On the impact of the biasing history on the characterization of Random Telegraph Noise

Pablo Saraza-Canflanca, Rafael Castro-Lopez, Elisenda Roca, Javier Martin-Martinez, Rosana Rodriguez, Montserrat Nafria and Francisco V. Fernandez, *Member, IEEE*

***Abstract*—Random Telegraph Noise is a time-dependent variability phenomenon that has gained increased attention during the last years, especially in deeply-scaled technologies. In particular, there is a wide variety of works presenting different techniques designed to analyze current traces in scaled FET devices displaying Random Telegraph Noise, and others focused on modeling the phenomenon using the parameters extracted through such techniques. However, very little attention has been paid to the effects that the biasing conditions of the transistors prior to the measurements may have on the extraction of the parameters that characterize this phenomenon. This paper investigates how these biasing conditions actually impact the extracted results. Specifically, it is demonstrated that the results obtained when Random Telegraph Noise is measured immediately after the device is biased may lead to an overestimation of the Random Telegraph Noise impact with respect to situations in which the device has been previously biased for some time. This fact is, first, presented from a theoretical point of view, and, after, demonstrated experimentally through measurements obtained from a CMOS-transistor array.**

***Index Terms*—Random telegraph noise (RTN), noise measurement, reliability, variability, characterization lab**

I. Introduction

The interest in Random Telegraph Noise (RTN) in deeply-scaled CMOS technologies [1] has considerably increased in recent years, due to its role as a source of time-dependent variability [2], [3]. At the transistor level, RTN is observed as discrete shifts of the drain current, which are caused by threshold voltage shifts associated to sudden and stochastic charge trapping/de-trapping events in/from device defects [4]. It is therefore critical to characterize this phenomenon and extract information about the statistical distributions of its main parameters to be able to predict and mitigate its potentially harmful impact on circuits [3],[5].

Fig. 1 shows an example of a measured drain current (I_D) vs. time trace in which one RTN defect can be detected (i.e., two distinct current levels are present) with its main parameters annotated. In a defect-centric framework [4], these parameters are the number of defects in the transistor (or, alternatively, the density of defects of the technology), the amplitude of the current shifts, ΔI (or analogously the amplitude of the threshold voltage shifts, ΔV_{th}) associated to the trapping/detrapping events of each of these defects, and their time constants. These time constants are the capture time (τ_c), i.e., the average time that a defect takes to capture a charge carrier when it is empty, and the emission time (τ_e), i.e., the average time that an RTN defect takes to emit the charge carrier once it is occupied. The actual time instant at which a certain emission (t_{e_i}) or capture (t_{c_i}) event occurs is random, as illustrated in Fig.1. All these RTN parameters are stochastic and should be described by the corresponding statistical distribution functions.

Among the extensive literature devoted to the study of the RTN phenomenon, a considerable number of works aims at developing methods for the extraction of RTN current levels from the current traces obtained experimentally. In [6], a 2D plot is obtained by plotting half of the current samples at each axis. An analogous approach is the time-lag plot (TLP)-based method [7]. The TLP is constructed by plotting the *i*-th current sample in the *x*-axis and the (*i*+1)-th sample in the *y*-axis for the full current trace. Then, if RTN is present in the trace, the points located in the diagonal of the graph will correspond to the distinct current levels, while the off-diagonal points will correspond to the RTN transitions. An enhanced TLP (eTLP) is used in [8] that assesses how often each point of the TLP is occupied. The weighted time lag plot (wTLP) method, an extension of the TLP, has been presented as robust even when

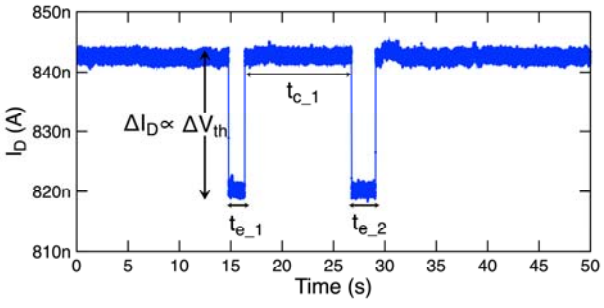

Fig. 1. Measured drain current trace displaying RTN transitions corresponding to a single defect.

P. Saraza-Canflanca, R. Castro-Lopez, E. Roca, and F.V. Fernandez are with Instituto de Microelectronica de Sevilla (CSIC and Universidad de Sevilla), 41092 Sevilla, Spain (e-mail: Francisco.Fernandez@imse-cnm.csic.es).

J. Martin-Martinez, R. Rodriguez, and M. Nafria are with Electronic Engineering Department (REDEC) group, Universitat Autònoma de Barcelona (UAB), 08193 Barcelona, Spain.

This work was supported by grant PID2019-103869RB-C31 and PID2019-103869RB-C32 funded by MCIN/AEI/ 10.13039/501100011033. P. Saraza-Canflanca´s research activity was supported by grant BES-2017-080160 funded by MCIN/AEI/ 10.13039/501100011033, and by "ESF Investing in your future". (*Corresponding author: Francisco V. Fernandez*)

the background noise (i.e., other noise sources associated to the device-under-test, noise associated to the access circuitry and the measurement equipment itself) is large [9]. A different approach is followed in [10], [11] where the maximum likelihood estimation (MLE) method is used to estimate the parameters of the statistical model. Although the previous methods are based on different principles, they all aim at the detection of the distinct current levels from which the RTN parameters, namely the number of defects, and their associated time constants and current shifts, can be determined. Then, the statistical distributions that these parameters follow can be determined and used to build a model that describes RTN. Such a model could then be integrated into simulation tools to predict the impact of RTN on circuit performances, which would allow circuit designers to perform RTN-aware design.

However, although the above-discussed methods are very useful for the analysis of RTN, they may lead to an inaccurate determination of the RTN parameters if not applied carefully. For instance, there are some works about the effect that undersampling during measurement (i.e., using a sampling rate which is not fast enough with respect to the RTN time constants), and not considering enough transitions when averaging the t_{c,e_i}, may have on the erroneous determination of the time constants $\tau_{c,e}$ [12], with measurement guidelines further developed in [13]. In this work, the effect of the bias conditions that the device has experienced prior to the measurement (i.e., its biasing history) may have on the RTN characterization is analyzed. In particular, it is shown how important it is to consider the biasing conditions not only during the measurement itself, which are generally taken into account, but also prior to it. It will be demonstrated, theoretically and experimentally, that different biasing histories can lead to a different RTN behavior during the actual measurement, and therefore to the extraction of different parameter values, so that these pre-measurement conditions should be accounted for when analyzing the measured trace. To the best of the authors' knowledge, it is the first time that the influence of the biasing history on the RTN measurements is analyzed.

