Advanced Data Analysis Algorithms for the Time-Dependent Defect Spectroscopy of NBTI

Michael Waltl, Paul-Jürgen Wagner*, Hans Reisinger°, Karina Rott° and Tibor Grasser**

*Institute for Microelectronics, Technische Universität Wien, Gusshausstrasse 27-29/E360, Austria
Phone: +43-1-58801/36050, Fax: +43-1-58801/36099, Email: waltl@iue.tuwien.ac.at

°Infineon Munich, Am Campeon 1-12, 85579 Neubiberg, Germany

***Abstract*—In order to identify the physical mechanisms behind the negative bias temperature instability (NBTI), the time-dependent defect spectroscopy (TDDS) has been recently proposed. The TDDS takes advantage of the fact that in nano-scaled devices only a handful of defects are present. As a consequence, degradation and recovery proceed in discrete steps, each of them corresponding to a charge capture or emission event. By repeatedly applying stress and recovery conditions, the TDDS analyzes the statistical properties of these discrete events. The measurement window of the TDDS is very large, but the occurrence of random telegraph noise (RTN) at certain biases/temperatures can limit its applicability. We have developed an advanced data analysis method which can also deal with data contaminated by RTN. The algorithm is based on the combination of a bootstrapping technique and cumulative sum charts. A benefit of the new method is the possibility to detect steps in a large class of different signals with a feasible amount of parameters. Moreover, de-/trapping parameters of the random telegraph noise (RTN) become accessible as well.**

I. Introduction

In modern ultra-scaled MOSFETs the negative bias temperature instability (NBTI) has become a major reliability issue [1–3] and is particularly observed in pMOS devices when stressed with negative gate voltages at elevated temperatures. In general, the degradation due to NBTI lowers the drain current and shifts the threshold voltage.

To study the recovery behavior of the devices after NBTI stress, the time-dependent defect spectroscopy (TDDS) has been recently introduced [4]. The TDDS exploits the discrete nature of degradation and recovery visible in nano-scaled devices. Therefore a fast measurement setup is necessary to record the recovery behavior of the devices after negative gate bias stress [5].

The detrapping events, visible in the recovery measurement data as discrete steps at time instants τ_e with step height d, are shown in the (τ_e, d) plane. We call such a graph a "spectral map", see Figure 1. To obtained sufficient statistics around 100 measurement/stress/measurement recording transitions have to be evaluated for each bias/temperature condition.

In such a spectral map the detrapping events from a number of relaxation traces for a constant stress time t_s, ambient temperature T, and stress voltage V_s, form clusters, which are the fingerprints of individual defects contributing to the recovery. In the case of RTN ($\tau_c \approx \tau_e$) two bands symetrically arranged around the abscissa are visible in the spectral maps. By repeating the experiment under different conditions, $\bar{\tau}_e$ and d of a handful of defects can be simultaneously extracted as a function of gate and/or drain voltage, and the ambient temperature.

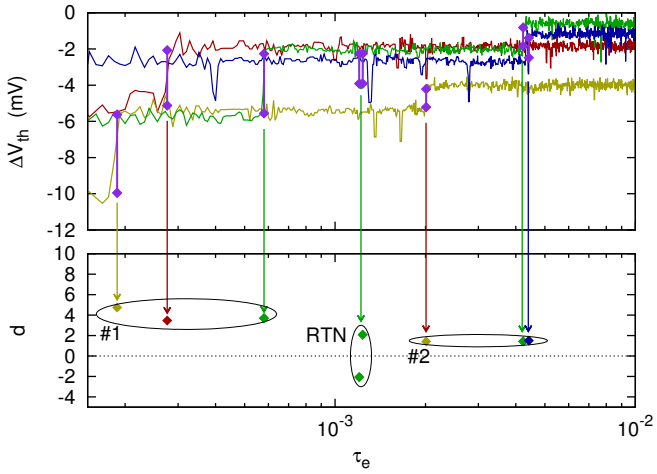

Fig. 1: Mapping of the single detrapping events from four recorded recovery traces of a PMOS device after NBTI stress (top) into the (τ_e, d) plane, called spectral map (bottom). From the discrete steps in the recovery traces the transition time instances, which are the defect emission times τ_e, and the step heights are extracted and contribute to single points in the resulting spectral map. After collecting the steps of a large number of traces, for example 100, clusters in the spectral maps are obtained. The defects #1 and #2 mark two individual defects with an emission time in the TDDS measurement window. Morevoer, a defect producing RTN leads to data points symetrically arrange arround the abscissa in the spectral map. Each cluster is the fingerprint of a single defect.

