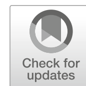
Advanced Electrical Characterization of Single Oxide Defects Utilizing Noise Signals



Bernhard Stampfer, Alexander Grill, and Michael Waltl

1 Introduction

A variety of electrical measurement techniques for investigating the functionality and performance of single semiconductor transistors has evolved over the recent decades. Among these, the simplest method is to measure an I_D-V_G characteristic of a pristine or stressed device from which important measures for the device quality like (1) the sub-threshold slope, (2) the leakage current, and (3) the on-state resistance can be extracted, as shown in Fig. 1. The change of parameters may not seem particularly critical for a single transistor, but it is of fundamental importance in the context of complex integrated circuits as any transistor operating outside of specification might lead to failure of the circuit.

Although the manufacturing processes for semiconductor structures used in modern circuits are continuously being optimized, their electrical properties still remain determined by a large number of electrically active defects. Such defects can either be located in the insulator, at the insulator–semiconductor interface, or in the semiconductor bulk material. In order to characterize the prevalent defect densities and their trap levels, measurement techniques such as capacitance voltage (CV) measurements [1], charge pumping (CP) [2, 3], deep level transient spectroscopy (DLTS) [4], direct-current current voltage (DCIV) measurements [5, 6], extended stress and recovery measurements (eMSM) [7], and noise measurements [8] are widely used. While most of these measurements require switching or sweeping

B. Stampfer (✉) · M. Waltl
Christian Doppler Laboratory for Single-Defect Spectroscopy in Semiconductor Devices at the Institute for Microelectronics, Vienna, Austria
e-mail: stampfer@iue.tuwien.ac.at; waltl@iue.tuwien.ac.at

A. Grill
Institute for Microelectronics, Vienna, Austria
e-mail: alexander.grill@imec.be

© Springer Nature Switzerland AG 2020

T. Grasser (ed.), *Noise in Nanoscale Semiconductor Devices*,
https://doi.org/10.1007/978-3-030-37500-3_7

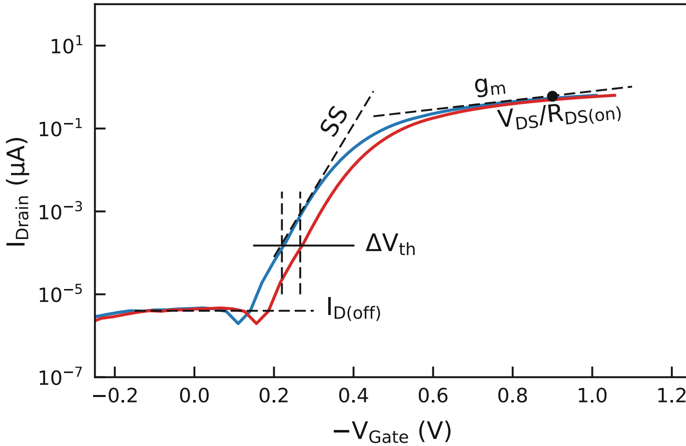


Fig. 1 Schematic of an I_D-V_G characteristics for a fresh (blue) and a stressed (red) device. Shown are the sub-threshold slope, the off-state leakage current, the on-state resistance, and threshold voltage shift

between different bias levels, noise measurements are considered an equilibrium measurement technique as the device current is recorded at constant biases.

To analyze the measurement data gathered from noise, either the time- or frequency-domain can be used. In the most general sense, the noise spectra of conventional MOS transistors typically follow a 1/f behavior, which is often referred to as flicker or pink noise. It was first proposed by McWorther that the nowadays well-known 1/f noise characteristics arise from defects in the transistors which dynamically change their charge state during operation [8]. However, at that time the feature size of transistors was too large to observe single charge transition events in the device current. With the continued decrease of the gate area of MOS devices it was observed that the 1/f noise power is roughly inversely proportional to the gate area [9]. Ralls et al. later showed that the origin for 1/f noise can be linked to individual defects modulating the resistance of the inversion channel depending on their charge state [10]. They observed that the changes in the charge state can be so large that they can be measured as discrete steps in the device current. This outstanding feature opens the avenue for device reliability characterization at the single defect level and can be directly observed on all small enough devices where only a few tens of single defects are active.

