Statistical threshold voltage shifts caused by BTI and HCI at nominal and accelerated conditions

J. Diaz-Fortuny[1*], Pablo Saraza-Canflanca[2], Rosana Rodriguez[1], Javier Martin-Martinez[1], Rafael Castro-Lopez[2], Elisenda Roca[2], Francisco V. Fernandez[2] and Montserrat Nafria[1]

[1]*Universitat Autònoma de Barcelona (UAB), Electronic Engineering Department, REDEC group. Barcelona, Spain*
[2]*Instituto de Microelectrónica de Sevilla, IMSE-CNM, CSIC and Universidad de Sevilla, Spain*
E-mail address of corresponding author: javier.diazfortuny@imec.be [*]currently affiliated with imec, Kapeldreef 75, 3001 Leuven, Belgium

***Abstract*— In nowadays deeply scaled CMOS technologies, time-zero and time-dependent variability effects have become important concerns for analog and digital circuit design. For instance, transistor parameter shifts caused by Bias Temperature Instability and Hot-Carrier Injection phenomena can lead to progressive deviations of the circuit performance or even to its catastrophic failure. In this scenario, and to understand the effects of these variability sources, an extensive and accurate device characterization under several test conditions has become an unavoidable step towards trustworthy implementing the stochastic reliability models and simulation tools needed to achieve reliable integrated circuits. In this paper, the statistical distributions of threshold voltage shifts in nanometric CMOS transistors will be studied at nominal and accelerated aging conditions. To this end, a versatile transistor array chip and a flexible measurement setup have been used to reduce the required testing time to attainable values.**

Keywords—CMOS; BTI; HCI; parameters; extraction; method; RTN; defects; aging;

I. Introduction

With nowadays CMOS technology downscaling, Bias Temperature Instability (BTI) [1-2] and Hot Carrier Degradation (HCI) [3-4] aging effects, as well as Random Telegraph Noise (RTN) [5] transient effects, have re-emerged as important time-dependent variability (TDV) phenomena that must be considered in order to achieve reliable designs of digital and analog integrated circuits (ICs) [6-7]. During circuit operation, the variability effects related to the trapping/detrapping of charge carriers in/from oxide defects could result in circuit malfunction due to the shift of key transistor parameters, such as the threshold voltage (V_{th}). Thus, it is critical for IC designers to consider TDV effects to implement reliability-aware circuits [6-8]. To this end, appropriate compact models, like the Probabilistic Defect Occupancy (PDO) model [8], are essential. These models need to account for and clearly distinguish the 'slow' defects, responsible for aging-induced degradation, from the 'fast' defects causing the RTN transient variations [5]. As shown in this paper, isolating the impact of the different aging phenomena sources in the form of transistor degradation is not a straightforward task.

Another important aspect is that an accurate transistor characterization requires a robust statistical characterization of the transistor response to accelerated stress conditions due to the stochastic nature that TDV phenomena reveal in deeply-scaled technologies [9]. Conventional BTI/HCI aging characterization techniques are based on the application of serialized measurement-stress-measurement (MSM) sequences. The MSM characterization technique consists of an initial measurement phase, in which the "fresh" characteristics of the device are measured, a stress phase, that accelerates the device degradation by means of the application of an overvoltage to the device terminals, and another measurement phase, where the effects of the induced aging can be measured in reasonable testing times. In order to obtain statistically-relevant information, a statistical aging test, with hundreds or thousands of transistors under the same stress and measurement conditions, must be performed. Typically, wafer-level device characterization implies that devices are tested one by one leading to a continuous experiment of days or months. In our work, we use our previously designed 65-nm array-based IC [10] in combination with a custom-designed characterization setup [11] that allows to execute BTI/HCI aging tests over hundreds of CMOS transistors with a stress parallelization technique that significantly reduces the total aging test time.

