Statistical simulations of oxide leakage current in MOS transistors and Floating Gate memories

L. Larcher and P. Pavan

DISMI, University of Modena and Reggio Emilia Via Fogliani 1, 42100 Reggio Emilia, Italy, larcher.luca@unimore.it

ABSTRACT

The purpose of this paper is to illustrate a physically-based model allowing the statistical simulations of oxide leakage currents in MOS transistors and Floating Gate memories. This model computes the leakage current through defects randomly generated in the oxide, in case accounting for the formation of percolation paths. Furthermore, a calculation procedure has been developed to calculate the threshold voltage of FG memories from the simulated oxide leakage current in some reliability conditions, thus allowing to investigate their actual Flash data retention issues and their future trends. To this regard, it will be shown how this simulation model can be used to investigate threshold voltage shift occurring in retention conditions in FG memories after both Program/Erase cycles, i.e. electrical stress and radiation exposure.

Keywords: Flash memories, semiconductor device reliability, modeling, oxide reliability, SILC.

1 INTRODUCTION

Leakage current through tunnel and gate oxides is a serious issue for MOS transistors and Flash memories. Therefore, the understanding of its physical origins and its experimental characterization and modeling have attracted much interest in the last years.

From the MOS transistor point of view, the very thin gate oxide thickness of state-of-the-art CMOS technologies have led to an exponential increase of tunneling and leakage currents [1]-[3]. The understanding of the physical mechanisms underneath thin oxide degradation and their effects on CMOS reliability is crucial for technology improvement. Furthermore, being capable to model the leakage current (soft- and hard breakdown current in modern devices) and their effects on common circuit functionality [4] is very important to assess more effective reliability criteria based on circuit operation. However, a correct approach to this issue requires statistical modeling tools, necessary when considering the large number of devices in modern ICs.

From a FG memory point of view, the scaling of the tunnel oxide thickness has been limited by reliability concerns due to Stress-Induced Leakage Current (SILC) increase. Flash memories are today one of the basic building blocks of modern electronic systems, both as

stand-alone devices and embedded products: understanding and modeling SILC and its dependence on process recipe as well as cell size parameters is crucial for process development and product manufacturing. The extremely tight requirements for Flash memory reliability (10 years data retention and 10⁵ cycle endurance) and the large capacity of actual products (up to 512 Mb and increasing) ask for new simulation-characterization tools capable not only to reproduce the single cell behavior, but also to comply with all the thousands of cells in a real product. Unfortunately, experimental methods usually adopted to characterize Flash memory oxide reliability either neglects (when SILC is measured on large area capacitors) or does not properly catch (when SILC is derived from threshold voltage measurements) the statistical aspects related to leakage current distributions [5]. Moreover, SILC models proposed in the literature do not feature full statistical capabilities, although they have been used to relate the leakage current distribution to oxide defect distributions

In this scenario, the purpose of this paper is to explain a statistical model that allows to calculate both leakage current distribution across tunnel and gate oxides and the threshold voltage (V_{TH}) distribution in FG memories in different reliability conditions. The physical-statistical SILC model, proposed in [9], will be briefly described in Section 2.1: the new calculation procedure allowing to determine the V_{TH} distribution (in FG memories) from the SILC distribution will be illustrated in Section 2.2. Section 3 shows the simulation results obtained. Firstly, the model capability to reproduce actual leakage currents will be tested on actual measures on different oxide layers. Then, V_{TH} and leakage current distributions in retention conditions after both electrical stress, i.e. P/E cycles, and radiation exposure will be shown.

2 SIMULATION MODEL

This model allows to calculate both leakage current and V_{TH} distributions and its flow-chart is sketched in Fig. 1. It is based on the physical-statistical model of leakage current proposed in [9]. This SILC model will be briefly described in Section 2.1, highlighting its main novel theoretical aspects compared to other oxide leakage models proposed in the literature. The new calculation procedure allowing to determine the V_{TH} distribution from the SILC distribution will be illustrated in Section 2.2.

