

Progress in the Simulation of Time Dependent Statistical Variability in Nano CMOS Transistors

Asen Asenov^{1,2}, Salvatore M. Amoroso¹, Louis Gerrer¹

¹Device Modelling Group, University of Glasgow, Glasgow G12 8LT, Scotland, UK

²Gold Standard Simulations Ltd, Oakfield Avenue- Rankine Building, Glasgow G12 8LT, Scotland, UK
Asen.Asenov@glasgow.ac.uk

Abstract— This paper presents an overview of state-of-the-art simulation methodologies to investigate statistical effects associated with charge trapping dynamics and their impact on the reliability projection in decanometer MOSFETs. By means of novel 3-D Kinetic Monte Carlo TCAD reliability simulation technology we track the time dependent variability associated with granular charge injection and trapping on pre-existing or stress generated oxide traps. For the first time we take into account the interactions between the statistical variability of the ‘virgin’ transistors introduced by the discreteness of charge and granularity of matter and the stochastic nature of the traps distribution and the trapping process itself. Throughout these 3D statistical TCAD techniques we derive the distribution of threshold voltage shift and degradation time constants in conventional bulk, SOI and FinFET transistors.

Keywords—Kinetic Monte Carlo, Reliability, nanoscale CMOS, NBTI, RTN, Device Modeling.

I. INTRODUCTION

The scaling down of the CMOS transistors advocates the adoption of Reliability-Aware circuits and systems design methodology [1]. In fact, charge trapping related issues such as random telegraph noise (RTN) [2-6] and bias temperature instabilities (BTI) [7-9] are major threat to SRAM yield and endurance [1]. Moreover, the interplay between statistical variability and the discrete oxide charge trapping related degradation in transistors necessitates the interpretation of their performance- and reliability- figures of merit as time dependent stochastic variables [3-9]. An important paradigm shift [9] has recently identified the discrete charge trapping in the gate oxide as unique phenomenon underlying both RTN and BTI. This has been confirmed by several new experimental studies [8-9]. The same experiments highlight that reliability is strongly affected by (and interlaced with) static variability. In this paper we present an overview of the cutting-edge TCAD methodology to address the study of the interplay between statistical variability and reliability degradation in conventional and novel nanoscale transistors.

II. SIMULATION METHODOLOGY

Our new approach to statistical reliability simulated is integrated with the Gold Standard Simulations (GSS) ‘atomistic’ simulator GARAND [10]. The stochastic dynamics of the threshold voltage shift amplitudes (ΔV_T) due to charge trapping

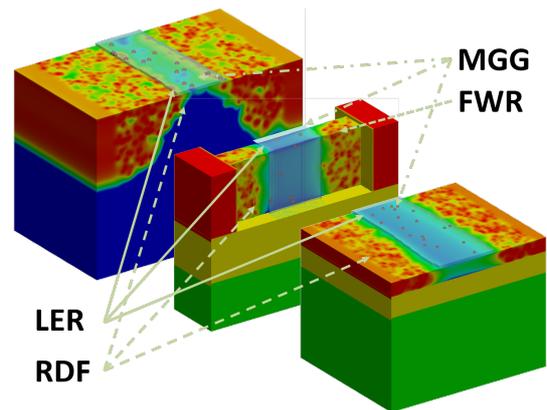


Fig. 1. Investigated bulk FDSOI and Fin FETs, all of them featuring a 60 nm wide effective gate; considered SV sources are Random Dopants Fluctuations (RDF), Metal Gate Granularity (MGG), Line Edge Roughness (LER) and Fin Width Roughness (FWR).

