DOI:
https://doi.org/10.1109/TVLSI.2014.2320307

Impact of Bias Temperature Instability on Soft Error Susceptibility

Daniele Rossi, *Member, IEEE*, Martin Omaña, Cecilia Metra, *Fellow, IEEE*,
Alessandro Paccagnella, *Senior Member, IEEE*

***Abstract*— In this paper we address the issue of analyzing the effects of aging mechanisms on ICs' soft error susceptibility. Particularly, we consider Bias Temperature Instability (BTI), namely Negative BTI (NBTI) in PMOS transistors and Positive BTI (PBTI) in NMOS transistors, that are recognized as the most critical aging mechanisms reducing the reliability of ICs. We show that BTI reduces significantly the critical charge of nodes of combinational circuits during their in-field operation, thus increasing the soft error susceptibility of the whole IC. We then propose a time dependent model for soft error susceptibility evaluation, enabling the use of adaptive SE hardening approaches, based on the IC's lifetime.**

***Index Terms*— Aging, bias temperature instability, critical charge, soft error.**

I. INTRODUCTION

THE continuous scaling of microelectronic technology enables to keep on increasing system complexity and performance. However, this growth comes together with a reduction in ICs power supply and, consequently, noise margins, thus increasing significantly their vulnerability to radiation induced errors [1, 2, 3]. Particularly, it is expected that single event transients (SETs) affecting combinational logic will soon become a concern. In fact, the combinational logic vulnerability to SETs will keep on increasing with the reduction of the charge stored on circuit nodes and the decrease in noise margins. Meanwhile, the operating frequency increase will augment the likelihood that signals affected by SETs are sampled, thus giving rise to soft errors (SEs) [4]. This is the case when a SET affecting an internal node of a combinational circuit propagates till the input of a sampling element. If this occurs and the SET satisfies the sampling element set-up and hold-time constraints, it gets latched, thus giving rise to a SE [3, 4, 5]. Intensive research has been devoted to the accurate modeling of SETs [1, 2, 6, 7], as well as to the development of approaches to tolerate them [1, 8].

Along with the susceptibility to SETs, aggressively scaled electronics is becoming increasingly prone to aging mechanisms. Particularly, bias temperature-instability (BTI), caused by both interface-state generation and charge trapping, is recognized as the primary parametric failure mechanism for modern ICs [9, 10]. Negative BTI and Positive BTI are observed in pMOS and nMOS transistors, respectively. They cause performance degradation of MOS transistors when they are biased in strong inversion. For instance, it has been proven that, due to NBTI, the absolute threshold voltage of pMOS transistors can increase by more than 50mV over ten years, thus resulting in more than 20% circuit performance degradation [11]. In data-paths of high performance systems, such a performance degradation may cause the violation of flip-flop set-up and hold times, so that an incorrect value is sampled, thus giving rise to an output SE. Therefore, in the last few years, significant efforts have been devoted to modeling circuit performance degradation due to BTI [12, 13, 14]), and to develop approaches to limit its effects [9, 11, 15, 16].

While both SET and BTI modeling have each received significant attention, less effort has been devoted to the analysis of their interaction. Particularly, to the best of our knowledge, so far the effects of aging on circuit Soft Error Rate (SER) have been analyzed only in [17, 18, 19, 20, 21]. In [17, 18], the impact of aging on the critical charge of SRAM cells has been considered, showing that NBTI has a limited impact on the memory cell critical charge, thus on the memory SER. Instead, in [19], it has been analyzed how the SER of some combinational benchmark circuits varies with circuit life time. However, no details have been provided on the impact of NBTI on the critical charge (and consequently soft error susceptibility) of the different circuit nodes. Finally, in [20], we have presented the results of some preliminary analyses showing that NBTI impacts considerably the critical charge of circuit nodes. This poses new challenges to SE susceptibility modeling, mandating for a time dependent modeling, different from the static modeling broadly considered so far.

Based on these considerations, in this paper we address the analysis of the effects of BTI on IC SE susceptibility. We will show that BTI reduces significantly the value of the critical charge of nodes of combinational circuits during their lifetime. In fact, as shown in [6], the critical charge of a node strongly depends on the value of the restoring current of its pull-

D. Rossi, M. Omaña and C. Metra are with the University of Bologna, Bologna, Italy (e-mail: d.rossi@unibo.it; martin.omana@unibo.it; cecilia.metra.@unibo.it).
A. Paccagnella is with the University of Padova, Italy (e-mail: alessandro.paccagnella@dei.unipd.it).

Manuscript received (insert date of submission if desired). Please note that all acknowledgments should be placed at the end of the paper, before the bibliography.

up/pull-down networks. Since BTI reduces the conductance of the circuit driving the affected node, the value of its critical charge is reduced as well [20]. Furthermore, we prove that NOR gates present a relative reduction of critical charge due to BTI higher than NAND gates. Therefore, NAND gates are preferable over NOR gates not only for their better area/performance trade-off, but also from the SE susceptibility and BTI-induced degradation perspective.

Our analyses have been performed considering elementary gates as NOT, NAND and NOR gates with different fan-ins. As highlighted in [6], the critical charge of a circuit node depends only on technology, as well as on the electrical characteristics of the gate driving the node and fan-out gates. Consequently, our performed analyses can be directly applied to any node, also of complex circuits. In this regard, we have considered the ITC'99 b02 benchmark circuit, composed by different gates, with different fan-ins and fan-outs.

Finally, we propose a dynamic model to estimate the drift of Q_{crit} with circuit aging, where an accurate Q_{crit} evaluation is mandatory to evaluate the probability to generate a SET possibly giving rise to a SE, thus the SER of complex ICs. Our proposed dynamic model achieves an average accuracy higher than 95% compared to HSPICE simulations. Our model will enable the adoption of proper adaptive solutions to counteract the detrimental effect of BTI on circuit SER.

The rest of the paper is organized as follows. In Section 2, we give some preliminaries on BTI and soft error susceptibility. In Section 3, we present the obtained results on the impact of BTI on the Q_{crit} of nodes of combinational circuits. In Sect. 4, we evaluate the impact of BTI in the SE susceptibility of a benchmark circuit. In Section 5, we propose a time-dependent model for SE susceptibility evaluation. In Section 6, we evaluate the accuracy of the proposed model with respect to HSPICE simulations. Finally, in Section. 6, we give some conclusive remarks.

II. BACKGROUND

A. BTI Modeling

Bias Temperature Instability causes significant threshold voltage shift in MOSFET, both using Hafnium-dioxide high-k dielectric material [22], and pure Silicon Dioxide (SiO_2) [14]. Negative BTI and Positive BTI are observed in pMOS and nMOS transistors, respectively. They cause performance degradation of MOS transistors when they are in ON states (stress phase), at elevated temperatures [12]. The BTI-induced degradation is partially recovered when the MOS transistors are polarized in their OFF state (recovery phase).