II. Experimental Characterization Framework

Transistor statistical characterization has been conducted using our versatile ENDURANCE transistor array IC (Fig. 1a), fabricated in a commercial 1.2-V, 65-nm bulk CMOS technology [5]. This IC allows individual access to thousands of devices of various geometries for their individual characterization. Full variability characterization can be accurately executed since each device under test (DUT) terminal can be connected/disconnected in parallel to individual analog signal paths depending on the tests to be performed. In this sense, DUT terminal connections are performed by means of appropriately sized full transmission gates (TG) designed using thick IO transistors. Each DUT is accompanied by an individual digital circuit that controls 8 TGs, i.e., 5 TGs for the drain and 3 TGs for the gate, which connect the DUT drain and gate to the corresponding on-chip analog paths. In this scenario, Fig. 1(a) illustrates the distribution of the stress, measure and stand-by analog signal paths for drain and gate DUT off-chip biasing.

The IC design harbours 3,136 unit cells, each yielding access to a single transistor with local force-&-sense connections to mitigate on-chip voltage drops, as shown in Fig. 1(b). The unit cell circuit design, consists of two separate blocks: the digital CONTROL core and the DUT block. The DUT block contains a single nMOS or pMOS transistor with two analog connections: the drain connection and the gate connection. The CONTROL core block consists of 3-bit individual memories to store the required operation mode of the unit cell, and a level shifter block that performs the necessary voltage level shift of all TG control signals from 1.2V of the digital circuitry up to the 3.3V operation voltage level of the I/O transistors.

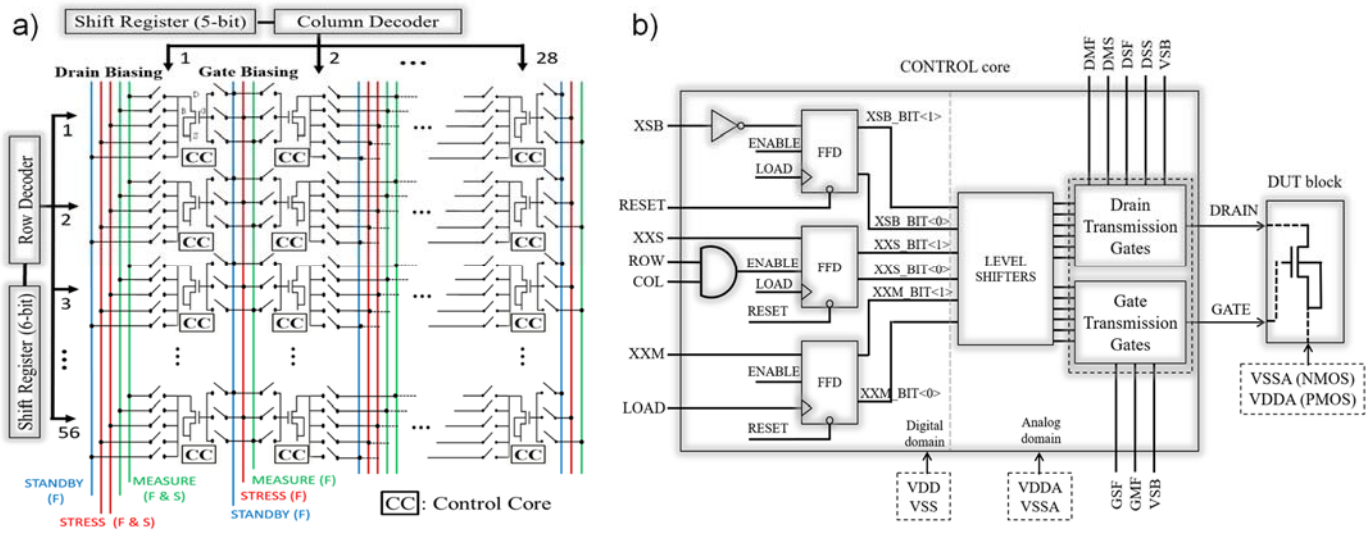

Fig. 1. (a) High-level schematic illustration of the ENDURANCE chip design showing the array structure of unit cells and (b) the unit cell design consisting of a single DUT with properly sized force-and-sense transmission gates and operation mode circuitry [5]

