supposed to fulfil a mission for a given life mission without failure, meaning without exceeding performance limits defined by the system functionality. In such condition a degraded device is defined as a device whose characteristics have changed to a certain extent above a given limit criteria fixed for the system to function properly. Let's call the final state to be the state the device is for the failure criteria to reach. So, the equations IV.1, 2 and 3 describe the probability the transition rates are associated to a "Sound device" to reach a failure limit.

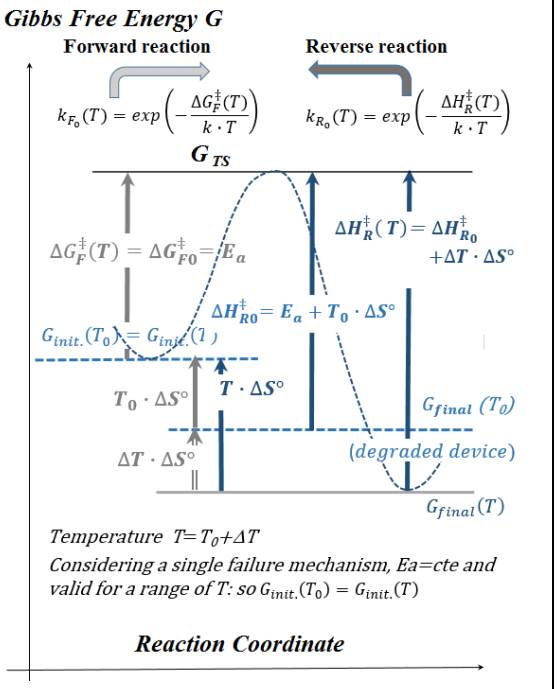


Figure II-9: Gibbs Free Energy diagram for *T* temperature increase applied, e.g. Arrhenius thermal activation energy diagram.

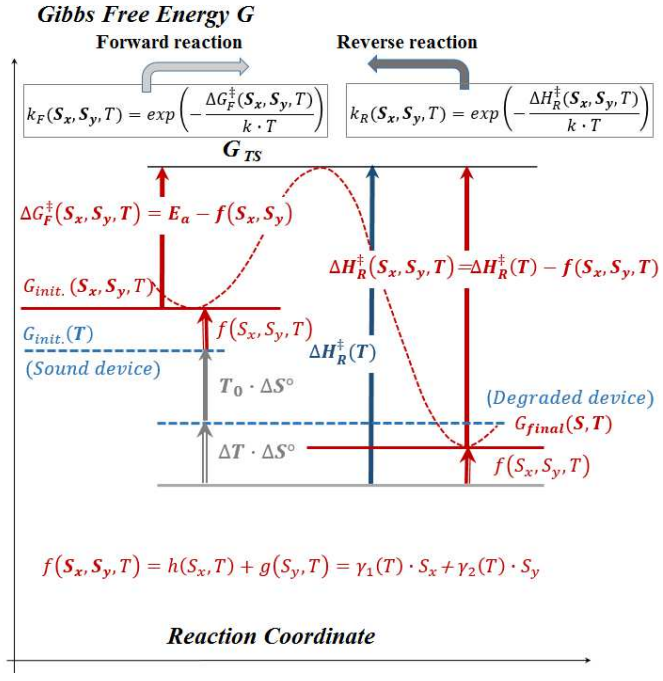


Figure II-10 : Gibbs Free Energy diagram considering reaction with two external stresses Sx and Sy applied and temperature T (diagram modified from ref **[23]**).

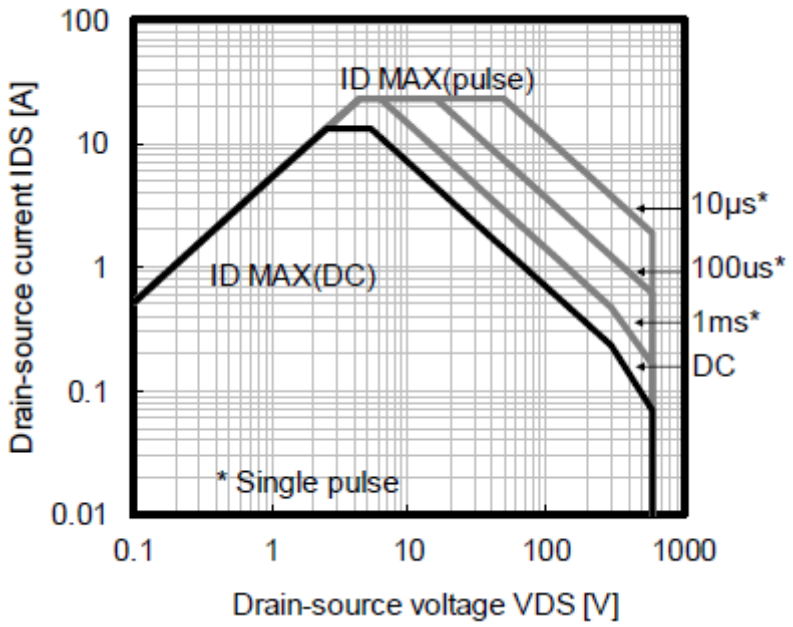


Figure IV.3: Example of Maximum rating and Safe Operating Area plot for a power transistor.

Figure II-10 is the Gibbs free energy diagram when considering multiple stress conditions and show how they impact the diagram. Note that the equivalent activation energy $\Delta G^{\ddagger}_F(S, T)$ is reduced when stresses are applied and can reduce to zero for extreme stress limit. So, a stress current may conduct to an instantaneous burnout catastrophic failure. As a consequence the energy barrier vanishes because of like a catalyst effect and the stress function $f(I_{burnout})$ compensate exactly the initial $\Delta G^{\ddagger}_{F0} = E_a$ at a given temperature T. So the equivalent activation energy range from E_a *(the Arrhenius pure thermal value)* to 0 for a current stress range for $0 \leq I \leq I_{burnout}$ or for $0 \leq I/Iburnout \leq 100\%$. Similarly, we can assume multiple stress to cumulate and the combination of them may reach the instantaneous catastrophic failure when the initial Arrhenius activation energy is fully compensated. These conditions can be compared when considering current, voltage and power dissipation parameters of a transistor defined by their limits called Safe Operating Area as shown in the example given in figure 3. The static characteristics limits related to current (vertical axis), voltage (horizontal axis) and DC or pulsed power dissipation are plotted in log-log scale. In this case, the *I(V)* curves are limited by the power dissipation capability of the device (strait lines of the top right corner of

the figure). The burnout can be reached either under high current or high voltage or high power overstresses. The combination of current and voltage stresses is seen through the power dissipation parameter: either one or the other or both lead to the destruction of the device. This paradigm can be modelled too by **Figure II-10**. It is worst to consider these parameters limits as static limits but similarly, dynamic and pulsed operation can be set.

From figure IV.2 we can observe the enthalpy $\Delta H_R^{\ddagger}(S = 0, T_0)$ is by definition independent of the temperature variation ΔT because the transition state G_{TS} as well as the final state G_{final} is only related to initial T_0 value. We assume the expression of the enthalpy is:

$$\Delta \boldsymbol{H_R^{\ddagger}(S_i = 0, T_0) = \Delta H_{R0}^{\ddagger} = E_a + T_0 \cdot \Delta S^{\circ}} \qquad \text{for i = 1 , n} \qquad \text{Eq. II-4}$$

Let consider the case to superpose n stress conditions S_i. For the simplest case of n=2 stressors S_x and S_y,, a generalized Eyring model was established in ref [57] describing thermally activated failure mechanisms in materials/devices under stress as supported by figure IV.2. The forward reaction described by the Gibbs Free energy (respectively Enthalpy for the reverse reaction) is a function of the generalized stress S, and we set a Taylor Series expansion around point $(S_x, S_y)=(S_a, S_b)$ expanded to (limited to 1st order):

$$\boldsymbol{f(S_x, S_y, z) = f(S_a, S_b, c) + \frac{\partial f(S_a, S_b, c)}{\partial S_x} \cdot (S_x - S_a) + \frac{\partial f(S_a, S_b, c)}{\partial S_y} \cdot (S_y - S_b)} \qquad \text{Eq. II-5}$$

We consider the two infinitely differentiable functions $\Delta G^{\ddagger}_F (S_x, S_y, T)$ and $\Delta H^{\ddagger}_R (S_x, S_y, T)$ and limited to the Taylor series expansion linear around $(S_x, S_y) = (S_{a0}, S_{b0})$:

$$\Delta \boldsymbol{G_F^{\ddagger}(S_x, S_y, T) \approx \Delta G_F^{\ddagger}(S_{a0}, S_{b0}, T_0) + \left[\frac{\partial\left(\Delta G_F^{\ddagger}(S_x, S_y, T)\right)}{\partial S_x}\right]_{S_{a0}, S_{b0}} \cdot (S_x - S_{a0}) + \left[\frac{\partial\left(\Delta G_F^{\ddagger}(S_x, S_y, T)\right)}{\partial S_y}\right]_{S_{a0}, S_{b0}} \cdot (S_y - S_{b0})} \qquad \text{Eq. II-6}$$

$$\Delta \boldsymbol{H_R^{\ddagger}(S_x, S_y, T) \approx \Delta H_R^{\ddagger}(S_{a0}, S_{b0}, T_0) + \left[\frac{\partial\left(\Delta H_R^{\ddagger}(S_x, S_y, T)\right)}{\partial S_x}\right]_{S_{a0}, S_{b0}} \cdot (S_x - S_{a0}) + \left[\frac{\partial\left(\Delta H_R^{\ddagger}(S_x, S_y, T)\right)}{\partial S_y}\right]_{S_{a0}, S_{b0}} \cdot (S_y - S_{b0})} \qquad \text{Eq. II-7}$$

According to figure 1 and 2, we observe:

$$\Delta \boldsymbol{G_{F0}^{\ddagger}(T_0) = E_a} \quad \text{and} \quad \boldsymbol{E_a = \Delta H_{R0}^{\ddagger} - T_0 \cdot \Delta S^{\circ}} \qquad \text{Eq. II-8}$$

$$\Delta \boldsymbol{H_R^{\ddagger}(S_x, S_y, T) = E_a + T \cdot \Delta S^{\circ} - \frac{1}{2} \cdot f(S_x, S_y, T)} \qquad \text{Eq. II-9}$$

$$\Delta \boldsymbol{G_F^{\ddagger}(S_x, S_y, T) = E_a - \frac{1}{2} \cdot f(S_x, S_y, T)} \qquad \text{Eq. II-10}$$

Let's assume simply $$\boldsymbol{f(S_x, S_y, T) = h(S_x, T) + g(S_y, T)} \qquad \text{Eq. II-11}$$

The factor ½ is set because we consider the same stress level affecting the entropy on the initial state of fig. 1.b and the enthalpy affecting the final state.

At zero stress the activation energy is converging to the standard Arrhenius thermal activation energy so we must assume (see fig. IV.2) $f(S_x=0, S_y=0) = 0$ and then

$$\boldsymbol{h(S_x = 0, T) = 0} \quad \text{and} \quad \boldsymbol{g\left(S_y = 0, T\right) = 0} \qquad \text{Eq. II-12}$$

According to Eyring law the stress function $f(S_x, S_y)$ is defined by :

$$\boldsymbol{h(S_x, T) = \gamma_1(T) \cdot S_x} \qquad \text{Eq. II-13}$$

and

$$\boldsymbol{g\left(S_y, T\right) = \gamma_2(T) \cdot S_y} \qquad \text{Eq. II-14}$$

We assume parameters $\gamma_1(T)$ and $\gamma_1(T)$ are defined by a temperature-dependence law of the form (see [57]:

$$\boldsymbol{\gamma_1(T) = a_0 + a_1 \cdot k \cdot T} \qquad \text{Eq. II-15}$$

$$\boldsymbol{\gamma_2(T) = b_0 + b_1 \cdot k \cdot T} \qquad \text{Eq. II-16}$$

Hence III.7.d, 8 and 9 combined give $f(S_{a0}, S_{b0}, T) = (a_0 + a_1 \cdot k \cdot T) \cdot S_{a0} + (b_0 + b_1 \cdot k \cdot T) \cdot S_{b0}$ Eq. II-17

From equations IV.7.c we get

$$\left[\frac{\partial\left(\Delta G_F^{\ddagger}(S_x,S_y,T)\right)}{\partial S_x}\right]_{S_{a0},S_{b0}} = \left[-\frac{\partial\left(f(S_x,S_y)\right)}{\partial S_x}\right]_{S_{a0},S_{b0}} = -\frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T) \qquad \text{Eq. II-18}$$

and

$$\left[\frac{\partial\left(\Delta G_F^{\ddagger}(S_x,S_y,T)\right)}{\partial S_y}\right]_{S_{a0},S_{b0}} = \left[-\frac{\partial\left(f(S_x,S_y)\right)}{\partial S_y}\right]_{S_{a0},S_{b0}} = -\frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T) \qquad \text{Eq. II-19}$$

The partial derivative of IV.7.b with respect to S_x *or* S_y gives

$$\left[\frac{\partial\left(\Delta H_R^{\ddagger}(S_x,S_y,T)\right)}{\partial S_x}\right]_{S_{a0},S_{b0}} = -\left[\frac{\partial\left(f(S_x,S_y)\right)}{\partial S_x}\right]_{S_{a0},S_{b0}} = -\frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T) \qquad \text{Eq. II-20}$$

And

$$\left[\frac{\partial\left(\Delta H_R^{\ddagger}(S_x,S_y,T)\right)}{\partial S_y}\right]_{S_{a0},S_{b0}} = -\left[\frac{\partial\left(f(S_x,S_y)\right)}{\partial S_y}\right]_{S_{a0},S_{b0}} = -\frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T) \qquad \text{Eq. II-21}$$

The parameters a_0, a_1, b_0, b_1 and $\Delta S°$ must be determined experimentally.
Consequently injecting equations IV.7, 10 and 11 into equations 6.a and 6.b, we obtain:

$$\Delta G_F^{\ddagger}(S_x,S_y,T) \approx E_a - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot S_{a0} - \frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T)\cdot S_{b0} - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot(S_x - S_{a0}) - \frac{1}{2}\cdot (b_0 + b_1 \cdot k \cdot T)\cdot(S_y - S_{b0}) \qquad \text{Eq. II-22}$$

$$\Delta H_R^{\ddagger}(S_x,S_y,T) \approx E_a + T\cdot\Delta S° - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot S_{a0} - \frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T)\cdot S_{b0} - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot(S_x - S_{a0}) - \frac{1}{2}\cdot (b_0 + b_1 \cdot k \cdot T)\cdot\left(S_y - S_{b0}\right) \qquad \text{Eq. II-23}$$

Which reduce to:

$$\Delta G_F^{\ddagger}(S_x,S_y,T) \approx E_a - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot S_x - \frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T)\cdot S_y \qquad \text{Eq. II-24}$$

$$\Delta H_R^{\ddagger}(S_x,S_y,T) \approx E_a + T\cdot\Delta S° - \frac{1}{2}\cdot(a_0 + a_1 \cdot k \cdot T)\cdot(S_x) - \frac{1}{2}\cdot(b_0 + b_1 \cdot k \cdot T)\cdot\left(S_y\right) \qquad \text{Eq. II-25}$$

Let's define $S_x = x \cdot S_{Br1}$ and $S_y = y \cdot S_{Br2}$ Eq. II-26

and $S_{a0} = x_0 \cdot S_{Br1}$ and $S_{b0} = y_0 \cdot S_{Br2}$ Eq. II-27

where S_{Br1} and S_{Br2} are respectively the breakdown or burnout experimental values of the two stress parameters considered for S_x and S_y. In such a case the stress are defined by the percentages x and y respectively for S_x and S_y with $0 \leq x \leq 100\%$ and $0 \leq y \leq 100\%$.

$$\Delta G^{\ddagger}_{F_{LS}}(S_x, S_y, T) \approx E_a - \frac{1}{2}\cdot(a_0 + a_1\cdot k\cdot T)\cdot x\cdot S_{Br1} - \frac{1}{2}\cdot(b_0 + b_1\cdot k\cdot T)\cdot y\cdot S_{Br2} \quad \text{Eq. II-28}$$

$$\Delta H^{\ddagger}_{R_{LS}}(S_x, S_y, T) \approx E_a + T\cdot\Delta S° - \frac{1}{2}\cdot(a_0 + a_1\cdot k\cdot T)\cdot x\cdot S_{Br1} - \frac{1}{2}\cdot(b_0 + b_1\cdot k\cdot T)\cdot y\cdot S_{Br2} \quad \text{Eq. II-29}$$

The net reaction rate under stress conditions for the breakdown process from equation IV.1 become:

$$k_{net_{LS}} = exp\left(-\frac{E_a - \frac{1}{2}(a_0 + a_1\cdot k\cdot T)\cdot x\cdot S_{Br1} - \frac{1}{2}(b_0 + b_1\cdot k\cdot T)\cdot y\cdot S_{Br2}}{k\cdot T}\right) - exp\left(-\frac{E_a + T\cdot\Delta S° - \frac{1}{2}(a_0 + a_1\cdot k\cdot T)\cdot x\cdot S_{Br1} - \frac{1}{2}(b_0 + b_1\cdot k\cdot T)\cdot y\cdot S_{Br2}}{k\cdot T}\right) \quad \text{Eq. II-30}$$

Hence, equation III.16 expresses with a thermal term related to the activation energy as a function of stress conditions in a like-catalyst thermodynamic process and a stress term depending of the stresses applied in the form (same conclusion as ref [24]):

$$k_{net} = k_1(x, y)\cdot exp\left(-\frac{E_{a_{equ}}(x,y)}{k\cdot T}\right) \quad \text{Eq. II-31}$$

with $$\boldsymbol{E_{a_{equ}}(x, y) = E_a - \frac{1}{2}\cdot a_0\cdot x\cdot S_{Br1} - \frac{1}{2}\cdot b_0\cdot y\cdot S_{Br2}} \quad \text{Eq. II-32}$$

and $k_1(x, y) = k_0\cdot exp\left(\frac{1}{2}\cdot a_1\cdot x\cdot S_{Br1} + 1/2\cdot b_1\cdot y\cdot S_{Br2}\right)$ and $k_0 = \left(1 - exp\left(-\frac{\Delta S°}{k}\right)\right)$ Eq. II-33

From these equations we can observe the **equivalent activation energy is dependent of the temperature T_0 and is increasing with temperature**. So, under the stress conditions, **the failure mechanism model is a non-uniform acceleration mechanism**.

We can also observe the equivalent activation energies is based on the **constant Arrhenius activation energy** E_a of the mechanism under investigation but modified by a series of factors related to the stress condition applied.

Accelerating factor can be expressed considering range stress conditions. We can simply write the *AF* considering operational conditions versus reference stress conditions (sub-labelled 'op' respectively 'ref') as follows:

$$AF(x_{op}, y_{op}, T_{op}, x_{ref}, y_{ref}, T_{ref}) = exp\left[-\frac{E_a}{k}\left(\frac{1}{T_{op}} - \frac{1}{T_{ref}}\right)\right]\cdot exp\left[\frac{a_0\cdot S_{Br1}}{2\cdot k}\left(\frac{x_{op}}{T_{op}} - \frac{x_{ref}}{T_{ref}}\right)\right]\cdot exp\left[\frac{b_0\cdot S_{Br2}}{2\cdot k}\left(\frac{y_{op}}{T_{op}} - \frac{y_{ref}}{T_{ref}}\right)\right]\cdot exp\left[\frac{a_1}{2}\cdot S_{Br1}\cdot(x_{op} - x_{ref}) + \frac{b_1}{2}\cdot S_{Br2}\cdot(y_{op} - y_{ref})\right] \quad \text{Eq. II-34}$$

Precautions associated with accelerated testing

The x and y values are assumed to range from no stress (e.g. x=y=0%) to highest stress (e.g. x=y=100%) and the AF expression is exact for any range of external stress applied. These theoretical calculations must be taken with extreme cautions. As we use to set the acceleration is supposed to be related to failure physics without changing the physics. Indeed, assuming a single failure mechanism of PoF, the accelerating factor calculated for various stress conditions superimposed, is non-uniform and the equivalent activation energy is stress and temperature dependent. This must be verified using Weibull (or lognormal) distributions when conducting proper test experiments and observing a single value of E_a activation energy parameter determined within a limited domain of range of temperature.

Based on failure rate expression the time-to-failure for reference stress accelerated conditions (REF) compared to operational condition (labelled 'op') becomes:

$$TTF_{50\%}\left(S_{x_{op}}, S_{y_{op}}, T_{op}\right) = TTF_{50\%}(S_{REF}, S_{REF}, T_{REF})\cdot exp\left[\frac{E_a}{k}\left(\frac{1}{T_{REF}} - \frac{1}{T_{op}}\right)\right]\cdot exp\left[\frac{a_0\cdot S_{Br1}}{2\cdot k}\left(\frac{x_{op}}{T_{op}} - \frac{x_{REF}}{T_{REF}}\right)\right]\cdot exp\left[\frac{b_0\cdot S_{Br2}}{2\cdot k}\left(\frac{y_{op}}{T_{op}} - \frac{y_{REF}}{T_{REF}}\right)\right]\cdot exp\left[\frac{a_1}{2}\cdot S_{Br1}\cdot(x_{op} - x_{REF}) + \frac{b_1}{2}\cdot S_{Br2}\cdot\left(y_{op} - y_{REF}\right)\right] \quad \text{Eq. II-35}$$

These mathematics and physic approaches show how the activation energy is related to the stress and temperature applied and can no-longer be considered as a constant to extrapolate some experiment under high stress to nominal mission operation profile. That's the reason why several end-user Industries and Institutions are very cautious to perform lifetest conditions as close as the nominal conditions because of the change of activation energy attributed sometime wrongly to new (or different) failure mechanism while it is simply explained by the interaction of stress and temperature effect on the measured activation energy or to a best extend to Eyring law.

One may note the burnout limits S_{Br1} and S_{Br2} are characteristic parameters which may change during aging. These limits can drift also during aging time and hence as the stress applied S_x or S_y (eq. III.14) this means the x and y percentage stress values increase during aging from low stress level to high stress level.

The general statement that during useful life the failure rate is supposed to be constant can be wrong when approaching the

end of life. When devices get closer to the wearout occurrence, it is assumed they degrade faster and their burnout limit reduces when time elapse. In such a way, the applied stress cannot be considered constant and is reaching high-stress level comparatively to initial low stress.

This observation can give some insight to describe how the failure rate can be considered as non-constant and hence to propose a vision on EOL of the bathtub curve.

D. Reliability and probability mathemathics

Reliability is defined as the probability that a device will function over some period of time, and usually is measured in Failures-In-Time units (FIT units). The FIT is a rate, defined as the number of expected device failures per 10^9 parts-hours.

Failure Rate (λ) is calculated by dividing the total number of failures or rejects by the cumulative time of operation. In the HTOL model, the cumulative time of operation is referred to as Number of Device Hours (NDH):

$$NDH = D \cdot H \cdot A_F \quad \text{Eq. IV-36}$$

where: D = Number of Devices Tested
H = Test Hours per Device
A_F = Acceleration Factor derived from the Arrhenius equation and or other stressor.

A FIT is assigned to each component multiplied by the number of devices in a system as an approximation of the expected system reliability. For the reliability model of an entire system, the FIT rates of each component in the system are summed together.

The *Failure Rate (λ)* in FITs (Failures per billion unit-hours) is given by:

$$\lambda = \frac{\#\cdot failures}{NDH} \quad \text{Eq. IV-37}$$

The conventional chi-squared expression for failure rate, λ, is:

$$\lambda = \frac{\chi^2(2n+2,1-\alpha)\cdot 10^9}{2\cdot D\cdot H\cdot A_F} \quad \text{Eq. IV-38}$$

Where $\frac{\chi^2(2n+2,1-\alpha)}{2}$ is the upper confidence value for "n" failures and upper confidence limit, α (expressed in %). The value for 2n + 2 degrees of freedom and the probability, 1 – α, can be obtained from a table or calculated using Microsoft Excel functions.

According to this expression, to demonstrate zero failure rate must account either a sample size to go to infinity, AF to go to infinity or test time to go to infinity [61]. In such case we get no product to ship the accelerating factor cannot be infinity because physically impossible) or test time will end when products are too old to ship. Achieving a zero-failure rate goal using reliability testing is impossible.

A practical numerical application from Standard Device Qualification from HTOL qualification sequence leads to the following limitation:

For defect density limits:

let's assume 0 failures on 92 devices sample size (SS) under tests, a 60% confidence level applied on a Poisson distribution is expressed as:

$$F \leq \frac{1}{SS}\cdot ln\left(\frac{1}{1-P}\right) = \frac{1}{92}\cdot ln\left(\frac{1}{1-0.6}\right) = 0.025 = 2{,}5\% \text{ or } 25\,000 \text{ ppm.} \quad \text{Eq. IV-39}$$

Numerical application for a study case on a FinFET technology assuming 3 major failure mechanisms defined by the following Arrhenius reliability parameters (not considering other type of stressors like voltage or current for sake of simplification):

- EM- Electromigration kinetics (Ea=0.75eV (Al-Cu),
- NBTI kinetics (Ea=0.6eV),
- HCI kinetics (Ea=-0.25eV long channel or +0.25 short channel).

Assuming a qualification test experiment on 130 samples tested during 2000 hrs of 150°C HTOL with 0 failure observed.

What would be the device operating projected lifetime verified if the operational application is at T=45°C ?

In such case, what are the failure rates λ_i (for i=EM, NBTI or HCI) in FIT and the corresponding accelerating factors for each mechanism?

What are the Time-To-Failure for each failure mechanism considering a constant failure rate based on a Poisson statistic?

What would be the equivalent total λ_{total} and corresponding Time-To-Failure if all failure mechanisms are equally activated?

The hypotheses are summarized in table IV-1.a and some answers are shown to the questions above in Table IV-1.b.

Qualification test conditions		
Sample Size	130	
duration	2000	hrs
nb of hr.comp.	260000	hr.comp.
Tstress for high temp. (EM, BTI and HCI short channel)	150	°C
Tstress for HCI long channel	0	°C
Tuse	45	°C
k (Boltzmann factor)	8,62E-05	eV/°K
CL	90%	

Table IV-1.a: Example of qualification test condition for a CMOS FinFET technology with an example of associated list of failure mechanisms.

Reliability figures			Failure mechanisms	AF	Operational time verified (in years)	λ Poisson (in FIT) in operation	MTBF (in years) $1/\lambda$ Poisson	CDF *Fpoisson* (@ t = 30 ans)
Ea EM	0,75	eV	Ea EM Tref= 150°C	892	204	10	11501	0,26%
Ea NBTI	0,6	eV	Ea NBTI Tref= 150°C	229	52	39	2955	1,01%
Ea HCI	0,25	eV (short channel)	Ea HCI (short channel) Tref= 150°C	9,6	2,2	920	124	21,48%
Ea HCI	-0,25	eV (long channel)	Ea HCI (long channel) Tref=0°C	4,5	1,03	1969	58	40,39%
						λ total	MTBF equ.	CDF *Fpoisson* (@ t = 30 ans)
						2937	39	53,78%

Table IV-1.b: CMOS FinFET technology reliability assessment to predict operational time duration and CDF, MTTF for various failure mechanisms for CL=90%.

From table IV-1.a we consider a lot of 130 sample used to validate a qualification test sequence as defined by Quality Standards with two kind of lifetest conditions: one at high temperature (150°C) to stress EM, BTI and HCI for short channel transistors and one a low temperature (0°C) stress applicable to long channel based transistors (HCI mechanism has a negative activation energy).

According to the lifetest sequence duration, the number of hours.components is 260,000.

What should be the sample size and duration to demonstrate a cumulative CDF equal of less than 0.1% after 30 years (as needed in Space application)? Considering all failure mechanisms activated simultaneously (according to table IV-1.a), the answer is summarized in table IV-A.c and it is required to test a 200 Mhrs.comp (or 200,000 parts tested for 1,000 hrs) to guarantee the 0.1% CDF goal after 30 years at 90%CL. In such a case the failure rate λ_{total} is 3.8 FIT, MTBF is about 29.9 Kyears and the $F_{poisson}$ (CDF) after 30 years is as low as the 0.1% goal.

To demonstrate 30 years for CDF less than 0.1%	20000	Sample size	2E+08 hr.comp	λ_{total}	MTBF equ. (in years)	CDF *Fpoisson* (@ t = 30 years)
	10000	hrs		3,82	29898	0,10%

Table IV-1.b: CMOS FinFET technology. What should be the sample size and duration to demonstrate a cumulative CDF equal of less than 0.1% after 30 years with 90%CL.

The accelerating factors at 90% confidence level range broadly from 892 down to 4.5. The device operating projected lifetime (CL 90%) verified by the qualification stress test is projected for the operational application duration at T=45°C up to 204 years for EM mechanism, 52 years for NBTI, 2.2 years for Hot HCI and 1 year for Cold HCI mechanisms. This shows how much multiple failure mechanisms can affect the demonstrated operating time projection.

Respectively, the failure rate projected in operation assuming a Poisson distribution at 90% CL is assessed 10 FIT for EM, 39 FIT for NBTI, 920 FIT for hot HCI, and 1,969 FIT for Cold HCI. These results are not enough to prove application reliability goals.

If supposed disjointed, each failure mechanism leads to a wide range of TTF from 1,1501 years for EM, 2,955 years for NBTI, 124 years for Hot HCI and 58 years for Cold HCI. Again these figures must be considered from a different point of view to be representative of the theoretical truth.

Because of this large variety of values, we need to suppose all failure mechanisms EM, BTI and hot HCI are equally and simultaneously activated. So, we need to cumulate their corresponding $\lambda_{Poisson}$ distribution to assess a total equivalent failure rate. This sum gives λ_{total} (CL90%)= 2,937 FIT for Poisson distribution at 90%CL. Then the total MTTF is about 39 years for a Poisson distribution and not as supposed if we take into account the worst case of each failure mechanism.

We have to be careful considering these results because they are first extracted from a single lot testing and a reduced number of device. The statistic helps to predict how long a device from such a tested long sequence can be operational without any failure (still with 90% confidence level). The true λ and TTF are expected to be much greater than the values calculated. The occurrence of various failure mechanism must be considered with some caution as their activation is strongly dependent of the conditions of test which can be far from the operational one's.

A last comment is that these failures are supposed to be constant with time (the bottom of the bathtub curve) but this is wrong way to address end of life reliability figure. We know that the failure mechanisms must be considered as wearout (to the end of the bathtub curve) and we should take into consideration WEIBULL (or LOGNORMAL respectively) distribution to determine the $MTTF_{wearout}$ as 63% lot failure (respectively 50% lot failure). This causes that the $MTTF_{wearout}$ is not the MTTF deduced from the constant failure rate.

Therefore, Standard HTOL is relatively poor to predict MTTF associated constant failure rate and ineffective for wearout modelling as no failure are observed. Furthermore, the MTTF is not a zero-failure guarantee (is a 50% lot failure) and we should define rather a $MTTF_{0.1\%}$ for high reliability application.

Continuing we can question, how many devices would you need to test to assure that the fraction defective is F≤ 35 ppm?

Answering such a question need to invoke the "Stress-Resistance" concept. The implementation of the "Stress-Resistance" method usually refers to information called "lump sum". The most basic version of the "Stress-Resistance" method calculates the probability of failure which results from the interaction between two probability laws. In general, if any two laws are explicitly known, the probability of failure can be calculated by numerical resolution of a convolution integral. Only a few particular cases lead to purely analytical solutions, when the two laws belong to the same statistical family. Among the most classic examples, we can cite the interaction between pairs of counterpart laws, normal or log-normal, which allow to obtain exact analytical solutions as presented by L. Pierrat [62], and D. Delaux. [63]. Nevertheless, assuming correct hypotheses, approximation can be set [64]. At the 60% confidence level (P=0.6) and with accepting on finding zero defects (x=0) the sample size (SS) is deduced from various statistical distributions:

$$\text{Poisson: } SS = \frac{1}{F} \cdot ln\left(\frac{1}{1-P}\right) = \frac{1}{35ppm} \cdot ln\left(\frac{1}{1-0.6}\right) = 26{,}180\ devcies$$

$$\text{Chi Square: : } SS = \frac{[\chi^2(P,2)]/2}{F} = \frac{[\chi^2(0.6,2)]/2}{35\ ppm} = \frac{1.8326/2}{35\ ppm} = 26{,}180\ devices$$

So above sample size SS>50, little or no differences is observed in statistical distributions. In fact, such situations are not entirely realistic, because in general, the laws resulting from statistical estimates deviate from the theoretical laws to which they are assimilated.

The acceleration factor, AF, is the most problematic consideration. The industry recognizes that an AF is given only for each failure mechanism individually by specific testing from the manufacturer, and is defined as the ratio between the time it takes a certain fraction of the devices to fail, under a certain amount of stress or use conditions, and the corresponding time under more severe stress or use conditions. In the case of multiple mechanisms, the TTF for each mechanism must be considered separately for each associated AF per mechanism.