By examining the experimental data of the noise measurements on small MOS structures, one can observe many remarkable features of single defects which are not measurable on large MOS structures. The diversity of the behavior of the single defects poses a particular challenge for the development of physical models for device simulators. However, exact physical models can only be ensured if the various characteristics of individual processes can be exactly reproduced.

1.1 Random Telegraph Noise

For the most common case, an RTN signal may look as depicted in Fig. 2a. At a constant gate- and drain-source bias, the amplitude of the current through the conducting channel exhibits a step-like decrease as soon as the active defect captures a charge, and increases by the same amount as soon as the same defect emits its charge. The time period from the last discrete current decrease to the next increase is termed *charge capture time*, and the reverse is called the *charge emission time*. It has to be noted that both, the charge capture and the charge emission processes, are stochastic events. The time periods between the charge capture and successive charge emissions events follow an exponential distribution.

$$p(t) = \exp(-t/\tau_e), \tag{1}$$

with the average emission time τ_e, usually referred to simply as *emission time*. Likewise, the average time between charge emission events and successive charge capture events is called the *capture time* (τ_c), and follows an exponential distribution too. An interesting feature of the charge transition times is that the values of τ_c and τ_e depend on the applied gate bias, as shown in Fig. 2a. The average step height—which is the third parameter which can be extracted from an RTN signal—

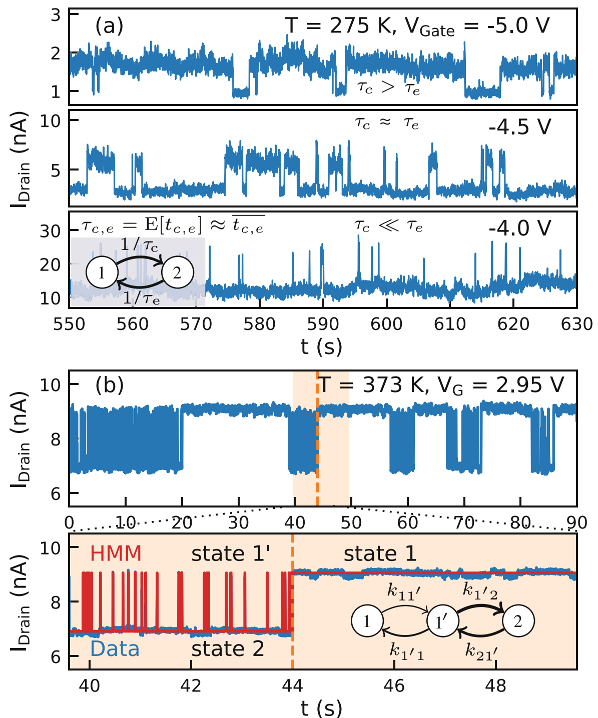


Fig. 2 Exemplary RTN traces for a simple defect (**a**) and a more complex defect (**b**). The currents shown in (**a**) decrease once the defect captures a charge and increases again once it emits. The average dwelling times of the RTN signals change with gate voltage. The more complex defect affecting the current in (**b**) shows two different capture times in its neutral state, effect termed anomalous RTN (aRTN). Markov models as shown in the insets can be used to model such defects. Modified from [11]

