Statistical Characterization of BTI and RTN using Integrated pMOS Arrays

B. Stampfer[°,†], M. Simicic[•], P. Weckx[•], A. Abbasi[°], B. Kaczer[•], T. Grasser[*], and M. Waltl[°]

[°]Christian Doppler Laboratory for Single-Defect Spectroscopy at the

[*]Institute for Microelectronics, TU Wien, Gusshausstraße 27-29, 1040 Vienna, Austria

[•]imec, Kapeldreef 75, 3001 Leuven, Belgium

[†]E-Mail: stampfer@iue.tuwien.ac.at

***Abstract*—To study charge trapping kinetics of oxide and interface defects, BTI and RTN measurements are typically performed. However, characterizing and investigating a statistically meaningful set of single defects is time consuming and inefficient at the single device level. To mitigate this, here we employ integrated arrays of nano-scale devices and characterize several thousands of single devices on a custom-made chip. We extract defect statistics from the threshold voltage shifts arising from single defects in individual transistors using the defect-centric approach. Finally, we also perform TCAD simulations to replicate the measurements and verify the array measurement scheme.**

***Index Terms*—Bias temperature instability (BTI), random telegraph noise (RTN), single defects, array chip, defect centric, non-radiative multi phonon (NMP) model, technology computer aided design (TCAD)**

I. Introduction

With the continued scaling of MOS transistors the influence of single defects on device performance has increased [1]. These defects, located in the gate oxide or at the interface between the bulk material and the gate dielectric, are the root cause of multiple effects negatively impacting the stable operation of MOSFETs. As a consequence, proper characterization of single defect behavior is crucial for understanding the reliability of modern devices [2]–[4]. In this work we focus on the characterization of bias temperature instabilities (BTI) [5] and random telegraph noise (RTN) [6]. On small devices these effects are often studied at a single defect level, which enables precise monitoring of the charge transition kinetics. However, given its time-consuming nature, only a limited number of defects can be assessed. In order to draw conclusions from a statistically meaningful dataset we carefully analyze distributions in ΔV_{th} from measurements performed on more than 3000 individual devices. From our experiments we extract the average ΔV_{th} caused by single defects as well as and the average number of active defects in each device, by explaining the measured distributions with theoretical distributions using the defect centric approach [7], [8]. This approach allows to extract parameters of single defects as obtained from measurements on nano-scale devices, while at the same time a large number of defects can be studied to achieve meaningful statistics, which typically requires large area devices, albeit in an averaging manner.

The research leading to this work has received funding from the Take-off program of the Austrian Research Promotion Agency FFG (projects no. 861022 and 867414).

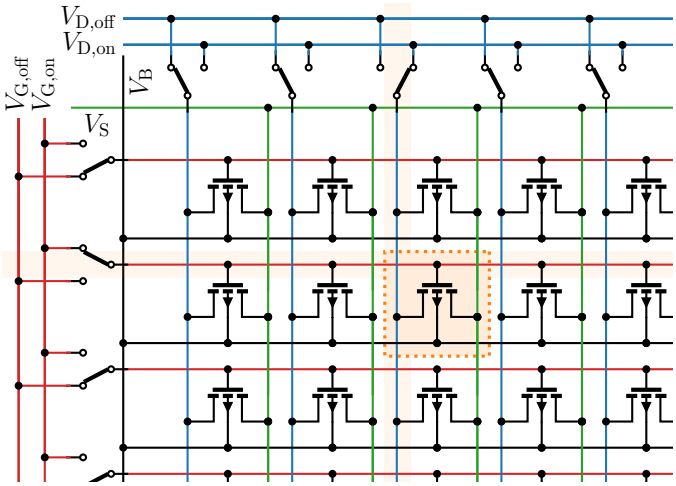

Figure 1. Layout of the signal lines of the array used for the defect characterization. The gate terminals of the transistors in each row can be switched between externally supplied on- or off-biases using on-chip logic. Likewise, the drain terminals can be switched for each transistor column. This allows to address and thus characterize each individual device in the array. The bulk and source terminals are common for all devices. More details about the array structures can be found in [9].

