7th RTD Framework Program

REALITY

Reliable and Variability Tolerant System-on-a-Chip Design in More-Moore Technologies

Contract No 216537

Deliverable D1.1

Variability simulation and comparison for the 45 nm technology generation

Editor:	Andrew Brown
Co-author / Acknowledgement:	Asen Asenov, Scott Roy, Binjie Cheng, Miguel Miranda, Paul Zuber
Status - Version:	V0.0
Date:	11/12/2008
Confidentiality Level:	Public
ID number:	IST-216537-WP1-D1.1-v0.4.doc

The REALITY Consortium consists of:

Interuniversity Microelectronics Centre (IMEC vzw)	Prime Contractor	Belgium
STMicroelectronics S.R.L. (STM)	Contractor	Italy
Universita Di Bologna (UNIBO)	Contractor	Italy
Katholieke Universiteit Leuven (KUL)	Contractor	Belgium
ARM Limited (ARM)	Contractor	United Kingdom
University Of Glasgow (UoG)	Contractor	United Kingdom

1. Disclaimer

The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

2. Acknowledgements

The editor Andrew Brown acknowledges contributions by Asen Asenov, Scott Roy and Binjie Cheng at UoG, and Miguel Miranda and Paul Zuber at IMEC.

3. Document revision history

Date	Version	Editor/Contributor	*Comments*
11/12/2008	V0.0	Andrew Brown	Partial first draft
04/01/2009	V0.1	Miguel Miranda	Contribution from IMEC
08/01/2009	V0.2	Andrew Brown	Almost complete first draft
09/01/2009	V0.3	Paul Zuber	IMEC contribution update
09/01/2009	V0.4	Andrew Brown	Complete first draft

4. Preface

The scope and objectives of the REALITY project are :

- Development of design techniques, methodologies and methods for real-time guaranteed, energy-efficient, robust and adaptive SoCs, including both digital and analogue macro-blocks“

The Technical Challenges are :

- To deal with increased static variability and static fault rates of devices and interconnects.
- To overcome increased time-dependent dynamic variability and dynamic fault rates.
- To build reliable systems out of unreliable technology while maintaining design productivity.
- To deploy design techniques that allow technology scalable energy efficient SoC systems while guaranteeing real-time performance constraints.

Focus Areas of this project are :

- “Analysis techniques” for exploring the design space, and analysis of the system in terms of performance, power and reliability of manufactured instances across a wide spectrum of operating conditions.

- “Solution techniques” which are design time and/or runtime techniques to mitigate impact of reliability issues of integrated circuits, at component, circuit, architecture and system (application software) design.

The REALITY project has started its activities in January 2008 and is planned to be completed after 30 months. It is led by Mr. Bart Dierickx and Mr. Miguel Miranda of IMEC. The Project Coordinator is Mr Peter Lemmens. Five contractors (STM, ARM, KUL, UoG, UNIBO) participate in the project. The total budget is 2.899 k€.

5. Abstract

This report describes the physical modelling of variability at the 45nm technology node including the generation of statistical compact models. The work is based on the 45nm low-power technology devices provided by the project partners at STMicroelectronics. The modelling is performed by the University of Glasgow using their advanced 3D quantum-corrected drift-diffusion simulator which can account for the different sources of variability that play important role at this technology node. The simulator is calibrated against the data provided by STM and statistical samples of device characteristic, including the appropriate sources of variability, are obtained. From this simulation data, statistical compact models are extracted which allow propagation of the variability information up the 'food-chain' to the circuit and system design stages. IMEC has provided data on the temporal degradation of devices related to formation of defect states. An analysis of the reliability of devices due to different levels of degradation has been performed.

6. List of Abbreviations

REALITY	Reliable and Variability tolerant System-on-a-chip Design in More-Moore Technologies
BTF	Body Thickness Fluctuations
DIBL	Drain-Induced Barrier Lowering
GSM	Gate Stack Morphology
HCD	Hot Carrier Degradation
LER	Line Edge Roughness
LP	Low Power
NBTI	Negative Bias Temperature Instability
OTF	Oxide Thickness Fluctuations
PSG	Poly-Silicon Granularity
RDD	Random Discrete Dopants
SBD	Soft-Breakdown
S/D	Source/Drain
SV	Statistical Variability
TCAD	Technology Computer-Aided Design
VAM	Variability-Aware Modelling

7. List of Tables

8. List of Figures

9. Table of contents

10. Introduction

The aim of WP1 is the physical modelling and understanding of the variability at 45/32 nm technology nodes including the generation of statistical compact models. The statistical variability stems from the discreteness of charge and matter and the statistics of small numbers. The major sources of statistical variability include random discrete dopants (RDD), poly-silicon granularity (PSG), line edge roughness (LER), oxide and body thickness fluctuations (OTF/BTF) and the high-κ gate stack morphology (GSM) [1]. Realistic reliability measurements and accurate estimates of statistical variability are only available for mature technologies. In this work package (WP1) we use 3D statistical numerical simulations of large statistical samples of microscopically different devices in order to estimate statistical variability. We provide reliable quantitative data on which the design technologies and solutions in WP2-WP6 will be based. The simulations are carried out with the University of Glasgow (UoG) 3D physical statistical device simulator which incorporates RDD, PSG, LER, OTF, BTF and GSM as sources of statistical variability [2,3,4]. The simulator can also estimate the gate leakage variability associated with the above statistical variability sources.

Deliverable (D1.1) is based on the 45nm technology of STMicroelectronics (STM). STM have provided technology data files of the doping profiles of their n- and p-channel 45nm low power (LP) devices and calibrated I-V characteristics from TCAD simulation. UoG have included the provided device structure within their simulator and calibrated it against this data as discussed in section 11.

Investigation of variability requires the simulation of a statistical ensemble of microscopically different devices – different random distribution of discrete dopants; different line edge roughness pattern along the gate edge; different pattern of poly-silicon grains, etc. For this work we simulate 200 different devices based on each of the nominal n- and p-channel devices and this is presented in section 12.

Extraction of statistical compact models first requires the extraction of compact models for the continuously doped n- and p-channel devices. This involves the simulation of I_D-V_G characteristics at high and low drain voltages and with a different substrate biases. This is done for the nominal device and also devices of longer channel length. For each of the 200 devices in the statistical sample a compact model can be generated by extracting 7 parameters which can account for the variability, while the remaining parameters are kept from the initial extracted model. Compact model extraction is described in section 13.

Most device reliability problems are associated with generation of fixed charges or the trapping of electrons and/or holes in defect states in the gate stack during circuit operation. IMEC have provided data on expected trap sheet densities for different levels of degradation allowing UoG to run statistical simulations to investigate the impact of stress on device performance at different points in the lifetime of the devices. This allows propagation of reliability information along the design food chain – combining pure TCAD-like simulations and reliability measurements as reliability modelling. Section 14 looks at the statistical reliability issue.

11. Calibration of the 'atomistic' simulator

STM have provided TCAD process simulation data of the doping profiles of their 45nm LP technology, along with I_D-V_G characteristic of their TCAD device simulations calibrated to experimental data. Net doping profiles of the n- and p-channel devices are shown in Figure 1 and Figure 2 respectively. The first step in the work here is to introduce the provided device structures into the UoG 'atomistic' device simulator. The profiles of each dopant species are mapped onto the simulation mesh used by the simulator, and saved as doping files to be read in upon execution of the simulator.

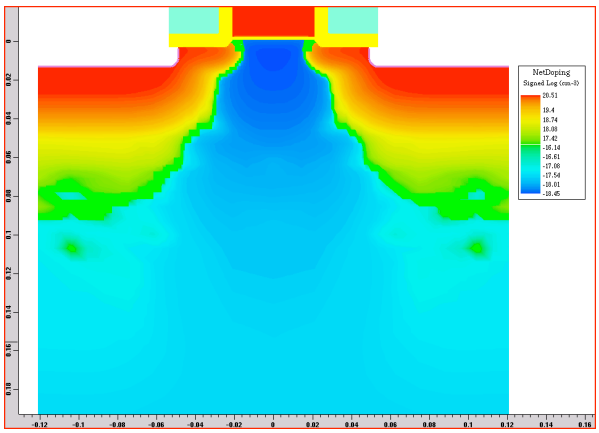


Figure 1: Structure of the n-channel device provided by ST and obtained from TCAD process simulation

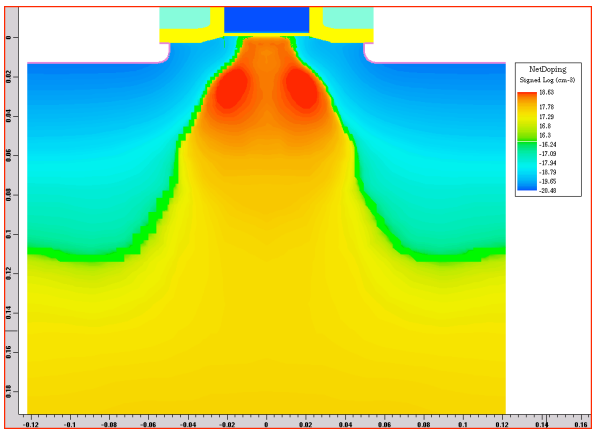


Figure 2: Structure of the p-channel device provided by ST and obtained from TCAD process simulation

The second step is to run the simulator to produce I_D-V_G characteristics at low and high drain voltage and calibrate the simulations to match the TCAD simulation results provided by STM. This is done by adjusting parameters in the mobility models (such as the saturation velocity). The results of this calibration are shown in Figure 3 and Figure 4 for the n- and p-channel devices. Good agreement is achieved in both cases.

Once the simulator has been calibrated to match the provided data, the different sources of variability which are important for this particular bulk MOSFET technology can be introduced. For the n-channel device we include random discrete dopants (RDD), line edge roughness (LER) and Fermi-level pinning at poly-silicon grain boundaries (PSG). For the p-channel device we include RDD and LER. The absence of donor-type interface states at the grain boundaries of p-type poly-silicon means that Fermi-level pinning along these boundaries will not be a source of variability in p-channel devices [5]. Figure 5 shows the electrostatic potential in the n- and p-channel devices with these sources of variability included. The surface potential is shown above each device demonstrating the source-to-drain barrier due to the doping being modulated by the different sources of variability. In particular, the sharp potentials due to RDD are clear with the attractive potential wells in the source/drain and the repulsive barrier spikes in the channel. The reason for the variation in device characteristics is clear from this picture as the potential in each device will be affected differently depending on the particular configuration of the RDD, LER and PSG.

