Circuit reliability prediction: challenges and solutions for the device time-dependent variability characterization roadblock

M. Nafria, J. Diaz-Fortuny, P. Saraza-Canflanca, J. Martin-Martinez, E. Roca, R. Castro-Lopez, R. Rodriguez, P. Martin-Lloret, A. Toro-Frias, D. Mateo, E. Barajas, X. Aragones, F. V. Fernandez

Abstract— **The characterization of the MOSFET Time-Dependent Variability (TDV) can be a showstopper for reliability-aware circuit design in advanced CMOS nodes. In this work, a complete MOSFET characterization flow is presented, in the context of a physics-based TDV compact model, that addresses the main TDV characterization challenges for accurate circuit reliability prediction at design time. The pillars of this approach are described and illustrated through examples.**

Keywords—Time-dependent variability, CMOS technology, MOSFET, RTN, aging, BTI, HCI degradation, characterization

I. INTRODUCTION

A Reliability-Aware Design (RAD) approach must be adopted to sustain system reliability in ultrascaled CMOS technology nodes. In there, reliability, together with performance, has to be optimized during the early circuit/system design stage [1, 2]. Circuit reliability is strongly affected by the time-dependent variability (TDV) of the underlying devices. TDV is observed as shifts of the device electrical characteristics (measured, for example, as variations of the threshold voltage), which, together with the process-related shifts (i.e. Time-Zero Variability, TZV), will negatively impact circuit performance and may eventually end in the circuit failure [3, 4]. So, to bridge the gap between the device and circuit levels, compact models that accurately describe the device TDV must be developed and implemented into circuit simulation tools [5, 6]. Actually, several CAD commercial tools, such as RelXpert [7] or MOSRA [8], already evaluate the IC reliability, although they do not fully address all the challenges [9].

In deeply-scaled MOSFETs, TDV has been linked to a handful of defects in the device [10], which introduce a stochastic operation-dependent uncertainty in the electrical performance of devices. Then, for the development of the models and the extraction of the corresponding parameters, thousands of devices must be characterized under different operation conditions (voltages, temperature, time), to obtain statistically meaningful information. This poses serious challenges to the development of sufficiently accurate compact models, since, on the one hand, the experimental time may become unfeasible and, on the other, model parameter extraction becomes extremely intricate, since thousands of complex traces must be analyzed to obtain the relevant data.

In this work, a complete characterization flow of the TDV of MOSFETs in advanced nodes will be presented, in the context of a physics-based TDV compact model. Key in this approach is an array-based IC, whose architecture allows an optimized fully-automated stress-measurement procedure, to reduce the testing time. Smart data analysis algorithms have been also developed, to quickly and accurately obtain the device TDV model parameters. Finally, to complete the RAD flow, a reliability circuit simulation tool is presented.

II. TDV PHENOMENOLOGY AND MODELLING

TDV is observed on top of TZV and comprises transient phenomena, such as Random Telegraph Noise (RTN) [11], and aging mechanisms, such as Bias Temperature Instability (BTI) and channel Hot-Carrier Injection (HCI) degradation [12, 13]. TDV effects on the electrical performance of MOSFETs are usually measured as changes in their threshold voltage (ΔV_{th}), although other parameters (mobility, subthreshold current…) may be also affected. RTN is observed as sudden and random discrete shifts in the threshold voltage of the device and, equivalently, in the drain current of the transistor (Fig. 1(a)). Bias Temperature Instability aging is gate-voltage and temperature activated, whereas Hot-Carrier Injection degradation is also drain-voltage dependent. The associated ΔV_{th} can be decomposed into permanent and recoverable components. Because of the last property, the threshold voltage can be partially recovered when the bias decreases (Fig. 1b). Moreover, in ultrascaled devices, ΔV_{th} becomes stochastic (Fig. 1.b and 1.c).