II. Devices and Measurements

The devices under test (DUTs) are HKMG (High-k metal gate) planar pMOSFETs with $W = 100\,\mathrm{nm}$ and $L = 30\,\mathrm{nm}$ (*short devices*) and $L = 150\,\mathrm{nm}$ (*long devices*). The pMOSFETs are organized in an array structure [9] providing over 3000 individually addressable devices per geometry. The gate and drain connections form the rows and columns of the array and can be switched electrically by double transmission gates controlled by on-chip logic, as shown in Figure 1.

Devices which are not selected are supplied with a gate off-bias of $0.15\,\mathrm{V}$ and a drain off-bias of $0.0\,\mathrm{V}$ to reduce their cumulative leakage currents. All measurements are performed sequentially by a custom built defect probing instrument, which also controls the device selection.

978-1-7281-2203-8/19/$31.00 ©2019 IEEE

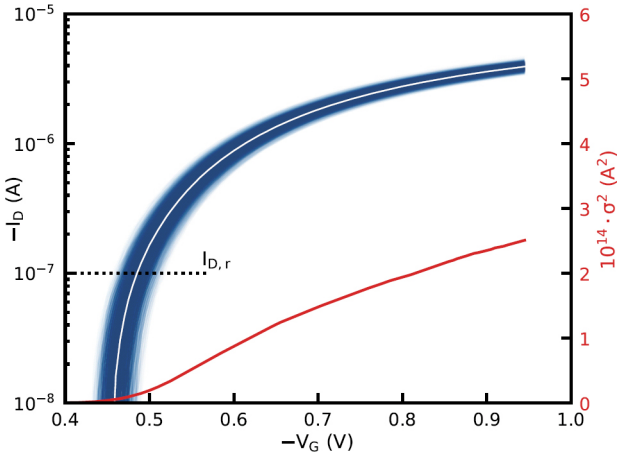

Figure 2. $I_D(V_G)$ curves of all long devices recorded prior to stress, used to map the recorded drain currents to ΔV_{th}. The white line represents the mean value of I_D, while the red line depicts the variance among the recorded curves. The dashed line shows the targeted relaxation current ($I_{D,r}$). Relaxation measurements were recorded at constant voltages corresponding to this current.

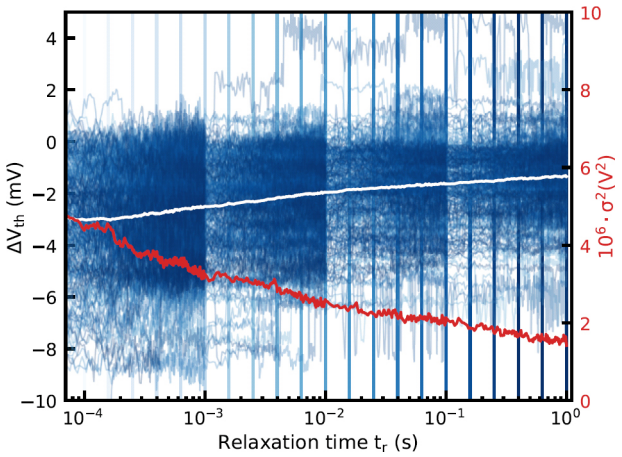

Figure 3. Recorded drain current data from 3066 devices mapped to ΔV_{th} for an exemplary set of measurements. Momentary distributions of ΔV_{th} are drawn from the set at specific points in relaxation time, as indicated by the vertical lines, to extract defect parameters. The red line depicts the variance and the white line the average among the recorded traces.

