A Unified Computational Scheme for 3D Statistical Simulation of Reliability Degradation of Nanoscale MOSFETs

F. Adamu-Lema, S. Amoroso, S. Markov, L. Gerrer
School of Engineering, University of Glasgow
Glasgow, United Kingdom
fikru.adamu-lema@glasgow.ac.uk

A. Asenov
Gold Standard Simulations, Ltd,
Glasgow, G12 8LT,
United Kingdom

Abstract— **In this paper[1] we present a comprehensive simulation methodology for oxide reliability degradation considering bias temperature instability, trap assisted tunnelling, and random telegraph noise in contemporary CMOS transistors. The impacts of device variability on capture and emission time-constants, and trap assisted tunnelling are discussed in conjunction with the dispersion in threshold voltage shift due to trapped charge.**

Keywords: Relibility; BTI; RTN; TAT; time constants

I. INTRODUCTION

Statistical variability (SV) is one of the major challenges of the semiconductor industry. It dramatically affects transistor the performance, and ultimately the yield of VLSI systems [1–2]. The statistical variability is manifested in a dispersion of the electrical characteristics of the transistors. The dispersion is additionally broadened by the presence of randomly distributed defects in the oxide, existing natively, or created during operational wear-out [3]. It is well known that oxide defects lead to an enhancement of trap-assisted tunnelling (TAT) gate leakage [4, 14]. Charged oxide traps are also responsible for random telegraph noise (RTN) and bias temperature instability (BTI) widely observed in contemporary MOSFETs [5–6]. Random trapped charge fluctuations (RTF) and SV eventually lead to stochastic distribution of the device-lifetimes in ensembles of ultra-scaled devices, e.g. as in SRAMs, and make reliability projections very difficult [6, 7]. Here we present a 3D modelling framework allowing the simulation of RTN, BTI and TAT in unified simulation environment, based on the dynamics of charge trapping and de-trapping. Our discussion focuses on the distribution of the capture and emission time-constants (τ_c, and τ_e respectively), threshold voltage shift ΔV_T, and the TAT-component I_{TAT} of gate leakage, in a 25 nm n-channel bulk MOSFET illustrated in Fig. 1, considering the discrete nature of both doping atoms and oxide traps as the primary sources of the dispersion.

II. COMPUTATIONAL SCHEME

Simulation methodology expands on the 3D drift-diffusion simulator GARAND, featuring density-gradient quantum corrections and physics-based modelling of the major sources of statistical variability [8], to which we add the modelling of oxide trap dynamics, as in [9]. The computational scheme for evaluating the statistical variability in τ_c, τ_e, ΔV_T, and I_{TAT} follows the algorithm shown in Fig. 2. In order to map the features associated with an individual trap and its interplay with RDF, we define an auxiliary lateral grid of n_{max}=*25x25* locations, at position *(z)* above the Si-oxide interface. The following steps of the computational loop shown in Fig. 2 are

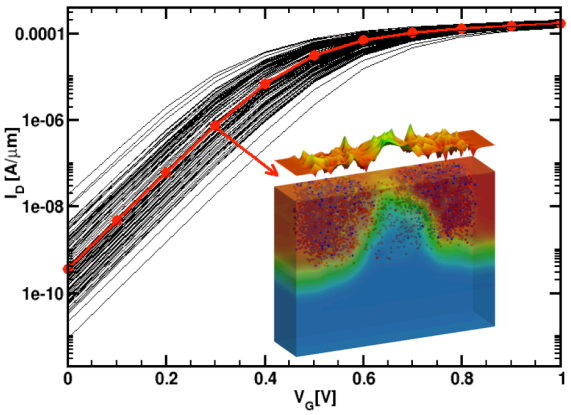

Figura 1. Transfer characteristics of 100 atomistic devices. One device with average characteristics is selected for this work and its potential distribution is shown in 3D, its surface potential as a 2D elevated plot.