Extensive and accurate statistical transistor characterization has been conducted by means of our automated flexible characterization setup that allows executing parallel TDV tests over hundreds of devices simultaneously. To the best of the authors' knowledge, this is the first array IC chip that allows HCI parallel testing. Our novel all-DUT-parallel-stress-pipeline measurement (ADPSPM) measurement technique (Fig. 2), designed to accommodate Measurement Stress Measurement (MSM) test sequences in a parallel-pipelined scheme, has been utilized to significantly reduce the required aging test time from several months or years to a few days [11]. Notice that the first measurement cycle does not appear in Fig. 2, since the measurement of the pristine characteristics of the devices can be performed at any point and does not need to be accommodated within the ADPSPM timing scheme. In the first SM cycle (C1, in orange in Fig. 2), the SM pattern of each DUT is delayed ensuring that all the measurement phases in all DUTs are correctly pipelined. Once all the SM patterns of cycle C1 have been temporarily allocated, the algorithm starts the parallel distribution of the next cycle (C2). Since the setup and the chip allow the parallelization of the stress, but not of the measurement, the critical condition is that the last measurement phase of cycle C1 (for DUT#5) cannot overlap with the first measurement phase of cycle C2 (for DUT#1). To comply with this critical condition, the algorithm computes the total SM time needed for DUT#1 in cycle C2 (time B) and compares it to the time required for measurement of the remaining DUTs in cycle C1 (time A). If the difference is negative (B<A), the algorithm inserts the necessary standby period (SB, in green in Fig. 2) to avoid any measurement overlap.

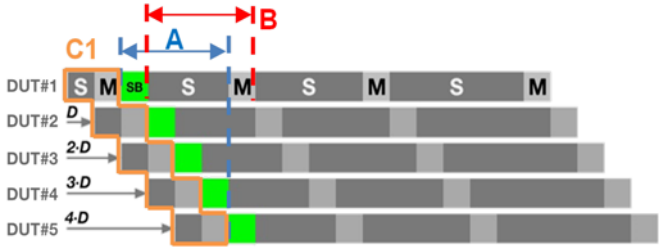

Fig. 2. Illustration of the ADPSPM test technique utilized in this study to drastically reduce the aging testing time. 5 DUTs are shown during a 4-cycle SM test. C1 illustrates the first SM pattern. Interval A denotes the time reserved for other DUTs measurements after the end of the DUT#1 SM cycle and interval B is the required time for the SM pattern execution of DUT#1 in the next cycle [6].

III. SAMPLES AND STATISTICAL TESTS

Prior to the realization of any TDV test, more than 40,000 transistors from 15 ENDURANCE chips have been evaluated at their nominal operating conditions ($|V_{GS}|\leq 1.2V$ and $|V_{DS}|\leq 1.2V$) to assess the intra- and inter-die process variability (i.e., time-zero variability, TZV). In this regard, Fig. 3 shows the empirical Cumulative Distribution Function (CDFs) of the initial device threshold voltage |Vth0|, which properly fits gaussian distributions. Statistical die-to-die |Vth0| dispersion exhibits small variability as shown by the compactness of the CDFs measured on different chips (~16mV/~23mV for nMOS/pMOS for 50% of devices). However, considerable device-to-device variability is observed (~350mV/~250mV for nMOS/pMOS).

After the TZV tests, extensive BTI and HCI accelerated tests have been conducted on these devices at different test

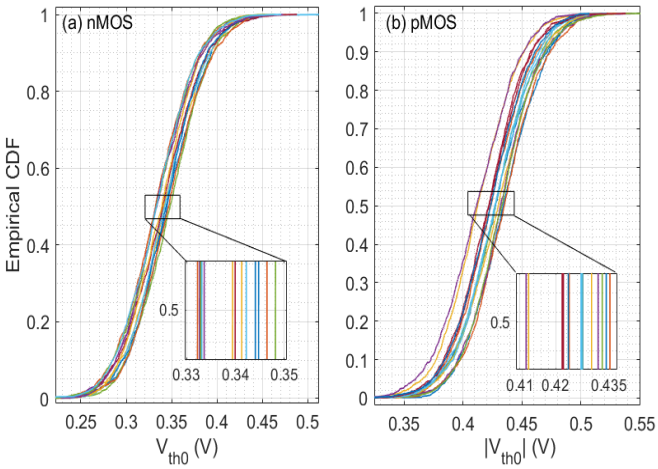

Fig. 3. Empirical CDFs of the |Vth0| parameter for the 80nm/60nm pristine nMOS (a) and pMOS (b) transistors of 15 different ICs. Die-to-die variation is much smaller than the intra-die variability.