In total, 39 sets of stress-recovery measurements are performed at different gate, drain and bulk biases. Each set is repeated for five stress times $t_s = \{2, 10, 100, 1000, 10000\}$ms, with a relaxation time of $t_r = 1$ s. The relaxation traces are recorded with 200 samples per decade in time and with the first sample taken at $t_r = 100\,\mu$s after stress. All measurements are performed slightly above room temperature at 35 °C.

Prior to each stress phase, $I_D(V_G)$ curves are recorded for each device, as shown for one set of measurements in Figure 2. It should be noted that characterization of the off-currents is not directly possible in these structures, as all off-currents of devices in the column with active drain contacts contribute to the measurement. This, however, does not preclude defect characterization as the relaxation currents of the DUTs are measured in the sub-threshold regime at a relaxation voltage around $V_G = -0.5$ V (set to correspond to -10^{-7} A in the virgin $I_D(V_G)$), where the currents significantly exceed the leakage currents and the extracted ΔV_{th} is not affected.

After each measurement sequence, the drain currents recorded during the relaxation phase are mapped to ΔV_{th} employing the respective virgin $I_D(V_G)$, see Figure 3. The sets of recorded ΔV_{th} curves then provide the distribution of ΔV_{th} at any moment in relaxation time. Although the average drift of ΔV_{th} is similar compared to large area counterparts, a significant device-to-device variability can be observed for scaled devices. As can be seen from the recovery traces shown in the figure, the data recorded exhibit visible noise, which would be challenging if individual traces were to be characterized for RTN. Additional effort might have to be spent on shielding and external noise reduction to make individual characterization possible, which is not required for the statistical characterization performed here.

To evaluate a possible fabrication related inhomogeneity of device performance, the distribution of ΔV_{th} over the transistor array after 10 s of stress at -1.45 V is given in Figure 4a. In Figure 4b, the same data is averaged over slices of the array to visualize spatial variation in device performance over the array area. It can be seen that the degradation among the devices is homogeneous over the entire area, and no formation of clusters indicating unusual device behavior can be observed.

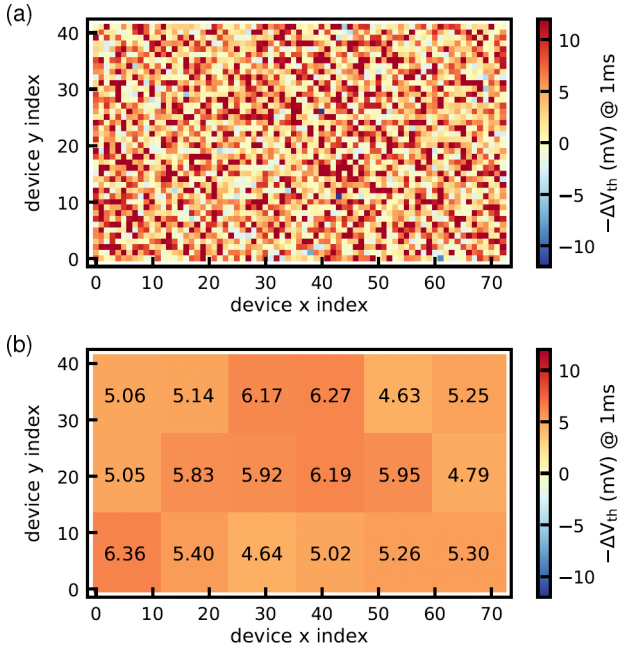

Figure 4. **(top):** ΔV_{th} extracted after $t_r = 1$ ms shown for all short devices as aligned on the chip, measured after 10 s stress at -1.45 V. **(bottom):** Average ΔV_{th} (in mV) calculated over slices of the array. The plots indicate that device degradation is homogeneous over the array area.