The Time-to-Failure (TF) of each mechanism can also be defined from the physical models of the mechanisms, which are described on the JEDEC publication JED-122G [5] (see also ref [23]):

TDDB anode hole injection exponential 1/E model (Chen & Hu 1985) from Fowler-Nordheim tunnelling current:

$$TF_{TDDB_{\frac{1}{E}model}} = e^{G(T)\cdot\frac{1}{E_{0x}}} \cdot e^{\frac{E_{a_{TDDB}}}{k\cdot T}} \qquad \text{Eq. IV-36}$$

TDDB thermoelectrical exponential E model (McPherson & Baglee 1985)

$$TF_{TDDB_{Emodel}} = e^{\gamma(T)\cdot E_{ox}} \cdot e^{\frac{Ea_{TDDB}}{k\cdot T}} \quad \text{Eq. IV-40}$$

TDDB Power Law voltage V-model for hyper-thin Si02 dielectrics (<40Å)

$$TF_{TDDB_{Power\ lawmodel}} = B_0(T) \cdot {V_{ox}}^{-n} \cdot e^{\frac{Ea_{TDDB}}{k\cdot T}} \quad \text{Eq. IV-41}$$

TDDB exponential $E^{1/2}$ model

$$TF_{TDDB_{E1/2model}} = C_0(T) \cdot e^{-\alpha\cdot\sqrt{E}} \cdot e^{\frac{Ea_{TDDB}}{k\cdot T}} \quad \text{Eq. IV-42}$$

Hot Carrier Injection (Takeda 1983)

$$AT = f \cdot \left(\frac{I_{sub}}{W}\right)^{-N} \cdot e^{\frac{Ea_{HCI}}{k\cdot T}} \quad \text{Eq. IV-43}$$

Negative-Bias Temperature Instability

$$TF_{NBTI} = V^{\frac{\alpha}{m}} \cdot e^{\frac{Ea_{NBTI}}{k\cdot T}} \quad \text{Eq. II-44}$$

Electromigration (Black 1969)

$$TF_{EM} = A_0 \cdot (J - J_{crit})^{-p} \cdot e^{\frac{Ea_{EM}}{k\cdot T}} \quad \text{Eq. IV-45}$$

Stress migration (McPherson & Dunn 1987) due to mechanical stress inducing plastic deformation of metal with time

SM in Aluminium or Copper interconnects

$$Creep(voiding)rate = B_0 \cdot (T_O - T)^n \cdot e^{\frac{Ea_{SM_{Al}}}{k\cdot T}} \quad \text{Eq. IV-46}$$

Humidity model (HM) exponential reciprocal-humidity

$$TF_{HM_{exp-reciprocal}} = A_0 \cdot e^{\frac{b}{RH}} \cdot e^{\frac{Ea_{HM}}{k\cdot T}} \quad \text{Eq. IV-47}$$

Humidity model (HM) power law humidity model

$$TF_{HM_{power-law}} = A_0 \cdot RH^{-n} \cdot e^{\frac{Ea_{HM}}{k\cdot T}} \quad \text{Eq. IV-48}$$

Humidity model (HM) exponential humidity model

$$TF_{HM_{power-law}} = A_0 \cdot e^{-a\cdot RH} \cdot e^{\frac{Ea_{HM}}{k\cdot T}} \quad \text{Eq. IV-49}$$

Thermal-cycling/Fatigue induced mechanisms

Coffin-Manson model for low-cycle fatigue

$$\#Cycle - to - Failure = A_0 \cdot \left(\Delta\varepsilon_p\right)^{-s} \quad \text{Eq. IV-50}$$

Where $\Delta\varepsilon_p$ is the plastic strain range.

Coffin-Manson model for temperature cycling

$$\Delta\varepsilon_p \propto (\Delta T - \Delta T_0)^{\beta} \quad \text{Eq. IV-51}$$

Failure rates are often expressed in term of failure units (FITs): 1*FIT* = 1 failure in 10^9 device-hours.

TTF is statistical and a distribution of time-to-failure is observed. Life distributions are defined based on three mathematical functions:

a) the probability density function for device or system failures *f(t)* which relate the relative frequency of failure to time, *t* ;
b) the cumulative distribution function, *F(t)*, which gives the summation of failure proportion to that time.
c) The curve of failure rate *λ(t)*

The instantaneous failure rate, the hazard rate function *λ(t)*, is the ratio of the number of failures during the time period Δt,

for the devices that were healthy at the beginning of testing (operation) to the time period Δt.

$$\lambda(t) = \frac{f(t)}{1-F(t)} \quad \text{Eq. IV-52}$$

The cumulative probability distribution function (CDF) *F(t)* for the probability of failure is related to the probability density distribution function *f(t)* as

$$F(t) = \int_0^t f(x) \cdot dx \quad \text{Eq. IV-53}$$

and the reliability function *R(t)*, the probability of non-failure is defined as

$$R(t) = 1 - F(t) \quad \text{Eq. IV-54}$$

Probability data obtained when performing accelerated tests can be modeled by various distribution models, such as exponential law, Weibull law, normal or log-normal distributions, beta distributions, etc. The main life distribution functions are:

- Exponential
- Normal
- Lognormal
- Weibull distribution

Exponential distribution summary:

Probability density function $f(t) = \lambda \cdot e^{-\lambda \cdot t}$ Eq. IV-55

Cumulative density function: $F(t) = 1 - e^{-\lambda \cdot t}$ Eq. IV-56

Instantaneous Failure Rate, hazard rate: $\lambda(t) = \lambda$ Eq. IV-57

Significant properties

Mean (average) of MTBF: $MTBF = \frac{1}{\lambda}$

Normal distribution summary:

Probability density function $f(t) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{t-\mu}{\sigma}\right)^2}$ Eq. IV-58

Cumulative density function: $F(t) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_0^t e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2} \cdot dx$ Eq. IV-59

Instantaneous Failure Rate, hazard rate: $\lambda(t) = \frac{\frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}}{1 - \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_0^t e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot dx}$ Eq. IV-60

Significant Distribution properties

Median (50% failure): $t = t_{50\%} = \mu$

Mean (average): $t = \mu$

Location parameter: μ

Shape parameter: σ

s, estimate of σ can be calculated as $t_{50\%} - t_{16\%}$

Weibull distribution summary:

Probability density function $f(t) = \frac{\beta \cdot (t-\gamma)^{\beta-1}}{\alpha^{\beta}} \cdot e^{-\left(\frac{t-\gamma}{\sigma}\right)^{\beta}}$ Eq. IV-61

Cumulative density function: $F(t) = 1 - e^{-\left(\frac{t-\gamma}{\sigma}\right)^{\beta}}$ Eq. IV-62

Instantaneous Failure Rate, hazard rate: $\lambda(t) = \frac{\beta \cdot (t-\gamma)^{\beta-1}}{\alpha^{\beta}}$ Eq. IV-63

Significant properties

Location parameter (63% failure): $t_{63\%} = \alpha$

Shape parameter: β

Time-delay parameter = γ, a factor used only when the data do not fit the distribution, except with the use of a time delay.

Lognormal distribution summary:

Probability density function: $$f(t) = \frac{1}{\sigma \cdot t \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2}$$ Eq. IV-64

Cumulative density function: $$F(t) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_0^t \frac{1}{x} \cdot e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} \cdot dx$$ Eq. IV-65

Instantaneous Failure Rate, hazard rate: $$\lambda(t) = \frac{\frac{1}{\sigma \cdot t \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma}\right)^2}}{1 - \frac{1}{\sigma \cdot \sqrt{2\pi}} \int_0^{t1} \frac{1}{x} \cdot e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} \cdot dx}$$ Eq. IV-66

Significant properties

Median (50% failure): $t = t_{50\%} = e^{\mu}$

Mean (average): $t = e^{\mu + \frac{\sigma^2}{2}}$

Location parameter: e^{μ}

Shape parameter: σ

s, estimate of σ can be calculated as $\ln \frac{t_{50\%}}{t_{16\%}}$

Their schematic representation behaviors are given in Figure IV.4.

When considering devices or system life, three failure types can be distinguished and are commonly represented by the bathtub curve (see Fig. IV.7).

Failures that appear during the early period of component life and are called early (infantile) failures. They can be explained through a faulty manufacture and an insufficient quality control in the production. They can be eliminated by a systematic screening test. But some component weaknesses can pass through this filter for several reasons either because the screening stress test conditions are not high enough effective to screen the latent defects or because the failure criteria set is defined for eliminating catastrophic failures and some level of drift failures.

Random failures, the second category, can't be eliminated neither by a screening test, nor by an optimal use politics (maintenance). They can be provoked by sudden voltage increases that can strongly influence the component quality and reliability. These failures appear erratically, accidentally, unforeseeably. In paper written in 2019, by Ed Sperling, the editor in chief of Semiconductor Engineering [65], presented automakers' interviews and as said by Gert Jørgensen, vice president of marketing at Delta Microelectronics "You can't test for the random failures random failures and latent defects add their own challenges. Sperling argue "the problem is worse in automotive due to both the expected lifespan of systems and the harsh environmental conditions. True random failures are rare. A stray alpha particle hitting a circuit and causing damage is known to happen, and the chances of that occurring increase with denser circuits and thinner insulation. So a single-event upset affecting 7nm device with FinFETs packed tightly together is more likely than at 28nm. The same is true for random contaminants, which may affect one part differently than another. But delineating which failures are truly random from those that are not is time-consuming, and that adds to the cost and slows down time to market."

Wearout failures, the third category, constitute an indicator of the component ageing.

Two methods are generally used to make reliability estimates: (i) parts counts method and (ii) parts stress analysis method. The parts counts method requires less information, generally that dealing with the quantity of different part types, quality level of the parts, and the operational environment.

Parts stress analysis requires the greatest amount of details and is applicable during the later design phases where actual hardware and circuits are being designed.

Whichever method is used, the objective is to obtain a reliability estimate that is expressed as a failure rate. Calculation of failure rate for an electronic assembly, unit or system requires knowledge on the failure rate of each part contained in the item of interest. If we assume that the item will fail when any (all in series) of its parts fail, the failure rate of the item will equal the sum of the failure rate of its parts.

Parts count reliability prediction [4], [66]: the information needed to use the method is: (i) generic part types (including complexity of microelectronics) and quantities; (ii) part quality levels; and (iii) equipment environment. The general expression for equipment failure rate with this method is for given environmental conditions:

$$\lambda = \sum_{i=1}^{n} N_i \cdot \left(\lambda_G \cdot \pi_Q\right)_i$$ Eq. IV-67

Where :

λ = total equipment failure rate (in FIT = 1 failure in 10^9 hrs)

λ_G = generic failure rate for the i^{th} generic part

π_Q = quality factor for the i^{th} generic part

N_i = quantity of i^{th} generic part

n = number of different generic parts.

In Parts stress analysis method, part failure models vary with different part types, but their general form is:

$$\lambda_i = \lambda_B \cdot \pi_E \cdot \pi_A \cdot \pi_Q \ldots \pi_N \qquad \text{Eq. IV-68}$$

Where:

λ_B = reference failure rate without environmental stress

π_E = environmental factor for accounting for other than temperature stressors (e.g. external stressors like vibration, humidity, etc

π_A = adjustment factors as for example biasing conditions (either static or dynamic or signal stressors, etc)

π_Q = quality factor referring to the quality control applied during manufacturing, screening, lot assurance testing before shipment to user.

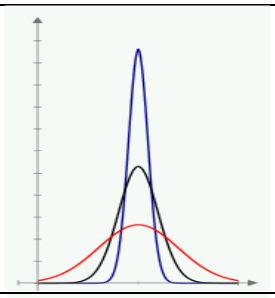
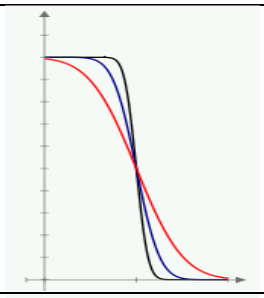
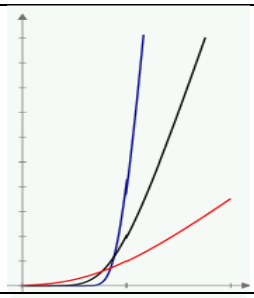
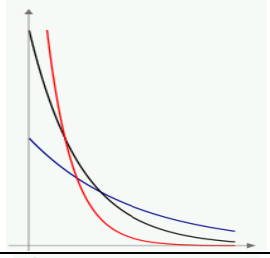
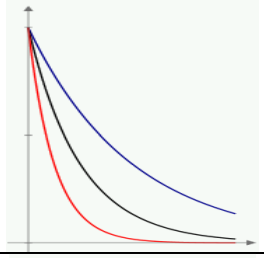
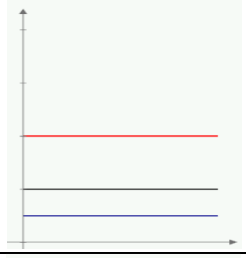
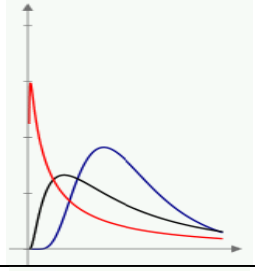
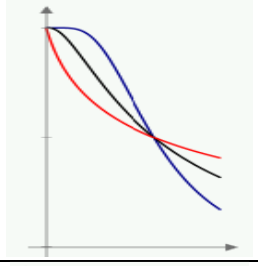
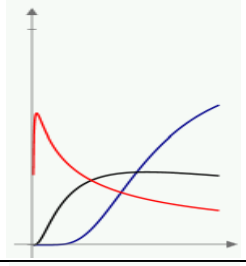
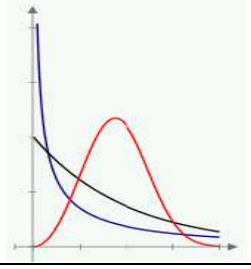
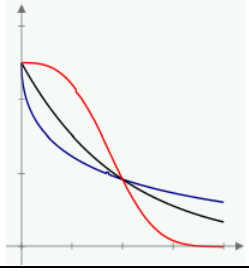
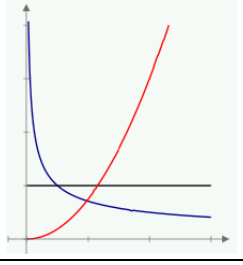
Life distributions	Probability density function *f(t)*	Reliability function *R(t)*	Hazard Failure Rate $\lambda(t)$
Normal			
Exponential			
Lognormal			
Weibull			

Figure IV.4: Life distribution functions represented schematically.

One can remark a series of multiplicative factors is assuming the effect of each stressor is counted independently. This appears to be somewhat of a contradiction because the effects are not really independent, so we express the failure rate $\lambda(x,y)$ from equation IV.1 to IV.3 as:

$$\lambda = \lambda_0 \cdot exp\left(\frac{E_a}{k \cdot T}\right) \cdot exp\left[-\frac{a_0 \cdot S_{Br1} \cdot x}{2 \cdot k \cdot T}\right] \cdot exp\left[-\frac{b_0 \cdot S_{Br2} \cdot y}{2 \cdot k \cdot T}\right] \cdot exp\left(\frac{-a_1}{2} \cdot S_{Br1} \cdot x\right) \cdot exp\left(\frac{-b_1}{2} \cdot S_{Br2} \cdot y\right) \quad \text{Eq. IV-69}$$

with $$\lambda_0 = \frac{1}{k_0} = \left(1 - exp\left(-\frac{\Delta S^\circ}{k}\right)\right)^{-1} \quad \text{Eq. IV-70}$$

Comparing equation V.4 and V.9, we observe the π factors can be defined as:

a) the listed in [4] or [66]

$$\pi_{T,T_{REF}} = exp\left[\frac{E_a}{k}\left(\frac{1}{T_{REF}} - \frac{1}{T_{op}}\right)\right] \quad \text{Eq. IV-71}$$

$$\pi_{S_{1_{op}},S_{1_{REF}}} = exp\left[\frac{a_1}{2} \cdot S_{Br1} \cdot \left(x_{op} - x_{REF}\right)\right] \quad \text{Eq. IV-72}$$

$$\pi_{S_{2_{op}},S_{1_{REF}}} = exp\left[\frac{b_1}{2} \cdot S_{Br2} \cdot \left(y_{op} - y_{REF}\right)\right] \quad \text{Eq. IV-73}$$

b) the one's not considered in [4] or [66]

$$\pi_{S_{1_{op}},S_{REF},T_{op},T_{REF}} = exp\left[\frac{a_0 \cdot S_{Br1}}{2 \cdot k}\left(\frac{x_{op}}{T_{op}} - \frac{x_{REF}}{T_{REF}}\right)\right] \quad \text{Eq. IV-74}$$

$$\pi_{S_{2_{op}},S_{REF},T_{op},T_{REF}} = exp\left[\frac{b_0 \cdot S_{Br2}}{2 \cdot k}\left(\frac{y_{op}}{T_{op}} - \frac{y_{REF}}{T_{REF}}\right)\right] \quad \text{Eq. II-75}$$

From this, we observe the equivalence with actual Standards is uncomplete because of missing terms, those related to external stressors like environmental (vibration, mechanical, radiation, humidity, etc).

E. Sedyakin principle

It would be very useful if one could somehow convert a rather complicated dynamical stress pulse, over some time interval t_1–t_2, into a rectangular pulse effective stress which would produce an equivalent amount of material/device degradation over this same time interval t_1–t_2.

Sedyakin principle

Assuming an Accelerated Failure Time Model (AFT) (e.g. only the time scale is affected under stress), in both cases, accelerating factors (AF) can apply based on equ. III.19. Quality Standards define also the similar rate function models and related AF. In 1966, Sedyakin [19] formulated his famous physical principle in reliability which states that for two identical populations of units functioning under different stresses S_1 and S_2, two moment's t_1 and t_2 are equivalent if the probabilities of survival until these moments are equal. The Figure IV.3 show how this principle also relates to the resulting reliability function $R(t)$ for a population submitted to a step stress conditions (these can be generalized to a superposition of various stressors as temperature, voltage, etc).

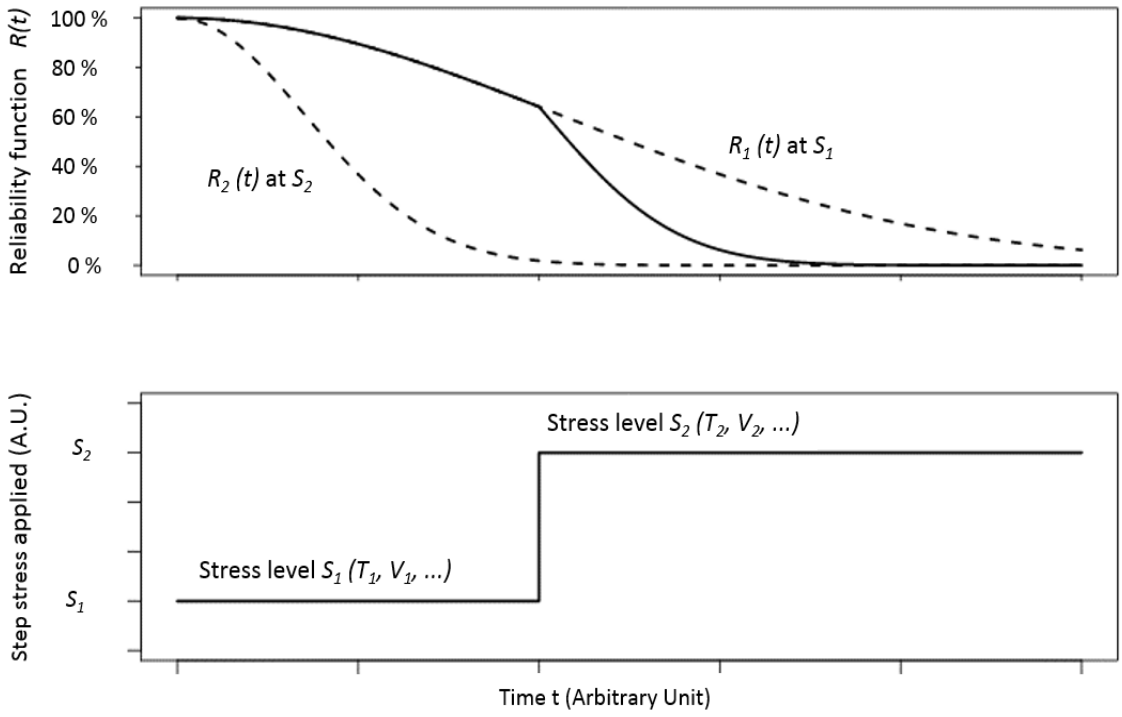


Figure IV.3: Reliability function $R(t)$ as a function of step stress applied and Sedyakin principle.

As presented by J. McPherson in ref [23] in his book chapter 14, and invoking potentially the Sedyakin principle, the time-to-failure models developed assume that the stress remains constant with time until the product fails. In semiconductor products (transistors, integrated circuits or any active electronic device) the applied signal induce voltage, current and power stresses changing during operation and are generally frequency dependent. So it is convenient to convert dynamical (time-dependent) stress to an equivalent effective static stress. So basic mathematics to determine the effective stress $\xi_{effective}$ such that it produce an equivalent amount of degradation and thus the same time-to-failure as the dynamical stress $\xi(t)$ is developed in ref [23] using the following concept of **compliance equation** for periodic (period P) dynamical stress and equivalent to power-law TF models :

$$\frac{1}{P} \cdot \int_0^P AF_{\xi(t),\xi_{effective}} \cdot dt = 1 \qquad \text{Eq. II-76}$$

Where the accelerating factor for the power-law TF model is for a yield stress $\xi_{yield} < \xi(t)$:

$$AF_{\xi(t),\xi_{effective}} = \left(\frac{\xi(t) - \xi_{yield}}{\xi_{effective} - \xi_{yield}}\right)^n$$

Eq. II-77

Gives:

$$\xi_{effective} - \xi_{yield} = \left[\frac{1}{P} \cdot \int_0^P \left(\xi(t) - \xi_{yield}\right)^n \cdot dt\right]^{\frac{1}{n}}$$

Eq. II-78

It is interesting to consider also how to calculate the effective static temperature equivalent proposed by J. McPherson. He mentioned that similar to stress, the temperature T of a device is not constant but depend on biasing on-off operation, pulse operation, thermal cycling, as well as on dynamic signal applied.

To assess the effective static temperature $T_{effective}$ which produces an equivalent amount of device degradation versus the temperature variation $T(t)$. The compliance equation to determine $T_{effective}$ over a time interval of ($t_a - t_b$)

$$\frac{1}{(t_a - t_b)} \cdot \int_0^P AF_{T(t),T_{effective}} \cdot dt = 1$$

Eq. II-79

III. SYSTEM RELIABILITY

A. *Series systems*

How to estimate the reliability distribution parameters and failure probabilities for complex DSM components operating at harsh condition as applied during their mission profiles?

The newest component technologies deploy transistors having very small gate lengths (less than 10 nm) with new materials, reduced dimensions, and new technologies. Once the emerging technologies for numerical applications under high environmental constraints for true and real-time mission profiles are validated, the issue is not yet completed.

In Part II, we have identified CMOS Bulk reliability distribution parameters accounting electrical aging mechanisms, such as high-temperature degradation so-called Negative or Positive Thermal Instability (NBTI, PBTI), or Hot Carriers Injection (HCI) and the hard or soft breakdown of dielectric gates. These mechanisms can cause significant reductions in performance and software errors that are very important to quantify under nominal operating conditions. We need to determine their impact on the lifetime of the single sensitive active area of a device and then to look on how it can degrade the performance of the design. It is important to note that most of these mechanisms do not directly lead to a straightforward failure but only a gradual parametric drift of the elementary functions which will only induce a failure when the parameters exceed a critical threshold strongly dependent on the architecture. For the dominant mechanisms in recent technologies, this failure rate increases over time, a characteristic property of Wearout mechanisms. Moreover, the reliability of the current circuits often leads to not observe any failure associated with these mechanisms for the duration of the tests (a few months), which does not make it possible to evaluate the probability of failure beyond the equivalent duration of the test, this probability being likely to increase strongly then because of the increasing nature of the failure rate.

How can these probabilities be used to predict total component DSM system performance? How does the DSM design affect reliability? How redundancy design in a DSM affect the total reliability?

What are the hypotheses to consider a system reliability model like a DSM?

How component failure rate characteristics are described: multiple failure mechanisms superimposed dependent or independent, constant (cataleptic failures) vs non-constant. What if a failure results in an open path or in a short circuit path and what could be such an impact for series or parallel system modelling?

This chapter will derive the formula for the reliability of a series system, parallel system or a mix series/parallel system as fully detailed in ref [67] in Applied Reliability Third Ed. book from P. Tobias and D. Trindade.

The most commonly used model for system reliability assumes that the system is made up of n independent components which all must operate in order for the system to function properly. But this series model is applied specifically when a single integrated circuit with several independent failure modes is analogous to a system with several independent elementary constituents. The failure mechanisms are competing with each other in the sense that the first to reach a failure states causes the component to fail: the open question is still to consider what a failure state is? Is it for catastrophic failure or related to a failure criteria or performance? As argued by P. Tobias and D. Trindade, "*the more general competing risk model where the failure processes for different mechanisms are not independent can be very complicated since one must know how the random times of failure for different mechanisms are correlated*".

Assuming the first hypothesis, and the i^{th} element have a cumulative distribution function (CDF) $F_i(t)$, the probability the serie system fails at time t is the probability that one or more of the n independent components have failed at time t. In term of total probability, we consider the product of the n CDFs:

$$R_s\,(t) = Pr(E) = \Pr(E_1, E_2, \dots, E_n) = \Pr(E_1) \cdot \Pr\,(E_2) \cdot \dots \cdot \Pr\,(E_n) = \prod_{i=1}^{n} R_i(t)$$

Eq. III-1

With

$$R_i\,(t) = e^{-\lambda_i \cdot t}$$

Eq. III-2

where λ_i is the failure rate of i^{th} component. Hence,

$$\lambda_s\,(t) = \sum_{i=1}^{n} \lambda_i$$

Eq. III-3

If the components have the same reliability R_{s0}, meaning $\lambda_i = \lambda_0$ for i = 1 to n, this becomes:

$$R_s(t) = R_{s0}^n = \left(e^{-\lambda_0 \cdot t}\right)^n = e^{-n \cdot \lambda_0 \cdot t}$$

Eq. III-4

LEMMA: A series system constituted of *n* identical and independent elements, each described by a POISSON distribution (λ_0) reliability model can be approximated by a general equivalent POISSON distribution with parameter $n \cdot \lambda_0$.

B. Parallel systems

A system that operate with n elements in parallel survives until the last of its element fails considering in this case a failure mode definition to be an open circuit. We will not consider the second hypothesis case if the failure mode of each element is a short circuit. In such a case the system with n elements in parallel survives if only one element fails.

Assuming the first hypothesis, and the i[th] element have a cumulative distribution function (CDF) $F_i(t)$, the probability the parallel system fails at time t is the probability that all the components have failed at time t. In term of total probability we consider the product of the n CDFs:

$$F_p\,(t) = \prod_{i=1}^{n} F_i(t)$$

Eq. III-5

$$R_p\,(t) = 1 - \prod_{i=1}^{n} [1 - R_i(t)]$$

Eq. III-6

Poisson distribution function:

If the CDF is defined by Poisson or exponential distribution function, the probability the system fails, e.g. one at least of the n elements fails, is the reliability or probability of success at time *t*: as expressed by equation VII.3

$$R_p\,(t) = 1 - \prod_{i=1}^{n} \left(1 - e^{-\lambda_i \cdot t}\right)$$

Eq. III-7

If $\lambda_i = \lambda$ for any i = 1 to n

$$R_p\,(t) = 1 - \left(1 - e^{-\lambda \cdot t}\right)^n$$

Eq. III-8

Considering the life mission t is defined by the plateau of the bathtub, for $\lambda \cdot t << 1$, and we can approximate the reliability at time *t* by:

$$e^{-\lambda \cdot t} \approx 1 - \lambda \cdot t \qquad \text{or} \qquad \lambda \cdot t \approx 1 - e^{-\lambda \cdot t}$$

Eq. III-9

and merging equ. VII.8 with VII.9 gives:

$$R_p(t) \approx 1 - (\lambda \cdot t)^N \approx e^{-(\lambda \cdot t)^n} = e^{-\left(t/\theta\right)^\beta}$$

Eq. III-10

The error generated by the approximation of equation VII.7 is assessed and plotted figure 4 showing equations VII.8 and VII.10 versus parameter $\lambda \cdot t$ in unit of MTTF. The error generated by the approximation is less than 1% for $\lambda \cdot t \leq 0.3$ or in other word at time *t* lower or equal to 30% of MTTF (for λ = 100 FIT and n=3, MTTF=1/λ = 10^7 hrs). This error is decreasing for higher values of n.

As a major fact, such parallel system on *n* independent and identical elements having the same constant Poisson failure rate λ (e.g. same equivalent $MTTF_i = MTTF = 1/\lambda$ for i=1 to n) show an equivalent Weibull distribution for short times t $\leq$ 30%·MTTF range for a error lower than 1%.

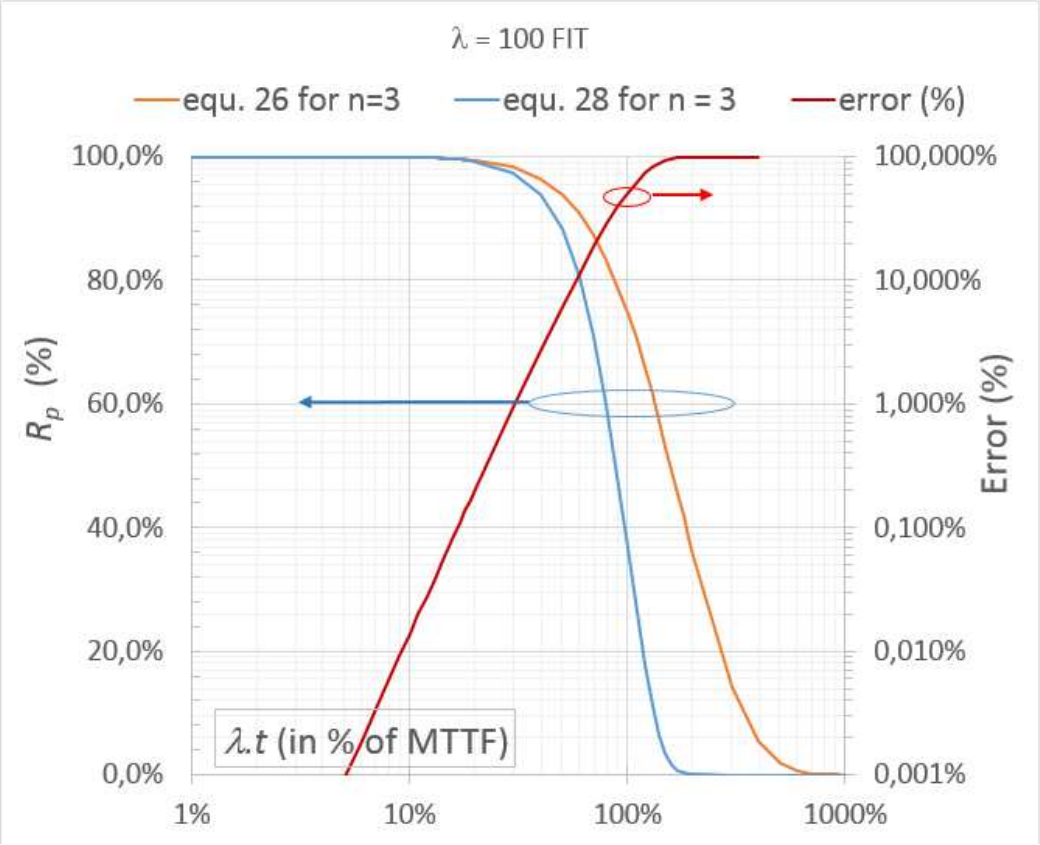


Figure V-6: Plot error (in red) as the difference between plot of equation VII.8 and VII.10 for λ=100FIT and n=3.

LEMMA: A parallel system constituted of n identical and independent elements, each described by a POISSON distribution (λ) reliability model can be approximated by a general equivalent WEIBULL distribution with parameters λ and $\beta = n$ with an error lower than 1% for time operation lower than 30%·MTTF for the example shown.

Weibull distribution function:

The random variable T has a Weibull distribution with parameters α_i and β_i if the reliability or probability of success at time t is given by:

$$R_{i_W}(t, \alpha_i, \beta_i) = e^{-\left(\frac{t}{\alpha_i}\right)^{\beta_i}} \quad , t \geq 0 \qquad \text{Eq. III-11}$$

$$R_{p_W}(t, \alpha_i, \beta_i) = 1 - \prod_{i=1}^{n}\left(1 - e^{-(\beta_i \cdot t^{\mu_i})}\right) \qquad \text{Eq. III-12}$$

For identical independent elements we set $\alpha_i = \alpha$ and $\beta_i = \beta$ for any i = 1 to n, so:

$$R_{p_W}(t, \alpha, \beta) = 1 - \left[1 - e^{-\left(\frac{t}{\alpha_i}\right)^{\beta_i}}\right]^n \qquad \text{Eq. III-13}$$

Considering the life mission t is defined by the plateau of the bathtub, for $\lambda \cdot t << 1$, and we can get for a parallel system and short time t:

$$R_{p_W}(t, \alpha, \beta) \approx e^{-\left(t/\alpha\right)^{\beta \cdot n}} \qquad \text{Eq. III-14}$$

The error generated by the approximation of equation VII.14 is assessed and plotted figure 5 showing equations VII.13 and VII.14 versus parameter $\lambda \cdot t$. The error generated by the approximation is less than 1% for $\lambda \cdot t \leq 0.3$ or in other word at time t lower or equal to 80% of MTTF (for example λ = 100 FIT and n=5). This error is decreasing for higher values of n and β.

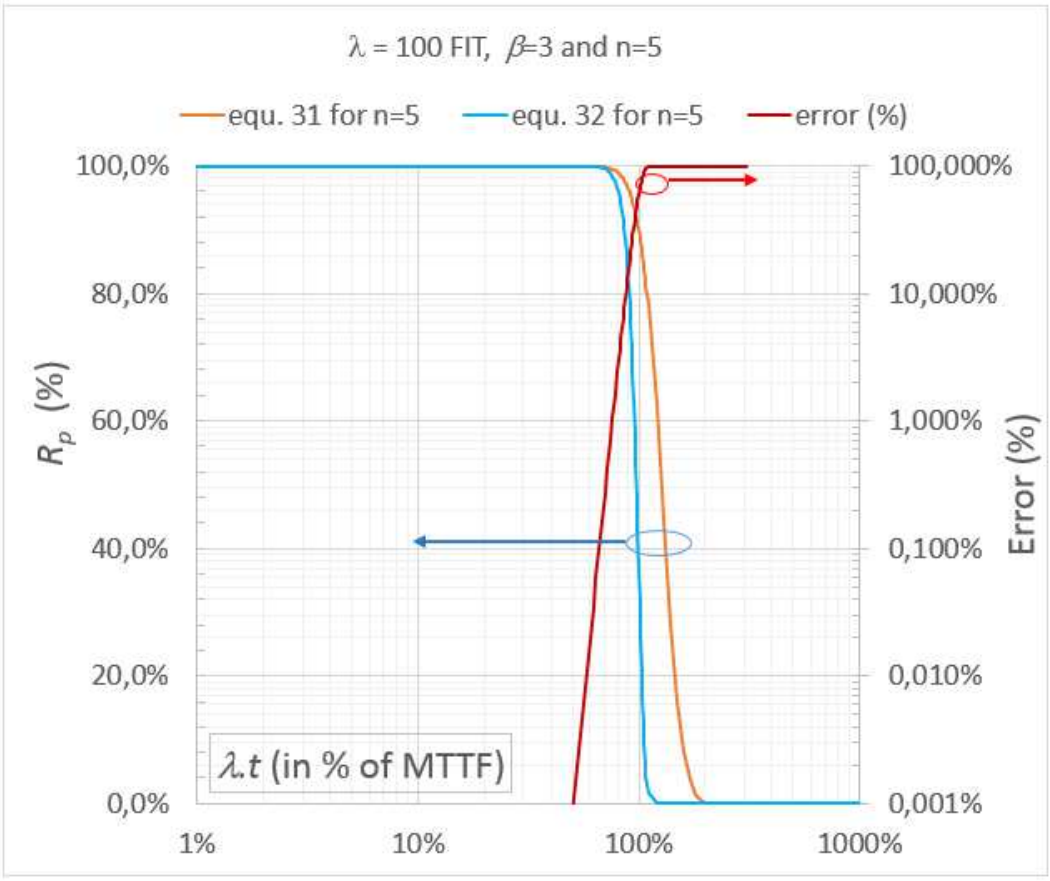


Figure V-7: Plot error (in red) as the difference between equation VII.13 and VII.14 for λ=100FIT, β=3 and

LEMMA: A parallel system constituted of *n* identical and independent elements, each described by a reliability WEIBULL distribution (α, β) model can be approximated by a general equivalent WEIBULL distribution with parameter (α, $\beta \cdot$ n) with an error lower than 1% for time operation lower than 75%·MTTF.