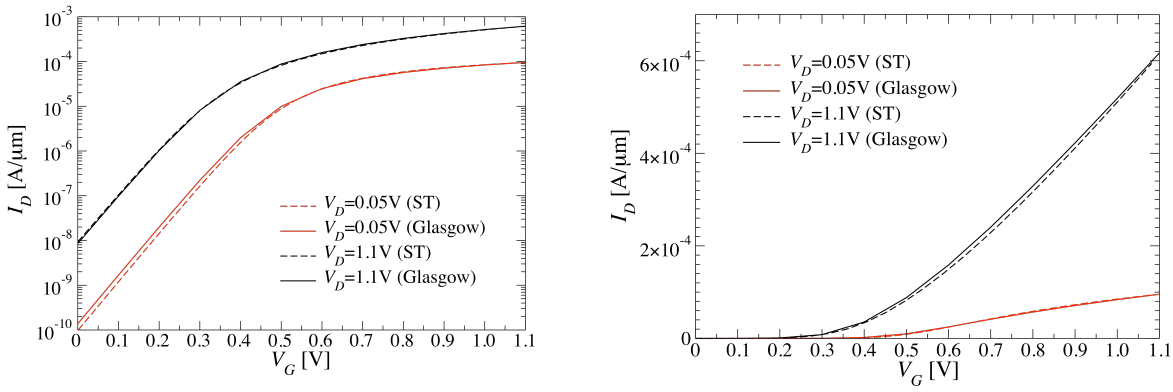


Figure 3: Calibration of the Glasgow 'atomistic' simulator against TCAD simulation of the ST 45nm n-channel device

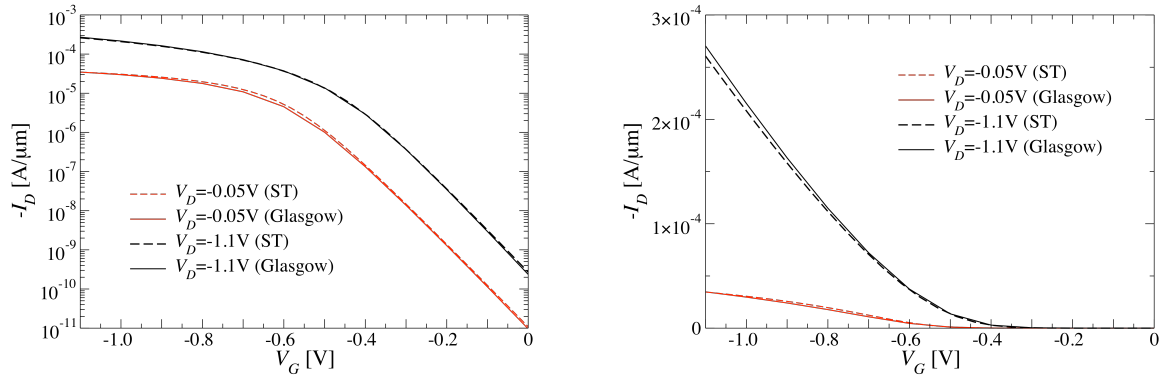
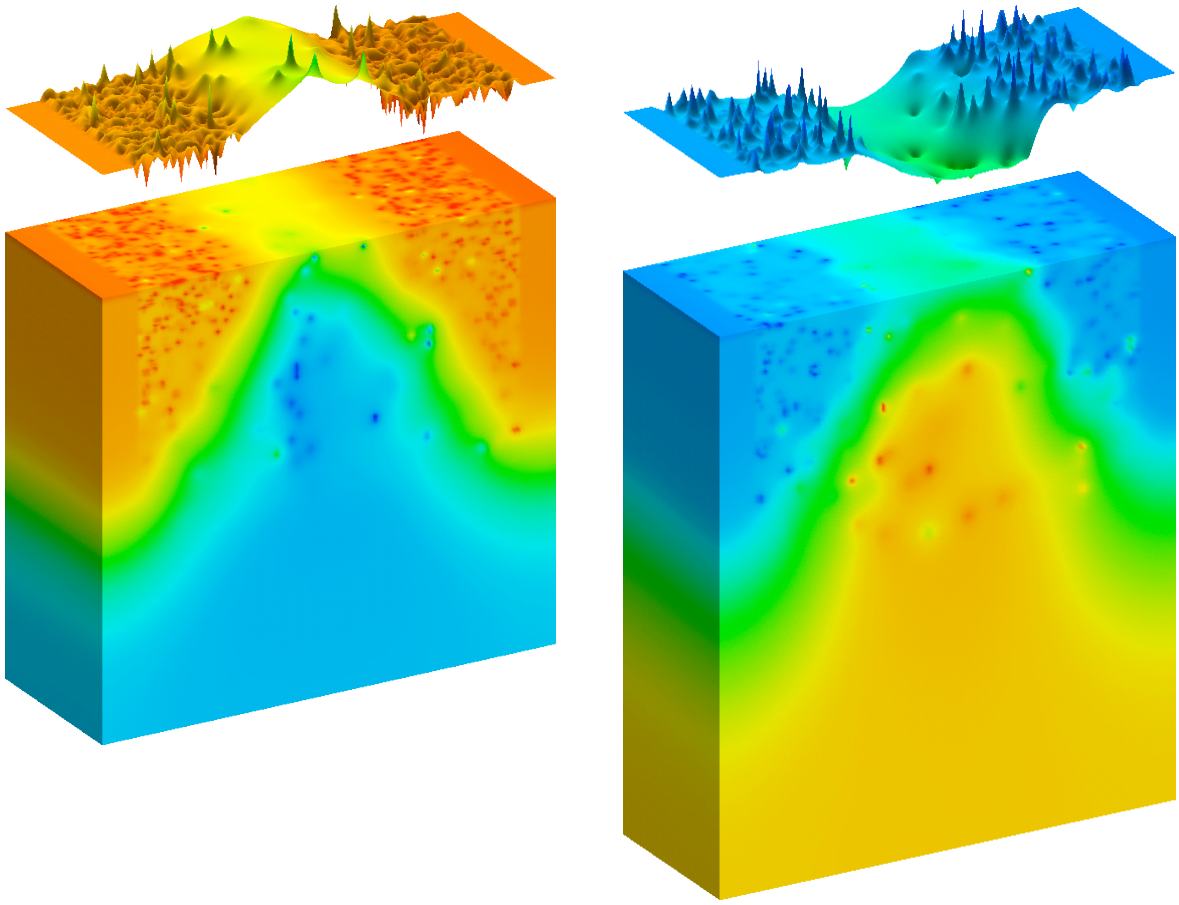


Figure 4: Calibration of the Glasgow 'atomistic' simulator against TCAD simulation of the ST 45nm p-channel device



(a) (b)

Figure 5: Electrostatic potential in the (a) n-channel and (b) p-channel devices with sources of variability included.

12. Simulation of statistical ensembles of devices

In order to extract statistical compact models it is necessary to simulate an ensemble of devices; in this case we simulate 200 microscopically different devices. Unlike in the PULLNANO project where only the variation in threshold voltage was investigated, here we require full I_D-V_G curves in order to allow extraction of the compact model for each different device. Figure 6 to Figure 9 show the ensemble of I_D-V_G curves for the n- and p-channel devices at low (V_D=0.05V) and high (V_D=1.1V) drain voltages. Each set of 200 I-V curves requires approximately 3-5 CPU-years of simulation time depending on the drain bias, highlighting the computational expense of the task, and demonstrating the need for large computer clusters in order to perform such work.

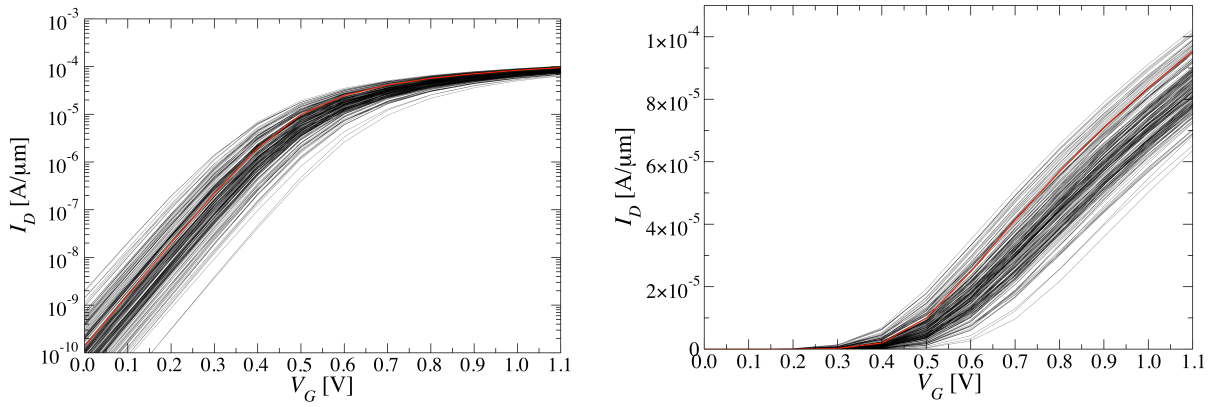


Figure 6: I_D-V_G curves on (a) log and (b) linear scale for a statistical ensemble of 200 n-channel devices. V_D=0.05V

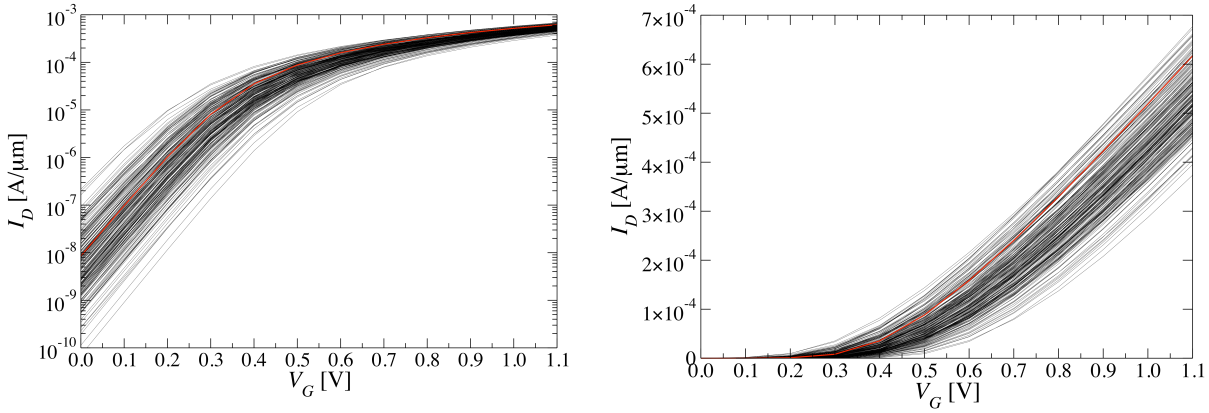


Figure 7: I_D-V_G curves on (a) log and (b) linear scale for a statistical ensemble of 200 n-channel devices. V_D=1.1V

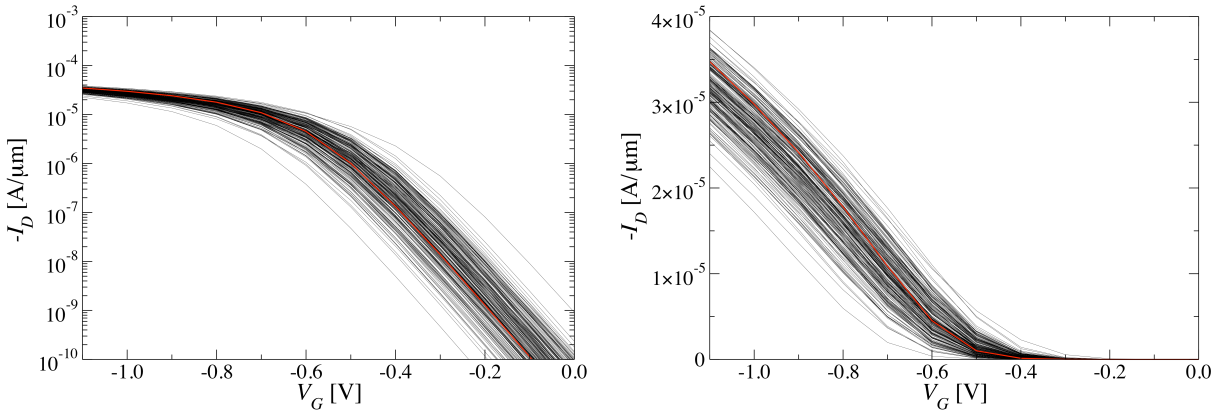


Figure 8: I_D-V_G curves on (a) log and (b) linear scale for a statistical ensemble of 200 p-channel devices. V_D=-0.05V