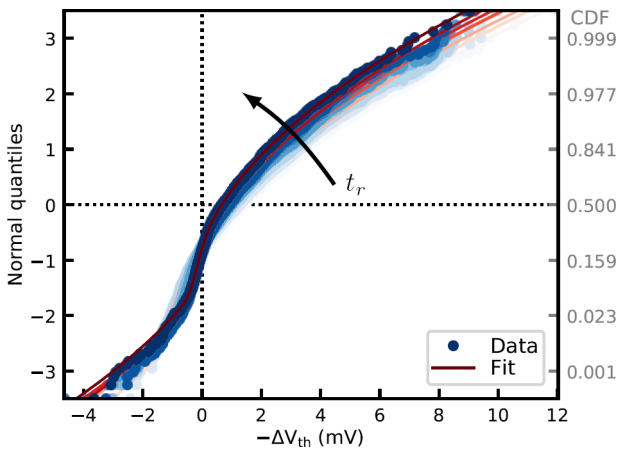

Figure 5. CDFs (points) of ΔV_{th} at a number of points in time during relaxation, for all long devices, after 10 ms of stress at $-1.45\,\mathrm{V}$, with fitted theoretical distributions (lines). The shapes of the recorded data agree well with the theoretical CDFs.

III. Defect Characterization

Cumulative density functions (CDFs) of ΔV_{th} extracted from the relaxation measurements can be seen in Figure 5. To explain the shape of these distributions, one has to consider the drift of the threshold voltage of a single device, which is modeled as the sum of three individual contributions [10], [11]:

- Positive shifts of (-)ΔV_{th} due to BTI defects which captured a charge during stress but have not yet emitted it
- Positive and negative shifts due to RTN-active defects, which captured or emitted between the initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ and the point in relaxation time which is evaluated
- Normally distributed shifts due to Gaussian noise on the recorded data

Each of these contributions can be described statistically, and the total ΔV_{th} distribution, as obtained from the measurement, follows from the convolution of the individual contributions.

Assuming BTI as the dominant effect, a simple analytical method to extract the contribution can be used to obtain the expected number of defects N and average step height η from the moments of the distribution [7], [8]:

$$N = \frac{\langle \Delta V_{\mathrm{th}} \rangle}{\eta} \tag{1}$$

$$\eta = \frac{\sigma^2}{2\,\langle \Delta V_{\mathrm{th}} \rangle} \tag{2}$$

In this work, however, we aim to characterize both BTI and RTN. For this, we fitted the total theoretical CDF to the CDF obtained from the measured data.

As outlined above, the PDF of ΔV_{th} ($p(\Delta V_{\mathrm{th}})$) consists of contributions due to BTI, RTN and Gaussian measurement noise and can be obtained by convolving their individual PDFs (p) [10], [11]:

$$\begin{aligned} p(\Delta V_{\mathrm{th}}) = {} & p_{\mathrm{Discharge}}(\Delta V_{\mathrm{th}}|N,\eta) \\ & * p_{\mathrm{RTN}}(\Delta V_{\mathrm{th}}|N_{\mathrm{RTN}},\eta) \\ & * p_{\mathrm{Noise}}(\Delta V_{\mathrm{th}}|m,\sigma) \end{aligned} \tag{3}$$

The PDF due to BTI results from a Poisson distributed number of independent defects per device, each with an exponentially distributed effect on ΔV_{th}. Mathematically, this can be described with a sum of Poisson-weighted Gamma distributions (γ), with the expected number of active defects (N) and their average effect on $\Delta V_{\mathrm{th}}(\eta)$ as parameters. $\mathcal{P}_N$ is the Poisson distribution with mean N.

$$p_{\mathrm{Discharge}}(\Delta V_{\mathrm{th}}|N,\eta) = \sum_{k=0}^{\infty} \mathcal{P}_N(k)\gamma(k,\Delta V_{\mathrm{th}}/\eta) \tag{4}$$

RTN, generally thought to be caused by the same kind of defects which are responsible for BTI [12], is described similarly. The average effect on ΔV_{th} is the same as for BTI, while the differences are the number of RTN active defects (N_{RTN}) and the fact that RTN can cause both positive and negative contributions to the measured ΔV_{th}.

