# Statistical Simulation of Leakage Currents in MOS and Flash Memory Devices With a New Multiphonon Trap-Assisted Tunneling Model

Luca Larcher, *Member, IEEE*

***Abstract*—A new physics-based model of leakage current suitable for MOS and Flash memory gate oxide is presented in this paper. This model, which assumes the multiphonon trap-assisted tunneling as conduction mechanism, calculates the total leakage current summing the contributions of the percolation paths formed by one or more aligned traps. Spatial positions and energetic levels of traps have been randomly generated within the oxide by a random number generator which has been integrated into the model. Using this model, statistical simulations of leakage currents measured from both MOS and Flash EEPROM memory tunnel oxides have been carried out. In this way, experimental leakage current distributions can be directly reproduced, thus opening a wide range of useful applications in MOS and Flash EEPROM memory reliability prediction.**

***Index Terms*—Flash memories, leakage currents, modeling, MOS device, oxide reliability, semiconductor device reliability, stress-induced leakage current (SILC).**

## I. Introduction

In the last years, the interest in studying oxide leakage current has grown significantly for two main reasons.

First, the continuous scaling of gate oxide in MOS transistors has leaded to the exponential increase of the leakage current, from stress-induced leakage current (SILC) [1]–[3] on relatively thick oxides (4–6 nm) to soft breakdown in the thinner ones ($<$ 4 nm) [4]–[6]. The understanding of the physical mechanisms of thin oxide degradation and how they affect the overall CMOS reliability is a key factor in technology improvement process. Many efforts have been recently devoted to find a unique picture of the whole gate oxide degradation process (from SILC spot generation, to prebreakdown current jumps, and finally, to breakdown current [7]), to evaluate the location and hardness of oxide breakdown in nMOS and in pMOS transistors [8], and to model the hard- and soft-breakdown current [4], [9]. Further, besides traditional studies on the physical generation mechanisms of defects driving the leakage current conduction [10], in the literature people has started to investigate how MOSFET oxide breakdown affects common circuit functionality (a ring oscillator circuit in [11]). To this regard, new reliability criteria based on circuit functionality after a breakdown event have been introduced [12].

Second, Flash memory scaling (to reduce write voltages and enhance area density) has been strongly limited due to tunnel oxide thickness reduction, leading to reliability concerns due to the SILC increase. SILC was usually measured only on large area capacitors [13], but unfortunately such measurements are not suitable to accurately evaluate retentions and disturbs occurring in Flash memory, since all statistical aspects related to leakage current variations among single cells are completely neglected [14]–[16]. Therefore, many efforts have been devoted to investigate this phenomenon developing new experimental methodologies to measure directly the anomalous charge leakage phenomenon occurring in Flash memories [17]–[19]. Temperature and voltage dependencies of Flash memory leakage current have been widely investigated to gain deeper insights on their physical mechanism [16], [20], [21]. At the same time, the research activity has focused on modeling the anomalous leakage currents flowing through the tunnel oxide. The theoretical models proposed in the literature, that assume the trap-assisted tunneling (TAT) (through one or two traps) as conduction mechanism [22], [23] (in [23], the phonon contribution is also taken into account) have been used to link the statistical distribution of SILC in a Flash array to the spatial distribution of defects generated by electric stress during program/erase (P/E) operations [24], [25].

In this scenario, the aim of this paper is twofold: 1) to present a new physically based model of the leakage current flowing through the gate oxides of MOS transistors and Flash memories; 2) to integrate this leakage current model with a random trap generator to perform the statistical analysis of leakage current distributions. The theory of the physical model of oxide leakage current, which assumes multiphonon trap-assisted tunneling (PTAT) as conduction mechanism, will be explained in details in the next Section II. The model capability to reproduce actual leakage currents will be tested comparing measurements and simulations of Micro Breakdown currents on small area MOSFET [26], SILC on large area MOS capacitors, and anomalous charge loss occurring in several Flash memory cells after P/E cycle stress (Section III). The wide range of applications of this model, that span from the statistical evaluation of the impact of defect features and device size on the leakage current distribution to the deep investigation of the physical oxide damaging mechanism, will be briefly summarized in the conclusions.

Manuscript received November 7, 2002; revised February 27, 2003. The review of this paper was arranged by Editor C. McAndrew.

The author is with the Dipartimento di Scienze e Metodi dell'Ingegneria, CUMEC and INFM, Università di Modena e Reggio Emilia, 42100 Reggio Emilia, Italy.

Digital Object Identifier 10.1109/TED.2003.813236

## II. Leakage Current Model: Phonon and TAT Conduction Mechanism

The new leakage current model is based on the multiphonon assisted tunneling model presented by Hermann *et al.* in [27], [28]. The main difference between this model and the common TAT models presented in [2], [3], [29], and [30], is given by electron coupling to oxide phonons, which results in a series of virtual states in the oxide energy band-gap broadening the trap energy level $E_T$. Coupling to oxide phonons is represented by an effective phonon energy $\hbar\omega_0 = 60$ meV, and two coupling constants: the lattice relaxation energy and the Huang-Rhys factor $S = 6$ [28].

Starting from this physical model describing trap-assisted conduction in the oxide, some important changes have been introduced to: 1) calculate the current driven by conductive paths comprised of two or more traps, i.e., to percolation paths (whereas models reported in the literature addressed mainly TAT through one [2], [3], [29], [30] or at most two traps [22], [23]); 2) extend the leakage current calculation to the case of positively charged traps, simplifying also the mathematical apparatus of the original model. Furthermore, the obtained trap-assisted current model has been integrated with a random number generator supplying spatial and energetic coordinates of defects generated by the high-field stress in the oxide, thus allowing for statistical simulations of leakage current distributions among samples stressed in the same bias conditions.

### A. Percolation Path Transmission Probability Calculation

To extend this model to the case of a percolation path, we have first to calculate the rate (per time) of electrons passing through the $n$-trap conductive path $R$. To determine $R$, let us consider $R_{c,j}$ and $R_{e,j}$, that are the rates electrons are captured by the $j^{\text{th}}$ and leave from $j^{\text{th}}$ trap in the percolation path ($j = 0$ and $j = n + 1$ correspond to the cathode and the anode, respectively) [28].

$$R_{c,j} = \tau_{c,j}^{-1} \cdot (1 - f_{t,j}) \tag{1a}$$

$$R_{e,j} = \tau_{e,j}^{-1} \cdot f_{t,j}. \tag{1.b}$$

$f_{t,j}$ is the occupation probability of the $j^{\text{th}}$ trap, and $\tau_{c,n}$ and $\tau_{e,n}$ are the time constants of the capture and the emission of electrons by and from the $j^{\text{th}}$ trap, respectively.