BTI degradation originates from the creation of charges at the Si-dielectric interface. During the stress phase, the Si-H bonds at the Si-dielectric interface breaks. The broken bonds act as interface traps, while the released Hydrogen, in the form of both atoms (H) and molecules (H_2), diffuse toward the gate [12, 14]. As described in the reaction-diffusion model [12], the interface traps concentration N_{IT} depends on the initial Si-H bond density (N_0), on the Si-H (forward) bond dissociation rate constant (k_f), on the Si bond annealing rate constant (k_r), as well as on the H and H_2 diffusion coefficient (D_H and D_{H2}, respectively). We can assume that the diffusion mechanism of the H_2 molecules prevails over that of the H atoms [12].

During the stress phase, the interface trap concentration N_{IT} varies with time (t) as follows:

$$N_{IT}(t) = \left(\frac{k_f N_0}{k_r}\right)^{2/3} \left(\frac{k_H}{k_{H_2}}\right)^{1/3} \left(6D_{H_2}t\right)^{1/6}, \qquad (1)$$

where parameters k_H and k_{H2} represent the H to H_2 conversion rate, and the H_2 to H conversion rate inside the dielectric [14], respectively.

During the recovery phase, the Hydrogen diffuses back and recombines with the Si dangling bonds, annealing them [12]. Assuming that the stress phase ends at time t_0, and denoting by $N_{IT}(t_0)$ the concentration of the interface traps generated during the stress phase, the trend over time of N_{IT} during the following recovery phase is given by [12]:

$$N_{IT}(t) = N_{IT}(t_0)\left[1 - \left(\frac{\xi(t-t_0)}{t_0}\right)^{1/2} \Big/ \left(1 + \frac{t-t_0}{t_0}\right)^{1/2}\right], \quad (t > t_0) \qquad (2)$$

where ξ=1/2 for one side diffusion.

The traps generated at the Si-dielectric interface shield the applied gate voltage, thus resulting in a threshold voltage increase, denoted by ΔV_{th}. The dependency of ΔV_{th} on N_{IT} is given by [14]:

$$\Delta V_{th}(t) = \frac{q(1+m)}{C_{ox}} \chi N_{IT}(t), \qquad (3)$$

where the coefficient χ allows to distinguish between PBTI and NBTI effects on nMOS and pMOS transistors, respectively. Particularly, χ equals 0.5, for nMOS transistors, and 1 for pMOS transistors [14, 23], showing that PBTI is a less severe problem than NBTI [23].

The threshold voltage shift can considerably degrade transistor performance, since it may reach 50mV over ten years [11], with a consequent reduction of the overdrive voltage, thus of the provided current. Moreover, as shown in Section 4, it can noticeably impact the SE susceptibility of logic circuits, since it reduces the ability of a gate to maintain the correct voltage value on a node hit by an energetic particle.

B. SET Modeling

As known, when an energetic particle strikes the silicon, it travels and loses energy along a straight path. Considering an alpha-particle hitting an IC die, while it loses kinetic energy, it allows more time for its positive charge to induce electron-hole pairs through Coulombic interaction [3]. Consequently, the charge generation rate increases with the distance traveled by the alpha-particle and reaches its maximum near the end of the alpha-particle path. If an electric field is present in the region hit by the particle, such as the depletion region of a reversed biased p-n junction (usually referred to as *critical*

area), the electron-hole pairs are separated. The electrons drift to a more positive potential area, whereas the holes drift to a more negative one, causing a SET.

The current induced by alpha-particles hitting CMOS circuits has been modeled in [24] by a double-exponential current pulse:

$$I(t) = \frac{Q}{\tau_\alpha - \tau_\beta}\left(e^{-t/\tau_\alpha} - e^{-t/\tau_\beta}\right). \quad (4)$$

Q is the total amount of charge collected by the affected node (dependent on the particle energy and trajectory), τ_α is the collection time-constant of the junction, and τ_β accounts for the ion-track establishment time constant. These time constants depend on several process-related factors. For simulation purposes, we will hereinafter consider the values given in [25]: $\tau_\alpha = 1.64\times10^{-10}$ sec and $\tau_\beta = 5\times10^{-11}$ sec.

If the collected charge reaches a critical value (denoted by Q_{crit}) high enough to result in a SET with an amplitude exceeding the fan-out gate logic threshold, an incorrect logic value can propagates. In combinational logic, the generated SET may propagate through the downstream logic and get captured by a sampling element, thus resulting in an SE. Instead, in sequential elements, the generated SET may directly result in an SEU [3].

The Q_{crit} is employed to quantify the SER of an electronic circuit. In fact, it allows to estimate the probability that a particle hitting a node has an energy sufficient to generate a SET. The probability that an alpha-particle striking a node i gives rise to a SET exceeding the fan-out gate logic threshold is given by [8, 29]:

$$P_{gen}(i) = k_i \frac{\alpha}{\beta} e^{-\beta Q_{crit}(i)}, \quad (5)$$

where the coefficient k_i accounts for the area of node i and the alpha-particle flux, while α and β are fitting parameters. From (5) it derives that the Q_{crit} plays a dominant role in the SET generation mechanism. The estimation of the Q_{crit} is therefore of utmost importance to evaluate the SER of a circuit.

In [6], Rossi et al. have proposed an accurate linear model expressing the Q_{crit} of a circuit node as a linear function of the transistor sizes of both the driving and fan-out gates. The proposed model enables to derive the Q_{crit} value from the nominal size of the above mentioned transistors, with no need to perform time consuming electrical level simulations. According to this model, Q_{crit} is given by:

$$Q_{crit}(W_{DR}, W_{FO}) = Q_{min} + a(W_{DR}^G - W_{min}) + b(W_{DR}^C - W_{min}) + b(W_{FO} - W_{min}), \quad (6)$$

where: W_{DR}^G is the equivalent channel width of the driving gate (modeled as an equivalent inverter), accounting for the driving conductance; W_{DR}^C is the equivalent channel width of the driving gate, accounting for the output capacitance; W_{FO} is the equivalent channel width of the fan-out gates (modeled as an equivalent inverter); W_{min} is the minimum channel width for the considered technology; Q_{min} is the Q_{crit} when $W_{DR} = W_{FO} = W_{min}$ (constant for a given technology and power supply); a and b are fitting parameters, weighting the contribution to Q_{crit} of the conductance of the driver and the load capacitance, respectively. The model in (6), however, does not account for aging effects on the SER during circuit lifetime.

III. Analysis of the BTI Impact on the Critical Charge

In this section, we analyze the impact of the n/pMOS transistor threshold voltage shift (ΔV_{th}) induced by BTI on the Q_{crit} of combinational circuits. Particularly, by means of HSPICE simulations, we evaluate the Q_{crit} variation at the output of NOT, NAND and NOR gates, with a variable number of inputs (up to 4).