$$\begin{aligned} p_{\mathrm{RTN}}(\Delta V_{\mathrm{th}}|N_{\mathrm{RTN}},\eta) = {} & \sum_{k=0}^{\infty} \mathcal{P}_{N_{\mathrm{RTN}}/2}(k)\gamma(k,+\Delta V_{\mathrm{th}}/\eta) \\ & * \sum_{k=0}^{\infty} \mathcal{P}_{N_{\mathrm{RTN}}/2}(k)\gamma(k,-\Delta V_{\mathrm{th}}/\eta) \end{aligned} \tag{5}$$

Finally, measurement noise (and drift if necessary) can be described by a normal distribution with its mean and variance parameters.

$$p_{\mathrm{Noise}}(\Delta V_{\mathrm{th}}|m,\sigma) = \mathcal{N}(m,\sigma^2) \tag{6}$$

To obtain the CDF either the PDF or any of the components of the convolution may be integrated. This finally allows for an extraction of the parameters describing the CDF, the average number of defects exhibiting BTI and RTN (N, N_{RTN}), the average step height (η) and Gaussian noise parameters (m,σ), from the experimental CDFs of ΔV_{th}.

IV. Results

An overview of the extracted defect parameters for a relaxation time of 100 ms from our measurements performed on the short devices is shown in Figure 6. It can be seen that both the step height and the RTN activity is largely independent of the measurement conditions.

In contrast, the average number of defects exhibiting BTI activity strongly depends on the stress time and bias conditions. Measurement sets C, D, E, and J with $V_{\mathrm{G}} = -1.3\,\mathrm{V}$ and $V_{\mathrm{D}} = 0.0\,\mathrm{V}$, $-0.15\,\mathrm{V}$, $-0.25\,\mathrm{V}$ and $-0.45\,\mathrm{V}$ indicate that the number of defects charged during stress is largely independent of the drain bias. The measurement data further indicate that the influence of both gate and bulk bias have a similar effect on degradation, as can be observed for example

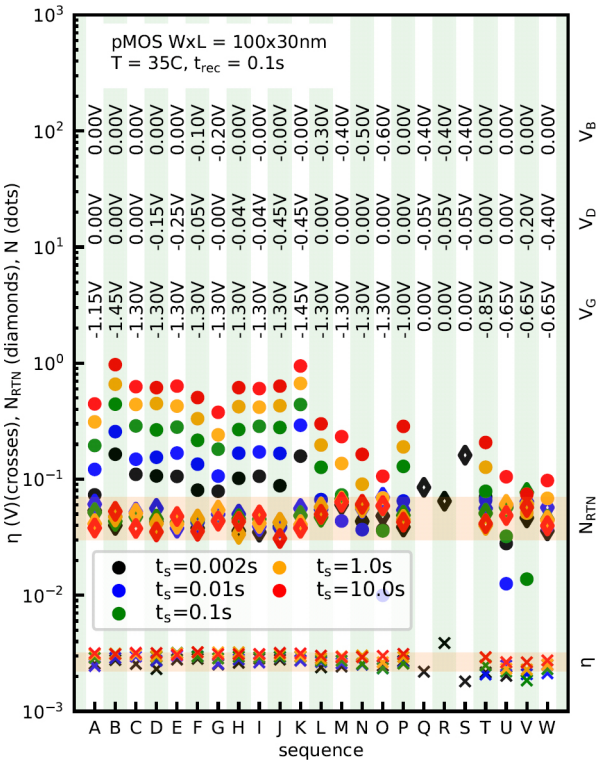

Figure 6. Overview of the extraction results for all measurement sets performed on the short devices, at a relaxation time of 0.1 s. The extracted average step heights (η, ≈3 mV) and the average number of RTN active defects (N_{RTN}, ≈4e-2) are largely independent of the bias conditions applied during stress (V_{G}, V_{D}, V_{B}).

in L ($V_{\mathrm{G}} = -1.3\,\mathrm{V}, V_{\mathrm{B}} = -0.3\,\mathrm{V}$) and P ($V_{\mathrm{G}} = -1.0\,\mathrm{V}$). Measurement sets Q, R and S were performed at very low V_{GB} and did not result in any meaningful degradation, which is why no proper fits could be obtained for them.