If no charge build-up occurs in the traps under steady-state conditions, $R_{c,j} = R_{e,j}$. Thus, the occupation probability of the $j^{\text{th}}$ trap is given by $f_{t,j} = \tau_{e,j} \cdot (\tau_{c,j} + \tau_{e,j})^{-1}$, and by replacing it in (1a) or (1b), the rate charge passes through the $j^{\text{th}}$ trap of the percolation path $R_j$ can be calculated.

$$R_j = R_{c,j} = R_{e,j} = (\tau_{c,j} + \tau_{e,j})^{-1}. \tag{2}$$

Moreover, since in a percolation path the capture rate of the $j^{\text{th}}$ trap is the emission rate of the $j-1^{\text{th}}$ trap, $R_{c,j} = R_{e,j-1}$, and by applying (2), $R_j = R_{j-1}$. Since this holds for every trap, rates charge passes through all the traps of the percolation path are all equal, as intuitively expected since no charge pile-up occurs in the traps under steady-state conditions. Then, the rate $R$ charge flows through the percolation path is given by the minimum $R_j$, that is

$$R = \frac{1}{\max\limits_{j} (\tau_{c,j} + \tau_{e,j})}. \tag{3}$$

In other words, the rate the charge passes through faster traps is limited by the rate of the slowest trap (it differs from the formula used in [23]). Accordingly, with (1) and (2) still valid, the occupation probability of faster traps increase determining the global reduction of $R_j$, thus taking into account the decrease of the electron supply not included into the physical description of the single trap capture and emission time [28]. Finally, note that since (3) has been derived independently of the physical mechanism assumed to calculate $\tau_{c,j}$ and $\tau_{e,j}$, this formula holds in any percolation path transmission probability calculations.

### B. Trap Capture and Emission Time Calculation

To relieve the computational weight of the model, we have simplified the calculation of the capture and emission time constants. In [28], $\tau_{c,j}$ and $\tau_{e,j}$ are calculated summing over the discrete energies $E_{j,n} = E_{C,j} + n \cdot \hbar\omega_0$ all the single phonon time constant contributions, $\tau_{c,j,n}$ and $\tau_{e,j,n}$, where $E_{C,j}$ is the conduction band edge for $j = 0$ or $j =$ trap number $+ 1$, or the $j^{\text{th}}$ trap energy level $E_{T,j}$ for $0 < j <$ trap number $+ 1$ [23], [28]. See (4) and (5) at the bottom of the page. $N(E_j)$ 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; $Ca_{j,n}$ and $Em_{j,n}$ are the trap capture and the emission rates; $P_T$ is the tunnel probability, where $D_{j,i}$ is the distance between the $j^{\text{th}}$ and the $i^{\text{th}}$ trap, and $F_{j,i}$ is the equivalent oxide field given by $F_{j,i} = F_{\text{OX}}(z_j - z_i)/D_{j,i}$ ($z$ is the trap coordinate with respect to the axis perpendicular to the Si/$SiO_2$ interface, and $F_{\text{OX}}$ is the oxide field) [28].

Assuming parabolic bands and no states in the energy gap, the density of states is given by (6).

$$N(E) = \frac{\sqrt{kT}}{2\pi^2} \left(\frac{2m_e}{\hbar^2}\right)^{3/2} \sqrt{\frac{E - E_{C,j}}{kT}} \cdot \delta_{-1}(E - E_{C,j}). \tag{6}$$

$$\tau_{c,j} = \sum_n \tau_{c,j,n} = \sum_n N(E_{j-1,n}) \cdot f(E_{j-1,n}) \cdot P_T(E_{C,j} - E_{j-1,n}, F_{j-1,j}, D_{j-1,j}) \cdot Ca_{j,n} \tag{4}$$

$$\tau_{e,j} = \sum_n \tau_{e,j,n} = \sum_n N(E_{j+1,n}) \cdot P_T(E_{C,j} - E_{j,n}, F_{j,j+1}, D_{j,j+1}) \cdot Em_{j,n}. \tag{5}$$

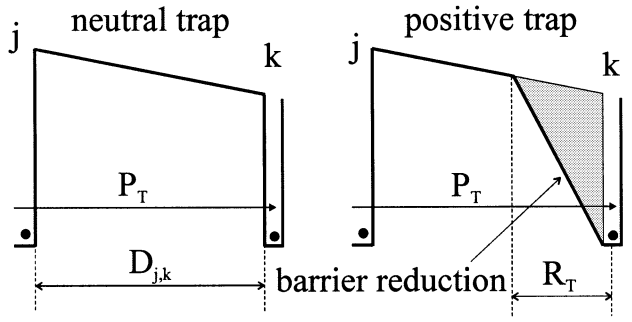



Fig. 1. Schematic picture of the oxide barrier deformation (thick line) due to the presence of a neutral and a positively charged trap: $r_T$ is the trap radius, and $P_T$ is the tunnel probability.

$k$ is the Boltzmann's constant; $T$ is the temperature; $m_e$ is the electron effective mass; $E$ is the electron energy. The capture and emission rate are given by (7) and (8), respectively.

$$Ca_{j,n} = \frac{(4\pi)^2 r_t^3}{\phi_{g,0}} \frac{q^2 F^2 \hbar}{2m_e} \left(\frac{f_B+1}{f_B}\right)^{m/2} \times \exp\left[-2\,(2f_B+1)\right] \cdot I_n\left(2S\sqrt{f_B\,(f_B+1)}\right) \quad (7)$$

$$Em_{j,n} = Ca_{j,n} \exp\left(-\frac{n \cdot \hbar\omega_0}{kT}\right). \quad (8)$$

$r_t$ is the trap capture radius, which is given by $r_T = \sqrt{\sigma_T/\pi}$ ($\sigma_T$ is the trap cross section); $q$ is the electron charge; $\hbar$ is the Planck's constant; $\phi_{g,0}$ is the oxide energy gap; $I_n$ is the modified Bessel function of order $n$; $f_B = 1\exp(\hbar\omega_0/kT) + 1$ is the Bose function giving the phonon occupation number.

The tunnel probability, $P_T$, has been calculated using the WKB method, and its calculation has been extended to the case of positively charged traps adopting the approach reported in [26]. Differently from a neutral trap ($10^{-16}$ cm$^{-2}$ $< \sigma_T <$ $10^{-14}$ cm$^{-2}$ [31]), a positive trap ($10^{-14}$ cm$^{-2}$ $< \sigma_T <$ $10^{-12}$ cm$^{-2}$ [31]) deforms and reduces the oxide potential barrier, strongly enhancing the tunneling probability. To model in a simple way the oxide barrier deformation induced by a positive trap, the capture radius has been assumed as a measure of the oxide barrier reduction. Particularly, the oxide conduction band profile has been assumed to drop linearly with the oxide depth when the distance from the trap is smaller than $r_T$, as schematically depicted in Fig. 1. Obviously, when increasing $r_t$, the barrier gets smaller and smaller, and, therefore, $P_T$ rises steeply.

