

# Towards Interdependencies of Aging Mechanisms

Hussam Amrouch, Victor M. van Santen, Thomas Ebi, Volker Wenzel and Jörg Henkel

Chair for Embedded Systems (CES), Karlsruhe Institute of Technology (KIT), Germany

{amrouch, vantsanten, thomas.ebi, volker.wenzel, henkel}@kit.edu

**Abstract**—With technology in deep nano scale, the susceptibility of transistors to various aging mechanisms such as *Negative/Positive Bias Temperature Instability* (NBTI/PBTI) and *Hot Carrier Induced Degradation* (HCID) etc. is increasing. As a matter of fact, different aging mechanisms simultaneously occur in the gate dielectric of a transistor. In addition, scaling in conjunction with high-K materials has made aging mechanisms, that have often been assumed to be negligible (e.g., PBTI in NMOS and HCID in PMOS), become noticeable. Therefore, in this paper we investigate the key challenge of providing designers with an abstracted, yet accurate reliability estimation that combines, from the physical to system level, the effects of *multiple simultaneous* aging mechanisms and their interdependencies. We show that the overall aging can be modeled as a superposition of the interdependent aging effects. Our presented model deviates by around 6% from recent industrial physical measurements. We conclude from our experiments that an isolated treatment of individual aging mechanisms is insufficient to devise effective mitigation strategies in current and upcoming technology nodes. We also demonstrate that estimating reliability due to an individual dominant aging mechanism together with solely considering a single kind of failures, as currently is a main focus of state-of-the-art (e.g., [28], [22]), can result in 75% underestimation on average.

## I. INTRODUCTION

The International Technology Roadmap for Semiconductors states that upcoming technology nodes introduce reliability challenges at an increased pace compared to the last decade [1] because devices below 45 nm are increasingly susceptible to various aging mechanisms. We therefore focus within this work on the negative impact of aging on the probability of failures.

**Aging Effects:** Shrinking feature sizes leads to higher electric field strengths, as well as higher current densities, which both accelerate device aging and thus increase degradation of transistor parameters which can ultimately turn into failures. NBTI, PBTI and HCID have become the most prominent aging mechanisms impeding reliable transistors. Their effects on aging are more significant than others including Time-Dependent-Dielectric Breakdown (TDDB) [15], even in current high-K transistors [10]. While understanding the physical processes of aging mechanisms is not entirely required at the system level, there is still a substantial need to analyze their impact on degradations to accurately estimate reliability – *this holds even more when multiple aging mechanisms interdepend i.e. when they interact with each other.*

Of the two forms of BTI, NBTI degrades PMOS transistors and PBTI degrades NMOS transistors, whereas HCID degrades both. Over time, aging-induced degradations ultimately cause transistor malfunctions and increase a circuit's susceptibility to failures. Such failures are mainly due to timing violations and data corruption caused by voltage noise or radiation. We focus within the paper on how simultaneous occurring aging mechanisms jointly increase the probability of these failures.

**The Challenge of Combining Aging Effects:** Recently introduced physical-based aging models such as [11], [20] describe the detailed underlying physical process behind aging mechanisms to interpret them. Additionally, measurements have shown that these processes occur simultaneously [6], [12], [20]. Unlike higher-level aging models (e.g., [24]), physical-based models are more accurate but complex as they are highly

device-dependent and computationally intensive which is due to the large number of chemical bonds which need to be modeled along with their varying properties (e.g., the *Si-H* and *Si-O* bonds affected by BTI exhibit a wide range of density variability due to locally higher breaking rates induced by the interaction with HCID). As these models aim to fully capture the actual underlying physical processes along with modeling them in-depth (e.g., interpreting aging in the order of  $\mu\text{sec}$ ), their computationally intensive solutions are limited to a single transistor device – especially when aiming to study multiple mechanisms results in a significant increase in the complexity. This makes such solutions not feasible for designers at the system level dealing with tremendous number of transistors to estimate the impact of aging on the entire system lifetime (e.g., years). Finally, manufacturing variability also plays an important role as it varies transistors properties leading to different degradations. Therefore, paying attention to it is inevitable when estimating reliability.

**In summary:** Analyzing failures due to isolated individual aging mechanism is insufficient in order to estimate the overall reliability because the interdependencies do matter. Demonstrating this is our goal along with showing how the effects of multiple simultaneous mechanisms can be combined towards providing a more accurate reliability estimation.

## II. RELATED WORK

As reliability becomes a more imminent challenge, researchers increasingly focus on understanding and modeling the physical processes behind NBTI, PBTI [14] and HCID [20] aging mechanisms. When combining the effects of multiple aging mechanisms is targeted, state-of-the-art concepts (as also presented in Fig 1) can be categorized into:

1) System-level (e.g. [30], [31]): Where the mean-time-to-failure (MTTF) of each aging mechanism is individually calculated and, then the overall reliability is estimated based on the SOFR rule within the RAMP model [32]. The key problem behind this concept is the assumption that each failure mechanism proceeds independently, is not valid anymore because recent observations through measurements [20], [12] established that BTI aging mechanism *simultaneously* occurs with HCID. Additionally, [6] showed that HCID models have intrinsic BTI components and thus treating these mechanisms *separately* can lead to overestimating the overall aging-induced degradations due the twofold consideration of BTI (which itself is significant). Moreover, estimating the reliability degradations relying on MTTF (which is for the sake of simplicity and to keep the computational time low) can result in either ignoring (e.g., [9]) or oversimplifying (e.g., [17]) modeling the recovery mechanism of BTI leading to underestimating or overestimating, respectively, the self-healing impact on reliability when the voltage stress is ceased.

2) Circuit-level (e.g., [29], [13]): There, mainly the dominant aging mechanism in each transistor device is considered to estimate the overall reliability degradation of the entire circuit. In practice, it considers solely NBTI in PMOS and solely HCID in NMOS. This is due to the assumption that other aging mechanisms like HCID in PMOS and PBTI in NMOS

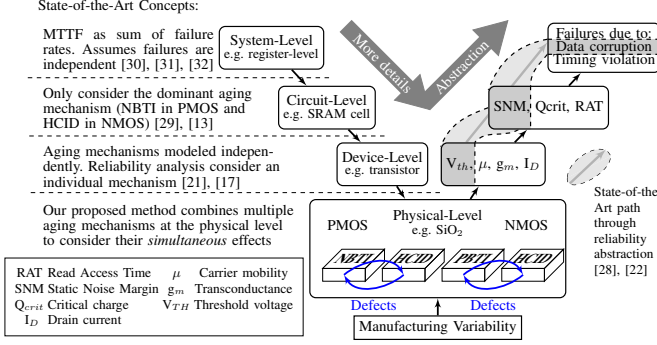


Fig. 1: Abstracting aging degradations

are negligible. This assumption was reasonable in the past with 65nm technology nodes but the recent introduction of high-K materials along with technology scaling (as found in 22nm) has strengthened the HCID mechanism in PMOS and the PBTI mechanism in NMOS [6]. Therefore, such a concept can lead to underestimating the induced degradations as it will be demonstrated in Section V. Another concept in [16] proposed to integrate multiple failure-equivalent circuits which model multiple aging mechanisms into the studied circuit. While SPICE can model the interrelations between these equivalent circuits e.g. the overall  $\Delta V_{TH}$  due to  $N$  aging mechanisms (which can shift  $V_{TH}$ ) will be represented as  $\sum_i \Delta V_{THi}$ , the assumption that aging mechanisms are independent still applies because each individual aging mechanism is represented with its own equivalent circuit, whereas in fact these aging mechanisms simultaneously interact at the physical level (which SPICE cannot model) and they need to be represented as one comprehensive failure-equivalent circuit rather than multiple ones to avoid overestimating the overall degradations.