The gate bias dependence of the number of charged defects is shown in detail in Figure 7 for a recovery time of 2 ms – shortly after stress release. It should be noted that N is often below one in the results shown, indicating that many of the devices do not contain defects which capture at these measurement conditions.

To verify the array measurement scheme, numerical physics based device simulations are performed. For this we use the open-source compact physical simulator *Comphy* [13], which models the charge trapping kinetics of single defects using the non-radiative multi phonon (NMP) model, see Figure 8.

In the two-state NMP model used for the simulations, defects are modeled using a Markov chain with a charged and an uncharged state. The transition rates from one of the states to the other are calculated from the number of available start and end states for the carrier, a WKB tunneling factor, and an factor accounting for the energy barrier necessary to be overcome to change the defect configuration. This energy barrier is calculated by determining the intersection point between two potential energy surfaces for the charged and discharged state of the defect. The potential energy surfaces

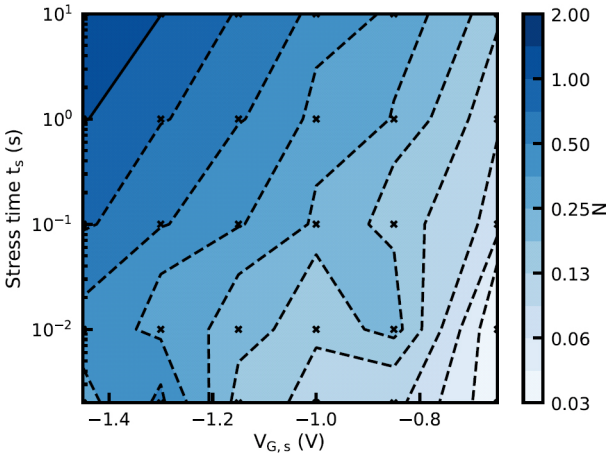

Figure 7. Average number of charged defects (N) measured on the short devices and extracted at $t_{\mathrm{r}} = 2$ ms after stress release, over stress time and gate bias. Crosses indicate the stress conditions at which sets of measurements are performed and the colors indicate regions with similar degradation, separated by contour lines. Both stress time and absolute gate bias increases the average number of charged defects. For combinations of low gate bias and stress time, values far below unity are observed, indicating that in most of the devices no defect captured at these conditions.

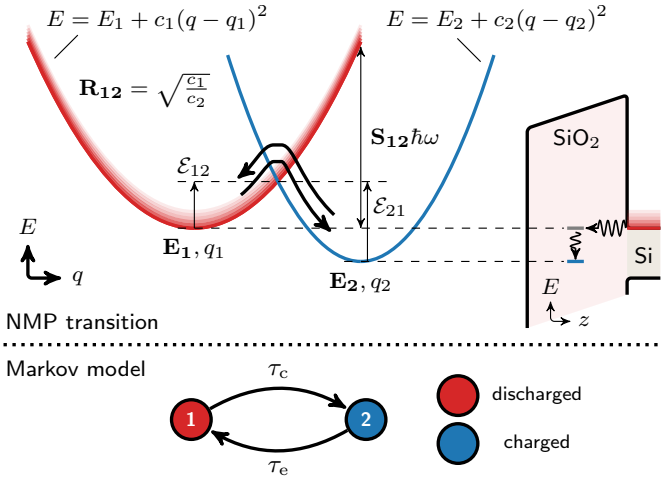

Figure 8. Potential energy surfaces for the two-state NMP model used in the simulations. For the defect to capture a charge, the carrier has to tunnel from the oxide to the defect, while at the same time the defect has to be excited by phonons to overcome a barrier (defined by $\mathbf{R_{12}}, \mathbf{S_{12}}, \mathbf{E_1}, \mathbf{E_2}$) between the two defect configurations. The energetic offset between the two parabolas changes with the applied gate bias as the alignment between the band and the defect energy levels changes.