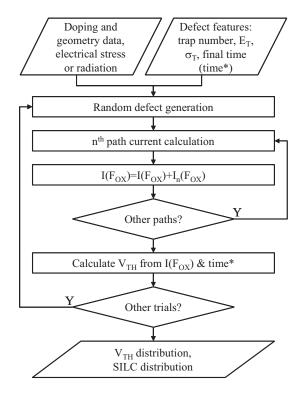


Figure 1: Flow-chart of the simulation model.

2.1 The SILC statistical model

The physical-statistical SILC model presented here differs in some important aspects from usual Trap-Assisted Tunneling (TAT) models presented in the literature [10, 11]: 1) from the calculation point of view, it features statistical simulation capabilities; 2) from the physics, i.e. conduction mechanism point of view, this model assumes that electrons are coupled to oxide phonons: this results in a series of virtual states in the oxide energy band-gap broadening the trap energy level, E_T [12].

If we focus on the conduction mechanism assumed in this model, there are two aspects that are worth to be stressed. First, this SILC model calculates leakage current contribution driven by conductive paths comprised of two or more traps, i.e. to percolation paths, whereas models reported in the literature are limited to two traps [13]-[14]. To calculate the leakage current contribution through every percolation path, the rate (per time) of electrons passing through a n-trap conductive path, R, should be determined. Under steady-state conditions (no charge build-up in any traps of the percolation path), R is given by

$$R_{j} = \frac{I}{\max_{j} \left(\tau_{c,j} + \tau_{e,j}\right)}.$$
 (1)

 $au_{e,j}$ and $au_{e,j}$ are the time constants of the capture and the emission of electrons by and from the j^{th} trap of the

percolation path, respectively. Then, the total leakage current is determined by summing all the leakage current contributions due to every percolation paths

$$I = \sum_{j} q \cdot R_{j} \,. \tag{2}$$

The capture and emission time constants are calculated by the maximum of the single phonon time constant contributions, $\tau_{c,j,n}$ and $\tau_{e,j,n}$, evaluated over the discrete energies $E_{j,n} = E_{C,j} + n \cdot \hbar \omega_0$, where $E_{C,j}$ is the conduction band edge for j=0 or j=trap number+1, or the jth trap energy level $E_{T,j}$ for 0 < j < t rap number+1.

$$\tau_{c,j} = \sum_{n} \tau_{c,j,n} = \sum_{n} N(E_{j-l,n}) \cdot f(E_{j-l,n}) \cdot P_{T}(E_{C,j} - E_{j-l,n}, F_{j-l,j}, D_{j-l,j}) \cdot Ca_{j,n}$$

$$\tau_{e,j} = \sum_{n} \tau_{e,j,n} = \sum_{n} N(E_{j+l,n}) \cdot P_{T}(E_{C,j} - E_{j,n}, F_{j,j+l}, D_{j,j+l}) \cdot Em_{j,n}$$
(4)

 $N(E_i)$ is the density of states at the cathode (j=0), in the trap states (0<j<trap number+1), and at the anode (j=trap number+1); f is the Maxwell-Boltzmann occupation probability; $Ca_{i,n}$ and $Em_{i,n}$ are the trap capture and the emission rates; P_T is the tunnel probability, where $D_{i,i}$ is the distance between the j^{th} and the i^{th} trap, and $F_{j,i}$ is the equivalent oxide field given by $F_{j,i}=F_{OX}(z_j-z_i)/D_{j,i}$ (z is the trap coordinate with respect to the axis perpendicular to the Si/SiO_2 interface, and F_{OX} is the oxide field). Interested readers are suggested to refer to [9] for details in N, $Ca_{j,n}$ and $Em_{i,n}$ calculations. The calculation of the tunnel probability, P_T , performed applying the WKB method is extended to the case of positively charged traps by adopting a simplified approach. It is known from the Poisson equation solution that a positive charged trap deforms and reduces the oxide potential barrier compared to a neutral trap, thus enhancing the tunneling probability. To model in a simple way the oxide barrier deformation induced by a positive trap, it is assumed that the oxide conduction band profile drops linearly with the oxide depth when the distance from the trap is smaller than the capture radius. Thus, the barrier gets smaller and smaller when increasing r_t , and therefore P_T rises steeply.