are quantitatively evaluated by means of 3-D Kinetic Monte Carlo (KMC) simulations of large ensembles of microscopically different transistors, accounting for the discrete nature of both doping and oxide traps and reproducing the stochastic process ruling the discrete charge injection into the gate oxide. Fig.1 shows the transistor architectures that will be used as test-bed in the remaining of the paper. Fig.2 shows the KMC simulation procedure developed for the statistical analysis of time dependent device degradation. An outer Monte Carlo loop is used to gather results on a statistical ensemble of thousands of microscopically different transistors: after defining the stochastic configuration of atomistic dopants and oxide traps, the 3-D electrostatics and the drift-diffusion (DD) equations are solved to obtain the ‘time zero’ V_T of each microscopically different transistor. An inner KMC loop is then used to simulate the stochastic charge-injection process from substrate to oxide traps: once cell electrostatics is solved for typical BTI or RTN stress bias conditions, the tunneling current density $J(x,y)$ reaching each trap is calculated over the channel area within the Wentzel-Kramers-Brillouin approximation and the average capture time constant $\langle\tau_c\rangle$ is computed for each trap integrating $J(x,y)$ on an area equal to the trap capture cross section ($\sigma=10^{-14}\text{cm}^2$) as in [11-12]. An activation energy ($E_A=0.6\text{eV}$) is added to the tunneling process to empirically reproduce the multiphonon assisted model proposed in [9]. This activation energy is the main source of variability of $\langle\tau_c\rangle$ [9,12]. In this work we considered an E_A variability range of 1.2eV,

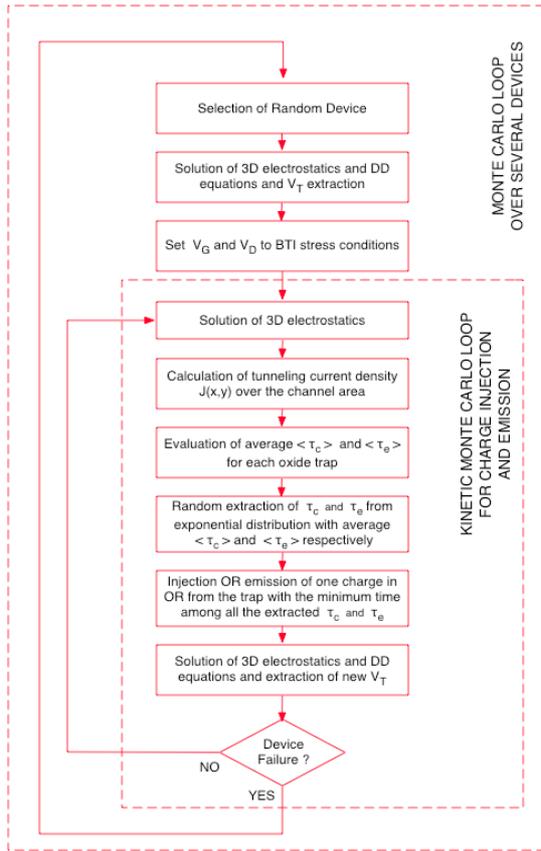


Fig. 2 Block diagram for the simulation procedure used to statistically investigate BTI charging on nanoscaled MOSFETs.

as we have demonstrated in [12] that this value is necessary to explain the experimental results reported in [13]. The average emission time constant $\langle \tau_e \rangle$ can be derived for each trap accordingly to several models –namely Shockley-Read-Hall [14], Kirton [2] or Grassler [9], in order of increasing accuracy. For each trap, the stochastic capture times $\tau_{c,i}$ are then drawn from exponential distributions with average value $\langle \tau_{c,i} \rangle$. Note that the same methodology can be adopted to calculate the capture and emission times from gate to traps [15], which allows the simulation of trap-assisted tunneling [16]. After each trapping/detrapping event the transistor V_T is calculated again and, if the failure limit (in terms of ΔV_T) is not reached, the internal Monte Carlo loop is repeated. Purpose of this simulation procedure is to study in detail the stochastic performance evolution dynamics for a nanoscale transistor under BTI or RTN conditions (Fig.3). Please note that the transistor degradation is not only different from transistor to transistor (due to the difference in dopants and traps concentrations), but also has a marked statistical distribution for a single transistor (due to the stochastic nature of the charge injection/emission process).

III. RESULTS AND DISCUSSION

A. Planar Bulk Transistor

We start our investigation considering a planar bulk 25nm channel length template MOSFET representative for the 20nm CMOS technology generation [17]. The transistor has a gate oxide thickness (T_{ox}) of 1.2 nm and width (W) and length (L) equal to 25 nm. Fig. 4a shows the dynamic simulation results obtained with the KMC engine schematically illustrated in Fig.