are approximated as parabolic in this model and described using their curvatures [14]. The gate bias dependence as seen in the measurements is captured in this model by the relative shift the potential energy surfaces experience as the effective energy of the defect in the oxide changes relative to the channel carrier energy when the gate bias is changed. Using the obtained transition rates, the simulator then calculates the occupation for each simulated defect. Finally, from the occupations the threshold voltage shift can be calculated using the charge sheet approximation.

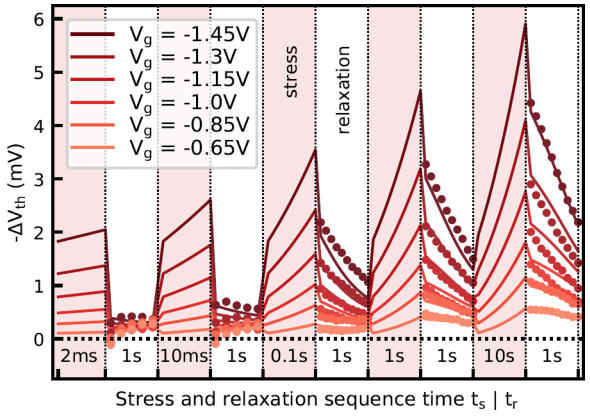

Figure 9. Simulated ΔV_{th} (lines) for measurement sequences at different gate stress biases ($V_{\mathrm{D}} = V_{\mathrm{B}} = 0\,\mathrm{V}$), compared to the respective measurement results averaged over all devices in the array (dots). The SiO_2 and HfO_2 trap bands used are taken from [13] where long term measurements on larger devices of the same technology have been calibrated. Overall, the simulations show excellent agreement with the measurements.

Results comparing the average degradation obtained during the measure-stress-measure sequences for multiple stress voltages are shown in Figure 9. The simulations have been performed using the same defect bands as previously extracted on large area devices in [13] for the same technology. It can be seen that there is good agreement between the measurement and simulation results. For the shortest two stress times and low gate biases the measurements show negative spikes of degradation shortly after the switch to recovery. Further testing needs to be done to verify whether this behavior arises from the devices themselves or from the employed array measurement scheme.

V. Conclusions

Stress-relaxation measurements employing integrated arrays containing small pMOSFETs in combination with the defect centric methodology allow extracting parameters for single defects, while at the same time obtaining meaningful statistics for a variety of stress conditions. The array measurement scheme enables extensive characterization of a large number of devices with little user interaction and minimal use of measurement equipment. With the defect centric model, the statistical effect of the defects on measurements among the whole set of transistors can be explained. Using the model, the statistical properties of the defects and their behavior can be extracted, thus allowing defect parameter extraction from large amounts of data. BTI activity in the tested devices is heavily affected by gate and bulk stress conditions while RTN activity and average step heights are not affected, consistent with [15]. The NMP model is used to describe the behavior of the individual defects and their charge transition rates. TCAD simulations using the NMP model agree with the measurements and validate the array measurement scheme.

References

[1] A. Ghetti, C. Compagnoni, A. Spinelli, and A. Visconti, "Comprehensive analysis of random telegraph noise instability and its scaling in deca–nanometer flash memories," *IEEE Transactions on Electron Devices*, vol. 56, no. 8, pp. 1746–1752, 2009.