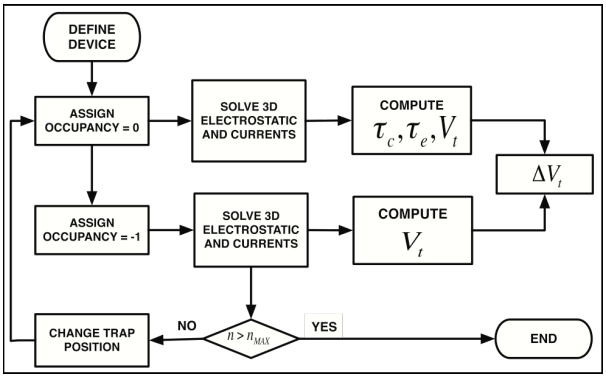

Figura 2. Flow chart of the computational scheme used to map the lateral dependence of capture/emission time constants (τc/τe), threshold shift ΔV_T, and TAT-component of gate leakage I_{TAT}, in the presence of RDF.

[1] This work was supported by the European 7th framework collaborative project entitled "Modelling of the reliability and degradation of next generation nanoelectronic devices" (MORDRED, Grant number 261868). S.M. acknowledges support by the U.K. Engineering and Physical Sciences Research Council's "Molecular-Metal-Oxide-nanoelectronics: Achieving the Molecular Limit", under grant EP/H024107/1.

ISBN 978-0-615-71756-2

executed for each location *(x,y)* of the auxiliary grid, after placing a trap in it. First, 3D device electrostatic and drift-diffusion equations are solved self consistently without a charge in the trap, to obtain the initial threshold voltage V_T. Then, the direct tunnelling current density *J(x,y)* into the trap is computed using the WKB approximation. The corresponding *average* capture time of the trap is calculated by integrating *J(x,y)* over the cross-section σ_s of the trap, and applying a multi-phonon activation energy term to take into account the experimentally observed temperature dependence [9], [10]:

$$< \tau_c >= \exp\left(\frac{E_a}{kT}\right) \frac{q}{\int J(x,y,z)dxdy} \qquad (1)$$

The corresponding *average* emission time constant $< \tau_e >$ is obtained from the SRH statistics, [11]:

$$< \tau_e >= \tau_c \exp\left(\frac{E_T - E_F}{kT}\right) \qquad (2)$$

In (1) and (2) E_a is the activation energy, E_T the trap energy level, and E_F the electron quasi-Fermi level. Note that the trap energy level with respect to the oxide conduction band $E_{T,0}$ = 3.12 eV, the trap cross-section σ_s =10^{-12} cm^{-2}, and the activation energy E_a = 0.5 eV, are kept constant in all the simulations reported here. Finally, a charge is placed in the trap, and the 3D device electrostatic and drift-diffusion equations are solved self consistently again, to obtain the corresponding single-trap-induced threshold voltage shift ΔV_T.

Equations (1) and (2) are solved assuming capture (emission) from (to) the substrate, and from (to) gate. Knowing the time constants for both cases allows us to model the TAT current through an individual trap *i* as:

$$I_{TAT,i} = q/(\tau_{c,C} + \tau_{e,G}) \qquad (3)$$

where $\tau_{c,C}$ ($\tau_{e,G}$) is the *average* time constant for capture (emission) from channel (to gate).

Unless explicitly mentioned, results in the following sections pertain to simulations of the selected device from Fig. 1, performed at low drain bias of 50 mV. The template transistor is a well-scaled 25nm MOSFET with retrograde substrate doping and halo. Average impurity concentration in the channel is approximately 5x10^{18} cm^{-3}, and 10^{20} cm^{-3} in the source and drain regions. The oxide thickness is 1.2 nm, and a dielectric constant of 5.2 is assumed, yielding EOT of 0.9 nm.

III. BTI /RTN

Biased temperature instability and random telegraph noise are two of the most relevant reliability issues in semiconductor industry. BTI is manifested in terms of a threshold voltage increase with the device operational time, due to channel carriers trapping on existent defects and additional oxide defects generation. This results in a reduction of the lifetime span of nanoscale devices [12]. On the other hand RTN occurs as a consequence of the cyclic trapping and de-trapping on single defect state at particular gate bias conditions. Due to the percolative nature of the source-to-drain conduction, the trapped charge will induce different values of RTN ΔV_T amplitude depending on its position over the channel and on the underlying RDF in the channel [13]. This fluctuation will have a negative impact on device and circuit operations, compromising the achievement of a reliability-robust design.