As a major fact and a generalisation point of view, such parallel system based on *n* independent and identical elements having the same constant Weibull failure rate parameters (α, β) is shown to be modelled by an equivalent Weibull distributions with reliability parameter (α, $\beta \cdot n$). The error induced by such approximation is estimated for short times $t \leq 80\% \cdot MTTF$ and range for an error lower than 1% for the example shown for λ=100FIT, β=5 and *n*=3.

C. Complex systems

Physical configurations in series or parallel do not necessarily indicate the same logic relations in terms of reliability. An integrated circuit (DSM) composed of LUTs connected in logical paths contains billions of elementary structures inter-connected. From a reliability perspective the LUT includes multiple transistors in series-parallel configurations but a LUT is said to have failed if one or more transistor failed. So, the transistors in a LUT are considered in series from a reliability perspective.

There are systems that require more than one component to succeed in order for the entire system to operate. In addition, the performance of a product is usually measured by multiple characteristics. In many applications, there is one critical characteristic, which describes the dominant degradation process. This one can be used to characterize product reliability. The failure of a product can be defined in terms of performance characteristic or called indicator crossing a specific threshold. Figure 6, depicts the relation of degradation path, pseudo-life, and life distribution. The maximum likelihood method should be used for estimation of distribution parameters. Once the estimates of reliability parameters we can use a Monte Carlo simulation to generate a large number of degradation paths. The probability of failure *F(t)* is approximated by the percentage of simulated degradation paths crossing a specified threshold after a stress time applied.

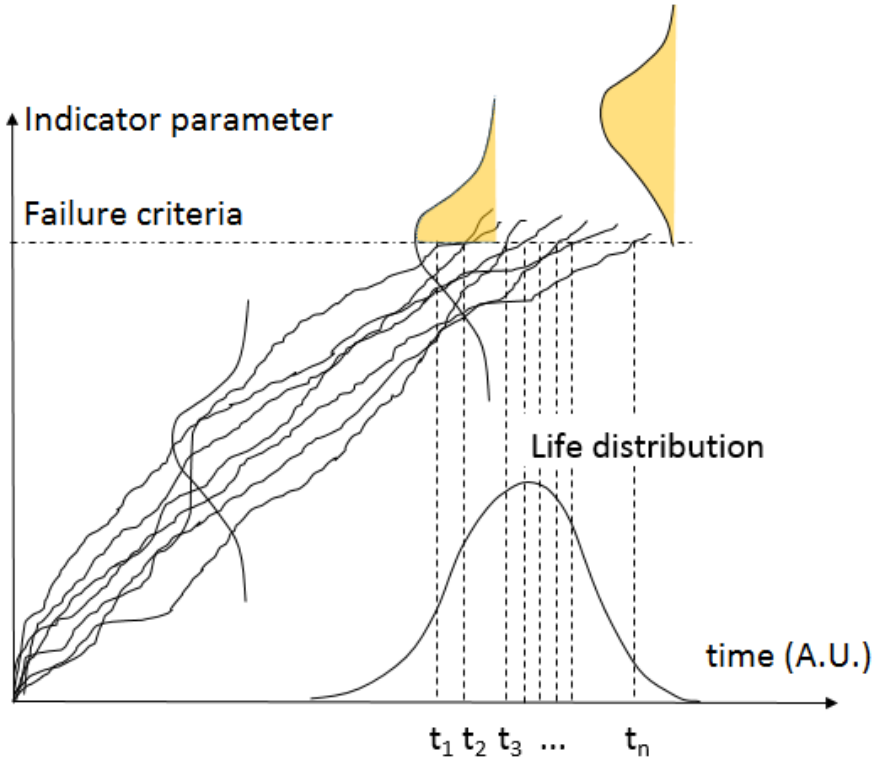


Figure V-8: Relation of degradation path, pseudo-life, and life distributions.

When considering catastrophic failure paradigm, a general series system can be observed in a configuration for a DSM FPGA test structure based on ring oscillator (RINGO) test structure represented in Figure V-9.

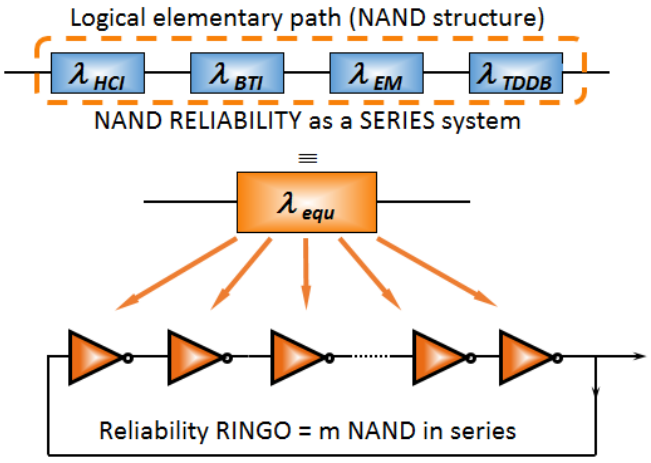


Figure V-9: Series reliability configuration for DSM RINGO catastrophic failure mode modelling. (4 failure mechanisms and m= RINGO size defined as number of NAND interconnected).

Nevertheless, if we consider parametric failure mechanisms hypothesis, another representation should be proposed as shown in **Erreur ! Source du renvoi introuvable.**. In such a case a single component with 4 independent parametric failure mechanisms (HCI, BIT, EM or TDDB) and occurring simultaneously, is analogous to a series system. Each failure mechanism 'competes' with the others to cause a failure and failure rates are additive, mechanism by mechanism, and so drifts are cumulated. In a RINGO structure, a failure is observed when all NAND fail in a cumulated drift failure reaching the failure criteria. In such a case, the reliability model is similar to a parallel system because by definition only the failure of all components within the system results in the failure of the entire system. In other words, we know a parallel system succeeds if one or more components are operational. Considering a parametric drift failure, the NANDs work together pending to reach the cumulated parameter drift limit and because the contribution of each failure mechanism is accumulated.

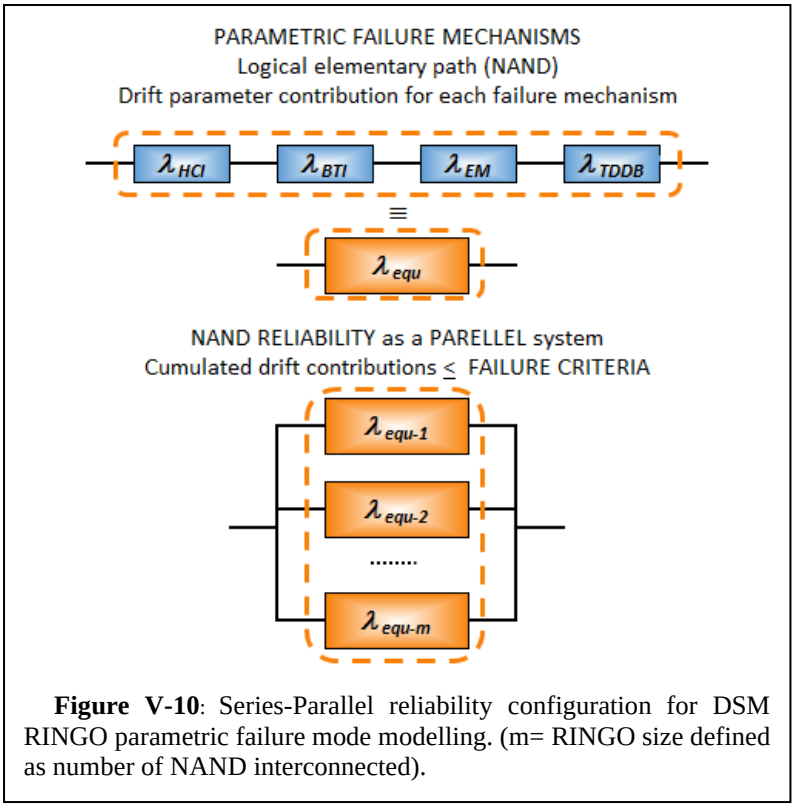


Figure V-10: Series-Parallel reliability configuration for DSM RINGO parametric failure mode modelling. (m= RINGO size defined as number of NAND interconnected).

A reliability of a series-parallel system defined **Erreur ! Source du renvoi introuvable.** and for *a **POISSON distribution function,*** is written as:

$$R_{p-s_P}\left(t,\alpha_{i,j},\beta_{i,j}\right) = 1 - \prod_{j=1}^{m}\left(1 - \prod_{i=1}^{4} e^{-\lambda_i \cdot t}\right) \qquad \text{(Eq. 15)}$$

Considering *m* NAND elements in parallel each with 4 types of failure mechanisms in series, we get

$$R_{p-s_P}\left(t\right) = 1 - \prod_{j=1}^{m}\left(1 - e^{-t\cdot\sum_{i=1}^{4}\lambda_i}\right) \qquad \text{(Eq. 16)}$$

Because all NAND elements are similar, we can write:

$$R_{p-s_P}\left(t\right) = 1 - \left[1 - e^{-t\cdot\sum_{i=1}^{4}\lambda_i}\right]^{m} \qquad \text{(Eq. 17)}$$

If $\sum_{i=1}^{4}\lambda_i \cdot t << 1$, e.g. assuming true up to few percent's of MTTF, we can approximate the reliability at time *t* by:

$$\left(1 - e^{-\sum_{i=1}^{4}\lambda_i \cdot t}\right)^{m} \approx \left(t\cdot\sum_{i=1}^{4}\lambda_i\right)^{m} \qquad \text{(Eq. 18)}$$

And (Eq. **17** reduce to:

$$R_{p-s_P}\left(t\right) \approx 1 - \left(t\cdot\sum_{i=1}^{4}\lambda_i\right)^{m} \qquad \text{(Eq. 19)}$$

Allowing us to express:

$$R_{p-s_P}\left(t\right) \approx 1 - \left(t\cdot\sum_{i=1}^{4}\lambda_i\right)^{m} \approx e^{-\left(t\cdot\sum_{i=1}^{4}\lambda_i\right)^{m}} = e^{-\left(t/\theta\right)^{\beta}} \qquad \text{(Eq. 20)}$$

as a WEIBULL general equivalent representation where

$$\frac{1}{\theta} = \sum_{i=1}^{4}\lambda_i \quad \text{and} \quad \beta = \mathrm{m} \qquad \text{(Eq. 21)}$$

which is indeed the number of NANDs incorporated in a RINGO test structure.

LEMMA #04: In a series-parallel system constituted of *n* identical and independent elements, each described by a reliability POISSON distribution (λ), its reliability model can be approximated by a general equivalent WEIBULL distribution with parameter (θ, β= m) with an error lower than 3% for time operation lower than 50%·MTTF (see **Erreur ! Source du renvoi introuvable.**).

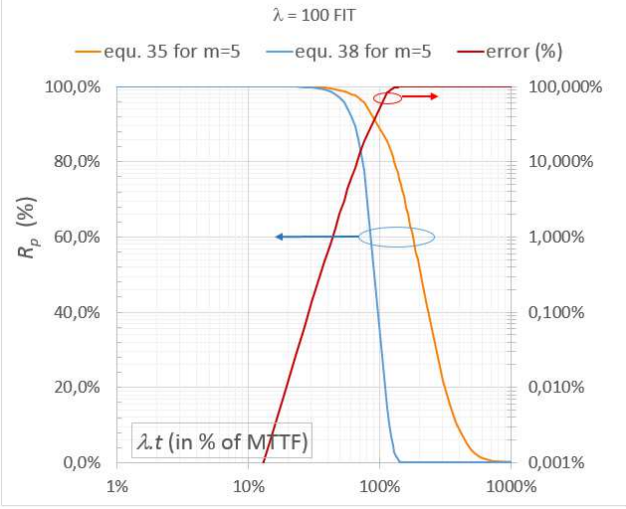


Figure V-11: Plot error (in red) as the difference between plot of Eq. 35 and Eq.38 for a POISSON Distribution with λ=100 FIT, β=m=5.

IV. Conclusion and prospectives

Using emerging microelectronic COTS appear mandatory to achieve today's product and system performances with the drawback of lack of inadequate Quality Standards which don't address in a proper manner the way to qualify them in a proper short period of time. Indeed, in the last end of XXth century, the introduction of a new component in a Space system required few years of evaluation and complete qualification. Such timing was compatible with the technology evolution and the satellite development and manufacturing phases. Today, with the explosion of satellite constellation and their short development and production phase in less than 2 to 3 years, such Standards are not adapted in some case.

To balance this inconvenience, we need to better understand the physics of semiconductors and failure mechanisms in order to quickly identify and select "reliable COTS" technology and lot screening before to authorize their implementation in complex systems. By "reliable COTS" we consider we have enough information consolidated by short time adapted experiments to identify how we can guarantee mission profile as existing for Low Earth Orbit satellite (LEO) constellation or long term Geostationary Earth Orbit telecommunication satellite. Same kind of considerations must be fixed based on skill and knowledge of Physics of devices to draw optimized experiments and deep analyses related to the Automotive or Aeronautic industries.

A reliability engineer team needs to gather multi-skills people having valuable knowledge on (not exhaustive):

- electronic design, to be able to understand how the devices are biased (DC and digital or analog signals) and stressed from normal to extreme mission profile conditions,
- semiconductor and assembly process manufacturing, characterization and metrology technics, to understand how devices are packaged and tested with the aim of determine and validate their Safe Operating Area limits and to define sequence of screening for efficient freak distributions purging, and their ability to survive operating life environmental and application conditions,
- physics of semiconductor, physics of new microelectronic device, induced failure mode and mechanisms to understand how devices are designed and manufactured so as to determine what are the key electrical stressors and indicators to monitor degradation patterns influenced by internal and external stress conditions (respectively biasing and environmental),
- reliability concepts and modelling, to identify and consolidate hypotheses of degradation based on series of experiments (as short as possible) representative of true mission conditions.

In the paradigm of Physics of Failure, people focus on how a product is failing either catastrophically or gradually.

In the Physics of Healthy we propose to focus on the characterization and physical reliability knowledge of lot devices able to assure the system mission in operation before failure. The goal is to provide advices to the equipment engineer on associated quality and reliability risks.

The very short development of new technologies into the custom market seized the microelectronic COTS products, and consequently induce a change in the technique high reliability application are designed.

This Part I is dedicated to understand how PoH concept can be defined. To do so, we first focused to describe two kind of emerging technologies DSM (very narrow node size, e.g. < 10 nm and GaN Power DC switch transistor in a short State of The Art review. Letting aside the random failure of device, system or equipment breakdowns generally due to handling, packaging or external overstress, we concentrate on wearout mechanisms description taking advantage of J. McPherson's book on Reliability Physics and Engineering implementing the basics of reliability modelling [23], we presented and detailed the mathematics based on Gibbs Free Energy diagram considering reaction with several external stresses applied and temperature. The multi-failure mechanisms are also considered to elaborate a reliability model and to establish accelerating factor multi stress and multi-mechanism expression based on a non-constant activation energy.

Reliability models were also recalled with respect to random failure rate modelling and system reliability are presented from formula of series systems, parallel systems or a mix series/parallel system as fully detailed in book from P. Tobias and D. Trindade.

Establishing and reminding these preliminary set of mathematics, we propose to develop the next Part II to show how these tools can be applied and used experimentally to predict pre-failure occurrence and elaborate predictive reliability SHM protocols. A set of sequence is described to help to diagnostics PoH capability on new COTS devices through a series of control steps also named Prognostics and Health Management Protocol. Such protocol can be easily applied both on new emerging or existing technologies.

Part II: Applied Engineering on Physics-of-Healthy and SHM of microelectronic equipment for aeronautic, space, automotive and transport operations.

Bensoussan Alain
Dr.-Ing., Dr.-HDR, IEEE senior member
Thales Alenia Space
Meta Competence Center
Toulouse, France
alain.bensoussan@thalesaleniaspace.com

Bernstein Joseph
Professor, IEEE senior member
Laboratory for Failure Analysis and Reliability of Electronic Systems
Ariel University
Ariel, Israel
josephbe@ariel.ac.il

Bravaix Alain
PhD, HDR, IEEE senior member
REER : Radiation Effects & Electrical Reliability, IM2NP, ISEN
Yncrea Mediterrannée, Toulon, France
alain.bravaix@isen.fr

"Experience without theory is blind, but theory without experience is mere intellectual play."

"L'expérience sans théorie est aveugle, mais la théorie sans expérience n'est qu'un simple jeu intellectuel."

Immanuel Kant

I. INTRODUCTION

In Reliability Standards the MTTF is related to random mechanisms which are supposed to reach sooner than the weraout MTTF. This is mainly because the considered technologies are rather mature and the applied mission stress conditions are low enough to push the limits of the wearout events.

Several end-user Industries and Institutions are very cautious to perform lifetest conditions as close as the nominal conditions. Automotive OEMs last year began demanding that electronics components last 18 years with zero failures.

Nevertheless, new technologies, and innovative components have shown possible shorter life-time occurrence [68] because of their immaturity. Especially when procured as COTS but used in long term mission application or sometimes in High Rel equipment, it is necessary to perform screening to eliminate freak distribution and run lot qualification customized programs adapted to the foreseen application. It is not easy and the simplest way to ask the component manufacturer to help quantifying these parameters. Sometime, manufacturers may provide information about the life-time and generally based on cumulative failures observed during burn-in and qualification test sequences (considered rather as Random defects).

The concept of stressor parameters is introduced to quantify how the stress will impact the equivalent activation energy described in Part I for the principle of BAZ model [25] and Mc Pherson [57] of the energy band diagram. Indeed the external stress (any kind: electrical, signal, or external as thermal cycling or radiation for example) is generally (but not always) reducing the height of the thermal activation energy measured from pure thermal stress conditions.

Physics of Healthy (PoH) concept: we define Stressor and Predictors Indicators electrical parameters.

- A Stressor parameter is a setting condition applied to the component of interest which stress the device and age it. Such parameter can be either a electrical nominal setting lifetest operation (outlined in percentage of the limit of the capability of the component defined as in the Maximum Rating conditions: thermal, DC bias, AC or RF signal) or an external environment stress condition (thermal chocs, thermal cycling, radiation, ionic contamination or gaz poisoning).
- A Predictor (or Indicator) an in-situ measurable and quantifiable electrical parameters to assess the healthiness of a component introduced in an equipment. Such a component is supposed to drive the reliability of the final card installed in an equipment. To do so, the card should be outfitted of dedicated electrical point of access or accessible through the bias access in order to be monitored during operation (see for example DSM ASIC based on 10 nm node described in the chapter). The aim is to quantify how the internal characteristics and performances are affected during time of stress applied.

This Part II presents context and particular issues for Automotive application using emerging technologies. We describe the methodology implemented by GIFAS consortium [69] (Groupement des Industries Françaises Aéronautiques et Spatiales) and FIDES group [66]. Predictive Reliability and Health monitoring methodology concept for immature technologies are deeply considered by GIFAS consortium. We show the 3-loop methodology and show schematics of the main steps involved in the Reliability determination of DDR3 (FF20) and FPGA (FF16) CMOS technology nodes. We take advantage of the Prognostic Failure Model (PFM) level 3 recommendations and develop how to apply reliability prediction on DSM technologies in harsh environment.

Failure Modes and Effects Analysis (FMEA) is a method for reliability analysis trying to assess and identify existing failure modes of a system. When applying Standards method at component level, the failure mode rate is considered as a constant. Furthermore, its value is an estimation depending of the database used or extracted from few experiments. The more realistic approach proposed by FIDES group is based on reliability prediction models but considering one mechanism involved time to time. They have shown that the failure mode rate is not a constant but depend on mission profile conditions. The FIDES tool developed is considering failure rate as stress dependent and can be expressed by COX models.

When multiple failure mechanisms co-exist the Standard and FIDES methodology must be revisited.

We present a case study on application for Automotive and Aerospace. It affords series of numerical application supported by experiments for study cases on a FinFET technology assuming HCI (Hot Carrier Injection) , BTI (Bias Temperature Instability) and EM (Electro-Migration) as major failure mechanisms. It is shown how the activation energy is related to the stress and temperature applied and proof activation energy can no-longer be considered as constant to extrapolate some experiment under high stress to nominal mission operation.

In a second paragraph, we give highlights (advantage and drawbacks compared to existing Standards (MTOL) and recommendations when implementing a true case study on DRAM IC DDR3. In particular how to consider occurrence of multiple failure mechanisms on a component used under operational real harsh environment profiles. For the Aerospace Electronics reliability, we relate a pragmatic application of MTOL based on J. Bernstein approach. The prediction of a system

reliability can be described using a linear matrix solution.

In the last paragraph, we develop the Prognostic Failure Model (PFM) background.

How to build a methodology to forecast the reliability of a device, board or system in harsh operating conditions?

What are the clues and the evidences to answer correctly and objectively if there is a risk or not to operate a new technology, in a repeatable, safe and precise analysis process?

How to automatize such a method in order to reduce uncertainty and avoid influence of many expert positions or even give to the single expert the right to come back to his analysis later in years and to improve it because of new evidence of data?

How to discriminate the concepts of constant vs non-constant rates when people are locking a pragmatic solution to assess the risk of failure during the entire mission life of an equipment, a product or a component?

A reliability risk analysis of the lifetime of an electronic device, versus the mission profile of the final product is defined as a dedicated Prognostic Failure Model (PFM) guideline. The method proposes four levels, depending on the data and service available from the manufacturer. Level 1 is the simplest to implement, and level 4 is the utmost complex.

II. Component Health Monitoring - A case study for Automotive and Aerospace application.

A. Context and particular issues for Automotive application using emerging technologies :

Several end-user Industries and Institutions are very cautious to perform lifetest conditions as close as the nominal conditions. Automotive OEMs last year began demanding that electronics components last 18 years with zero failures. If we consider a car manufacturer reporting about 7,000 semiconductor devices in its high-end models, and it turns out about 4,000 cars every day, then failure in the range of one part per million equals 24 defective cars per day. For another one is which uses about half as many electronic components but manufactures 10,000 cars every day, that equates to 54 defective cars.

"Normally it takes five to six years before you know if there is a problem," said Jim McLeish, senior quality/reliability consultant manager at DFR Solutions. "With 5nm to 7nm technology, there is no experience. We don't know what the variation will be." The use of 10/7nm chips in automotive applications is making test more challenging. There are a lot more transistors in each area. That requires higher test coverage, and a lot of testers are in the beginning phase of bringing technology up to an acceptable level.

It is not always clear what is a killer defect and what is not. Some latent defects may never cause failures, while others that are less obvious can develop into more serious issues under various environmental conditions or excessive electrical surge current. As shown in Figure IV.4.a below, this is the example of an MMIC MESFET technology in mid-1990 [70] where micro-particles were generated during integration of MIM capacitor manufacturing. Based upon inputs from customers, the failure rate was estimated to be about 200 defectives per million devices. At that rate, we would have expected a 50% chance of seeing three capacitor failures in the 15,000 devices that were life tested between 1985 and 1995 – but we did not experience any (probability was not helping us). Over the course of two years, the particle density had been reduced by the MMIC manufacturer by almost two orders of magnitude (see figure IV.4.b) – still without generating a single accelerated reliability test failure! The size of these micro-defects ranges below 1 µm diameter but considering the dielectric thickness of 0.2 µm as a conformal deposition process, some of them generated early failure short circuits while a small portion remains as latent defects. Safe operating conditions were demonstrated and validated because we conducted 100% screening applying maximum voltage (destroying killer defects) prior to release devices for a high-reliability application.

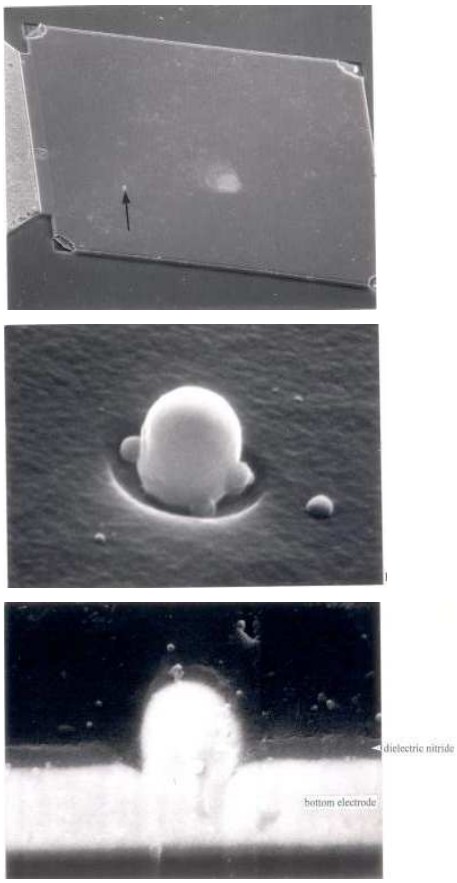


Figure IV-4.a: Capacitor with a microparticle (1 µm diameter) on bottom electrode (Courtesy Alcatel Espace: Serma Technologies Constructional Analysis report **[130]**).

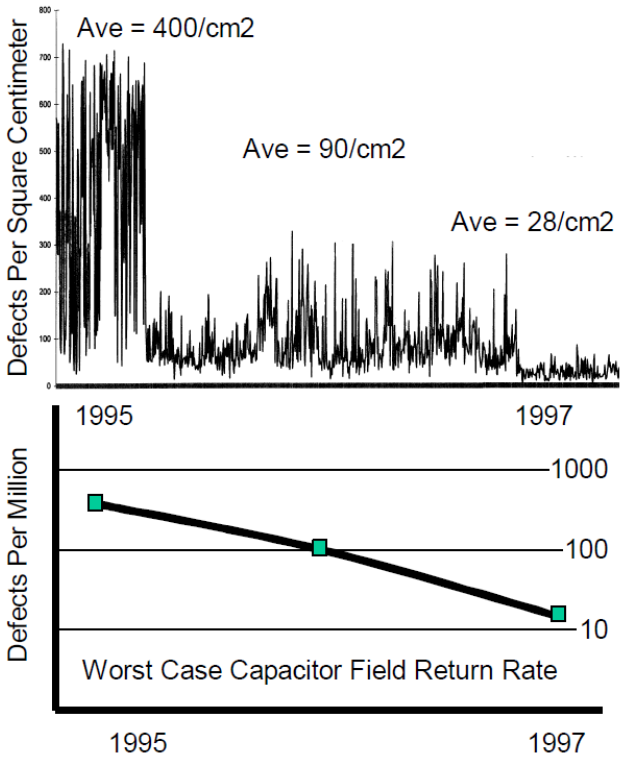


Figure IV-4.b: Defects and Failure Rates of Capacitors **[70]**

Current predictive reliability model tools such as MIL, 217+ or FIDES are based on Arrhenius' law which expresses the lifetime to failure (Time-To-Failure or TTF) as a function of an invariable activation energy and temperature. It considers only one failure mechanism and temperature as a major stress. These tools are limited to predicting the constant failure rate.

The TTF assessment depends if we consider failure mechanisms either related to random failure or to wearout failure mecanisms. For emerging DSM technology based on nanometric scale the TTF calculation strongly depends on such overlapping properties between random and wearout mechanisms as we will see in the next paragraph B developed. In addition to this confusion, the activation energy is not a true constant in the range of temperature and mission biasing condition applied as we will illustrate in paragraph C.

A predictive reliability paradigm should consider these limitations and propose how to handle:

The multi-failure mechanisms which can occur simultaneously either at random or at end of life

- Assign correct activation energy values to random and wearout failure mechanisms
- Express activation energy in term of temperature and operating biasing condition ranges
- Deduce extrapolation of TTF under validated hypotheses (be careful to design accelerating stress conditions in adequation with true operating mission profile (e.g. mitigate and validate the representativeness of failure mechanisms activated under high stress).

During environmental tests, we need to demonstrate that the probability of failure of a part is less than a given value, which leads showing with n tests, that the mean of the resistance (particular quantile) has (at least) a given value. With n tests, the average should be demonstrated to be within a certain confidence interval with a given level of confidence.

B. Predictive Reliability and Health monitoring methodology concept for immature technologies :

The lifetime of a component and thus of an electronic card, is quantified using a constant random failure rate (λ). This law of probability applies to rare events or a situation where the manifestation of an event occurs randomly over time with a low probability. The reliability estimation problem turns out to be different if we consider a product subjects to a degradation process that can be quantified over time and on which a "criterion of acceptability" can be set. It is then possible to track the evolution of this drift during stress tests, in order to evaluate the forecast reliability of the product, or during its operational life, to know its state of health from day to day, and to estimate its remaining potential.

In Reliability Standards the MTTF is related to random mechanisms which are supposed to reach sooner than the weraout MTTF. This is mainly because the considered technologies are rather mature and the applied mission stress conditions are low enough to push the limits of the wearout events.

Nevertheless, new technologies, and innovative components (by definition less mature), have shown possible shorter life-time occurrence [68].

Consequently, in parallel to the constant failure rate evaluation, it must be verified that the wearout of a component will occur far after the lifetime determined by random failure law. So, the concern to predict reliability and implement health monitoring on emergent technologies need to answer this tricky issue discrimination between random and wearout failure mechanisms. It is not easy and the simplest way to ask the component manufacturer to help quantifying these parameters. Sometime, manufacturers may provide information about the life-time and generally based on cumulative failures observed during burn-in and qualification test sequences (considered rather as Random defects).

The definition of MTTF (either Random or Wearout) is becoming a major issue and must be well discriminated as each of them assess its own value. Several major hypotheses must be described and considered:

- Failure mechanisms are based on Physics of Failure (PoF) models in terms of failure rate (FR) and Accelerating Factors (AF) compared to a reference set of stress conditions.
- The Standard Model of Accelerated Life applies which assumes that only the scale parameter of the reliability distribution is modified by the level of stress and that the shape of the failure distribution remains unchanged according to the following relationship and as shown in figure V-1:

$$R_{stress}(t) = R_{mission}(AF \cdot t) \qquad \text{Eq. V-1}$$

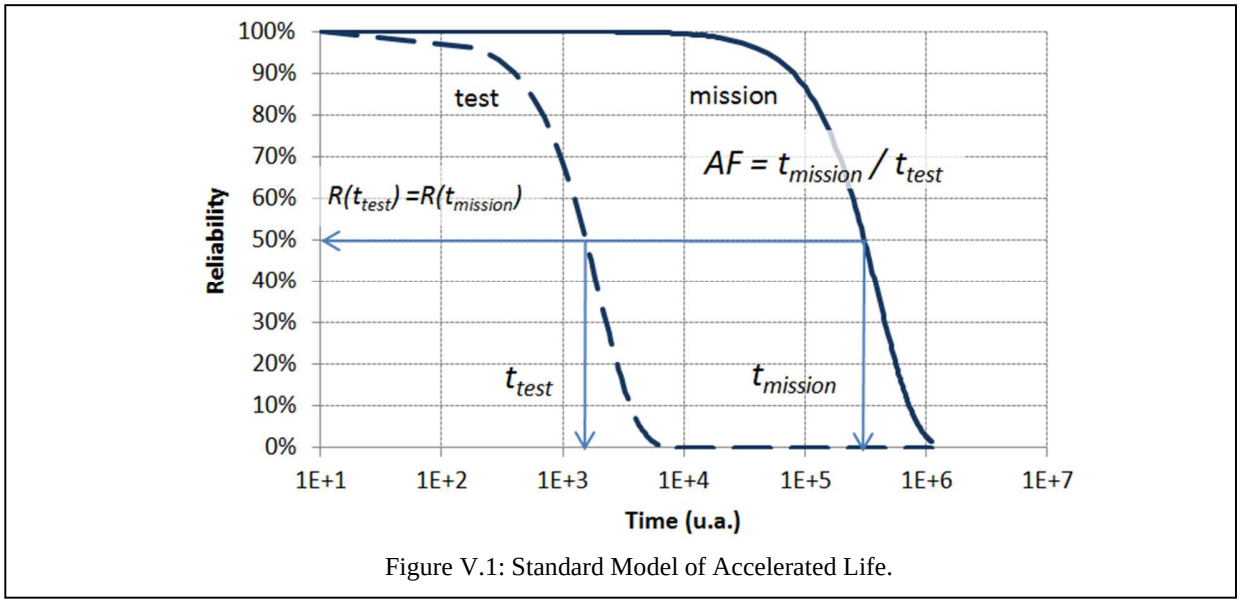


Figure V.1: Standard Model of Accelerated Life.

- The modeling approach assumes the same physical mechanisms can cause early "random" fails and intrinsic wearout, but act on different parts of a strength distribution (e.g. abnormal/defective sites [71]).
- Wearout assessments are expressed in terms of "time to failure" or "failure fraction" due to any or all competing PoF mechanisms if occurring simultaneously (see for example NBTI, HCI, TDDB and EM on Si DSM technologies. Failure models are characterized by the instantaneous failure rate drafted by the bathtub curve. It is a simple and comprehensive description of how failure distributions are observed during ageing.