As long as the device-under-test contains only defects for which $\bar{\tau}_e \gg \bar{\tau}_c$ holds, the TDDS data is straight-forward to analyze. The intensity of the marked clusters reflect the probability of a defect to relax within the observation time window and also gives information about the defect capture probability $P_c(t_s)$ at a certain stress time t_s. In general $\bar{\tau}_e \gg \bar{\tau}_c$ gives a single detrapping event for each defect in a relaxation trace. The capture probability then follows an exponential law

$$P_c(t_s) = A(1 - \exp(-t_s/\bar{\tau}_c)) \quad (1)$$

with A the occupancy, t_s the stress time, and $\bar{\tau}_c$ the average capture time of the single defect [3, 4, 6].

By carefully selecting the device bias and temperature, the above condition can be usually satisfied, c.f. Figure 3. However, under certain bias or temperature conditions, $\bar{\tau}_e \approx \bar{\tau}_c$ for a certain defect, resulting in the occurrence of random

978-1-4673-2752-7/12/$31.00 ©2012 IEEE

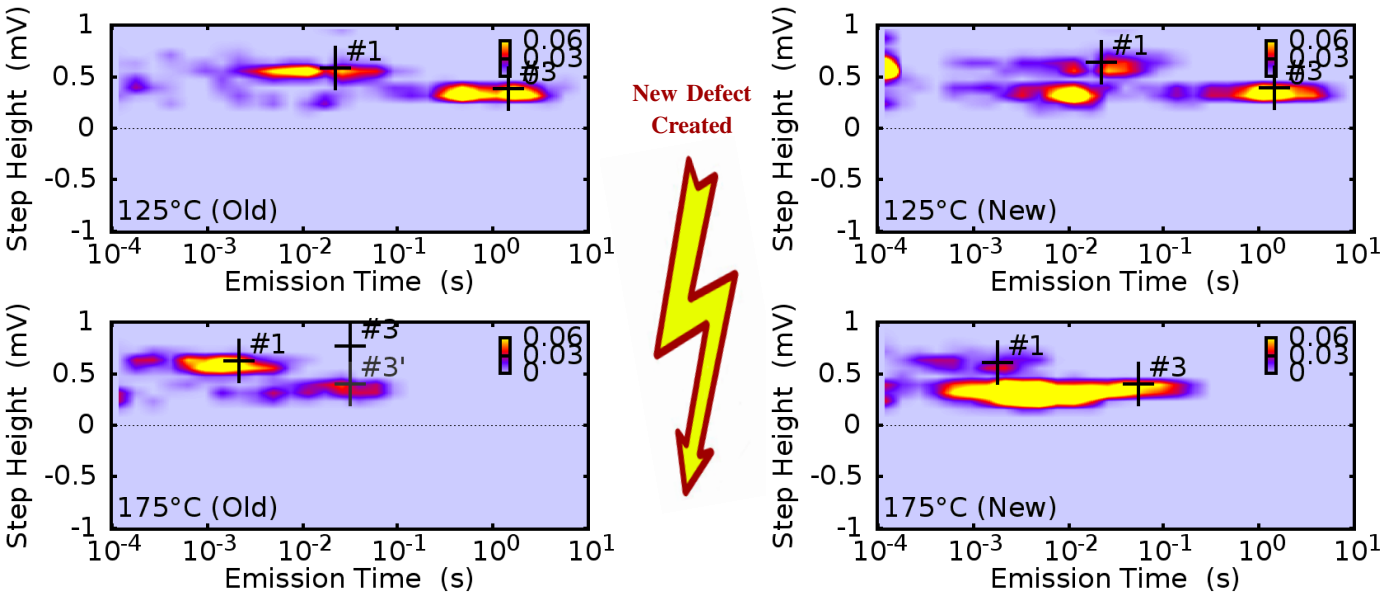

Fig. 2: Comparison of typical spectral maps obtained from past measurements in 2010 (left) and the latest ones with a new defect producing RTN (right) for 100 traces. The defects from previously recorded data result in clean clusters. In contrast, the new defect produces RTN visible as a band between the defects #1 and #3. In the case of the measurement data at $T = 175°$ (bottom right) the previously obtained defects #1 and #3 are visibly blurred and joined due to the new RTN defect. As a consequence, a new advanced algorithm is necessary to analyze the recovery behavior of our 'Golden Device'. Moreover, the clusters move to smaller τ_e with increased temperature ($T = 125°$ top spectral maps and $T = 175°$ bottom spectral maps). When varying the stress time the clusters remain in their positions with increased intensity.

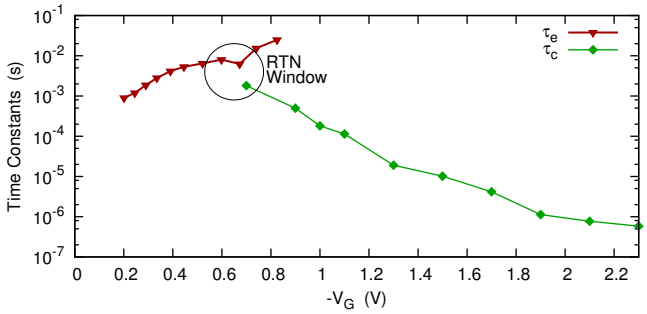

Fig. 3: The capture times $\bar{\tau}_c$ typically decrease with increased magnitude of $V_G = V_s$. In the case of $\bar{\tau}_e \approx \bar{\tau}_c$ the defect produces RTN and the simple extraction algorithm fails. The above shows the data extracted for defect #1.

telegraph noise (RTN), which then severely contaminates the data.