A. BTI-Induced Q_{crit} Variation

We have implemented elementary gates by a High Performance Metal Gate, High-K, Strained-Si, 32nm CMOS technology (with 1V power supply) from PTM [27]. The gates have been designed in order to be with minimum area and symmetric. We have evaluated the threshold voltage shift (ΔV_{th}) of n/pMOS transistors induced by PBTI/NBTI degradation by utilizing the model described in Section II. Particularly, we have simulated alternating stress and recovery phases, considering a 50% switching activity of inputs, and an operating time up to 10 years. The estimated voltage shifts ΔV_{th} for the considered operating time have been adopted to customize the HSPICE device model employed for simulation. For each gate, we have considered the worst case, that is the degradation effect inducing the larger decrement of Q_{crit}.

Consider a generic symmetric gate described by an equivalent inverter [28]. The currents provided by the pull-up and pull-down equivalent transistors are equal to each other. Assuming that the transistors always work in velocity saturation condition [29], at time 0 it is:

$$I_{Dn}^0 = I_{Dp}^0 \Rightarrow W_{eqn} C_{ox}(V_{GSn} - V_{Tn}^0) v_{satn} = W_{eqp} C_{ox}\left(V_{SGp} - \left|V_{Tp}^0\right|\right) v_{satp}, \quad (7)$$

where V_{GSn}, v_{satn}, (V_{SGp}, v_{satp}) are the gate-source (source-gate) voltage difference and the velocity saturation of the nMOS (pMOS) transistors, respectively; V_{Tp}^0 (V_{Tn}^0) is the threshold voltage of pMOS (nMOS) transistors at time 0; C_{ox} is the oxide capacitance (per unit of area); W_{eqn} (W_{eqp}) is the equivalent channel width of the nMOS (pMOS) transistors of the equivalent inverter. Denoting $K_{eqn(p)} = W_{eqn(p)}\, C_{ox}\, v_{satn(p)}$, we can write:

$$K_{eqn}(V_{GSn} - V_{Tn}^0) = K_{eqp}\left(V_{SGp} - \left|V_{Tp}^0\right|\right) = I_D^0. \quad (8)$$

Considering the BTI-induced degradation, at the generic time instant time t, the drain current for the nMOS and pMOS transistors are:

$$I_{Dn}(t) = K_{eqn}\left[V_{GSn} - \left(V_{Tn}^0 + \Delta V_{Tn}(t)\right)\right] = I_D^0 - K_{eqn}\Delta V_{Tn}(t)$$
$$I_{Dp}(t) = K_{eqp}\left[V_{SGp} - \left(\left|V_{Tp}^0 - \Delta V_{Tp}(t)\right|\right)\right] = I_D^0 - K_{eqp}\Delta V_{Tp}(t), \quad (9)$$

where the values of the transistor threshold voltage variations $\Delta V_{Tn(p)}$ can be calculated by means of (3).

Since it is $V_{Tn}^0 \cong \left|V_{Tp}^0\right|$, (8) implies that $K_{eqn} \cong K_{eqp}$. Moreover, from (3) it is $\Delta V_{Tp}(t) = 2\Delta V_{Tn}(t)$. Then, from (9), it can be derived that it is always $I_{Dp}(t) \leq I_{Dn}(t), \forall t > 0$. Consequently, it is $G_p(t) \leq G_n(t), \forall t > 0$, where $G_n(t)$ ($G_p(t)$) denotes the conductance of the nMOS (pMOS) transistors as a function of time.

As an example, Fig. 1 reports the simulation results obtained when two alpha particles, with the same energy, hit the same node within a NOT chain, at two different instants during circuit life time: i) at the beginning of the gate operating life (Fig. 1(a)); ii) after 10 years of operating time (Fig. 1(b)). We have simulated the worst case scenario, when the hit node is driven by a pMOS transistor. As can be seen, even if the energy of the two hitting particles is the same, the glitch induced after 10 years of circuit operation has a higher amplitude than that at the beginning of circuit lifetime. This means that the Q_{crit} is considerably smaller after 10 years of circuit operation than at the beginning of circuit lifetime.

1) Symmetric, Minimum-Sized NOT Gate

Figure 2 reports the values of Q_{crit} at the output of a minimum sized, symmetric inverter, for different operating times. It depicts the case of a hitting particle temporarily charging (discharging) the output node, when it is driven by the pull-down (pull-up) network affected by PBTI (NBTI) degradation. At t=0 (non-aged circuit), the Q_{crit} values are approximately the same, whether the pull-up (solid line), or the pull down network (dashed line) is active. This is in accordance to the symmetric design of the NOT. As the circuit ages, the Q_{crit} obtained when the output is driven by the pMOS transistor decreases with a higher rate than when it is driven by the nMOS transistor. The difference in ΔV_{th} reaches -9.3% for t=10 years.

The value of Q_{crit} is a decreasing, monotonic function. During the first 2-3 years of circuit operation, Q_{crit} decreases with a much higher rate than during the remaining operating time. After the third year, the curve tends very slowly to its lowest value, which is reached after 10 years of circuit operation.

2) Symmetric Minimum-Sized NOR gate

Table I reports the values of the Q_{crit} at the output of a NOR gate with 2, 3 and 4 inputs, for several values of operating time, in case of both NBTI affecting the pull-up network and PBTI affecting the pull-down network. For this latter case, we have considered only one nMOS transistor ON, since this represents the worst case condition for the value of Q_{crit} when the output is driven by the pull-down network [20]. Likewise the NOT gate, the worst case condition is when the output is driven by the pull-up network. In this case, the larger degradation induced by NBTI is reinforced by the fact that the pMOS transistors are connected in series, thus making the gate suffer from cumulative performance degradation.

In case of no degradation (t=0) and symmetric gate, the Q_{crit} can be equivalently computed by considering the pull-up network, or a single nMOS transistor of the pull-down network. In this case, it is $G_{p0}/m = G_{n0}$, where G_{p0} (G_{n0}) is the conductance of the pMOS (nMOS) transistors at the beginning of circuit life time (t=0), while m is the number of inputs.

TABLE I
CRITICAL CHARGE (FC) FOR MINIMUM SIZED SYMMETRIC NOR GATES WITH 2, 3 AND 4 INPUTS

Operating Time	2 IN		3 IN		4 IN	
	NBTI	PBTI	NBTI	PBTI	NBTI	PBTI
0	10.16	10.16	10.34	10.34	10.68	10.68
1y	9.01	9.62	9.18	9.80	9.23	9.91
5y	8.47	9.32	8.54	9.52	8.58	9.64
10y	8.25	9.26	8.38	9.45	8.40	9.58

As the NOR gate starts aging (t>0), it is $G_p(t)/m < G_n(t)$, therefore $Q_{crit}(NBTI) < Q_{crit}(PBTI)$. Of course, the Q_{crit} increases with the increase in the number of inputs, since the parasitic capacitance at the output node increases as well.