• Random failure rate are expressed in Failures in Time (FITs). Mathematically, wearout failure rates tend to increase with time and are characterized by Weibull distributions with shape parameter, $\beta > 1$ while random failure rates are defined to be constant with time, characterized by the exponential distribution (the exponential distribution is a special case of the Weibull distribution, in which the Weibull slope is one), or decreasing in time for infant mortality, characterized by a Weibull $\beta < 1$ [72].

• Mitigating measures and drift parameter criteria for each failure type may also differ. Wearout can be moderated by limiting exposure to fatigue environments, but this may not help prevent random failures.

• Both random and wearout failures are described by different Weibull shapes parameters need to be combined since these differences limit the usefulness of a single, combined metric. Nevertheless, both metrics must be considered in competition for new emerging COTS devices due to their nanoscale geometries and low voltage biasing range.

• If random phenomena are existing only, reliability can be estimated using the chi-squared distribution as a function of the cumulative operation and the number of failures observed.

• If wear phenomena exist, it is necessary to adjust the reliability or accelerated degradation model from the operating time hypotheses or the degradation levels observed in tests using the maximum likelihood method.

• It is important to consider if high level accelerated test conditions far away from the use condition they may lead to the activation of different failure mechanism because their order rank can be altered (see for example HCI mostly activated at low temperature compared to EM activated at high temperature). More details are given in the methodology described in chapter Aerospace Electronics Reliability: Practical Application of MTOL.

It becomes necessary to describe a process flow to assess the wearout time-to-failure a component can achieve before the end of lifetime of the product using it. This tool is managing and implementing document history (return of experiment, heritage) and store it to assess the risk a new device will bring to a future equipment in construction and design. This method is based on data and models available about a given technology, so it won't give a confident lifetime of component, only a level of risk that the lifetime is too close to product lifetime.

Questions raised in this chapter will address some very specific aspects:

How to build a methodology to forecast the reliability of a device, card, or system in its harsh operating condition? Forecast and not predict because predict is something related to a "crystal ball".

What are the clues and the evidence to answer correctly and objectively if there is a risk or not to operate a new technology, in a repeatable, safe and precise analysis process?

How to automatize such a method in order to reduce uncertainty and avoid influence of many expert positions or even give to the single expert the right to come back to his analysis later in years and to improve it because of new evidence of data?

Prognostic Failure Methodology (PFM)

Current reliability predictive standards such as the 217^{+} [8] or FIDES [66] quality guideline are based on the Arrhenius and Eyring laws which express random failure rate and the time to failure (TTF) as a function of a constant activation energy and power law of stressor parameters. These tools involve accelerated tests aiming to quantify the life time-to-failure characteristics of a product, equipment, system, electronic card or a component under specific environmental operating conditions as labelled "Mission Profiles". They are focused to predict the constant Failure Rate (FR) associated to each failure mechanism.

Confusion may exist as failure rate of a given mechanism can be associated to two concepts:

- either the end of life or so-called wearout region of the Bathtub plot considering the failure rate is by definition a non-constant failure rate process,
- the random failure rate period where the devices under mission profile or equivalent accelerated tests may fail randomly but are still related to a single failure mechanism. Because of the exponential law, the time-to-failure is simply deduced as the inverse failure rate.

The two kind of TTF must be considered and quantified. Generally, the wearout is supposed to occur later than the TTF deduced from constant FR.

Today, the advanced microprocessor PMOS or NMOS FinFET technologies continue to shrink with very narrow node dimension and process variability close to the operating voltage limits, that the early wearout have been confirmed since a decade by Boeing [68] or Thales [73] (see figure V.1).

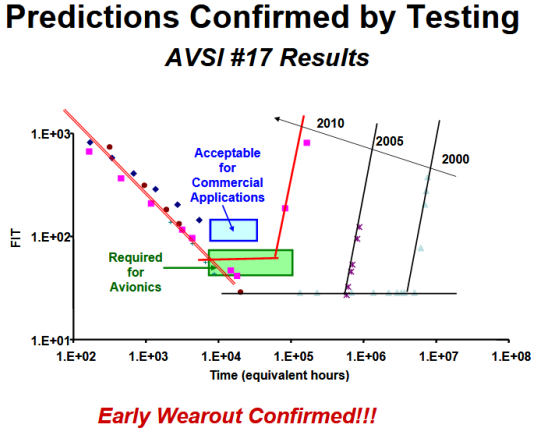


Figure V.1: Extract of L. Condra et al. presentation at National Software and Complex Electronic Hardware Standardization conference in 2005 (ref [68]).

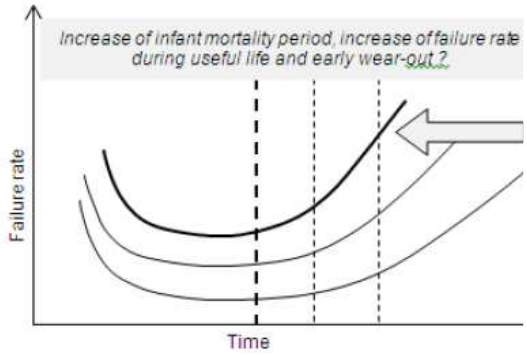


Figure V.2: Bathtub plot evolution (ref [73]).

Indeed, as noted by [73] if most publications agree to say that wear-out failures occur more and more prematurely and reduce strongly the useful life, it is more difficult to distinguish how the failure rate behavior will evolve in the next years. The authors mentioned from all recent studies and publications on Deep Sub-Micron components state at least two facts e.g. on one side, an expansion of the early life and on the other side, a shortening of the useful life.

Failure rates evolves as shown in figure V.2: a bathtub curve describing the relative failure rate of an entire population of devices over time.

The lifetime of a component or of an electronic card, is quantified using a constant random failure rate (λ). This law of probability applies to rare events or a situation where the manifestation of an event occurs randomly over time with a low probability. The reliability estimation problem turns out to be different if we consider a product subjects to a degradation process that can be quantified over time and on which a criterion of acceptability can be set. It is then possible to follow the evolution of this drift during stress tests, in order to evaluate the forecast reliability of the product, or during its operational life, to know its state of health from day to day, to estimate its remaining potential and act in consequence.

Saying this, how to discriminate these concepts (constant vs non-constant rates) when people are locking a pragmatic solution to assess the risk of failure during the entire mission life of an equipment, a product or a component?

To answer this question requires to elaborate a failure risk analysis methodology as preliminary stated in a guideline proposed by GIFAS (Groupement des Industries Françaises Aéronautiques et Spatiales) [74]. This document describes a method for selecting digital components to ensure that the reliability of these components is compatible with the requirements in aeronautics and aerospace environments. Guidelines are provided to assess the long-term reliability of COTS semiconductors components in such applications. Those guidelines mainly apply during the electronic design phase when selecting electronic components and assessing the application reliability. It focuses on the intrinsic wearout of DSM COTS semiconductor components processed in below 90 nm and puts aside, at this time, packaging wearout and random failure mechanisms. In this view, Physics of failure (PoF) is at the heart of the approach.

To address this demand, a reliability risk analysis of the lifetime of an electronic device, versus the mission profile of the final product is defined as a special guideline document named Prognostic Failure Methodology (PFM).

The general overview method is shown Figure V.3. Three loops are drawn from the most general to the most detailed. The external ring start from the Technology under analysis, how the Design at system level is defined? what are the constraints and stress environment?, for which applications?, what are the criteria associated to the failure modes and what are the End of Life (EOL) prediction or maintenance requirements. Each step is then detailed through data processing sequences and 4 risk's levels are defined and depend on the reliability and design knowledge the manufacturer can share with the user. The goal isn't to calculate the failure rate of a component, but to assess the level of risk of the lifetime being shorter than the mission duration of COTS electronic component, under its nominal environment.

Reliability concerns can be seen and analyzed according to the level of knowledge and parameter (signature or indicator) measurements:

a) either from a pure statistical point of view when a population of devices is described by catastrophic failures counted at random in a stochastic process (no history of the past is taken into account). In this case constant failure rate λ_c is defined based on *chi²* test and number of hour.components (e.g. number of devices multiplied by duration of the observation). So λ_c is expressed in FIT (Failure in Time or number of failure in 10^9 hrs).

b) or from signature parameters drift measured and degrading when time elapse thus considering history of the degradation induced by stress

As pointed out by M. Nikulin et al. from chapter 8 of ref. [75] it is natural to consider that a component experiences a wear effect characterized by a growing monotonous failure rate. The contributing link between degradation and particular types of failure can be quantified by two considerations:

- the failure occurs when the degradation exceeds the threshold limit,
- the failure has an instant risk function that depends on the history of component degradation.

The corresponding mathematic, supporting these two kinds of observation (random or wearout degradation), are quite different and gives distinctive time to failure estimations. On one way the constant failure rate λ_c allow to calculate one equivalent Time-To-Failure (TTF) as the inverse of λ_c while the other way is to characterize the start of failure mechanism occurrence by analyzing series of test experiments under various condition of stress (assumed representative to accelerate the same failure mechanisms as the one existing under nominal use of a device). In this case the parameter drift population is described by appropriate statistical models (Weibull, Lognormal, Beta function or Gamma functions, etc). Each model is represented by a quantile value parameter time to reach failure criteria which is in general defined as the time for half of the population to fail (the most usual one). This time is noted $t_{50\%}$. Depending on the mathematical model applied the MTTF is calculated according to the following expression:

$$MTTF = E(t) = \int_0^\infty t \cdot f(t) \cdot dt = \int_0^\infty R(t) \cdot dt$$

Eq. II-2

Where *E(t)* is the mean expected value of the variable t. Refer to chapter V for f(t) and R(t) definitions. For an exponential distribution

$$R(t) = e^{-\lambda \cdot t}$$

Eq. II-3

And IV.1 reduces to:

$$MTTF = \int_0^\infty e^{-\lambda_c \cdot t} \cdot dt = \frac{1}{\lambda_c}$$

Eq. II-4

It is also easy to demonstrate the expression of MTTF for a lognormal distribution ($t_{50\%} = e^{\mu}$, μ the location parameter and σ the scale parameter) is given by:

$$MTTF = t_{50\%} \cdot e^{\frac{\sigma^2}{2}}$$

Eq. II-5

The failure rate for an EOL mechanism is typically a non-constant failure rate increasing with time and then decreasing when the lot population reduce to zero: this is because by definition the failure rate is the ratio number of failed devices divided by the number of remaining sound devices. This ratio reduces when the remaining population.

Consequently, in parallel to the constant failure rate evaluation, it must be verified that the wearout of a component will occur far after the lifetime of the final product of which it is apart.

Sometime manufacturers provide information about the lifetime, but unfortunately, they often not communicate on this subject. The methodology describes a process flow to assess the wearout time-to-failure of a component can achieve before the end of lifetime of the product using it. This method is based on data and models available about technology, so it won't give a confident lifetime of component, only a level of risk that the lifetime is too close to product lifetime.

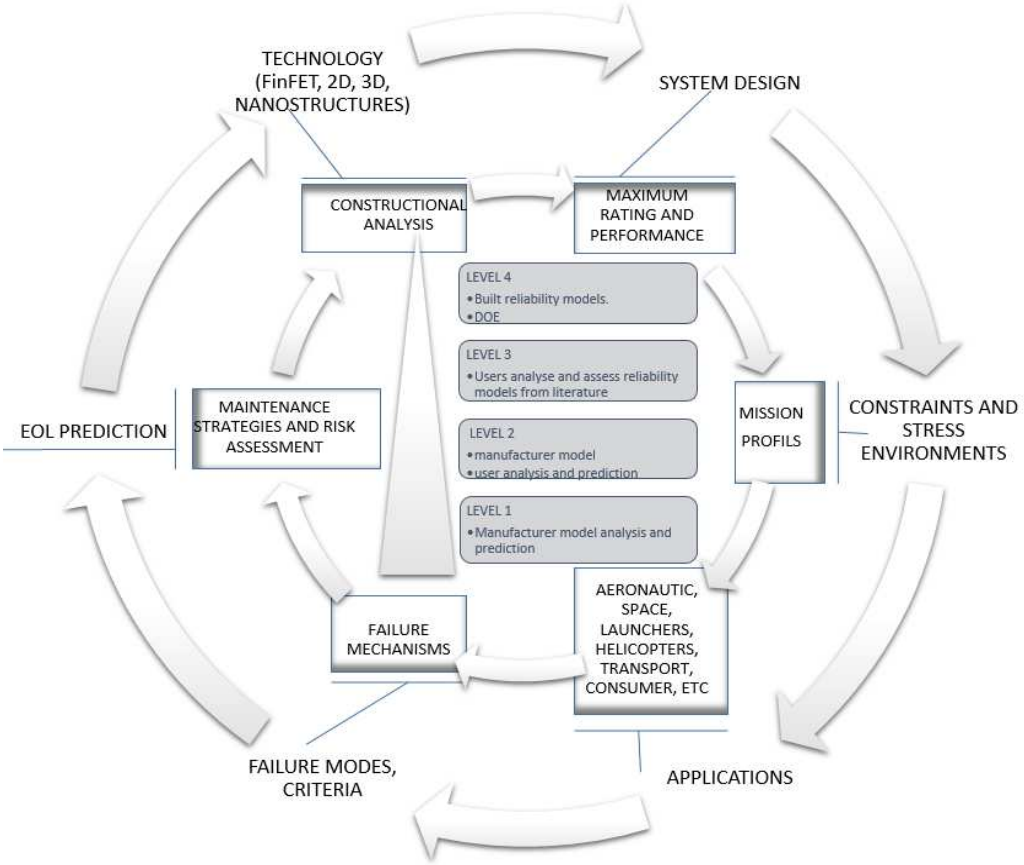


Figure V.3: General PFM methodology overview (External ring related to PFM method loop from Technology to EOL prediction, second ring describing definition/content of each topic, center relates to 4 levels of reliability risk assessment.

The method proposes four levels as defined in [69], depending on the data and service available from the manufacturer. Level 1 is the simplest to implement, and level 4 is the most complex:

Level 1: the Original Component manufacturer (OCM) provides the assessment of the lifetime of its product, using the mission profile the Original Equipment Manufacturer (OEM) proposes.

Level 2: the OCM gives means and tools to assess the lifetime, with data or mathematics physic models. Then the OEM compute and assess the time-to-failure of the product under the mission environment.

Level 3: the OCM declines any information related to data on the reliability models of its products. Then the OEM must assess the lifetime, based on component technologies, bibliography, manufacturing analysis and Physic of Failure models.

Level 4: the product COTS is not mature enough and no reliability data is existing or very few not consolidated. The OEM must perform accelerating lifetest sequence and evaluation program to construct reliability models and identify Physics of Failure and characterize failure mechanisms.

C. Prognostic Failure Model (PFM) level 3: Reliability prediction applied to DSM technologies in harsh environment

This chapter gives an overview on how to conduct the Prognostic Failure Model (PFM) methodology applied to a specific industrial case.

The goal is to assess the reliability of one component from industry and qualitatively assess the risk to implement such new technologies in Aeronautic or Space mission if 15 to 30 years application. Let's take the following example of a component based on advanced 3D technologies as three gate CMOS FinFETs (FF) designed at 20nm nodes and is dedicated to SDRAM Memory. The proposed study is on:

- DDR3 MICRON which contains Synchronous DRAM (SDRAM) with $L_G = L_{nom}$ = 20nm (FF20)

The final objective is to determine the reliability of such a component with respect to wearout mechanisms and hard failure probability in the context of avionics and space applications, *i.e.* under high constraints environments and mission profiles. These harsh constraints arise from the technology parameters themselves, as the components may be placed in boost configurations depending on the supply voltages (multicore processor), operation frequency and bit rate transmission.

The accent is done on how to validate or improve a technical expertise in the calculations of worst case degradation modeled under each mechanism, which may be involved in each industrial components and the corresponding Lifetime value associated with the most relevant failure with respect to the optimized FinFET technology, and to the use case of memory and FPGA circuits.

A synthetized view on how we can clarify some concept for a comprehensive harmonization of existing reliability model of failure mechanism and associated failure mode definition:

- ***Internal electrical*** stresses labelled ***Stressor parameters*** are responsible of the wearout failure rate (Weibull β greater than 1). They are only of four types of applied and imposed stress conditions: they are voltage, current, dissipated power and input signal or ESD/EOS/EMC energies and can be either static, dynamic, transient or surge. They are quantified with respect to their level of stress applied compared to their level of burnout instantaneous failure mode. But for sake of standardization and normalization they are limited by the maximum values allowed by the technology.
- When device operates under ***External stress*** (thermal management constraints, packaging and assembly constraints, atmosphere contaminants, radiations environments), such stressor parameters level are modified with respect to their maximum burnout and breakdown limits thus accelerating wearout failures compared to temperature and biasing stress in the absence of external environment.
- ***Failure modes*** of interest are ***electrical or mechanical signatures*** related to failure mechanisms observed and are ***Predictor parameters***. Such parameters can be measured as absolute drift value of electrical parameter or as relative percentage of drift.
- ***Constant failure rate (Random)*** are caused by random defects and random events. The Failure rate is modeled by a Weibull shape parameter close to 1 which is equivalent to an exponential distribution law.
- ***Lot-to-lot production variation*** (respectively device-to-device) and performance dispersion from a single manufacturing lot (respectively device) will affect the burnout limits, inducing in return a change of percentage of stress applied on a given lot (device). Statistic dispersion will affect the time to failure on similar way (producing the same statistical effect). Such dispersion at lot and device level will impact the Remaining Useful Life (RUL) for some part of the population.
- ***Infant mortality failure*** population are caused by "defects" and correlates with defect-related yield loss. They are reduced by improved quality manufacturing and by screening.

Stressor definition and normalization

In a similar way considering a population of devices submitted to heating will only degrade continuously up to malfunction and failure. But when superposing high (or low) temperature and adequate stressors, the time-to-failure of such alike population will reduce. The term "stressors" here is defined as the electrical factors applied to the device of concern. Stressors are all limited by technology boundaries defined by the burnout values of each related electrical parameter (breakdown voltage, current overstress and burnout, power burnout, input signal overstress). These stressors can be normalized with respect to their burnout limits and strains are weighted as percentage of breakdown limits. The main hypotheses, verified by experiments on electronic devices and population of similar devices, are:

i. the physical instantaneous degradation phenomena due to electrical stress above the limits is observed at any temperature and depend of the active zone temperature of the device under test (Sze, S. M [56])
ii. the relative drift of a predictor parameter is a function of time (for example square root for diffusion mechanisms) and relate to a failure mechanism activated by temperature and biasing.
iii. For a biasing set higher and close to the breakdown limit, the two failure mechanisms (e.g. the diffusion and the

instantaneous catastrophic ones) are in competition and occurred simultaneously; for sake of simplicity it is assumed they are progressively and linearly combined from a pure diffusion mechanism at nominal biasing to a pure burnout at high bias (voltage or current of power dissipation).

Predictor definition

As mentioned previously, an electrical predictor parameter ξp is defined as the electrical signature (failure mode) of a failure mechanism of interest. Such a parameter can be normalized with respect to its initial value at time zero as a relative drift $\Delta\xi p/\xi p_0$ (in %). Figure 4 is a schematic drawing showing how predictors are defined and change when stress time elapse. The predictor relative drift shown is an example of actual measurements performed on microwave transistors when submitted to steady state aging testing [58]. The predictor of each single device is normalized with respect to its initial measurement (mean value) and the failure criteria was 20% drift reached. So, the drawing is set to consider failed devices for all drift greater than 20%.

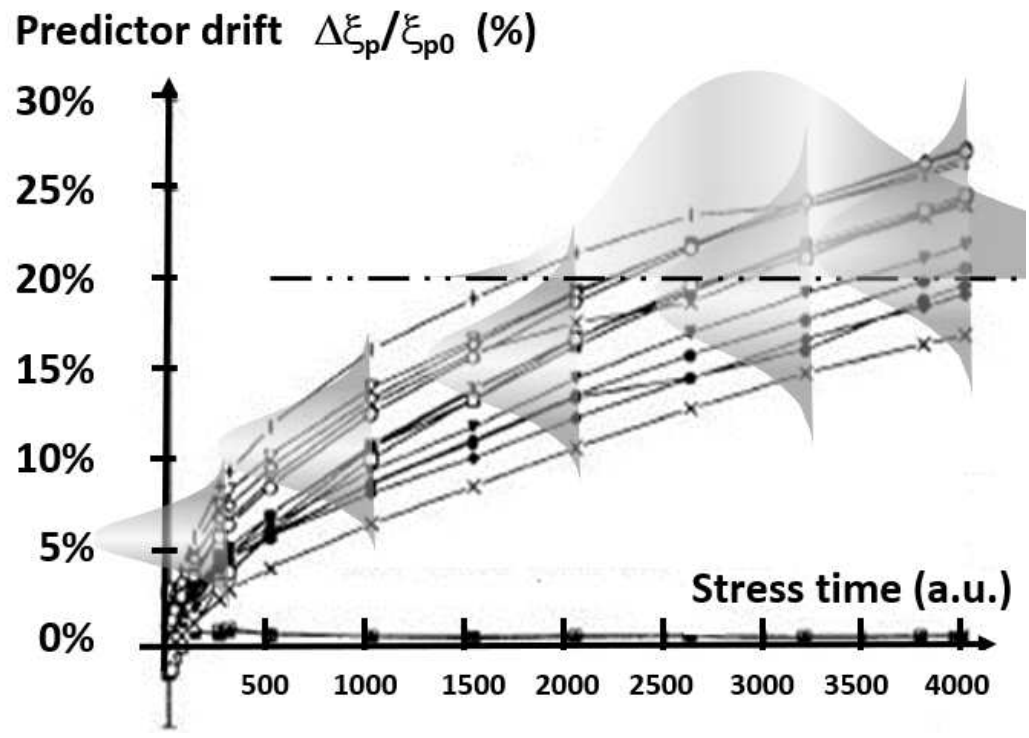


Figure VI.1: Schematic drawing showing how predictors are defined and varies when stress time elapse.

We have defined a special industry case application with the aim:

1 To validate the degradation models firstly identified on the basics of provided data in the modeling, and de-processing analysis (Constructional Analysis - CA) for each component. Description of the models adapted to the present cases about TDDB, HCI, BTI and EM in order to do a ranking of the Wearout mechanisms as a function of the operation mode and in relation to the CA previously done which has given details of the structures, dimensions and type of materials.

2 To relate the Physics of Failure (PoF) to these components considering their own parameters by giving the necessary details coming from recent references for justification in correlation to technological data given by the manufacturers. Summary will be given in tables with respect to selected mechanisms, pointing out the sensitivity and worst-case accelerating values. The goal is to determine a generic structure of elementary FinFET transistors in terms of dimensions and materials as an "intrinsic architecture" which will help to select the weakest points (at front end first) about lifetime calculations according to the operation mode.

3 To understand the data previously published and chosen to elaborate a strategy of data exploitation by the justification of the different acceleration factors in AC mode versus DC mode. Validation of the models for avionic and space applications using their respective mission profile. And if it's possible, to extend the results of the study by giving new recommendations for design rules carried out by the manufacturers.

Therefore, we study in a first part the validity of the development focusing on the distinct acceleration degradation modeling, consider the most adapted framework based on recent developments in the literature obtained on the main degradation mechanisms. In a second time we correlated this part to the FinFET technologies and their own specificities for the use case, which will concentrate on the extraction of sensitive parameters with the help of the accessible data from the manufacturers. Hence, the first part deal with the validity of the chosen accelerated models with respect to published results about Wearout and failure mechanisms that are still in progress with the continuous scaling of the CMOS technology.

1) Component selection

As the ultimate goal is to select a component with regards to searched functionality and their level of performance, the risk that the lifetime of a component does not complies with the lifetime of a product (regarding its mission profile) becomes highly probable with shrinking the device geometry, design rules and biasing condition applied. The tested technologies are developed from recent FinFET 20nm and 16nm processing development, but they have distinct application scheme. Then, it is necessary to consider the main differences about the general application type and extract the main parameters and operation conditions given by the manufacturer that are involved in the wearout damage.

2) Construction Analysis principle with Prognostic Failure Model (PFM)

The principle of cutting the Prognostic Failure Model (PFM) methodology in sub sections between level 1 and level 4 is relevant considering the low level of knowledge we may have from the manufacturers. The difficulty is indeed about the correlation to the limited results published on new FinFET structures with the theoretical acceleration laws, which may be applied to these 3D devices pointing out the good correspondences and limitations of the techniques. This is also linked to the type of devices (N-channel vs. P-Channel), the shrink of channel length (L_{eff} / L_{nom}), and the processing flow (Dummy gate, Replacement metal gate, Trench isolations for the gate and for lateral isolations) that may continuously change with respect to the trade-off between performance and reliability [76] done by the manufacturer.

Many aspects are treated in detail in the whole Prognostic Failure Model (PFM) methodology and the principle of lifetime calculation considering realistic mission profile that accounts for on/off phases standby phase and aircraft maintenance. For example, to express the TDDB, the known reference acceleration factor, the temperature, electric field and device geometry acceleration factors can be determined separately. This is a rule of thumb that will give us the technology quality and resistance under a given degradation mechanism, considering the effective temperature. However, we have too much unknown parameters from the technology side, that are confusing the determination of an accurate lifetime value for a given device and functionality. That is why a sort of shortcut is implemented in order to simplify the lifetime extractions for all mechanisms that need to be selected in order of importance for memory application (DDR3L).

Consequently, we propose lifetime extractions based first on the selection of the mandatory elements required for an accurate analysis of reliability by predictive calculations applied to these recent technologies, *i.e.* based on 3D FinFETs 20nm down to 16nm. This can be categorized by numerical calculations of a minimum, typical and maximum reliability parameter extractions in states of dimensions, voltage, temperature that will be related in a first step to uncertainties arising from processing quality and technological spread in designing FF at such dimensions. Indeed, this architecture is often obtained by a simple shrink from previous CMOS nodes, as we do not know the full process flow. For example, in some CMOS products, TSMC has derived its 16nm FinFET (FF16) CMOS from the shrink of the previous 20nm node for Xilinx. That's why the presented results have the finality to be compared to the measurement steps (Fig. IV.3) within a simplified building block framework (similar to level 4) in order to validate the PFM approach which is becoming strongly complex with the large number of parameters and failure mechanisms (FEOL/BEOL) with the range of uncertainties extracted from the CA step. This explains why we can propose a 3-loop methodology that can be derived from PFM flow, where for step 1 in Fig. VI.2, these dimensional uncertainties can be straightly transferred to lifetime margins once one knows the nominal supply voltage (V_{DD}) and boost conditions in speed (V_{DDmax}) between core and IOs.

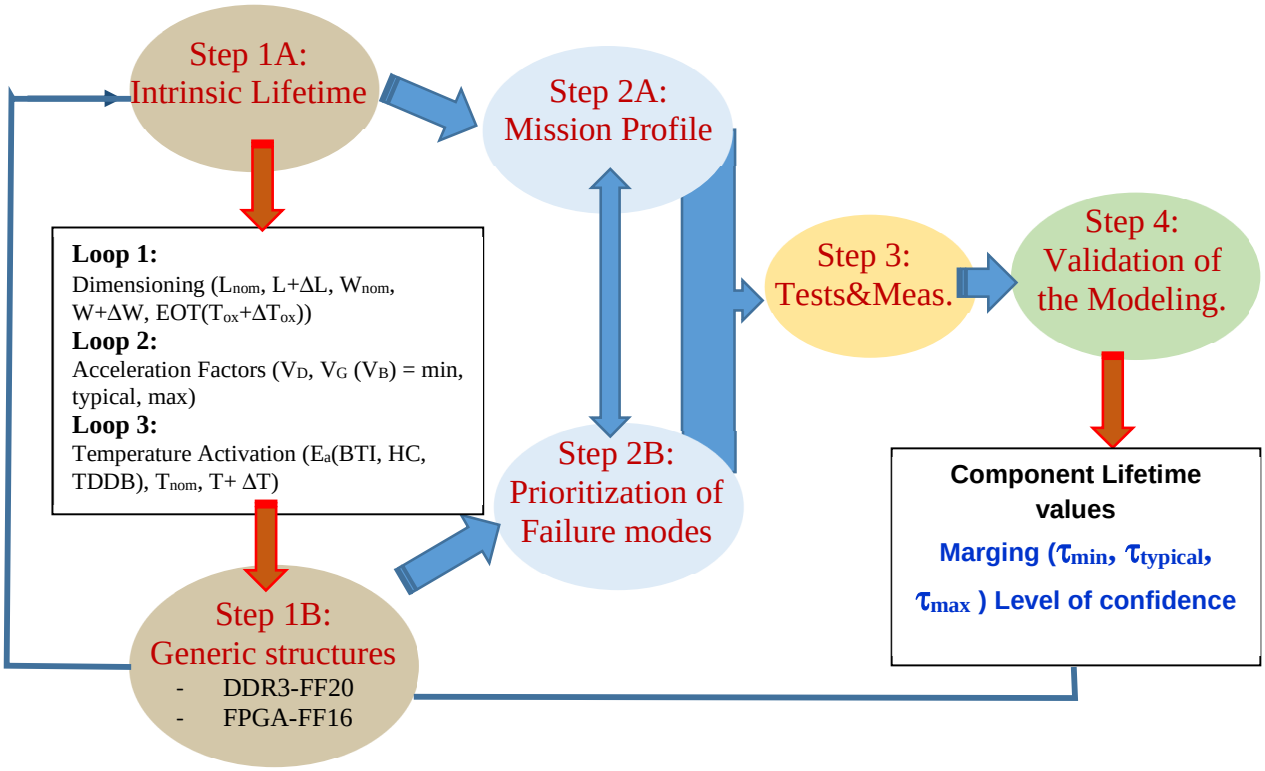


Figure VI.2: Schematics of the main steps involved in the Reliability determination of DDR3 (FF20) and FPGA (FF16) CMOS technology nodes.

Hence, this study is primarily devoted to the realization of step1 and step 2, according to the tasks previously cited above (work plan) in coordination to IRT results and the possible comparison later to measurements performed at IRT (step 3) under realistic temperature and cycles conditions. It is important to see that the Prognostic Failure Model (PFM) approach and lifetime determination for components requires this comparison step, which can bring good calculation accuracy through the temperature dependence and supply voltage acceleration.

3) Degradation mode monitoring

As the applications requiring zero failures like Aerospace, Vehicles, Implantable Medical devices, it is needed to take into account the reliability assurance of the future must be based on the transitioning from Time-To-Failure modeling to degradation modeling. The Gibbs free energy model presented in chapter IV is a method for reaching a lower potential and describing degradation physics. Because of Thermodynamic laws (First law of Energy Conservation, Second law of Entropy as description of disorder of isolated systems tends to increase with time. The Gibbs Potential report change in Enthalpy and change in Entropy, and so degradation will always lower the Gibbs Potential because all forces on electrons/atoms/molecules are derived from gradients [77]. Degradation models commonly used are power-law or exponential or logarithmic laws. As of concern of NBTI, this mechanism has been a persistent reliability concern for Silicon ICs since mid-1960s. A review of studies of various aspects of NBTI degradation have been reported by M.A. Alam et al. [78]. The Reaction–Diffusion (R–D) formulation of NBTI degradation, assumes that NBTI arises due to hole-assisted breaking of Si–H bonds at the Si/SiO2. M.A. Alam discussion addresses a careful analysis of the absolute value of time-exponent n, and resolved that diffusion of H2 molecule would consistently interpret experimental data regarding quasi-saturation, frequency-insensitive degradation and n = 1/6 exponent with ‘‘no-delay’’ measurements. Yet he considered too, “instead of absolute values of n, the groups of studies who focused on variation of n as a function of temperature (determined by standard delay-based measurement) reached an entirely different conclusion [79]: that the transport of H-specie is dispersive and most likely specie is H+(proton), not H_2”. The classic exponent of n = 1/4 is the signature of NBTI degradation based on wide variety of experimental results. But it is well-known that N_{IT} (the fraction of Si–H bonds broken at time t due to NBTI stress) generated during the stress phase of the degradation recovers as soon as the stress is removed: NBTI is a self-annealing process.

Considering, interface trap generation with stress time for different models of Hydrogen diffusion, the RD model summarizing interface trap generation in all the regimes of stress phase is shown in fig. VI.3 from [80] and also plotted the various cases of resulting power laws. The exact value of the slope depends on the underlying mechanism of Hydrogen flow inside oxide and the mechanism of reaction.

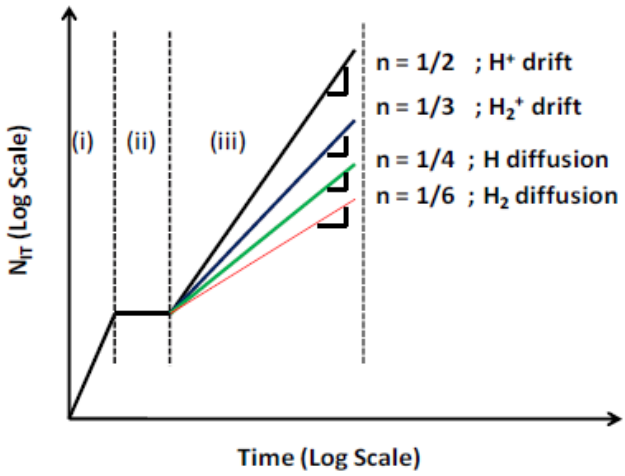


Figure VI.3: Interface trap generation with stress time for different models of Hydrogen diffusion.

In a paper published in Microelectronics Reliability journal in 2018, titled "Controversial issues in negative bias temperature instability" J.H. Stathis et al; [28] concluded different approaches to the fundamental study of NBTI naturally lead to different conclusions. *There is still no consensus on the basic physics of Negative Bias Temperature Instability. Two competing models, Reaction-Diffusion and Defect-Centric, currently vie for dominance. The differences appear fundamental: one model holds that NBTI is a diffusion-limited process and the other holds that it is reaction-limited. Basic issues of disagreement were summarized, and the main controversial aspects of each model were reviewed and contrasted.*
Broadly speaking, the critiques of these two models can be summarized thusly:
(1) While the Reaction Diffusion (RD) model can now very successfully describe a large variety of observed data over a broad set of experimental conditions, the validity of the underlying physical interpretation is questioned because the model parameters conflict with established literature on H in Si/SiO2 systems;
(2) The Defect-Centric (DC) model has a strong basis in physics, supported by microscopic measurement of discrete defects, but until recently it has paid scant attention to interface state generation, and the ability to comprehensively describe NBTI stress data over a broad set of process and stress conditions is questioned.