To simplify the calculation of the capture and emission time constants in (4) and (5), we have analyzed the energy dependence of $\tau_{c,j,n}$ and $\tau_{e,j,n}$, finding that they decrease exponentially on increasing the phonon energy. This is due to the energy dependence of the occupation probability and the capture rate in (4) and of the emission rate in (5), that reduce by around 40 times $\tau_{c,j,n}$ and $\tau_{e,j,n}$ on increasing by one the phonon number, $n$. Definitely, this demonstrates that both (4) and (5) can be approximated by the maximum $\tau_{c,j,n}$ and $\tau_{e,j,n}$ without degrading significantly the simulation accuracy, and strongly reducing the computational effort to carry out simulations.

### C. Statistical Simulations of Trap-Assisted Current

The above described trap-assisted current model has been integrated with a random number generator supplying spatial and energetic coordinates of defects generated during electrical stresses, so that the model can be used to perform statistical simulation of leakage current distributions of samples stressed in the same bias conditions.

To this purpose, a procedure has been developed to calculate automatically the leakage current driven by every randomly generated traps, checking eventually if some multiple trap conductive paths are formed within the oxide. Contributions due to single percolation paths are summed to calculate the total leakage current, thus reducing the complex 3-D problem involved in the trap-assisted leakage current calculation to a sum of simpler 1-D problems.

For each trap, the procedure for calculating the current driven by the $j^{\text{th}}$ trap works as schematically illustrated here:

1) The rate $R_{j\to\text{anode}}$ of every trap not belonging to other percolation paths is calculated assuming that electrons arrive directly to the anode without passing through other traps; 2)$R_{j\to\text{anode}}$ is then compared to rates $R_{j\to k}$ calculated assuming that electrons travel from the $j^{\text{th}}$ trap to every trap located in the oxide portion between the $j^{\text{th}}$ trap and the anode; 3) if the maximum $R_{j\to k}$ is greater than $R_{j\to\text{anode}}$, a percolation path including the $j^{\text{th}}$ and the $k^{\text{th}}$ trap is formed, and the procedure returns to point 1) considering the $k^{\text{th}}$ trap; 4) otherwise, the rate electrons flow through conductive paths including the $j^{\text{th}}$ trap is calculated using (3).

In this way, the current contributions due to every single or multiple trap conductive paths are automatically computed.

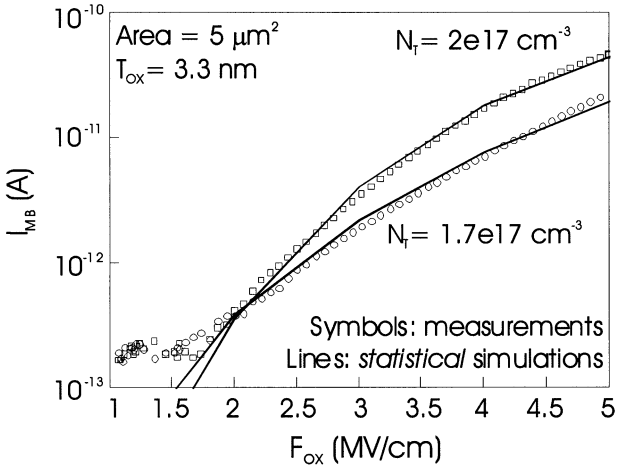
## III. Simulation Results

To test the capability of this model to reproduce experimental leakage currents, we have compared model simulations to: 1) Micro Breakdown current measured on small area MOSFET ($T_{\text{OX}} = 3.3$ nm) (Section III.A); 2) SILC measured on large area MOS capacitors ($T_{\text{OX}} = 4.4$, 5.2 nm) (Section III.B); 3) leakage currents measured across different Flash memory tunnel oxides ($T_{\text{OX}} = 5.7$, 6.5, 7, 10.6 nm) after P/E cycling stress (Section III.C).

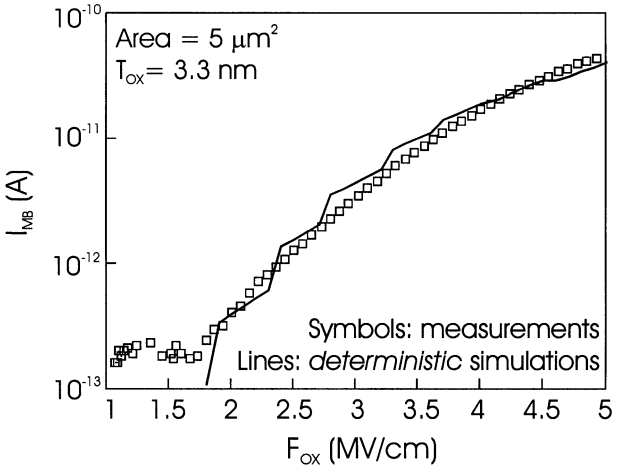
### A. Micro Breakdown Current Simulations

Micro Breakdown (MB) currents, namely leakage currents measured after gate oxide has undergone a MB event, have been measured on small area MOSFETs (Area $= 5\ \mu$m$^2$, $T_{\text{OX}} = 3.3$ nm) stressed by a staircase-like voltage waveform, as described in [26]. Using the PTAT model described in Section II, we have simulated the MB currents adopting two different approaches, that for clarity in the following we will call *statistical* and *deterministic*.

MB simulations shown in Fig. 2(a) are performed following the *statistical* approach: given a trap density $N_T$, $N_T \cdot \text{Area} \cdot T_{\text{OX}}$ statistically uncorrelated defects are randomly generated within the oxide. Spatial coordinates of defects have been generated assuming a uniform statistical distribution, whereas a Gaussian distribution has been adopted for the trap




(a)




(b)

Fig. 2. MB current $I_{\rm MB} - F_{\rm OX}$ curves simulated (solid lines) and measured (symbols) on MOS capacitors (Area $= 5\ \mu$m$^2$, $T_{\rm OX} = 3.3$ nm) after a staircase voltage pulse stress (see [26]). (a) Simulations are performed adopting the *statistical* approach, $10^3$ simulation trials, assuming a distribution of traps ($\sigma_T = 10^{-14}$ cm$^2$) uniform in space and normal in energy ($E_T = 2.8$ eV and $\Delta E_T = 0.1$ eV, and $E_T = 3.0$ eV and $\Delta E_T = 0.3$ eV, for the $N_T = 1.7 \cdot 10^{17}$ cm$^{-3}$ and $N_T = 2 \cdot 10^{17}$ cm$^{-3}$ curves, respectively). (b) Simulations are performed adopting the *deterministic* approach, i.e., assuming that all the current is driven a unique conductive path comprised of two aligned positive traps ($\sigma_T = 6 \cdot 10^{-14}$ cm$^{-2}$), whose energy depth from the oxide conduction band $E_T$ is 2.9 eV and the distances from the cathode are $z_1 = 1.3$ nm and $z_2 = 2.6$ nm.