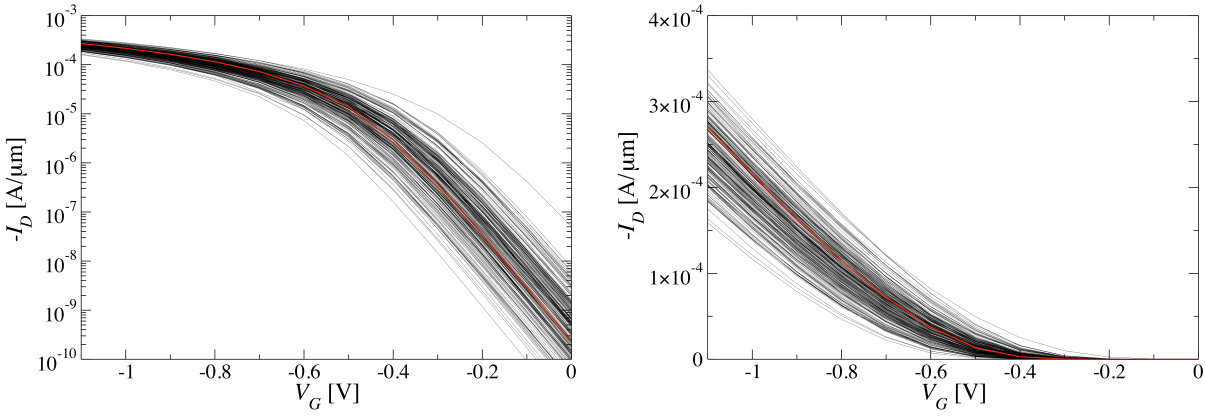


Figure 9: I_D-V_G curves on (a) log and (b) linear scale for a statistical ensemble of 200 p-channel devices. V_D=-1.1V

13. Extraction of compact models

Compact models are the interface between technology and design. The approach adopted here for this interface is to generate a different compact model card for each of the 200 different devices. Then, during the statistical circuit simulation stage a Monte Carlo approach is used, whereby for each transistor in the circuit, one compact model card is selected at random from the ensemble and used to characterise that particular transistor.

Here, a two-stage statistical compact model parameter extraction strategy is implemented [6]. During the first stage, a combination of local optimisation and group extraction strategy is employed to extract the complete set of BSIM4 parameter using the Synopsys tool Aurora. This is based on the simulation of a set of transistors with *continuous doping concentration* and different channel lengths, focusing on those critical to long channel behaviour, the threshold voltage in the short channel regime, and drain current response in the presence of high fields. Figure 10 shows the quality of the BSIM4 parameter extraction results obtained for the n-channel MOSFET, and Figure 11 for the p-channel, at the 45nm LP technology node, where the symbols represent the I-V points obtained from the device simulation and the curves are the compact model fitted to this data.

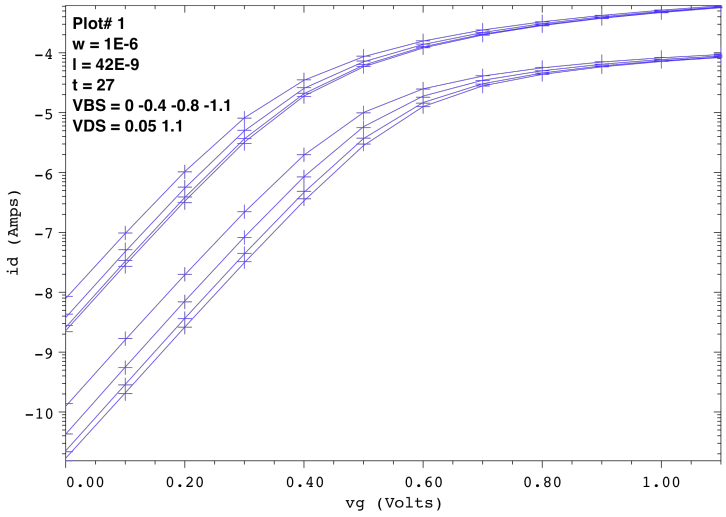
From a device operation aspect, statistical variability (SV) sources have the following consequences: firstly, they cause off-current and subthreshold slope variations in the weak inversion regime; secondly, they cause the well documented threshold voltage variation; thirdly, they cause variation in device characteristics at moderate inversion; fourthly, they cause S/D resistance variation; finally, they cause variation in DIBL effects. The parameters extracted from the first stage are divided into two groups: those which have been found to be relatively independent of statistical variation that are then fixed during the next stage, and those sensitive to the statistical variation which will be re-extracted for each 'atomistically' different device at the second stage.

Seven key BSIM4 parameters are chosen to represent the effect of statistical variability sources, where:

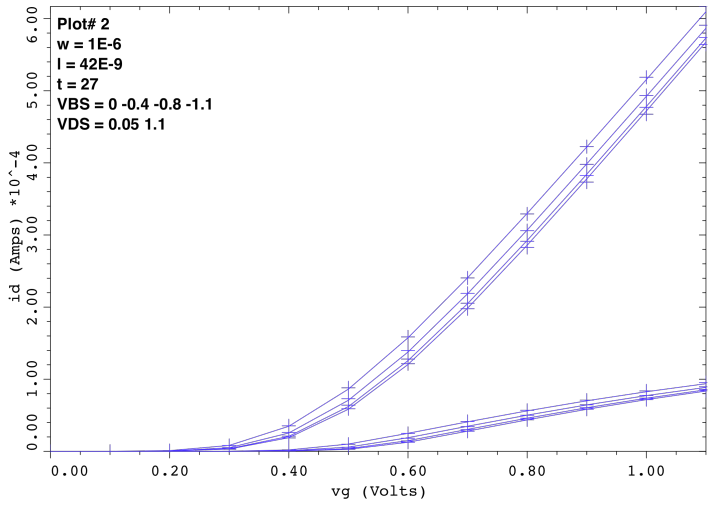
- *Vth0* is selected to account for traditional threshold voltage variation
- *U0* is selected to account for current factor variation
- *Voff* and *Nfactor* account for off-current and subthreshold slope variation
- *Minv* is selected to account for variation in the moderate inversion region
- *Rdsw* is a basic S/D resistance parameter, and is selected to account for dopant variation in the S/D region
- *Dsub* is employed to account for DIBL variation.

Only I_D–V_G characteristics are used for second extraction stage.

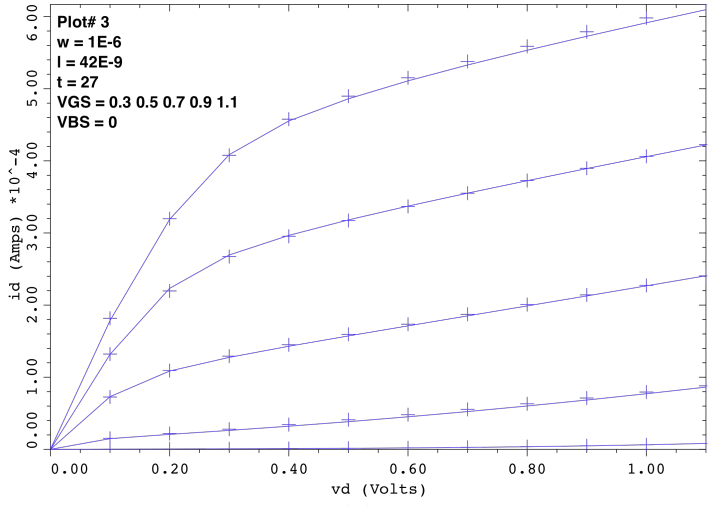
Figure 12 and Figure 13 show the distribution of RMS error for the statistical extraction of n- and p-channel devices respectively. The error in the uniform extraction of the first stage is included for comparison. The error is the linear difference between the I-V data obtained from the device simulation (to which the compact model is being fitted) and the I-V data produce by the final compact model obtained after re-extracting the 7 above parameters. The calculated RMS error includes all points throughout the range of applied gate voltage and both low and high drain bias characteristics. These figures show that the RMS error is below 5% for the n-channel device and below 7% for the p-channel device which good when considering the variation in characteristics observed in section 12.



(a)

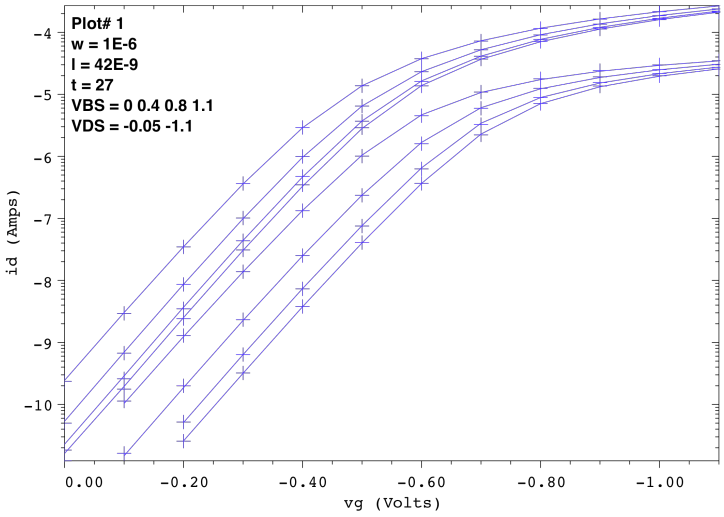


(b)

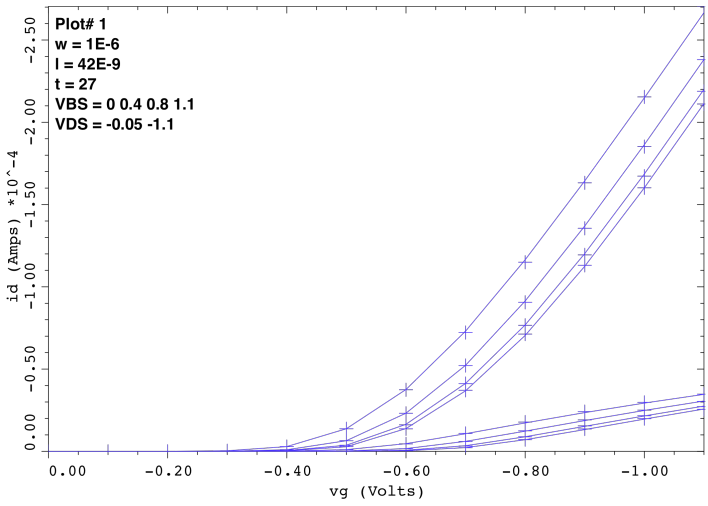


(c)

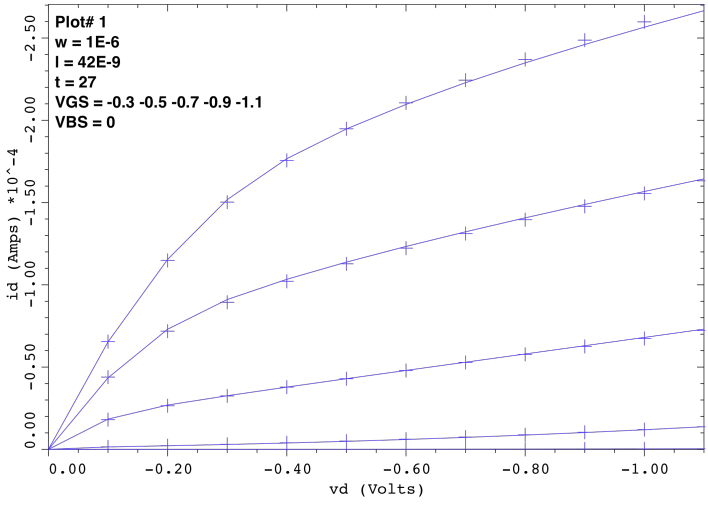
Figure 10: Extraction of uniform n-channel compact model. (a) I_D-V_G curves on a log scale, (b) I_D-V_G curves on a linear scale and (c) I_D-V_D curves.