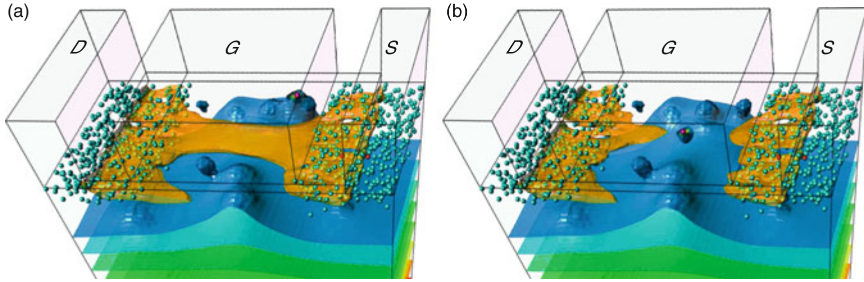


Fig. 3 TCAD simulation showing the current percolation path in a MOS transistor for two defect configurations (**a,b**). In a small-area device, a single defect changing its charge state may heavily influence the percolation path (orange) and thus the conductivity of the device. Cyan balls: source and drain donor ions, red dots: charged defects. From [12]

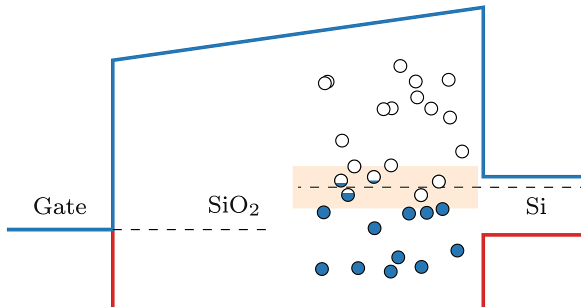


Fig. 4 Band diagram illustrating the energetical window of defects producing RTN signals at a given gate bias. Defects located far above or far below the Fermi level remain neutral or charged during the entire measurement, while defects located closer to the Fermi level have an occupancy approaching 0.5 (at $E_{\mathrm{T}} = E_{\mathrm{F}}$), and thus produce RTN

is determined by the position of the single defect with respect to the percolation path of the inversion channel (see Fig. 3). On average, the step height is observed to increase for devices with a smaller active channel area [13]. To model the defect behavior shown in Fig. 2a, a simple two-state Markov process can be used, as shown in the inset of Fig. 2a. The transition rates ($k_{\mathrm{c,e}}$) between the two states are directly related to the average charge capture and emission time by $k_{\mathrm{c,e}} = 1/\tau_{\mathrm{c,e}}$.

Defects that can be explained by two-state Markov processes are among the simplest cases. In addition, detailed RTN analysis has shown that the temporal behavior of individual defects can be very complex, as shown in Fig. 2b. Such intricate characteristics can be attributed to defects which can exist in more than one atomistic configuration at a given charge state [14]. A signal as given in Fig. 2b is typically referred to as *anomalous RTN* (aRTN). To model this behavior, an extension of the two-state Markov process to a three- or even four-state Markov chain as shown in the inset of Fig. 2b is necessary and will be discussed in more detail in Sect. 4.

The effect of RTN for a given defect is most pronounced if the defect captures and emits its charge at a similar rate (i.e., $\tau_c \approx \tau_e$), which is the case if the energetic trap level of the defect is close to the Fermi level of the channel or the gate at the given gate bias, as indicated by the shaded area in Fig. 4.

In the case of $E_T \approx E_F$, the probability of the defect being charged (referred to as *occupancy*) becomes $\approx$0.5 and thus, its average capture time equals its average emission time, which at the same time maximizes the noise power of the signal for the given defect. Changing the gate bias from this point shifts the relative position of the defect with respect to the Fermi level, which in turn impacts the defect's capture and emission rates and thus its occupancy. By varying the gate bias, the bias dependence of the charge transition times, i.e., $\tau_c(V_G)$ and $\tau_e(V_G)$, can be determined. This voltage dependence provides important information on the location and the atomistic nature of single defects.