energy depth ($E_T$ and $\Delta E_T$ are the mean and the standard deviation of the energy distribution, respectively) [32], [33]. Then, the PTAT leakage current driven by all the randomly generated traps ($\sigma_T = 10^{-14}$ cm$^2$) has been automatically calculated applying the procedure described in Section II-C. To check how the random space and energy trap coordinates affect the total leakage current distribution, we have performed $10^3$ simulation trials assuming the same energy and space trap statistical distribution. On increasing the simulation trial number, we have directly verified that the mean and variance of the output leakage current simulations do not vary significantly, demonstrating that the choice of $10^3$ trials is a reasonable tradeoff between simulation accuracy and computation time, al least for a first-order approximated analysis. Since the number of traps generated in the oxide is quite large (2800 and 3300 for the $N_T = 1.7 \cdot 10^{17}$ cm$^{-3}$ and $N_T = 2 \cdot 10^{17}$ cm$^{-3}$ leakage curves depicted in Fig. 2(a)), differences due to the reciprocal energy and space positions of defects compensate each other (this does not happen when considering a small amount of traps), resulting in a very narrow log-normal leakage current distribution coinciding with its mean value. As shown in Fig. 2(a), leakage currents simulated following the *statistical* approach agree excellently with the experimental MB currents, without adopting additional parameters to adjust the fitting.

On the other hand, as aforementioned, we have simulated the MB current shown in Fig. 2(b) also following a different approach, that we have called *deterministic*. Following also [7], [22], [23], [26], we have supposed that the MB current is driven by only two aligned traps forming a percolation path through the oxide. The trap capture cross section has increased to $\sigma_T = 6 \cdot 10^{-14}$ cm$^2$, compared to the *statistical* approach, all the current being driven by a unique two-trap conductive path. The energy levels ($E_{T1} = E_{T2} = 2.9$ eV) and distances from the cathodic interface ($z_1 = 0.4 \cdot T_{\rm OX} - z_2 = 0.8 \cdot T_{\rm OX}$) of the two traps have been chosen to maximize the P-TAT probability, i.e., the percolation leakage current. A zero trap misalignment $X_D$, which is defined as trap-to-trap distance projected to the plane orthogonal to the oxide thickness direction, has been assumed. To this regard, it has been observed that $X_D$ does not affect significantly the simulated leakage current curves provided $X_D < 4$ nm, whereas $X_D > 4$ nm reduces and bends more and more the simulated $I_G - F_{\rm OX}$ curves at high fields. As shown in Fig. 2(b), although leakage current simulations agree well with the experimental measurements, the fitting quality is worse compared to the *statistical* approach. Particularly, *deterministic* $I_G$ simulations feature abrupt step-like increases not shown by experimental MB currents. Such abrupt $I_G$ increases are due to the electron coupling to oxide phonons, which results in a series of virtual trap energy states at discrete levels. Particularly, every abrupt increase of the simulated leakage current corresponds to the transition from a virtual trap energy state to the next one, which is due to the fact that on increasing $F_{\rm OX}$, the energy level of the arrival trap (i.e., the trap whereto electrons travel) shifts downward compared to the departure trap (i.e., the trap wherefrom electrons leave), thus determining the step-like increase of the simulated leakage current. Of course, on increasing the number of conductive paths contributing to the total leakage current, this typical signature characterizing the single conductive path disappears in the total leakage currents, suggesting that MB [and also Soft Breakdown, (SB)] phenomenon could not be due to a single percolation path. On the other hand, if MB current were due to a multi-percolation path conduction, the typical MB oscillations at low field [26] would not be ascribed any more to the opening and closing of the unique conductive path driving all the leakage current, thus leaving open the discussion on SB and MB noise.

### B. SILC Simulations

By adopting the *statistical* approach, we have simulated SILC measured on large area MOS capacitors ($T_{\rm OX} = 4.4$, 5.2 nm, area $= 10^{-3}$ cm$^2$) after a constant current stress (CCS), with stress current density $J_{\rm STR}$ ranging from 1 to 30 mA/cm$^2$, and the cumulative injected charge $Q_{\rm INJ}$ from 0.016 to 16 C/cm$^2$.

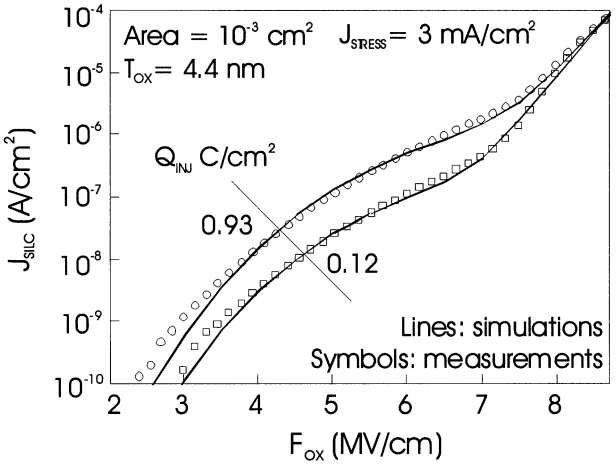



Fig. 3. SILC density curves simulated (solid lines) and measured (symbols) on large area MOS capacitors (Area $= 10^{-3}$ cm$^2$, $T_{\rm OX} = 4.4$ nm) that have undergone a CCS. $J_{\rm STRESS}$ and $Q_{\rm INJ}$ are the stress current density and the cumulative dose of charge injected across the oxide during the stress. Trap parameters assumed in the simulations are: $E_T = 2.4$ eV, $\Delta E_T = 0.15$ eV, $\sigma_T = 1 \cdot 10^{-14}$ cm$^2$, $N_T = 1 \cdot 10^{16}$ cm$^{-3}$, and $N_T = 1 \cdot 10^{17}$ cm$^{-3}$ for the $Q_{\rm INJ} = 0.12$ C/cm$^2$ and the $Q_{\rm INJ} = 0.93$ C/cm$^2$ curves.