The rest of this paper is organized as follows. Section II introduces the probability of occupation, an essential magnitude to assess the state of each defect according to the biasing conditions and their timings. This magnitude is used to simulate the impact that the biasing history has on the actual measurement. Then, in Section III, experimental measurements are reported that validate the findings in Section II. Finally, Section IV presents the concluding remarks with recommendations to correctly account for the impact of the biasing history.

II. Probability of Occupation

Let us consider a defect with emission and capture times τ_e and τ_c. The capture/emission process can be described by a two-state Markov chain [14]. Then, the emission probability of the occupied defect, p_e, and the capture probability of the empty defect, p_c, in a given time interval Δt, are given by:

$$p_e = \frac{\Delta t}{\tau_e} \quad p_c = \frac{\Delta t}{\tau_c} \tag{1}$$

This holds true as long as the considered time interval is much shorter than the time constants of the defect ($\Delta t \ll \tau_e, \tau_c$). Then, the probability that the defect is occupied (probability of occupation, p_{occ}) at the end of Δt is the probability that it was empty at the beginning of the interval (1-$p_{occ}(t)$), times the probability that a charge is captured during Δt (p_c), plus the probability that the defect was occupied at the beginning of Δt ($p_{occ}(t)$), times the probability that the charge is not emitted during Δt, $(1 - p_e)$:

$$p_{occ}(t + \Delta t) = \left(1 - p_{occ}(t)\right)\frac{\Delta t}{\tau_c} + p_{occ}(t)(1 - \frac{\Delta t}{\tau_e}) \tag{2}$$

The underlying assumption in (2) is that the emission (p_e) and capture (p_c) probabilities have not changed along the time interval Δt. Therefore, (2) can be expressed in differential form as:

$$\frac{dp_{occ}(t)}{dt} + p_{occ}(t)\left(\frac{1}{\tau_e} + \frac{1}{\tau_c}\right) - \frac{1}{\tau_c} = 0 \tag{3}$$

Another aspect to consider is that the emission and capture times change with the bias conditions. The following exponential dependencies on the gate and drain voltages are assumed in this work (with source and bulk short-circuited):

$$\begin{aligned} \tau_e &= \tau_{e0} 10^{\beta_e |V_{gs}|} 10^{\gamma_e |V_{ds}|} \\ \tau_c &= \tau_{c0} 10^{\beta_c |V_{gs}|} 10^{\gamma_c |V_{ds}|} \end{aligned} \tag{4}$$

The absolute values have been used for the voltages so that they take positive values for both nMOS and pMOS. This model is common in the literature and fits well with the experimental measurements [15]. There is also general agreement that β_e is positive and β_c is negative, i.e., the emission times increase with $|V_{gs}|$ while the capture times decrease [15]. For the sake of illustration, let us analyze what (4) means for a given defect when $|V_{gs}|$ is varied while $|V_{ds}|$ is kept constant. At "low" $|V_{gs}|$, its time constants imply that this defect has a high probability of being empty, i.e., a small p_{occ}. Then, as $|V_{gs}|$ increases, its τ_e will increase, while its τ_c will decrease. This means that the defect will take, on average, more time to emit a charge carrier when occupied, and less time to capture one when empty. Therefore, its p_{occ} increases with $|V_{gs}|$, and there will be certain $|V_{gs}|$ values for which the defect will have a similar probability of being empty and occupied. Then, as $|V_{gs}|$ further increases, the defect's time constants continue to change, and its p_{occ} continues to increase, so that at "high" $|V_{gs}|$, the defect has a high p_{occ}. The previous case can be generalized: taking the dependences in (4) into account, (3) can be solved to obtain the temporal evolution of the probability of occupation p_{occ} for any time-varying bias conditions applied to the transistor.

The characterization of RTN in a device is usually carried

out in the following manner: first, the device is biased at time $t = 0$, which will be referred to as biasing instant, by applying certain voltages $|V_{gs}|$ and $|V_{ds}|$ to its terminals, and then, at some instant $t \geq 0$, its drain current is measured during some time. Usually, no attention is paid to how much time has passed between the biasing instant and the start of the measurement. However, if these instants and the voltage values $|V_{gs}|$ and $|V_{ds}|$ are known, it is possible to analytically solve (3) to study how the evolution of the p_{occ} depends on the biasing history. From the unlimited number of possible scenarios in terms of biasing history, two will be considered here: the scenario in which the measurement starts right after the biasing occurs (at $t = 0$), and the scenario in which the device is biased, and the measurement starts 10,000s after the biasing instant (at $t = 10{,}000$s). This second scenario emulates the impact of RTN on a device that has been powered on for a long time.

A. Measurement starts right after biasing

In this scenario, the devices have been off for a long time and then, at $t = 0$, some biases are applied, and the drain current is measured to monitor RTN during a certain time t_{meas}. In this scenario, the biasing instant and the beginning of the measurement concur. It can be assumed that when the transistor was powered-off, there was no voltage difference between its terminals and, therefore, $\tau_e = \tau_{e0}$ and $\tau_c = \tau_{c0}$. Also, in this scenario, the time that the transistors have been powered-off before $t = 0$ is long enough so that the probability of occupation is not changing with time and, therefore, from (3), the probability of occupation at $t = 0$ is:

$$p_{occ}(t = 0) = \frac{\tau_{e0}}{\tau_{e0} + \tau_{c0}} \tag{5}$$

Fig. 2 depicts the map of the probability of occupation of the defects at the start of the measurement window $(t = 0)$ as a function of the emission (τ_{e0}) and capture (τ_{c0}) time constants of the defects. Notice that, to ease the visualization and the comparison between maps at different conditions, the axis of these maps do not correspond to the time constants at the biasing conditions of each measurement. Instead, they always correspond to the time constants at $|V_{gs}| = |V_{ds}| = 0$V, i.e., (τ_{e0}, τ_{c0}). In Fig. 2, the dark blue area represents defects that are almost always empty; hence, they usually do not undergo any charge trapping/detrapping event, and do not cause discrete RTN current shifts at the device drain current. The dark red area represents defects that are almost always occupied; analogously to the previous case, these defects do not cause RTN current shifts. However, the intermediate region represents defects with a probability of occupation significantly different from 0 or 1.