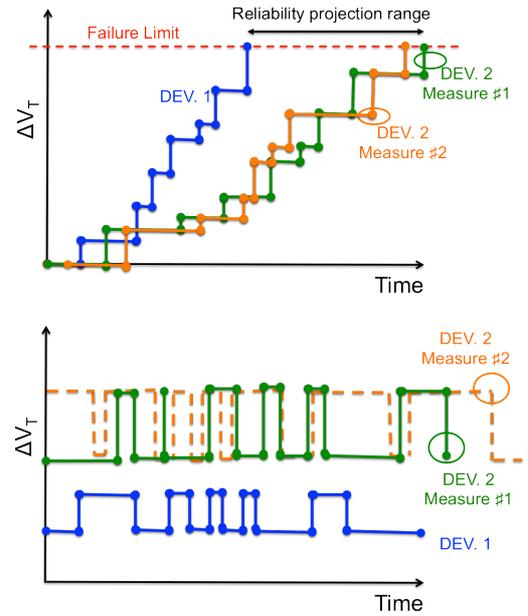


Fig.3 Schematic BTI (top) and RTN (bottom) stochastic traces. The V_T evolution not only differs from transistor to transistor but it is also different when repeating the same experiment on the same transistor.

2, considering only RDD as statistical variability source, without E_A variability. Only pre existing traps with average trap density of $1.6 \times 10^{12} \text{ cm}^{-2}$ are considered in these simulations. Fig. 4b shows analogues results considering the trapping activation energy as additional source of variability. The BTI stochastic charging behavior is evident when KMC simulations are performed on many transistors (the dark traces in Fig. 4a,b), due to the atomistic differences from device to device, but also when simulations are repeated many times on a single transistor (the light traces in Fig. 4a,b), due to the statistical nature of discrete charge injection mechanism [11].

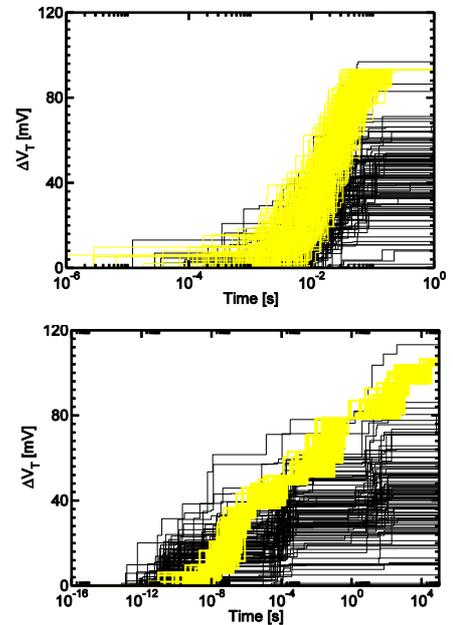


Fig.4 Simulated BTI charging transients for 100 'atomistic' devices in presence of (a, top) RDD variability, (b, bottom) RDD+ E_A variability.

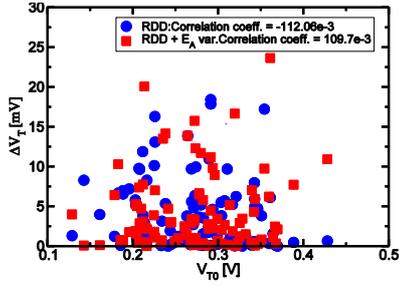


Fig.5 Single charge trapped-induced ΔV_T as a function of the initial V_T .

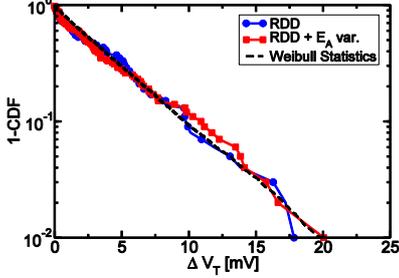


Fig.6 Complementary cumulative distribution of single charge trapped-induced ΔV_T for RDD and E_A variability. A Weibull approximation featuring $\beta=0.95$ is also reported.