(a)



(b)



(c)

Figure 11: Extraction of uniform p-channel compact model. (a) I_D-V_G curves on a log scale, (b) I_D-V_G curves on a linear scale and (c) I_D-V_D curves.

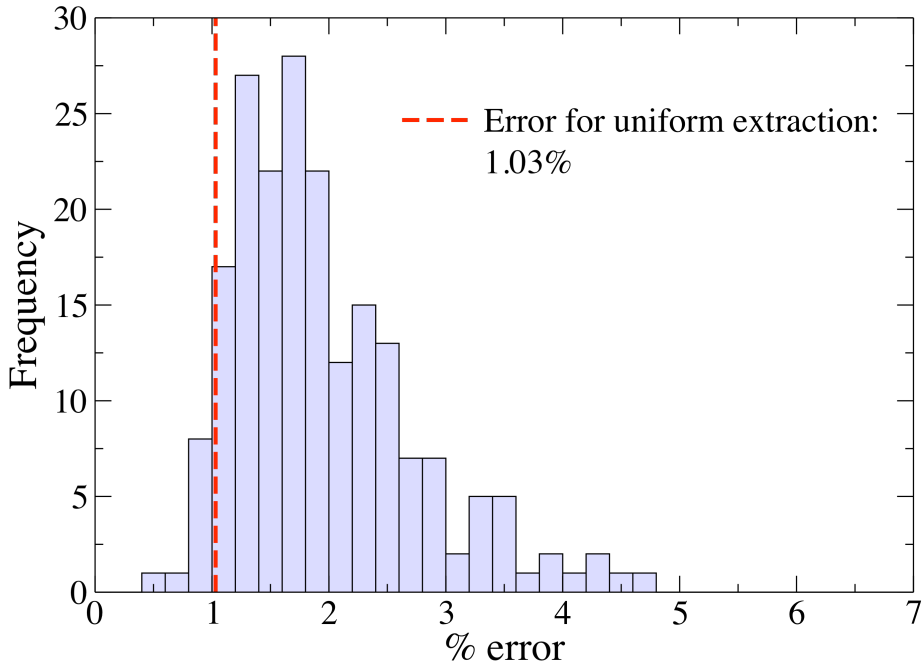


Figure 12: Distribution of RMS errors in the extraction of 200 statistical compact models for the n-channel device.

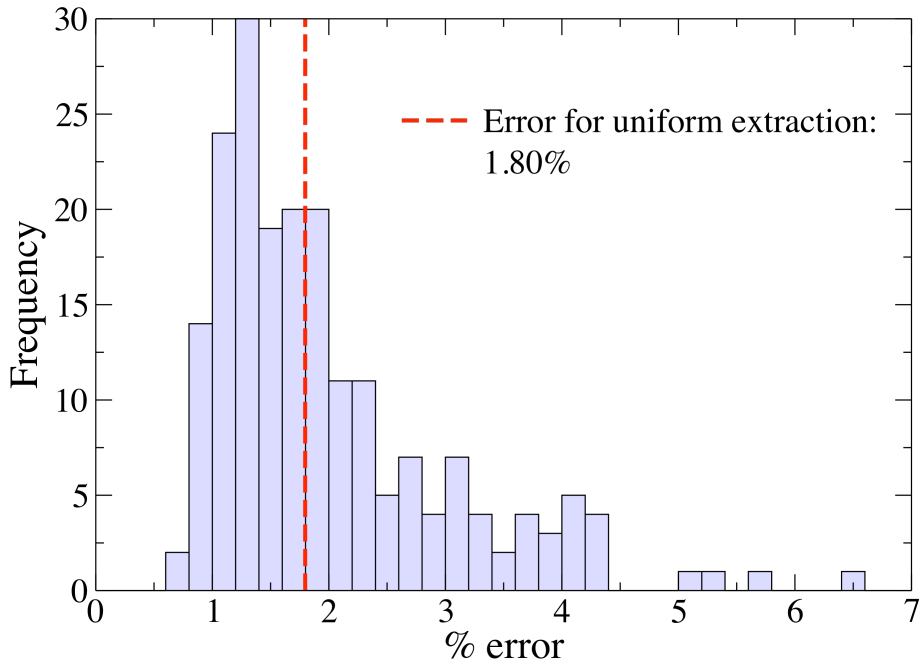


Figure 13: Distribution of RMS errors in the extraction of 200 statistical compact models for the n-channel device.

14. Statistical reliability

Most device reliability problems are associated with generation of fixed charges or the trapping of electrons and/or holes in defect states in the gate stack during circuit operation. The expected introduction of high-κ dielectrics will exacerbate these problems due to higher defect densities and lower dielectric stability. Due to the discrete nature of the fixed charges and trapped carriers, reliability problems in sub 45 nm technology devices will have a distinct statistical manifestation. For a particular density of hot carrier/NBTI/stress generated defects the change in the characteristics of any nano-CMOS device during circuit operation will vary depending on the actual position of single or multiple defect charges in the channel region of the transistor. Using experimental data for the stress dependence of the average density of the stress generated charge and defect states, and their spatial positions, we carry out a detailed 3D simulation of the statistical distribution of the corresponding changes in device characteristics due to microscopic manifestations of particular device defect configurations. In combination with measured data for the energy distribution and the capture cross section of the defect states, this will eventually allow the development of statistical time dependent reliability models.

IMEC has provided data to UoG that an appropriate level of degradation to investigate for this technology would be with traps densities in the range of $10^{11}cm^{-2}$ to $10^{12}cm^{-2}$. Therefore three levels of degradation have been investigated with trap densities of $10^{11}cm^{-2}$ $5\times10^{11}cm^{-2}$ and $10^{12}cm^{-2}$.

In a statistical ensemble of devices it is expected that the number of traps in a device, for any given level of degradation, will follow a Poisson distribution. Figure 14 shows the distribution in the number of traps in 1000000 instances of a 100nm × 45nm region at the interface of the p-channel device.

The effect of these fixed charges cannot be considered alone as their effect is inextricably linked to the other sources of variability, in particular the other discrete charges due to RDD [7]. An additional stress-induced charge will have a large impact if it occurs along a favoured current path where there are no discrete dopants, but will have little effect if it occurs within a cluster of dopants which already exclude current flow from that region of the channel. For this reason the starting point is the 200 devices already simulated with sources of variability, and to each device are added additional fixed charges, randomly within the channel, the number of which will follow the distribution observed in Figure 14. I_D-V_G curves are simulated for the three levels of degradation noted above and these are shown in Figure 15 to Figure 17 for both n- and p-channel devices.

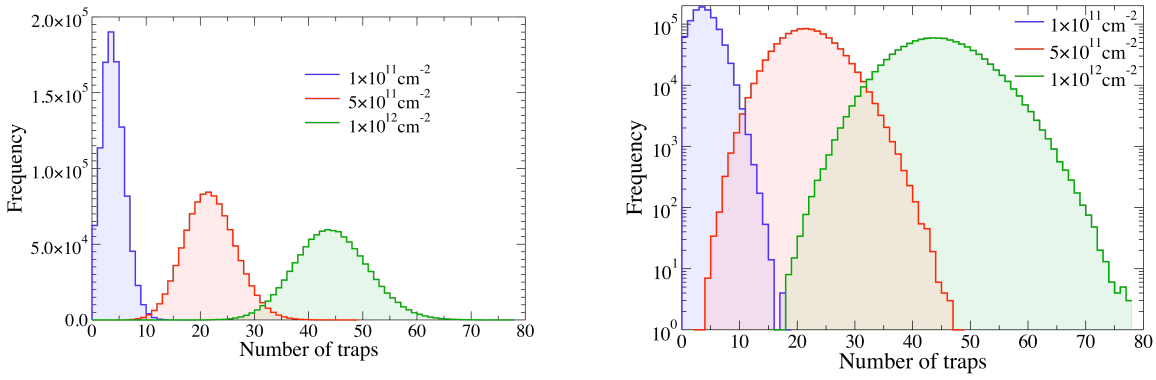


Figure 14: Distribution of the number of traps at the interface of the simulated p-channel device for different levels of degradation on a (a) linear and (b) log scale. 1,000,000 devices were generated.

This additional channel charge will act to increase the absolute threshold voltage, however the shift in V_T will depend on the number and location of the additional charges. As NBTI is the major reason for degradation in this technology, and it affects p-channel devices, we examine in detail the results of the p-channel simulations. Figure 18 and Figure 19 show the distributions of V_T and the shift in V_T (ΔV_T) for the different levels of degradation. It is clear that not only does the shift increase with increasing trap density but also the spread in V_T. This is quantified in Figure 20 and Figure 21 which show a linear increase in the average V_T along with an increase in the standard deviation, σV_T.

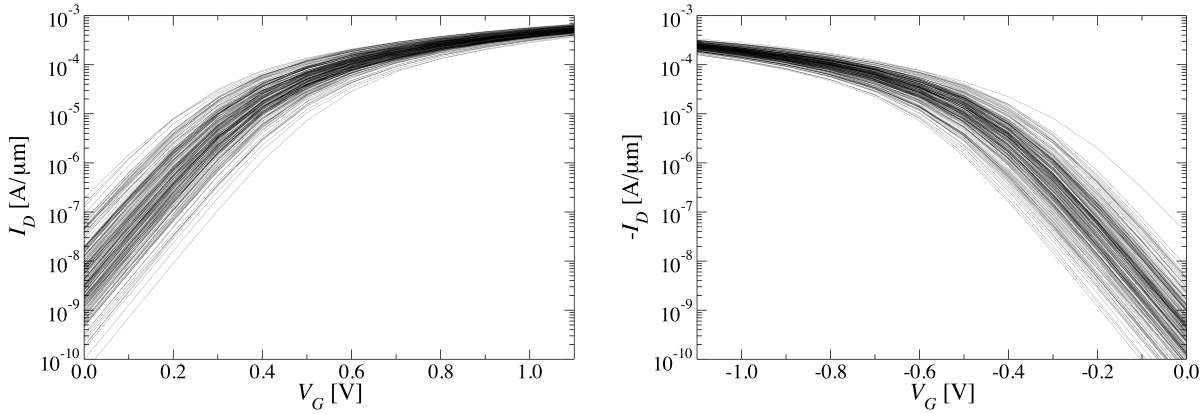


Figure 15: I_D-V_G curves for the 200 simulated (a) n-channel and (b) p-channel devices with degradation trap density of $1\times10^{11}cm^{-2}$.