If the biasing voltages have remained constant during the measurement window, the emission and capture times for those bias conditions can be obtained from (4) and the probability of occupation at the end of the measurement window can be obtained from the solution of (3):

$$p_{occ}(t = t_{meas}) = \\ = \frac{\tau_e}{\tau_e + \tau_c} + \left(p_{occ}(t = 0) - \frac{\tau_e}{\tau_e + \tau_c}\right) e^{-\left(\frac{1}{\tau_e} + \frac{1}{\tau_c}\right) t_{meas}} \tag{6}$$

Let us consider the case in which the bias conditions are $|V_{gs}| = 1.2V$ and $|V_{ds}| = 0.1V$, and $t_{meas} = 100s$. For this case, the plot in Fig. 3 is obtained for the probability of occupation as a function of the time constants at the end of the measurement window $(t = 100s)$. In this figure, it can be observed that the probability of occupation of the defects with certain emission and capture times has changed significantly with respect to the probability of occupation at the beginning of the measurement window. For instance, consider the hypothetical defect with certain time constants at the measurement bias conditions, indicated with a white dot in Fig. 2 and Fig. 3. Before powering on the transistor, that defect was probably empty. Therefore, when the transistor is powered on and the measurement starts at $t = 0$, the defect is empty (i.e., its p_{occ} is very close to 0). However, at the end of the

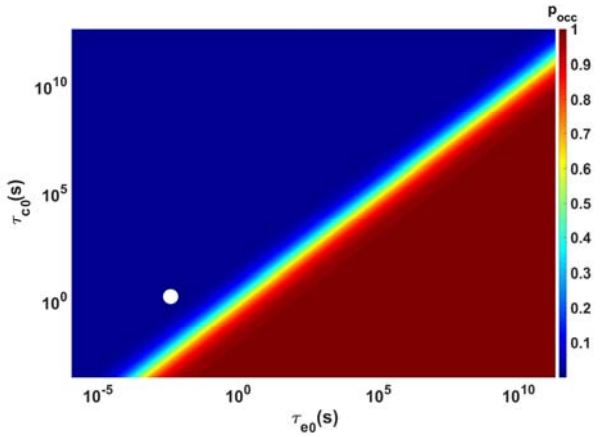

Fig. 2. Probability of occupation at the start of the measurement window when the measurement starts right after the biasing of the devices.

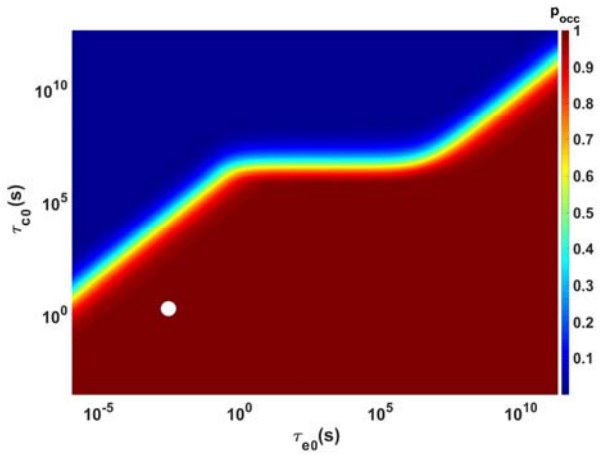

Fig. 3. Probability of occupation at the end of the measurement window when the measurement starts right after the biasing of the devices.

measurement, after 100s biased with $|V_{gs}| = 1.2V$ and $|V_{ds}| = 0.1V$, the defect has a very high probability of being occupied (i.e., its p_{occ} is close to 1).

To generalize the case of how the p_{occ} of that individual defect has changed after 100s to all the possible defects, it is possible to subtract the probability of occupation between the end and the beginning of the measurement window, as shown in Fig. 4. In this case, the white area represents defects that have not significantly changed their probability of occupation during the measurement window. On the other hand, the shadowed regions correspond to defects that have significantly increased their probability of occupation. These include, for example, defects that were empty at the beginning of the measurement window and now are observed either as charge trapping/detrapping RTN defects, or as permanently occupied defects.

B. Measurement starts a certain time after biasing

In this second scenario, the devices have been off for a very long time, and then have been biased during a certain time t_{bias} = $10{,}000s$ in the same conditions of the measurement before the actual measurement starts. The duration of the measurement is again $t_{meas} = 100s$. Let us consider the same biasing conditions as in the previous case $|V_{gs}| = 1.2V$ and $|V_{ds}| = 0.1V$. Equation (6) can be used to evaluate the probability of occupation at the beginning ($t = t_{bias}$) and at the end ($t = t_{bias} + t_{meas}$) of the measurement window. The probability of occupation at the start of the measurement window ($t = t_{bias}$) is shown in Fig. 5(a). It can be observed that this map is very different from the one of the first scenario, shown in Fig. 2. This is to be expected, since the start of the measurement in the first scenario occurred immediately after the transistor is biased, while in this second scenario the transistor has already been biased for 10,000s when the measurement starts, and during those 10,000s the probability of occupation map has evolved according to (6). Then, the probability of occupation at the end of the measurement window can be calculated again and this one is plotted in Fig. 5(b). It can be observed that in this case, unlike in the first scenario, there are no significant differences between the maps at the beginning and at the end of the measurement window.

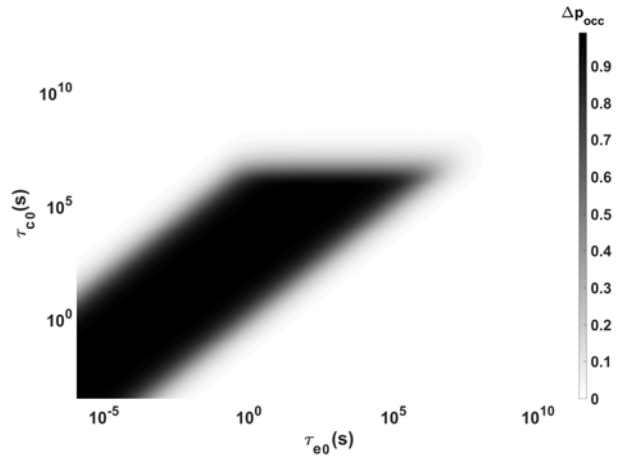

Fig. 4. Difference of the probability of occupation maps at t=100s and t=0.