Second, the leakage current model is integrated with a random number generator supplying spatial and energetic coordinates of defects generated in the oxide during both electrical stress and radiation exposure, thus allowing for statistical simulations of leakage current distribution among samples subjected to the same stress/radiation conditions. A calculation procedure has been developed to evaluate the leakage current driven by every randomly generated trap, checking in case if some multiple trap conductive paths are formed within the oxide. All current contribution due to single or multiple trap paths are summed to calculate the total leakage current, thus reducing the complex 3-D

problem involved in the leakage current calculation to the sum of simpler 1-D problems. To determine whether a trap belongs to a percolation paths, the following procedure is carried out, for every trap: $R_{j\rightarrow anode}$, which is the rate the electrons arrive directly to the anode, is compared to $R_{j\rightarrow k}$, which is the rate the electrons pass from the j^{th} to the k^{th} trap, k comprising all the traps located between the j^{th} trap and the oxide anodic interface. If the maximum $R_{j\rightarrow k}$ is greater than $R_{j\rightarrow anode}$, a percolation path including the j^{th} and the k^{th} trap is identified.

2.2 The V_{TH} calculation procedure

By taking advantage of the statistical capabilities of the above described SILC model, we developed a procedure to calculate the threshold voltage shift occurring in retention conditions (which is the most interesting reliability condition) in Flash cells that underwent either electrical stress or radiation exposure.

Once the spatial coordinates of traps have been generated according to cell geometry/doping, defect features and stress type (i.e. radiation or electrical), the leakage current is calculated for several oxide field values, so that the I-F_{OX} curve can be determined. The relation between the floating-gate voltage, V_{FG} , and the oxide field is calculated using a model taking into account both charge quantization and poly depletion effects [15], hence the I- V_{FG} curve can be calculated.

Thus, the threshold voltage shift occurring in the retention experiment can be calculated by solving $I(V_{FG}) = C_T \cdot dV_{FG} / dt$, C_T being the total FG capacitance [16], which can be written as

$$\frac{t^*}{C_T} = \int_{V_{FG,FN}}^{V_{FG,FN}} \frac{dv_{fg}}{I(v_{fg})}$$
 (5)

 t^* is the retention time; $V_{FG,IN}$ and $V_{FG,FIN}$ are the FG voltages and the beginning and at the end of the retention experiment. Then, the basic FG transistor theory [16] allows to calculate the threshold voltage, V_T , from V_{FG} in retention conditions: $V_T = V_{T0} - V_{FG}/\alpha_G$, V_{T0} and α_G being the UV threshold voltage and the control gate capacitive coupling, respectively. Since defects assisting the current conduction through the oxide are randomly generated starting from defect features and stress type, the statistical V_T distribution can be calculated by repeating the simulations in the same defect conditions.

3 SIMULATION RESULTS

In this section, we will show the simulation capability of the model to reproduce oxide leakage current measured on oxide layers with different thickness. Furthermore, we will illustrate how the statistical capabilities of this model can be used to predict the V_{TH} distribution in retention

conditions in Flash memory cells that underwent electrical and/or radiation stress.