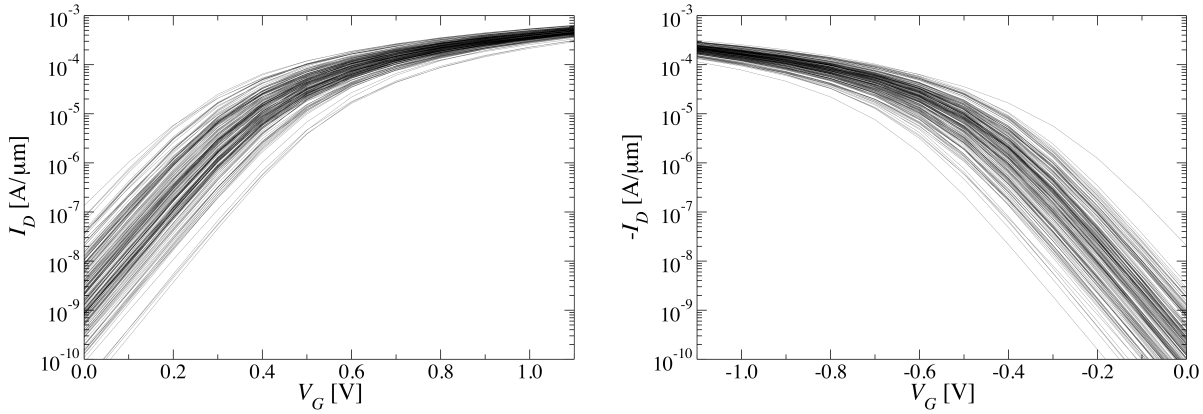


Figure 16: I_D-V_G curves for the 200 simulated (a) n-channel and (b) p-channel devices with degradation trap density of $5\times10^{11}cm^{-2}$.

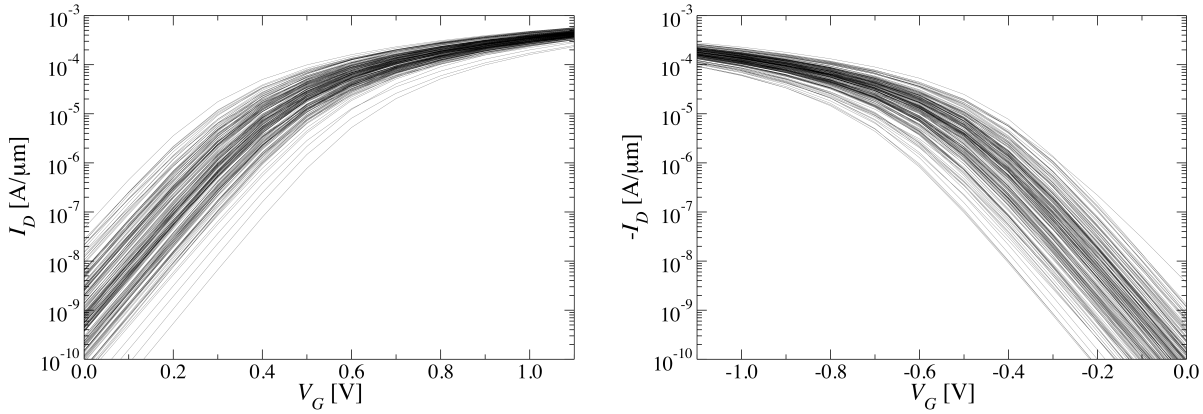


Figure 17: I_D-V_G curves for the 200 simulated (a) n-channel and (b) p-channel devices with degradation trap density of $1\times10^{12}cm^{-2}$.

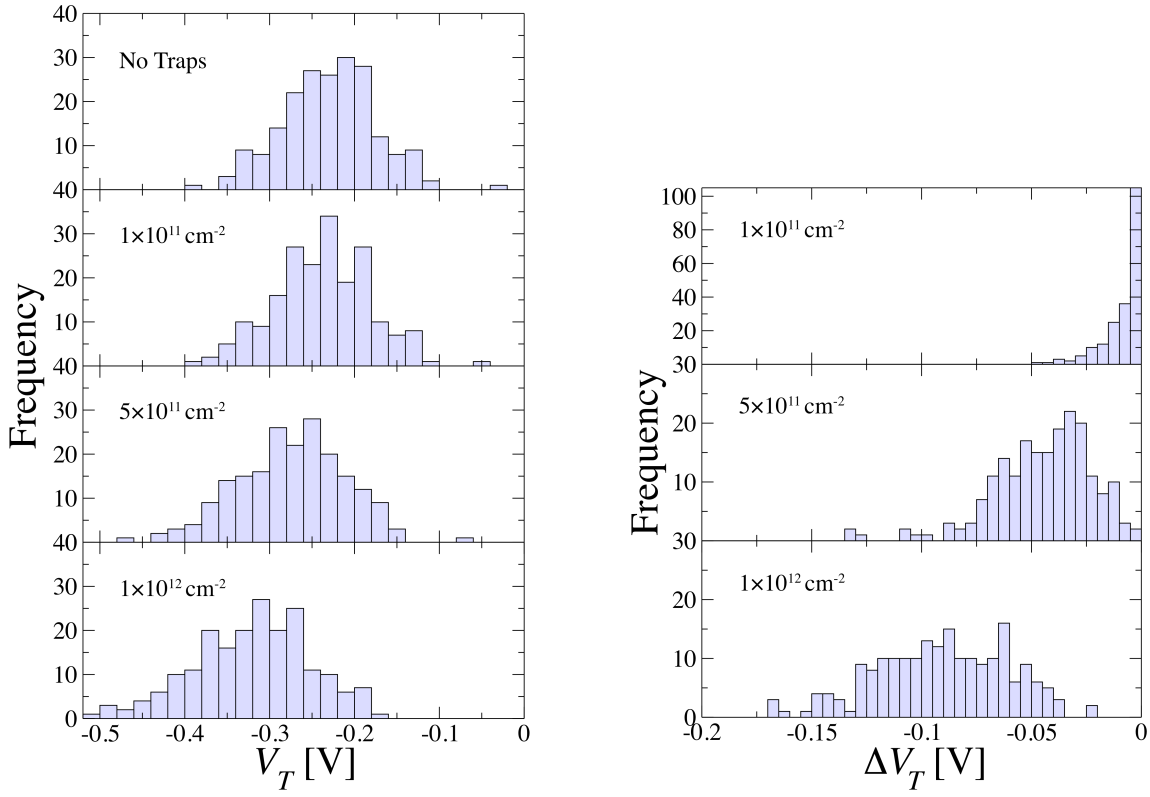


Figure 18: Distribution of threshold voltages for different levels of degradation of the p-channel device and the associated shift in V_T from the trap-free device. V_D=-1.1V

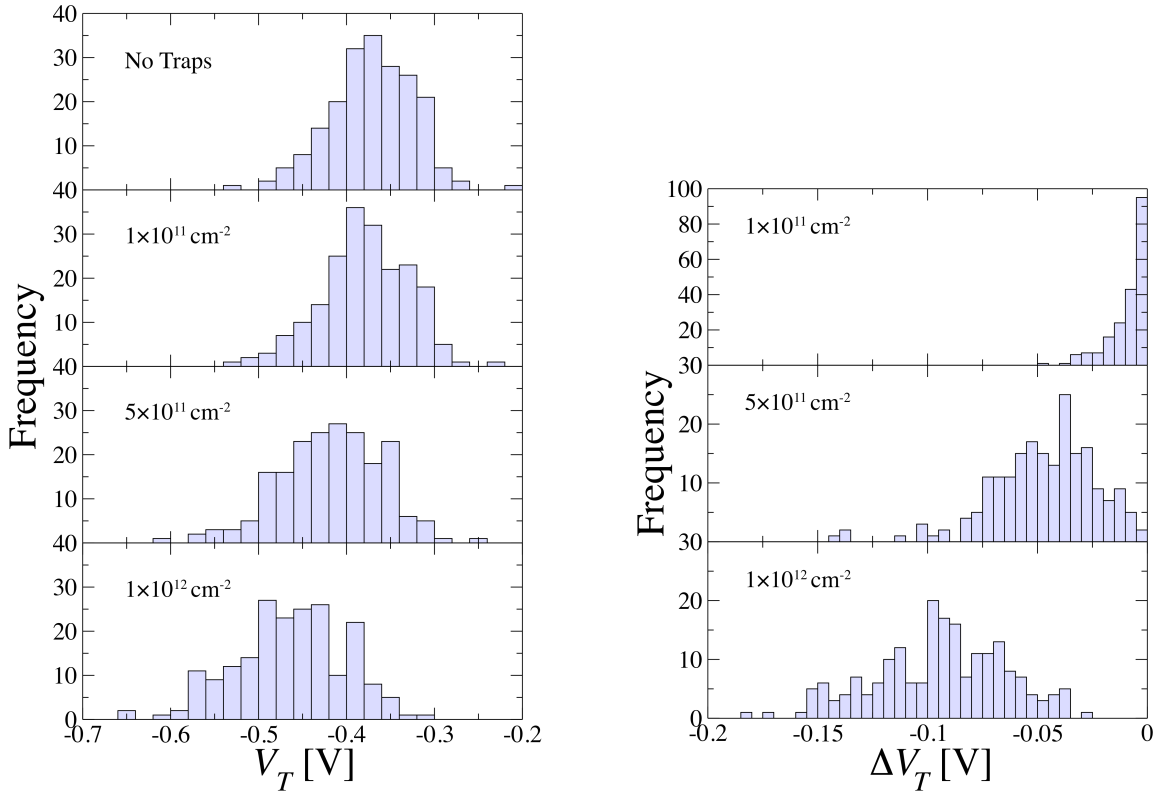


Figure 19: Distribution of threshold voltages for different levels of degradation of the p-channel device and the associated shift in V_T from the trap-free device. V_D=-0.05V

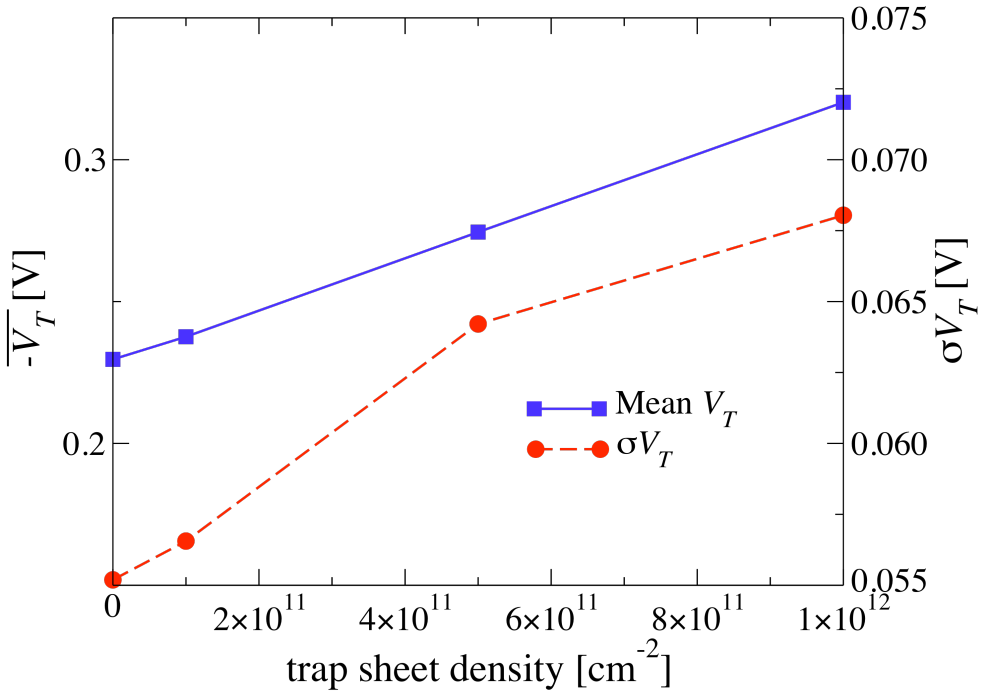


Figure 20: Average and standard deviation of V_T for different levels of degradation of the p-channel device. V_D=-1.1V