3) Device-level: The Berkeley Reliability Tool (BERT) [21] and similarly the RelXpert [17] from Cadence (which is based on BERT) combines different aging simulators into a modular reliability framework. Each simulator module individually models an individual aging mechanism. While it can provide a rough guidance for design space exploration, it assumes, for the sake of enabling the modularity, that aging mechanisms are independent and thus existing interdependencies are missed.

Regarding exploring the impact of aging-induced degradations on systems, [18] showed how aging increases the susceptibility to timing violations, [22], [28] illustrated how the susceptibility to noise can also be increased and [8] reported the relation between aging and the susceptibility to radiation. While these works focus only on one particular kind of failure at a time when an individual aging mechanism (NBTI) is considered, they omit other kinds of failures – especially when multiple aging mechanisms occur simultaneously.

**Distinguishing from existing work**, we combined the effects of *multiple simultaneous* aging mechanisms from the physical level to accurately analyze the induced degradations at the circuit level and to provide an abstracted, yet sufficiently accurate reliability estimation at the system level summarizing how aging-induced defects in the transistors gate dielectric will ultimately increase probability of failures of the entire system.

### III. PROBLEM FORMULATION

System designers aim to estimate the lifetime of the systems in order to determine the cost of sustaining reliable operation during runtime (e.g., employing aging mitigation

techniques). The challenge is that there are several interdependencies between aging mechanisms that need to be carefully analyzed to correctly estimate transistor parameter degradations over time and their effects on reliability. Given various aging mechanisms  $M = \{m_1, m_2, \dots\}$ , the set of initial transistor parameters<sup>1</sup>  $\mathfrak{P}_{t_0} = \{p_1(t_0), p_2(t_0), \dots\}$  (e.g., threshold voltage, etc.), environment parameters  $E = \{\varepsilon_1(t), \varepsilon_2(t), \dots\}$  (e.g. temperature, voltage noise, radiation<sup>2</sup>), and a stress condition  $S(t)$ , aging can be expressed as a function  $A_{\mathfrak{P}} : M^n \times \mathfrak{P}^{\mathfrak{P}} \times E^{|E|} \times S \rightarrow \mathfrak{P}^j$ , where  $\mathfrak{P}^j$  is the set of  $j$  affected transistor parameters degraded by  $n$  mechanisms. Assuming that the time dependencies for  $E$  and  $S$  are given, the time dependency for  $p_i \in \mathfrak{P}$  can be expressed as

$$p_i(t) = \int_{t_0}^t A_{p_i}(m_1, \dots, m_n, \mathfrak{P}, E, S)(\hat{t}) d\hat{t} \quad (1)$$

Due to this recursive dependency of  $\mathfrak{P}$  on transistor parameters,  $\mathfrak{P}_1, \mathfrak{P}_2 = \mathfrak{P}_{t_k}$  at time  $t_k > 0$  for  $A_{\mathfrak{P}}$  considering two non-empty subsets  $M_1, M_2 \subset M$ , respectively, are generally only equal ( $\mathfrak{P}_1 = \mathfrak{P}_2$ ) if  $M_1 = M_2$ .

*Modeling aging mechanisms separately results in a different estimation of transistor parameter degradation than when modeling them simultaneously.* For a given circuit state  $\mathfrak{S}$  and behavior  $\mathfrak{B}$ , transistor parameters can be *abstracted to a failure probability* and, by extension, this probability can be expressed through  $M_i \subset M$ :

$$P_f = P_{\mathfrak{S}, \mathfrak{B}, E, A_{\mathfrak{P}}}(M_i) \quad (2)$$

Analog to the parameters,  $\forall M_i, M_j \subset M; M_{i,j} \neq \{\emptyset\}$ ,  $P_{\mathfrak{S}, \mathfrak{B}, E, A_{\mathfrak{P}}}(M_i) = P_{\mathfrak{S}, \mathfrak{B}, E, A_{\mathfrak{P}}}(M_j) \Leftrightarrow M_i = M_j$ , at time  $t_k > 0$ . A key problem is that it is impossible for system-level designers to conceive how degradation of various  $p_i(t)$  over time will interdepend to ultimately degrade the entire system's reliability, i.e. increasing the  $P_f(Total)$ .

**Our novel contributions within this paper are as follows:**

- (1) We combine, from the physical to system level, the effects of multiple aging mechanisms occurring *simultaneously* based on their interdependencies and show how considering a sole individual mechanism results in a non-negligible reliability underestimation.
- (2) We abstract the various degradations induced by aging (see Fig 1) along with radiation towards a probabilistic fault analysis which has a more meaningful interpretation of reliability, unlike state-of-the-art that employ other quantification metrics e.g.,  $SNM$  and  $\Delta V_{TH}$  which are hard for interpreting the overall reliability degradation of the entire system.

### IV. DEGRADATION MODELING

In this section we illustrate our proposed concept of combining the effects of multiple aging mechanisms showing that the degradation of transistor parameters is a superposition of multiple interdependent aging effects. For instance, while the defects induced by BTI and HCID at a given time (and given transistor parameters) can be considered independent due to their different location in the dielectric (see Fig 2), the overall degradation of  $V_{TH}$  relies on the total number of induced defects. Over time, this will lead to interdependency between BTI and HCID since the amount of defects induced by each is recursively dependent on  $V_{TH}$ . Finally, we present our abstraction of different kinds of failures caused by aging along with our implementation.

<sup>1</sup>Initial values of parameters at  $t = t_0$  may vary from transistor to transistor due to manufacturing variability.

<sup>2</sup>While not important for transistor parameter degradation, radiation plays a role in determining failure probabilities.

Stress induced by running application stimulates multiple aging mechanisms *simultaneously* at the physical level

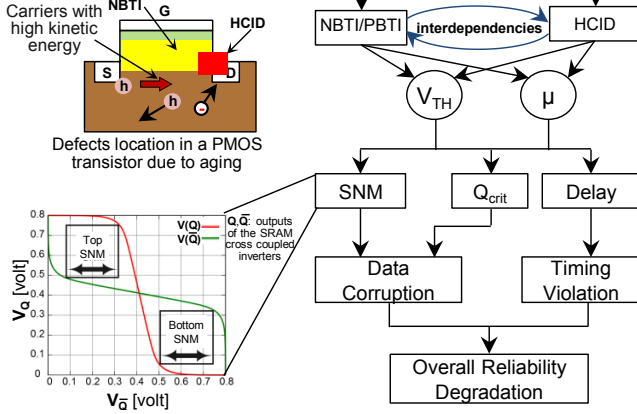


Fig. 2: Our flow for superposing multiple aging effects

### A. Defects due to Aging

BTI is caused by the electric field through the gate dielectric [14], whereas the key source of *hot carriers* is the acceleration of carriers within the electric field in the transistor channel [20]. The following kinds of traps are responsible for the observable defects in a transistor during its lifetime:

**Interface Traps** ( $N_{IT}$ ) are induced by breaking the  $Si-H$  bonds at the  $Si-SiO_2$  interface. One physical  $Si-H$  dissociation mechanism is through the BTI mechanism. Additionally, *interface trap* defects can also be induced by HCID, when  $Si-H$  bonds are dissociated because *hot carriers* deposit their kinetic energy due to Coulomb scattering.