In advanced nodes, RTN effects can be relevant at short operation times under nominal operation conditions, but aging observation requires very long times at the nominal biasing. Therefore, accelerated tests (i.e., temperatures and/or voltages higher than the nominal) are used to induce visible device degradation in reasonable experimental times. Typical aging tests involve on-wafer measurement-stress-measurement

This work was supported by the VIGILANT Project (PID2019-103869RB / AEI / 10.13039/501100011033) and the TEC2016-75151-C3-R Project (AEI/FEDER, UE).

M. Nafria, J. Diaz-Fortuny, J. Martin-Martinez and R. Rodriguez are with Dep. Electronic Engineering. Universitat Autonoma Barcelona. 08193 Bellaterra. Spain (email: montse.nafria@uab.es).

P. Saraza-Canflanca, E. Roca, R. Castro-Lopez, P. Martin-Lloret, A. Toro-Frias and F. V. Fernandez are with Instituto de Microelectronica de Sevilla, IMSE-CNM (CSIC/Universidad de Sevilla), 41092 Seville, Spain.

D. Mateo, E. Barajas and X. Aragones are with Dep. Electronic Engineering, Universitat Politecnica de Catalunya, 08034 Barcelona, Spain.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

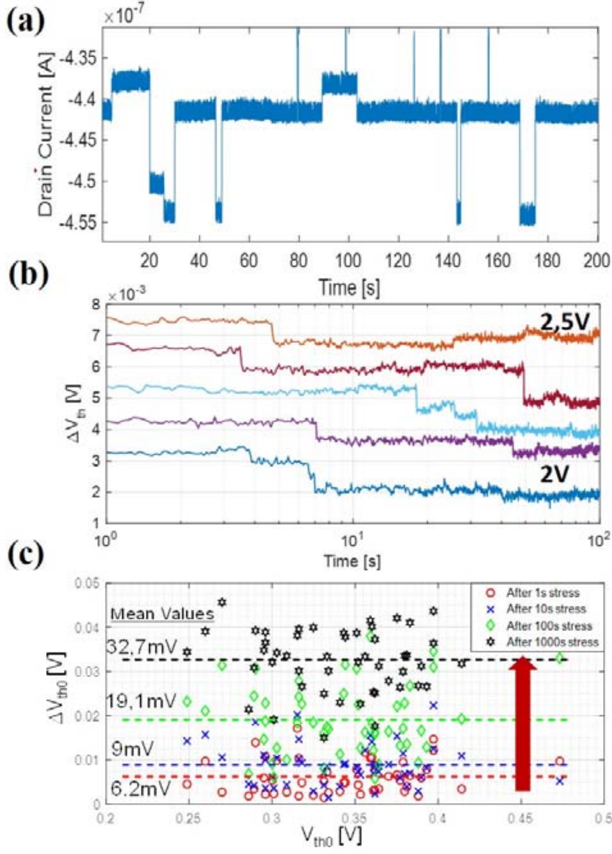

Figure 1. Example of a RTN current trace measured in a 80x60nm^2 pMOS transistor at V_G=0.4V and V_D=0.1V. More than two current levels are observed, pointing to the presence of more than one defect in the device. (b) Recovery trace measured on a 1000x60nm^2 pMOS at V_G=0.4V and V_S=0.1V after 100s of BTI stress at different gate stress voltages. The V_{th} shift increases with the stress voltage. In such small devices, sudden charge emission events from single defects can be observed. (c) Threshold voltage shifts observed in a set of 80x60nm^2 pMOS measured at V_G=0.4V, V_D=0.1V after HCI stress at V_G=V_D=2V after different stress times. The shifts increase with the stress time. In these small devices, ΔV_{th} becomes stochastic.

(MSM) sequences, where stress stages (of different durations) are followed by sensing at lower voltages to measure the impact of the previous stress [14].