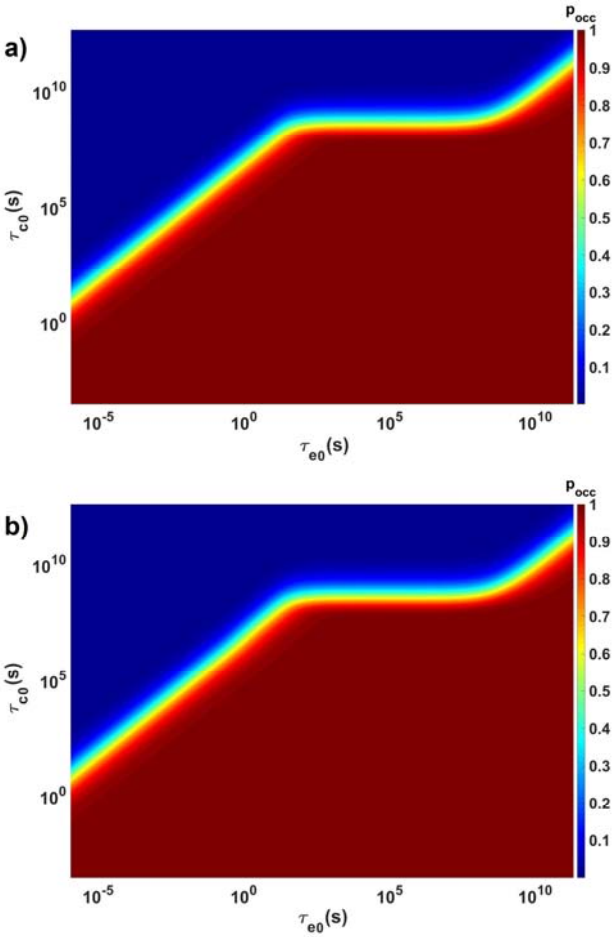

Fig. 5. Probability of occupation at the start (top) and at the end (bottom) of the measurement window when it starts 10,000s after the biasing.

The reason behind is that, in the first scenario, the occupancy state of many defects at the beginning of the measurement window does not match what their time constants during biasing would dictate in a stationary situation, but rather their time constants when the device is not powered on. The consequence is that part of the defects that were empty at the beginning of measurement will exhibit RTN behavior, while others will experience a capture event and remain occupied during the whole measurement window, since their time constants at the measurement biasing conditions cause them to be predominantly occupied. As stated above, the capture and emission process are commonly described by a two-state Markov chain. This means that, for a defect with a given emission or capture time, the corresponding probability distribution function of the real time at which its emission or capture events occur spans with a significant probability up to three decades of time around the emission or capture time (see Fig. 6). Therefore, it is logical that in the second scenario there

are minor differences, since after 10,000s it is expected that all defects that may have an impact on a measurement window of 100s (two decades smaller) statistically have an occupancy state according to their current time constants.

Since the evolution of the probability of occupation depends on the defect time constants, and these depend on the biasing voltages at the device terminals (see (4)), it follows that the probability of occupation map after a given biasing time t_{bias} will also depend on the applied bias values. Fig. 7(a) and Fig. 7(b) show the probability of occupation before and after the same measurement window when the biasing voltages are $|V_{gs}| = 0.6V$ and $|V_{ds}| = 0.1V$ and the measurement starts immediately after the transistor is biased. Fig. 7(c) and Fig. 7(d) show the equivalent probability of occupation before and after the measurement window when these biasing voltages are applied for $t_{bias} = 10{,}000s$ before the measurement window starts. When comparing these occupation probability maps with those obtained at larger $|V_{gs}|$ (Figs. 2, 3 and 5), it can be concluded that the same qualitative behavior is observed for lower voltages. However, the higher $|V_{gs}|$the more different the situation will be with respect to the first scenario in which t_{bias} = 0s.

Up to this point, it has been shown through theory how the biasing conditions that devices experience prior to being measured (i.e., voltage values and timing) can considerably impact the map of probability of occupation of the defects during the measurement. In particular, some transitions that would be observed during the measurement in the first scenario (i.e., the measurement starts immediately after biasing) and would be therefore considered while constructing a model, correspond to defects that were empty before the device was biased, and tend to be occupied when biasing is applied, and therefore will undergo only one capture event. These defects would not be observed in a measurement performed in the second scenario (i.e., the measurement starts after the devices have been biased for a long time), since they remain permanently occupied at those biasing conditions. If the analysis and modelling performed do not account for this, those defects may lead to an overestimation of the impact of RTN. In

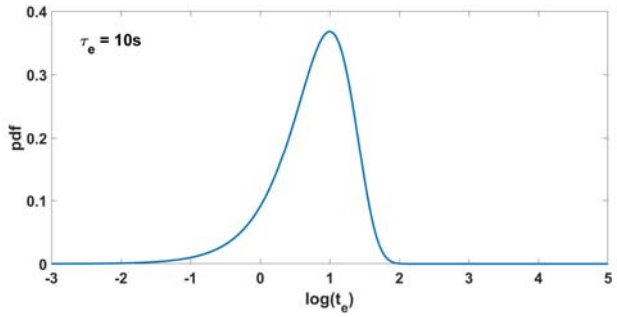

Fig. 6. Probability density function of the times at which real emission events occur (t_e) for a defect with an emission time constant of 10s.

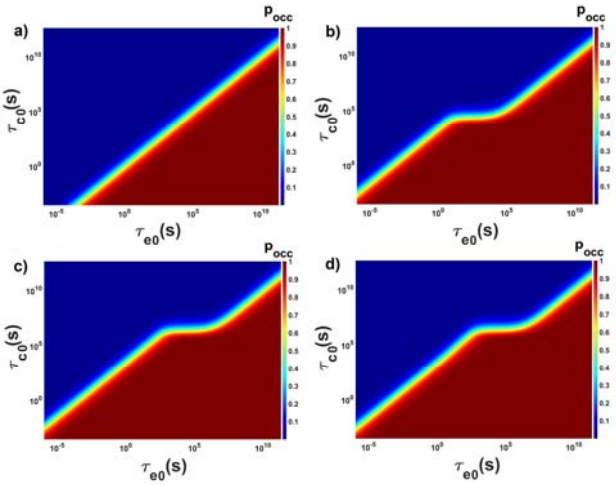

Fig. 7. Probability of occupation at the start (a) and end (b) of the measurement window when the measurement starts immediately after the biasing, and at the start (c) and end (d) when the measurement starts 10,000s after the biasing.

the next Section, experimental results will be presented to validate these theoretical findings.