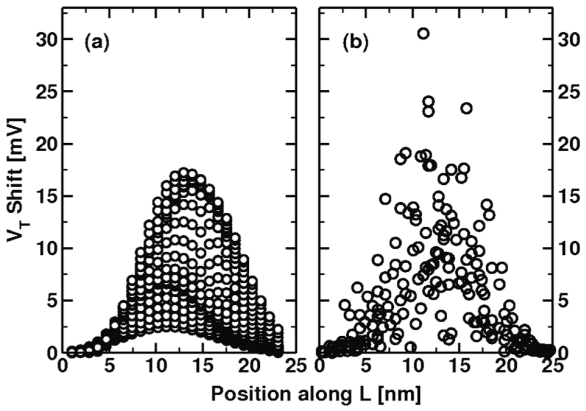

Figura 3. Single-trap-induced threshold shift versus the distance of the trap position *x* from the source-edge of the gate (a) in a signle device with RDF (for all *y* of the auxiliary grid), and (b) in an ensemble of 200 transistors with RDF and one trap at random (x,y) position. All traps 3Å from interface.

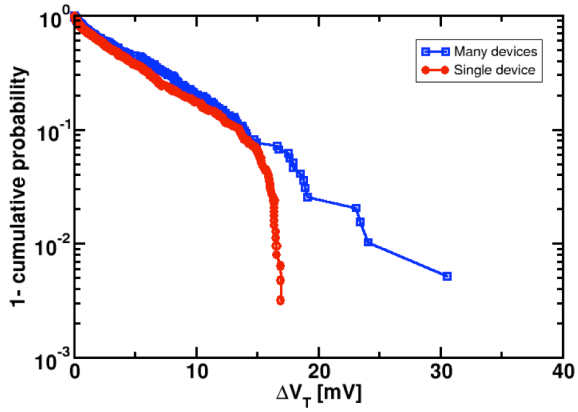

Figura 4. Cumulative probability plot of the statistical data from Fig. 3 (a)–circles, and (b)–squares, demonstrating the need for large device ensemble for the evaluation of the tail of the ΔV_T distibution.

Figure 3(a) shows the dispersion of ΔV_T induced by a single electron trapped at a varying location within the extent of the gate, for the transistor identified in Fig. 1. The envelope shape is due to the 2D non-uniform potential distribution between the source and drain, and reflects the well-known fact that a trap near the peak of the potential barrier has the largest influence on V_T. However, the dispersion at a fixed position, in Fig. 3(a), is entirely due to RDF in the substrate. Figure 3(b) shows once more the single-trap-induced ΔV_T versus longitudinal position, but this time the data is obtained from an ensemble of 200 devices subjected to RDF and one trap with random (*x, y*) coordinates. Similar envelope shape can be perceived, but the dispersion is much broader. Figure 4 compares the cumulative distributions of ΔV_T from Fig. 3. Although Fig. 3(a) gives the global trend of ΔV_T distribution, it does not capture the extreme values that arise from particular combination of trap position and dopants configurations, as clearly shown from the comparison of the cumulative distributions in Fig. 4. This result highlights the importance of performing statistical simulations on large ensembles of devices in order to provide information regarding the low probability tails, as necessary for design of Terascale systems.

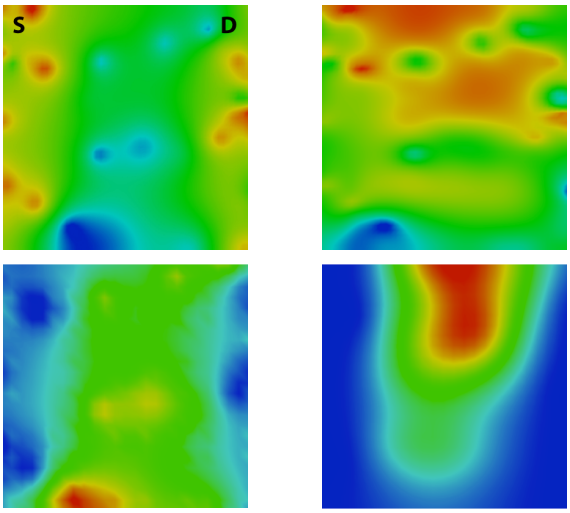

Figura 5. 2D color maps in (*x,y*) of (a) the carrier concentrations at the Si–oxide interface, (b) the current density in the channel, (c) average capture time τ_c, and (d) ΔV_T due to a single trapped charge.