Though the physical mechanisms behind aging are still under debate [15, 16], nowadays there is a large consensus on its link with the capture/emission of charges in/from defects in the device, which lead to the observed threshold voltage shifts. Actually, it is widely accepted that the same kind of defects is responsible of RTN and BTI, though RTN or BTI can dominate depending on biasing [17, 18].

One of the most accepted pictures for the description of RTN and BTI aging is provided by defect-centric models, which are able to account for the observed stochasticity in the threshold voltage shifts [19]. In these models, ΔV_{th} depends on the number of defects in the transistor, N, and their major parameters, namely, the amplitude of the threshold voltage shift associated to each defect (η) and their capture (τ_c)/emission (τ_e) time constants. The exact values of these parameters depend on the actual operation conditions of the device (V, T and time). One of the models within this category is the Probabilistic Defect Occupancy model (PDO) [20]. The average ΔV_{th} for particular V, T and stress, t_s, and recovery,

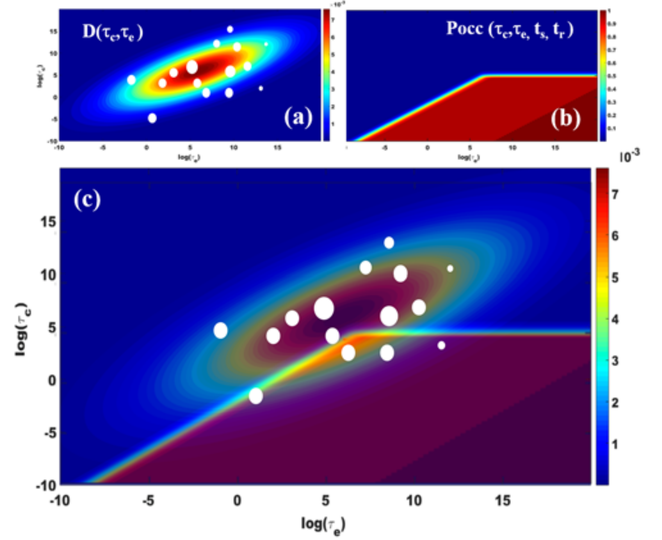

Figure 2. (a) Typical representation of the defect distribution, $D(\tau_c,\tau_e)$, which is voltage and temperature dependent. In the case of small area devices, defects can be individually identified (dots). The size of the dots is indicative of η. (b) The occupation probability of the defects is also time-dependent. (c) According to the PDO model, ΔV_{th} is calculated by integrating the product of the two magnitudes.

t_r, times is given by equation (1), where the integral part accounts for the fast recoverable component of the degradation and P_p for the 'permanent' or slower part.

$$\Delta V_{th}(t_s,t_r) = N < \eta > \int_0^{\infty}\int_0^{\infty} D(\tau_e,\tau_c)\cdot P_{occ}(\tau_e,\tau_c;t_s,t_r)d\tau_e d\tau_c + P_p(t_s,t_r) \quad (1)$$

The voltage and temperature dependences of the fast recovery component are included in N, τ_c and τ_e. In equation (1), $<\eta>$ is the average η of the defects. $D(\tau_c,\tau_e)$ corresponds to the defect density distribution in the τ_c-τ_e space, which has been determined to be log-normal bivariate and is usually displayed as the ellipses in Fig. 2.a. In the case of small area devices, individual defects can be identified (dots in Fig. 2.a and 2.c). P_{occ} corresponds to the defect occupation probability, which depends on time and operation conditions (Fig. 2.b). According to equation (1), to evaluate ΔV_{th} (for given time and operation conditions), the product $D(\tau_c,\tau_e)\cdot P_{occ}(\tau_c,\tau_e,t_r,t_s)$ must be integrated (Fig. 2.c).