Due to the strong bias- and temperature-dependencies of $\bar{\tau}_e$ and $\bar{\tau}_c$, the occurrence of this scenario is difficult to predict while pre-screening the device. Even worse, new defects can be created during the measurements. In a particularly unfortunate case, the 'Golden Device' of our previous studies [4, 7], which we had been analyzing for over three years, suddenly showed a defect producing RTN, thereby obscuring the analysis of defects #1 and #3, see Figure 2. In order to salvage the data and continue our analysis, more sophisticated data analysis algorithms are required.

Moreover, the measurement window of the TDDS can get very large. Defects show emission time contants from the μs regime, directly after stress release, up to weeks, month or even years. A measurment window spread over several decades in time goes hand in hand with the necessity of finding a trade-off between the sample intervals and the amount of measurement data recorded. An adjustment of the sampling time intervals to higher values for larger recovery times is therefore necessary. The Fourier transform scaling property reads [8]

$$r(at) \overset{\mathcal{F}}{\Longleftrightarrow} \frac{1}{a} R\left(\frac{\omega}{a}\right) \tag{2}$$

where a compression in the time domain is transfered into a dilatation in the frequency domain and vice versa, $\mathcal{F}$ denotes the Fourier transform operator and $R(\omega) = \mathcal{F}\{r(t)\}$ is the Fourier transform of the time signal $r(t)$. An on-the-fly adjustment of the sampling rate directly decreases/increases the noise power of the measurement signal when the sampling rate is decreased or increased, respectively. It is therefore necessary that the step detection of the advanced algorithm can operate on non-uniformly sampled measurement data as well.

To detect fast RTN signals with $\tau_c \approx \tau_e$ together with changes in the sampling rates, very short pulses, have to be detected.

In the following the evaluation of heavily contaminated data for the analysis of hole capture events for single defects and fast carrier capture and emission events stemming from RTN is presented.

II. The Advanced Detection Algorithm

The step detection algorithm has to extract the unknown step-heights d given an unknown mean value of the signal, $\mu(t_r)$. Based on the mean shift model given by

$r = \mu + d\sigma(t_\mathrm{r} - \tau_\mathrm{e})$ with the step function σ and the emission time τ_e, a statistical treatment of this problem is possible using the bootstrapping and cumulative sum (BCSUM) algorithm [9].