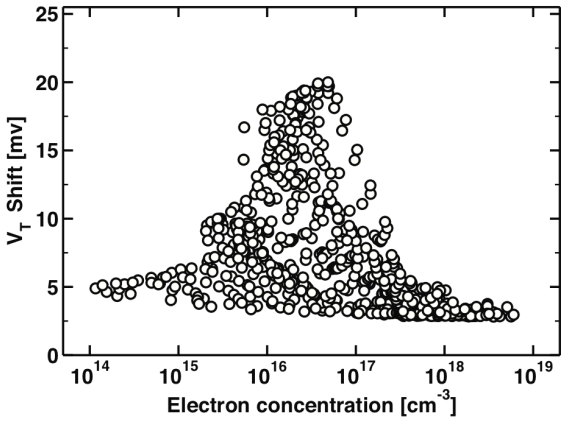

Figura 6. De-correlation of the single-trap-induced threshold voltage shift and the local carrier concentrations at the Si/SiO2 interface below the trap.

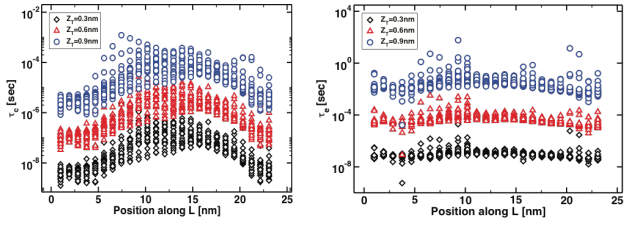

Figura 7. Variation of the average (a) τ_c and (b) τ_e corresponding to individual trap positions along the channel of the transistor, for three vertical positions in the oxide, relative to the Si-oxide interface.

Figures 5(a) and 5(c) compare the 2D lateral distribution of electron concentration at the Si–oxide interface, and average capture time constant τ_c, showing their anti-correlation. Figures 5(b) and 5(d) depict correlation between the threshold voltage shift (ΔV_T) and current density, confirming that trapped charges over a current percolation path between the source and drain have the strongest impact on V_T shift. Note the lack of correlation between τ_c and ΔV_T, further clarified in Fig. 6.

Finally, Fig. 7 demonstrates that the time constants (τ_c,τ_e) also vary significantly over the channel of the device. More importantly, the RDF-induced variation (for example if we focus at a fixed position in *x-position* in Fig7) is of similar order or greater than the variations arising from the non-uniform 2D electrostatics between the source and drain.

IV. Trap Assisted Tunnelling Gate Leakage

Trap assisted tunnelling (TAT) consists in an increase of the gate leakage at low applied biases due to the presence of an intermediate tunnelling state offered by a trap in the band-gap of the gate oxide [14]. Our simulation framework enables us to understand the significance of the electrostatic environment of each individual trap on the corresponding contribution to the TAT leakage current. Again, results here pertain to the device illustrated in Fig. 1.

Figure 8 is a scatter-plot of the TAT-current per trap, versus the local (*in x, y*) inversion carrier concentration, for two different vertical positions of traps. The ensemble of traps reflects the nodes of the auxiliary grid discussed in Section II. The variation in the inversion carrier density reflects the RDF-induced fluctuations in the channel. It is clear from Fig. 8 that I_{GTAT} can be either correlated, or uncorrelated to the surface carrier concentration in the channel, depending on the distance of the trap from the Si–oxide interface. This effect is understood by considering the dominant mechanism that modulates the tunnelling current through a trap – whether it is the capture from channel into the trap, or the emission from the trap into the gate – see equation (3). For a trap that is closer to the Si–oxide interface ($z = 3\text{Å}$), τ_c (capture) is smaller and the variability in I_{TAT} is dominated by the variability of τ_e (emission). The converse is true for a trap closer to the gate ($z = 9\text{Å}$). Accordingly, Fig. 9 shows a strong negative correlation between the average τ_c and the local carriers density in the channel, while the average τ_e is not correlated to that.