In these models, the technology-dependent parameters are N, $<\eta>$ and $D(\tau_c,\tau_e)$, for large area devices or, equivalently, for small areas, N and the triad (η,τ_c,τ_e) of each defect. These parameters can be extracted from recovery traces as those plotted in Fig. 1.b, measured in test structures of the technology of interest. Note that, to build the operation-dependent defect distributions, large sets of devices (to get statistically representative values) must be characterized for many stress/relaxation times under many stress conditions (T, V_D and V_G). However, taking into account the typical duration of just one of those tests, it becomes clear that standard on-wafer serial tests become experimentally unaffordable. To illustrate this, one 6-cycle MSM test on 100 devices subjected to a concrete (T,V_D,V_G) stress triad, for total t_s=1.1 10^{5}s (~31h) and t_r=600s can typically take more than 120 days.

III. Array-based TDV characterization solutions

To allow the statistical characterization of TDV, as needed for model parameter extraction, the use of ICs that contain arrays of devices has been proposed [21-23]. Moving from

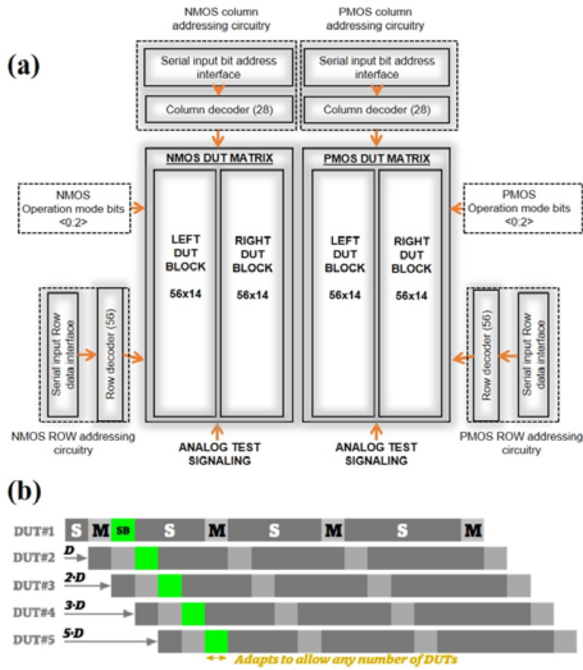

Figure 3. (a) Block diagram of the ENDURANCE chip, which contains thousands of devices, distributed into 4 arrays [21]. Its architecture allows pipelined tests, to reduce the testing time. (b) Example of parallelized test scheme of 5 devices in a four-cycle test, where 4 devices are simultaneously stressed while one is measured [24].

wafer level tests to IC measurements implies a large reduction of area (contact pads are shared by a large number of devices), but the IC must be carefully designed, for example, to withstand the large currents in the IC under the high-voltage stresses or to avoid voltage drops along the lines, among others [21]. An IC, which we named ENDURANCE, that fulfills all the requirements for massive TDV statistical characterization, has been fabricated in a 1.2V commercial 65-nm CMOS technology. The IC contains 3136 MOS devices (nMOS and pMOS of different areas) and allows measuring RTN, BTI and HCI aging. The architecture of the chip is such that individual access to the terminals of each device in the IC is possible, to stress or measure, depending on the applied voltages (Fig. 3.a).

But individual device access is not enough. TDV measurements in the IC must be fully automated, since the manual alternative becomes unattainable when hundreds or thousands of devices have to be measured at varying operation conditions with exact timing. As a consequence, together with the IC, a complete customized experimental setup (hardware + software) is a must. A software tool, which we called TARS, has been implemented, which, from a user-friendly interface where the high-level test information is provided, controls the generation of the thousands of properly scheduled instructions needed to control the hardware. More importantly, a pipelined testing procedure has been developed, to simultaneous stress multiple devices while only one device is measured at a time (Fig. 3.b) [24]. With this parallelization scheme, automatically executed by our TARS software, testing times are strongly reduced, reaching attainable values (considering the previous example, testing time is reduced from more than 120 days to less than 2 days). With our set-up, to the best of our knowledge, parallelized HCI tests have been executed for the first time (Fig.1.c).