The BCSUM detection algorithm has the ability of detecting either positive or negative steps in a data set. From the BCSUM procedure the decision that either a change point in the underlying data samples exists or not is obtained by selection of a detection sensitivity parameter ϵ. In the following the cumulative sum charts [10, 11] are discussed, the bootstrapping mechanism is discussed and finally the advanced detection method is presented.

A. Cumulative Sum Charts

First consider a series of independent samples $\mathbf{r} = [r_0, r_1, \ldots, r_{\mathrm{N}-1}]$ with N the number of samples. The signal distributions before and after a change in the signal mean are given by $p_0(r)$ and $p_1(r)$, respectively, with their corresponding means μ_0 and $\mu_1 = \mu_0 + d$, where d is the step height. Introducing the log-likelihood ratio given by

$$s_i = \ln \frac{p_1(r_i)}{p_0(r_i)} \tag{3}$$

a change in the signal mean is reflected by a change of the sign of the mean value of the log-likelihood ratio S_m^n [11] of

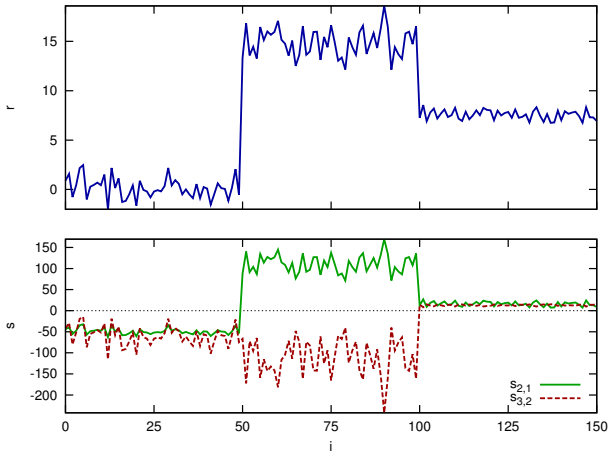

Fig. 4: The log-likelihood ratio (bottom) is calulated for Gaussian distributed data for $N_1(0,1)$, $N_2(15,1.5)$ and $N_2(7.5,0.5)$ (top). When the test signal r changes its underlying distribution function from N_1 to N_2 the resulting log-likelihood ratio $s_{2,1} = \ln p_2(r)/p_1(r)$, whereas p_1 is the Gaussian distribution function for N_1 and p_2 the Gaussian distribution function for N_2, shows a change in the sign at the transition time instance, called change point. When the underlying data distribution of r goes from N_2 to N_3 a sign change is observed in the log-likelihood ration $s_{3,2} = \ln p_3(r)/p_2(r)$.

the data set $[r_\mathrm{m}, \ldots, r_\mathrm{n}]$ given by

$$S_m^n = \sum_{i=m}^{n} s_i. \tag{4}$$

This is illustrated in Figure 4.

Together with the assumption of Gaussian distributed samples with the probability density function

$$p_{\mu,\sigma}(r_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(r_i-\mu)^2}{2\sigma^2}\right), \tag{5}$$

the log-likelihood ratio for $d = \mu_1 - \mu_0 > 0$ and $\sigma = \sigma_0 = \sigma_1$ is

$$s_i^{\uparrow} = A\left(r_i - \mu - \frac{d}{2}\right) \quad \text{with} \quad A = \frac{\mu_1 - \mu_0}{\sigma^2}, \tag{6}$$

and the change point position is at the sample with index

$$i_{\tau_\mathrm{e}}^{\uparrow} = \{k\colon \min_{0<k<N-1} S_0^k\}. \tag{7}$$

In the case of trapping events, a negative mean shift is obtained and the log-likelihood ratio for $d < 0$ reads

$$s_i^{\downarrow} = A\left(r_i + \mu - \frac{d}{2}\right) \quad \text{with} \quad A = \frac{\mu_1 - \mu_0}{\sigma^2}. \tag{8}$$

with the change point at the sample index position

$$i_{\tau_\mathrm{e}}^{\downarrow} = \{k\colon \max_{0<k<N-1} S_0^k\} \tag{9}$$

Since the change of the mean is unknown the detection of either positive or negative changes has to be provided by the algorithm. Combining both cases, $d > 0$ and $d < 0$, the log-likelihood ratio for Gaussian distributed samples can be formulated as

$$s_i = r_i - \nu \tag{10}$$

where the threshold value ν already considers the prefactor A. With the choice of $A = 1$ the CSUM chart function is

$$\mathcal{C} = \sum_N r_i - \nu \quad \text{with} \quad \nu = \frac{1}{N}\sum_N r_i \tag{11}$$