III. Experimental Results

A. Impact on measurements

To demonstrate that the theoretical framework presented in the previous Section has a significant impact in practical RTN characterization, experimental measurements have been performed on a transistor-array chip. This integrated circuit contains 3,136 MOS transistors fabricated in a commercial, bulk CMOS, 65-nm, 1.2-V technology. The transistors are arranged in four matrices, two of which only contain nMOS transistors and two of which only contain pMOS transistors. The design of the array enables the accurate biasing of the transistors and a precise timing during the characterization tests [16], both fundamental features to replicate the two different scenarios presented in Section II. One pMOS matrix, containing 784 Devices Under Test (DUTs) with minimum dimensions (i.e., W=80nm and L=60nm), was selected to perform RTN measurements using the experimental setup in Fig. 8. The setup includes a full-custom printed circuit board (PCB), where the IC is inserted for DUT measurements. This IC is digitally controlled by a USB Digital Acquisition System (DAQ), model USB- 6501 from National Instruments, equipped with 24 digital IO channels. Excitations and measurements are performed with the Keysight Semiconductor Parameter Analyser (SPA) model B1500A. This SPA has 4 High Resolution Sense Measurement Units (HRSMU), with Force-&-Sense outputs for precise voltage application and current measurement. The complete setup is controlled by a personal computer equipped with Matlab®. More details can be found in [17].

The characterization strategy has been to replicate the two scenarios introduced in Section II. To this end, first 392 DUTs

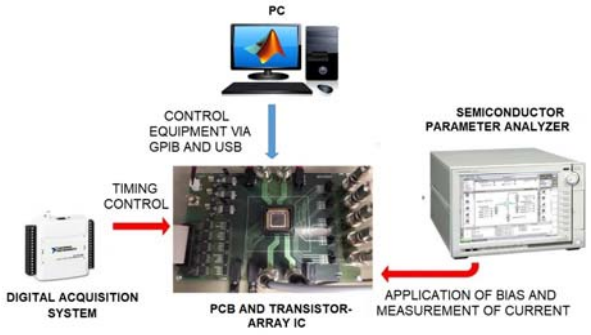

Fig. 8. Simplified representation of the experimental setup.

(i.e., half a matrix of the IC) have been biased with $|V_{gs}| = 1.2V$ and $|V_{ds}| = 0.1V$ and their drain current has been measured for 100s, starting right after the biasing conditions above are applied. Each DUT has remained biased at those same conditions after the measurement was finished and, after being biased for 10,000s, the drain current of those same devices have been measured again for 100s. Fig. 9 shows some of the measured current traces for the first (a) and the second (b) scenarios. For ease of visualization, only the traces of the first seven DUTs are shown in each case. The same color is used for the same device in both plots. It can be seen that the current traces of the seven different devices have different mean values, which is caused by process variability. Additionally, the traces reveal a typical RTN behavior, i.e., discrete current shifts associated to charge trapping/detrapping in/from defects. As expected, the amplitude shifts and time constants of the different defects are not equal since these are stochastic parameters. At first sight, no significant difference can be appreciated between the traces obtained for the two different scenarios. However, it is difficult to observe a trend from just a few traces as each device will exhibit very few active defects in a typical measurement window and for certain resolution of the measurement equipment.

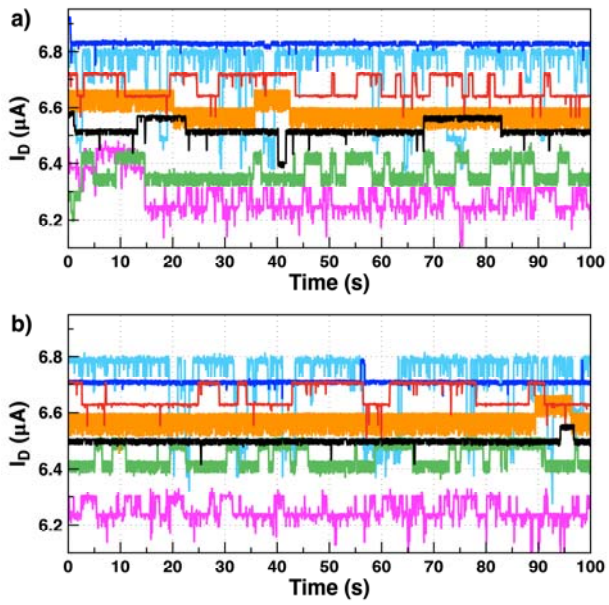

Fig. 9. Drain current traces measured for seven devices when the first (a) and the second (b) scenarios are considered.

To obtain a deeper insight into the difference between both scenarios, the following test was performed. The currents of the 392 measured devices were added together for each of the two scenarios. Fig. 10(a) shows the total current flowing through the 392 devices vs. time for the first scenario. The contribution of the individual defects is not visible any more due to the very high number of defects that appear when 392 DUTs are considered. However, it can be observed that there is a decreasing trend of the current, corresponding to the occupation of defects that have significantly increased their probability of occupation from the beginning to the end of the measurement window. That is, there are more capture than emission events. This phenomenon is much better observed if the time axis is plotted in logarithmic scale (see Fig. 10(b)), which clearly corresponds to the fact that the time constants of defects are naturally distributed on a logarithmic time scale [7], [18]. Fig. 10(c) and Fig. 10(d) show the equivalent graphs for the second scenario, where devices have been biased in the same conditions than during the measurement for $t_{bias} = 10{,}000s$ before the measurement starts. In this case, no clear temporal trend is observed for the total current, which means that, on average, the defects have not changed considerably their probability of occupation. Therefore, while in this second scenario there is approximately the same number of capture and emission events, in the first scenario there are more capture events than emission ones.