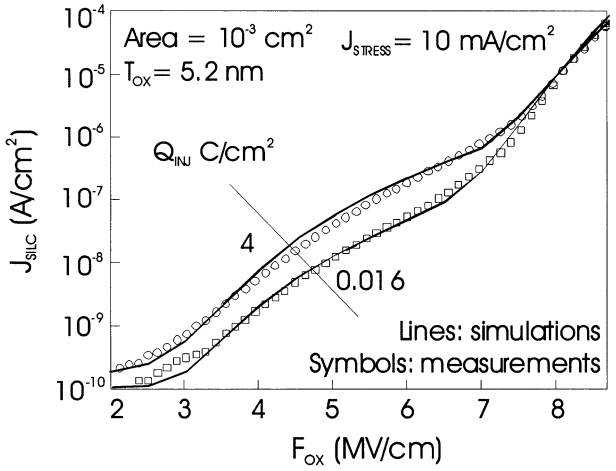



Fig. 4. SILC density curves simulated (solid lines) and measured (symbols) on large area MOS capacitors (Area $= 10^{-3}$ cm$^2$, $T_{\rm OX} = 5.2$ nm) that have undergone a constant current stress (CCS). $J_{\rm STRESS}$ and $Q_{\rm INJ}$ are the stress current density and the cumulative dose of charge injected across the oxide during the stress. Trap parameters assumed in the simulations are $E_T = 2.4$ eV, $\Delta E_T = 0.1$ eV, $\sigma_T = 1 \cdot 10^{-14}$ cm$^2$, $N_T = 5 \cdot 10^{16}$ cm$^{-3}$, and $N_T = 2 \cdot 10^{17}$ cm$^{-3}$ for the $Q_{\rm INJ} = 0.016$ C/cm$^2$ and the $Q_{\rm INJ} = 4$ C/cm$^2$ curves.

As shown in Figs. 3–4, the agreement between measurements and simulations is excellent. It is worth noting that this is the first time that SILC is simulated by adopting a statistical approach. In [2], [3], [29], and [30] SILC simulations were usually performed taking into account only the average contribution of the defect distribution by means of the product $N_T \cdot \sigma_T$. Unfortunately, this *classical* SILC modeling strategy, which holds for large area MOSFET capacitors, falls when considering small area devices [15]. When scaling device area to commercial device sizes, the total amount of defects reduces as well as trap-assisted conductive paths through the oxide. Thus, since the current driven by the single trap becomes a larger fraction of the total current, the variations induced by single trap positions have a greater influence on the total leakage current, thus contributing to increase the standard deviation of the SILC distribution.

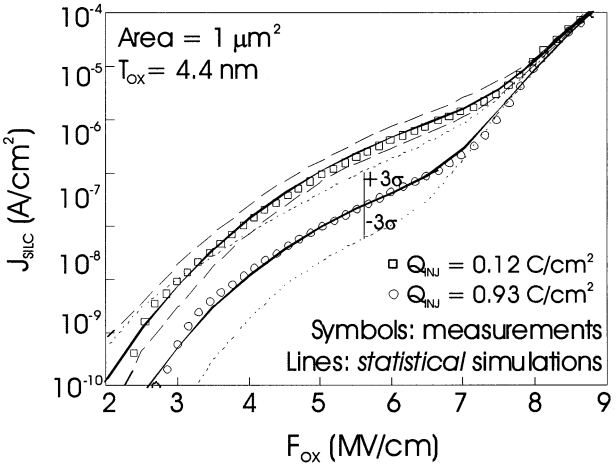



Fig. 5. SILC density curves depicted in Fig. 2, which are measured (symbols) on large area MOS capacitors (Area $= 10^{-3}$ cm$^2$, $T_{\rm OX} = 4.4$ nm), are compared to simulations (lines) performed adopting the *statistical* approach on smaller area MOS capacitors (Area $= 1$ $\mu$m$^2$, $T_{\rm OX} = 4.4$ nm). Solid lines depict the mean of the simulated SILC distributions; dotted and dashed lines depict the $\pm 3\sigma$ curves of the $Q_{\rm INJ} = 0.12$ C/cm$^2$ and $Q_{\rm INJ} = 0.93$ C/cm$^2$ SILC distribution, respectively. Trap parameters assumed in the simulations are $E_T = 2.4$ eV, $\Delta E_T = 0.15$ eV, $\sigma_T = 1 \cdot 10^{-14}$ cm$^2$, $N_T = 1 \cdot 10^{16}$ cm$^{-3}$, and $N_T = 1 \cdot 10^{17}$ cm$^{-3}$ for the $Q_{\rm INJ} = 0.12$ C/cm$^2$ and the $Q_{\rm INJ} = 0.93$ C/cm$^2$ curves.

This is not the case of Figs. 3–4, where the large MOSFET capacitor area, i.e., the large amount of traps in the oxide volume narrows the SILC distribution to coincide with its mean value, as confirmed by the excellent reproducibility of SILC measurements on several samples stressed in the same conditions. On the contrary, SILC measured on smaller area MOSFETs can vary more significantly because of the relative weight of the single trap-assisted conductive path contribution on the total leakage current. This is clearly demonstrated by the SILC simulations shown in Fig. 5, that have been performed adopting the *statistical* approach and considering smaller area MOS capacitors ($T_{\rm OX} = 4.4$ nm, area $= 1$ $\mu$m$^2$). Symbols refer to measurements (the same shown in Fig. 3), whereas solid lines correspond to simulations. Continuous lines depict the mean of the SILC density distribution, computed averaging $10^3$ simulation trials performed assuming the same defect densities ($N_T = 1 \cdot 10^{16}$ and $1 \cdot 10^{17}$ cm$^{-3}$ for $Q_{\rm INJ} = 0.12$ and 0.93 C/cm$^2$ SILC curves, respectively) and trap capture cross section ($\sigma_T = 1 \cdot 10^{-14}$ cm$^2$) considered in the large area (area $= 10^{-3}$ cm$^2$) SILC simulations depicted in Fig. 3. The dotted (dashed) line curves shown in Fig. 5 have been built by adding and subtracting $3\sigma$ ($\sigma$ is the standard deviation calculated at every $F_{\rm OX}$ value) to the mean $J_{\rm SILC}$ of the $Q_{\rm INJ} = 0.12$ C/cm$^2$ ($Q_{\rm INJ} = 0.93$ C/cm$^2$) SILC distribution. Probability densities of the SILC distributions have been derived from statistical simulations. As shown in Fig. 6(a), (b), despite of the relatively small number of trials, such probability densities (symbols) are fitted quite well by a theoretical log-normal probability density (thick lines), although some tails appear at low and high SILC [see the shadowed areas in Fig. 6(a), (b)]. Furthermore, note that the SILC distribution standard deviation is larger for samples that have undergone