The sample index of the change point is given by

$$i_{\tau_\mathrm{e}} = \{k\colon \max\left(|i_{\tau_\mathrm{e}}^{\downarrow}|, |i_{\tau_\mathrm{e}}^{\uparrow}|\right)\} = \{k\colon \max\left(|\mathcal{C}|\right)\} \tag{12}$$

B. Bootstrapping

A frequently used method for parameter estimation of a set of data samples for unkown underlying distributions is bootstrapping. Its big advantage is that it is fully automatic and it does not matter how complicated the mathematical model for the probability distribution is.

For a given set of samples $\mathbf{r}$ of length N a bootstrap sample $\mathbf{r}^*$ is obtained by randomly choosing N samples from $\mathbf{r}$ with replacement. Since bootstrapping is a resample technique with replacement, the samples out of $\mathbf{r}$ can occur never, once or more often in the bootstrap estimate $\mathbf{r}^*$ [12].

The statistical nature of bootstrapping includes the necessity of a huge number B of bootstrap samples $\mathbf{r}_b^*$. This circumstance inevitably leads to computationally expensive procedures.

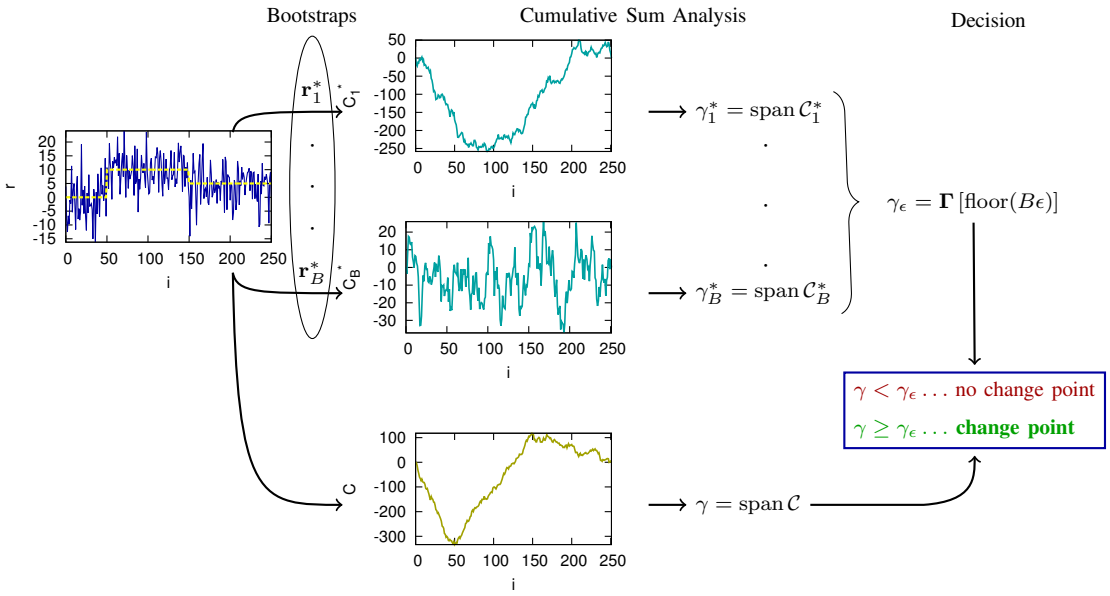

Fig. 5: Illustration of the bootstrap and cumulative sum algorithm to detect one step in a very noisy data set. Even though the data set contains two steps the algorithm has to be applied recursively because at most one, that is, the biggest change point is detected. Based on an initial data set, B bootstraps $\mathbf{r}_\mathrm{b}^*$, that are resampled data sets with replacement of the initial data set $\mathbf{r}$, are calculated. Afterwards, the cumulative sum charts of the bootstrapped and the original data are created. The spans of the single CSUM charts of the bootstrapped data are collected by a sorted list, from the minimum to the maximum, $\mathbf{\Gamma}$. The decision threshold γ_ϵ equals the element of $\mathbf{\Gamma}$ with index $i = B\epsilon$, with $\epsilon \in [0, 1]$ called the detection sensitivity. Finally, the decision threshold γ_ϵ is compared to the span of the CSUM chart obtained from the initial data set γ. For $\gamma > \gamma_\epsilon$ the initial data set contains a change point. In that case the initial data set has to be split at the change point and the algorithm is applied recursively to both sequences. This devide and conquer procedure is continued until no further change points larger than a threshold are detected in the subsequences.