Figure 10 compares the dependence of the individual TAT current on the position of the trap along the channel of the transistor, for the two vertical placements of traps discussed above. This dependence is very strong for near-gate traps, the TAT of which is dictated by the time to capture from the channel (z=0.9nm). In this case, traps near the source and drain junctions exhibit larger I_{GTAT} due to the high accumulation density in the S/D overlap regions, but dispersion is not so wide. Traps in the channel however, exhibit very large fluctuations in I_{GTAT}, attributable to RDF in the substrate, in both cases (near-gate or near-interface traps). Further insight is obtained from Fig. 11 reporting both the carrier distribution at the Si–oxide interface, and the tunnelling transparency, obtained from the WKB approximation, as a function of the trap position along the channel area. It clearly appears that the carrier concentration is the main contributor to the fluctuations of I_{GTAT} across the channel, due to the non-uniform 3D potential distribution, while the tunnel-barrier transparency remains relatively uniform, except for several cases of RDF induced large deviations. Overall, for the selected device the variation in I_{GTAT} arising from the 2D non-uniform electrostatics between source and drain is broader than the dispersion due to RDF.

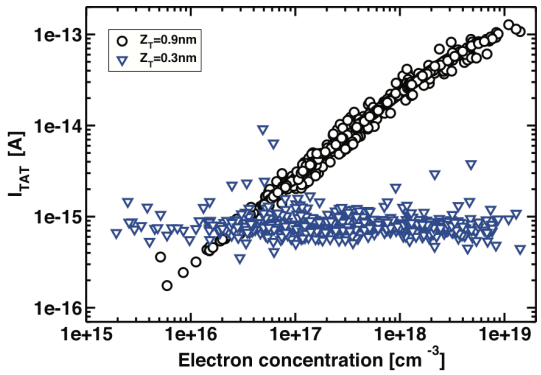

Figura 8. Scatter plot between the TAT current per trap and the local carrier concentration at Si/SiO_2 interface below the trap, showing high correlation only for the for the trap near the gate (*z=0.9nm*).

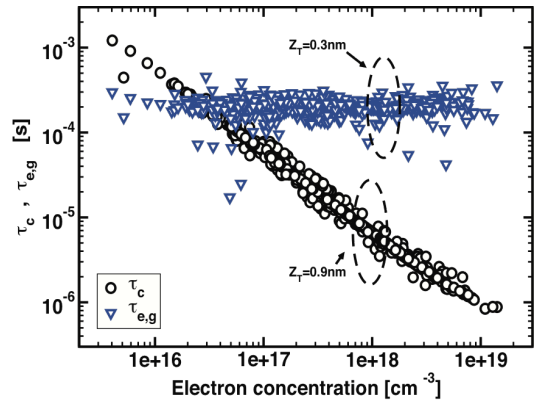

Figura 9. Scatter plot of the dominant time constant and the local carrier concentration in the channel. Only capture time from channel is correlated to the carrier density below the trap, explaining Fig. 8.

V. CONCLUSIONS

A unified computational framework to simulate reliability related to device degradation due to RTN, BTI and TAT has been presented. The dispersion in threshold voltage shifts and capture/emission time constants within the device channel area have been simultaneously related to the stochastic doping configuration and the complex 3D potential distribution. Our results demonstrate the importance of 3D statistical simulations for an accurate projection of reliability performance in modern nanoscale MOSFETs.