Finally, it is interesting to observe that the relative difference between Q_{crit} at t=10 years, and Q_{crit} at t=0 also increases with the number of inputs. It ranges from -17.8% for a 2-input NOR, to -21.3% for a 4-input NOR. This can be explained by considering that the number of pMOS series transistors increases with the number of inputs, thus giving rise to a cumulative degradation effect on Q_{crit}.

3) Symmetric, Minimum-Sized NAND gate

We evaluated the Q_{crit} at the output of a minimum sized, symmetric NAND gate with 2, 3 and 4 inputs, when the output is driven by either the pull-up or the pull-down network. The considered active pull-up network consists of a single transistor for all three cases, since this represents the worst case condition for Q_{crit} evaluation. Table II reports the obtained results. In case of no degradation (time 0), the Q_{crit} at the output of the symmetric NAND gate can be equivalently evaluated by considering the output driven either by a single

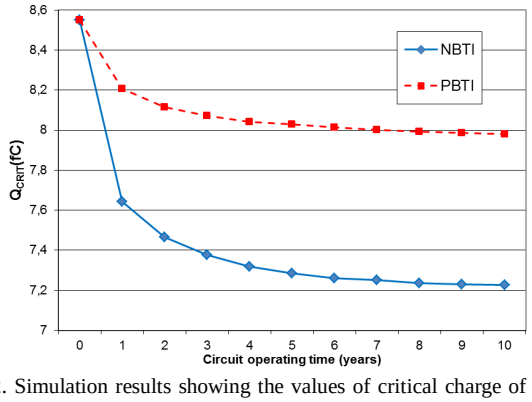


Fig. 2. Simulation results showing the values of critical charge of a NOT gate, as a function of circuit lifetime.

pMOS transistor of the pull-up network (worst case condition), or by the pull-down network. In fact, for a symmetric NAND gate, it is $G_{n0}/m = G_{p0}$. Moreover, also in this case we expect that, when the circuit starts aging, the worst case Q_{crit} scenario is encountered when the output of the NAND is driven by the pull-up network, since it is $G_{np}(t) < G_n(t)/m$. Therefore, it is $Q_{crit}(NBTI) < Q_{crit}(PBTI)$ for t>0.

TABLE II
CRITICAL CHARGE (FC) FOR MINIMUM SIZED SYMMETRIC NAND GATES WITH 2, 3 AND 4 INPUTS

Operating Time	2 IN		3 IN		4 IN	
	NBTI	PBTI	NBTI	PBTI	NBTI	PBTI
0	5.13	5.13	3.97	3.97	3.60	3.60
1y	4.71	4.88	3.61	3.83	3.27	3.53
5y	4.40	4.75	3.42	3.73	3.10	3.44
10y	4.34	4.73	3.39	3.71	3.06	3.42

The Q_{crit} decreases with the increase in the number of inputs. In fact, considering minimum-sized, symmetric gates, the conductance of the pull-up/pull-down network under worst case condition decreases with the number of inputs. In Fig. 4 we can notice that the Q_{crit} degradation is less dependent on the number of inputs than for the NOR gate. For all considered number of inputs, the relative difference between Q_{crit} at t=10 years and at t=0 is 15%-16%. This can be explained considering that, differently from the case of the NOR gate, here the pull-up network considered for Q_{crit} evaluation consists of a single pMOS transistor independently of the number of inputs.

B. NOR and NAND Gates Q_{crit} Comparison

In the previous subsection we have shown that the Q_{crit} of minimum-sized symmetric NOR gates exceeds that of minimum-sized symmetric NAND gates. This might lead to the conclusion that NOR gates are more robust to SE and aging than NAND gates: this is not true. In fact, the difference shown in the previous subsection depends on the sizes of the two gates. Minimum-sized symmetric NANDs are considerably smaller than minimum-sized symmetric NORs. We have estimated the gate area in terms of squares (□), where a square represent an area equal to 0.32μm × 0.32μm = 0.1024μm^2. From the sizes reported in Sect. 3.1, we can derive the following areas:

A_{NOR-2}= 11.4□; A_{NOR-3}= 24.3□; A_{NOR-4}= 42□

A_{NAND-2}= 4.7□; A_{NAND-3}= 6.2□; A_{NAND-4}= 8.4□

Moreover, as known, NAND gates outperform NOR gates of the same area. We will show here that NAND gates are also more robust to BTI aging than NOR gates. For comparison purposes, we will consider NAND and NOR gates with the same area. The obtained simulation results are depicted in Figs. 3 (a), (b) and (c) for 2, 3 and 4 input gates, respectively. As can be seen, the Q_{crit} values obtained for the NAND gates are considerably higher than those for NOR gates of the same area. The relative difference slightly increases with time and number of inputs. It ranges from 1.4X at t=0, for 2-inputs gates, to 2.3X at t=10 years, for 4-input gates. Therefore, we can conclude that NAND gates are considerably more robust against soft errors and BTI aging effects.

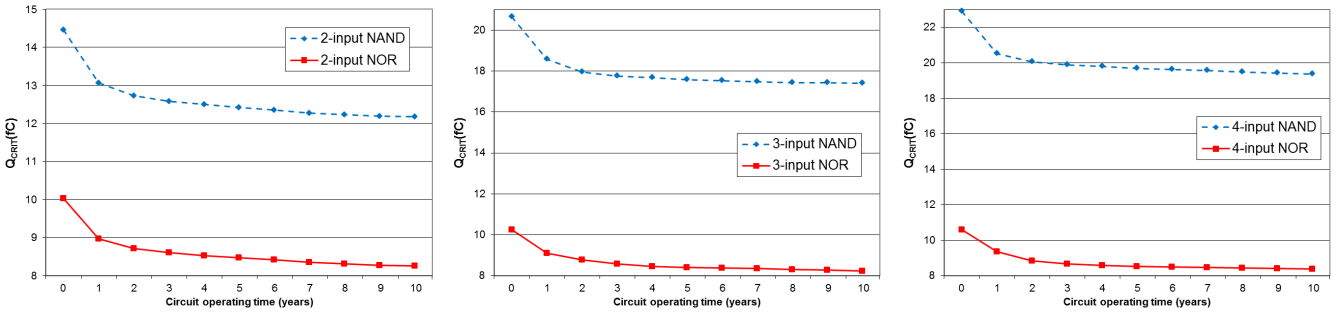
We have also considered the case of NAND and NOR gates exhibiting an identical delay while driving an identical load. As expected, in this case, NAND and NOR gates, having the same conductance, present approximately the same Q_{crit} at their outputs.