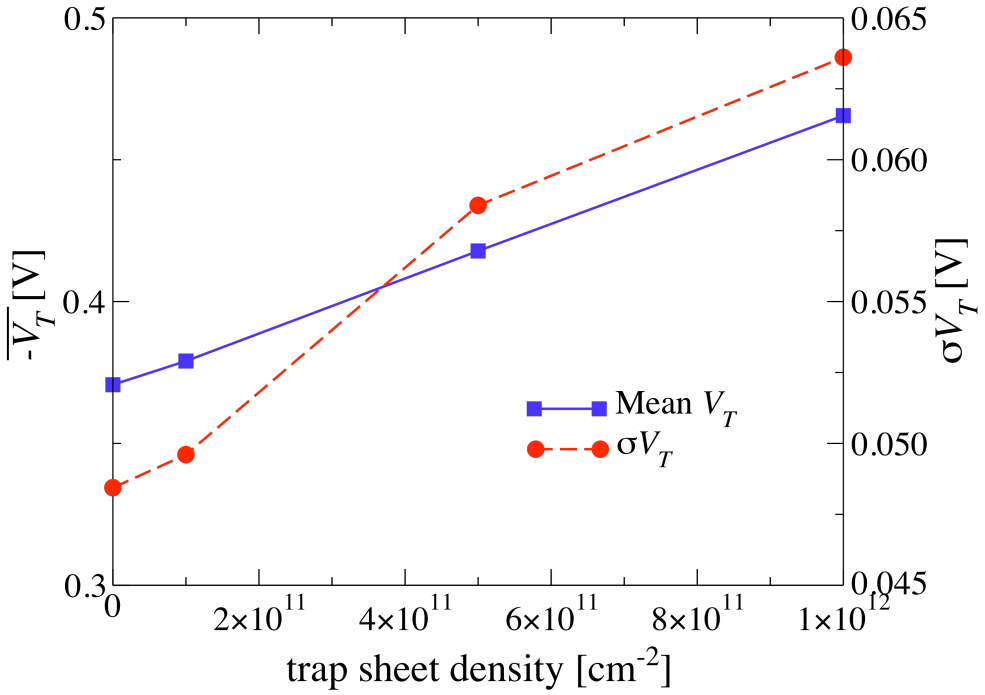


Figure 21: Average and standard deviation of V_T for different levels of degradation of the p-channel device. V_D=-0.05V

15. Modelling of Variability Injectors

Most variability parameters and several degradation mechanisms are condensed into a ΔVth and other netlist components. For that purpose every MOSFET in a SPICE netlist is reparsed as depicted in Figure 22.

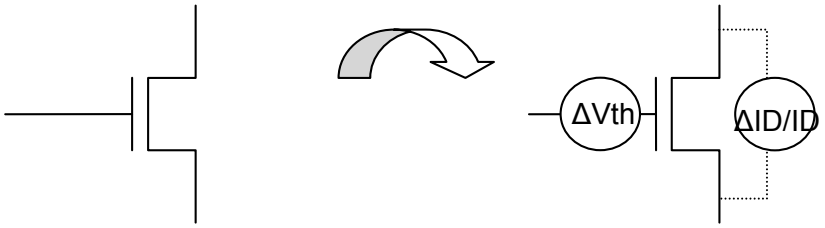


Figure 22: Automatic insertion of voltage source before gate to mimic ΔVth variation and current-controlled current source along source drain to mimic $\Delta I_D/I_D$ variation.

The ΔVth in this scheme is a series voltage source added to the netlist; the $\Delta I_D/I_D$ (Δbeta) Is a current dependent current source added to the netlist.

Conventions for denominating distributions:

- Vth, W, T: means the nominal, average or invariable value
- ΔVth, ΔW, ΔT: static shifts away from the above value

If a parameter is called “delta_vth”, it means that it is referred to a certain “vth” in a relative fashion. The invariable case of a “delta_something” is exactly zero.

- σVth, σW, σT: “distribution” of the parameter.

In the case of a Gaussian distribution, “σ” means standard deviation, but here we use the symbol σ in a generalised way, implying local and/or global types of variations. Any variation type distribution *may* be given explicity such as σ_MVth, σ_LVth, σ_CVth, σ_WVth, σ_BVth. We advise to refrain from using algebra on σs that is valid for true Gaussians only.

Note that in the variability injector modelling approach, the MOSFET, or any transistor, itself is a blackbox. Its *modelcard* is NOT changed. This allows using any compact model, macro, subcircuit, with minimal invasion in the existing simulation flow.

Pelgrom’s rule for W and L scaling *applies to local random variations only.* $\sigma(\Delta Vth)_{ref}$ and $\sigma(\Delta I_D/I_D)_{ref}$ are the variability for the nominal transistor size. ΔVth and Δbeta in this approach are assumed to already include the variability effects of dopant and interface state fluctuations, CD variations, LER variations and layer thickness variations, and more. If one chooses to ropagate one or more of DF, ISF, CDV, LER or LTV separately, we should make sure that their effect is taken out of the ΔVth and Δbeta parameters. In such case, one could e.g. implement CD variations and LER by directly impacting the L and W parameters of the MOSFET modelcard. The IMEC approach for variability injection and baseline however is that ΔVth and Δbeta do include all mentioned effects.

For devices of different sizes, Pelgrom’s rule for W and L scaling applies. However this is only in the case of local variations

$$\sigma(\Delta Vth) = \sigma(\Delta Vth_{ref}) \frac{\sqrt{(W_{ref}+\Delta W)(L_{ref}+\Delta L)}}{\sqrt{(W+\Delta W)(L+\Delta L)}}$$

And a similar law applies to $\Delta I_D/I_D$:

$$\sigma\left(\frac{\Delta I_D}{I_D}\right)=\sigma\left(\frac{\Delta I_D}{I_D}\right)_{ref}\frac{\sqrt{(W_{ref}+\Delta W)(L_{ref}+\Delta L)}}{\sqrt{(W+\Delta W)(L+\Delta L)}}$$

having nominal design size W_{ref} and L_{ref}. Often the nominal size is 1um. For all other variation types (e.g., global die to die, wafer to wafer and/or batch to batch), $\sigma(\Delta Vth)$ and $\sigma(\Delta\beta)$ are independent of W and L.

At this moment we assume that there is no significant scaling of $\sigma(\Delta Vth)$ and $\sigma(\Delta\beta)$ with temperature and V_{DS}.

Extraction of variability injectors for the Glasgow 45nm device simulations

The method to extract deltaVt and deltaBeta is not a simple vertical and horizontal cut of the variability set of I/V curves at some value of Vgs and Id. This is because we start on the assumption that in reality deltaVt and deltaBeta are symmetrically distributed around the typical device. We can then calibrate this typical device and compare it with the device corresponding to the atomistic average.

Figure 23 depicts the transfer Vgs/Id characteristic and scatter plot deltaVt/deltaBeta for nmos at high Vdd=1.1V. The blue circles of the scatter plot represent variability cloud deltaVt/deltaBeta with respect to the atomistic average. The blue lines represent mean values. The black circles represent the same scatter plot but moved to be centred around the mean at zero points (deltaVt = 0V, deltaBeta = 0), which corresponds to the assumption of symmetry. Finally, shown on the transfer characteristic are the atomistic average (blue curve) and the calibrated typical device (black curve).

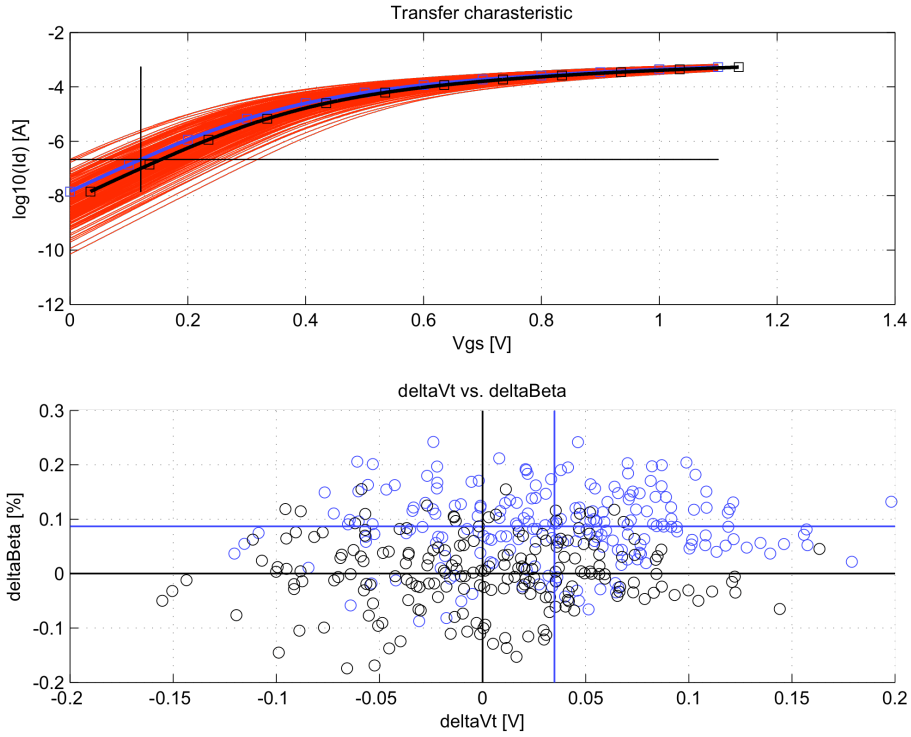


Figure 23: Transfer characteristic for Vgs vs. I_D (above) and scatter plot for ΔVth vs. $\Delta I_D/I_D$ (below).

16. Modelling of Reliability Mechanisms

This Section describes how a reliability model must be set up in order to be implementable using the concept of injectors described in Section 15.
Essentially such a description consists of 4 parts.

1. Analytical or algorithmic model (say, programming language source-code), containing design values and technology parameters, which describes (a) network element(s) for insertion in a SPICE netlist.
2. Statistical distribution of the parameters in that model
3. Scaling rules allowing translation of the effect to a different case that is not in the measured set
4. Scaling the cumulative effect of multiple different stress conditions

Hot Carrier Degradation

We propose to follow the approach of Chittoor Parthasarathy [8] compiled in his PhD.
The degradation relating to created interface states, and hence to saturation current, gm, weak inversion slope and Vth, is expressed as:

$$\Delta D(t) = \left(\frac{I_D}{WH} \left(\frac{I_B}{I_D} \right)^m t \right)^n$$

Where ΔD is proportional to the "damage" in terms of interface states, n is about ½, m is about 3 (depending on V_{GD}) and H is a technology constant. We assume that we can calibrate the formula with a reference measurement, thus:

$$\frac{\Delta V_{th}(t)}{\Delta V_{th\,REF}(t_{REF})} \; or \; \frac{\frac{\Delta gm(t)}{gm}}{\frac{\Delta gm_{REF}(t_{REF})}{gm}} = \frac{\Delta D(t)}{\Delta D_{REF}(t)} = \frac{\left(\frac{I_D}{WH} \left(\frac{I_B}{I_D} \right)^m t \right)^n}{\left(\frac{I_{DREF}}{W_{REF} H} \left(\frac{I_{BREF}}{I_{DREF}} \right)^m t_{REF} \right)^n}$$

which allows the prospect of straightforward implementation in the VAM IF, as ΔV_{TH} and $\Delta g_m/g_m$ can be modelled with an analog netlist element. The variability is then easily brought in via variability of these parameters themselves.
We *silently assume* that we can interchange $\Delta g_m/g_m$ and $\Delta I_D/I_D$, thus:

$$\frac{\Delta V_{th}(t)}{\Delta V_{th\,REF}(t_{REF})} \; or \; \frac{\frac{\Delta I_D(t)}{I_D}}{\frac{\Delta I_{DREF}(t_{REF})}{I_D}} = \frac{\left(\frac{I_D}{W} \left(\frac{I_B}{I_D} \right)^m t \right)^n}{\left(\frac{I_{DREF}}{W_{REF}} \left(\frac{I_{BREF}}{I_{DREF}} \right)^m t_{REF} \right)^n}$$