Test	Temp	$\|V_{GS}\|;\|V_{DS}\|$ stress conditions			
		1.2V; 0V	1.5V; 0V	2.0V; 0V	2.5V; 0V
NBTI	25°C	500	400	400	400
NBTI	80°C	196	396	396	396
PBTI	25°C	100	100	100	100
PBTI	80°C	196	196	196	196
Total devices		992	1092	1092	1092
Serial		133.3d	146.7d	146.7d	146.7d
ADPSPM		5.7d	6.4d	6.4d	6.4d

Test	$\|V_{GS}\|, \|V_{DS}\|$			
	1.2V, 1.2V	1.5V, 1.5V	2.0V, 2.0V	2.5V, 2.5V
HCI NMOS	50	250	200	210
HCI PMOS	498	392	150	400
Total devices	548	642	350	610
Serial	67.8d	57.05d	23.66d	62.42d
ADPSPM	3.114d	3.426d	1.794d	3.107d

Fig. 4. Summary tables of the BTI (top) and HCI (bottom) conducted aging tests in this study. Each table specifies the total number of tested devices for each test, temperature and voltage stress conditions, the amount of devices for each $|V_{GS}|;|V_{DS}|$ condition and the test time comparison in days (d) between serial and ADPSPM methods.

conditions. Fig. 4 summarizes the aging tests, which consisted in 5-cycles MSM sequences, where stress times are increased logarithmically, i.e., 1s, 10s 100s, 1ks and 10ks, while measurement times are fixed to 100s. In total, 6,418 transistors have been tested using the ADPSPM technique. The total test time was 37 days, which is experimentally feasible, instead of the 2.36 years needed if serial tests had been performed. The advantage of our approach becomes even more evident when an additional sixth stress cycle of 100ks is added to this test: 45 days versus 22 years with the standard serial approach.

To evaluate device degradation after each stress phase, device recovery current traces were measured at $|V_{DS}|$ = 0.1V and $V_{GS} \approx V_{th0}$ for 100s. Some examples are shown in Figs. 5 and 6, which correspond to threshold voltage shift vs, time ($\Delta|Vth|$-t) traces measured on pMOS after NBTI tests at $|V_{DS}|$ = 0V and $|V_{GS}|$ = 1.2V (Fig. 5), and HCI tests at $|V_{DS}| = |V_{GS}|$ = 1.2V (Fig. 6). Note that the nominal operation voltage of the technology was applied to 'stress' the devices. However, even in this case with relatively short stress periods, the NBTI-biased devices show changes in their threshold voltage during the measurement phase, with multiple defect discharges that drive Vth towards its pristine value once that the applied voltage is

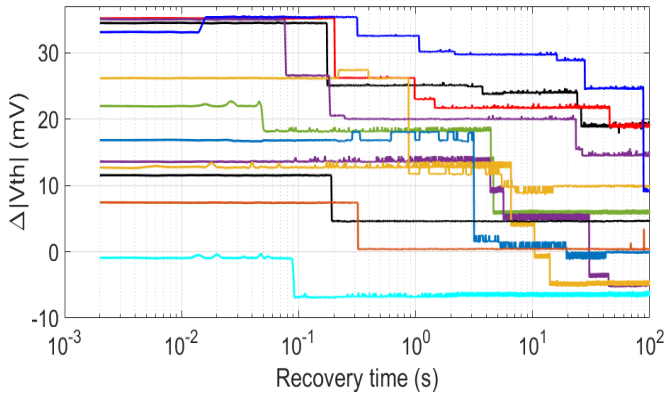

Fig.5. NBTI recovery traces measured on pMOS after a stress time of 10ks at $|V_{DS}|$ = 0V, $|V_{GS}|$ = 1.2V and T = 25ºC showing multiple discrete defect discharges combined with RTN and background noise.