IV. EVALUATION OF SER VARIATION DUE TO BTI

Consider a particle hitting a node j of a combinational circuit, and generating a SET. To generate a SE, the SET must propagate up to a storage element m, satisfying its set-up and hold times. The probability that a SE is generated is [30]:

$$P_{SE,jm} = P_{GEN,j} P_{SENS,jm} P_{PROP,jm} P_{LATCH,jm}, \quad (10)$$

where: $P_{GEN,j}$ is the probability that a particle hitting the node j generates a collected charge higher than $Q_{crit,j}$; $P_{SENS,jm}$ is the probability that the path between node j and the storage element m is sensitized by the input configuration; $P_{PROP,jm}$ is the probability that the glitch generated at node j arrives at the input of the storage element m with amplitude and duration large enough to be sampled; $P_{LATCH,jm}$ is the probability that the generated voltage glitch satisfies the set-up and hold times of the storage element and gets latched. The overall P_{SE} is:



(a) (b) (c)

Fig. 3. Worst case Q_{crit} as a function of circuit lifetime, for NANDs (dashed lines) and NORs (solid lines) with 2 (a), 3 (b) and 4 (c) inputs.

$$P_{SE} = \sum_j \left(\sum_m P_{SE,jm} \right) = \sum_j P_{GEN,j} \left(\sum_m P_{SENS,jm} P_{PROP,jm} P_{LATCH,jm} \right).$$

It is worth noting that P_{SENS} and P_{LATCH} are not affected by BTI, thus they do not vary with time. As for P_{PROP}, it can slightly decrease with aging, since the ability of a gate to propagate a voltage glitch depends on the gate conductance [30]. However, in this work, we focus on the generation probability, which enables to determine the maximum impact of BTI on SE generation probability. Therefore, we assess only BTI effect on P_{GEN}. For the node j, it is [5]:

$$P_{GEN,j}(t) = k_j \phi \frac{\gamma}{\sigma} e^{-\sigma Q_{crit,j}(t)}, \quad (11)$$

where ϕ is a parameter depending on the alpha-particle flux, while γ and σ are fitting parameters. The parameter k_j accounts for the probability that a particle impacting the considered circuit hits the critical area at node j, denoted by A_j. It is: $k_j = A_j / A_{TOT}$, where A_{TOT} is the total area of the circuit, evaluated as the summation of the transistors' gate and junctions area.

The term A_j has been evaluated as the critical area of the gate driving the node j. Particularly, when the pull-up network of the driver is ON, the critical area is the drain junction of the nMOS transistor in the OFF state connected to the output node of the gate. Similarly, when the pull-down is ON, the area of the drain junction of the pMOS transistor in the OFF state connected to the output node of the gate should be considered. Therefore, not only different critical charges, but also different critical areas must be accounted for when a generic node j is driven by the pull-up or pull-down network.

Considering equal to 0.5 the probability that a generic node j is driven by the pull-up/pull-down network, the SET generation probability in (11) can be written as:

$$P_{GEN,j}(t) = \frac{1}{2} \phi \frac{\gamma}{\sigma} \left(k_j^{(n)} e^{-\sigma Q_{crit,j}^{(p)}(t)} + k_j^{(p)} e^{-\sigma Q_{crit,j}^{(n)}(t)} \right). \quad (12)$$

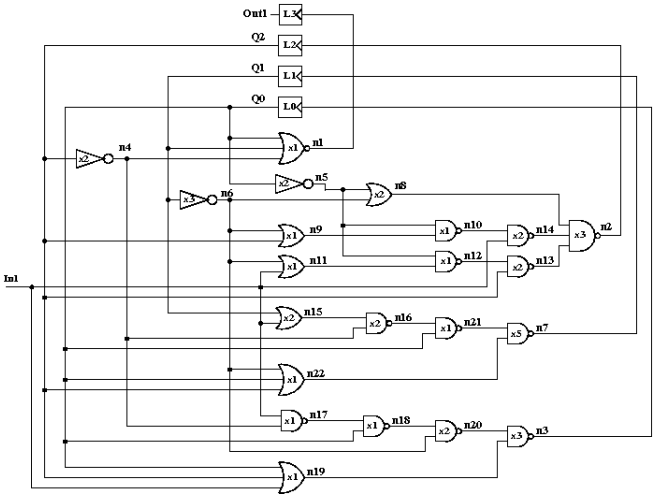
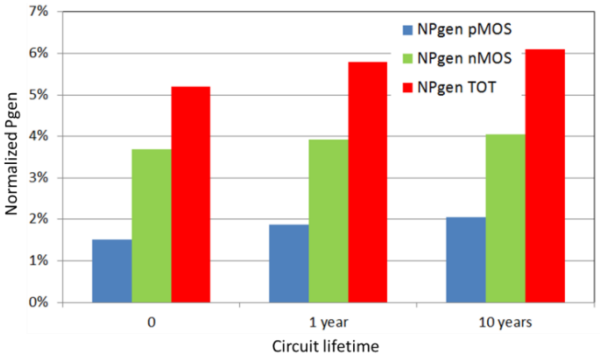


Fig. 4. ITC'99 b02 benchmark.

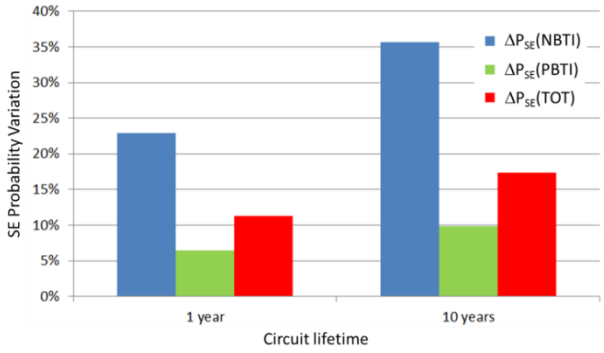
Therefore, the maximum SER variation induced by BTI during circuit lifetime (t>0), with respect to its value at t=0, is:

$$\Delta SER^{max}(t) = \Delta P_{SE}^{max}(t) = \Delta P_{GEN} = \sum_j \frac{P_{GEN,j}(t) - P_{GEN,j}(0)}{P_{GEN,j}(0)} = \sum_j \frac{k_j^{(n)} \left[e^{-\sigma Q_{crit,j}^{(p)}(t)} - e^{-\sigma Q_{crit,j}^{(p)}(0)} \right]}{k_j^{(n)} e^{-\sigma Q_{crit,j}^{(p)}(0)} + k_j^{(p)} e^{-\sigma Q_{crit,j}^{(n)}(0)}} + \sum_j \frac{k_j^{(p)} \left[e^{-\sigma Q_{crit,j}^{(n)}(t)} - e^{-\sigma Q_{crit,j}^{(n)}(t)} \right]}{k_j^{(n)} e^{-\sigma Q_{crit,j}^{(p)}(0)} + k_j^{(p)} e^{-\sigma Q_{crit,j}^{(n)}(0)}}. \quad (13)$$