**Oxide Traps** are partially induced by *pre-existing* oxide vacancies in the amorphous  $SiO_2$  of the gate dielectric during manufacturing [19]. When electrically activated, these are *hole traps* ( $N_{HT}$ ). Importantly, the number of *hole traps* is limited to the number of unsatisfied bonds (created by manufacturing). Additional, *oxide traps* ( $N_{OT}$ ) can be induced over time due to the slow and irreversible dissociation of  $Si-O$  bonds [7]. Despite measurements show *oxide traps* during HCID stress, it has been proven that these traps are only induced by the BTI mechanism that *simultaneously* occurs [6].

### B. Superposition of Aging Effects

Over time, the induced defects at the physical level will degrade the following key transistor parameters.

**Threshold Voltage Shift** ( $\Delta V_{TH}$ ): The induced defects result in undesirable charges in the gate dielectric of a transistor weakening the electric field between the gate and bulk. Therefore, the degradation manifests itself as an increase of  $V_{TH}$ .  $N_{IT}$ ,  $N_{HT}$  and  $N_{OT}$  defects will contribute to  $\Delta V_{TH}$  and the role of each one in weakening the electric field depends on the number of present defects. To model this, we modified the analytical solutions of the differential equations [14] that describe trapping mechanisms, with the factor  $d$  to take into consideration the recovery of *interface traps* (that occurs when the voltage stress ceases) based on [5]. Additionally, we introduced HCID on top of BTI via the combination of their simultaneous occurring physical process (i.e.  $N_{IT}$  generation):

$$\Delta V_{TH} = \frac{q}{C_{ox}} \cdot (\Delta N_{IT} + \Delta N_{HT} + \Delta N_{OT}) \quad (3)$$

$$\text{where } \Delta N_{IT} = \Delta N_{IT.BTI} + \Delta N_{IT.HCID} \quad (4)$$

Interface Traps:
$\Delta N_{IT.BTI} = \frac{A(V_{GS} - V_{TH} - \Delta V_{TH})^{\Gamma_{IT}} e^{-\frac{E_{AIT}}{kT}} \cdot t^{n_1} d^{n_1}}{1 + \sqrt{\frac{1-\Lambda}{2}}}; E_{AIT} = \frac{2}{3}(E_{Akr} - E_{Akr}) + \frac{E_{ADH2}}{6}$ $\Delta N_{IT.HCID} = B \left[ t \cdot \frac{I_{DS}}{W} \cdot e^{-\frac{\Phi_{IT,e}}{q\lambda_e E_m}} \right]^{n_2}$ <p>with <math>E_m = \frac{V_{DS} - V_{DSAT}}{\sqrt{\frac{\epsilon_{Si}}{\epsilon_{SiO_2}} \cdot t_{ox} \cdot x_j}}</math> and <math>V_{DSAT} = \frac{(V_{GS} - V_{TH}) \cdot L \cdot E_{cr}}{V_{GS} - V_{TH} + L \cdot E_{cr}}</math></p>
Oxide Traps:
<p><b>Stress Phase:</b></p> $\Delta N_{HT} = C(V_{GS} - V_{TH} - \Delta V_{TH})^{\Gamma_{HT}} \cdot e^{-\frac{E_{AHT}}{kT}} \cdot (1 - e^{-(\frac{t}{\tau})^{\beta_{HT}}})$ $\Delta N_{OT} = D(1 - e^{-(\frac{t}{\tau})^{\beta_{OT}}})$ $n = \eta(V_{GS} - V_{TH} - \Delta V_{TH})^{-\frac{\Gamma_{OT}}{\beta_{OT}}} \cdot e^{\frac{E_{AOT}}{kT\beta_{OT}}}$ <p><b>Recovery Phase:</b></p> $\Delta N_{HT} + \Delta N_{OT} = E(e^{-(\frac{t}{\tau_r})^{\beta_{HTR}}}) \quad \text{details in Section VII}$

As this is not our main scope, in-depth explanation of the employed physical aging models can be found in [5], [14], [16]. It is worthy to note that BTI-induced defects are evenly distributed across the  $Si-SiO_2$  interface, where the electric field is homogeneous, contrary to the HCID-induced defects which are concentrated near the drain (see Fig 2).

**Carrier Mobility ( $\mu$ ) Degradation:** The induced defects negatively impact the mobility of carriers within the transistor channel, as the charged defects can interact with the carriers impeding their passage through the channel leading to degrading the transistor's carrier mobility. Because *hole/oxide traps* ( $N_{HT}/N_{OT}$ ) are located deep within the gate dielectric away from the channel, they have a negligible impact on the carrier mobility and, thus, only *interface traps* have to be considered here. Similarly as [16], we model the  $\mu$  degradation as follows:

$$\mu = \frac{\mu_0}{1 + \alpha \cdot \Delta N_{IT}} \quad (5)$$

Therefore, the degradation of transistor parameters is due to both BTI and HCID and the overall aging is a superposition of their interdependent effects. State-of-the-art approaches [29], [22], [28] look at  $V_{TH}$  solely when analyzing reliability. We show later in Section V (see Fig 9) why it is necessary to additionally take  $\mu$  into account to avoid underestimating aging-induced degradations.

### C. Reliability Abstraction

In the following, we present how we deal with the aging-induced degradations occurring at the device/circuit level to provide a *reliability abstraction* summarizing their impacts on the susceptibility to failures at the system level (see Fig 2). As an example, we apply our reliability abstraction to SRAMs<sup>3</sup> due to their susceptibility to different reliability aspects (e.g., noise, radiation, etc.) and due to their total chip area that may reach up to 70% [23].

**Data Corruption:** Data in SRAM cells can be corrupted because of the voltage noise from neighboring circuits transferred over parasitic capacitances, supply voltage, etc. The *Static Noise Margin (SNM)* quantifies the resiliency of the SRAM against noise. The *SNM butterfly* curve that describes the transfer characteristics of the cross coupled inverters within the SRAM is shown in Fig 2. Aging-induced  $V_{TH}$  degradation will shift the *butterfly* curve shrinking the size

<sup>3</sup>Our proposed work is not restricted to a specific kind of circuits and can also be applied to others such as computational units.