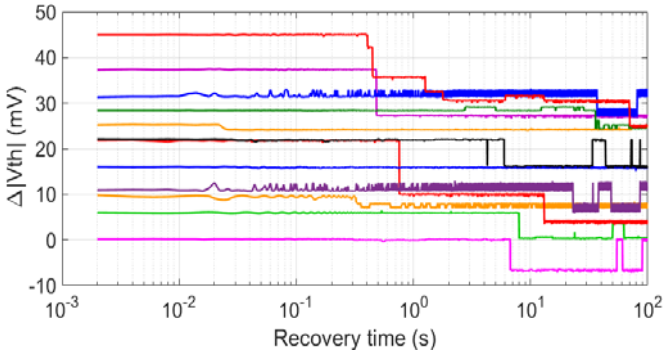

Fig. 6. HCI recovery traces of pMOS transistors after a stress time of 10ks at $|V_{DS}| = |V_{GS}|$ = 1.2V, T = 25ºC showing less discrete $\Delta|Vth|$ discharges than for NBTI (Fig. 5), while RTN is also present in almost all traces.

decreased. HCI-stressed devices (Fig. 6) can show larger $\Delta|V_{th}|$ and reveal less discrete defect discharges, resulting in a more 'permanent' damage that NBTI-stressed ones. Moreover, RTN is present in almost every trace.

The statistical $\Delta|Vth|$ variability for the different stress voltage/temperature conditions has been evaluated, as a function of the stress time. Two examples of CDFs are shown in Fig. 7 for BTI and HCI tests on pMOS at nominal (1.2V) (Fig. 7a and 7b) and accelerated (2.5V) (Fig. 7c and 7d) conditions. Fig. 7a and Fig. 7b show that $\Delta|Vth|$ increases with time, being the shifts

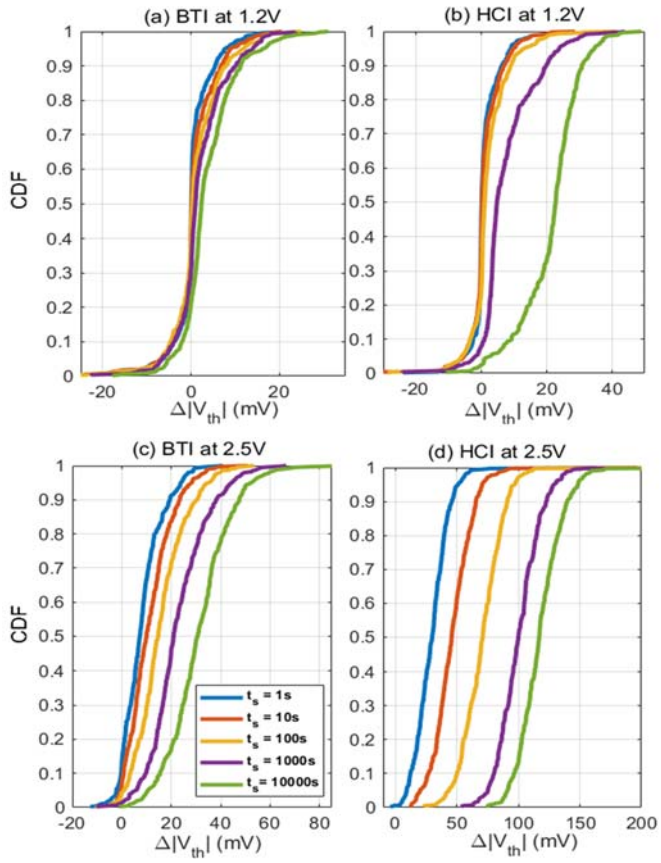

Fig.7: CDFs of the $\Delta|Vth|$ values measured on the pMOS after 100s of recovery for several stress times at T=25ºC. (a) NBTI and (b) HCI at 1.2V (nominal conditions). (c) NBTI and (d) HCI accelerated stresses at 2.5V. $\Delta|Vth| < 0$ is also present because of the significant impact of RTN.

larger when $|V_{DS}|\neq 0$. However, the observed shifts can be comparable to those introduced by TZV. Moreover, the strong presence of RTN before and after device stress, results in $\Delta|Vth| < 0V$ for some devices, as observed in Fig. 7a, b and c. As expected, the most severe accelerated stress conditions (HCI at 2.5V, Fig. 7d) clearly show larger $\Delta|Vth|$ shifts. These results suggest that the voltage dependence of TDV threshold voltage shifts can be accurately studied incorporating data obtained at nominal operation conditions.