C. The Algorithm

The BCSUM algorithm is a Monte Carlo based combination of bootstraps and cumulative sum charts to detect changes in the signal mean. The divide and conquer procedure to detect one change point in the measurement data set $\mathbf{r}$ is:

1) Calculate the B bootstrap samples $\mathbf{r}_1^*, \mathbf{r}_2^*, \ldots \mathbf{r}_B^*$ from $\mathbf{r}$.
2) Create the cumulative sum charts $\mathcal{C}_b^*$ for each bootstrap sample $\mathbf{r}_b^*$ with $b \in [1, B]$.
3) Estimate the spans of each chart $\mathcal{C}_b^*$ and collect them into a sorted list from the minimum to the maximum $\mathbf{\Gamma}$.
4) With the detection sensitivity ϵ the decision threshold is obtained as $\gamma_\epsilon = \mathbf{\Gamma}[\mathrm{floor}(B\epsilon)]$, where $[\cdot]$ denotes the list index operator. For a data set $\mathbf{r} = [r_0, r_1, \ldots, r_{N-1}]$ the list index operator is

$$\mathbf{r}[i] = r_\mathrm{i}. \tag{13}$$

5) Create the CSUM chart $\mathcal{C}$ from the initial data set and evaluate $\gamma = \operatorname{span}\mathcal{C}$
6) A change point in the underlying sequence occurs if $\gamma_\epsilon < \gamma$. Split the data set at the change point position and recursively continue with 1) for each subsequence until all change points are detected.

Critical parameters for the detection outcome of the BCSUM algorithm is the number of bootstraps B and the detection sensitivity ϵ. For statistical relevance, the bootstrapping and the calculation of $\mathcal{C}_b^*$ has to be performed many times, $B = 10^5$ is recommended.

A choice of $\epsilon \approx 1$ means that just the most dominant changes in mean are detected. For very fast events, which often just span two or three measurement data samples, the detection sensitivity is a crucial parameter. To extract small and very fast changes $\epsilon \approx 0.6$ is recommended.

Moreover, the position of the change point in the measurement data is directly accessible via the CSUM chart obtained from the initial data series.

Last, but not least it has to be noted that the algorithm can be apply to uniform and non-uniformly sampled data as well. This property is very importand in the context of the TDDS. To cover several decades in time a compromise between the sampling rate and the accumulated recovery time has to found in order to achieve a feasable amount of measurement data.

III. Results

The benefit of the new BCSUM algorithm is the possibility to consider trapping events (negative steps) occurring in the case of defects producing RTN. For the measurements recorded on a pMOSFET with a 2.2 nm thick oxide stressed at $V_\mathrm{s} = -1.3$ V and $T = 175°\mathrm{C}$, as depicted in Figure 6, the spectral maps are contaminated with RTN. To remove the RTN from the spectral maps, the detection of positive and negative mean shifts is necessary. When a simple step detection algorithm is applied to the signal from Figure 7, the first event is detected at $\tau_{\mathrm{e},1} = 1.43$ ms with $d_1 = 3.73$ mV followed by an event at $\tau_{\mathrm{e},2} = 3.19$ ms with $d_2 = 3.26$ mV. Obviously, the event at $(\tau_{\mathrm{e},1}, d_1)$ is due to RTN and so should not contribute to the statistics for the defects #1 or #3.