1.2 Link Between RTN and 1/f Noise

In Fig. 5, typical noise measurements recorded from a large-area and a small-area device are compared. As stated earlier, RTN can be seen as the manifestation of 1/f noise in small devices—or likewise—1/f noise as the sum of many single RTN signals. To show this relation, the power spectral density (PSD) of a two-state RTN signal is calculated from its auto-correlation function using the Wiener–Khintchine theorem, which yields [15, 16]:

$$S_{\text{single}}(f) = \frac{(2d\tau_0)^2}{(\tau_c + \tau_e)(1 + (2\pi f \tau_0)^2)} \tag{2}$$

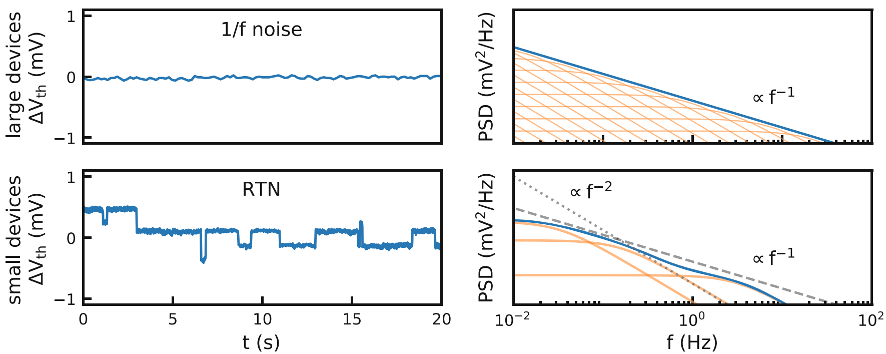


Fig. 5 Exemplary noise measurements shown for a large-area (top) and a small-area (bottom) device at static operating conditions. Illustrations of the frequency behavior are given to show the link between small and large area transistors. In large devices, many Lorentzian power spectral densities (PSDs) produced by individual defects add up to 1/f noise

with the step height d and the mean frequency

$$1/\tau_0 = 1/\tau_\mathrm{c} + 1/\tau_\mathrm{e}. \quad (3)$$

The resulting Lorentzian PSD is depicted for three defects in the bottom right plot in Fig. 5. Assuming the defects as causing only a small perturbation to device electrostatics, the noise spectral density for a number of defects is given by the superposition

$$S(f) = \sum_i \frac{(2d_i\tau_{0,i})^2}{(\tau_{\mathrm{c},i} + \tau_{\mathrm{e},i})(1 + (2\pi f\tau_{0,i})^2)}. \quad (4)$$

For a large number of defects with a uniform distribution of $\log(\tau)$, as is apparently often the case for large-area transistors, the distribution $S(f)$ converges to a 1/f shape, as indicated in Fig. 5. This can be shown by simplifying Equation 4 by replacing all defects within ranges of $\Delta\log(\tau_0)$ by equivalent defects with $\tau_\mathrm{c} = \tau_\mathrm{e} = \tau$ and the summed step heights of $d = d_\mathrm{e}$. This simplification leads to:

$$S(f) \approx \frac{d_\mathrm{e}^2}{2}\sum_i \frac{\tau_i}{1 + (\pi f\tau_i)^2} \quad (5)$$

Now, at any given frequency f, the PSD of one defect will be dominating the sum. Thus, we can simplify the equation further by considering only these dominant defects. Maximizing the inner function with τ_i as the parameter gives $\tau_m = 1/(\pi f)$, i.e., the dominant defect is the defect with its corner frequency at πf. This approximation then reveals the overall 1/f frequency dependence:

$$S(f) \approx \frac{d_\mathrm{e}^2\tau_m}{4} = \frac{d_e^2}{4\pi f}. \quad (6)$$

From this, one can see that RTN may be characterized both in time- and frequency-domain. In the following sections, the RTN and single defect analysis is performed using the time-domain data. Advantages of this approach include that the charge capture and emission times can be extracted separately and individually, and multi-state defects can be extracted with relative ease. Analysis using frequency-domain data is discussed in Chapter 4 of this book [17].