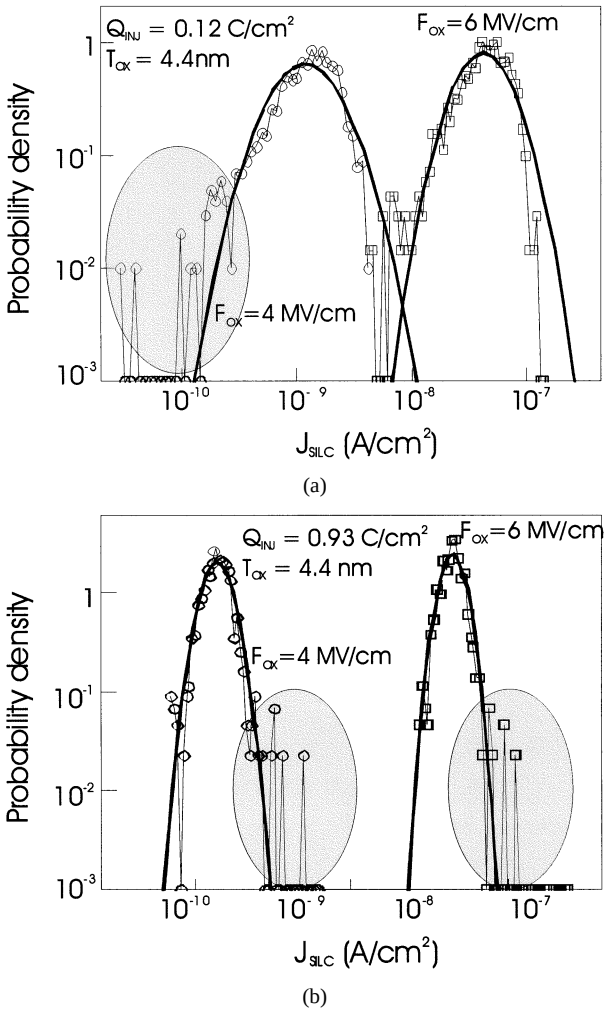



Fig. 6. Probability density of the (a) $Q_{\rm INJ} = 0.12$ C/cm$^2$ and (b) $Q_{\rm INJ} = 0.93$ C/cm$^2$ SILC distributions plotted in Fig. 4, directly evaluated from the SILC simulation data (symbols) for two $F_{\rm OX}$ values. Thick lines depict the theoretical log-normal probability density extrapolated from the SILC simulation data.

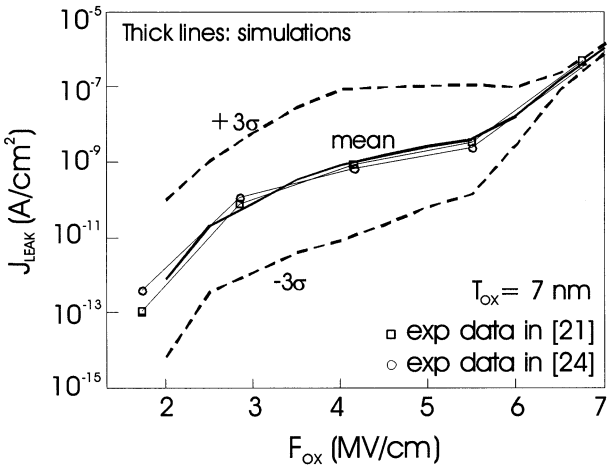



Fig. 7. Experimental leakage current curves derived from threshold voltage data measured on Flash memories having a $T_{\rm OX} = 6.5$ and $T_{\rm OX} = 7$ nm tunnel oxide [21], [24] are compared to simulations (thick lines) performed adopting the *statistical* approach. The solid line and the dashed 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: $\sigma_T = 1 \cdot 10^{-14}$ cm$^2$ and $N_T = 7 \cdot 10^{17}$ cm$^{-3}$; the trap energy level has taken constant at $E_T = 2.4$ eV.

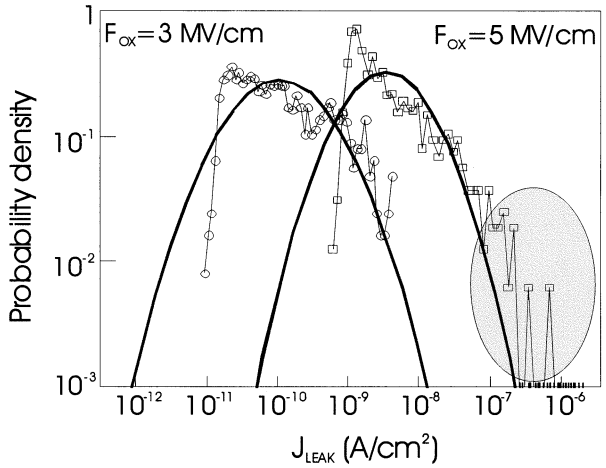



Fig. 8. Probability density of the $J_{\rm LEAK}$ distributions plotted in Fig. 6, directly evaluated from the leakage current simulation data (symbols) for two $F_{\rm OX}$ values. Thick lines depict the theoretical log-normal probability density extrapolated from the leakage current simulation data.

less stress, because the smaller amount of traps generated during the stress enhance the effect of single random trap coordinates on the total leakage current.

### *C. Flash Memory Leakage Current Simulations*

The new leakage current model we present allows to simulate the leakage current flowing across the tunnel oxide of Flash memories that have undergone PE cycle stresses. The experimental $I_G - F_{\rm OX}$ or $I_G - V_G$ curves we have considered to test the simulation capabilities of the model have been previously published in [14], [16], [21], and [24].

The first data considered are the experimental SILC density – $V_G$ curves reported in [21] and [24] (see [24, Fig. 3] and [21, Fig. 13], respectively). Such current curves, that are derived from threshold voltage ($V_T$) measurements performed on Flash memory cell having $T_{\rm OX} = 6.5$ and $T_{\rm OX} = 7$ nm tunnel oxides, are similar, and both can be fitted accurately by the model adopting the *statistical* approach (see Fig. 7). In fact, assuming $N_T = 7 \cdot 10^{17}$ cm$^{-3}$ and $\sigma_T = 10^{-14}$ cm$^2$ and taking the trap energy level fixed at 2.4 eV below the oxide conduction band, experimental currents (symbols) are reproduced excellently by the mean value of the log-normal leakage current distribution calculated averaging $10^3$ simulation trials. Further, the large standard deviation derived from the distribution of the simulated leakage current density $J_{\rm LEAK}$ makes the model capable to simulate also the statistical variations among different samples. The probability density derived from the statistical $J_{\rm LEAK}$ distribution is shown in Fig. 8 for two $F_{\rm OX}$ values, and in both cases it is well fitted by a log-normal probability density (thick lines in Fig. 8). However, as found also in the previous section, $J_{\rm LEAK}$ probability densities calculated from the $10^3$ simulation trials feature some populations tails at high current (see the shadowed region of Fig. 8), that can represent a serious threat for the Flash memory retention. In fact, the larger than expected percentage of devices featuring a high leakage current (the so-called fast or moving bits [16], [21]) can degrade the information data of cells, sometimes leading to the failure of the whole memory array. In