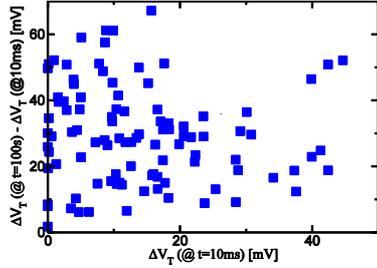


Fig.7 Long-term vs. short-term threshold voltage shift during the BTI stress. The short term component is removed from the y-axis in order to emphasize the absence of correlation.

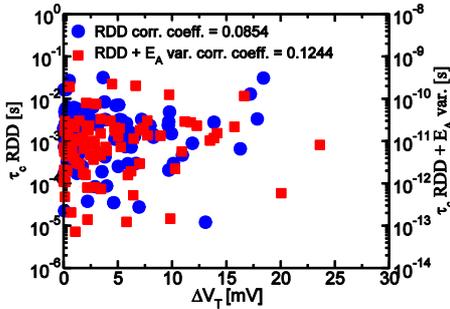


Fig.8 Time constants for the first steps of BTI transients in Fig.3 as a function of the ΔV_T induced after the charge trapping event.3

The device-to-device variability is larger than single device variability when we consider E_A variability. It is important to highlight that the ΔV_T distribution due to a single trapped charge is completely uncorrelated to the initial V_T distribution, as shown in Fig. 5. Moreover, in Fig. 6 we show that the single step ΔV_T distribution is nearly exponential with average 4.05 mV and standard deviation 4.78 mV. Another important feature highlighted by our analysis is the complete lack of correlation between long-term and short-term degradation, as

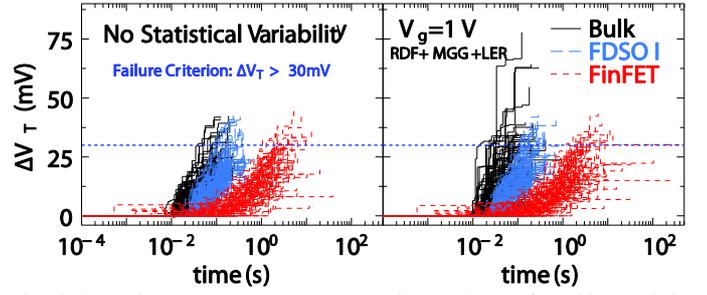


Fig. 9 BTI charge traces at $V_g=1$ V and $V_d=50$ mV for 100 atombic transistors featuring a Poisson distribution averaged on $N_T=10^{12}$ cm⁻², for bulk, FDSOI and Fin FETs, with (up) and without (down) Statistical Variability (SV).

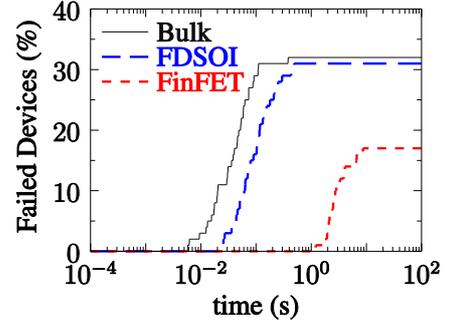


Fig. 10 Percentage of failed device, criterion being $\Delta V_T > 30$ mV

shown in Fig. 7 where we report the cumulative ΔV_T value reached at 100s as a function of the ΔV_T value reached at 10ms for each single device. Finally, Fig.8 shows that the ΔV_T steps and the capture time constants are totally uncorrelated. Therefore, the path that each device follows towards reaching failure can be considered a proper ‘random walk’ and the nanoscale MOSFET reliability can be treated as stochastic feature.