VI. REFERENCES

[1] Asenov, A.; Brown, A.R.; Davies, J.H.; Kaya, S.; Slavcheva, G. A. "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs", IEEE Transaction on Electron Devices, Volume 50 (9), pp. 1837, 2003

[2] Gielen, G., *et al*, "Emerging Yield and Reliability Challenges in Nanometer CMOS Technologies", EDAA, pp. 1322, 2008,

[3] Toledano-Luque, M., *et al,* "From mean values to distributions of BTI lifetime of deeply scaled FETs through atomistic understanding of the degradation", pp. 152-153, VLSI 2011

[4] Ielmini, D: Spinelli, A.S; Lacaita, A.L; Modelli, A; "A statistical model for SILC in flash memories", IEEE Trans, Elec. Dev, 49, pp. 1955, 2002

[5] A. Ghetti, C. M. Compagnoni, A. Spinelli, A. Visconti, "Comprehensive analysis of random telegraph noise instability and its scaling in deca–nanometer flash", IEEE Trans, Elec. Dev, 56, pp. 1746, 2009

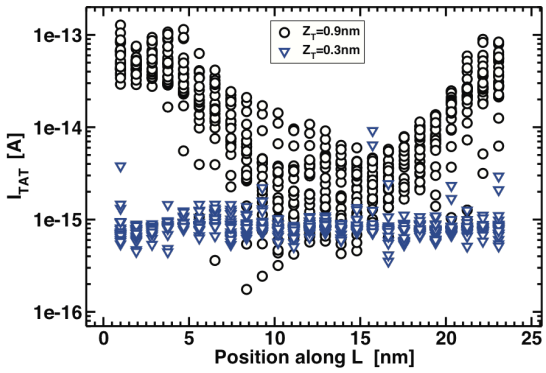

Figura 10. TAT current per trap along the channel of the transistor for all y-positions on the auxiliary grid, Showing larger dispersion for traps nearer to the gate.

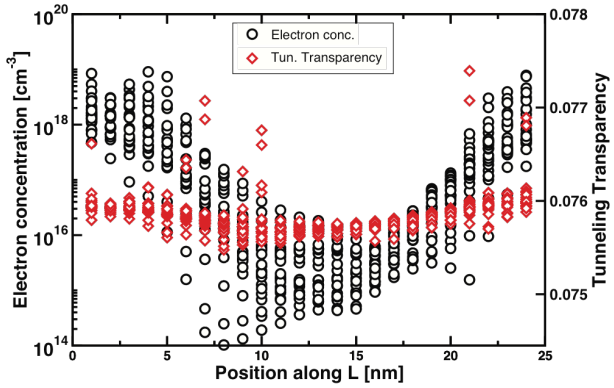

Figura 11. Longitudinal distribution of the electron concentration at the Si/SiO_2 interface and tunneling transparency from the WKB approximation.

[6] Grasser, T; *et al*, "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction–Diffusion to Switching Oxide Traps", IEEE Trans. Elec. Dev., 58, pp. 3652, 2011

[7] T. Grasser, *et al*., "Recent Advances in Understanding the Bias Temperature Instability", *IEDM*, pp. 4.4.1, 2010

[8] Roy G, Brown, A; Adamu-Lema,F; Roy, S; Asenov, A., "Simulation Study of Individual and Combined Sources of Intrinsic Parameter Fluctuations in Conventional Nano-MOSFETs", IEEE Trans. Elec. Dev., 53, pp. 3063, 2006

[9] S. Amoroso, *et al*, "Three-Dimensional Simulation of Charge-Trap Memory Programming—Part I: Average Behavior", IEEE Trans. Elec. Dev., 58, pp. 1864, 2011

[10] Kirton, M.;Uren, M., "Noise in solid-state microelectronics: A new perspective on individual defects, interface and low frequency (1/f) noise", Advances. In Phys., 38, pp. 367, 1989

[11] Shockley, W.; Read, W, "Statistics of Recombination of Holes and Electrons", Phys. Rev. Vol. 87 pp. 835, 1952

[12] Cheng B., Brown, A. and Asenov, A., "Impact of NBTI/PBTI on SRAM Stability Degradation", IEEE Electron Device Lett., Vol. 32, pp. 740, June 2011

[13] Kurata, H., *et al*, "The impact of Random Telegraphic Signals on the Scaling of Multilevel Flash Memories", *VLSI Circuits 2006, Digest of Technical papers*, pp. 112-113, 2006

[14] Jiménez-Molinos, F., Palma, A., Gámiz, Banqueri, J., and Lopez-Villanueva, J. A., "Physical model for trap-assisted inelastic tunneling in MOS structures," *J.Appl.Phys.*, vol. 90, pp. 3396, 2001.