In order to assess the SER variation with circuit aging, we have evaluated the Q_{crit} of the nodes of the benchmark circuit b02 (Fig. 4) from the ITC'99 benchmark set [31]. We have considered gates with different sizes and 1 year and 10 years aging. For each node, we have estimated the Q_{crit} and the critical area when the node is driven either by the pull-up or pull-down network. The results are depicted in Fig. 5. Particularly, Fig. 5(a) shows the generation probability, normalized with respect to the coefficient $\phi\gamma/\sigma$, which is not impacted by aging: NPgen-p represents the normalized generation probability obtained when the nodes are driven by the pull-up networks; analogously for NPgen-n and the pull-down network. NPgen-TOT represents the cumulative probability. The probability values obtained at the beginning



(a)



(b)

Fig. 5. (a) Alpha particle induced voltage glitch generation probability; (b) SE probability relative variation.

of circuit lifetime, as well as after 1 year and 10 years of operating time have been depicted. As can be seen, NPgen-n is always higher than NPgen-p. This is in contrast with the fact that the Q_{crit} decrease due to NBTI is larger than that due to PBTI, as discussed in the previous section. However, in a symmetric gate, the critical area when the pull-up network is ON is usually smaller than when the pull-down network is ON. As a result, in (12) the contribution of the terms k_j prevails over that of the exponential terms, and NPgen-n exceeds NPgen-p for the whole circuit lifetime.

In Fig. 5(b), we represent the relative variation of the SE probabilities $\Delta P_{SE} = \Delta P_{GEN}$ due to NBTI, PBTI and the total one, after 1 and 10 years of circuit operation, over the respective values at the beginning of circuit lifetime. As can be seen, ΔP_{SE}(TOT) ranges from more than 11% after one year, to more than 17% after ten year of circuit operation.

V. Proposed Time-Dependent Model for SE Susceptibility Evaluation

The results presented in the previous section have highlighted that the Q_{crit} decreases considerably with circuit aging due to BTI. This poses new challenges to IC SER modeling, mandating a time dependent modeling. Based on these results, in this section we propose a new time-dependent model of Q_{crit}, enabling to account for BTI-induced degradation.

The static model in (6), allowing us to estimate the node Q_{crit} as a function of load capacitance and driver conductance, has been considered here as the starting point for our dynamic model. Let us express Q_{crit} as a function of time as follows:

$$Q_{crit}(W_{DR}, W_{FO}, t) = Q_{min}(t) + a(t)(W_{DR}^{G} - W_{min}) + b_0(W_{DR}^{C} - W_{min}) + b_0(W_{FO} - W_{min}), \quad (14)$$

where: $Q_{min}(t)$ is a time-dependent variable accounting for the critical charge at the output of a minimum sized inverter loaded by another minimum sized inverter; $a(t)$ is a time-dependent coefficient accounting for the contribution to Q_{crit} of the conductance of the driving circuit; b_0 is a constant coefficient taking into account the contribution to Q_{crit} of the load capacitance. It coincides with the correspondent b parameter in (6), as well as the parameters W_{DR}^{G}, $W_{DR}^{C}$and W_{FO}.

Differently from the static model in (6), the terms including the contribution of the transistors' conductance (Q_{min} and a) vary over time, since they are affected by BTI degradation. Instead, the term b_0 is constant, since the node capacitance value is not affected by BTI. As recalled in Sect. 2.1, BTI degradation induces a transistor threshold voltage increase (in absolute value) with a time dependency of $t^{1/6}$. Therefore, considering the linear relationship between the transistor current and the threshold voltage in (9), we can expect that also the transistor conductance degrades with the same time dependence. Based on these considerations, the time dependent terms in (14) can be written as:

$$Q_{min}(t) = Q_{min0}\left(1 - k_{1(n,p)}t^{1/6}\right)$$
$$a(t) = a_0\left(1 \mp k_{2(n,p)}t^{1/6}\right), \quad (15)$$

where Q_{min0} and a_0 are the values assumed by such parameters at time 0, and coincide with the respective values in (6). The coefficients $k_{1(n,p)}$ and $k_{2(n,p)}$ are fitting parameters that depend on technology and Si-dielectric interface trap dynamics, accounting for PBTI and NBTI, respectively.

The sign of in the time-dependent term should be chosen as follows. Considering $\Delta W_{eq} = (W_{DR}^{G} - W_{min})$, if: i) $\Delta W_{eq} \geq 0$, the negative sign has to be chosen; ii) $\Delta W_{eq} < 0$, the positive sign must be taken. This way, the contribution given to Q_{crit} by the term accounting for the driver conductance properly diminishes with circuit aging. Let us analyze in details the case of $\Delta W_{eq} < 0$ (e.g., the case of a 2-input NAND gate with minimum sized transistors). In this case, (14) can be written as:

$$Q_{crit} = Q_{min0}\left(1 - k_1 t^{1/6}\right) + a_0\left(1 + k_2 t^{1/6}\right)\Delta W_{eq} + \cdots .$$

Since it is $\Delta W_{eq} < 0$, the contribution $a_0 \Delta W_{eq}$ correctly subtracts to Q_{min0}, at $t = 0$. Moreover, for $t > 0$, the term $a_0 k_2 t^{1/6} \Delta W_{eq}$ is negative, accounting for the decrease of the Q_{crit} as circuit ages.

In order to estimate Q_{crit}, the coefficients in (14) and (15) have been fitted with the results of the HSPICE simulations, considering the device model as reported in Sect. 3.1. The coefficients Q_{min0}, a_0 and b_0 have been obtained by simulating a chain of two symmetric inverters I1 (driver) and I2 (load), with the following configurations: a) I1 and I2 both minimum sized (to estimate Q_{min0}); b) I1 3X and I2 minimum sized (to estimate a_0); c) I1 minimum sized and I2 3X (to estimate b_0). The coefficient $k_{1(n,p)}$ and $k_{2(n,p)}$ in (11) have been fitted considering a chain of two inverters I1 (driver) and I2 (load), 1 year aging, and the configurations a) and b) described above. The fitting parameters should be recalculated for each process corner. However, only a few simulations of two cascaded inverters are required.

The current induced by the alpha-particle hit has been modeled by a double exponential pulse as in (4). The values obtained are reported in Table III. As can be seen, the parameter a_0 is two orders of magnitude greater than b_0. As highlighted also in [8], this means that the Q_{crit} of a circuit node depends much more on the driving gate conductance, rather than on the node capacitance.