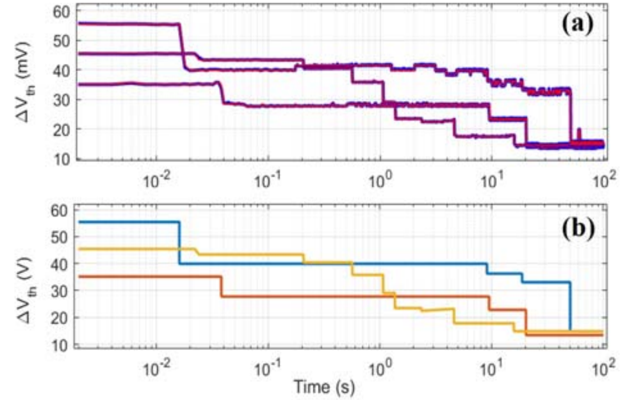

Figure 4. (a) Typical BTI recovery traces, where experimental noise and RTN effects are superposed. To extract the BTI defect parameters, these noises must be filtered, using automated smart algorithms, leading to traces as those shown in (b) [25].

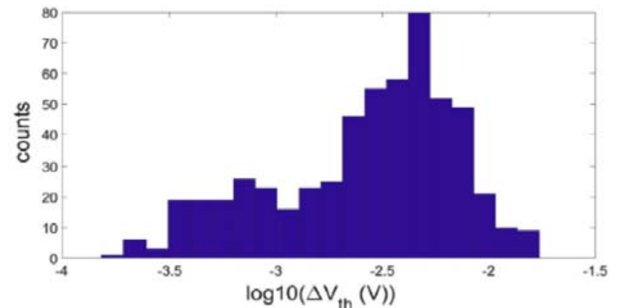

Figure 5. Experimental distribution of the threshold voltage shifts (η) measured from RTN traces in 80x60nm^2 pMOS in the ENDURANCE chip, at V_G=-1V and V_D=-0.1V, obtained with TiDeVa.

IV. SMART DATA ANALYSIS

The characterization of the devices in the IC results in thousands of traces (which correspond to different device types, areas, stress and recovery times, gate and drain voltages and temperatures), from which the parameters of the compact model (PDO model in our case) can be extracted. In addition, complex signals will have to be analyzed, where background noise, aging and RTN may coexist. Then, an initial 'cleaning' of the traces may be required for accurate parameter extraction, to isolate the different phenomena under analysis (Fig. 4) [25]. It becomes obvious that, once more, the 'manual' treatment of the data becomes unfeasible, so that the fully-automated extraction of the RTN, BTI and HCI parameters is mandatory. Recently, the authors developed such kind of tool, which was named TiDeVa [26]. As an example of the kind of output provided, Fig. 5 shows the distribution of η measured from RTN traces, automatically extracted with TiDeVa.

V. CIRCUIT RELIABILITY PREDICTION

At this point, all the device-level 'ingredients' are available for the prediction of the circuit reliability. For this, a circuit reliability simulator controls and exploits a regular circuit simulator (to determine device biasing) and the device TDV compact model, together with the parameters corresponding to the considered technology (to calculate the device aging in the circuit). Recently, a circuit reliability simulator of such kind, which we named CASE, has been presented [27], being the first reported that takes into consideration the combined TZV and TDV stochasticity.

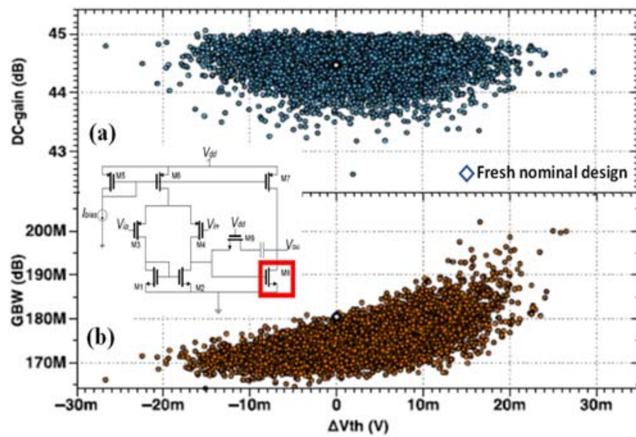

Figure 6. Sensitivity of two performance parameters of the Miller amplifier shown in the inset to the aging of transistor M8 after 1 year of operation, calculated using CASE. (a) DC-gain is not sensitive to this aging, but the bandwidth certainly is (b) [27].