The degradation under bias conditions that have been different during different time spans t_1, t_2, … is a straightforward extension:

$$\frac{\Delta V_{th}(t)}{\Delta V_{th\,REF}(t_{REF})} or \frac{\frac{\Delta I_D(t)}{I_D}}{\frac{\Delta I_{DREF}(t_{REF})}{I_D}} = \frac{\left(\frac{I_{D1}}{W}\left(\frac{I_{B1}}{I_{D1}}\right)^m t_1 + \frac{I_{D2}}{W}\left(\frac{I_{B2}}{I_{D2}}\right)^m t_2 + \ldots\right)^n}{\left(\frac{I_{DREF}}{W_{REF}}\left(\frac{I_{BREF}}{I_{DREF}}\right)^m t_{REF}\right)^n}$$

This formula is usable as such when I_B or I_B/I_D is known by the simulator. Often this is not the case. In order to tackle that situation, we include in the rule Parthasarathy's approach (with some pragmatic simplification) to calculate I_B/I_D:

$$I_B = \frac{A_i}{B_i} E_m l_c I_D \exp(-\frac{B_i}{E_m})$$

Where

$$E_m = \frac{V_{DS} - V_{DSAT}}{l_c} \text{ and } V_{DSAT} \cong V_{GS} - V_{TH}$$

Hence

$$\frac{I_B}{I_D} = \frac{A_i}{B_i}(V_{DS} - V_{GS} + V_{TH})\exp(-\frac{B_i l_c}{V_{DS} - V_{GS} + V_{TH}})$$

for $V_{DS}>V_{GS}-V_{TH}$, (for an NMOSFET). Otherwise I_B is just zero, and we reduce to two technology constants A and V_B:

$$\frac{I_B}{I_D} = A(V_{DS} - V_{GS} + V_{THref})\exp(-\frac{V_B}{V_{DS} - V_{GS} + V_{THref}})$$

$$\log\left(\frac{(I_B/I_D)}{A(V_{DS} - V_{GS} + V_{THref})}\right) = -\frac{V_B}{V_{DS} - V_{GS} + V_{THref}}$$

A and V_B are approximately constant for a given type of MOSFET in a given technology. A is not of importance as it is eliminated in the HCD formula. V_B must be obtained from a I_B/I_D calibration measurement plotted as in Figure 24. Temperature dependence is neglected.

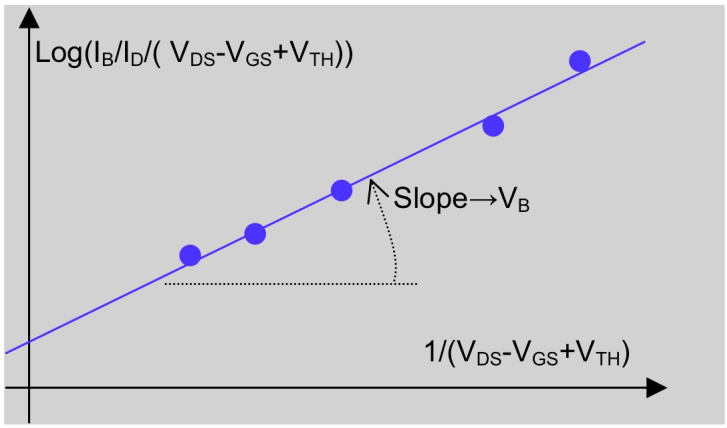


Figure 24: Calibration of the Hot Carrier model.

Although conceived for NMOS, the overall HCD formula might be applied to PMOS. In that case, $V_{DS}>V_{GS}-V_{TH}$. In a PMOS the absolute value of V_{TH} decreases due to HCD, and the g_m decreases. Typically one chooses to neglect the effect of HCD in PMOS.

Scaling rules for design parameters

$$\frac{\Delta V_{th}(t)}{\Delta V_{th\,REF}(t_{REF})} \; or \; \frac{\dfrac{\Delta I_D(t)}{I_D}}{\dfrac{\Delta I_{DREF}(t_{REF})}{I_D}} = \frac{\left(\dfrac{I_D}{W}\left(\dfrac{I_B}{I_D}\right)^m t\right)^n}{\left(\dfrac{I_{DREF}}{W_{REF}}\left(\dfrac{I_{BREF}}{I_{DREF}}\right)^m t_{REF}\right)^n}$$

Allows scaling for different W. L is not explicit in this formula, I_D/I_B depends on L. One must of course know I_D and I_B either from the simulations or from

$$\log\left(\frac{(I_B/I_D)}{A(V_{DS}-V_{GS}+V_{THref})}\right) = -\frac{V_B}{V_{DS}-V_{GS}+V_{THref}}$$

Scaling rules for multiple stress conditions

$$\frac{\Delta V_{th}(t)}{\Delta V_{th\,REF}(t_{REF})} \; or \; \frac{\dfrac{\Delta I_D(t)}{I_D}}{\dfrac{\Delta I_{DREF}(t_{REF})}{I_D}} = \frac{\left(\dfrac{I_{D1}}{W}\left(\dfrac{I_{B1}}{I_{D1}}\right)^m t_1 + \dfrac{I_{D2}}{W}\left(\dfrac{I_{B2}}{I_{D2}}\right)^m t_2 + \ldots\right)^n}{\left(\dfrac{I_{DREF}}{W_{REF}}\left(\dfrac{I_{BREF}}{I_{DREF}}\right)^m t_{REF}\right)^n}$$

indicates that the stress histories can be added before taking the result to the power n.

Time-Dependent Dielectric Breakdown of MOSFETs

Piece-wise approximation model

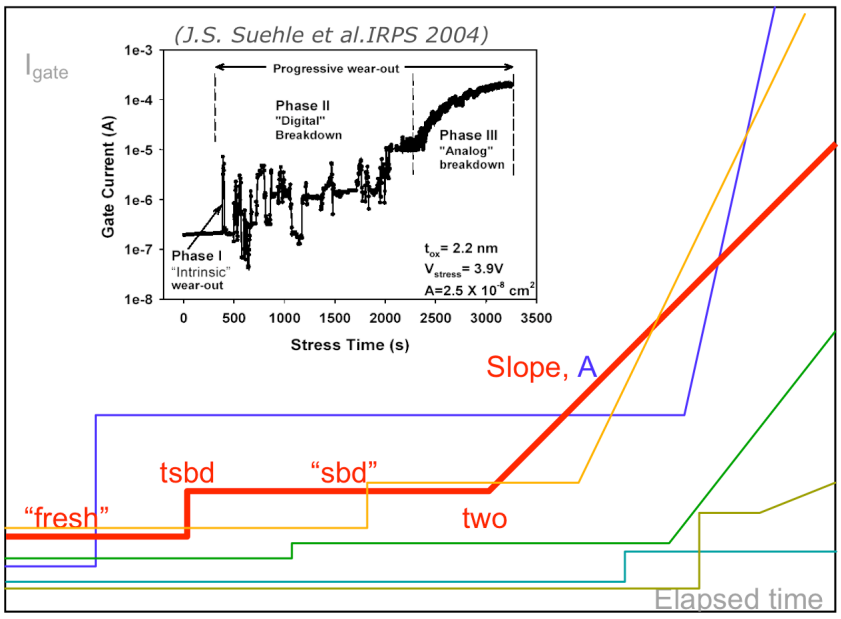


Figure 25: Gate current vs. stress time. Measurement data from literature (above) and model proposed (below). The model considers variations in the effect by considering variations in the parameter set (fresh resistance, time-to-wear-out, soft-breakdown resistance, etc.).

- Soft breakdown (sbd) happens after a time *tsbd*. After that the initial "fresh" I-V characteristic changes to an "sbd" I-V characteristic. This I-V characteristic is strongly non-linear.
- Wear-out (wo) happens after a time *two*, after which the device goes in "hard breakdown" (hbd). The I-V behavior becomes that of a simple time-dependent progressively-decreasing resistance R, obeying:
  - $\Delta(1/R) = \Delta(t) \times slope \times \exp(A \times \max(abs(VGS), abs(VDS))$
  - Asymptotically R tends to zero. The observed current limitation is due to external series resistances to the gate.x

How is this injected in a Spice net list?
In fresh state and in soft breakdown, gate leakage is a time dependent nonlinear resistance between G, S, D and B, as shown below

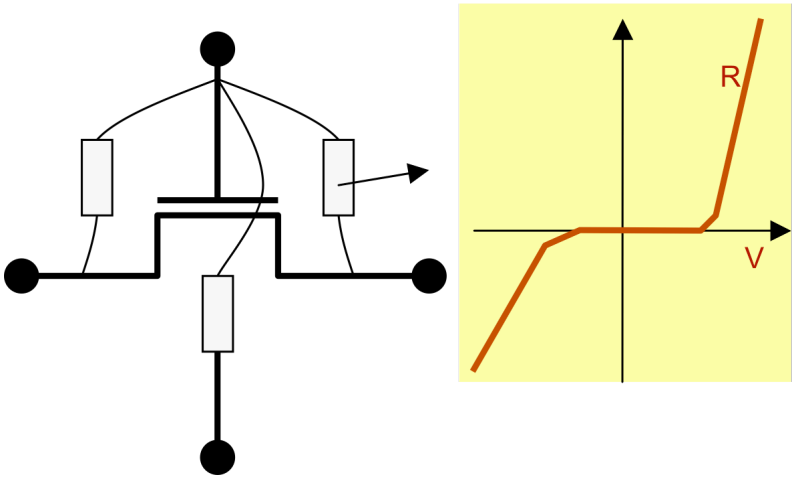


Figure 26: In the fresh state, gate leakage (in the graph to the right) is a nonlinear resistance (injected in schematics to the left).

In *hard breakdown* the resistance is linear, and obeys.

$$\Delta(1/R) = \Delta(t) * slope * \exp(A * \max(abs(V_{GS}), abs(V_{DS}))$$

Herein *slope* is a "parameter" and, there is one technology specific "value" *A*, which has units [V]

We simplify the SBD model further: we disregard the small asymmetry between inversion and accumulation, and we completely disregard the leakage to Bulk. The model will be implemented as a Verilog-A model, which is identical in both quadrants.

The proposed model [9] for Fresh and SBD, for one quadrant is:

$$I = I_0 \cdot \left(\exp\left(\frac{V - I \cdot R}{\eta \frac{kT}{q}} \right) - 1 \right)$$

with 3 parameters that must be fitted to each measurement curve.
R fits to the I-V curve for high I. η ("eta") and I_0 determine the low current behaviour. Physically the "ideality factor" η must be larger than 1.
One can revert to a simple, piece-wise linear approximation, using R and a V_{on} in case the analog simulator can not handle the full model. V_{on} is then calculated from the formula, where $V_{on}=2R.I(V_{on})$

Table of parameters

This table contains a population of measured devices (or created devices). Some parameters are given, duplicated for G-S and G-D.
These are for a given technology, NMOS/PMOS, tox, W, L, T, t, …, example:

Table 1: Example population of breakdown parameter sets.

ptoir	tsbd	two	slope_s	slope_d	rfresh	rsbd_s	rsbd_d	etafresh	etasbd_s	etasbd_d	i0fresh	i0sbd_s	i0sbd_d
1	2022	23456	23.3	22.1	1.23e9	3.4e7	3.4e7	3,7	1.3	1.3	1.23e-19	3.4e-7	3.4e-7
1	678	10987	24.5	23.7	9.88e8	1.1e7	1.1e7	3.5	1.6	1.6	9.88e-18	1.1e-7	1.1e-7
1	13900	19765	21.7	19.0	7.88e8	2.9e7	2.9e7	3.2	1.5	1.5	7.88e-18	2.9e-7	2.9e-7

In this table separate values for source and drain are given. We might consider simplifying further and making these identical.