To test the simulation capability of the model, we have compared simulations to leakage currents measured on oxide layers with thickness ranging from 3.3 nm to 10.7 nm. In all cases, the agreement between measurements and simulation is excellent [9]. An example of the simulation quality is reported in Fig. 2, showing SILC vs FOX curves measured on large area MOS capacitors (TOX= 4.4 nm, area=10⁻³ cm²) after a Constant Current Stress, CCS. It is worth noting that simulations, that agree excellently with measurements, have been performed by adopting the statistical approach, that means: N_T·Area·T_{OX} defects are randomly generated within the oxide, N_T being the defect density. Spatial coordinates have been generated assuming a uniform distribution, whereas a Gaussian distribution has been adopted for the energy level (E_T and ΔE_T are the mean and the standard deviation of the energy distribution, respectively).

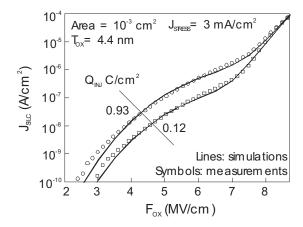


Figure 2: SILC density curves simulated (solid lines) and measured (symbols) on large area MOS capacitors (Area= 10^{-3} cm², $T_{\rm OX}$ =4.4 nm) that have undergone a CCS. $J_{\rm STRESS}$ and $Q_{\rm INJ}$ are the stress current density and the cumulative dose of charge injected across the oxide during the stress. Trap parameters assumed in the simulations are: E_T =2.4eV, ΔE_T =0.15eV, σ_T =1·10⁻¹⁴ cm², N_T =1·10¹⁶ cm⁻³ and N_T =1·10¹⁷ cm⁻³ for the $Q_{\rm INJ}$ =0.12 C/cm² and the $Q_{\rm INJ}$ =0.93 C/cm² curves.

Note that SILC simulations were usually performed taking into account only the average contribution of the defect distribution by means of the $N_T \cdot \sigma_T$ term. Unfortunately, this modeling strategy, which holds for large area MOSFET capacitors, falls when considering small area devices, as the reduction of total amount of defects (i.e. the trap-assisted conductive paths) makes the total leakage current more sensitive against single variations of spatial-energetic trap positions, thus increasing the standard deviation of the SILC distribution. This is not the case of

Fig. 2, where the large device area, i.e. the large amount of oxide traps, results in SILC distributions so narrow to coincide with its mean. On the other hand, a standard deviation increase when considering 1 μm^2 area capacitors can be clearly observed in Fig. 3, where the probability densities of the SILC distributions have been derived for two $F_{\rm OX}$ values from statistical simulations performed on smaller area capacitors. As shown, probability densities (symbols) are fitted quite well by a theoretical log-normal probability density (thick line), although tails appear at high currents.

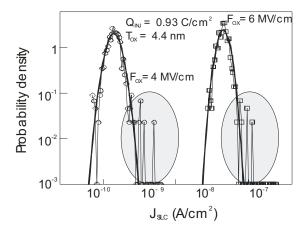


Figure 3: Probability density curves of the $Q_{\rm INJ}$ =0.93 C/cm² SILC distributions plotted in Fig. 2, directly evaluated from the SILC simulation data (symbols) for two $F_{\rm OX}$ values. Thick lines depict the theoretical lognormal probability density extrapolated from the SILC simulation data.

The above SILC model has been used also to simulate leakage currents flowing across the tunnel oxide of Flash memories after P/E cycles. Experimental $I_{\rm G}-F_{\rm OX}$ curves (thin lines) shown in Fig. 4 have been derived from $V_{\rm TH}$ measurements performed on Flash memories with $T_{\rm OX}{=}6.5{-}7$ nm tunnel oxides [9]. SILC simulations (thick lines) have been performed by considering a uniform trap distribution within oxide. As shown, the mean of the leakage current simulation distribution, calculated averaging 10^3 simulation trials, reproduces correctly SILC measurements. Moreover, the large standard deviation derived from the simulated SILC distribution makes the model capable to account also for the statistical variations among different samples.

Once the model capacity to reproduce experimental SILC curves has been assessed, we can show how its statistical capabilities can be used to predict the V_{TH} distribution in Flash memory left unbiased (i.e. in retention conditions) after either electrical or radiation stresses.