Based on the PDO model, the simulator generates a number of transistor samples (with defect parameters that obey the statistical distribution experimentally obtained). For each of these samples, considering the operation conditions determined by the circuit simulator at a given time, the occupancy probability of each of the defects in the device is evaluated and the device ΔV_{th} computed. With the new device V_{th}, another circuit simulation will determine the device aging effect on the circuit performance. As an example, Figure 6 shows the sensitivity of two performance parameters of a Miller amplifier to the aging of one of its transistors. It must be noticed that, due to the transistor aging during operation, its biasing (i.e. stress) changes, so that a number of intermediate time steps have to be considered to update the device biasing. CASE uses an algorithm that adapts the size of these time steps with the progressive aging of the circuit [27]. In this manner, an accurate result is achieved while keeping the computational time as low as possible.

VI. CONCLUSIONS

In a nutshell, solutions to the demanding TDV MOSFET characterization become a must for the implementation of the RAD approach in advanced CMOS nodes. Some examples of key elements, which cover from TDV physics description to circuit reliability prediction, have been described. The core is the PDO physics-based TDV MOSFET compact model, capable of properly describing the TDV effects. At hardware level, the central element is the ENDURANCE chip, an array-based IC that allows the massive statistical characterization of TDV phenomena. At software level, complementary to the IC, the TARS toolbox, through a fully-automated pipelined testing of thousands of devices, reduces testing times to attainable values; and TiDeVa, with its smart extraction algorithms, provides the PDO model parameters. Finally, CASE, the circuit reliability simulator, evaluates the effects of device aging on circuit performance, thus facilitating the design of variability-resilient circuits.

REFERENCES

[1] M.A. Alam, K. Roy and C. Augustine, “Reliability- and process-variation aware design of integrated circuits-a broader perspective”, Proc. Int. Reliab. Phys. Symp. (IRPS), 353-363, (2011)

[2] V.J. Reddi, D.Z. Pan, S.R. Nassif and K.A. Bowman, “Robust and resilient designs from the bottom-up: technology, CAD, circuit and system issues”, Proc. ASP-DAC, 7-16, (2012)

[3] S. Ghosh and K. Roy, “Parameter variation tolerance and error resiliency: new design paradigm for the nanoscale era”, Proc. IEEE, v.98, 1718-1751, (2010)

[4] B. Vaidyanathan and A.S. Oates, “Technology Scaling Effect on the Relative Impact of NBTI and Process Variation on the Reliability of Digital Circuits” IEEE Trans. Dev. Mat. Reliab., v.12, 428-436 (2012)

[5] E. Maricau and G. Gielen, “Computer-aided analog circuit design for reliability in nanometer CMOS”, IEEE Trans. Emerging and Selected topics in circuits and systems”, Vol. 1 (1), pp. 50-58 (2011)

[6] M. Nafria et al., "Time-dependent variability of high-k based MOS devices: Nanoscale characterization and inclusion in circuit simulators," Int. Electr. Dev. Tech. Dig. (IEDM), pp. 6.3.1-6.3.4 (2011)

[7] D. Mahajan and V. Ruparelia, "Reliability Simulation and Analysis of Important RF Circuits Using Cadence Relxpert," IEEE Int. Conf. Electronics, Computing and Communication Tech., pp. 1-6 (2018)