There is a need to be selective with respect to all degradation mechanisms (assumed here to be described by the constant failure rate mechanisms) that can be involved at the front end and back end. Generally, FEOL reliability requires mainly TDDB, BTI, HC and EM at first level of metallization (M1) but may also consider also cosmic rays (CR) through the gate reaching active channel and S/D regions. BEOL Reliability analysis generally focus on interconnect stacks looking to Electromigration (EM), stress migration/voiding (SM/SV), thermal mechanical stability, low K dielectric breakdown (TDDB) and chip/package interactions (CPI), while cross talk and Electro Magnetic Compatibility (EMC) can have also a net impact for a given packaging.

We focus in this present work on FEOL. Till now, only few works deals with interactions between degradation mechanisms, considering that each progressive degradation can be treated separately, leading to a direct simplification that the wearout issue of a system (MTTF) examined by a steady state failure rate (λ_{SS}) that can be considered as a generic failure rate λ_G for device "*i*" expressed as a function of its own acceleration factors (*AF*) depending on voltage, temperature, quality (variability) and active surface with:

$$\lambda_{SSi} = \lambda_G.AF_V.AF_T.AF_Q.\ AF_S \qquad \text{Eq. V-6}$$

leading to the sum of a total failure rate $MTTF = 1/\lambda_{SS}$, with:

$$\lambda_{SS} = \pi_E \sum_{i=1}^{N} n_i\ \lambda_{SSi} \qquad \text{Eq. II-7}$$

Where n_i is the quantity of device "*i*", *N* the total number of devices, π_E the environmental factors which can be related to mission profile (t_{on}/t_{off}). Our goal is to determine the component reliability regarding its generic reliability criterion based on a (simplified) generic structure extracted from the DDR3 and FPGA components.

a) TDDB modeling

Time-Dependent Dielectric Breakdown (TDDB) in dielectrics is one of the more important failure mechanisms for digital circuits as the scaling of metal-oxide-silicon field effect transistors (MOSFETs) is become a figure of merit for performance gain with miniaturization. For this reason, Physics-based models for TDDB and Physics of Failure (PoF) include many different approach through gate-oxide thinning with field-based models [81], current-based models [82], or a combination of field and current-based models [83], and as well, as a voltage power law acceleration model [84] [85] and multi vibration H release (MVHR) model [86].

The Prognostic Failure Model (PFM) development for TDDB study at FEOL is well organized taking into account the most important parameters that are primarily involved in lifetime determination and pointing out the difference between use condition and accelerated stress conditions. Good references are chosen throughout the historical TDDB modeling, giving a clear picture of the recent trends used for lifetime extrapolation. For practical use of TDDB calculation, the way by which slope Weibull plot (β) is obtained as well as the level of confidence at $t_{63\%} = \eta$ factor extracted from publication, can bring a certain level of accuracy depending on the used publications. The difficulty arises from the lack of recent publications on FinFET TDDB coming from foundries as TSMC, Intel, and Samsung. Even if the last published models are not referred as for example the clustering [87] and filamentary modeling [88] that leads to bi-modal distributions (β variation with T_{ox}), these recent modeling bring more accuracy for high level of percentile, non-Poisson area scaling. This is explained by non-uniform distribution of defects (deviating from Weibull statistics) between interfacial layer mode and bulk oxide behavior that result in a change of slope and related level of confidence through η factor variation, that leads to saturating behavior in AF_V influence. As we want accuracy towards small supply voltage condition, we can indeed remain on the E model for our subject of interests, especially for TDDB.

One focus is to determine the lifetime margin (TDDB) with typical, min and max values (that can be related to voltage conditions between derating and boost) we can indeed (as proposed by Prognostic Failure Model (PFM)) choose the most severe lifetime criterion as the E-model (thermochemical) which is most restrictive with respect to its direct acceleration with oxide field (exponential law), and is satisfactorily applied to low field and long term. The 1/E model is rather straightly related to the injected (gate tunneling) current at high field, and was historically used in thicker gate-oxide MOSFET (T_{ox} > 5nm), leading to a direct determination of t_{BD} in relation to the Anode Hole Injection (AHI) mechanism that puts the back-injected holes into the modeling (NMOS device).

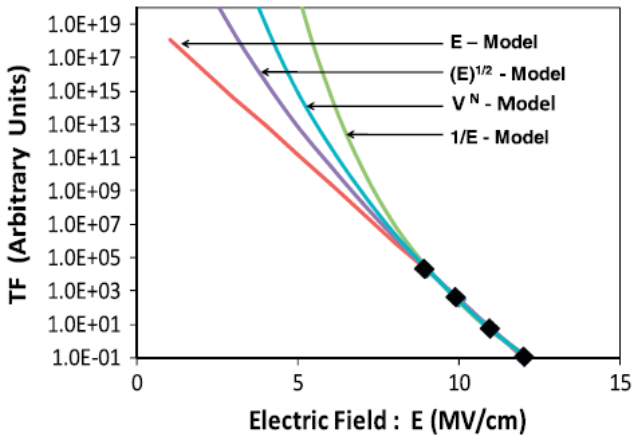


Fig. VI.4 : Comparison of different TDDB acceleration modeling [89].

However, with T_{ox} thinning (< 3.5nm), the charge at breakdown (Q_{BD}) is strongly reduced due to direct tunneling leakage currents. That's why some works turned out to a direct voltage dependence either as an exponential law [90], [86], or a power law in gate-voltage [91] known as anode H release modeling. Therefore, one can indeed prefer the E model as is remains more restrictive for lifetime estimation in products, especially devoted to avionics and space.

Moreover, with the reduction of supply voltage, many results have shown that hard breakdown (HBD) from the gate-oxide field has a reduced probability to occur, and that when it may occur (coming also from a triggering of field spikes or to the correlation to other failure mechanisms as ESD, EM at first metallization level M1), this would first induce a soft breakdown (SBD) that can reach HBD, but it still maintains the functionality particularly in ring oscillator or digital cells, showing that the first breakdown event does not necessarily spell the immediate failure of the entire CMOS application [92].

This points out that HBD may be more closely related to BEOL reliability and the scaling challenges of thinning ILD for metallization with the use of the Low-K that has to be examined as a function of type of material (SiOCH) and its level of porosity, the mechanical weaknesses as cracking, delamination that lower the TDDB breakdown strength, and also the critical spacing rules for metallization M_x to M_{x+1} level. All these mechanisms with scaling can have stronger effect for failure analysis due to resistance variation with material quality (grain boundary, surface scattering, and increase of liner / Cu volume ratio).

That's a reason why HBD for FEOL and core using low supply voltage is not relevant applied directly to the device from M_0, but rather applied to IO with larger supply voltage. In a way, the more we have on current, the more we meet dissipation process and hot spots that link us to FEOL reliability issues at the M1 level with ILD and Low-K that suffer from smaller critical field conditions to trigger eventually HBD to gate-drain overlap region of a biased device.

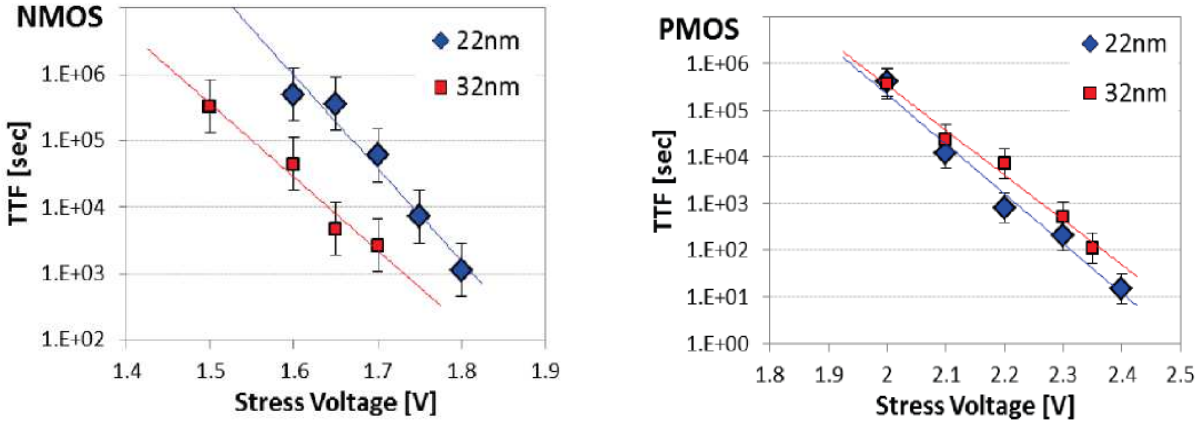


Fig.VI.5 : NMOS and PMOS TDDB in 22nm FinFET shown as better improved with respect to 32nm that may be used for DDR3 [93]

An example can be obtained with the E-model in 22nm FinFET that can be used for DDR3, comparison following the exponential law in figure VI.5 a and b data given from INTEL [93].

In this case we can extract directly the exponential law parameter (equivalent to E-model) usually expressed as $exp(-\gamma E_{ox})$ if one wants to retain the restrictive evaluation criterion. The practical use is to plot the results expressed as a function of the stressing gate voltage VGS, (as $E_{ox} = (V_G - \phi_{ms} - 2\phi_p)/E_{OT} \cong (V_G - V_{TH}+0.2)/E_{OT}$ which shows the TTF_{BD} at low field is obtained as :

$$TTF_{BD} = A_i \cdot e^{-\gamma \cdot V_G} \cdot e^{-\frac{E_a}{kT}} \cdot e^{[Ln(-Ln(1-F))/\beta]} \qquad \text{Eq. V-8}$$

NMOS	PMOS:
$\gamma = 34.5\ V^{-1}$	$\gamma = 32.8\ V^{-1}$
$E_{a,N} = 0.6\ eV$	$E_{a,P} = 0.6\ eV$

Where the last exponential factor is to mitigate the time to breakdown value by the distribution function. We find also the power law in V_G that can be useful but for different technology from FinFETs as in Bulk Si and standard CMOS components [91], [94].

b) HC modeling

HC modeling may be applied in an adapted way to FinFETs depending on the supply voltage range (V_{DDmin}, V_{DDmax}) and the chosen accelerated conditions (10% to 35% from V_{DD}) that can be observed from the last references on this topic [95], as

devices operate at Low voltage operation:

- V_{DDmax}= 1.5V down, V_{DDmin}=1.28V for DDR3L core memory
- V_{DDmax}= 0.9V down, V_{DDmin}=0.75V for FPGA core memory
- V_{DDmax}= 2V for Output drivers for HP IO banks of the FPGA

As we depend on the accessible data of FF Reliability for 16nm and 20nm CMOS nodes, one can choose the voltage dependence, drain current dependence or both laws including the gate-length acceleration factor, when the data are available.

For the voltage condition of periphery NMOS devices with longer gate-length (Leff= 0.1µm), common HC determination can be done as this corresponds to planar CMOS bulk with thin SiO2 gate-oxide (1.5nm), neglecting the possible use of SiON in this stack. Then in this later, one may consider the standard HC lifetime as by using a Takeda plot [96] where this useful modeling links the dominant (permanent) degradation mechanism (here ΔN_{IT}) to the maximum lateral field (E_{Max}) with the measurement of Impact Ionization (II) rate, which is directly proportional to (I_{SUB}/I_{DS}). This is generally done as a function of the stressing condition (which is maximum in the AC transient during operation for the DC/AC transfer function) corresponding to the maximum of HC rate ($V_{GS} \approx V_{DS}/2$). It results that the lifetime expressed as τ- (I_{SUB}/W_G) is equivalent to τ- $exp(A/(V_{DS} - V_{DSat}))$ where the most common approximation gives the simplified form as τ - $exp(B/V_{DS})$, if one does not know the exact value of V_{TH} is saturation (including DIBL effects [97]). This modeling is referred as the Lucky Electron model (LEM) where if one can have the data of I_{SUB} and I_{DS} we have the classical relationship:

$$\frac{\tau \cdot I_{DS}}{W_G} = C_1 \cdot \left(\frac{I_{SUB}}{I_{DS}}\right)^{-m_1} \quad \Leftrightarrow \quad \frac{1}{\tau_{M_1}} = C_1' \cdot \left(\frac{I_{DS}}{W_G}\right) \cdot \left(\frac{I_{SUB}}{I_{DS}}\right)^{-m_1} \qquad \text{Eq. VII-9}$$

That we call Mode 1 (M_1) pour "high voltage" condition, i.e. $V_{DD} \geq 2.5$V (L_G= 0.25μm). For voltage with $T_{ox} \geq$ 5nm, i.e. with thick gate-oxide where charge trapping occurs. In medium T_{ox} thickness, high-field assisted charge detrapping modifies trapping efficiency and kinetics, while in thin to ultrathin insulators T_{ox} < 3.2nm tunneling effects largely increase, modifying the classical HC scenario because of slow traps. We must use the M_2 modeling taking into account the electron-electron scattering (EES) [98] [99] for medium voltages, while in our case, we consider the low voltage modeling related to the multivibrational excitation of SiH bonds at the interface of the dielectric and silicon, as shown in Fig.VI.6 in Si bulk.

$$\frac{\tau \cdot I_{DS}^2}{W_G} = C_2 \cdot \left(\frac{I_{SUB}}{I_{DS}}\right)^{-m_2} \quad \Leftrightarrow \quad \frac{1}{\tau_{M_2}} = C_2' \cdot \left(\frac{I_{DS}}{W_G}\right)^2 \cdot \left(\frac{I_{SUB}}{I_{DS}}\right)^{-m_2} \qquad \text{Eq. V-10}$$

$$\frac{1}{\tau_{M_3}} = C_3' \cdot \left[(V_{DS} - \hbar\omega)^{0.5} \cdot \left(\frac{I_{DS}}{W}\right)\right]^{\frac{E_B}{\hbar\omega}} \cdot e^{\left(-\frac{E_{emi}}{kT}\right)} = C_3' \cdot \left[(V_{DS})^{a_3/2} \cdot \left(\frac{I_{DS}}{W}\right)\right]^{a_3} \cdot e^{\left(-\frac{E_{emi}}{kT}\right)} \qquad \text{Eq. V-11}$$

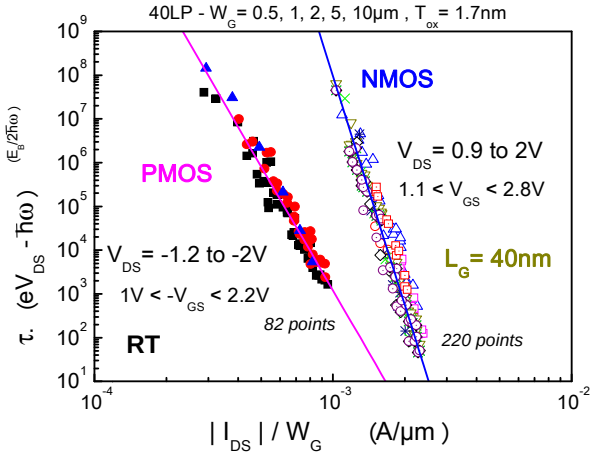


Fig. VI.6: Mode M_3 in C_{40} under CC damage according to equ. VI.6 for a large set of data in PMOS and NMOS. The M_3 degradation dependence gives a_3= 9.4 and 17.4, with C_3, V_D= $7.5 \cdot 10^{-26}$ and $6 \cdot 10^{-45}$, respectively. Same lifetime criterion 10% ΔI_{DSat} FWD and various stressing V_{GD} that can be used for the long channel Peri COS standard from transistors in inter blocks (Fig from ref [99]).

To summarize the different parameter extraction for NMOS and PMOS under mode M3 with relation VI.4 to VI.6. We have put the needed values of the degradation modelling in Table II for the complete lifetime modelling based on degradation rate between mode M_1 and M_3.

Degradation Modes	NMOS	PMOS
m ($m_1 = m_2$)	2.7	1.3
$\mathbf{a_1}$	1	1
$\mathbf{a_2}$	2.5	3
$\mathbf{a_3}$	17 - 20	9 - 12
E_{emi}(eV)	0.22-0.26	0.22-0.26

Table VI.I: Summary of the parameter in the full lifetime modeling in Silicon bulk planar technologies as a function of VDD range and Thick to medium and ultra-thin gate-oxides [95].

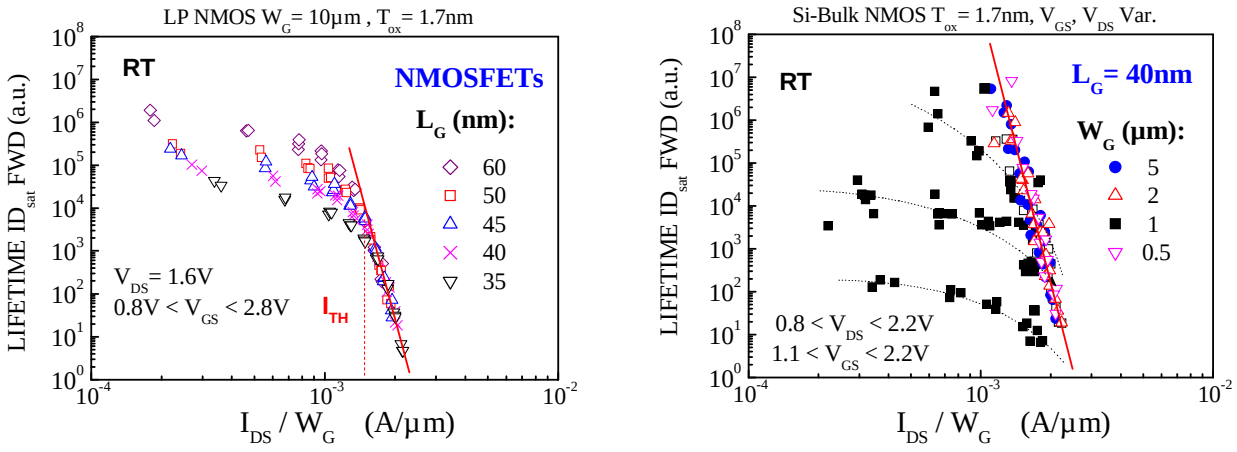


Fig. VI.7: Low voltage HC lifetime as a function of the drive current **(a)** when the channel length is reduced **(b)** when the channel width is reduced. This can offer the basics of comparison for standard CMOS (Peri CMOS) of the DDR3L component.

One has to point out that a recent publication of HC damage in FinFETs [100] has shown that our low voltage modeling (M3) is clearly followed for 14nm n-FinFETs, with very similar process as HfO2/TiN gate stack where n-FinFETs were fabricated by 14nm technology with a gate-last process with thermal (IL) oxide (SiO2) with thickness of 1nm. The TiN layers were deposited by ALD as work function metal (WFM) layers. The sizes of the FinFET used in were W= 154 nm and L_G= 60 nm and 14 nm. The authors found a slope -23.4 (Fig VI.8a-d) for L_G= 14nm (W_G = 154 nm) according to the low energy range for NMOS device lifetime in mode 3, where MVE is dominant.

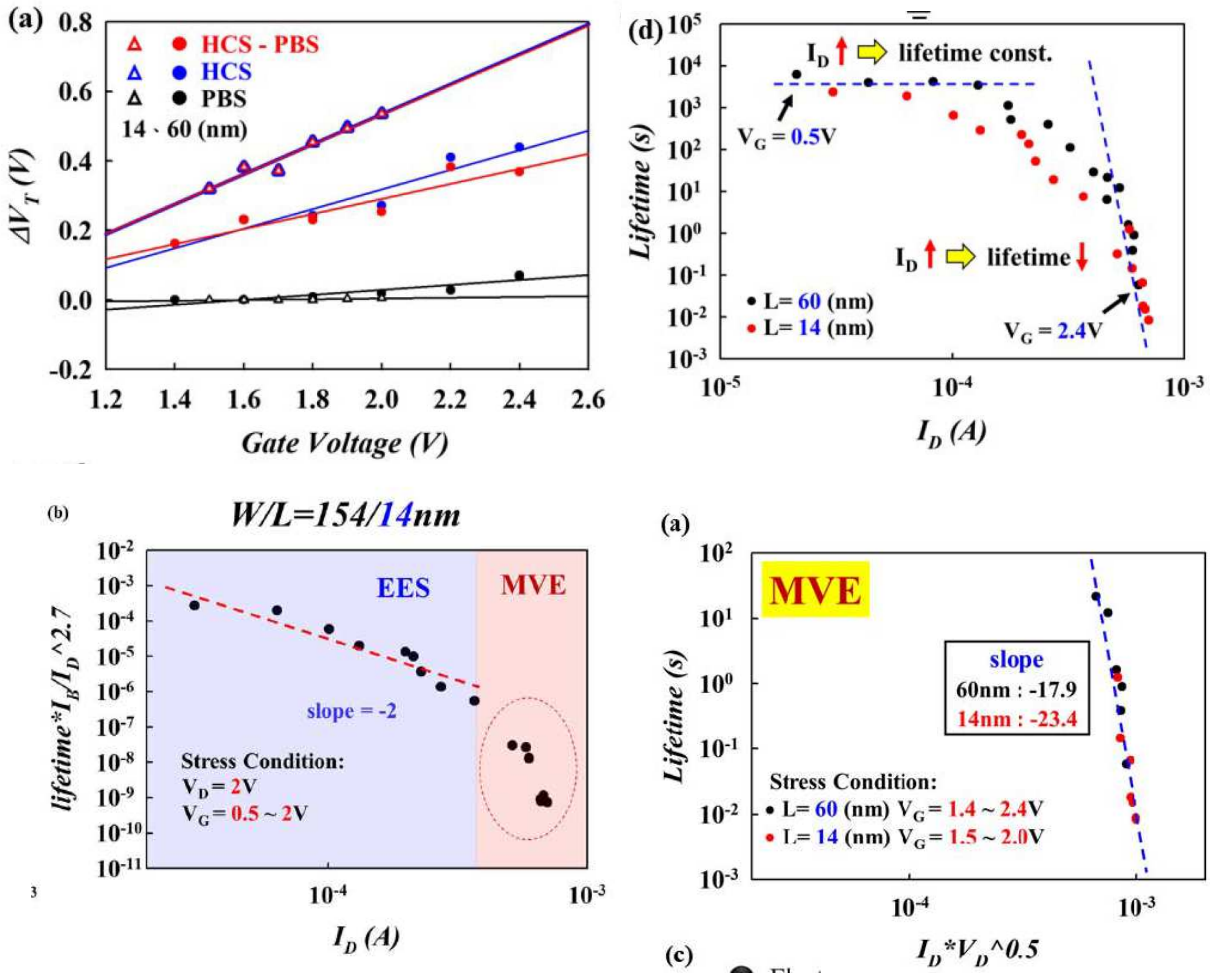


Fig.VI.8: Up left, comparison of degradation kinetics between HC, PBT, and HC+PBT (VG > VD). Up right the lifetime vs. ID dependence as a function of VG. Down left, combined Lifetime plot in under M3 mode between EES [98] and MVE [95]. Down right Lifetime as a funtion of long channel and short channel showing the phonons number dependence. [100]

This graph (down, right) [100] will be used for low voltage HC for FF 16nm as well as for FinFET 20nm as the stressing voltage has been used down to $V_G = V_{DD} = 1.5V$, (DD3R) and V_{DD}= 0.9V to 0.85V in FPGA as a function of I_{On} enabling the process constant extraction Ci required for lifetime modeling ad extraction.

c) BTI Modeling

BTI degradation is most of the time analyzed in PMOS under negative gate-voltage (NBT) as this represents the worst issues in CMOS technologies compared to smaller damage under positive gate-voltage PBT.

Among several modified approaches, three models for BTI have historically been developed through technology scaling:

- The Reaction limited model with dispersive bond breaking energies [83] and uncoupled (hole) trapping that is [N] dependent from interface traps (ΔN_{IT}) [101] this later being sensitive to H2/D2 anneals
- The composite Reaction Diffusion (RD) modeling [80] [102]
- The Stochastic Reaction Diffusion (RD) model [95] improved with two stage model and triple well transitions [103] with volatile defects [104]

The general way is to simplify the approach by a level of damage for the threshold voltage related to a Lifetime defined at ΔV_{Th} = 50mV with the permanent and recoverable contributions. This corresponds to the static degradation mechanism under NBT accelerated by the vertical field (no channel carriers), meaning that NBT damage is involved with a permanent part and a recoverable part $\Delta V_{Th}(t_s) = \Delta V_{Th,P} + \Delta V_{Th,R}$. If we consider the Huard modeling for total damage, we have the permanent (ΔN_{IT}) and recoverable part (ΔN_{ot} hole trapping and slow traps) as a function of time constant for capture and emission, stress and recovering time:

$$\Delta D_T = \Delta D_p + \Delta D_R \quad \text{Eq. V.12.a}$$

$$\Delta D_P = K_p \cdot V_G^a \cdot t_S^b \cdot e^{-\frac{E_{a,P}}{kT}} \quad \text{Eq. V.12.b}$$

$$\Delta D_R = K_R \cdot V_G^c \cdot Ln\left(1 + \frac{t_S \cdot \tau_e}{t_R \cdot \tau_c}\right) \cdot e^{-\frac{E_{a,R}}{kT}} \quad \text{Eq. V.12.c}$$

Grasser has a full (atomistic) description but a relatively similar dependence for the recoverable part expressed as [103] [104] $\Delta V_{Th}(t_s) = P + R$ with $R = R_o/(1+B.(t_r / t_s)^{\alpha})$.

Let us simplifies the approach to the main role of the permanent damage that is involved at long term in order to determine the NBT impact by the use of the compact model. This is can be first considered as a power law for TTF $\propto A.V_G^{-m}$ or else similarly to the TDDB as following an exponential law TTF $\propto \exp(-B'.E_{ox}) \propto \exp(-B.V_G)$ with the usual thermal factor $\exp(-E_a/kT)$. If one considers the shift as the power law in $\Delta V_{TH} = A\ t_s^n$ (n= 1/4 for H^o, n= 1/6 for H_2, n= 1/5 as a typical value retained for a mix of both contribution), the pre-factor $A = \Delta V_{TH}.\ t_s^{-n}$ can be normalized to the reached criterion ΔV_{TH}= 0.05V (or ΔID_{Sat} = 10%), leading at 125°C to:

$$TTF_{NBT} = A_i \cdot \left(\frac{\Delta}{0.05}\right)^{1/n} \cdot e^{-B \cdot V_G} \cdot e^{-\frac{E_a}{kT}} \quad \text{Eq. V.13}$$

with:

core	IO
n = 0.16 (core ΔV_{TH}= 50mV)	n= 0.15 (ΔID_{Sat} = 10%)
A_i = 1.48 10^8 , B= 12,	A_i = 5.02 10^7 , B= 4.6

With apparent activations:

E_a = 0.35 to 0.25 eV (Si Bulk)
E_a = 0.6 eV (FDSOI)
E_a= 0.15 eV (FinFETs)

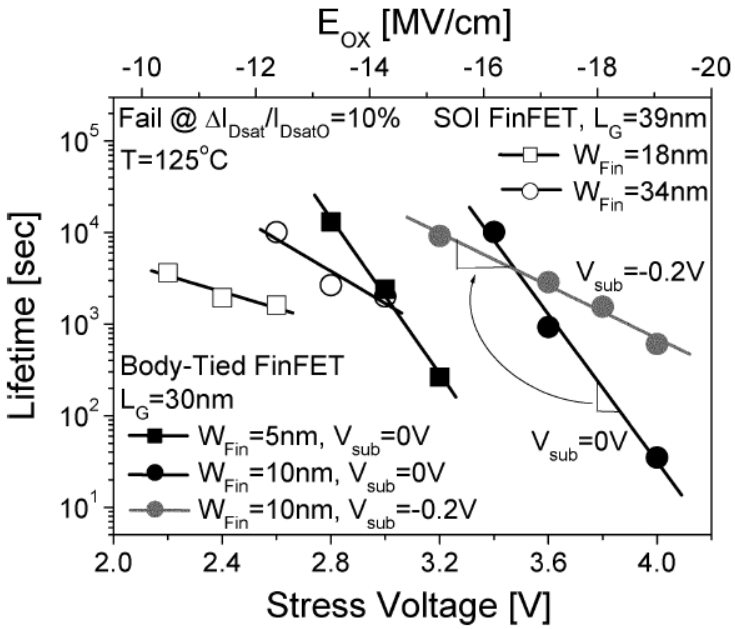


Fig.VI.9: Device lifetime versus stress voltages and E_{ox} of 18 and 34 nm fin-width SOI FinFETs (blank symbols) as well as 5- and 10-nm fin-width body-tied FinFETs (filled symbols) at 125 C substrate temperature. The gray symbol represents the body-tied FinFET with 0.2-V substrate bias. With minus substrate bias, the quasi-floated body effect is shown. [82]

Finally, many types of FinFET structures exist and rend the parameters very strongly technology dependent. Results from Lee [82]. Fig. VI.9 shows the device lifetime as a function of stress voltages and of 18 and 34 nm fin-width SOI FinFETs as well as 5 and 10nm fin-width body-tied FinFETs at 125° C substrate temperature. Both SOI and body-tied FinFETs show that the lifetime is more seriously degraded with a narrower fin. The slope of the predicted body-tied FinFET lifetime is steeper than that of the SOI FinFET. This implies that the body tied FinFET used in DDR3 is more robust to NBT stress than is the SOI FinFET. A body-tied FinFET shows a smaller NBTI degradation that becomes more effective due to the presence of grounded substrate.

With the principle of Quasi-Static (QS) transfer [99], inverter waveform can be used in Fig.VI.10 to include the Off-mode phases as it represents a general form of digital operation. With the QS scheme, the results of previous sections have evidenced the high V_{DG} conditions required to observe Off mode degradation in contrast to CHC damage which needs smaller V_{DS} and much reduced V_{GD} in MOSFETs [105] [106]. Hence Off-mode damage can be included in the full modeling for digital operation. This is performed by completing the 3 mode On-state lifetime (τ_{ON}) developed here for NMOS and PMOS by the Off-state contributions (τ_{OFF}) in two phases $1/\tau_{Off} = 1/\tau_{VG=0} + 1/\tau_{SubVT}$ useful for any pulse waveform shapes. Device lifetime extraction (τ) is then deduced for all stressing conditions by assuming that all degradation modes compete in parallel leading to a general device lifetime modeling expressed by:

$$R_{it} = \frac{1}{\tau} = \sum_i \frac{1}{\tau_{on,i}} + \sum_j \frac{1}{\tau_{OFF,j}} \quad \text{Eq. V.14}$$

$$\frac{1}{\tau_{ON}} = C_1 \cdot \left(\frac{I_{ds}}{W}\right)^{a_1} \cdot \left(\frac{I_{bs}}{I_{ds}}\right)^{m_1} + C_2 \cdot \left(\frac{I_{ds}}{W}\right)^{a_2} \cdot \left(\frac{I_{bs}}{I_{ds}}\right)^{m_2} + C_3 \cdot V_{DS}^{a_3/2} \cdot \left(\frac{I_{ds}}{W}\right)^{a_3} \cdot e^{-\frac{E_{emi}}{kT}} \quad \text{Eq. V.15.a}$$

$$\frac{1}{\tau_{OFF}} = C_4 \cdot \left(\frac{I_{Sub}}{W}\right)^{m_2} + C_5 \cdot \left(\frac{I_{ds}}{W}\right) \cdot \left(\frac{I_{Sub}}{I_{ds}}\right)^{m_3} \quad \text{Eq. V.15.b}$$

where C1 (SVE), C2 (EES) and C3 (MVE), C4(Off-state), C5(Sub-V_T) are all the subsequent constants determined in their respective dominant mode. This modeling has been validated for various Tox= 5nm, 3.2nm and 1.7nm, under a large set of voltages conditions (VGS, VDS), temperatures and device geometries in both NMOSFET and IO PMOSFET.

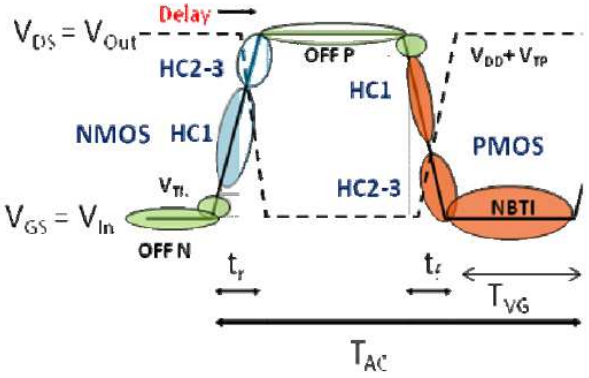


Fig.VI.10: Inverter-like pulse waveform used for the DC-AC transfer of each accelerating degradation rates in three mode energy modeling devoted to core and IO CMOS nodes. The Off-mode phases are included both in NMOS and PMOS for the use of QS device lifetime extrapolation [99].

4) Risk analysis method applied to DDR3 – FF20 (Manufacturer A).

This Part copes with the memory analysis proposed by Manufacturer A for data storage as a SDRAM circuit processed on a 20nm FinFET (FF20) technology. The first section aims to determine the sensitive parameters that have an accelerating effect on Wearout and Failure modes. Section 2 will propose the extraction of a generic structure from CA (see Fig.VI.1 step 1B) at the front end while Section 3 concerns the lifetime determination using the chosen accelerated failure mechanisms (Fig.VI.1 Steps 2A-2B) according to the referred modeling from literature and the mission profile (MP) detailed in part 1. Finally, a strategy of data exploitation will be given in section 4 with respect to AC operation.

a) Section 1: Extraction of the sensitive Technology parameters

This technology is developed from recent FinFET 20nm and 16nm processing developments, but they have originally distinct application scheme. Thus, it is necessary to consider the main differences about the general application type of the mentioned component and extract the main parameters and operation conditions given by the manufacturer that are involved in the Wearout damage. That will offer the fundamentals parameters required for the lifetime modeling.

DD3R-FF20 General Temperature and Supply Voltage Characteristics

The DDR3 is a Synchronized dynamic random-access memory (DRAM) by an external clock so the central operation type is related to DRAM semiconductor memory functioning that can run at faster speeds than conventional DRAM. So the synchronization is done at double rate linked to the leading (CK) and falling edge (CK#) of the clock cycle, where this memory type is volatile which means that when once the power is removed, or a supply voltage brought by degraded metallization is fallen down, it loses all data.