[2] M. Toledano-Luque, B. Kaczer, P. Roussel, M. Cho, T. Grasser, and G. Groeseneken, "Temperature dependence of the emission and capture times of SiON individual traps after positive bias temperature stress," in *Book of Abstracts of Workshop on Dielectrics in Microelectronics (WODIM)*, pp. 1–2, 2010.

[3] J. Franco, B. Kaczer, M. Toledano-Luque, P. J. Roussel, J. Mitard, L.-Å. Ragnarsson, L. Witters, T. Chiarella, M. Togo, N. Horiguchi, *et al.*, "Impact of single charged gate oxide defects on the performance and scaling of nanoscaled FETs," in *2012 IEEE International Reliability Physics Symposium (IRPS)*, pp. 5A–4, IEEE, 2012.

[4] G. Rzepa, M. Waltl, W. Goes, B. Kaczer, and T. Grasser, "Microscopic oxide defects causing BTI, RTN, and SILC on high-k FinFETs," in *Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 144–147, 2015.

[5] J. H. Stathis and S. Zafar, "The negative bias temperature instability in MOS devices: A review," *Microelectronics Reliability*, vol. 46, no. 2-4, pp. 270–286, 2006.

[6] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, "Discrete resistance switching in submicrometer silicon inversion layers: Individual interface traps and low-frequency ($\frac{1}{f}$) noise," *Phys. Rev. Lett.*, vol. 52, pp. 228–231, Jan 1984.

[7] B. Kaczer, T. Grasser, P. J. Roussel, J. Franco, R. Degraeve, L. . Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger, "Origin of NBTI variability in deeply scaled pFETs," in *2010 IEEE International Reliability Physics Symposium*, pp. 26–32, May 2010.

[8] B. Kaczer, J. Franco, P. J. Roussel, G. Groeseneken, T. Chiarella, N. Horiguchi, and T. Grasser, "Extraction of the random component of time-dependent variability using matched pairs," *IEEE Electron Device Letters*, vol. 36, pp. 300–302, April 2015.

[9] M. Simicic, V. Putcha, B. Parvais, P. Weckx, B. Kaczer, G. Groeseneken, G. Gielen, D. Linten, and A. Thean, "Advanced MOSFET variability and reliability characterization array," in *2015 IEEE International Integrated Reliability Workshop (IIRW)*, pp. 73–76, IEEE, 2015.

[10] P. Weckx, B. Kaczer, J. Franco, P. J. Roussel, E. Bury, A. Subirats, G. Groeseneken, F. Catthoor, D. Linten, P. Raghavan, and A. Thean, "Defect-centric perspective of combined BTI and RTN time-dependent variability," in *2015 IEEE International Integrated Reliability Workshop (IIRW)*, pp. 21–28, Oct 2015.

[11] P. Weckx, B. Kaczer, C. Chen, J. Franco, E. Bury, K. Chanda, J. Watt, P. J. Roussel, F. Catthoor, and G. Groeseneken, "Characterization of time-dependent variability using 32k transistor arrays in an advanced HK/MG technology," in *2015 IEEE International Reliability Physics Symposium*, pp. 3B.1.1–3B.1.6, April 2015.

[12] T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, and B. Kaczer, "A unified perspective of RTN and BTI," in *2014 IEEE International Reliability Physics Symposium*, pp. 4A.5.1–4A.5.7, June 2014.

[13] G. Rzepa, J. Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl, P. J. Roussel, D. Linten, B. Kaczer, , and T. Grasser, "Comphy — a compact-physics framework for unified modeling of BTI," *Microelectronics Reliability*, vol. 85, pp. 49–65, 2018.

[14] W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. Shluger, and T. Grasser, "Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence," *Microelectronics Reliability*, vol. 87, pp. 286 – 320, 2018.

[15] A. Kerber, "Impact of RTN on stochastic BTI degradation in scaled metal gate/high-k CMOS technologies," in *2015 IEEE International Reliability Physics Symposium*, pp. 3B–3, IEEE, 2015.