To this regard, it is worth noting that the model has proven to be a very effective simulation tool to catch the behavior of Flash memory cells irradiated by heavy ions. It is known from the literature that FG devices hit by heavy ions feature degraded data retention properties [17]. In fact, once hit cells are programmed, their V_{TH} cumulative probability plot, initially very steep and centered at the programmed V_{TH} , features (after some time) a clear tail extending at lower threshold voltage, which increases with time, see Fig. 5. Such V_{TH} reduction affecting hit cells is due to very small leakage currents related to the ion impact, which produces some oxide defects concentrated near the ion track, i.e. the trajectory followed by impinging ions.

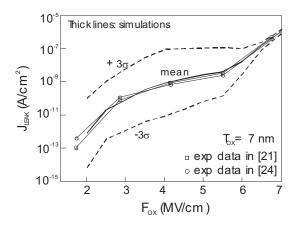


Figure 4: SILC measured (thin lines) on a $T_{\rm OX}$ =6.5/7 nm tunnel oxide (see [9]) are compared to statistical leakage current simulations (thick lines). The solid line and the dashed lines depict the mean and the $\pm 3\sigma$ curves of the simulated leakage current ($J_{\rm LEAK}$) distributions, respectively. Trap parameters assumed in the simulations are: $\sigma_{\rm T}$ =1·10⁻¹⁴ cm² and $N_{\rm T}$ =7·10¹⁷ cm⁻³; the trap energy level has taken constant at $E_{\rm T}$ =2.4eV.

Provided that the random defect generation mechanism of the model is modified to account for radiation stress peculiarities, the evolution of the V_{TH} cumulative probability over time can be reproduced. Figure 5 shows measurements and simulations of threshold voltage cumulative distributions observed in 10nm thick oxide Flash memories irradiated with Iodine. The threshold voltage was measured by using proprietary equipments and DMA algorithms. FG devices were irradiated by using the SIRAD irradiation facility of the 15MV Tandem accelerator at Laboratori Nazionali di Legnaro, INFN, Italy, by using (E=286MeV,Linear Energy LET=64MeV/cm²mg in SiO₂), Silver (E=266MeV, LET=57MeV/cm²mg in SiO₂), and Bromine (E=250MeV, LET=41MeV/cm²mg inSiO₂) ion beams. Devices were kept unbiased during irradiation: the field across the tunnel oxide during irradiation depends on the electron charge stored in the FG, i.e. V_{TH}.

Simulations shown in Fig. 5 were performed by adopting a new defect generation scheme. To reproduce the defect generation mechanism caused by the impinging ion, defects are generated by considering a cylindrical

coordinate system (r, θ , z): θ and z are uniformly distributed, whereas r is a Gaussian variable, with a variance ~3nm. We used 12 oxide defects, probably originated by the electron-hole recombination, each having a cross section $\sigma_T = 10^{-14} \text{cm}^2$ (neutral traps). As shown, simulations are in excellent agreement with measurements, without any fitting parameters.

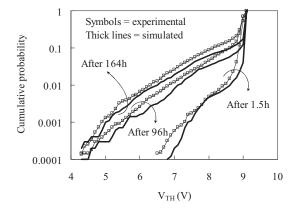


Figure 5: V_{TH} cumulative probability for Flash memories (T_{OX} =10.5nm) irradiated with iodine, after being re-programmed: simulated (thick lines) vs. experimental (symbols) data.

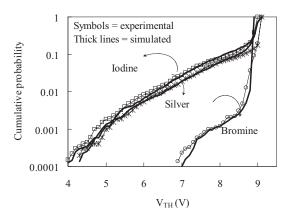


Figure 6: Cumulative probability of V_{TH} for Flash memories irradiated with different ions, 164hour after being re-programmed: simulated (thick lines) vs. experimental (symbols) data.