An analogous experiment has been performed for $|V_{gs}| = 0.6V$ and $|V_{ds}| = 0.1V$ using the remaining 392 DUTs of the matrix that were not biased during the first experiment. The corresponding results are shown in Fig. 11. These display the same qualitative behavior as those shown in Fig. 10. However,

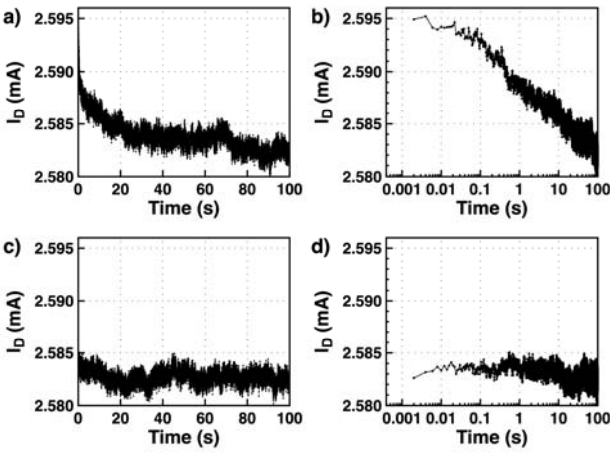

Fig. 10. Total current for the 392 DUTs with $|V_{gs}| = 1.2V$ and $|V_{ds}| = 0.1V$ in the first scenario in (a) linear and (b) logarithmic time scales, and in the second scenario in (c) linear and (d) logarithmic time scales.

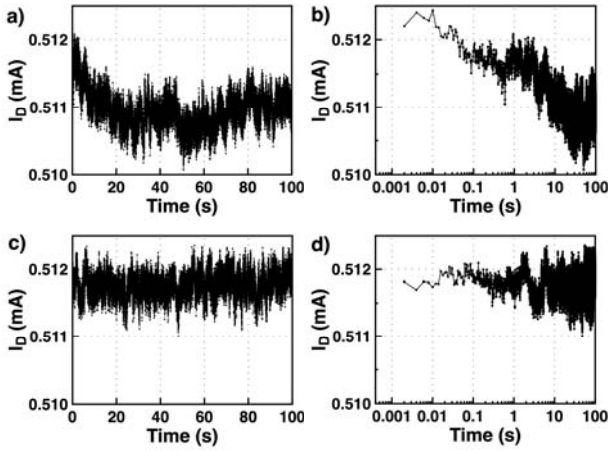

Fig. 11. Total current for the 392 DUTs with $|V_{gs}| = 0.6V$ and $|V_{ds}| = 0.1V$ in the first scenario in (a) linear and (b) logarithmic time scales, and in the second scenario in (c) linear and (d) logarithmic time scales.

this time the decreasing current trend in the first scenario is considerably smaller than in the previous case. This could be forecasted from the comparison of the probability of occupation before and after the measurement window in Fig. 7 when a lower $|V_{gs}|$ is used.

In the following, a different metric will be used to quantify the impact of the biasing history on the measurable RTN activity. This is the Maximum Current Fluctuation (MCF), which requires a very simple processing of the current traces measured in the lab and is very useful to evaluate the impact of RTN in them [19].

Consider the current trace shown in Fig. 12 (a) corresponding to a pMOS device with the same dimensions as the previously discussed ones, biased with $|V_{gs}| = 0.6$V and $|V_{ds}| = 0.1$V. It is possible to define the cumulative maximum current (*CMAXC*) at any time instant t within the experimental window as the maximum current within the interval $0 \leq t' \leq t$:

$$CMAXC(t) = \max_{\forall t' \in [0,t]} I(t') \tag{7}$$

CMAXC is plotted in the same figure with the upper orange line. Similarly, the cumulative minimum current (*CMINC*) at any time instant t within the experimental window is defined as:

$$CMINC(t) = \min_{\forall t' \in [0,t]} I(t') \tag{8}$$

and is plotted in the same figure with the lower red line. Then, the *MCF* is defined as the difference between both:

$$MCF(t) = CMAXC(t) - CMINC(t) \tag{9}$$

Therefore, the *MCF* of a given current trace accounts for the maximum range of current values that has occurred during that time window. Thus, the discrete current shifts induced by charge trapping/detrapping in/from defects will contribute to this metric (i.e., a trace with more/bigger current shifts will have a larger *MCF* than a trace with fewer/smaller ones). Fig. 12 (b) shows the temporal evolution of the *MCF*s of 20 pMOS transistors. From these data, the cumulative distribution function (cdf) of the *MCF* in a set of DUTs can be obtained for a particular time instant, as shown in Fig. 12(c). The cdf of the *MCF* at any time instant can be related to the distribution of the current shift amplitudes and to the distribution of the number of active defects up to that point [19].

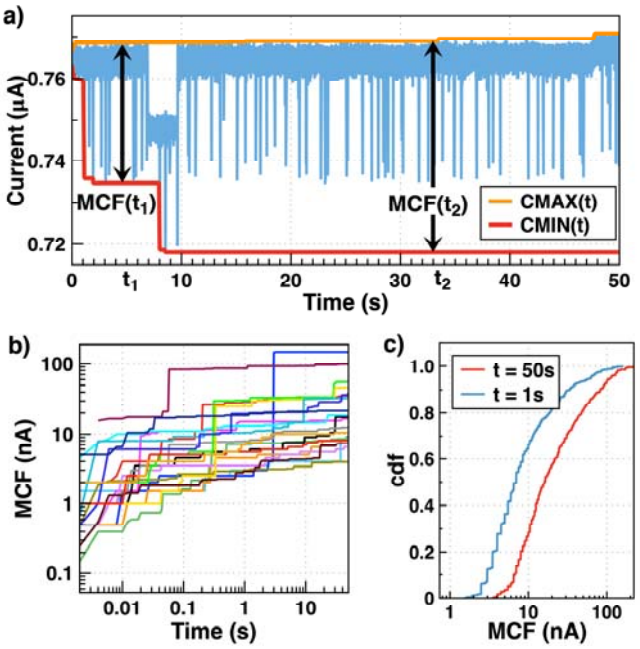

Fig. 12. a) Experimental current trace displaying RTN with its *MCF* at two different times indicated, b) *MCF* evolution for twenty transistors and c) the cdf of the *MCF* values for two time instants for 500 transistors.