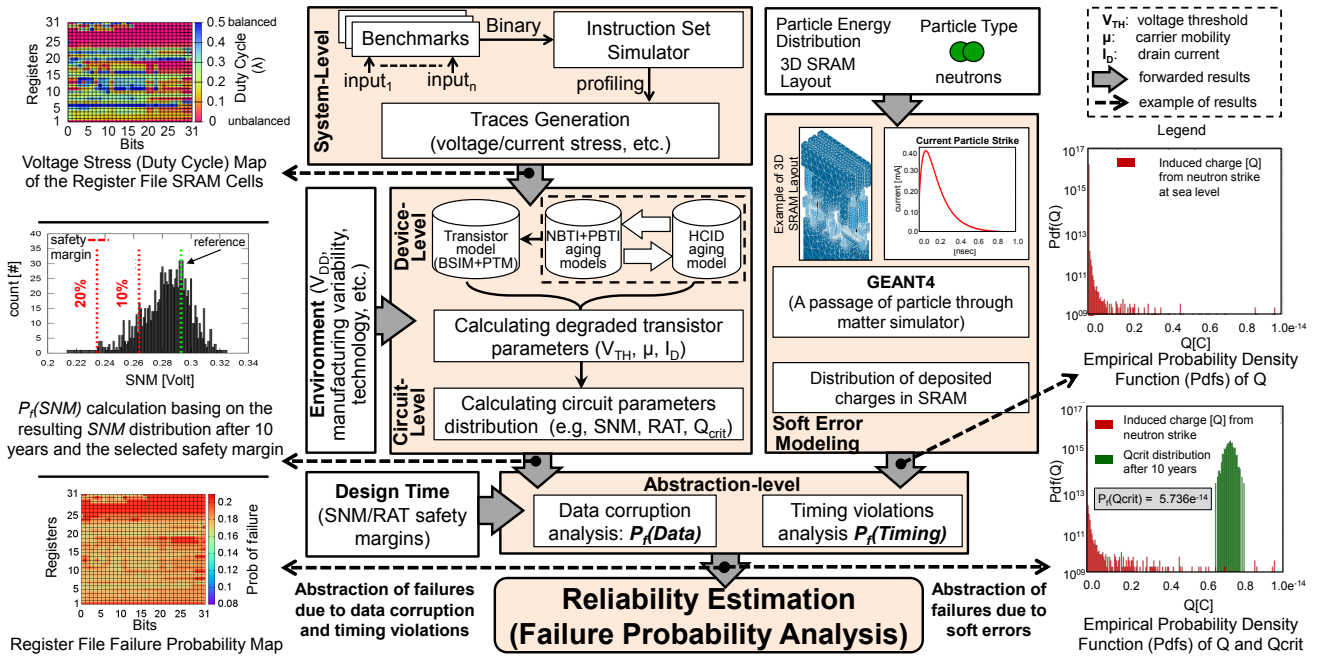


Fig. 3: Our proposed in-house implementation to bridge the gap between application-induced stress at the system level and induced defects at the physical level together with abstracting the corresponding induced failures to estimate the overall reliability

of the square within and, thus, degrading the  $SNM$  of the aged SRAM. Indeed,  $SNM$  degradation reduces the SRAM resiliency resulting in an increased failure probability due to data corruption ( $P_f(Data)$ ). Additionally, radiation can also corrupt the SRAM data when a particle deposits its charges through an SRAM resulting in an electrical current spike (see Fig 3). The *transconductance* ( $g_m$ ) of a transistor determines if the generated spike will induce charges above the *Critical Charge* ( $Q_{crit}$ ) (i.e. the minimum amount of charge required to flip/corrupt stored data). Both shifts in  $V_{TH}$  and  $\mu$  degrade  $g_m$  and, therefore, aging increases the probability of failure due to soft errors ( $P_f(Q_{crit})$ ).

**Timing Violations:** An SRAM in a (synchronous) design can cause timing violations i.e. it fails to provide correct data in time. The capability of the SRAM (for a given sense amplifier) to drive its bitlines within timing constraints depends on the  $I_D$  of the SRAM transistors which is reduced by the degradation of  $\Delta V_{TH}$  and  $\mu$ . In other words, aging will result in a longer *read access time* ( $RAT$ ) in the SRAM<sup>4</sup> increasing the failures due to timing violations ( $P_f(Timing)$ ). Fig 2 summarizes how aging increases the system's susceptibility to failures.

**Implementing our reliability abstraction:** Fig 3 illustrates our implementation to abstract the impact of multiple simultaneous aging mechanisms on the probability of failures. As shown, the model of NBTI/PBTI together with HCID is employed to degrade the affected transistor parameters based on the Predictive Technology Model [25] and the BSIM4 [4] that is utilized to address the interrelation between these parameters (e.g., how  $\Delta V_{TH}$  &  $\mu$  influence  $I_D$  &  $g_m$  etc.). Then, device-level parameters and the corresponding circuit-level metrics (i.e.  $SNM$ ,  $RAT$ , and  $Q_{crit}$ ) are computed. There is a variation in device/circuit-level metrics due to manufacturing variability, resulting in varying susceptibilities to failures. The developed manufacturing variability modeling was provided from semiconductor industry and corresponds

to a normal distribution for the transistor dimensions. To calculate charges deposited in an SRAM by natural neutron radiation, we employed the *Geant4* simulator [3]. The obtained deposited charges distribution in conjunction with the  $Q_{crit}$  distribution due to aging after the targeted lifetime (e.g., 10 years) is then used to compute the failure probability due to soft error  $P_f(Q_{crit})$  as shown in the right side of Fig 3 (further details in Fig 8(a)). When analyzing failure probabilities due to noise or timing, it is important to consider the employed safety margins. These are chosen by the system designer to allow for variability in the design (either at the beginning due to manufacturing or later on due to aging degradations). If the resulting  $SNM/RAT$  degradation exceeds the corresponding noise/timing safety margin, failures may occur more frequently. We call the probability of these failures by  $P_f(SNM)/P_f(RAT)$ , respectively. Fig 3 (left) presents an example in terms of data corruption failures and how the selected safety margin of  $SNM$  can be applied to the aging-induced  $SNM$  distribution to calculate  $P_f(SNM)$  after the targeted lifetime. Similarly,  $P_f(RAT)$  can be obtained. Then, the total failure probability is expressed as:

$$P_f(Total) = \underbrace{P_f(Data)}_{P_f(SNM) + P_f(Q_{crit})} + \underbrace{P_f(Timing)}_{P_f(RAT)} - \underbrace{P_f(SNM \cap Q_{crit}) - P_f(SNM \cap RAT) - P_f(RAT \cap Q_{crit})}_{+ P_f(SNM \cap Q_{crit} \cap RAT)}$$

Finally, the aforementioned steps are analogously repeated for additional inputs (e.g., temperatures, voltage stresses,  $V_{DD}$ , etc.) to build a database that provides designers with fast lookup-based reliability estimation at the system level through a wide range of operating conditions.

## V. EVALUATION

In the following, we first show a validation against industrial measurements, then we illustrate the impact of combining multiple simultaneous aging mechanisms. Finally, comparisons

<sup>4</sup>Unlike write access time which is improved by aging [35].