VI. Proposed Model Validation and Accuracy

We have evaluated the accuracy of our proposed model by determining the Q_{crit} at the outputs of different gates, with different fan-ins and fan-outs, for a circuit operating time up to 10 years. The gates have been implemented by the same 32nm CMOS technology considered before.

TABLE III
VALUES OF THE COEFFICIENTS OF THE PROPOSED MODEL

Q_{min0} [fC)]	a_o [fC/ nm]	b_0 [fC/ nm]	$k_{1(p)}$ $k_{1(n)}$ [$s^{-1/6}$]	$k_{2(p)}$ $k_{2(n)}$ [$s^{-1/6}$]
8.55	306.4×10^6	1.781×10^6	5.89×10^{-3} 2.26×10^{-3}	6.57×10^{-3} 2.96×10^{-3}

The Q_{crit} at the output node of each gate has been determined by means of HSPICE simulations and by our model. For each node, the driving and fan-out gates have been mapped into equivalent symmetric inverters [6]. Fig. 6 reports the trend in time of the relative error in the Q_{crit} prediction of our model over HSPICE simulations. As can be seen, the error is always lower than the 5%, but for the case of NOR gates with 3 inputs (max relative error equal to 5.3%), and 4 inputs (max relative error equal to 7.1%).

The proposed model has been validated considering also the benchmark circuit ITC'99 b02. Particularly, Fig. 7 reports the values of the normalized *Pgen* (as defined and calculated in Sect. 4) obtained by HSPICE (Sim) and by adopting our proposed model for Q_{crit} estimation (Mod). As can be seen, the relative error is approximately 3%, for 1 year and 10 years of circuit operating time.

Therefore, the proposed time-dependent model is able to predict with a very good accuracy over electrical level simulations the Q_{crit} at the outputs of elementary gates during their life time, as well as the impact of BTI on the SER of more complex circuits. This represents a great advantage over the models proposed so far in literature, none of which accounts for the BTI-induced degradation of Q_{crit} during circuit lifetime.

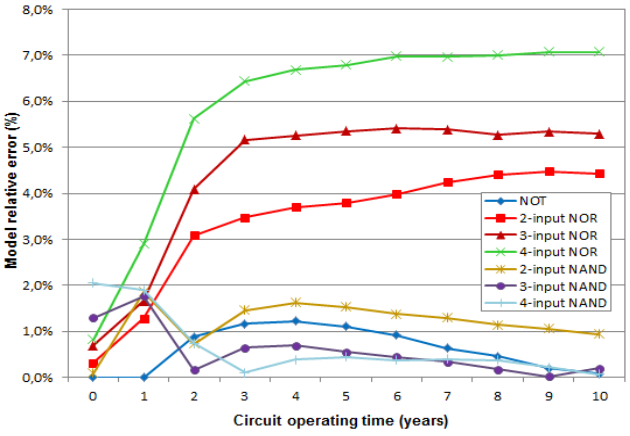


Fig. 6. Relative error of the proposed model predictions over HSPICE simulations.

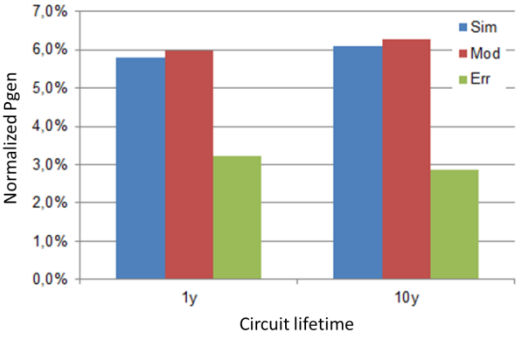


Fig. 7. Normalized Pgen obtained by HSPICE simulations (Sim) and by the proposed model(Mod), and relative error (Err).

VII. CONCLUSIONS

We have analyzed the effects of BTI on elementary gates, as well as on SER of complex circuits. We have shown that BTI may reduce significantly the Q_{crit} of their nodes, during their in-field operation. Particularly, NOR gates present a reduction of Q_{crit} higher than NAND gates of the same area.

We have then proposed a time dependent model of Q_{crit} that, differently from the static models developed so far, is able to account for the impact of BTI degradation on the Q_{crit}, thus on the IC SER. The proposed model features a very good accuracy (below 4%average error and 7.1% maximum error below in Q_{crit} estimation) over HSPICE simulations, for circuit lifetime up to 10 years.

REFERENCES

[1] N. Seifert, B. Gill, V. Zia, M. Zhang, V. Ambrose, "On the Scalability of Redundancy Based SER Mitigation Schemes", in Proc. of Int. Conf. on Integrated Circuit design and Technology, 2007.

[2] D. G. Mavis, P. H. Eaton, "SEU and SET Modeling and Mitigation in Deep Submicron Technologies", in Proc. of Annual IEEE Int. Reliability Physics Symp, 2007, pp. 293-305.

[3] R. C. Baumann, "Radiation-Induced Soft Errors in Advanced Semiconductor Technologies", IEEE Trans. on Device and Materials Reliability, Vol.5, No.3, September 2005, pp. 305 - 316.

[4] D. Rossi, M. Omaña, C. Metra "Transient Fault and Soft Error On-die Monitoring Scheme", in Proc. of IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, 2010, pp. 391-398.

[5] M. Omaña, D. Rossi, C. Metra, "High Performance Robust Latches", IEEE Trans. on Computers, vol. 59, no. 11, Nov. 2010, pp. 1455-1465,

[6] D. Rossi, J. M. Cazeaux, M. Omaña, C. Metra, A. Chatterjee, "Accurate Linear Model for SET Critical Charge Estimation", *IEEE Trans. on VLSI Systems*, Vol. 17, No. 8, Aug. 2009, pp. 1161 – 1166.

[7] J. M. Cazeaux, D. Rossi, M. Omaña, A. Chatterjee, C. Metra, "On-Transistor Level Gate Sizing for Increased Robustness to Transient Faults", in Proc. of 11th IEEE International On-Line Testing Symposium, 2005, pp. 23-28,

[8] Q. Zhou, K. Mohanram, "Gate Sizing to Radiation Harden Combinational Logic", IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, Vol. 25, No. 1, January 2006, pp. 155-166.

[9] J. Keane, T-H. Kim, C. H. Kim, "An On-Chip NBTI Sensor for Measuring pMOS Threshold Voltage Degradation", IEEE Trans. On Very Large Scale Integration (VLSI) Syst., 2009.

[10] V. Huard, M. Denais, "Hole Trapping Effect on Methodology for DC and AC Negative Bias Temperature Instability Measurements in PMOS Transistors", in Proc. of IEEE Int. Rel. Physics Symp., pp 40-45, 2004.