Excellent fitting can be obtained also for FG devices irradiated with different ions, see Fig. 6. Note that the number of defects one should consider in this case to accurately fit experimental data depends on the ion, and particularly on the ion LET: 11 traps were used for Silver, and 5.5 for Bromine (in this case, the software randomly chooses to use 5 or 6 traps, at each run). This is not

surprising: with the increase of the ion energy transferred to the oxide, i.e. LET, it is reasonable to expect that also the number of defect generated raises.

As reported in [5], the model described in Section 2 constitutes a powerful tool to investigate the reliability of actual Flash memories subjected to conventional P/E cycle stress. Particularly, the model can be used to correlate at statistical level the V_{TH} reduction observed in some reliability conditions (data retention, gate/drain disturbs, ...) to the typical outputs coming from oxide characterization, that are density, cross section, and energy level of defects. Besides, the model can be usefully employed to analyze in depth the physical mechanisms leading to the largest V_{TH} degradation (for example, the correlated/uncorrelated nature of the defect generation process) and also to the effects of TOX scaling on Flash memory retention. To this regard, Fig. 7 shows a typical exaples of the simulator output: the cumulative probability curve derived from the simulation of 10⁶ cells subjected to the same stress, i.e. the same P/E cycles. Simulations in this figure are performed by considering a typical 0.18μm Flash memory cell (W=0.3μm, L=0.3μm, T_{OX}=10.5nm, V_{TH,UV}≈3V) left unbiased for 10 years. Due to the random generation of defects, whose amount depends on their density, N_T, there is a very small percentage of cells (those having V_{TH}<9V, i.e. in the tail of the cumulative probability curve) that features a small but significant leakage current, leading to the partly FG discharge and V_{TH} reduction.

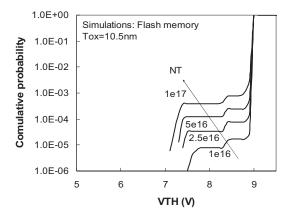


Figure 7: Simulated cumulative probability of V_{TH} for Flash memories (W=0.3µm, L=0.3µm, T_{OX} =10.5nm, $V_{TH,UV}$ ≈3V) with a uniform defect concentration within the oxide (corresponding to a P/E cycle stress), after 10 years. Simulation parameters are: σ_T =1·10⁻¹⁴ cm² and E_T =2.4eV

This relatively large SILC curves are believed to be due to the random alignment of two or more defects, that form a percolation path, driving much more current than simple one-trap conductive paths. Such phenomena threaten severely Flash memory data retention, and therefore are extremely interesting for non-volatile industry people.

It is clear by now that this model can be a powerful tool to evaluate and predict the reliability not only of MOS transistors in future scaled generations, but also of nonvolatile memories based on Floating Gate devices. Both technologies are leading the scaling of microelectronics and any tool that can give results on device reliability that can be trusted is of strategic importance in the industry.

4 CONCLUSIONS

In this paper a physically-based model allowing the statistical simulations of SILCs in MOS transistors and Floating Gate memories has been presented. This model has been integrated with a calculation routine developed to determine the threshold voltage of FG memories in retention conditions. The agreement between simulation and measurements is excellent without any fitting parameter to adjust the fitting quality. Furthermore, some useful examples on how this model can be used to investigate and predict $V_{\rm TH}$ distribution in retention conditions for cells after either electrical stress and radiation exposure have been reported.