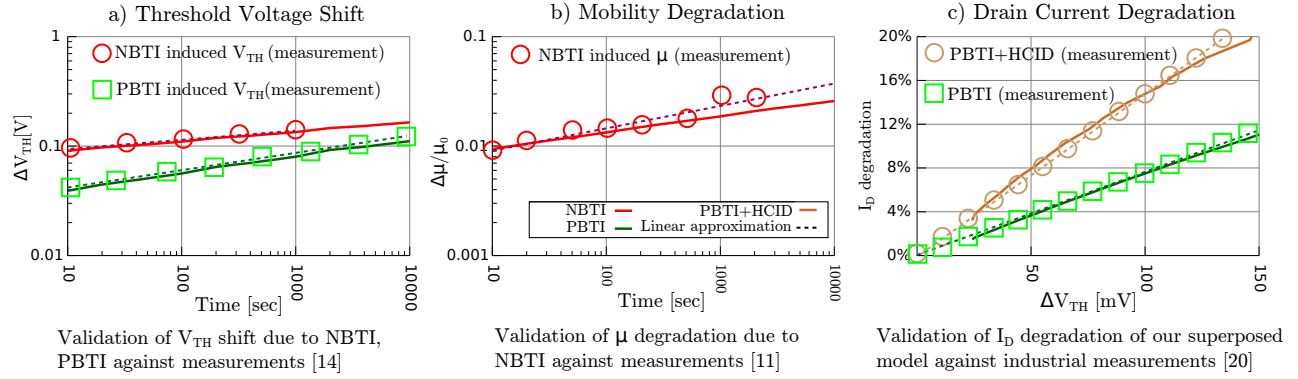


Fig. 4: Validation of our superposed model of combining BTI and HCID aging mechanisms

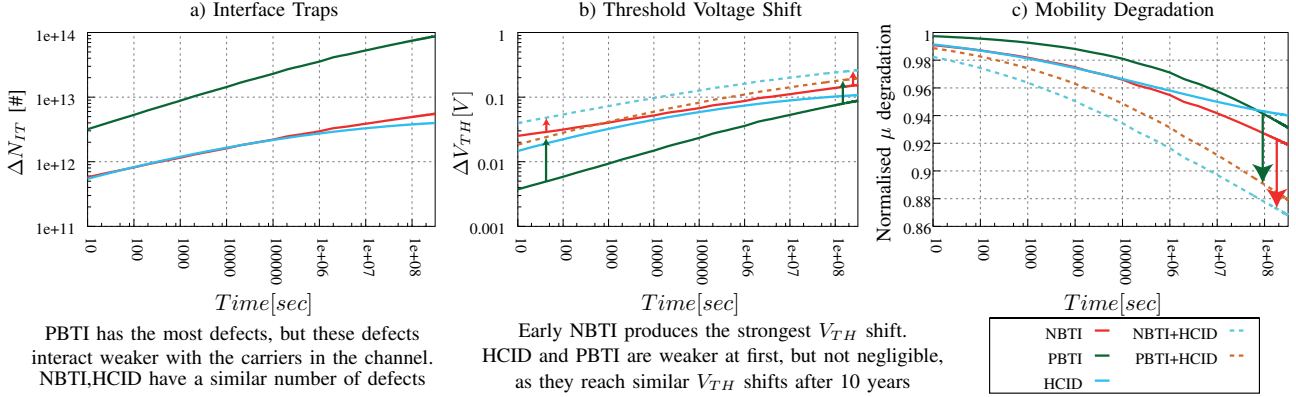


Fig. 5: Transistor degradations due to BTI and HCID aging mechanisms *separately / simultaneously* considering.

to state-of-the-art under different scenarios are presented<sup>5</sup> to demonstrate the reliability underestimation that comes from not taking multiple simultaneous aging mechanisms into account and/or not considering the different kinds of failures.

**Validation:** Comparisons of the impact of BTI-induced *interface traps* on  $V_{TH}$  and  $\mu$  against measurement data obtained from [14], [11], respectively, are presented in Fig 4 (a, b). Fig 4(a) shows a good agreement with the measurements and in Fig 4(b) a slight mismatch. However, the comparison in Fig 4(c) of the  $I_d$  degradation (which covers both  $V_{TH}$  as well as  $\mu$  parameters) against measurements conducted by *STMicroelectronics* [20] shows that our superposition model deviates by around 6% from the measurements<sup>6</sup>. We selected the latter to validate our results as it is the only available industrial experiment for current nano-scale high-k technology nodes that measures multiple simultaneous aging phenomena.

**Device/Circuit-level Evaluation:** Fig 5 presents the corresponding  $V_{TH}$  and  $\mu$  degradations with respect to the induced defects under different cases. As shown, the negative impact of PBTI on the transistor is quite small (but not negligible) in comparison to other aging mechanisms, even though the number of induced defects is higher (see Fig 5(a)). This is because these defects interact weaker with the carriers in the channel. As shown in Fig 5(b) NBTI initially shifts  $V_{TH}$  more than HCID because of the *pre-existing oxide traps* which come from manufacturing and, additionally, inducing

*oxide traps* as well as *interface traps* (see Section IV-A). Because the electrical activation of *hole traps* saturates over time [14], the BTI-induced  $\Delta V_{TH}$  is dominated by *interface traps* in the long term. On the other hand,  $\mu$  degradation only depends on charges in the proximity of the transistor's channel (i.e. *interface traps*). Thus, BTI-induced *oxide traps* play a weaker role here. Therefore, both BTI and HCID have a similar  $\mu$  degradation over time (see Fig 5(c)). Arrows in Fig 5 indicate the shifts (up to 6%) due to multiple simultaneous aging effects. It is worthy to note that the impact of simultaneous aging effects cannot be fully grasped at this abstraction level as motivated in section II and thus further evaluation at the system level is done later in Figs (9, 10, 11). The impact of aforementioned device-level degradations on reducing the circuit reliability in terms of  $SNM$ ,  $Q_{crit}$  and  $RAT$  for both *duty cycle* (i.e. voltage stress  $\lambda$ )<sup>7</sup>, and temperature cases are reported in Fig 6 (a, b)<sup>8, 9</sup>. As observed, the circuit-level metrics exhibit a similar degradation in the temperature case (see Fig 6(b)) but they behave either similarly or contradictively with respect to  $\lambda$  (see Fig 6(a)). For instance, balancing the voltage stress mitigates  $SNM$  degradations (even more than the  $Q_{crit}$  degradations), whereas, it worsen  $RAT$  degradations. This is because both  $SNM$  and  $Q_{crit}$  are dominated by the most degraded transistors, and, thus, stress-balancing mitigates their degradations. In contrast,  $RAT$  is determined by the

<sup>7</sup>The voltage stress will be uniformly distributed among the cross coupled inverter transistors when  $\lambda = 0.5$  [22]

<sup>8</sup>The high degradation is due to the high  $V_{DD}$  (1.85V) which is used to amplify the tendencies for clarity. However, the observations still apply at lower  $V_{DD}$ .

<sup>9</sup>The accuracy of  $SNM$ ,  $RAT$  and  $Q_{crit}$  is 1mV,  $10^{-24}$ s and 0.01fC, respectively, based on settings within our estimation e.g. SPICE stepping size.

<sup>5</sup>22nm,  $V_{DD} = 1.0$  V,  $125^\circ C$ , and 10 years lifetime have been targeted.