B. Comparison of Novel Transistor Architectures

Combined effects of SV and Reliability lead to the introduction of dopant free channel transistors to continue downscaling below the 22 nm technology node. Fig. 9 present a comparison of 100 devices BTI charges for three devices from the following technologies: a bulk transistor, a Fully Depleted Silicon On Insulator (FDSOI) FET and a FinFET. All of them feature 1.2 nm equivalent gate oxide thickness. The fin height is 25 nm, its width 10 nm and its buried oxide thickness is 20 nm and 10 nm for the FDSOI under a 6 nm layer of silicon channel. Doping levels are detailed in [18]. Equivalent gate width for the Fin is 60 nm and so are the planar devices widths. SV sources affecting these devices are: random discrete dopants (RDF), Metal Gate Granularity (MGG), defined by two grains with an average diameter of 5 nm and respective work functions of 4.021 and 4.221 eV. Line Edge Roughness (LER) and Fin Width Roughness (FER) with a roughness mean square of 3.8 nm and a 25 nm correlation length are introduced to take into account variability coming from lithography. Comparing the BTI traces with and without SV stresses out the huge impact of channel dopant interaction with charged traps; time constants are impacted as well, as detailed in [18-19], leading to the dispersion of projected percentage of failed device with time, as shown in Fig.10. From this picture is clear the superiority of

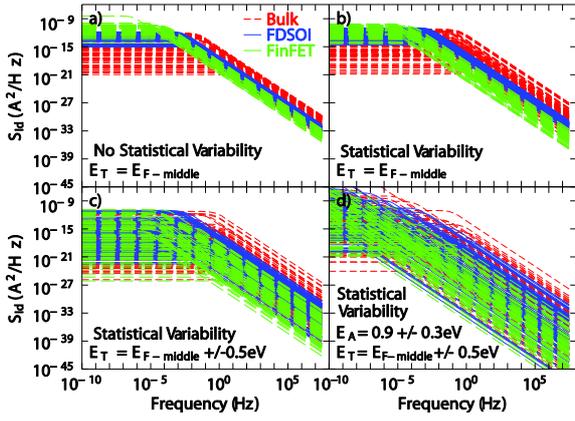


Fig. 11 Noise spectral density for Bulk, FDSOI and FinFET architectures: a) and b) for fixed trap levels with and without statistical variability, c), for a random trap level and d) for a random activation energy.

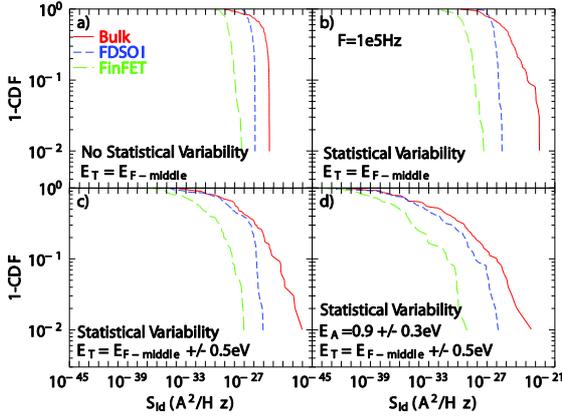


Fig. 12 Noise spectral density cumulative distribution at 100 kHz for Bulk, FDSOI and FinFET architectures: a) and b) for fixed trap levels with and without statistical variability, c), for a random trap level and d) for a random activation energy.

Fin transistors, both in terms of V_T step height and charging times. By means of this simulation methodology we can study the device performance not only under BTI stress but also under RTN conditions. Once the capture/emission time constants are obtained, we can study the device behavior both in the time domain and in the frequency domain. Fig. 11 illustrates the RTN noise spectral density for these three architectures; their distribution at high frequency is given in Fig. 12. Several cases were considered, with and without statistical variability, with a distributed trap energy level E_F and finally with a distributed activation energy E_A . The approach used to estimate noise densities is given in [6, 20]. Actually the SV is not playing a dominant role in the noise spectra dispersion, affecting mainly the V_T step height. On the contrary E_T and E_A affect directly the time constants and therefore the noise density spectra.

IV. CONCLUSIONS

This paper presents an overview of the cutting-edge simulation methodology developed by a collaborative effort of Glasgow Device Modelling Group and Gold Standard Simulation, Ltd to

investigate time dependent statistical variability in nano-CMOS transistors. By means of novel 3-D Kinetic Monte Carlo TCAD reliability simulation technology we have shown how to track, using physical simulations, the time dependent variability associated with granular charge injection and trapping on pre-existing or stress generated oxide traps. We have shown in details the stochastic distribution of threshold voltage shift and time constants involved in the performance degradation of conventional bulk, SOI and FinFET transistors, highlighting how to study the transistor reliability both in time and frequency domain.