The component DDR3 is referred from Manufacturer A and means part under analysis has a data rate referred as 1866 for a density of 4 Gb (4x1Gb) and the memory is in fact a DDR3L SDRAM with 256Mb x 16. The package contains 96 balls 8x14 mm FBGA.

The aimed applications are automotive market with temperature range -40°C ≤ TC ≤ 120°C and for industrial (commercial) market with TC = -40°C to 95°C. TC is the maximum temperature measured at the center of the package. The junction temperature is given by the data sheet to be (Automotive) as [a 4Gb_DDR3L]:

$T_{Junction}$ (°C)	SLOW	TYPICAL	FAST
Commercial	110	50	-40
Automotive	120	50	-40
Automotive UH	140	50	-40

The timing cycle time for SDRAM 1866 is T_{AC} = 1.07ns with programmable CAS (READ) latency CL = 13 corresponding to a frequency F_{AC} = 934 MHz.

The given supply voltages are V_{nom} = 1.35V with V_{DDmin} = 1.283V and V_{DDmax} = 1.425V which is backward compatible to V_{DD}= 1.5V ± 75mV application with DDR3 SDRAM. The same supply voltages are supplied for internal logic (V_{DD}) and for the data bus IO pin drivers (V_{DDQ}).

This allows to extract the spice netlists for automotive constraints as [4Gb_DDR3L]:

LIB	Data rate	Process	Die Cap	V_{DD}	Temperature
*SLOW_1066_AAT	1066-1866 Mbps	Weak	Maximum	1.283 V	120°C
***TYP_1066_AAT**	**1066-1866 Mbps**	**Typical**	**Typical**	**1.350V**	**50°C**
*FAST_1066_AAT	1066-1866 Mbps	Strong	Minimum	1.425V	-40°C

The main block to be considered is the memory arrays with sense amplifier and with IO gating with mask logics and column decoder. As we want to pick up a generic structure representative of DDR3 unitary cell, we first have to extract the main elements helped by CA of the double data rate synchronous RAMs (burst mode) which achieve high-speed operation.

DD3R-FF20 Structural Characteristics

The General way of operation of DRAM block to DRAM cell is illustrated in Fig.VI.11 where a row decoder selects the word line (WL) x_i of the row in which the addressed word (or cell) is located. The row decoder is also known as an x-decoder. A column decoder selects the bit line (BL) select line y_j of the column, in which the addressed word (or cell) is located. Address buffers connected to the address inputs drive the row and column decoders. A sense amplifier detects the contents of the selected cell via the complementary bit lines b_j and its complement, and data bus lines and its complement.

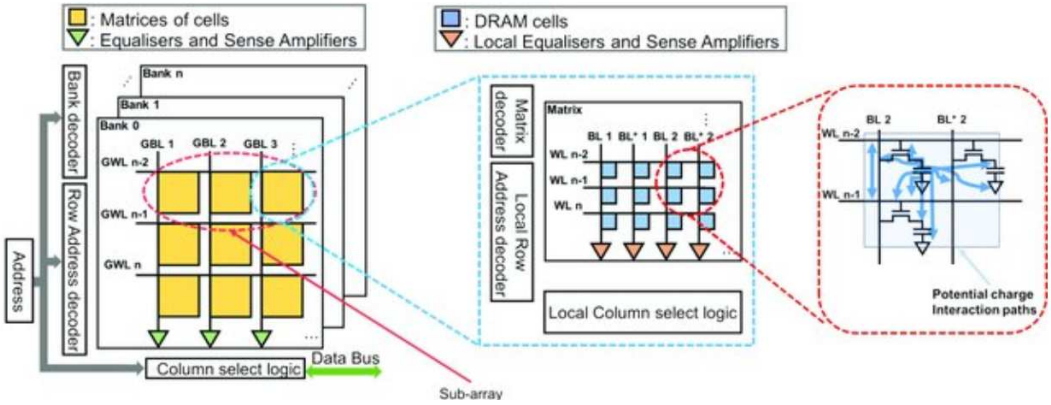


Fig. VI.11: Illustration of the DRAM block as a matrix of elementary cells and the potential charge interaction between storage capacitors and active transistor.

DRAM cell is generally fabricated with trench capacitors in planar technology but rather with stacked (3D) capacitors for higher density RAM. Here, DDR3-FF20 component is composed by an innovative three-dimensional cell structure such as trench and stacked which makes its realization possible with much higher density DRAM cells up to 4 Gb. Either trench or stacked structure, capacitance had to be boosted without increasing projected area (or even decreasing the area) with increasing DRAM density.

From CA report it is important to identify and distinguish the different element:

- Vertical capacitors are localized between M2 and Poly layers. Capacitor top electrodes contact each other to form a common top plate. This top plate is composed of Ti.
- Below this common top plate, vertical capacitors seem to have a doubled structure. They are formed by successive concentric tubes of SiTi top electrode (inner tube), ZrO dielectric, Ti intermediate electrode (intermediate tube), ZrO_2 dielectric, and Ti bottom electrode (outside tube). The intermediate electrode is connected to a single active area. A de-processing has been performed to confirm this doubled structure capacitor.
- Inside memory table, metal M1 cubes are used to connect capacitor intermediate electrodes to transistor active areas. Metal M1 cubes are made with W deposited on Co (W/Co).
- Dielectric ILD0 is a SiO single layer (SiO) supposed assimilated to SiO_2
- Poly layer is localized below capacitors. Poly lines are covered by two silicides (SiW on SiTi) and a SiNC cap. Inside memory table, the minimum (drawn) Poly length is measured at $L_G \approx$ 20nm.
- Inside memory table, Poly cubes are used to connect capacitor intermediates electrodes to transistor active areas.
- Inside memory table, an unusual metal M0 layer is present below Poly layer. Metal M0 is made with W deposited on a bottom Ti barrier.

Generic DRAM cell for DDR3L-FF20

From the CA analysis, the DRAM cell exhibits a 3D structure with capacitance storage on top and 3D transistors at the bottom (Fig. VI.12) controlled by M0 lines which drive bit line (BL) on their drains, and Poly lines that drive word line (WL) on their gates. The example Fig. VI.11 gives a DRAM cell but here with trench capacitor storage. Beyond 50 nm only stacked-capacitor memories are manufactured. Hence, in our case, the capacitor seems to be a double-sided cylinder capacitor with a Ti top plate M2 electrode and a Ti bottom electrode to Poly gate. This has to be examined in detail as High-K dielectric (ZrO_2) is used in order to increase the stored charge capability.

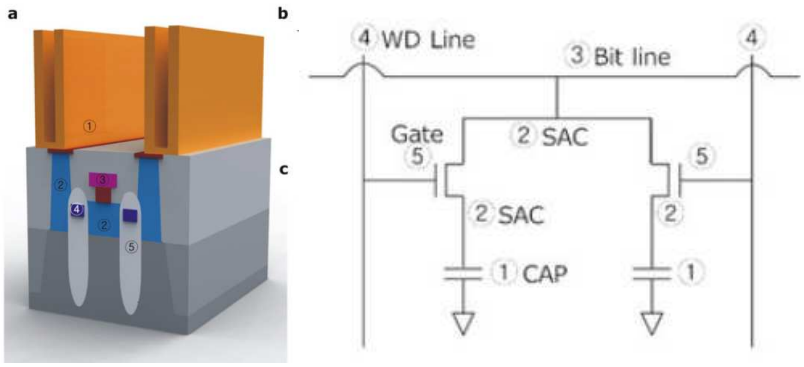


Fig. VI.12: Analysis in DRAM cell structure. (a) The schematics of a DRAM cell with the storage capacitors (Capacitor SAC Bit line Word line Gate). (b) Equivalent circuit of the DRAM cell structure, (c) The sample structure and scheme of the DRAM structure [107].

Vertical storage capacity

In order to produce a dynamic random access memory 20 nm DRAM or less, the most important concern regarding operation is to reduce the leakage current degradation of the capacitor using high-k dielectrics eliminated by removing external impurities of boron and hydrogen without any change in the structure or materials of the capacitor. What seems to be clear is that the storage capacitor (Fig. VI.12).

2) exhibits a double-sided capacitor at that dimension. Stacked capacitor DRAMs are built from twin cells, meaning that two cells share the same bit line contacts. Trench-capacitor areas (8 F^2/bit) are therefore usually larger in area than stacked-capacitor cells (6 F^2/bit), with F being the half pitch (in our case F = 21 nm). Because of the use of vertical pillar transistors (VPT) (drain stacked on top of source Fig. VI.12, separated by the channel length) a single memory cell area is expected to reduce to 4 F^2/bit.

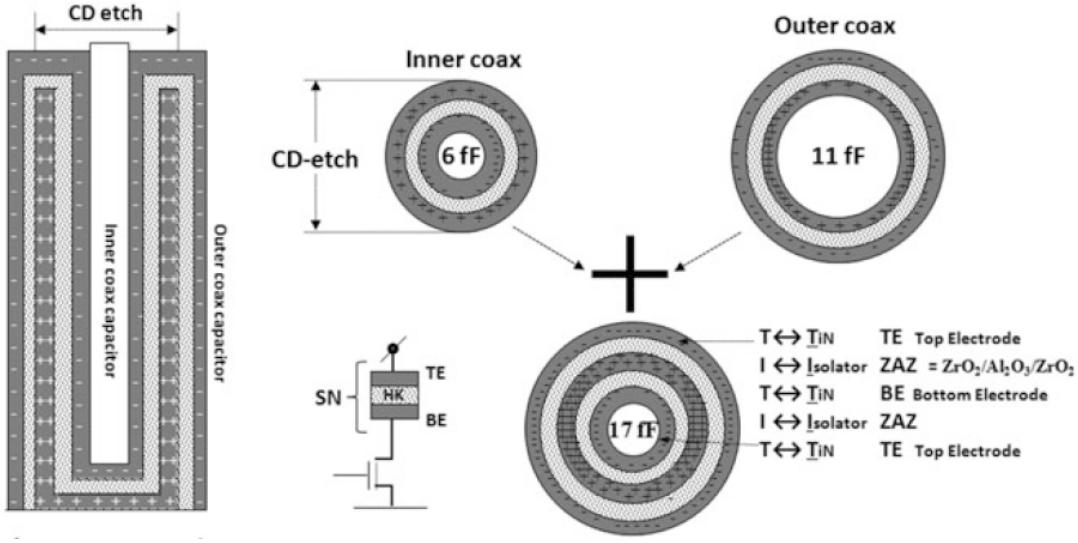


Fig. VI.13: Double-sided charge capacitor referred as crown capacitor for DRAM cells [[108] source Vreugdenhil; ASML semiconductors]

We show in Fig.VI.13 an example which was used (in 2013) in a 25nm DRAM technology [108].This capacitor represents a capacitance value of 17 fF, which is built from two cylinders. The inner and outer cylinder represent respectively 6 *fF* and 11 *fF* of storage capacitance. This tri-layer High-K dielectric ZAZ (Fig. 4) has been processed as a stack of Zirconium - Aluminum - Zirconium oxides giving a thickness td of around 9 nm. We have the same kind of structure cell in DDR3, excepted that only one material is used for the High-K.

From CA extraction there is only a single dielectric material ZrO(2) identified as the High-K dielectric:

- we have a height H= 882nm and Si mesh of height h= 615nm
- inner coax tube is in SiTi
- High-K in ZrO_2 with dielectric constant ε_R= 40 (to 45) for T_{ox}= 6 (to 9nm), respectively

This simplifies the structure in DDR3L as:

- the inner coax Si/Ti with thickness 18.42nm (first diameter= $2R_1$)
- the successive tubes of ZrO / Ti / ZrO with thicknesses 6.14nm / 6.14 nm / 6.70nm
- the inner electrode and the outside electrode are connected to the common top plate.

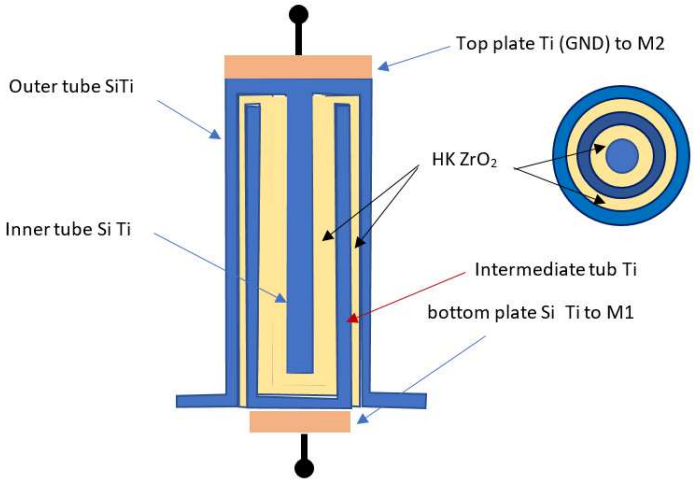


Fig. VI.14: Schematics of the double-sided capacitor for DDR3 – FF20.

This tubular cell capacitance is known as much more robust in terms of process variability. One disadvantage of the cell is its larger capacitor area, but its aspect ratio is 3.5 times smaller, allowing faster DRAM operation. This crown capacitor cell is used since 2015 for the 20 nm node DRAM cell in DDR3 and DDR3L.

For our HK thickness range, the dielectric constant of zirconium oxide (ZrO2) is around 45 to 35 depending on its amorphous phase, or between 30 and 20 as a crystalline phase when it is mixed as in ZAZ. Below a thickness of 5 nm its value dramatically reduces with thickness. CA gives a 6.1nm to 6.7nm thickness. We will choose a medium value ε_R = 40.

The cylinder capacitance is $C = 2\pi.\ \varepsilon_o.\varepsilon_R\ .H\ /\ Ln(R_2/R_1)$ taking into account ε_R = 40, the height H and R_1 and R_2 = 15.3 nm measured from the CA. This gives C_1 = 1.96 *fF* for the inner capacitance. For the outer capacitance we have C_2 = 7.23 *fF*. Then, for the double charge (parallel) capacitance, we have: C_{tot} = 9.19 *fF*.

The bottom of the storage capacitor is connected to the cell transistors thanks to M1 cubes. One observe ILD1 in SiO2, M1 level with WCo, then silicide and Poly/SiO then M0/Gate Oxide and active + STI in the shape as "Saddle FinFET". This means that the 3D FF has been rotated by 180° giving an inverse gate control with respect to the Fin oriented between S/D.

3D cell analysis of DDR3L

In the DDR3L mode of operation, DDR3 bus rates is given at 1.5GHz (800–1600MHz) at a maximum of 1.5V. The bandwidth of DDR3 is doubled compared to DDR2, with no major speed change of the DRAM core, i.e. the memory cell array. To achieve this double bandwidth, DDR3 uses a prefetch of eight words instead of a four-word prefetch for DDR2. This means that for every single read or write operation, a total of eight words are accessed in parallel (shown by the parallel Fins) in the DRAM core to enable the high data rate at the interface.

Thanks to CA study, the 3D core cell has been identified showing that this technology uses FinFETs with Buried Word Line. The active in the cell matrix area is disorientated of 30°shift vs. Bit Line and Word Line. Each active block is isolated from the adjacent one with STI oxide.

On each active island, two Buried Word Lines are observed, two capacitor contacts and a single central Bit Line contact Pictures as shown in the schematic Fig. VI.15.

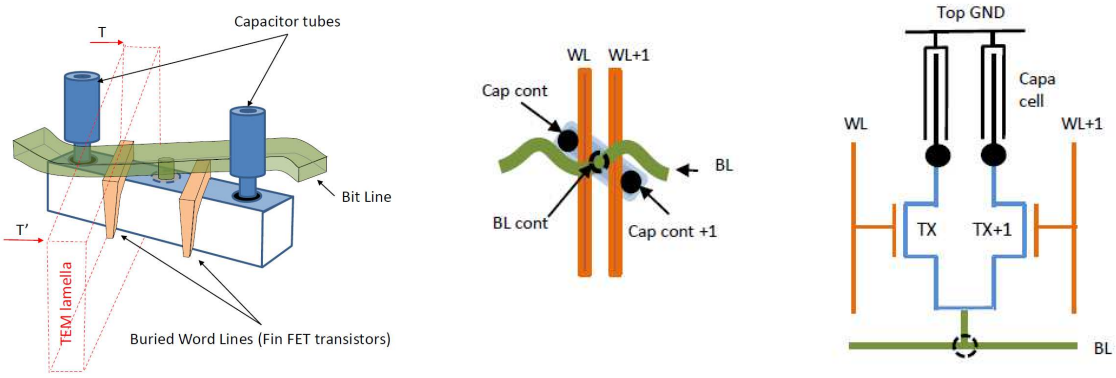


Fig.VI.15: Schematics of the cell structure synthesis obtained from TEM lamella axis [Report RCA_0123] showing the double sided capacitors and paired transistors connected to WL and WL+1.

With decreasing DRAM cell size, a recessed channel array transistor (RCAT) has been proposed to overcome the short channel effect (SCE) of conventional MOSFETs with planar channels. To use the fabrication process simplicity of planar cell array transistor, a vertical transistor, RCAT structure was given in stack capacitor cells. The basic concept of the RCAT is to increase the effective length (L_{eff}) by recessing the channel from silicon surface. So, channel doping density can be reduced, thereby reducing S/D resistance of memory cell transistor, and enhancing carrier mobility, in addition to the main advantage of reduced leakage.

Although the recessed channel of RCAT has improved short channel effect (SCE), RCAT suffers from low driving current and V_{TH} sensitivity due to the shape of the bottom corner of the recessed channel [107]. To solve these problems, a saddle FinFET (S-FinFET) has been proposed with a tri-gate that wraps both the recessed channel surface and the side surface [109], [110] as shown Fig. VI.16. S-FinFET not only exhibits excellent short channel effect characteristics, but also maintains excellent subthreshold swing (SS), high Ion, and nearly constant V_{TH}.

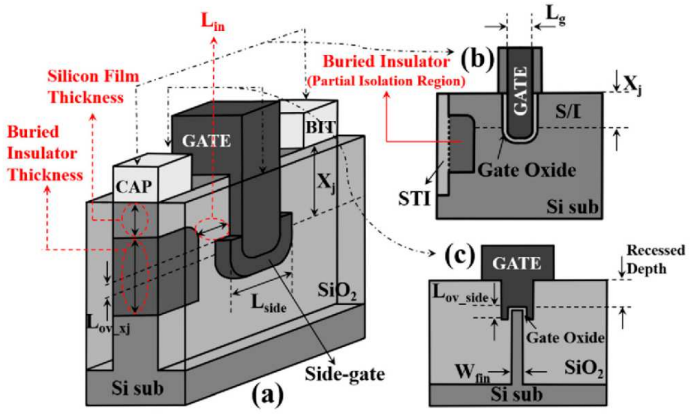


Fig.VI.16: (a) 3D schematic of Pi-FinFET; (b) Cross-sectional view across the gate; (c) Cross-sectional view of the thin body.

The gate wraps three surfaces of the recessed channel, similarly to a FinFET. The buried insulator material is used with SiO_2 below the storage node. The buried insulator is penetrated from the shallow trench isolation (STI) region. X_j of the

source/drain (S/D) is located 112 nm from the top surface of the S/D region with a Gaussian profile. The peak concentration of the Gaussian S/D doping profile is $1.5 \cdot 10^{20}$ cm^{-3}, the uniform body doping concentration is $5 \cdot 10^{17}$ cm^{-3}. [109].

Extraction of the Generic structure for DD3RL

The complete analysis of the SDRAM array shows that the full cell matrix is very complex in design and processing at that level of integration and memory density. A lot of uncertainties still remain although the CA study has brought a lot of improvements and helps in determination of the nature of materials and dimensions. This leads us to simplify the approach to a generic (simpler) digital cell based on the CA dimensioning. As the objective is to consider a generic structural device at which a Reliability methodology can be applied, leading to the determination of an "intrinsic reliability" examination with failure rate λ_G, we shall consider that the cell is composed by paired N-channel transistor very similar to FinFETs wrapped around the Recessed channel, with buried WL connected to the gates of the FinFETs and BL connected to the (common) drain to reach the required drivability of the cell.

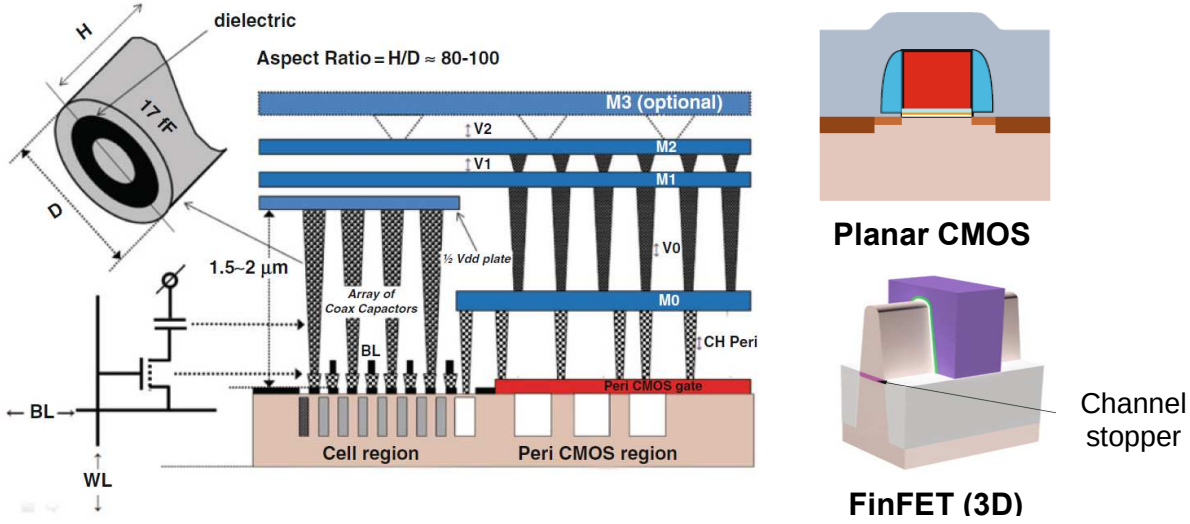


Fig.VI.17: Equivalent 3D cylindrical stacked memory cell in 22nm DRAM technology proposed here for DDR3L SDRAM and extractions of 3D FinFET (DRAM) and planar (Peri) CMOS [108].

Fig.VI.17 shows a schematics of matrix 3D cell of the DRAM with paired transistors and double-sided cylindrical (coax) capacitors (22nm node). In periphery CMOS region (interface area) we have inter-block area between M1 in inter block and M0 in memory cells, with standard planar transistors (isolated by STI) with poly gates and thicker gate oxide (T_{ox} = 2.8nm).

Consequently, we propose that for DDR3L, the generic structures can be chosen as a memory point composed of N-channel FinFET with its advanced processing (RCAT, recessed channel, saddle FinFET) and storage capacitance, and the other case of standard planar transistors. This will provide two points of examinations:

- One with very short channel representing the worst-case reliability with respect to FEOL at the length of the technology node (T_{op} part L_{eff}= 16*nm*).
- Another one with relaxed design rules for Periphery CMOS with planar process (poly gate and STI isolation) at longer device length (L_{eff}= 102*nm*).

The CA study parameters can be summarized in the table below with a summary of the physical process parameters obtained for memory cell and peri standard CMOS:

Device type	**WxL drawn**	**L_{nom}**	**L_{eff}**	**T_{ox} = EOT**	**V_{DD} (V) min/ Typ. / Max**
FinFET in SDRAM (3D)	33x37nm	**20nm**	**16nm**	2.8nm	1.283 / **1.35** / 1.425 (1.5V)

Std Peri CMOS (Planar)	NA	**130nm**	**102nm**	1.5nm	1.283 / **1.35** / 1.425 (1.5V)

Note that for the FinFET, we can check the rule in width where $W_{eff} = 2\,H_{Fin} + W_{Fin}$ to determine the effective channel width with respect to the drawn width, and if we do observe multi Fins in parallel into the pictures, one have to account for an effective width as $W_{eff} = [2\,H_{Fin} + W_{Fin}]\cdot N_{Fin}$ that would give for a supposed $H_{Fin} = 50nm$:

$$\boldsymbol{W_{eff}} = [2{\cdot}50nm + 33nm] = \boldsymbol{133nm} \text{ per fin.}$$

Reliability study for DD3RL

We study in this section the SDRAM Reliability under the worst case of aging, that we analyze through the Reliability under TDDB, HC and NBTI in a first step and the CA reports about that device. We have to use the FinFET dimensions for DRAM cell at nominal L_{nom} and a single effective W_{eff} , i.e. for one Fin width, extracted from the DPA for core N-channel FinFET of the memory cell.

TDDB Reliability

The TDDB Reliability of the core cell can be determined by the 22nm results from the literature on similar FinFET from other foundries (INTEL) where the dimensions are very close to DDR3 dimensions [93] (fin Height *H~34nm*, the fin width *Wsi~8nm*, the gate length is drawn at 30nm, and gate pitch 90nm, HKMG stack). Ramey indicates that the threshold voltage of FF is $V_{TH} = V_{TH}$(planar) – 0.1V in that case. The TDDB modeling is done by the E-model which is the most restrictive. The exponential dependence gives the following parameter usually required for TDDB, but the lack of a reference area (A_G) and TDDB distribution in the published results [93] avoids the AF_{AG} factor evaluation. As we know the area scaling relationship if one knows the beta value of the distribution giving:

$$TTF_{BD1}/TTF_{BD2} = \left(A_{G2}/A_{G1}\right)^{1/\beta} \qquad \text{Eq. V.16}$$

where the Weibull parameter β, range from 1.1 to 1.7.

So with this assumption, we found that using for DDR3 FF with $L_G = 20nm$ at typical 1.35V and maximum 1.5V operation, the direct extrapolation from the results Fig.VI.4 give: $TTF_{BD} \propto A_i \,.\, exp(-\gamma V_G) \,.\, exp(-E_a/kT)$.

$\tau_{TDDB} \propto \exp(-\gamma E)$ model	NMOS	PMOS
V_{DD}= 1.35 V	5.22 10^9 s = 165 years	6.65 10^{11}s = 2.1 10^4 years
V_{DD} = 1.5 V	2.58 10^7 s = 0.81 years	4.21 10^{10} = 1.33 10^3 years

This shows (with the afore mentioned assumption) that the DDR3 at typical value of 1.35V is well above the lifetime criterion for PMOS with no TDDB sensitivity, but the NMOS exhibits a larger acceleration factor with V_{DD} that puts the boost condition at $V_{DDmax} = 1.5V$ the most sensitive factor with respect to TDDB.

HC Reliability

As stated before, the difficulty in extrapolation lifetime data is that we don't know yet the process dependent constant value C_i which differs from batch to batch and from silicon bulk to FinFET CMOS node. What can be done first, is to determine the lifetime value from I_{On} performance from the single transistor or from a block of (idealized) identical geometry as because I_{On}

current per device is much smaller in this FF 20nm process, that has to be analyzed for its typical nominal supply voltage V_{DD} = 1.35V and (a) for EOT= 2.8 nm and for (b) L_G= 130nm (L_{eff} = 102nm) for EOT= 1.5nm.

(a) Short channel L_G = 20nm from the memory point

A first direct evaluation could be done by considering the average current delivered by the memory block DDR3 – 1866 giving 12mA / 16 banks = 750µA per bank as row addressing denoted by A[n:0] is for 4Gb: n = 14 (x16).

This leads to a level of nominal current at I_{On}= *750µA/µm* for lifetime calculation (Fig. VI.7) as $\tau_{HC} = C_i\,(V_{DS}^{0.5} \cdot I_{DS})^{-24.3}$ which provides the relationship for C_i calculation This may be related directly to the experimental lifetime plot of the (close) 14nm FF node giving a DC lifetime [100] as a function of :

$$I_{On}\,(\mu A/\mu m) = ID_{Sat}\,(V_{DD}, V_{DD}) / W_{eff} \qquad \text{Eq. V.17}$$

This makes the nominal I_{On} as the reference value giving $I_{DS} = 99.75\ \mu A \cong 100\ \mu A$. Due to the quadratic dependence on $(V_{GS} - V_{TH})^2$, the value of $I_{DS}(V_{DD2})$ can be obtained assuming identical V_{TH} increasing V_{DS} (with negligible Drain Induced Barrier Lowering effect if V_{DD} is increased by 150mV) as:

$$I_{DSat1}(V_{DD1}) / I_{DSat2}\,(V_{DD2}) = (V_{DD1}\text{-}V_{TH1})^2 / (V_{DD2}\text{-}V_{TH2})^2 \approx (V_{DD1}/V_{DD2})^2 \text{ , giving :}$$

$$\boldsymbol{I_{DSat2} \approx I_{DSat1}\,(V_{DD1})\,(V_{DD2}/V_{DD1})^2} = 1.291\ I_{DSat1}\,(V_{DD1})$$

thus leading to a 30% increase,

W_{eff} = 0.133 µm **per fin** I_{On}= 750µA/µm, V_{DD1}= 1.35V	$I_{DSat}\,(V_{DD}) \cdot V_{DD}^{0.5}$	**Lifetime (s)**
ID_{Sat1} = 100 µA, V_{DD1}= 1.35V	$1.16 \cdot 10^{-4}$	$1.33 \cdot 10^{17}$
ID_{Sat2} = 129µA V_{DD2}= 1.50 V	$1.58 \cdot 10^{-4}$	$3.24 \cdot 10^{14}$

This very large lifetime value for DC serves indicates that the HC do not seem to be an issue at such low voltage and small dimensions.

If now one considers two and four fins width in parallel for example in an inverter, results will be quickly changed but the power law as I_{On} *(1 Fin)* ⇨ $I_{DSat} = I_{On} \cdot W_{eff} \cdot N_{Fin}$ translate the results with respect to the x-coordinate:

W_{eff} = 0.133 µm, ***Inverter x2Fins*** I_{On}= 750µA/µm, V_{DD1}= 1.35V	$I_{DSat}\,(V_{DD}) \cdot V_{DD}^{0.5}$	***Lifetime (s)***
ID_{Sat1} = 200 µA, V_{DD1}= 1.35V	$2.320 \cdot 10^{-4}$	$1.46 \cdot 10^{11}$
ID_{Sat2} = 258µA V_{DD2}= 1.50 V	$3.159 \cdot 10^{-4}$	$2.662 \cdot 10^{8}$

W_{eff} = 0.133 µm, ***Inverter x4 Fins*** I_{On}= 750µA/µm, V_{DD1}= 1.35V	$I_{DSat}\,(V_{DD}) \times V_{DD}^{0.5}$	***Lifetime (s)***
ID_{Sat1} = 400 µA, V_{DD1}= 1.35V	$4.647 \cdot 10^{-4}$	$1.03 \cdot 10^{5}$
ID_{Sat2} = 516 µA V_{DD2}= 1.50 V	$6.319 \cdot 10^{-4}$	$2.46 \cdot 10^{2}$

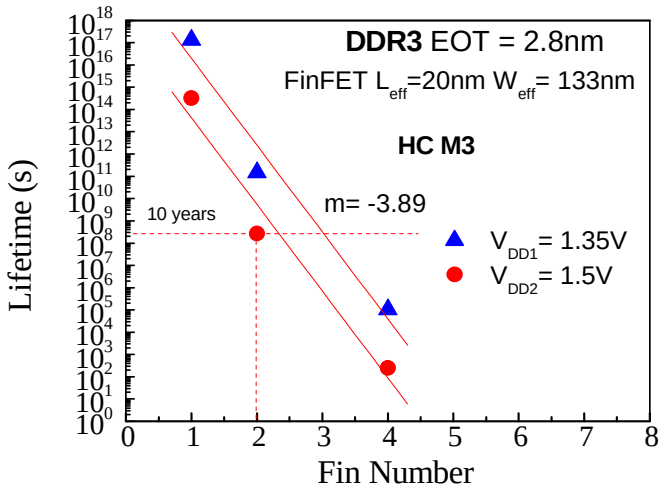


Fig.VI.18: Summary of the HC (DC) Lifetime in DDR3 as a function of the number of Fins for NMOS inverter side for equivalent N-channel FinFET transistor.

Then it is clear that the FF 20nm technology is limited in the number of Fin for the current drivability optimization as the lifetime strongly decreases as a function of 2 Fins per device for $V_{DD}=1.5V$ (< 10 years) and for both supply voltage if one put 4 Fins in parallel.

As these last cases would correspond to the case of a digital inverter it can be translated to the real AC operation condition, taking into account the time for pull-up (and pull down) of the equivalent stage as a function of Fan Out (FO).