[8] B. Tudor et al., "MOSRA: An efficient and versatile MOS aging modeling and reliability analysis solution for 45nm and below", IEEE Int. Conf. Solid-State Integrated Circuit Tech., pp. 1645-1647 (2010)

[9] A. Toro-Frias, et al., “Reliability simulation for analog ICs: goals, solutions and challenges”, Integration 55, pp. 341–348, (2016)

[10] B. Kaczer et al., “A brief overview of gate oxide defect properties and their relation to MOSFET instabilities and device and circuit time-dependent variability”, Micr. Reliab., Vol. 81, pp. 186-194 (2018)

[11] C. Claeys et al., “Random Telegraph Signal Noise in Advanced High Performance and Memory Devices”, Symposium on Microelectronics Technology and Devices (SBMicro), (2016)

[12] T. Grasser (editor), “Bias Temperature Instability for Devices and Circuits, Springer (2014)

[13] T. Grasser (editor), “Hot carrier degradation in semiconductor devices”, Springer (2015)

[14] B. Kaczer et al., “Ubiquitous relaxation in BTI stressing—new evaluation and insights”, Proc. Int. Reliab. Phys. Symp. (IRPS), pp. 20-27 (2008)

[15] S. Mahapatra and U. Sharma, “A Review of Hot Carrier Degradation in-Channel MOSFETs—Part I: Physical Mechanism”, IEEE Trans. Electron Devices, Vol. 67, No. 7, pp. 2660-2671, (2020)

[16] S. Mahapatra and N. Parihar, “A review of NBTI mechanisms and models”, Microelectr. Reliab., Vol. 81, pp 127-135 (2018)

[17] T. Grasser et al., “A Unified Perspective of RTN and BTI”, Proc. Int. Reliab. Phys. Symp. (IRPS), pp. 4A.5.1-4A.5.7 (2014)

[18] N. Ayala, J. Martin-Martinez, R. Rodriguez, M. Nafria, X. Aymerich, “Unified characterization of RTN and BTI for circuit performance and variability simulation”, Proc. ESSDERC, pp. 266-269 (2012)

[19] B. Kaczer et al., "The defect-centric perspective of device and circuit reliability — From individual defects to circuits," Proc. European Solid State Device Research Conference (ESSDERC), pp. 218-225 (2015)

[20] J. Martin-Martinez et al., “Probabilistic defect occupancy model for NBTI”, Proc. Int. Reliab. Phys. Symp. (IRPS), pp. XT.4.1-4.6 (2011)

[21] J. Diaz-Fortuny et al., “A Versatile CMOS transistor array IC for the statistical characterization of time-zero variability, RTN, BTI and HCI”, IEEE J. Solid State Circ., 54 (2), pp. 476–488 (2018)

[22] P. Weckx et al., “Characterization of time-dependent variability using 32k transistor arrays in an advanced HK/MG technology”, Proc. Int. Reliab. Phys. Symp. (IRPS), pp. 3B11–3B16 (2015)

[23] C. Schlünder et al., “A new smart device array structure for statisticalinvestigations of BTI degradation and recovery”, Proc. Int. Rel. Phys. Symp. (IRPS), pp. 56–60 (2011)

[24] J. Diaz-Fortuny et al., “Flexible setup for the measurement of CMOS time-dependent variability with array-based integrated circuits”, IEEE Trans. Instrumentation and Measurement, Vol. 69, pp.853-864 (2020)

[25] J. Diaz-Fortuny et al., “A noise and RTN-removal smart method for parameter extraction of CMOS aging compact models”, Solid State Electronics, 159, pp. 99-105 (2019)

[26] P. Saraza-Canflanca et al., “A robust and automated methodology for the analysis of Time-Dependent Variability at transistor level”, Integration, Vol. 72, pp. 13-20, 2020.

[27] P. Martin-Lloret et al., “CASE: a reliability simulation tool for analog ICs”, Proc. Int. Conf. on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), pp. 1–4 (2017)