<sup>6</sup> $V_{DD}$  in measurements (1.85 V) is typically used to accelerate observing aging effects (i.e. hours) rather than months/years in the normal operation condition (e.g.,  $V_{DD} = 1.0$  V).

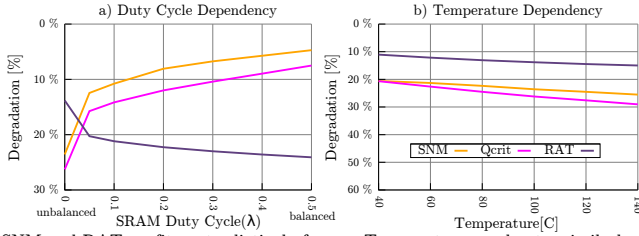
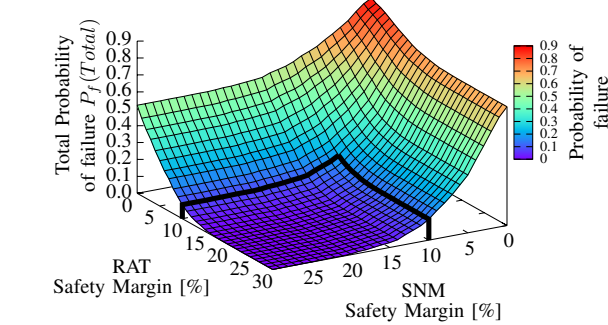
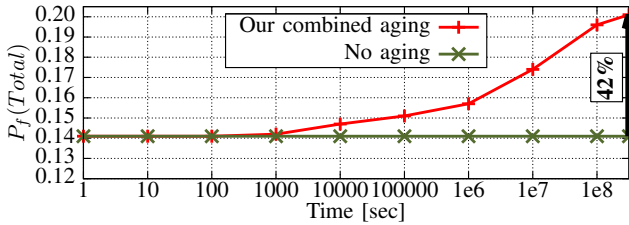


Fig. 6: Behavior of aging-induced degradations at the circuit level

Failures due to data corruption and timing violations, when multiple simultaneous aging mechanisms are considered



(a)  $P_f(Total)$  analysis under different safety margins



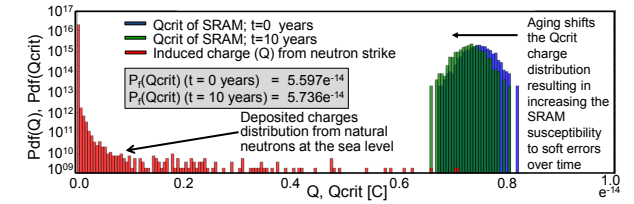
(b) Impact of aging in the presence of manufacturing variability

Fig. 7: Failure analysis from our proposed implementation

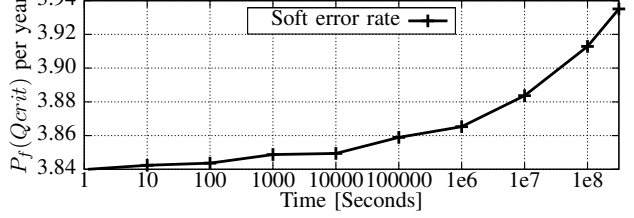
least degraded transistors which decide how long the SRAM requires to charge its bitlines. As soon as the least degraded transistors pull its bitline above the sensing threshold, the SRAM sense amplifier will drive both bitlines.

As a result, stress-balancing aging-aware mitigation techniques (e.g., [22], [28]) can reduce the susceptibility to only one particular kind of failures i.e.  $P_f(Data)$ .

**Analysis of Failure Probability:** Fig 7(a) shows, for the case of  $\lambda = 0$ , the increase of  $P_f(Total)$  depending on the chosen safety margins that determine if the induced degradations can be tolerated or not (see Section IV-C). As shown,  $P_f(Total)$  exponentially decreases with higher safety margins which, in turn, directly influences the device's cost. For instance, a higher  $SNM$  safety margin to cope with aging-induced  $SNM$  degradation necessitates building more robust SRAM sense amplifiers, which can negatively affect the area/power budget, and higher  $RAT$  safety margin leads to selling the device at lower frequency to avoid aging-induced timing violations during its lifetime. Therefore, such an analysis in Fig 7(a) can guide the designer to choose Pareto-optimal safety margins that maintain a reliable operation in the presence of aging. Fig 7(b) clarifies, for the case of 10% safety margins, that multiple simultaneous aging mechanisms can increase  $P_f(Total)$  up to 42% over 10 years on top of the failures due to manufacturing variability.



(a) Empirical probability density functions (PDFs)



(b) Effect of aging on the rate of radiation induced-soft errors

Fig. 8: Soft error analysis in the presence of aging

In Fig 8(a), we show the resulting degradation in  $Q_{crit}$  along with the distribution of electrical charge deposited when a high-energy particle strikes the device. To obtain the latter, we employed from [2] a realistic distribution of energy and flux of neutrons at the sea level together with SRAM layers information based on [36]. These distributions can be used to derive the probabilities of soft error during the device's lifetime in the presence of aging. Due to just a small number of charges generated above the  $Q_{crit}$ , the failure probability is very low. This is mainly due to analyzing soft error under typical operation conditions i.e. neutrons at the sea level<sup>10</sup> together with  $V_{DD} = 1.0$  V that results in low  $Q_{crit}$  shifts and, therefore, just a small change in soft error sensitivity. In Fig 8(b), these probabilities are combined with the neutron flux to compute the expected number of soft errors per year. As observed, aging can increase soft error rate by 2.4%.

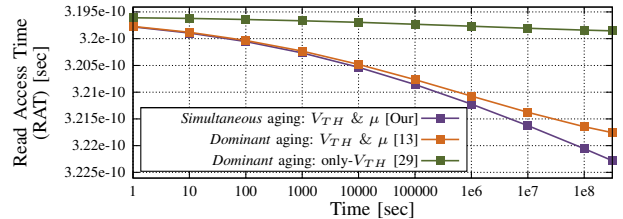


Fig. 9: Comparison between our proposed combination of multiple simultaneous aging mechanisms and state-of-the-art

**Comparison to state-of-the-art:** For a fair comparison, we select the work [13] as it follows similar goals. Additionally, other state-of-the-art often employ its concept (e.g., [29]) to estimate reliability when multiple aging mechanisms are targeted. As explained in Section II, [13], [29] mainly consider the *dominant* aging mechanism in each transistor of the studied circuit. Fig 9 presents a  $RAT$  degradation<sup>11</sup> of an SRAM at  $\lambda = 0.5$ . It shows a noticeable deviation over time compared to our proposed *simultaneous* aging combination that ignore neither PBTI in NMOS nor HCID in PMOS. This establishes why it is vital to not rely only on the dominant aging mechanism when reliability analysis is performed to avoid underestimation. Additionally, we show how solely considering

<sup>10</sup>Analyzing soft error due to other kinds of particles or at higher altitudes, where higher fluxes are available, can result in higher  $P_f(Q_{crit})$ .