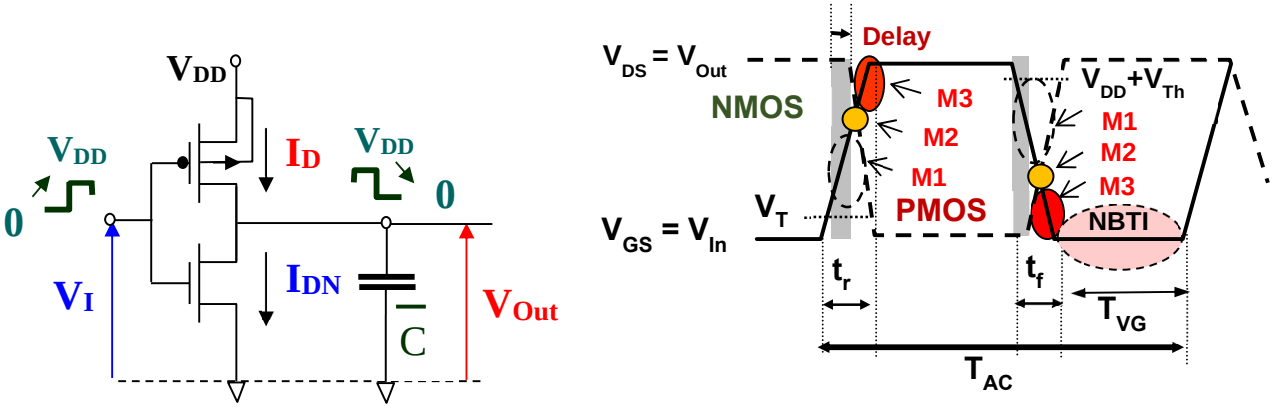
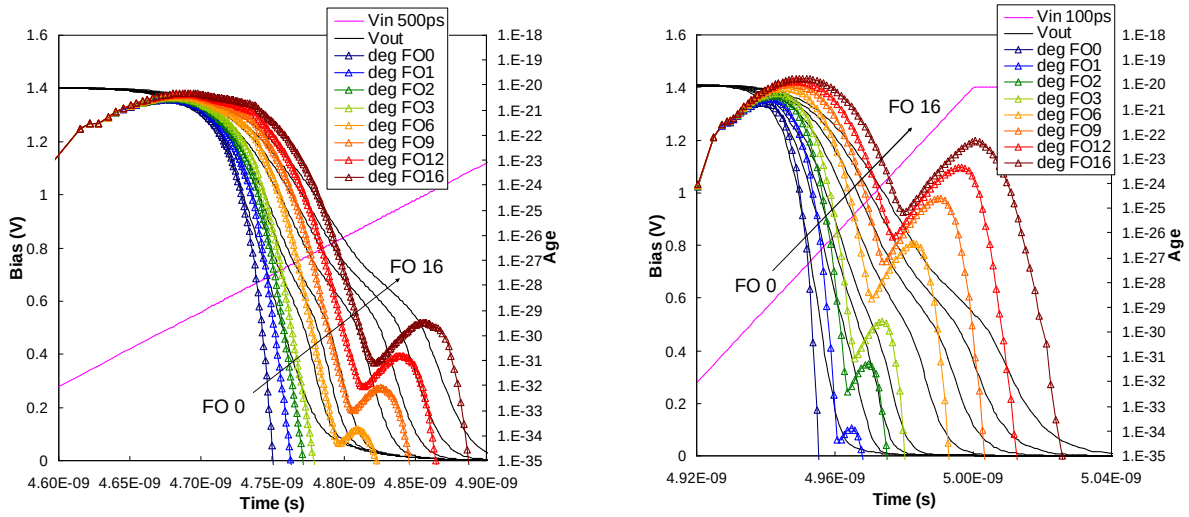


Fig. VI.19: (a) CMOS inverter operation schematics as a function of V_{In}, V_{Out} and its load capacitance CL (b) Signals applied to the inverter $V_{Out} = f(V_{In})$ related to the distinct bias region where HC to CC are involved (i) for NMOS transistor where $V_{IN}= V_{GS,N}$ and $V_{OUT} = V_{DS,N}$ while (ii) for PMOS transistor where we have $V_{GS,P} = V_{IN} - V_{DD}$ et $V_{DS,P} = V_{OUT} - V_{DD}$ due to the inverted potential reference for the PMOS source with $V_{S,P}= V_B = V_{DD}$. [111]



(a) (b)

Fig. VI.20: Function age defined as *Age(t)* = t_s / τ in *40nm* NMOS inverter as a function of FO= 1 to 16 (a) For t_r= *500ps* (b) for t_r= *100ps* (c) NAF, PAF (Full, red) vs. t_r with FO variable in core inverters showing a NAF (and PAF) fall down with t_r. [111]

M3 mode effect is enhanced during the transient through the Age increase when tr is reduced from 500ps to 100ps. A similar AC behavior has been demonstrated in PMOSFETs [111] as both CHC N- and P- MOSFETs worst-case degradation has moved to the high V_{GD} region. This transfer to our AC condition allows to define HC design guidelines due to the assignments of main degradation in M2 to M3 modes, that can be merged to M3 only for low voltage operation. The reduction of *NAF* (*PAF*) factors which are defined here as a function of AC-DC time factors *NAF= NTF·*f_{AC} with *NTF* = x_n / (f_{AC}.t_r) where x_n is the proportion of the CHC to CCC damage in the rising transient, i.e. with $T_{hot,n}$ = t_r / x_n and $T_{hot,p}$ = t_f / x_p with the corresponding relationships looking at Fig. VI.19b:

$$TF = \frac{T_{AC}}{T_{Hot}} = \frac{1}{f_{AC}\left(t_{r,f}/X_{N,P}\right)} \Leftrightarrow NTF = \frac{X_N}{f_{AC}\cdot t_r} \; and \; PTF = \frac{X_P}{f_{AC}\cdot t_r} \qquad \text{Eq. V.18.a}$$

$$T_{Hot}(NMOS) = \frac{t_r}{X_N} \; and \; T_{Hot}(PMOS) = \frac{t_r}{X_P} \qquad \text{Eq. V.19.b}$$

which are defined at shorter $t_{r,f}$ and larger Fan Out (FO), in relation to the larger impact of M3 mode showing a trade-off to optimize between frequency f_{AC}, t_r, t_f and the FO.

We can simply M3 duration to the input signal time between V_{TH} and *0.9·*V_{DD} over *x* factor that can be obtained as for FO= 1 we have $t_{r,f}$ = *0.03·*T_{AC} = 32 *ps*, in order to be closer to real AC stressing configurations at the device level. Then the effective (M3) damaging mode could be simplified to *0.9·*$t_{r,f}$*/2* duration enables to calculate the NTF (DC-AC) factor giving:

$$\text{NTF} = 1.07\text{ns} / (0.9 \cdot 32\text{ps} /2) = 74.30 \cong 75 \qquad \text{for typical estimate}$$

This consequently relax the AC lifetime for 2 Fins case, at V_{DD}= *1.5V* which becomes higher than the lifetime criterion (10 years) as we may apply:

$$\tau_{AC} = \text{NTF} \cdot \tau_{DC}$$

But the 4 Fins case remains below the 10 years lifetime criterion independently of V_{DD}.

(b) Long channel LG = 130 nm from standard Peri CMOS

Same approach is applied to DC lifetime evaluation in L_G= *130nm* (L_{eff}= *102nm*, with thinner gate oxide *EOT = 1.5nm*) that can be considered as a "long channel" behavior with respect to the one with L_{eff} = *16nm* if we know the current value.

In that case the Lifetime slope exponent Fig. VI.8 reduces to -20 till -17.9 [100] but would deviate from the M3 mode due to long channel dependence and smaller HC damage, even for this thin EOT. Then, as we do not know neither I_{On} (µA/µm) nor channel width for the Peri CMOS, we can apply a standard Lifetime technique coming from the classical HC in order to include the accelerating L_G dependence which is sensitive for that Long channel case, as well as for IO devices which operates at higher voltages.

Indeed assuming a time power law for the reduction of the I_{On} current (Δ_{IDSat} *= 10% FWD=* Δ) the normalized failure criterion follows $\Delta_{IDSat} = A\, t_s^n$ where exponent n*= 0.4,* we can express the Lifetime :

$$\tau_{HC,\,Long} = C_i\,(\Delta / 0.1)^{1/n} \cdot L_G^{m} \cdot \exp(B(1/V_{DS} - 1/V_{DD}))$$

Calculated for L_G*= 130nm, m= 10.7, Ci=* $1.81\cdot10^5$ *years =* $5.708\cdot10^{12}$ *s, B = 80 (Long channel),* V_{DD}*= 1.35V* (Typical) to maximum V_{DS}*= 1.5V*, giving :

$$\tau_{HC} = \mathbf{5.03\ years}\ \text{in DC}$$

As a consequence, the pull-up of a block (device) designed with the FF 20nm DDR3 would result in an AC lifetime at :

$$\tau_{AC} = NTF \cdot \tau_{DC} = 75 \cdot 5.03\ years = \mathbf{377\ years}\ \mathbf{(AC)}$$

This shows that with a lot of assumptions, the standard lifetime (AC) value might satisfy in any case the 10 years lifetime for inter block standard NMOS transistor of the SDRAM cell at V_{DD} *= 1.35V to 1.5V* (Typical to maximum value). With the same assumed frequency for operation and shape of the pulse for FO = 1, this leads to AC lifetime for L_G*= 130nm* at 1.35V, τ_{HC} *= 377* years in AC (ΔI_{DSat} *= 10%*).

NBT Reliability

Very few results show quantitative TTF under NBT mode for FinFETs. Most of results are on SOI FinFETs. As we have body-tied FinFETs for DDR3 component, we can use the acceleration modeling presented in section B.3 (Fig. VI.8) with the normalization trick with respect to the lifetime criterion (ΔV_{TH} *= 50mV*). Using eq. VI.9:

$$TTF_{NBT} = A_i \cdot \left(\frac{\Delta}{0.05}\right)^{1/n} \cdot e^{-B\cdot V_G} \cdot e^{-\frac{E_a}{kT}} \qquad \text{Eq. V.20}$$

with:

core	IO
n = 0.16 (core ΔV_{TH}= 50mV)	n= 0.15 (ΔID_{Sat} = 10%)
Ai = 1.48 10^8 , B= 12,	Ai = 5.02 10^7 , B= 4.6

With apparent activations:

E_a = 0.25 to 0.35 eV (Si Bulk)
E_a = 0.55 to 0.65 eV (FDSOI)
E_a= 0.15 to 0.20 eV (FinFETs)

We have supposed from CA that the width of a single FF is 10nm leading to $W_{eff} = [2.\ H_{Fin} + W_{Fin}]$ effective width:
W_{eff} (per Fin) = *[2 x 45 + 10] = 100 nm.*

Results of [82] do not exhibit NBTI issue (*T= 125°C*) as the extrapolated lifetime for FinFET (Bodie-Tie) per Fin width shows for:

For W_{eff} *= 5nm* V_{DD}*= 1.35V* ⇨ $\tau_{NBT} = 2\cdot10^{10}s = 634$ *years*
For W_{eff} *= 5nm* $\mathbf{V_{DD}= 1.5V}$ ⇨ $\boldsymbol{\tau_{NBT} = 4\cdot10^{9}s}$ *= 126 years*
For W_{eff} *= 10nm* ⇨ $\tau_{NBT} > 2\cdot10^{10}$s independently of V_{DD} (1.5V or 1.35V)

This arises from the much smaller vertical field condition for the DDR3 FF 20nm as we have

$$E_{ox} \cong V_{DDmax} / EOT = 1.5/2.8\cdot10^{-7} = 5.35\ MV/cm$$

down the minimum value giving

$$Eox \cong V_{DDmin} / EOT = 1.35/2.8 \cdot 10^{-7} = 4.8\ MV/cm$$

In conclusion, from the investigations shown, we have shown that DDR3 does not exhibit large damage acceleration factor for the typical $V_{DD}= 1.35V$ operation if one bases the calculation of a single FinFET with one Fin. For all three damage mechanisms no issue is found neither about HC, TDDB nor NBTI at that (knowledge) of right design margin and accessible technological parameters at the minimum length and effective width (16nm, 133nm).

The transfer from typical $V_{DD} = 1.35V$ to $V_{DDmax}= 1.5V$ shows an increased sensitivity to TDDB in that geometry of NMOS FinFETs where the lifetime is reduced to 0.81 year.

Moreover, the design of a digital cell emphasizes that the number of Fins has to be limited (with that current analysis of CA results) to a number of 2, as the increase at $V_{DDmax}= 1.5V$ (corresponding to the boost condition) makes appear a HC sensitivity for 2 Fins and more largely for 4 Fins, where the HC criterion is not fulfilled anymore for any (manufacturer) supply voltage conditions down to $V_{DD}= 1.35V$ (Typical value).

In conclusion we have shortly described our work devoted to a quick analysis in the context of the Prognostic Failure Model (PFM) methodology. We have projected a simplified framework analysis.

The much larger effects have been observed in the HC voltage at low voltage as it is a current driven phenomenon. Hence the effective width per Fin and the number of Fins exhibits a threshold (x4) as it has a strong impact on HC device to cell Lifetime value between 1.35V and the boost condition of 1.5V in the DDR3 component [106], [112]. The HC issue is shifted as a larger number of Fins (x6) in FPGA by the reduced supply voltage between 0.9V and 0.85V from the high performance to typical operating condition.

With respect to classical TDDB and NBTI behaviors, the only analyzed results from the literature do not evidence a large Lifetime reduction through these failure mechanisms as we are at much lower oxide fields. Accent has also been pointed out by the width dependence and the larger difference between the accelerating range and the field for usage. This is explained by the low voltage in the use condition, particularly for FPGA in the range 0.9V to 0.85V and the enlarged oxide field condition from the accelerated (stressing) voltage range. for SDRAM for 1.35V up to 1.5V.

III. Aerospace Electronics Reliability: Practical Application of MTOL.

We present a method for predicting the failure rate, and thus the reliability of an electronic system by summing the failure rate of each known failure mechanism. We use a competing acceleration factor methodology by combining the physics of failure for each mechanism with their effects as observed by High/Low temperature and High/Low voltage stresses. Our method assumes that lifetime of each of its failure mechanisms follows constant rate distribution and each mechanism is independently accelerated by the stress factors, which include also frequency, current, and other factors that can be entered into a reliability model. The overall failure rate also follows an exponential distribution and is described in the standard FIT (Failure unIT or Failure in Time). The method combines mathematical models for known failure mechanism and solves them simultaneously at a multiplicity of accelerated life tests to find a consistent set of weighting factors for each mechanism. The result of solving the system of equations is a more accurate and a unique combination for each system model by proportional summation of each of the contributing failure mechanisms.

To this day, the users of our most sophisticated electronic systems that include opto-electronic, photonic, MEMS device, etc. are expected to rely on a simple reliability value (FIT) published by the supplier. The FIT is determined today in the product qualification process by use of High-Temperature Overstress Life (HTOL) or other tests, depending on the product. The manufacturer reports a zero-failure result from the given conditions of the single-point test and uses a single-mechanism model to fit an expected MTTF at the operator's use conditions. The zero-failure qualification is well known as a very expensive exercise that provides nearly no useful information. As a result, designers often rely on HALT testing and on handbooks such as Fides, Telecordia or Mil Handbook 217 to estimate the failure rate of their products, knowing full well that these approaches act as guidelines rather than as a reliable prediction tool. Furthermore, with zero failure required for the "pass" criterion as well as the poor correlation of expensive HTOL data to test and field failures, there is no communication for the designers to utilize this knowledge in order to build in reliability or to trade it off with performance. Prediction is not really the goal of these tests, however current practice is to assign an expected failure rate, FIT, based only on this test even if the presumed acceleration factor is not correct.

This paradigm seems unfortunate since the manufacturers of our electronic equipment actually put a great deal of effort and spend so much money and excellent personnel resources to learn and study each failure mechanism. Today's approach to reliability takes the intimate knowledge of the failure mechanisms and then not communicate this knowledge downstream to the users, usually for fear that perhaps the models or the probabilistic interpretation will not be realized. Hence, known and already characterized mechanisms that could lead to failure are left out of the communication. This leaves the final, sterilized, qualification test as the only available assessment on which the user can rely. Everyone recognizes that the resulting calculation is far from being a reliable value to predict anything about the life. Worse than that, it makes a joke of the reliability prediction process and has led to confusion at all levels. It also makes a joke of the very excellent and hard work of the reliability engineers who evaluate the failure probabilities and the underlying physics.

New emerging methods of reliability prediction, including Probabilistic Design for Reliability, Physics of Failure, No-MTBF and other tools are being developed. These methods take many variables into consideration and look at the failure distribution as well as multiple mechanisms and their interactions. There are new campaigns, and calls to throw away the handbooks, all of them, without at least appreciating the theoretical underpinnings that may still be valid. Then we are left with the other debate of making deterministic assessments versus probabilistic models for reliability. However, without the proper communication between manufacturer and user, we again ask the question if making predictions for reliability is possible when models are used in disconnection from the actual products.

Here, we address the pros and cons of various predictive approaches using the common language of Failure in Time or Failure unIT (FIT). We will evaluate the goal of finding MTBF and evaluate the wisdom of various approaches to reliability prediction. Our goal is reliability prediction based on the system environment including space, military and commercial. It is our intent to show that the era of confidence in reliability prediction has arrived and that we can, in fact, make reasonable reliability predictions. There is a major value of having consistent reliability metrics and offer communication between levels of manufacturing and practical ways that upstream models can be used by downstream users to make cost-effective reliability predictions.

We have found that a practical means of separating electronic device failure mechanisms at the system level. We tested FPGA's with a large range of frequency operation and tested them at extremely high and low temperatures with voltages ranging from nominal to more than 2 times nominal voltages. The result is the ability to completely distinguish the influences of hot carrier injection (HCI), Electromigration (EM), Bias temperature instability (BTI) and time dependent dielectric breakdown (TDDB). Our result shows that a meaningful reliability prediction can be made by a summation of distinct intrinsic failure mechanisms as measured at the system level. The result of our work will be a system qualification protocol that can actually predict the failure in time (FIT) with a much greater accuracy than a standard high temperature or low temperature overstress life qualification.

Chip and packaged system reliability is still measured by a Failure unIT (FIT). The FIT is a rate, defined as the number of expected device failures per billion part hours. A FIT is assigned for each component multiplied by the number of devices in a system for an approximation of the expected system reliability. The semiconductor industry provides an expected FIT for every product that is sold based on operation within the specified conditions of voltage, frequency, heat dissipation and etc. Hence, a system reliability model is a prediction of the expected mean time between failures (MTBF) for an entire system as the sum of the FIT rates for every component.

Failures in terms of an acceleration factor, AF, as:

$$FIT = \frac{\#failures}{\#tested \cdot hours \cdot A_F} \cdot 10^9 \qquad \text{Eq. V-49}$$

where #failures and #tested are the number of actual failures that occurred as a fraction of the total number of units subjected to an accelerated test. The acceleration factor, AF, must be supplied by the manufacturer since only they know the failure mechanisms that are being accelerated in the High Temperature Operating Life (HTOL) and it is generally based on a company proprietary variant of the 217^+ approach for accelerated life testing. The true task of reliability modeling, therefore, is to choose an appropriate value for AF based on the physics of the dominant failure mechanisms that would occur in the field for the device.

1) Standard HTOL

The standard high temperature operating life (HTOL) qualification test is usually performed as the final qualification step of a semiconductor manufacturing process. The test consists of stressing some number of parts, usually about 100, for an extended time, usually 1000 hours, at an accelerated voltage and temperature. Two features shed doubt on the accuracy of this

procedure. One feature is the lack of sufficient statistical data and the second is that companies generally present zero failures results for their qualification tests and hence stress their parts under relatively low stress levels to guarantee zero failures during qualification testing.

Unfortunately, with zero failures no statistical data is acquired. Another feature is their calculation of the acceleration factor AF. If the qualification test results in zero failures, which allows the assumption (with only 60% confidence!) that no more than ½ a failure occurred during the accelerated test. This would result, based on the example parameters, in a reported FIT = 5000/AF, which can be almost any value from less than 1 FIT to more than 500 FIT, depending on the conditions and model used for the voltage and temperature acceleration.

The accepted approach for measuring FIT would be reasonably correct if there were only a single dominant failure mechanism that is excited equally by either voltage or temperature. Additionally, this same mechanism is the only one that is accelerated by the burn-in or accelerated test. For example, electromigration is known to follow Black's equation and is accelerated by increased stress current in a wire or by increased temperature of the device. If, however, multiple failure mechanisms are responsible for device failures, each failure mechanism should be modeled as an individual "element" in the system and the component survival is modeled as the survival probability of all the "elements" as a function of time [113].

2) Multiple Mechanisms

If multiple failure mechanisms, instead of a single mechanism, are assumed to be time-independent and independent of each other, FIT (constant failure rate approximation) should be a reasonable approximation for realistic field failures. Under the assumption of multiple failure mechanisms, each will be accelerated differently depending on the physics that is responsible for each mechanism. If, however, an HTOL test is performed at an arbitrary voltage and temperature for acceleration based only on a single failure mechanism, then only that mechanism will be accelerated. In that instance, which is generally true for most devices, the reported FIT (especially one based on zero failures) will be meaningless with respect to other failure mechanisms.

Whereas the failure rate qualification has not improved over the years, the semiconductor industry understanding of reliability physics of semiconductor devices has advanced enormously. Every known failure mechanism is so well understood and the processes are so tightly controlled that electronic components are designed to perform with reasonable life and with no single dominant failure mechanism. Standard HTOL tests generally reveal multiple failure mechanisms during testing, which would suggest also that no single failure mechanism would dominate the FIT rate in the field. Therefore, in order to make a more accurate model for FIT, a preferable approximation should be that all failures are equally likely and the resulting overall failure distribution resembles constant failure rate process that is consistent with the mil-handbook, FIT rate approach.

The acceleration of a single failure mechanism FM_i is a highly non-linear function of temperature and/or voltage. The temperature acceleration factor (AF_{i_T}) and voltage acceleration factor (AF_{i_V}) can be calculated separately and is the subject of most studies of reliability physics. The total acceleration factor of a failure mechanism AF_{FMi} du to different stress combinations is the product of the acceleration factors of temperature and voltage (in the example as per an exponential law, but coul be a Power law or a linear law),

$$AF_{FM_i} = \frac{\lambda_{FM_i}(T_{ref},V_{ref})}{\lambda_{FM_i}(T_{use},V_{use})} = AF_{i_T} \cdot AF_{i_V} = e^{\left[\frac{E_a}{k}\left(\frac{1}{T_{use}}-\frac{1}{T_{ref}}\right)\right]} \cdot e^{[\gamma_1 \cdot (V_{ref}-V_{use})]} \qquad \text{Eq. V-50}$$

This acceleration factor model is widely used as the industry standard for device qualification. However, it only approximates a single dielectric breakdown type of failure mechanism and does not correctly predict the acceleration of other mechanisms.

To be even approximately accurate, however, electronic devices should be considered to have several failure modes degrading simultaneously. Each mechanism 'competes' with the others to cause an eventual failure. When more than one mechanism exists in a system, then the relative acceleration of each one must be defined and averaged at the applied condition. Every potential failure mechanism should be identified and its unique AF_{FMi} should then be calculated for each mechanism at a given temperature and voltage so the FIT rate can be approximated for each mechanism separately. Then the final FIT will be the sum of the failure rates per mechanism, as is described by:

$$FIT_{total} = FIT_1 + FIT_2 + \ldots + FIT_i \qquad \text{Eq. V.51}$$

whereby each mechanism leads to an expected failure unit per mechanism, FIT_i. Unfortunately, again, individual failure mechanisms are not uniformly accelerated by a standard HTOL test, and the manufacturer is forced to model a single acceleration factor that cannot be combined with the known physics of failure models.

3) Acceleration Factor

The qualification of device reliability, as reported by a FIT rate, must be based on an acceleration factor, which represents the failure model for the tested device. If we assume that there is no failure analysis (FA) of the devices after the HTOL test, or that the manufacturer will not report FA results to the customer, then a model should be made for the acceleration factor, AF, based on a combination of competing mechanisms. This will be explained by way of example. Suppose there are two identifiable, constant rate competing failure modes (assume an exponential distribution). One failure mode is accelerated only by temperature. We denote its failure rate as $\lambda_{FM\_1}(T)$. The other failure mode is only accelerated by voltage, and the corresponding failure rate is denoted as $\lambda_{FM\_2}(V)$.

By performing the acceleration tests for temperature and voltage separately, we can get the failure rates of both failure modes at their corresponding stress conditions. Then we can calculate the acceleration factor of the mechanisms FM_i (for i= 1, and 2). If for the first failure mechanism $i=1$ we have $\lambda_{FM\_1}(T_{use})$, $\lambda_{FM\_1}(T_{ref})$, and for the second failure mechanism $i=2$, we have $\lambda_{FM\_2}(V_{use})$, $\lambda_{FM\_2}(V_{ref})$, then the temperature acceleration factor for the temperature at any given V is:

$$AF_{1_T}(V) = \frac{\lambda_{FM_1}(T_{use},V)}{\lambda_{FM_1}(T_{ref},V)} \quad \text{and} \quad AF_{2_T}(V) = \frac{\lambda_{FM_2}(T_{use},V)}{\lambda_{FM_2}(T_{ref},V)} \qquad T_{use} < T_{ref} \qquad \text{Eq. V-52}$$

and the voltage acceleration factor at any given T is:

$$AF_{1_V}(T) = \frac{\lambda_{FM_{use}}(T,V_{use})}{\lambda_{FM_{ref}}(T,V_{ref})}, \quad \text{and} \quad AF_{2_V}(T) = \frac{\lambda_{FM_{use}}(T,V_{use})}{\lambda_{FM_{ref}}(T,V_{ref})} \qquad V_{use} < V_{ref} \qquad \text{Eq. V-53}$$

The system acceleration factor between the stress ref and use conditions of (T_{ref}, V_{ref}) and (T_{use}, V_{use}) is:

$$AF = \frac{\lambda_{FM_1}(T_{use},V_{ref})+\lambda_{FM_2}(T_{ref},V_{use})}{\lambda_{FM_1}(T_{ref},V_{ref})+\lambda_{FM_2}(T_{ref},V_{ref})} \qquad \text{Eq. V-54}$$

The above equation can be transformed to the following two expressions:

$$AF = \frac{AF_{1_T}(V)\cdot\lambda_{FM_1}(T_{ref},V_{ref})+AF_{1_V}(T)\cdot\lambda_{FM_2}(T_{ref},V_{ref})}{\lambda_{FM_1}(T_{ref},V_{ref})+\lambda_{FM_2}(T_{ref},V_{ref})} \qquad \text{Eq. V-55.b}$$

Or

$$AF = \frac{\lambda_{FM_1}(T_{use},V_{use})+\lambda_{FM_2}(T_{use},V_{use})}{\frac{\lambda_{FM_1}(T_{use},V_{use})}{AF_{1_T}(V)}+\frac{\lambda_{FM_2}(T_{use},V_{use})}{AF_{1_V}(T)}} \qquad \text{Eq. V-55.a}$$

These two equations can be simplified based on different assumptions.

When $\lambda_1(T_{ref}, V_{ref}) = \lambda_2(T_{ref}, V_{ref})$, (i.e. equal probability at-ref condition), eq. V-55.b gives:

$$AF = \frac{AF_{1_T}(V)+AF_{1_V}(T)}{2} \qquad \text{Eq. V-56}$$

Therefore, unless the temperature and voltage is carefully chosen so that AF_T and AF_V are very close, within a factor of about 2, then one acceleration factor will overwhelm the failures at the accelerated conditions. Similarly, when $\lambda_1(T_{use}, V_{use}) = \lambda_2(T_{use}, V_{use})$, then AF will take this form:

$$AF = \frac{2}{\frac{1}{AF_{1_T}(V)}+\frac{1}{AF_{1_V}(T)}} \qquad \text{Eq. V-57}$$

and the acceleration factor applied to at-use conditions will be dominated by the individual factor with the smallest acceleration. In either situation, the accelerated test does not accurately reflect the correct proportion of acceleration factors

based on the understood physics of failure mechanisms.

This discussion can be generalized to incorporate situations with more than two failure modes. Suppose a device has n independent failure mechanisms, and λ_{i_ref} represents the *ith* failure mode at reference condition, λ_{i_use} represents the *ith* failure mode at use condition, then the total acceleration factor *AF* can be expressed. If the device is designed that the failure modes have equal frequency of occurrence during the use conditions:

$$\left.\begin{aligned}
\lambda_i(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) &= \lambda_i(S_{1_{ref}}, S_{2_{ref}}, \dots, S_{j_{ref}}) \cdot AF_{i_{S_{j\neq 1}}}(S_1) \\
\lambda_i(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) &= \lambda_i(\mathrm{S_{1_{ref}}}, S_{2_{ref}}, \dots, S_{j_{ref}}) \cdot AF_{i_{S_{j\neq 2}}}(S_2) \\
&\vdots
\end{aligned}\right\}
\begin{aligned}
&\text{for } S_j = S_1 \text{ to } S_p \text{ stressors} \\
&\text{for } i = 1 \text{ to } n^{th} \text{ failure mechanism}
\end{aligned}
\qquad \text{Eq. V.58}$$

The system acceleration factor between the stress ref and use conditions of $(S_{1\_ref}, S_{2\_ref}, \dots, S_{p\_ref})$ and $(S_{1\_use}, S_{2\_use}, \dots, S_{p\_use})$ reduces to:

$$AF = \frac{\lambda_1(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) + \lambda_2(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) + \cdots + \lambda_n(S_{1_{use}}, S_{2_{use}}, \dots, S_{p_{use}})}{\frac{\lambda_1(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}})}{AF_{i_{S_{j\neq 1}}}(S_1))} + \frac{\lambda_2(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}})}{AF_{i_{S_{j\neq 2}}}(S_2)} + \cdots + \frac{\lambda_2(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}})}{AF_{i_{S_{j\neq p}}}(S_p)}} \qquad \text{Eq.V-59}$$

For $\lambda_1(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) = \lambda_2(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) = \cdots = \lambda_n(S_{1_{use}}, S_{2_{use}}, \dots, S_{p_{use}}) = \lambda(S_{1_{use}}, S_{2_{use}}, \dots, S_{p_{use}})$, the AF factor become:

$$AF = \frac{n \cdot \lambda(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}})}{\lambda(S_{1_{use}}, S_{2_{use}}, \dots, S_{j_{use}}) \cdot \left[\frac{1}{AF_{i_{S_{j\neq 1}}}(S_1))} + \frac{1}{AF_{i_{S_{j\neq 2}}}(S_2)} + \cdots + \frac{1}{AF_{i_{S_{j\neq p}}}(S_p)}\right]} = \frac{n}{\sum_{i=1}^{n} \frac{1}{AF_{i_{S_{j\neq i}}}(S_i)}} \qquad \text{Eq.V-60}$$

From these relations, it is clear that only if acceleration factors for each mode are almost equal, i.e. $AF_1 \approx AF_2$, the total acceleration factor will be $AF = AF_1 = AF_2$, and certainly not the product of the two (as is currently the model used by industry). If, however, the acceleration of one failure mode is much greater than the second, the standard FIT calculation (Eq. V-50) could be incorrect by many orders of magnitude.

Due to the exponential nature of acceleration factor as a function of V or T, if only a single parameter is changed, then it is not likely for more than one mechanism to be accelerated significantly compared to the others for any given V and T. As we will see in the next section, at least 4 mechanisms should be considered. Also, the various voltage and temperature dependencies must be considered in order to make a reasonable reliability model for electron devices. Until now, the assumption of equal failure probability at-use conditions is used since it is the most conservative approach assuming the correct proportionality cannot be determined.

4) Proportionality Matrix Solution

The basic method for solving the system of equations is described in the paper from Bernstein [1] and using the suggestion of a Sum-of-failure-rate method as described in JEDEC Standard JEP122G as published in a more recent book by Bernstein [114]. The matrix method forms the basis for this work. It is clear that the manufacturers of electronic components recognize the importance of combining failure mechanisms in a sum-of-failure-rates method. Also, the formula for each mechanism is well studied and published.

The matrix approach we use to model useful life failure rate (FIT) for components in electronic assemblies by assuming each component is composed of multiple sub-components, for example; a certain percentage is effectively ring-oscillator, static SRAM, DRAM, etc. Each type of circuit, based on its operation, can be seen to affect the potential steady-state (defect related) failure mechanisms differently based on the accelerated environment, for example: Electromigration, Hot-Carrier, NBTI, etc. Each mechanism is known to have its own acceleration factor with voltage, temperature, frequency, cycles, etc. Each sub-component will be modeled to approximate the relative likelihood of each mechanism per sub-component. Then, each component can be seen as a matrix of sub-components, each with its own relative weight for each possible mechanism.

Hence, the standard system reliability FIT can be modeled using traditional 217^+ type of algorithms and adapted to known system reliability tools, however, instead of treating each component as individuals, we propose treating each complex component as a series system of sub-components, each with its own reliability matrix.

The prediction of a system reliability can be described using a linear matrix solution. Although until today, the methodology was consolidated on microelectronic device failure mechanism, it applies directly to additional mechanisms including thermal and mechanical stresses due to wafer bonding and on any failure mechanism that can be modelled by physics of failure, including wide bandgap semiconductors and even packaging failures.

Whereas each intrinsic mechanism is known to have different statistical distributions, the combination of distributions becomes, at the ensemble level, approximately constant rate as demonstrated by R.F. Drenick [17]. In its theorem, Drenick suggests and justifies the summation of failure rate approach also as explained in the JEDEC handbook.

The mechanism matrix is described in Table V-2. Each row of the matrix describes various operating conditions under which the system is tested. Each experiment, *i*, is operated with its unique voltage, frequency and temperature. The ''results'' column, FIT_i is the average time when the failure occurs under the experimental condition, which is associated with a pre-determined failure point. The example studied uses 10% performance degradation as the failure point, however any reasonable value will work as long as it is consistent with the application. The result FIT_i is a failure rate (λ) and measured as $10^9/MTTF$.

	HCI	**BTI**	**EM**	**Results**
V_1, f_1, T_1	$X \cdot A_1$	$Y \cdot B_1$	$Z \cdot C_1$	FIT_1
V_2, f_2, T_2	$X \cdot A_2$	$Y \cdot B_2$	$Z \cdot C_2$	FIT_2
V_3, f_3, T_3	$X \cdot A_3$	$Y \cdot B_3$	$Z \cdot C_3$	FIT_3

Table V-2: M-TOL matrix used to solve models with measured times to fail [115]

Let's assume each mechanism (A–C) affects the system linearly with its own acceleration factor (*AF*) for a given frequency. The Acceleration factor formulas are in Table V-3. Each equation is calculated with the experimental condition of each result on the right-hand side.

Hot carrier injection	$A_i \equiv AF_{HCI} = \frac{f}{f_0} \cdot \left(\frac{V}{V_0}\right)^N \cdot e^{-\frac{Ea_{HCI}}{k}\left(\frac{1}{T_0}-\frac{1}{T_{ac}}\right)}$
Negative bias temperature instability	$B_i \equiv AF_{NBTI} = e^{\gamma_{BTI}\cdot(V_G - V_{G,0})} \cdot e^{-\frac{Ea_{BTI}}{k}\left(\frac{1}{T_0}-\frac{1}{T_{ac}}\right)}$
Electromigration	$C_i \equiv AF_{EM} = \frac{f}{f_0} \cdot \left(\frac{V}{V_0}\right)^M \cdot e^{-\frac{Ea_{EM}}{k}\left(\frac{1}{T_0}-\frac{1}{T_{ac}}\right)}$

Table V-3: The equations for the acceleration factors matrix [115]

Then the matrix is solved to find a set of constants, P_i, shown here as *X*–*Z*, across the whole matrix that matches the experimental results with calculated acceleration factors. This linear matrix is solved by multiplying the inverse matrix, AF^{-1}, with lambda at each condition, as shown in Table V-4. The solution gives the coefficients (X–Z), which make up the relative contribution of each failure mechanism on the system.

$$\overbrace{\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{bmatrix}}^{AF} \cdot \overbrace{\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}^{P_i} = \overbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}}^{\lambda}$$

$$(AF) \cdot (Pi) = (\lambda) \;\rightarrow\; (Pi) = (AF)^{-1} \cdot (\lambda)$$

Table V-4: Matrix solution [115].