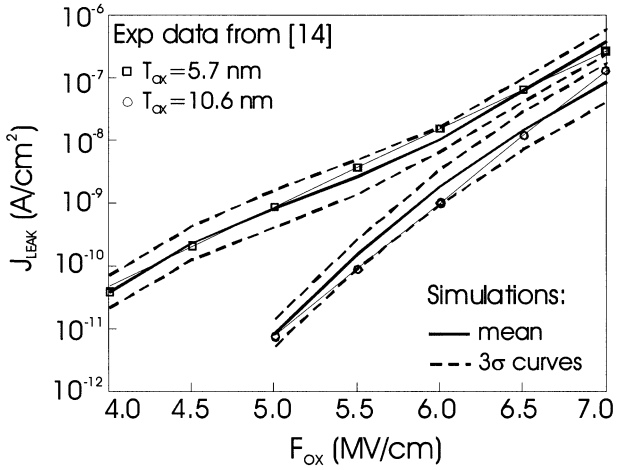



Fig. 9. Experimental leakage current curves measured on a $T_{\rm OX} = 5.7$ and $T_{\rm OX} = 10.6$ nm tunnel oxide Flash memories reported in [14] are compared to simulations (thick lines) performed adopting the *statistical* approach ($10^3$ trials). The solid line and the dashed lines depict the means and the $\pm 3\sigma$ curves of the simulated leakage current ($J_{\rm LEAK}$) distributions, respectively. Trap parameters assumed in the simulations are $\sigma_T = 1 \cdot 10^{-15}$ cm$^2$ and $N_T = 7 \cdot 10^{16}$ cm$^{-3}$ for $T_{\rm OX} = 5.7$ nm; $\sigma_T = 3 \cdot 10^{-15}$ cm$^2$ and $N_T = 3 \cdot 10^{16}$ cm$^{-3}$ for $T_{\rm OX} = 5.7$ nm; the trap energy level has taken constant at $E_T = 2.2$ eV for both $T_{\rm OX}$.

this respect, this model constitutes a powerful tool in the prediction and statistical analysis of leakage current distributions, that are key issues in the scaling and technology improvement process of Flash memory cells and that cannot be addressed satisfactorily only by traditional experimental methods.

The second set of data we have considered to test our model are the $I_G - F_{\rm OX}$ curves depicted in Figs. 2, 4 and 5 of Satoh's paper [14]. Such current curves have been derived from $V_T$ – time curves measured on Flash memories ($T_{\rm OX} = 5.7$ nm and $T_{\rm OX} = 10.6$ nm) after $10^6$ P/E cycles. Also in this case, the mean of the log-normal leakage current distribution simulated adopting the statistical approach reproduces accurately the experimental $J_{\rm LEAK}$ for both the oxide thicknesses considered (see Fig. 9). As aforementioned, values assumed for the density ($N_T = 7 \cdot 10^{16}$ cm$^{-3}$ and $N_T = 3 \cdot 10^{16}$ cm$^{-3}$) and the capture cross section ($\sigma_T = 1 \cdot 10^{-15}$ cm$^2$ and $\sigma_T = 2 \cdot 10^{-15}$ cm$^2$) of defects are physically reasonable, while the energetic trap level has been taken constant at 2.2 eV below the oxide conduction band.

Finally, we have noted also that on decreasing the trap capture cross section, the bending of the simulated $\log(J_{\rm LEAK}) - F_{\rm OX}$ curves reduces progressively, approaching the linear trend found by Belgal *et al.* when neutral traps are considered ($\sigma_T \approx 10^{-15}$ cm$^2$) [16].

## IV. Conclusion

In this paper we have presented a new physical model of the leakage current flowing through the oxides of MOS transistors and Flash memories. This model, which assumes the PTAT as conduction mechanism, has been implemented with a random trap generator, supplying spatial coordinates and energetic levels of defects generated in the oxide layer during a high-field stress. The total leakage current has been calculated summing the contributions of every single traps, checking if some percolation paths have been formed by two or more aligned traps.

This new leakage current model has been tested simulating the Micro Breakdown currents measured on small area MOSFET, SILC measured on large area MOS capacitors, and the anomalous charge loss currents derived from the $V_T$ – time curves of Flash memory cells after P/E cycle stresses. In all the cases, despite of the wide range of oxide thickness considered ($T_{\rm OX} = 3.2$–$10.6$ nm), the model has shown to be capable to fit excellently the experimental leakage current without any additional parameter, thus demonstrating to be a powerful tool for the statistical simulation of oxide leakage currents.

To this regard, this model has a wide range of useful applications in the field of MOS and Flash memory reliability. Particularly, it can be used to evaluate statistically the impact of defect features (density, energy level, and cross section) and device size (area and oxide thickness) on leakage current distributions, thus helping both to improve the technology (by understanding more deeply the physical mechanisms of the various degradation phenomena) and to predict the scaling limits and the failure rate of MOS and Flash memory arrays.

## Acknowledgment

The author thanks Prof. P. Pavan for the support and useful discussions, and Prof. G. Verzellesi for the review of the paper.

## References

[1] R. Moazzami and C. Hu, "Stess-induced current in thin silicon dioxide films," in *IEDM Tech. Dig.*, San Francisco, CA, 2002, pp. 139–142.

[2] S.-I. Takagi, N. Yasuda, and A. Toriumi, "A new I-V model for stress-induced leakage current including inelastic tunneling," *IEEE Trans. Electron Devices*, vol. 46, pp. 348–354, Feb. 1999.

[3] L. Larcher, A. Paccagnella, and G. Ghidini, "A new model of stress induced leakage current in gate oxides," *IEEE Trans. Electron Devices*, vol. 48, pp. 285–288, Feb. 2001.

[4] J. Suñe and E. Miranda, "Post soft breakdown conduction in $SiO_2$ gate oxides," in *IEDM Tech. Dig.*, San Francisco, CA, 2000, pp. 533–536.

[5] M. A. Alam, B. E. Weir, and P. J. Silverman, "A study of soft and hard breakdown–Part I: Analysis of statistical percolation conductance," *IEEE Trans. Electron Devices*, vol. 49, pp. 232–238, Feb., 2002.

[6] ——, "A study of soft and hard breakdown–Part II: Principles of area, thickness, and voltage scaling," *IEEE Trans. Electron Devices*, vol. 49, pp. 239–246, Feb. 2002.

[7] R. Degraeve, B. Kaczer, F. Schuler, M. Lorenzini, D. Wellekens, P. Hendrickx, J. Van Houdt, L. Haspeslagh, G. Tempel, and G. Groseneken, "Statistical model for stress-induced leakage current and pre-breakdown current jumps in ultra-thin layers," in *IEDM Tech. Dig.*, Washington, DC, 2001, pp. 121–124.