<sup>11</sup>Delay analysis has been chosen here for consistency's sake with [13]

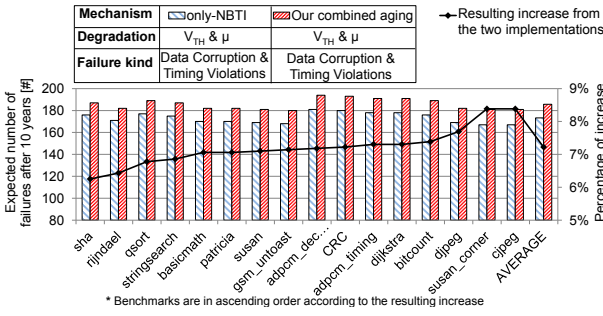


Fig. 10: Register file reliability estimation comparison showing the underestimation when only an individual aging mechanism is considered even through examining both kinds of failures

$V_{TH}$  in analysis [29], [22], [28] can significantly underestimate degradation. Therefore, examining  $\mu$  together with  $V_{TH}$  as our implementation does (see Section IV-B), is indeed essential.

**System-level Evaluation:** Register files in microprocessors are typically implemented using SRAMs and they are particularly susceptible to aging due to continuous voltage stress for prolonged intervals [22], [28]. They also have the highest average temperature [26] and as shown in Fig 6 elevated temperatures accelerate aging. Therefore, we selected the register file in our system-level evaluations. A 32-bit MIPS model<sup>12</sup> simulator and the Mibench benchmark suite [27] has been used to explore different stress scenarios in the register file. We assume in our experiments safety margins of 10% for both  $SNM$  and  $RAT$  which represent a good compromise as shown in Fig 7(a).

Fig 10 illustrates the discrepancy that arises from considering solely an individual, instead of multiple simultaneous, aging mechanisms even through examining both kinds of failures. As it can be observed, considering NBTI as the most dominant mechanism and, thus, ignoring PBTI and HCID results in an underestimation of 7%, on average. Importantly, we present in Fig 11 the serious impact of considering NBTI as an individual dominant aging mechanism together with looking at only failures due to  $SNM$  degradation as state-of-the-art techniques (e.g., [22], [28]) do when they estimate the reliability of register files. In such a case, the underestimation reaches, on average, 75% and up to 85%. An example of the register file failure map, obtained from our in-house reliability estimation (see Fig 3), is presented in Fig 12(a) for the *patricia* benchmark as well as the corresponding failure probability distribution in Fig 12(b). The samples with the higher probability of failures in Fig 12(b) correspond to a group of SRAM cells which suffer more from aging (also seen in the stress map of the register files SRAM cells, for the same benchmark, in Fig 3(left)).

*All in all, the above analysis can be used by the designer to obtain an abstracted, yet accurate reliability estimation of how defects, induced by multiple simultaneous aging mechanisms, at the physical level can ultimately increase the probability of failure at the system level. Moreover, it can be used to find a compromise between cost of the chip and its reliability by exploring different design choices (i.e. safety margins).*

**Limitations:** Attaining the reliability abstraction in terms of probability of failure through the modeling of physical defects comes at higher computational time compared to concepts discussed in II where the interdependencies of aging mechanisms are not taken into account. Having a meaningful interpretation of reliability at the system level together with examining

<sup>12</sup>Its register file consists of 32 registers. However, our reliability estimation can easily be performed for other architectures

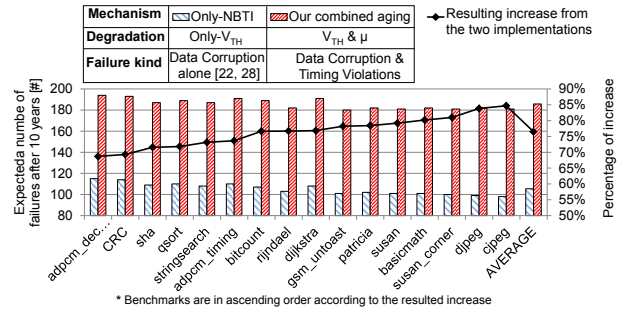


Fig. 11: Register file reliability estimation comparison showing the underestimation when only an individual aging mechanism is considered together with examining only a single kind of failures

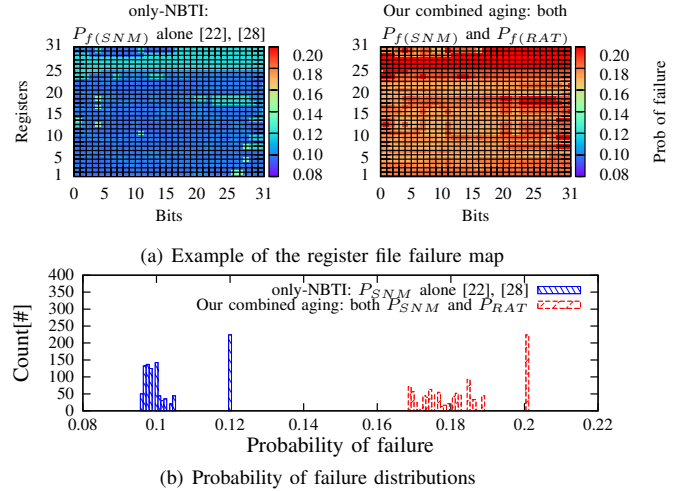


Fig. 12: Failure analysis of the register file SRAM cells

different kinds of failures compensate for this, as long the underlying transistor parameters (e.g., dimensions, manufacturing variability, etc.) do not exhibit a wide variance across the design, which would require multiple abstractions. Likewise, the approach benefits from regularity in the design (i.e. the reuse of circuits). Our implementation is currently limited to the most three dominant aging mechanisms (NBTI, PBTI and HCID). The interdependencies with, e.g. TDDB, which is less pronounced, as motivated in Section I, necessitates following our process presented in Section IV-B and Fig 2 analogously – especially measurements established that both BTI and TDDB share *oxide trap* defects [33]. However, there is an additional challenge here due to the geometrical connection of TDDB and BTI as stated in [34].

## VI. CONCLUSION

Abstracting degradations incurred by multiple aging mechanisms is a key challenge when estimating reliability at the system level. We presented one step by combining the simultaneous effects of multiple aging mechanisms examining their interdependencies from the physical level, in order to derive a probability of failure as a meaningful reliability abstraction, yet accurate for system-level designers enabling them to compromise between cost of the chip and its reliability and to additionally expose the most susceptible parts to aging in their systems. Through the abstraction via circuit-level metrics towards probability of failure, it becomes evident that targeting only one kind of failure tied to one aging mechanism results in a significant reliability underestimation.