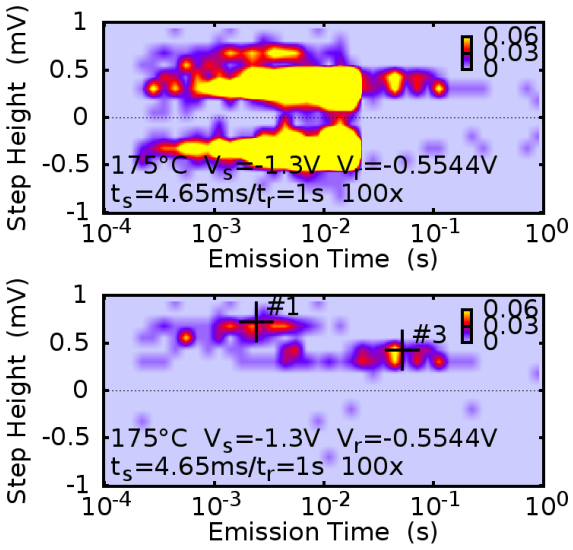

Fig. 6: Example for the improved quality of the spectral maps for 100 traces. After detection of the trapping and detrapping events with the BCSUM algorithm the full map (left) is obtained with blurred defects #1 and #3. After the noise stemming from the new RTN defect is removed, the denoised spectral map with separated defects #1 and #3 is obtained (right).

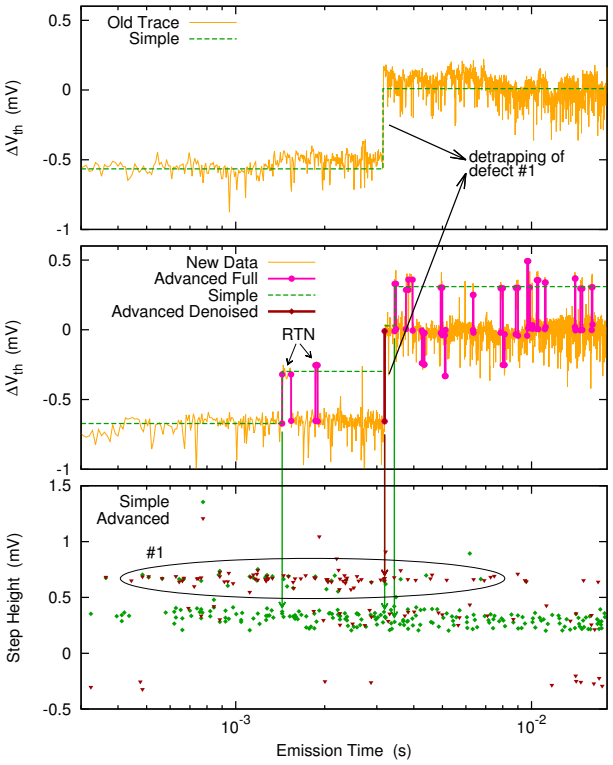

Fig. 7: Over the last three years the defects of a single device have been studied. An example for a previously recorded trace (top) show the detrapping event, in this case for defect #1, which is clearly visible and not contaminated with RTN. However, recent measurements are strongly influenced by a new defect producing RTN (middle). The so far used simple algorithm just detects three positive changes because only positive steps are considered (middle, *Simple*). The BCSUM method additionally detects negative steps stemming from RTN (middle, *New Full*). In the spectral map (below), here depicted for 100 traces, the steps detected with the simple method (green) show step heights below the step level for the defect #1 which equals the step heights produced by an RTN defect. Furthermore, the large step is extracted with a wrong amplitude and so it does not contribute to the defect #1 in the spectral map. The BCSUM algorithm extracts the amplitudes corresponding to the defect #1 correctly.

Moreover, the step at $(\tau_{e,2}, d_2)$ is a detrapping event attributed to #1, but is extracted with a spuriously reduced step height d_2 because negative mean shifts have not been considered so far. With the advanced BCSUM algorithm all trapping and detrapping events are detected. As such, the events according to #1 or #3 and the events stemming from the RTN can easily be separated and the correct step heights are obtained. The influence of RTN events in the measurement data is shown in Figure 7. Especially for the defect #1, lower step heights have been extracted so far. The fitted capture time $\bar{\tau}_c$ does not show a significant difference, cf. Figure 8, because the first detrapping event attributed to RTN occurred at time instances in the range of the ones for defect #1. However, there is a large difference in the number of extracted events, the simple approach considered two steps instead of one, and there is also a significant variation in the steps size of both methods.