REFERENCES

- [1] R. Moazzami and C. Hu, "Stess-induced current in thin silicon dioxide films," in *Proc. IEDM*, San Francisco (USA), pp. 139-142, 2002.
- [2] M. A. Alam, B. E. Weir, P. J. Silverman, "A study of Soft and Hard Breakdown – Part I: Analysis of statistical percolation conductance," *IEEE Trans. Electron Devices*, Vol. 49, N. 2, pp. 232-238, 2002.
- [3] M. A. Alam, B. E. Weir, P. J. Silverman, "A study of Soft and Hard Breakdown Part II: Pinciples of area, thickness, and voltage scaling," *IEEE Trans. Electron Devices*, Vol. 49, N. 2, pp. 239-246, 2002.
- [4] B. Kaczer, R. Degraeve, M. Rasras, K. Van de Mieroop, P. J. Roussell, and G. Groseneken, "Impact of MOSFET gate oxide breakdown on digital circuit operation and reliability," *IEEE Trans. Electron Dev.*, Vol. 49, N. 3, pp. 500-506, 2002.
- [5] L. Larcher and P. Pavan, "Statistical simulations for flash memory reliability analysis and prediction," *IEEE Trans. Electron Devices*, Vol. 51, N. 10, pp. 1636 - 1643, 2004
- [6] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and A. Modelli, "A new two-trap tunneling model for the anomalous SILC in Flash memories," in Proc. *INFOS*, pp. 39-40, 2001.
- [7] R. Degraeve, F. Schuler, M. Lorenzini, D. Wellekens, P. Hendrickx, J. Van Houdt, L. Haspeslagh, G. Groseneken, G. Tempel, "Analitycal model for failure rate prediction due to

- anomalous charge loss of Flash memories," in *Proc. IEDM*, Washington DC (USA), pp. 699-702, 2001...
- [8] M. Hermann and A. Schenk, "Field and high temperature dependence on the long term charge loss in erasable programmable read only memories: Measurements and Modeling," *J. Appl. Phys.*, Vol. 77, N. 9, pp. 4522-4540, 1995.
- [9] L. Larcher, "Statistical simulation of leakage currents in MOS and Flash memory devices with a new multiPhonon Trap-Assisted Tunneling model," *IEEE Trans. Electron Devices*, Vol. 50, no. 5, pag. 1246-1253, 2003.
- [10] Shin-ichi Takagi, Naoki Yasuda, and Akira Toriumi, "A new I-V model for stress-induced leakage current including inelastic tunneling", *IEEE Trans. Electron Devices*, Vol. 46, N. 2, pp. 348-354, 1999.
- [11] B. Riccò, G. Gozzi, and M. Lanzoni, "Modeling and simulation of Stress-Induced Leakage Current in ultrathin SiO₂ films," *IEEE Trans. Electron Devices*, Vol. 45, N. 7, pp. 1554-1560, 1998.
- [12] M. Hermann and A. Schenk, "Field and high temperature dependence on the long term charge loss in erasable programmable read only memories: Measurements and Modeling," *J. Appl. Phys.*, Vol. 77, N. 9, pp. 4522-4540, 1995.
- [13] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and A. Modelli, "A new two-trap tunneling model for the anomalous SILC in Flash memories," in Proc. *INFOS*, pp. 39-40, 2001.
- [14] F. Schuler, R. Degreave, P. Hendrickx, D. Wellekens, "Physical description of anoumaluos charge loss in floating gate based NVM's and identification of its dominant parameter," in *Proc.* 40th IEEE-IRPS, Dallas (USA), pp. 26-33, 2002.
- [15] L. Larcher, P. Pavan, F. Pellizer, G. Ghidini, "A new model of gate capacitance as a simple tool to extract MOS parameters," *IEEE Trans. Electron Devices*, Vol. 48, N. 5, pp. 935-945, 2001.
- [16] P. Pavan, R. Bez, P. Olivo, E. Zanoni, "Flash memory cells – An overview", *Proc. Of the IEEE*, vol. 85, N. 8, pp.1248-1271, 1997.
- [17] L. Larcher, G. Cellere, A. Paccagnella, A. Chimenton, A. Candelori, and A. Modelli, "Data retention after heavy ion exposure of Floating Gate memories: analysis and simulation," *IEEE Transaction on Nuclear Science*, Vol. 50, no. 6, pag. 2176-2183, 2003.