## VII. GLOSSARY

$\lambda$	SRAM duty cycle	$t$	current time
$\Lambda$	transistor duty cycle	$V_{DSAT}$	potential at channel pinchoff
$\lambda_e$	mean free path of hot electrons	$x_j$	source/drain junction depth
$\Phi_{it,e}$	critical electron energy interface trap creation	$t_{ox}$	oxide thickness
$\tau$	stress time	$\epsilon_{Si}$	relative permittivity of $Si$ and $SiO_2$
$\tau_r$	recovery time	$\epsilon_{SiO_2}$	
$C_{ox}$	oxide capacitance	$n_1$	BTI/HCID time exponent
$E_m$	channel electric field	$n_2$	
$E_{AHT}$	activation energy of interface/hole/oxide traps	$ABC$	fitting parameters
$E_{AHT}$		$DE$	
$E_{AOT}$		$\alpha\eta$	
$E_{ADH2}$	activation energy of $H_2$ diffusion	$\Gamma_{IT}$	field acceleration factors
$E_{AKf}$	activation energy bond dissociation/ bond annealing	$\Gamma_{HT}$	
$E_{AKr}$		$\Gamma_{OT}$	
$E_{cr}$	critical field for velocity saturation	$\beta_{OT}$	hole trapping parameter
$N_{IT}$	number of interface/hole/oxide traps	$\beta_{HT}$	
$N_{HT}$		$\beta_{HTR}$	hole trap recovery parameter
$N_{OT}$		$\mu_0$	carrier mobility at $t=0$

## REFERENCES

- http://www.itrs.net/.
- http://www.jedec.org/sites/default/files/docs/jesd89a.pdf.
- S. Agostinelli et al. GEANT4a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003.
- Y. Singh Chauhan et al. BSIM compact MOSFET models for SPICE simulation. *Mixed Design of Integrated Circuits and Systems, International Conference*, pp. 23–28, 2013.
- M. Alam, K. Roy, and C. Augustine. Reliability- and process-variation aware design of integrated circuits. *Reliability Physics Symposium (IRPS), IEEE International*, pp. 1–11, 2011.
- E. Amat et al. A comprehensive study of channel hot-carrier degradation in short channel MOSFETs with high-k dielectrics. *Microelectron. Eng.*, pp. 144–149, 2013.
- N. Anderson et al. First-Principles Investigation of Low Energy E' Center Precursors in Amorphous Silica. *Phys. Rev. Lett.*, 2011.
- M. Bagatin et al. Impact of NBTI Aging on the Single-Event Upset of SRAM Cells. *Nuclear Science, IEEE Transactions on*, pp. 3245–3250, 2010.
- J. Bernstein et al. Electronic circuit reliability modeling. *Microelectronics Reliability*, pp. 1957–1979, 2006.
- G. Bersuker et al. Breakdown in the metal/high-k gate stack: Identifying the "weak link" in the multilayer dielectric. *Electron Devices Meeting, IEEE International*, pp. 1–4, 2008.
- A. Chaudhary and S. Mahapatra. A Physical and SPICE Mobility Degradation Analysis for NBTI. *Electron Devices, IEEE Transactions on*, pp. 2096–2103, 2013.
- V. Huard et al. CMOS device design-in reliability approach in advanced nodes. *Reliability Physics Symposium, IEEE International*, pp. 624–633, 2009.
- Wenping Wang et al. Compact Modeling and Simulation of Circuit Reliability for 65-nm CMOS Technology. *Device and Materials Reliability, IEEE Transactions on*, pp. 509–517, 2007.
- K. Joshi et al. A consistent physical framework for N and P BTI in HKMG MOSFETs. In *Reliability Physics Symposium (IRPS), IEEE International*, pp. 5A.3.1–5A.3.10, 2012.
- J. Keane, et al. An All-In-One Silicon Odometer for Separately Monitoring HCI, BTI, and TDDB. *Solid-State Circuits, IEEE Journal of*, pp. 817–829, 2010.
- X. Li, J. Qin, and J. Bernstein. Compact Modeling of MOSFET Wearout Mechanisms for Circuit-Reliability Simulation. *Device and Materials Reliability, IEEE Transactions on*, pp. 98–121, 2008.
- Z. Liu et al. Design tools for reliability analysis. *Design Automation Conference*, pp. 182–187, 2006.
- Abirshami, H. et al. Multi-corner, energy-delay optimized, NBTI-aware flip-flop design Quality Electronic Design, 11th International Symposium on pp.. 652–659, 2010.
- P. Magnone et al. 1/f Noise in Drain and Gate Current of MOSFETs With High- k Gate Stacks. *Device and Materials Reliability, IEEE Transactions on*, pp. 180–189, 2009.
- Y. Randriamihaja et al. Microscopic scale characterization and modeling of transistor degradation under {HC} stress. *Microelectronics Reliability*, 52(11):2513–2520, 2012.
- R. Tu et al. Berkeley reliability tools-BERT. *Trans. Comp.-Aided Des. Integr. Cir. Sys.*, pp. 1524–1534, 2006.
- H. Amrouch et al. Stress balancing to mitigate NBTI effects in register files. *Dependable Systems and Networks, 43rd Annual IEEE/IFIP International Conference on*, pp. 1–10, 2013.
- H. Yamauchi. A Discussion on SRAM Circuit Design Trend in Deeper Nanometer-Scale Technologies. *Very Large Scale Integration Systems, IEEE Transactions on*, 18(5):763–774, 2010.
- S. Zafar et al. Threshold voltage instabilities in high-k; gate dielectric stacks. *Device and Materials Reliability, IEEE Transactions on*, pp. 45–64, 2005.
- W. Zhao and Y. Cao. Predictive technology model for nano-CMOS design exploration. *J. Emerg. Technol. Comput. Syst.*, 2007.
- I. K. et al, Temperature aware floorplanning. Temperature-Aware Computer Systems, 2005.
- M. R. Guthaus et al. Mibench: A free, commercially representative embedded benchmark suite *Proceedings of the Workload Characterization, IEEE International Workshop*, pp. 3–14, 2001.
- S. Kothawade and K. Chakraborty. Analysis and mitigation of BTI aging in register file: An application driven approach *Microelectronics Reliability*, vol.53, no.1, pp. 105–113, 2013.
- Oboril, F. et al. ExtraTime: Modeling and analysis of wearout due to transistor aging at microarchitecture-level Dependable Systems and Networks, 42nd Annual IEEE/IFIP International Conference on pp. 1–12, 2012.
- Xiaojun Li, H. et al. A new SPICE reliability simulation method for deep submicrometer CMOS VLSI circuits *Device and Materials Reliability, IEEE Transactions on* vol.6, no.2, pp. 247–257, 2006.
- Ayse K. Coskun et al. Evaluating the impact of job scheduling and power management on processor lifetime for chip multiprocessors *International Joint Conference on Measurement and Modeling of Computer Systems* pp. 169–180, 2009.
- J. Srinivasan et al. RAMP: A model for reliability aware microprocessor design IBM Research Division, Yorktown Heights, NY, IBM Research Rep. 2003.
- Jiaqi Yang et al. Intrinsic correlation between PBTI and TDDB degradations in nMOS HK/MG dielectrics *Reliability Physics Symposium, IEEE International* pp. 5D.4.1,5D.4.7, 2012.
- Tous, Santi et al. A compact analytic model for the breakdown distribution of gate stack dielectrics *Reliability Physics Symposium (IRPS), IEEE International* pp. 792–798, 2010.
- Tae-Hyoung Kim T. and Hui Kong Z. Impact Analysis of NBTI/PBTI on SRAM  $V_{MIN}$  and Design Techniques for Improved SRAM  $V_{MIN}$  *Journal of Semiconductor Technology and Science* vol.13, no.2, pp. 87–97, 2013.
- K. Mistry et al. A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging *Electron Devices Meeting, IEEE International* pp. 247–250, 2007.