The utility of detecting and considering RTN events becomes visible when the full spectral maps, cf. Figure 6 (top), and the denoised spectral maps, cf. Figure 6 (bottom), are compared. As RTN can be described as a series of steps with alternating sign and only very small variations in the step amplitudes, RTN results in two symmetric extended clusters in the spectral maps. Furthermore, the full spectral map containing all events shows the clusters of the defects #1 and #3 blurred and joined due to the new RTN defect. After denoising the spectral maps, the clusters for the defect #1 and #3 become clearly visible again, cf. Figure 6 (bottom), and nicely agree with the previously extracted data.

IV. Conclusions

An advanced data extraction method for the time-dependent defect spectroscopy has been developed which allows for denoising of the spectral maps. The method is validated using our 'Golden Device' which has been under study for more than three years but became unusable due to the recent appearance of a new RTN defect. With the advanced method, the contaminated data can be fully analyzed.

The application of the BCSUM algorithm to the measurement data evaluation in context with the TDDS is a benchmark for the algorithm. The application field of the advanced data analysis algorithm can be expanded to any case where an extraction of transitions between discrete levels in heavily noisy measurement data is necessary.

V. Acknowledgment

The research leading to these results has received funding from the Austria Science Fond (FWF) project n°P23390-N24.

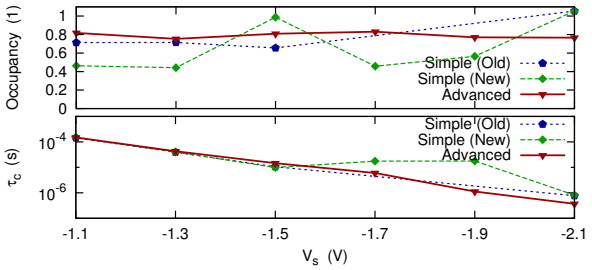

Fig. 8: The difference between both step detection methods is reflected in the occupancy extracted from the spectral maps [13]. For the defect #1 a higher occupancy is obtained from the BCSUM algorithm compared to the simple approach applied to the new measurement data because detrapping amplitudes corresponding to the defect #1 are incorrectly detected if an RTN appears before the defect #1 captures. As a result of the BCSUM algorithm more defects now contribute to the defect #1 and therefore the occupancy is increased. The denoised occupancy agrees well with the the simple analysis of the previously recorded data. The capture times are extracted from the clusters peak values for different stress times by fitting an exponential law for the capture probability $P_{\mathrm{c}}(t_{\mathrm{s}})$. Because the RTN occurs in the same time range as the emission time τ_{e} for the defect #1, the extracted capture times τ_{c} do not show a significant difference between the simple and advanced method.

REFERENCES

[1] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *Journal of Applied Physics*, vol. 94, no. 1, pp. 1–18, July 2003.

[2] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability*, vol. 46, no. 1, pp. 1–23, 2006.

[3] T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectronics Reliability*, vol. 52, pp. 39–70, 2012.

[4] T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, "The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability," in *Proceedings of the 2010 International Reliability Physics Symposium*, May 2010, pp. 16–25.

[5] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI degradation- and recovery-behavior based on ultra fast VT-measurements," in *Proceedings of the 2006 International Reliability Physics Symposium*, March 2006, pp. 448–453.

[6] T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. Luque, and M. Nelhiebel, "The paradigm shift in understanding the bias temperature instability: From reaction-diffusion to switching oxide traps," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3652 –3666, nov. 2011.

[7] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors," *Physical Review B*, vol. 82, pp. 245 318–1–245 318–10, 2010.

[8] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, 2000.

[9] W. A. Taylor, "Change-point analysis: A powerful new tool for detecting changes," Online, 2000. [Online]. Available: http://www.variation.com/cpa/tech/changepoint.html

[10] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

[11] M. E. Basseville, M. Basseville, and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Online, 1993.

[12] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall, 1993.

[13] P.-J. Wagner, T. Grasser, H. Reisinger, and B. Kaczer, "Oxide traps in MOS transistors: Semi-automatic extraction of trap parameters from time dependent defect spectroscopy," in *Proceedings of the 2010 International Symposium on the Physical and Failure Analysis of Integrated Circuits*, July 2010, pp. 1–5.