[8] F. Crupi, B. Kaczer, R. Degraeve, A. De Keersgieter, and G. Groseneken, "Location and hardness of the oxied breakdown in short channel n- and p- MOSFET's," in *Proc. 40th IEEE-IRPS*, 2002, pp. 55–59.

[9] B. Kaczer, R. Degraeve, A. De Keersgieter, K. Van de Mieroop, V. Simons, and G. Groseneken, "Consistent model for short-channel nMOSFET after hard gate oxide breakdown," *IEEE Trans. Electron Devices*, vol. 49, pp. 507–513, Mar. 2002.

[10] R. Degraeve, G. Groseneken, R. Bellens, J. L. Ogier, m. Depas, P. J. Roussell, and H. E. Maes, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Trans. Electron Devices*, vol. 45, pp. 904–911, 1998.

[11] 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 Devices*, vol. 49, pp. 500–506, Mar., 2002.

[12] R. Degraeve, B. Kaczer, A. De Keersgieter, and G. Groseneken, "Relation between breakdown mode and location in short-channel nMOSFET's and its impact on reliability specifications," *IEEE Trans. Device Mater. Rel.*, vol. 1, pp. 163–169, Sept., 2001.

[13] J. De Blauwe, J. Van Heudt, D. Wellekens, G. Groeseneken, and H. E. Maes, "SILC-related effects in flash $E^2PROM$'s-Part II: Prediction of steady-state SILC-related disturb characteristics ," *IEEE Trans. Electron Devices*, vol. 45, pp. 1751–1760, Aug. 1998.

[14] S. Satoh, G. Hemink, K. Hatakeyama, and S. Aritome, "Stress-induced leakage current of tunnel oxide derived from flash memory read-disturb characteristics," *IEEE Trans. Electron Devices*, vol. 45, pp. 482–486, Feb., 1998.

[15] L. Larcher, S. bertulu, and P. Pavan, "SILC effects on $E^2PROM$ memory cell reliability," *IEEE Trans. Device and Material Reliability*, vol. 2, pp. 13–18, Mar. 2002.

[16] H. P. Belgal, N. Righos, I. Kalastirsky, J. J. Peterson, R. Shiner, and N. Mielke, "A new reliability model for post-cycling charge retention of Flash memories," in *Proc. 40th IEEE-IRPS*, 2002, pp. 7–20.

[17] C. Lam, T. Sunaga, Y. Igarashi, M. Ichinose, K. Kitamura, C. Willets, J. Johnson, S. Mittl, F. White, H. Tang, and T.-C. Chen, "Anomalous low temperature charge leakage mechanism in ULSI flash memories," in *IEDM Tech. Dig.*, San Francisco, CA, 2000, pp. 335–338.

[18] H. Kameyama, Y. Okuyama, S. Kamohara, K. Kubota, H. Kume, K. Okuyama, Y. Manabe, A. Nozoe, H. Uchida, M. Hidaka, and K. Ogura, "A new data retention mechanism after endurance stress on flash memory," in *Proc. 38th IEEE-IRPS*, 2000, pp. 194–199.

[19] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and A. Modelli, "Equivalent cell approach for extraction of the SILC distribution in flash EEPROM cells," *IEEE Electron Devices Lett.*, vol. 23, pp. 40–42, Jan. 2002.

[20] P. J. Kuhn, A. Hoefler, T. Harp, and B. Hornung, "A reliability methodology for low temperature data retention in floating gate nonvolatile memories," *Proc. 39th IEEE-IRPS*, pp. 266–270, 2001.

[21] A. Modelli, F. Gilardoni, D. Ielmini, and A. S. Spinelli, "A new conduction mechanism for the anomalous cells in thin oxide flash EEPROM's," in *Proc. 39th IEEE-IRPS*, 2001, pp. 61–66.

[22] 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*, 2001, pp. 39–40.

[23] F. Schuler, R. Degreave, P. Hendrickx, and 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*, 2002, pp. 26–33.

[24] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and A. Modelli, "Statistical modeling or reliability and scaling projections for flash memories," in *IEDM Tech. Dig.*, Washington, DC, 2001, pp. 703–706.

[25] R. Degraeve, F. Schuler, M. Lorenzini, D. Wellekens, P. Hendrickx, J. Van Houdt, L. Haspeslagh, G. Groseneken, and G. Tempel, "Analitycal model for failure rate prediction due to anomalous charge loss of flash memories," in *IEDM Tech. Dig.*, Washington, DC, 2001, pp. 699–702.

[26] G. Cellere, L. Larcher, M. G. Valentini, and A. Paccagnella, "Micro breakdown in small-area ultra-thin gate oxide," *IEEE Trans. Electron Devices*, vol. 49, pp. 1367–1374, Aug. 2002.

[27] A. Schenk and H. Hermann, "A new model for long term charge loss in EPROM's," in *Ext. Abstract Int. Conf. Solid State Devices Materials (SSDM)*, 1994, pp. 494–496.

[28] 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, no. 9, pp. 4522–4540, 1995.

[29] D. Ielmini, A. S. Spinelli, M. A. Rigamonti, and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling. II. Steady-state," *IEEE Trans. Electron Devices*, vol. 47, pp. 1266–1272, June 2000.

[30] B. Riccò, G. Gozzi, and M. Lanzoni, "Modeling and simulation of stress-induced leakage current in ultrathin $SiO_2$ films," *IEEE Trans. Electron Devices*, vol. 45, pp. 1554–1560, July, 1998.

[31] G. Barbottin and E. Vapaille, *Instabilities in Silicon Devices*. Amsterdam, The Netherlands: North Holland, 1989, vol. 2.

[32] A. I. Chou, K. Lai, K. Kumar, P. Chowdhury, and J. C. Lee, "Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism," *Appl. Phys. Lett.*, vol. 70, no. 25, pp. 3407–3409, 1997.

[33] L. Larcher, A. Paccagnella, and G. Ghidini, "A model of the stress induced leakage current in gate oxides," *IEEE Trans. Electron Devices*, vol. 48, pp. 285–288, Feb. 2001.

**Luca Larcher** (S'99–M'01) received the Laurea degree in electronic engineering from the University of Padova, Padova, Italy, in 1998, working on the modeling of gate oxide currents in MOS devices. He received the Ph.D. degree in 2001 from the University of Modena and Reggio Emilia, Reggio Emilia, Italy, working on the compact modeling of nonvolatile ($E^2PROM$ and Flash) memories.

His research interests concern the reliability and the compact modeling of MOS and floating gate memories. In this field, he focused on the modeling of stress and radiation induced leakage currents, channel-hot electron currents, and MOS gate capacitance. He has authored and coauthored technical papers. He is currently Researcher of Electronics at the University of Modena and Reggio Emilia.