Rules for scaling W, L, T, t…

Should tell us “what is the effect on elapsed time” when stress condition V, T, etc change. (attention: the measurement conditions after stress do not change)

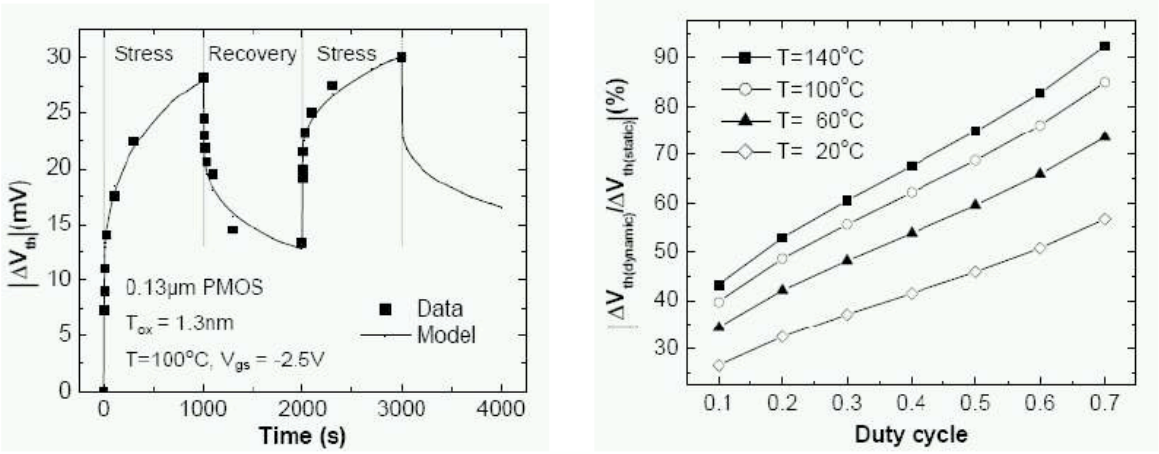
The Pelgrom rule applies to R, only for local random variations

Rules for scaling to multiple sequential stress conditions

Approach: each stress condition on its own creates an “elapsed time”. The different stress conditions just accumulate those times.

Negative Bias Temperature Instability of MOSFETs

The figure below shows the healing property of NBTI (negative-bias temperature instability) and effectiveness of duty-cycle in controlling Vth shift [10]



(a) (b)

Figure 27: (a) Threshold voltage shift over time considering interleaved times of stress and relax. (b) Influence of duty cycle of stress and temperature on threshold voltage shift.

NBTI creates an upper and lower boundary between which V_{th} will move. The actual value at evaluation time is not known (pick randomly?)

$$\Delta Vth_{\min} = f(time, duty\_cycle, toggle\_rate, VGS, VDS, ID, some\_parameters)$$
$$\Delta Vth_{\max} = f'(time, duty\_cycle, toggle\_rate, VGS, VDS, ID, some\_parameters)$$

We derive these relations from Cao's model [11], summarised in Table 2.

Table 2: Summary of the Predictive Model

$\lvert\Delta V_{th}\rvert$ under NBTI			
Static	$A\left((1+\delta)t_{ox}+\sqrt{C(t-t_0)}\right)^{2n}$		
Dynamic	Stress	$\left(K_v(t-t_0)^{0.5}+\sqrt[2n]{\Delta V_{th0}}\right)^{2n}$	
	Recovery	$\Delta V_{th0}\left(1-\frac{2\xi_1 t_e+\sqrt{\xi_2 C(t-t_0)}}{2t_{ox}+\sqrt{Ct}}\right)$	
K_v	$\left(\frac{qt_{ox}}{\epsilon_{ox}}\right)^3 K^2 C_{ox}(V_{gs}-V_{th})\sqrt{C}\exp\left(\frac{2E_{ox}}{E_o}\right)$		
C	$T_o^{-1}\cdot\exp(-E_a/kT)$		
t_e	t_{ox} if $t-t_0\geq t_1$; $t_{ox}\sqrt{\frac{t-t_0}{t_1}}-\frac{\sqrt{\xi_2 C(t-t_0)}}{2\xi_1}$ otherwise		
E_a (eV)	0.49	E_0 (V/nm)	0.335
δ	0.5	K $(s^{-0.25}\cdot C^{-0.5}\cdot nm^{-2})$	8×10^4
ξ_1	0.9	ξ_2	0.5
T_o	10^{-8}		

Figure 28 depicts the abstraction of the model

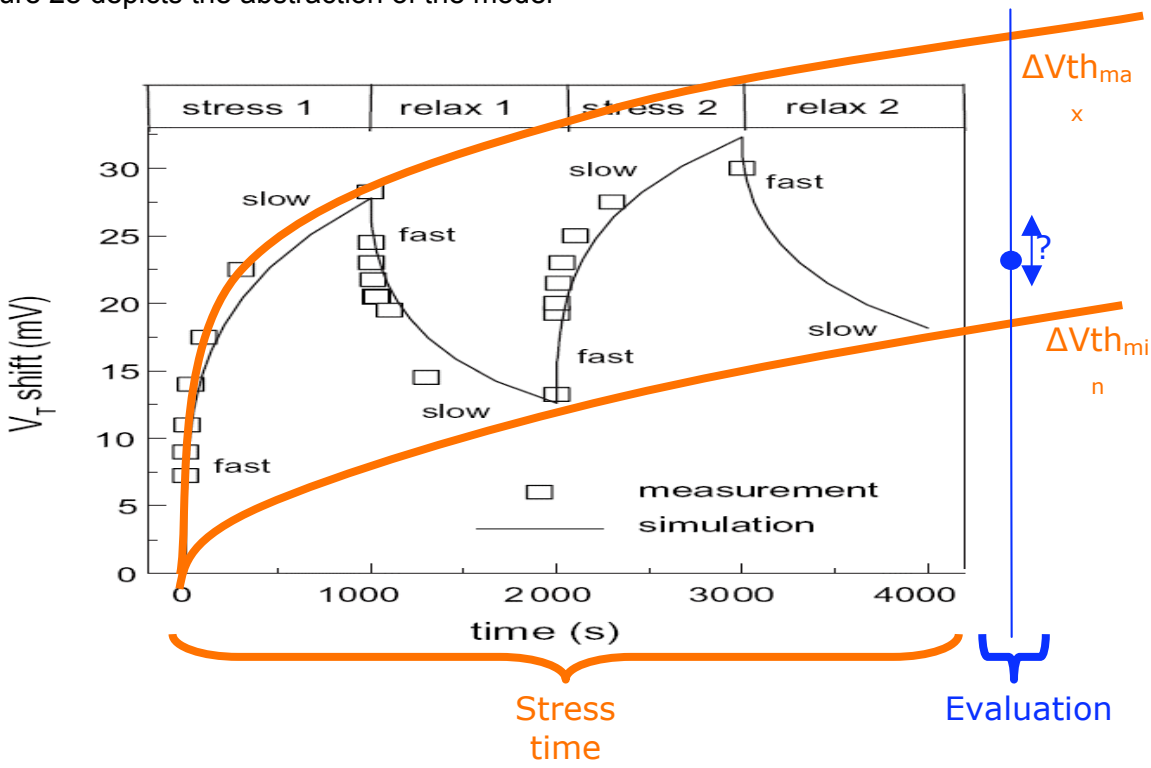


Figure 28: Abstracted NBTI model.

17. Conclusions

In this report we present the results of the statistical simulation of variability in the 45nm technology of STMicroelectronics. I_D-V_G curves for 200 microscopically different n- and p-channel devices have been produced using the University of Glasgow's 3D 'atomistic' simulator. Compact models of each of these devices have been extracted and can be used in statistical circuit simulations to investigate the effects of variability at the circuit and system level. Degradation due to increased density of defect states at the oxide interface during operation will be a major problem and we have simulated the effect for different levels of degradation. We show that increased trap density not only leads to a shift in the threshold voltage, but also to increased variability in device parameters.

18. References

[1] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya and G. Slavcheva, "Simulation of Intrinsic Parameter Fluctuations in Decananometer and Nanometer-Scale MOSFETs", IEEE Trans. on Electron Devices, Vol.50, No.9, pp.1837-1852 (2003)

[2] G. Roy, A. R. Brown, F. Adamu-Lema, S. Roy and A. Asenov, "Simulation Study of Individual and Combined Sources of Intrinsic Parameter Fluctuations in Conventional Nano-MOSFETs", IEEE Trans. on Electron Devices, Vol.53, No.12, pp.3063-3070 (2006)

[3] A. R. Brown, G. Roy and A. Asenov, Poly-Si Gate Related Variability in Decananometre MOSFETs with Conventional Architecture, IEEE Trans. on Electron Devices, Vol.54, No.11, pp.3056-3063 (2007)

[4] J. R. Watling, A. R. Brown, G. Ferrari, J. R. Barker, G. Bersuker, P. Zeitzoff and A. Asenov, "Impact of High-κ Gate Stacks on Transport and Variability in Nano-CMOS Devices", J. Computational and Theoretical Nanoscience, Vol.5, pp.1072-1088 (2008)

[5] A. Asenov, A. Cathignol, B. Cheng, K. P. McKenna, A. R. Brown, A. L. Shluger, D. Chanemougame, K. Rochereau and G. Ghibaudo, "Origin of the Asymmetry in the Statistical Variability of n- and p-channel Poly-Si Gate Bulk MOSFETs", IEEE Electron Device Letters, Vol.29, No.8 (2008)

[6] B. Cheng, S. Roy, and A. Asenov, "Statistical compact model parameter extraction strategy for intrinsic parameter fluctuations" in Proc. Simulation of Semiconductor Processes and Devices, T. Grasser and S. Selberherr, Eds., Vienna, Austria, Sept. 2007, pp. 301–304.

[7] M. F. Bukhori, S. Roy and A. Asenov, "Statistical aspects of reliability in bulk MOSFETs with multiple defect states and random discrete dopants", Microelectronics Reliability, Vol. 48, No 8-9, pp.1549-1552 (2008)

[8] C. Parthasarathy, "Etude de la fiabilité de technologies CMOS avancées: applications a la simulation de la fiabilité de conception de circuit numeriques et analogiques", PHD thesis, 9 Oct 2006, chapter 4.2. With acknowledgements for Guido Groeseneken.

[9] R. Fernández, J. Martín-Martínez, R. Rodríguez, M.Nafría, and X. H. Aymerich, "Gate Oxide Wear-Out and Breakdown Effects on the Performance of Analog and Digital Circuits", IEEE Transactions on ED, Vol. 55-4, p.997 (2008)

[10] Vattikonda R., Wang W., Cao Y., "Modeling and Minimization of PMOS NBTI Effect for Robust Nanometer Design", DAC 2006

[11] Sarvesh Bhardwaj, Wenping Wang, Rakesh Vattikonda, Yu Cao, Sarma Vrudhula, "Predictive Modeling of the NBTI Effect for Reliable Design", CICC 2006, p. 189