Conclusion

This work presents statistical MOSFET TDV data obtained by means of a versatile array-IC chip together with the ADPSPM test parallelization technique. This parallelization allows a large reduction in the necessary test times (i.e., from years to a few days). This enables the statistical characterization of thousands of devices under a large range of stress conditions, such as nominal and overvoltage operation conditions. Such tests have been performed both for BTI and HCI phenomena at different temperatures, and the obtained statistical results have been discussed. BTI and HCI experiments performed at nominal voltages also revealed Vth shifts in the devices, showing recovery when the applied voltages were decreased. However, at these conditions, the TDV related shifts are comparable to those introduced by TZV. . Moreover, RTN was observed simultaneously to the Vth shifts..These results suggest that, because the strong reduction of the testing time allowed by the set-up, data obtained at nominal operation conditions can be incorporated into the aging modelling.

Acknowledgements

This work has been supported in part by the Spanish TOGETHER Project TEC2016-75151-C3-R (AEI/FEDER, UE) and by the VIGILANT Project (PID2019-103869RB / AEI / 10.13039/501100011033). Pablo Sarazá Canflanca acknowledges MICINN for supporting his research activity through the predoctoral grant BES- 2017-080160.

References

[1] Kaczer B, Grasser T, Roussel PJ, Franco J, Degraeve R, Ragnarsson LA, et al. Origin of NBTI variability in deeply scaled pFETs. IEEE Int Reliab Phys Symp Proc 2010:26–32. doi:10.1109/IRPS.2010.5488856.

[2] Kaczer B, et al. "A brief overview of gate oxide defect properties and their relation to MOSFET instabilities and device and circuit time-dependent variability." Microelectronics Reliability 81 (2018): 186-194.

[3] Duan M, Zhang JF, Ji Z, Zhang WD, Kaczer B, Asenov A. Key Issues and Solutions for Characterizing Hot Carrier Aging of Nanometer Scale nMOSFETs. IEEE Trans Electron Devices 2017;64:2478–84. doi:10.1109/TED.2017.2691008.

[4] Martín-Martínez J, Gerardin S, Amat E, Rodríguez R, Nafría M, Aymerich X, et al. Channel-hot-carrier degradation and bias temperature instabilities in CMOS inverters. IEEE Trans Electron Devices 2009;56:2155–9. doi:10.1109/TED.2009.2026206.

[5] Diaz-Fortuny J, Martin-Martinez J, Rodriguez R, Nafria M, Castro-Lopez R, Roca E, Fernandez FV. A noise and RTN-removal smart method for parameters extraction of CMOS aging compact models. In 2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS) 2018 Mar 19 (pp. 1-4). IEEE.

[6] Toro-Frias A, Martin-Lloret P, Martín-Martínez J, Castro-López R, Roca E, Rodríguez R, Nafria M, Fernández FV. Reliability simulation for analog ICs: Goals, solutions, and challenges. Integration. 2016 Sep 1;55:341-8.

[7] Rzepa G, et al. "Comphy—A compact-physics framework for unified modeling of BTI." Microelectronics Reliability 85 (2018): 49-65.

[8] Martin-Martinez J, Kaczer B, Toledano-Luque M, Rodriguez R, Nafria M, Aymerich X, Groeseneken G. Probabilistic defect occupancy model for NBTI. In2011 International Reliability Physics Symposium 2011 Apr 10 (pp. XT-4), IEEE.

[9] Kaczer B, Grasser T, Roussel J, Martin-Martinez J, O'Connor R, O'Sullivan BJ, et al. Ubiquitous relaxation in BTI stressing-new evaluation and insights. Proc. Int. Reliab. Phys. Symp., IEEE; 2008, p. 20–7. doi:10.1109/RELPHY.2008.4558858.

[10] Diaz-Fortuny J, Martin-Martinez J, Rodriguez R, Nafria M, Castro-Lopez R, Roca E, et al. A transistor array chip for the statistical characterization of process variability, RTN and BTI/CHC aging. Proc. 14th Int. Conf. Synth. Model. Anal. Simul. Methods Appl. to Circuit Des., 2017. doi:10.1109/SMACD.2017.7981600.

[11] Diaz-Fortuny J, Saraza-Canflanca P, Castro-Lopez R, Roca E, Martin-Martinez J, Rodriguez R, Fernandez FV, Nafria M. Flexible setup for the measurement of CMOS time-dependent variability with array-based integrated circuits. IEEE Transactions on Instrumentation and Measurement. 2019 Apr 19;69(3):853-64.