In this matrix, the 3 left-hand columns describe the voltage, temperature and frequency used for the accelerated test. The normalized relative acceleration factor based on the formulae for each failure mechanism is given in each of the 3 columns that follow. The 2 right-hand columns show the measured extrapolated failure rate and the calculation based on the matrix solution. These values are extrapolated failure rates of the product after the accelerated test. This matrix will have a unique solution that will fit the percentages of each mechanism (P_i) with the measured failure rate, FIT. We found no influence of TDDB for these

experiments so we did not include them in the matrix, so we had only 3 equations solved with 3 measured FIT rates.

An absolute FIT value is determined in the next row based on the Mean time to fail. This allows calibration of the final results in operation. The column line is the expected FIT (failures per billion part-hours) at those conditions. By substituting these percentages into the matrix, the true acceleration factors are determined for not only the tested condition but also for any extrapolated condition.

A failure mechanism matrix can be solved at any given set of conditions, i.e. voltage, temperature and frequency, as a percentage at each stressed operating condition, there is a unique proportion of each mechanism for a given set of stressed conditions that will result in the given time to fail. For example, the matrix shown in Table V-1 shows the calculated acceleration factor for each mechanism contributing to an eventual failure. The columns are normalized to their relative calculated contribution out of 100% including all three; EM, HCI and NBTI. A calculated reliability curve is shown in Figure V-6 showing the full range of expected FIT versus Temperature for any set of operational conditions shown in Table V-6.

The 45 nm technology shows frequency related effects at both low temperatures (below 5°C) due to HCI and at high temperatures. It is observed the high voltage bias (@ 1.2 V) enhance the effect of frequency which reduce the overall HCI contribution at low frequency. The dominant failure mechanism at medium ambient temperature (range from 10°C to 150°C) is related to NBTI while EM failure mechanism is rather observed at high temperature.

Volt	T °C	Freq (MHz)	EM	HCI	NBTI	Measured FIT	Calculated FIT
2.4	-20	500	0.00%	100.00%	0.00%	8.00E+04	8.00E+04
1.2	140	500	100.00%	0.00%	0.00%	5.40E-01	5.40E-01
2.4	160	0.02	0.58%	3.40%	96.02%	1.78E-02	1.78E-02
3	0	500	0.00%	81.16%	18.84%	3.45E+07	3.43E+07
2.4	173	500	35.19%	61.50%	3.31%	1.76E+01	1.73E+01
2.4	160	500	14.56%	85.34%	0.10%	1.50E+01	1.77E+01
3	0	0.20	0.00%	0.02%	99.98%	6.37E+06	6.47E+06

Table V-5. *Test Results showing proportions of failure mechanisms for given V,T and F compared with the calculated as well as the measured failure rate (FIT).*

Volt	T °C	F (MHz)	EM	HCI	NBTI	Calculated FIT
1.1	30	0.02	99.76%	0.24%	0.00%	2.87E-10
1.2	70	500	100.00%	0.00%	0.00%	9.88E-04
1.5	80	500	98.58%	1.42%	0.00%	3.00E-03
2	-50	500	0.00%	100.00%	0.00%	3.12E+03
1.8	125	500	98.23%	1.77%	0.00%	1.86E-01
2	150	0.02	96.20%	3.80%	0.00%	5.16E-05

Table V-6. *Calculated FIT based on the solved matrix for typical use conditions.*

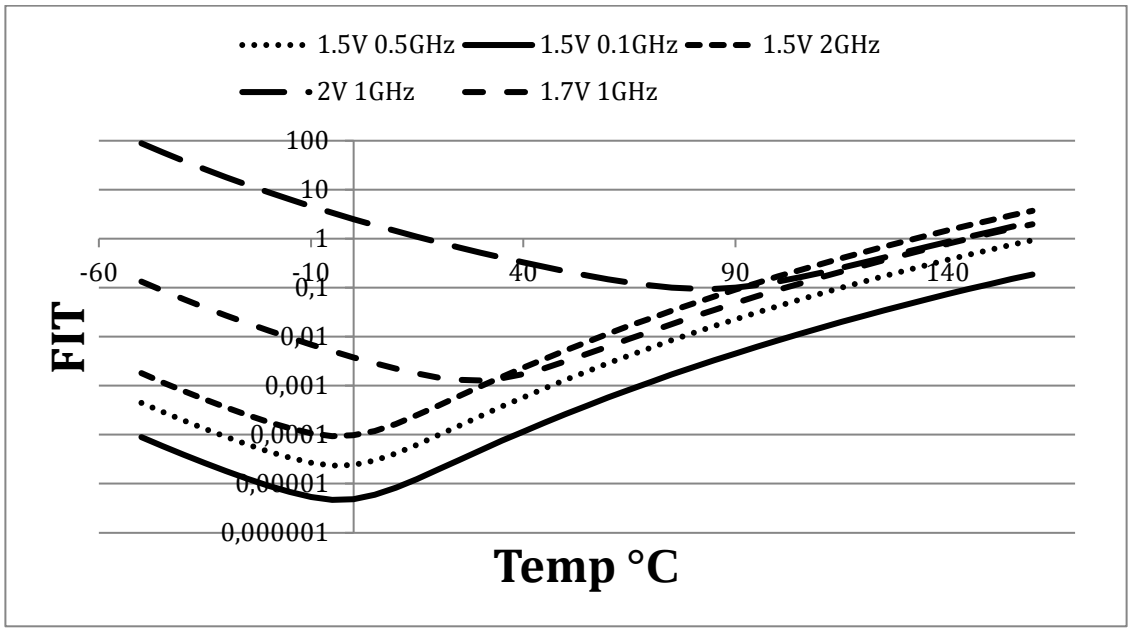


Figure V-5. FIT versus Temperature for various voltages and frequencies.

The unique solution that solves all 3 equations with the 3 extrapolated acceleration factors gives a percentage contribution for each of the failure mechanisms. We report the reliability as FIT, which is 10^9/MTTF for each condition. The percentages for each mechanism are shown, based on the relative contributions that were extrapolated from the physics of failure equations normalized to the measured FIT of each test.

The most important result from our study is that Electromigration and HCI are the most dominant failure mechanisms throughout the useful range of device operation. This is surprising since the standard HTOL test emphasizes only TDDB and BTI since those are most accelerated by high voltage and temperature, however under use conditions, the other two are most important. Furthermore, it is important to see that at very low temperature and high frequency, HCI is the most important failure mechanism and this could have very important implications for satellite and low-temperature military applications. Fortunately, very low FIT values are found, and reliability is confidently predicted.

We present here a simple and accurate way to combine the physics of failure equations for reliability prediction from accelerated life testing. We show that a matrix approach allows the reliability physics equations to be fit proportionally to the results of monitored accelerated life testing in order to extrapolate failure rate one would expect given actual operating parameters. This methodology can be extended to include radiation effects, frequency and even packaging and solder joint effects to give a complete system reliability evaluation framework. This matrix gives a very cost-effective way to predict reliability based on the Physics of Failure using only 3 tests as compared to the normal single-mechanism approach.

IV. Conclusion

In this chapter book we presented how basic concepts can be applied to high reliability application when using COTS devices or new emerging technologies in harsh environment conditions.

A short State of the Art was developed in first paragraph presenting DSM and GaN Technologies overview on deep scale integrated technology down to 3 nm and last developments on GaN on Si power switch technologies.

Multiple failure mechanisms and Physics of degradation in semiconductors may occur in a single set of time-to-failure data Physics of Healthy and associated Thermodynamics modeling and mathematical approach are completed.

Some case studies application for Automotive and Aerospace are presented. It affords series of numerical application supported by experiments for study cases on a FinFET technology assuming HCI, BTI and EM as major failure mechanisms. It is shown how the activation energy is related to the stress and temperature applied. It is obvious that activation energy can no-longer be considered as constant to extrapolate some experiment under high stress to nominal mission operation profile.

It is examined and shown that the equivalent activation energy is increasing with temperature and decreasing with increase of electrical stress condition (named stressors).

Some exercise based on PFM level 3 is proposed for Automotive application, starting from component selection, performing technology analysis sometimes supported by Constructional Analysis, established failure rate models identified and adjusted from the literature and provided by the Manufacturer.

Related to the Aerospace and Automotive Electronics reliability and a Practical application of MTOL based on J. Bernstein approach, the prediction of a system reliability is described using a linear matrix solution.

V. BIBLIOGRAPHY

[1] M. White and J. B. Bernstein, "Microelectronics Reliability: Physics-of-Failure Based Modeling and Lifetime Evaluation," *JPL Publication 08-5, Jet Propulsion Laboratory/California Institute of Technology,* February 2008.

[2] National Academy of Sciences, "Panel on Reliability Growth methods for Defense; Reliability Growth: Enhancing Defense System Reliability," The National Academies Press, Washington, DC. ISBN 978-0-309-31474-9, 2008.

[3] P. Lall, M. Pecht and E. B. Hakim, Influence of Temperature on Microelectronics and System Reliability, Boca Raton, NY: CRC Press, 1997.

[4] DoD, "MIL-HDBK-217, Military Handbook for Reliability prediction of Electronic Equipement," Washington, DC, USA, 1991, December.

[5] JEDEC, "JEDEC JEP-122G Failure Mechanisms and Models for Semiconductor Devices," JEDEC Solid State Technology Association, Arlington, 2011.

[6] J. B. Bernstein, M. Gabbay and O. Delly, "Reliability Matrix Solution to Multiple Mechanism Prediction," *Microelectronics Reliability Journal,* vol. 54, pp. 2951-2955, 2014.

[7] Robustness Group ZVEI, Handbook for Robustness Validation of Automotive Electrical/Electronic Modules, Frankfurt am Main, June 2013.

[8] QUANTERION, «Handbook of 217Plus Reliability Prediction Models,» *QUANTERION Solution Inc.,* 2015.

[9] T. S. Q. &. R. Publications, "http://www.triquint.com/shared/pubs/processes/Micro_Millimeter_Wave_Reliability_Overview.pdf," [Online].

[10] J. Berthon, D. Regis and G. Hubert, "Utilisation des composants DSM dans le contexte aéronautique.," in *19ème Congrés de Maitrise des Risques et Sûreté de Fonctionnement,* Dijon, 21-23 Oct. 2014.

[11] S. Arrhenius, "Arrhenius, Ueber den Einfluss des Atmosphärischen Kohlensäurengehalts auf die Temperatur der Erdoberfläche," *Proceedings of the Royal Swedish Academy of Science,* vol. 22, no. 1, 1896.

[12] L. Boltzmann, "Boltzmann, The Second Law of Thermodynamics. Populare Schriften," in *Essay 3, Address to a Formal Meeting of the Imperial Academy of Science,,* 1886.

[13] M. Evans and M. Polanyi, "Inertia and driving force of chemical reaction," *Faraday Society (London) Trans,* vol. 34, pp. 11-29, 1938.

[14] E. Wigner, "The transition state method," *Faraday Society (London) Trans,* vol. 34, pp. 29-41, 1938.

[15] S. Glasstone, K. J. Laider and H. Eyring, The theory of rate processs: the kinetics of chemical reaction, viscosity, diffusion and electrochemical phenomena., New York: Mc Graw-Hill, 1941.

[16] G. Hammond, "A correlation of Reaction Rates," *J. Am. Chem. Soc.,* vol. 77, pp. 334-338, 1955.

[17] R. Drenick, «Mathematical Aspects of the Reliability Problem,» *Journal of the Society for Industrial and Applied Mathematics,* vol. 8, n° %11, pp. 125-149, 1960.

[18] E. Snow, A. Grove, B. Deal and C. Sah, "Ion Transport Phenomena in Insulating Films," *J. Appl. Phys.,* vol. 36, p. 1664, 1965.

[19] N. Sedyakin, "On one physical principle in reliability theory (in Russian)," *Techn. Cybernetics,* pp. 80-87, 1966.

[20] D. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological), Vol. 34, No. 2. (1972), pp.,* vol. 34, no. 2, pp. 187-220, 1972.

[21] H. Eyring, S. Lin and S. Lin, Basic Chemical Kinetics, New York - Chichester - Brisbane - Toronto: John Willey & Sons, 1980.

[22] J. McPherson and D. Baglee, "Acceleration factors for thin gate oxide stressing," *Proc. 23rd Annual IEEE IRPS,* p. 1, 1985.

[23] J. W. McPherson, Reliability Physics and Engineering - Time-to-failure modeling; 3rd Edition, Plano (TX) USA: Springer Nature Switzerland AG, 2019.

[24] E. Suhir, "Probabilistic Design for Reliability," *ChipScale Reviews,* vol. 14, no. 6, 2010.

[25] E. Suhir, "Predicted Reliability of Aerospace Electronics: Application of Two Advanced Concepts," *IEEE Aerospace Conference,* 2-9 March 2013.

[26] D. Jin and J. Del Alamo, "Impact of high-power stress on dynamic On-resistance of high-volatge GaN HEMTs," *Microelectronics Reliability Journal,* vol. 52, pp. 2875-2879, 2012.

[27] G. Meneghesso, M. Meneghini, A. Tazzoli, N. Ronchi, A. Stocco, A. Chini and E. Zanoni, "Reliability issues of Gallium Nitride HEMTs," *International Journal of Microwave and Wireless Technologies,* vol. 2, no. 1, pp. 39-50, 2010.

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

[29] J. Black, "Black, Electromigration - A brief survey and some recent results," *IEEE, Trans. Electron Devices*, Vols. ED-16, p. 388, 1969.

[30] "RDF 2000 (UTE C 80-810, IEC-62380-TR Ed.1)," [Online]. Available: http://www.ute-fr.com/La-normalisation/UTE-and-standardisation.

[31] SIEMENS AG, SN29500, Reliability and Quality Specifications Failure Rates of Components, Siemens Technical Liaison and Standardisation, 1986.

[32] V. Huard, F. Cacho and X. Federspiel, "Technology scaling and reliability challenges in the multicore era," *IRPS*, 16-18 April 2012.

[33] R. Quay, Gallium Nitride Electronics, Berlin Heidelberg: Springer Series in Materials Sciences, 2008.

[34] J. Bernstein and A. Bensoussan, "Reliability Prediction with MTOL, proposed to IEEE System Qualification for Reliability," *Transaction on Devices and Material Reliability (TDMR)*, 2016.

[35] INTEL, "VDSM Issues and Design Methodology," Intel Microelectronics Services, January 2002.

[36] C. Lee, B. Welch and W. Fleming, "Reliability of AuGe/Pt and AuGe/Ni ohmic contacts on GaAs," *Electronics Letters*, vol. 17, pp. 407-408, 1981.

[37] M. Delaney, T. Wiltsey, M. Chiang and K. Yu, "Reliability of 0.25µm GaAs MESFET MMIC Process: Results of Accelerated Lifetests and Hydrogen Exposure," in *GaAs Reliability Workshop*, Philadelphia, PA, USA, 1994.

[38] K. Decker, "GaAs MMIC Hydrogen Degradation Study," in *GaAs Reliability Workshop*, Philadelphia, PA, USA, 1994.

[39] A. Bensoussan, P. Coval, W. Roesch and T. Rubalcava, "Reliability of a GaAs MMIC process based on 0.5µm Au/Pd/Ti gate MESFETs," in *32nd Annual Proceeding Reliability Physics, IRPS*, San Jose (CA), April 1994.

[40] J. Scotten, "/forum/content/8022-lithovision-2019-semiconductor-technology-trends-their-impact-lithography.html," SemiWiki.com, 24 Feb 2019. [Online]. Available: https://www.semiwiki.com. [Accessed 28 Feb 2019].

[41] M. Lapedus, "/big-trouble-at-3nm/," Semiengineering.com, 21 June 2018. [Online]. Available: https://semiengineering.com. [Accessed 28 February 2019].

[42] T. Ueda, "Reliability issues in GaN and SiC power devices," in *IEEE International Reliability Physics Symposium (IRPS)*, Waikoloa, Hawaii, USA, 2014.

[43] M. Meneghini and et al., "Normally-off GaN-HEMTs with p-type gate: Off-state degradation, forward gate stress and ESD failure," *Microelectronics reliability Journal*, vol. 58, pp. 177-184, 2016.

[44] G. Meneghesso, G. Verzellesi, F. Danesin, F. Rampazzo, F. Zanon, M. Meneghini and E. Zanoni, "Reliability of GaN High-Electron-Mobility Transistors: State of the Art and Perspectives," *IEEE Transactions on device and materials reliability,*, vol. 8, no. 2, pp. 332-343, 2008.

[45] Y. Wu, C.-Y. Chen and J. Del Alamo, "Temperature-accelerated degradation of GaN HEMTs under high-power stress: activation energy of drain-current degradation.," in *JEDEC Solid State Technology Association*, 2014.

[46] S. Wamock and J. Del Alamo, "OFF-state TDDB in high-voltage GaN MIS-HEMTs," in *IEEE International Reliability Physics Symposium (IRPS)*, Monterey, CA, USA, 2017.

[47] A. Guo and J. Del Alamo, "Positive-bias temperature instability (PBTI) of GaN MOSFETs," in *IEEE International Reliability Physics Symposium*, Monterey, CA, USA, 2015.

[48] M. Meneghini, G. Meneghesso and E. Zanoni, Power GaN Devices - Materials, Applications and Reliability, Springer, Power Electronics and Power Systems, Series editors J. H. Chow, A. M. Stankovic, D. Hill., 2017.

[49] B. Agarwala and al., "Dependence of Electromigration-induced Failure Time on Length and Width of Aluminium Thin-Film Conductors," *J. Appl. Phys.*, vol. 41, p. 3954, 1970.

[50] JEDEC, "JESD22-A108 HTOL Temperature, bias, and operating life," July 2017. [Online]. Available: www.jedec.org.

[51] K. Decker, "GaAs MMIC Hydrogen Degradation Study," in *GaAs Reliability Workshop*, Philadelphia, PA, USA, 1994.

[52] S. Kayali, G. Ponchak and R. Shaw, ""GaAs MMIC Reliability Assurance Guidelines for Space Applications," JPL Publication, Vol 96-25, Pasadena, CA, 1996.

[53] M. Meneghini, G. Meneghesso and E. Zanoni, Power GaN Devices - Materials, Applications and Reliability, Springer, Power Electronics and Power Systems, Series editors J. H. Chow, A. M. Stankovic, D. Hill., 2017.

[54] G. Meneghesso, G. Verzellesi, F. Danesin, F. Rampazzo, F. Zanon, M. Meneghini and E. Zanoni, "Reliability of GaN High-Electron-Mobility Transistors: State of the Art and Perspectives," *IEEE TRANSACTIONS ON DEVICE AND MATERIALS RELIABILITY,*, vol. 8, no. 2, pp. 332-343, 2008.

[55] A. K. M. C. a. M. M. H. Yunm, "Entropic Approach to Measure Damage with Applications to Fatigue," *Annual Reliability and Maintainability Symposium (RAMS), Reno, NV,*, pp. 1-6, 2018.

[56] J. McPherson, R. Khamankar and A. Shanware, "Complementary model for intrinsic TDDB in SiO2 dielectrics," *J. Semiconductor Sci. Technol.*, vol. 15, p. 462, 2000.

[57] J. McPherson, "Stress dependent activation energy," *Proc. IEEE 24th IEEE IRPS*, p. 12, 1986.

[58] A. Bensoussan and E. Suhir, "Design-for-Reliability (DfR) of Aerospace Electronics: attributes and challenges," *IEEE Aerospace Conference*, 2-9 March 2013.

[59] S. Zhurkov, "Kinetic Concept of the Strength of Solids," *Int. J. of Fracture Mechanics*, vol. 1, no. 4, 1965.

[60] D. Cox and D. Oakes, Analysis of Survival Data, Volume 21 Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press - 208 pages, 1984.

[61] J. W. PcPherson, ""Future of Reliability Testing - Can Zero Failures become reality"," in *ESREF 2019*, Toulouse, September 2019.

[62] L. Pierrat, «Estimation and validation of predictive reliability - Technical Report LJC-VAL/PTS,» VALEO Powertrain Thermal Systems (Internal report study) - LJ-CONSULTING, 2017.

[63] D. Delaux, «Fiablité Appliquée pour l'Industrie,» Valeo PTS (internal report), 2019.

[64] L. Pierrat et D. Delaux, «Analytical improvement of the Stress-Strength Method by considering a realistic strength distribution.,» chez *2013 ARS, Europe*, Berlin (Ge), 2013.

[65] E. Sperling, "https://semiengineering.com/gaps-emerge-in-automotive-test/," Semiconductor Engineering, 05 May 2019. [Online]. Available: https://semiengineering.com. [Accessed 08 May 2019].

[66] FIDES_Guide, "Reliability Methodology for Electronic Systems, DGA," 2004.

[67] P. A. Tobias and D. C. Trindade, Applied Reliability (3rd Ed.), Boca Raton, FL: CRC Press, 2012.

[68] L. Condra, G. Horan and B. Scofield, "As Semiconductor Devices Shrink so do their Reliability and Lifetimes," in *National Software and Airborne Electronic Hardware. Standardization Conference*, Denver (CO), 2005.

[69] GIFAS, «Guide for the selection of Deep Sub-Micron (DSM) for aeronautic equipment - Semiconductors lifetime - GIFAS/2015/5022,» Groupement Industries Françaises Aeronautiques et Spatiales, Paris, 2015.

[70] W. Roesch, «Volume & Quality Impacts on Reliability: A New Game for GaAs,» chez *GaAs REL Workshop*, Seattle, Washington, 2000.

[71] L. Bechtold, D. Redman et B. Tawfellows, «Semiconductor reliability using random and wearout failure models,» chez *Reliability and Maintainability Symposium (RAMS)*, Colarado Spring, CO, USA, 2014.

[72] VITA, «Physics of Failure Reliability Predictions, VITA (June) 2011,» ANSI/VITA 51.2-2011, June 2011.

[73] D. Regis, J. Berthon, G. Hubert and M. Gatti, "DSM Reliability concerns - impact on safety assessment," in *SAE International J. Aerospace, Vol 7, issue 1*, Cincinnati, September 2014.

[74] W. G. Components, "Guide for the selection of Deep Su-Micron (DSM) for Aeronautic equipment - Semiconductor lifetime," in *Réf : GIFAS/2015/5022*, Paris, 2015.

[75] M. Nikulin, L. Gerville-Reache and V. Couallier, Statistique des essais accélérés, Paris: Lavoisier; Hermes Science, 2007.

[76] H. H. Randomson, Y. Zhang, X. He, H. Cui, J. Li, J. Xiang, J. Liu, S. Gu and G. Wang, "The Challenges of Advanced CMOS Process from 2D to 3D.," *Applied Sciences (www.mdpi.com/journal/applsci)*, vol. 7, pp. 1-32, 2007.

[77] J. Mc Pherson, "Tutorial ESREF 2019: Future of reliability testing - Can Zero failures become reality?," in *ESREF, European Symposium on Reliability of Electron Devices*, Toulouse (France), 2019.

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

[79] B. Kaczer, V. Arkhipov, M. Jurczak and G. Groeseneken, "Negative bias temperature instability (NBTI) in SiO2 and SiON gate dielectrics understood through disorder-controlled kinetics," *Microelectron Eng.*, vol. 80, pp. 122-125, 2005.

[80] M. Alam, "Ch. 8: NBTI TIME DEPENDENCE: STRESS PHASE," in *Course Purdue Univ.*.

[81] J. W. McPherson, R. B. Khamankar and A. Shanware, "A complementary molecular-model (including field and current) for TDDB in SiO2 dielectrics,," *Microelectronics Reliability*, vol. 40, pp. 1591-1597, 2000.

[82] H. Lee, C.-H. Lee, D. Park and Y.-K. Choi, "A Study of Negative-Bias Temperature Instability of SOI and Body-Tied FinFETs," *IEEE Electron Device Letters*, vol. 26, no. 5, pp. 326-328, 2005.

[83] V. Huard, M. Denais and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability*, vol. 46, pp. 1-23, 2005.

[84] P. Nicollian, M. Rodder, D. Grider, P. Chen, R. Wallace and S. Hattangady, "Low Voltage Stress-Induced-Leakage-Current in Ultrathin Gate Oxides," in *IRPS, 37th Annual International Reliability Physics Symposium*, San Diego (USA), 1999.

[85] P. Nicollian, "Physics of trap generation and electrical breakdown in ultra-thin SiO2 and SiON gate dielectric materials," University of Twente, Twente, Netherlands, 2077.

[86] G. Ribes, S. Bruyere, D. Roy, M. Denais, J. Roux, C. Parthasarathy, V. Huard and G. Ghibaudo, "New extensive MVHR breakdown models," in *IRPS*, San Jose, (CA), 2006.

[87] E. Wu, B. Li, J. Stathis, R. Achanta, R. Filippi and P. McLaughlin, "A Time-Dependent Clustering Model for Non-Uniform Dielectric Breakdown," in *IEEE International Electron Devices Meeting*, 2013.

[88] S. Blonkowski, "Filamentary model of dielectric breakdown," *Journal of Applied Physics*, vol. 107, 2010.

[89] J. McPherson, "Time dependent dielectric breakdown physics - Models revisited," *Microelectronics Reliability Journal*, vol. 52, pp. 1753-1760, 2012.

[90] F. Monsieur, E. Vincent, D. Roy, S. Bruyere, J. Vildeuil, G. Pananakakis and G. Ghibaudo, "A Thorough Investigation of Progressive Breakdown in Ultra-thin Oxides. Physical Understanding and Application for Industrial Reliability Assessment.," in *IRPS*, Dallas (Tx), 2002.

[91] E. Wu and J. Sune, "Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability," *Microelectronics Reliability Journal*, vol. 45, pp. 1809-1834, 2005.

[92] B. Kaczer, R. Degraeve, P. Roussel and G. Groeseneken, "Gate oxide breakdown in FET devices and circuits: From nanoscale physics to system-level reliability," *Microelectronics Reliability Journal*, vol. 47, pp. 559-566, 2007.

[93] S. Ramey, A. Ashutosh, J. Clifford, M. Hattendorf, J. Hicks, R. James, A. Rahman, V. Sharma, A. St Amour and C. Wiegand, "Intrinsic Transistor Reliability Improvements from 22nm Tri-Gate Technology," in *IRPS*, Anahiem (CA), 2013.

[94] M. Saliva, F. Cacho, C. Ndiaye, V. Huard, D. Angot, A. Bravaix and L. Anghel, "Impact of gate oxide breakdown in logic gates from 28nm FDSOI CMOS technology," in *IRPS*, Monterey (CA), 2015.

[95] A. Bravaix, V. Huard, F. Cacho, X. Federspiel and D. Roy, Chapter book : Hot-Carrier Degradation in Decananometer CMOS Nodes: From an Energy-Driven to a Unified Current Degradation Modeling by a Multiple-Carrier Degradation Process, Wien (Austria): Springer, 2015.

[96] E. Takeda and N. Suzuki, "An empirical model for device degradation due to hot-carrier injection," *IEEE Elect. Dev. Lett.*, Vols. EDL-4, no. 4, pp. 111-113, 1983.

[97] A. K. Singh and J. Samanta, "Different Physical Effects in UDSM MOSFET for Delay & Power Estimation: A Review," in *IEEE Students' Conference on Electrical, Electronics and Computer Science*, Bhopal, India, 2012.

[98] S. Rauch and G. La Rosa, "The Energy Driven Paradigm of NMOSFET Hot Carrier Effects," in *IRPS*, San Jose (CA), 2005.

[99] A. Bravaix, C. Guérin, D. Goguenheim, V. Huard, D. Roy, C. Besset, S. Renard and Y. Mamy Randriamihaj, "Off State Incorporation into the 3 energy mode Device Lifetime Modeling for advanced 40nm CMOS node," in *IRPS*, San Jose (CA), 2010.

[100] K.-C. Chang, J.-C. Liao, T.-C. Chang, C.-H. Yeh, C.-Y. Lin, F.-Y. Jin, Y.-S. Lin, F.-M. Ciou, W.-C. Hung, Y.-H. Lin, Chen-Hsin Lien, O. Cheng, C.-T. Huang and a. Y.-H. Ye, "Abnormal Relationship Between Hot Carrier Stress Degradation and Body Current in High-k Metal Gate in the 14-nm Node," *IEEE ELECTRON DEVICE LETTERS*, vol. 40, no. 4, pp. 498-501, 2019.

[101] V. Huard, "Two independent components modeling for Negative Bias Temperature Instability," in *IRPS*, San Jose (CA), 2010.

[102] S. Mahapatra and S. Rashmi, "On the Universality of Hot Carrier Degradation: Multiple Probes, Various Operating Regimes, and Different MOSFET Architectures," *IEEE Transactions on Electron Devices*, vol. 65, no. 8, pp. 3088-3094, August 2018.

[103] T. Grasser, Hot Carrier degradation in semiconductor devices, Springer, 2015.

[104] T. Grasser, K. Rott, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, M. Waltl, M. Toledano-Luque and B. Kaczer, "Advanced characterization of oxide traps: The dynamic time-dependent defect spectroscopy," in *IRPS*, Apr. 2013.

[105] C. Guerin, V. Huard and A. Bravaix, "The Energy-Driven Hot-Carrier Degradation modes on nMOSFETs," *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 2, pp. 225-235, June 2007.

[106] A. Bravaix, "Part 1 - FRAME Analysis IRT-ISEN," Internal report IRT - ISEN, Toulouse, September 2019.

[107] C. Shin, K. Kim, J. Kim, W. Ko, Y. Yang, S. Lee, C. S. Jun and Y. S. Kim, "Fast, exact, and non-destructive diagnoses of contact failures in nano-scale semiconductor device using conductive AFM," *Scientific Reports*, vol. 3, no. Article number: 2088, p. DOI: 10.1038/srep02088, 2013.

[108] H. J. Veendrick, Nanometer CMOS ICs - From basics to ASICs, Heeze, The Netherlands: Springer, Jan. 2017 (second ed.).

[109] Y. K. Kim, J. S. Lee, G. Kim, T. Park, H. J. Kim, Y. P. Cho, Y. J. Park and M. J. Lee, "Partial Isolation Type Saddle-FinFET -Pi-FinFET - for Saddle FinFET Sub-30 nm DRAM Cell Transistors," *Electronics 2019*, no. doi:10.3390/electronics8010008, p. www.mdpi.com/journal/electronics, 2019.

[110] G. Crupi, D. M.-P. Schreurs, J.-P. Raskin and A. Caddemi, "A comprehensive review on microwave FinFET modeling for progressing beyond the state of art," *Solid-State Electronics*, vol. 80, pp. 81-95, 2013.

[111] A. Bravaix, C. Guerin, V. Huard, D. Roy, J. Roux and E. Vincent, "Hot-Carrier Acceleration Factors for Low Power Management in DC-AC stressed 40nm NMOS node at High Temperature," in *IEEE International Reliability Physics Symposium (IRPS) Proc.*,, Anaheim (CA), 2009.

[112] A. Bravaix, "Part 3 - FPGA FF16 Reliability Analysis - ISEN," Internal report IRT - ISEN, Toulouse, September 2019.

[113] J. Bernstein, M. Gurfinkel, X. Li, Walters, J., Y. Shapira and M. Talmor, "Electronic circuits reliability modelling," *Microelectronics Reliability*, vol. 46, pp. 1957-1979, 2006.

[114] J. Bernstein, Reliability prediction from Burn-in - Data Fit to Reliability models, ELSEVIER, 2014.

[115] J. B. Bernstein, M. Gabbay et O. Delly, «Reliability Matrix Solution to Multiple Mechanism Prediction,» *Microelectronics Reliability Journal*, vol. 54, pp. 2951-2955, 2014.

[116] BELLCORE, SR-332, Reliability Prediction Procedure for Electronic Equipment, BELLCORE Telcordia, 2011.

[117] "PRISM now 217Plus (TM) RIAC - The Reliability Analysis Center," [Online]. Available: http://www.theriac.org.

[118] C. Liu, H.-S. Sagong, H. Kim, S. Choo, H. Lee and Y. Kim, "Systematical study of 14 nm FinFET Reliability: from device level stress to product HTOL," in *IRPS*, Monterey (CA), 2015.

[119] JESD85, Methods for calculating failure rates in units of FITs, JEDEC Publication, JEDEC Solid State Technology.

[120] J. Srinivasan and S. V. Adce, "The case for microarchitectural awareness of lfetime reliability," Technical report UIUC CS n° UIUCDCS-R-2003-2391, 2003.

[121] A. Bensoussan, A. Durier and J. Bernstein, "M-STORM : A NEW APPROACH TO MODEL ELECTRONICS COMPONENTS RELIABILITY," in *Automotive Power Electronics (APE 2017)*, Paris (France), 2017.

[122] J. Bernstein, "Reliability Prediction for Aerospace Electronics," in *IEEE Aerospace Conference*, Big Sky (MN), 2015.

[123] A. Bravaix, "Part 2 - DDR3 FF20 Reliability Analysis - ISEN," Internal report IRT- ISEN, Toulouse, September 2019.

[124] C. Creveling, "Increasing the reliability of electronic equipment by the use of redundant circuits," *Proc. IRE*, vol. 44, pp. 509-515, 1956.

[125] B. Dhillon and S. Rayapati, "A method to evaluate reliability of three-state device network," *Microelectronics Reliability Journal*, vol. 26, pp. 525-554, 1986.

[126] M. Nikulin, V. Couallier and L. Gerville-Reache, "Accelerated Life and Degradation Models in Reliability and Safety: An Engineering Perspective," in *EA 2961, Statistique Mathematique et ses Applications, Univ. Bordeaux2*, Bordeaux, France.

[127] E. Takeda and N. Suzuki, "An empirical model for device degrada- tion due to hot-carrier injection," *IEEE Electron Device Lett.*, Vols. EDL-4, pp. 111-113, 1983.

[128] D. Verkest, "/cadence_blogs_8/b/breakfast-bytes/posts/imec2," IMEC Roadmap - Cadence Breakfast Bytes Blogs - Post from Paul McLellan, 7 June 2018. [Online]. Available: https://community.cadence.com/. [Accessed 28 February 2019].

[129] N. Loubet and et al., "Stacked Nanosheet Gate-All-Around Transistor to Enable Scaling Beyond FinFET - T17-5," in *Symposium on VLSI Technology Digest of Technical Papers*, Kyoto, Japan, 2017.

[130] Serma Technologies, «Alcatel Espace Internal failure analysis report EL 4103,» Toulouse, March 1996.