Fig. 13 shows the cdf of the *MCF* values at 10s (a) and 100s (b) after the start of the measurement obtained from the processing of the 392 devices measured at $|V_{gs}| = 1.2$V and $|V_{ds}| = 0.1$V in the two scenarios described in this paper. It can be observed that, although the biasing conditions during the measurement window are the same, the cdf of the *MCF* values differs significantly depending on the time that the devices have been biased before the measurement starts. In particular, the *MCF* values are in general larger when no biasing has been applied until the measurement starts, i.e., there is more current fluctuation in the measurements from the first scenario than in those from the second one. This divergence is caused by the differences between the maps of the probability of occupations, as shown in Section II.

B. Impact on parameter extraction

In this subsection, the impact on the extracted RTN parameters, obtained using the extraction method in [19], will be studied. That method relates the MCF at any time instant with the number of active defects (i.e., defects that suffer at least one trapping/detrapping event up to that time) and the amplitude shifts induced by such defects, as well as, the

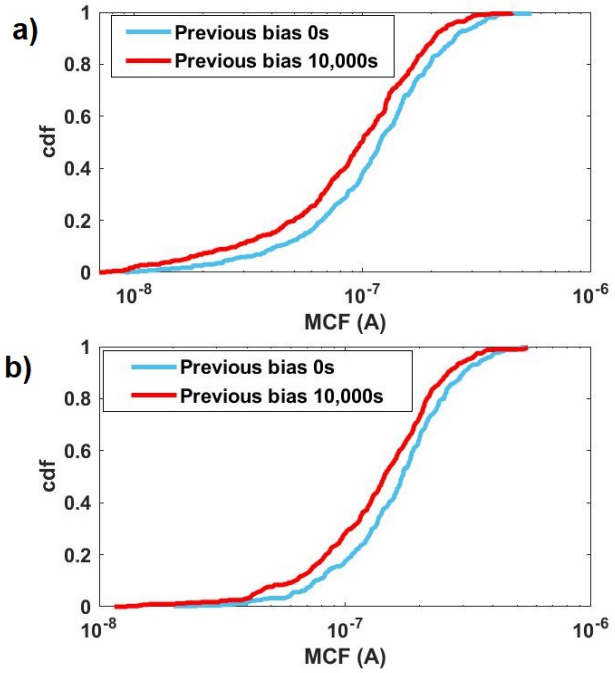

Fig. 13. Cumulative distribution functions of the *MCF* values for both scenarios after (a) 10s and (b) 100s of measurement.

background noise. The extraction procedure determines the parameters of the distribution functions of amplitude shifts, number of active defects and background noise that better fit the cdf of the MCF at different time instants.

For the amplitude shifts, a two-lognormal distribution is assumed [19]:

$$\text{cdf}(\delta I) = \frac{K}{2}\left[1 + erf\left(\frac{\log(\delta I) - \mu_l}{\sigma_l\sqrt{2}}\right)\right] + \frac{(1-K)}{2}\left[1 + erf\left(\frac{\log(\delta I) - \mu_u}{\sigma_u\sqrt{2}}\right)\right] \quad (10)$$

where *erf()* is the error function, μ_l, μ_u, σ_l and σ_u, represent the mean and standard deviation of the lower and upper lognormal function and K represents the relative amplitude of both functions. For the background noise, a Gaussian distribution centered at 0 and standard deviation σ_{NOISE} is assumed and for the number of defects the typical Poisson distribution is assumed. The mean number of active defects is assumed to follow a lognormal distribution with time [19]:

$$< \text{N} > (\text{t}) = \frac{N_0}{2}\left[1 + erf\left(\frac{\log(t) - \mu_N}{\sigma_N\sqrt{2}}\right)\right] \quad (11)$$

The parameters of all these functions were determined by fitting the cdf of the MCF at 22 different time instants and using as experimental input data the traces of the 392 PMOS devices measured at $|V_{gs}| = 1.2\text{V}$ and $|V_{ds}| = 0.1\text{V}$ in the two scenarios described in this paper. An implementation of the Particle Swarm Optimization algorithm was used to determine the function parameters that better fit the cdf of the MCF at all 22 time instants. The amplitude shift distribution was considered to be the same for both scenarios and the impact on the number of active defects was evaluated. The fitting procedure yielded the parameters shown in Table I. It can be seen that although the biasing conditions along the measurement window are exactly the same, the parameters corresponding to the number of active defects is different in both scenarios. Fig. 14 compares the evolution of the mean number of defects in both scenarios (continuous lines for the fitted function in (11) and asterisks for the mean number of defects at the 22 time instants in which the cdf has been calculated and fitted). It can be seen that as expected from Fig. 13, concurrence of the onsets of biasing and measurement yields an overestimation of the number of active defects compared to the case in which the same biasing was established long before the measurement window.

To sum up, the results shown in this section indicate that the biasing history of the devices prior to the RTN measurement has a clear and definite impact on their behavior during the measurement window, and therefore on the results of the algorithms that extract information to characterize the RTN. In algorithms like [6]-[10], this will lead to a determination of a

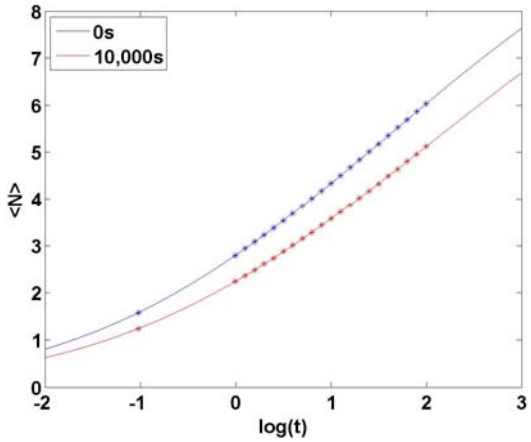

Fig. 14. Fitting of the time evolution of the mean number of active defects in both scenarios.

TABLE I. PARAMETERS EXTRACTED FOR THE δI DISTRIBUTION OF RTN DEFECTS, BACKGROUND NOISE AND THE TIME EVOLUTION OF <N>

Previous biasing	K	μ_l	σ_l	μ_u	σ_u	σ_{NOISE} (A)	N_0	μ_N	σ_N
0 s.	0.631	-7.975	0.311	-7.286	0.143	1.2e-9	10.785	1.624	2.515
10,000 s.	0.631	-7.975	0.311	-7.286	0.143	1.2e-9	10.487	2.075	2.617

different number of current levels under the same biasing conditions during the measurement in the different scenarios discussed above. This would also affect the calculation of the time constants, since a switching event may appear in one scenario and may not appear in the other scenario, or may happen at a different time. For time constants with a large number of switching events within the measurement window, the error will be small, but this error will increase as the number of switching events decreases. In techniques like [19], this could lead to a wrong determination of the number of active defects, as demonstrated with the experiment just described.