[11] M. Agarwal, V. Balakrishnan, A. Bhuyan, K. Kim, B.C. Paul, W. Wang, B. Yang, Y. Cao, S. Mitra, "Optimized Circuit Failure Prediction for Aging: Practicality and Promise", in Proc. *of IEEE Int. Test Conf.*, pp. 1-10, 2008.

[12] M.A. Alam, S. Mahapatra, "A Comprehensive Model of PMOS NBTI Degradation", Microelectronics Reliability vol. 45, 2005, Elsevier, pp. 71-81.

[13] W. Wang, Z. Wei, S. Yang, Y.Cao, "An Efficient Method to Identify Critical Gates under Circuit Aging", in Proc. of IEEE/ACM Int. Conf. on Computer-Aided Design, pp. 735-740, 2007.

[14] S. Khan, S. Hamidioui, H. Kukner, P. Raghavan, F. Catthoor, "BTI

Impact on Logical Gates in Nano-scale CMOS Technology" in Proc. of 2012 IEEE 15th International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS), pp. 348-353.
[15] M. Omaña, D. Rossi, N. Bosio, C. Metra, "Self-Checking Monitor for NBTI Due Degradation," in IEEE Proc. of Int. Mixed-Signals, Sensors, and Systems Test Workshop (IMS3TW'10), 2010.
[16] M. Omaña, D. Rossi, N. Bosio, C. Metra, "Low Cost NBTI Degradation Detection & Masking Approaches", *IEEE Transactions on Computers*, vol. 62, no. 3, March 2013, pp. 496-509.
[17] M. Bagatin, S. Gerardin, A. Paccagnella, F. Faccio, "Impact of NBTI Aging on the Single-Event Upset of SRAM Cells", IEEE Transactions on Nuclear Science, Vol. 57, No. 6, December 2010, pp. 3245-3250.
[18] E.H. Cannon, A. Klein Osowski, R. Kanj, D.D. Reinhardt, R.V. Joshi, "The Impact of Aging Effects and Manufacturing Variation on SRAM Soft-Error Rate", IEEE Transactions on Device and Materials Reliability, Vol. 8, Issue 1, 2008, pp. 145-152.
[19] K. Ramakrishnan, R. Rajaraman, S. Suresh, N. Vijaykrishnan, Y. Xie, M.J. Irwin, "Variation Impact on SER of Combinational Circuits", in Proc. of Int. Symp. on Quality Electronic Design, 2007, pp. 911 – 916.
[20] D. Rossi, M. Omaña, C. Metra, A. Paccagnella, "Impact of Aging Phenomena on Soft Error Susceptibility", in Proc. of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, 2011, pp. 18-24.
[21] R. Harada, Y. Mitsuyama, M. Hashimoto, T. Onoye, "Impact of NBTI-Induced Pulse-Width Modulation on SET Pulse-Width Measurement", IEEE Transactions on Nuclear Science, vol. 60, no. 4, Aug. 2013, pp 2630 - 2634
[22] S. K. Krishnappa, H. Mahmoodi, "Comparative BTI Reliability Analysis of SRAM Cell Designs in Nano-Scale CMOSTech nology", in Proc. of IEEE 12th Int'l Symp. on Quality Electronic Design, 2011, pp. 384-389.
[23] M. Toledano-Luque, B. Kaczer, J. Franco, Ph. J. Roussel, T. Grasser, T. Y. Hoffmann, G. Groeseneken, "From Mean Values to Distribution of BTI Lifetime of Deeply Scaled FETs Through Atomistic Understanding of the Degradation", in 2011Symposium on VLSI Technology, Digest of Technical Papers, pp. 152-153.
[24] G. C. Messenger, "Collection of Charge on Junction Nodes from Ion Tracks", IEEE Trans. on Nuclear Science, NS-29(6), December 1982, pp. 2024 - 2031.
[25] H. Cha and J. H. Patel, "A Logic-Level Model for α-Particle Hits in CMOS Circuits", in Proc. of IEEE Int. Conf. on Computer Design, 1993, pp. 538-542.
[26] P. Hazucha and C. Svensson, "Impact of CMOS Technology Scaling on the Atmospheric Neutron Soft Error Rate," IEEE Trans. on Nuclear Science, pp. 2586 – 2594, Dec. 2000.
[27] Predictive Technology Model (PTM), http://ptm.asu.edu/.
[28] N. Weste and D. Harris, *CMOS VLSI Design A Circuits and Systems Perspective*. New York: Addison-Wesley, 2004.
[29] J. M. Rabaey, A. Chandrakasan, B. Nikolic, *Digital Integrated Circuits* - Second Edition. Pearson - Prentice Hall, 2003.
[30] M. Omaña, D. Rossi, C. Metra, "Model for Transient Fault Susceptibility of Combinational Circuits", Journal of Electronic Testing, Theory and Application (JETTA), vol. 20, no. 5, Oct. 2004, pp.501-509.
[31] http://www.cerc.utexas.edu/itc99-benchmarks/bench.html

Daniele Rossi received the degree in electronic engineering and the PhD degree in electronic engineering and computer science from the University of Bologna in 2001 and 2005, respectively. He is currently a Post-Doctoral fellow at the same University. His research interests include fault modeling and fault tolerance, focusing on coding techniques for fault-tolerance and low-power, signal integrity for communication infrastructures, robust design for soft error resiliency.

Martin Omaña received the degree in electronic engineering from the University of Buenos Aires in 2000. In 2002 he was awarded a MADESS grant and joined the University of Bologna, where he obtained the PhD in Electronic Engineering and Computer Science in 2005. He is currently a Post-Doctoral fellow at the same University. His research interests are in the field of fault modeling, on-line test, robust design, fault tolerance and photovoltaic systems.

Cecilia Metra (F'14) is a Professor of Electronics with the University of Bologna, Bologna, Italy. Her current research interests include fault modeling, on-line test, robust design, fault tolerance, energy harvesting, and photovoltaic systems.
Prof. Metra is a Vice-President for Technical and Conference Activities of the IEEE Computer Society (CS) for 2014, and a member of the Board of Governors of the IEEE CS 2013-2015. Since 2013, she has been Editor-in-Chief of the IEEE CS online publication Computing. She is a Golden Core Member of the IEEE CS.

Alessandro Paccagnella is Full Professor of Electronics at the University of Padova. A. P. is author of more than 400 scientific papers, and about 300 of them have been published on international journals. In the past, his research activity has been directed to the study of different aspects of physics, technology, and reliability of semiconductor devices. In relation with this activity he spent some research periods at the University of California, San Diego, and at the IBM T.J. Watson Research Center, Yorktown Heights, New York. At present, he coordinates the research activity of a group at the Department of Information Engineering, University of Padova, working mainly on CMOS devices and technology. His research work is focused on the study of the effects produced by ionizing radiation in various environments on single devices and integrated circuits, with emphasis on non-volatile memories.