ACKNOWLEDGMENT

The authors would like to thank the support given by European Commission through the FP7 grant agreement 261868 MORDRED.

REFERENCES

- [1] K. Takeuchi, T. Nagumo, and T. Hase, VLSI Symp. Tech. Dig., vol. 2011, pp. 130–131, Dec. 2011.
- [2] M. Kirton and M. Uren, *Advances in Physics*, vol. 38, pp. 367–468, 1989.
- [3] A. Asenov, R. Balasubramaniam, A. Brown, and J. Davies, *IEEE Trans. Electron Devices*, vol. 50, pp. 334–336, Mar. 2003.
- [4] J. P. Campbell, J. Qin, K. P. Cheung, L. C. Yu, J. S. Suehle, A. Oates, and K. Sheng, *Proc. IRPS*, p. 382388, Apr. 2009.
- [5] A. Ghetti, C. M. Compagnoni, A. S. Spinelli, and A. Visconti, *IEEE Trans. Electron Devices*, vol. 56, pp. 1746–1752, Aug. 2009.
- [6] S. M. Amoroso, F. Adamu-Lema, S. Markov, L. Gerrer, and A. Asenov, *Proc IWCE*, vol. 2012, pp. 1–4, May 2012.
- [7] M.A.Alam and S.Mohapatra. *Journal of Microelectronics Reliability*, vol. 45, pp. 71–81, 2005.
- [8] B. Kaczer, T. Grasser, P. J. Roussel, J. Franco, R. Degraeve, L. Ragnarsson, E. Simoen, G. Groesenekenand, and H. Reisinger, *Proc. IRPS*, vol. 2010, pp. 26–32, May 2010.
- [9] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. T. Luque, and M. Nelhiebel, *IEEE Trans. Electron Devices*, vol. 58, pp. 3652–3666, Nov. 2011.
- [10] www.goldstandardsimulations.com
- [11] S.M. Amoroso, A. Maconi, A. Mauri, C.M. Compagnoni, E. Greco, E. Camozzi, S. Vigano, P. Tessariol, A. Ghetti, A.S. Spinelli and A.L. Lacaita, *Proc. of International Electron Devices Meeting (IEDM)*, pp. 540-543, 2010.
- [12] S. M. Amoroso, L. Gerrer, A. Asenov, *Proc. of SISPAD* 2013.
- [13] M. Toledano-Luque, B. Kaczer, J. Franco, Ph.J. Roussel, T. Grasser, T.Y. Hoffmann, and G. Groeseneken, *Symp. VLSI Tech.*, 2011, pp.152-153.
- [14] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, *IEEE Electron Device Lett.*, vol. 11, pp. 90–92, Feb. 1990.
- [15] F. A-Lema, C. Compagnoni, S. Amoroso, N. Castellani, L. Gerrer, *et al*, *IEEE Trans. Elec. Dev.* Vol. 60/2, p.833, 2013.
- [16] S. M. Amoroso, L. Gerrer, F. Adamu-Lema, S. Markov, and A. Asenov, *Proc. IRPS 2013*, pp. 3B.4.1 - 3B.4.6.
- [17] X. Wang, F. Adamu-Lema, B. Cheng, A. Asenov, *IEEE Trans. Elec. Dev.*, Vol. 60–5, pp. 1547–1554, 2013.
- [18] L. Gerrer, S. M. Amoroso, S. Markov, F. Adamu-Lema, and A. Asenov, *IEEE Transaction on Electron Devices*, pp. 4008-4013, 2013.
- [19] S. M. Amoroso, L. Gerrer, S. Markov, F. Adamu-Lema, and A. Asenov, *Solid-State Electronics*, pp. 120–126, 2013.
- [20] L. Gerrer, S. M. Amoroso, R. Hussin and A. Asenov, *Proc. of EUROSOI* 2014.