Furthermore, this finding may also determine the requirements of the experimental setup and of the characterization IC, if such IC is used. Since attaining statistically-sound RTN parameter distributions requires analyzing hundreds or thousands of devices, the experimental setup must then ensure a precise and equal biasing conditions for all the devices, and also guarantee that these conditions are applied exactly for the same time, both before and during the measurement window. Consider for example a test in which hundreds of DUTs are measured. A possibility to do this would be to bias all the devices at the same instant, and then successively measure the drain current of one device at a time. In this case, each device would experience a different biasing duration prior to its measurement, and therefore different devices would be measured in "different scenarios" and distorted parameter distributions would be extracted.

IV. Concluding remarks

In the characterization and modeling of the RTN phenomenon, like any characterization of intrinsically stochastic phenomena that claims to be sound and accurate, all devices under test should undergo exactly the same biasing conditions and these conditions should be known and perfectly controlled. Though this requirement is customary taken into account during the RTN measurement itself, the effect that the bias conditions that the device has experienced prior to this measurement is overlooked. In this work, the effect of this 'biasing history' (i.e. the biasing previous to the RTN measurement) on the RTN behavior is analyzed. This paper demonstrates, both theoretically and experimentally, that the condition of controlled biasing holds true not only during the measurement process itself, as it is generally acknowledged, but also during the time previous to the measurement.

Since the extraction of the characteristic parameters of the stochastic distribution functions (e.g., the amplitude shifts, time constants, etc.) require a statistical treatment of the current traces of hundreds or thousands of devices, it is essential that experimental setups guarantee not only the same biasing conditions for all devices, but also that these conditions are applied exactly for the same time, both before and during the measurement window. Furthermore, it is essential that the extraction algorithms that fit distribution functions of the different RTN characteristic parameters are aware of the historical biasing conditions during the lab experiments so that they can be mathematically accounted for.

V. References

[1] N. Tega et al., "Increasing threshold voltage variation due to random telegraph noise in FETs as gate lengths scale to 20 nm", in Symposium on VLSI Technology, IEEE, pp. 50, 2009.

[2] N. Tega et al., "Impact of threshold voltage fluctuation due to random telegraph noise on scaled-down SRAM", in Reliability Physics Symposium (IRPS), IEEE International, pp. 541-546, 2008.

[3] M. Luo, R. Wang, S. Guo, J. Wang, J. Zou, R. Huang, "Impacts of random telegraph noise (RTN) on digital circuits", IEEE Transactions on Electron Devices, vol. 62, pp. 1725-1732, 2015.

[4] T. Grasser et al., "Switching oxide traps as the missing link between negative bias temperature instability and random telegraph noise," in International Electron Devices Meeting (IEDM), IEEE, pp. 1-4, 2009.

[5] M. Simicic et al., "Understanding the impact of time-dependent random variability on analog ICs: from single transistor measurements to circuit simulations," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 27, no. 3, pp. 601-610, March 2019.

[6] A. Konczakowska, J. Cichosz and A. Szewczyk, "A new method for RTS noise of semiconductor device identification," IEEE Trans. Instrumentation and Meas., vol. 57, no. 6, pp. 1199-1206, June 2008.

[7] T. Nagumo, K. Takeuchi, T. Hase, and Y. Hayashi, "Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps," in Int. Electron Devices Meeting (IEDM), IEEE, pp. 28.3. 1-28.3. 4, 2010.

[8] S. Realov and K. L. Shepard, "Analysis of random telegraph noise in 45-nm CMOS using on-chip characterization system," IEEE Transactions on Electron Devices, vol. 60, pp. 1716-1722, 2013.

[9] J. Martin-Martinez, J. Diaz, R. Rodriguez, M. Nafria, and X. Aymerich, "New weighted time lag method for the analysis of random telegraph signals," IEEE Electron Device Letters, vol. 35, pp. 479-481, 2014.

[10] P. Saraza-Canflanca et al., "A robust and automated methodology for the analsis of time-dependent variability at transistor level," in Integration, the VLSI Journal, vol. 72, pp. 13-20, 2020.

[11] P. Saraza-Canflanca et al., "TiDeVa: a toolbox for the automated and robust analysis of Time-Dependent Variability at transistor level", in Int. Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), IEEE, pp. 197-200, 2019.

[12] G. Kapila, V. Reddy, "Impact of sampling rate on RTN time constant extraction and its implications on bias dependence and trap spectroscopy", IEEE Trans. Device and Materials Rel., vol. 14, no. 2, pp. 616-622, 2014.

[13] F. M. Puglisi and P. Pavan, "Guidelines for a reliable analysis of random telegraph noise in electronic devices", IEEE Trans. on Instrumentation and Measurement, vol. 65, no. 6, pp. 1435-1442, 2016.

[14] K. Ito et al., "Modeling of Random Telegraph Noise under circuit operation — Simulation and measurement of RTN-induced delay fluctuation", in Int. Symp. on Quality Electronic Design, pp. 1-6, 2011.

[15] P. Saraza-Canflanca et al., "A detailed study of the gate/drain voltage dependence of RTN in bulk pMOS transistors", Microelectronic Engineering, vol. 215, 2019.

[16] J. Diaz-Fortuny et al., "A versatile CMOS transistor array IC for the statistical characterization of time-zero variability, RTN, BTI and HCI", IEEE Journal of Solid-State Circuits, vol. 54, no 2, pp. 476-488, 2018.

[17] J. Diaz-Fortuny et al., "Flexible setup for the measurement of CMOS time-dependent variability with array-based integrated circuits", IEEE Trans. on Instrumentation and Measurement, vol. 69, no. 3, pp. 853-864, 2019.

[18] R. Wang et al., "Too noisy at the bottom?-Random Telegraph Noise (RTN) in advanced logic devices and circuits", in International Electron Devices Meeting (IEDM), IEEE, pp. 17-2, 2018.

[19] P. Saraza-Canflanca *et al*., "Statistical characterization of time-dependent variability defects using the maximum current fluctuation," IEEE Transactions on Electron Devices, vol. 68, no. 8, pp. 4039-4044, Aug. 2021.