2 Measurements

To characterize RTN, either general purpose instruments or specialized measurement configurations as shown in Fig. 6 are commonly used. In the first configuration depicted in Fig. 6a, the biases applied at the gate and drain terminals of the transistor

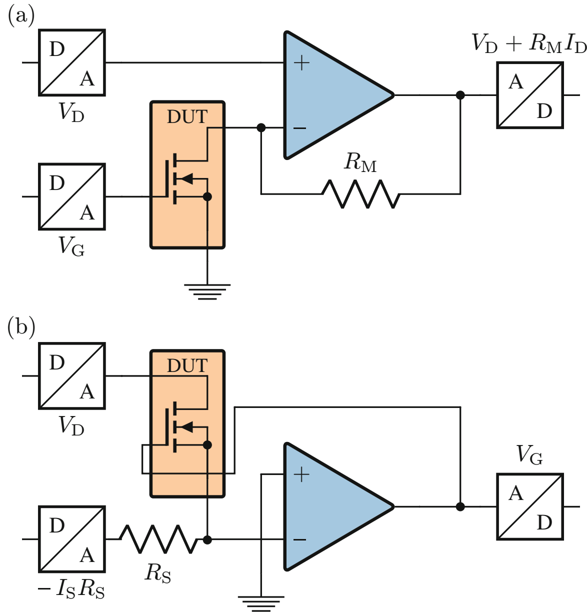


Fig. 6 Simplified measurement configurations used for RTN (and TDDS) measurements: (**a**) constant gate voltage or (**b**) constant source current. Practical designs will typically exhibit several feedback resistors in parallel to switch between drain current ranges, include additional passive components to adjust the bandwidth of the op-amp stage to achieve a maximum signal-to-noise ratio (SNR), and may further include secondary stage amplifiers. Notice that in the case of the constant current configuration (**b**) both the DUT and the operational amplifier feed back into each other, forming a closed loop. This may lead to instabilities if the circuit is not carefully designed

are held at a constant voltage during the experiment. If one considers an ideal operational amplifier (op-amp), with zero input currents on the positive and negative terminals ($I_+ = \mathrm{I}_- = 0\,\mathrm{A}$), infinite amplification, and zero offset voltage between the terminals ($V_{+-} = 0\,\mathrm{V}$), the current through the feedback loop equals the drain-source current, and the output voltage can be calculated as $V_{\mathrm{out}} = V_{\mathrm{D}} + R_{\mathrm{M}} I_{\mathrm{D}}$. Given a measurement circuit where these conditions can be matched closely, this voltage can be recorded and the drain current can be calculated. As the degradation of the device is typically considered in terms of an equivalent shift of the threshold voltage, the recorded drain current is subsequently mapped to a ΔV_{th} using a pre-recorded I_{D}-V_{G} characteristic [18, 19].

The second widely used configuration is shown in Fig. 6b, where the MOS transistor is operated at a constant source current. This can be achieved by placing the DUT in the feedback loop of the operational amplifier (op-amp) with the gate voltage being directly controlled by the op-amp [20]. The constant current configuration has the big advantage that the threshold voltage shift can be measured directly, which may lead to slightly different values for ΔV_{th} as compared to those measured using the first configuration [21]. A serious disadvantage of the constant

current configuration is that the transistor being part of the feedback loop can lead to oscillations. Thus, more attention has to be put on part selection and design before the scheme can be used.

One major advantage of RTN characterization is that the recording can be started at any time and does not have to occur highly synchronized with any bias switches. The sampling rate can be adjusted to the expected charge capture and emission times which allows to optimize the system for maximum signal-to-noise ratio (SNR) and highest step resolution. To maximize the usable bias range, the sampling unit recording the trace should be able to sample relatively fast, as well as to record a large number of points, which is necessary to monitor defects with significantly different charge transition times, i.e., $\tau_c \ll \tau_e$ or vice versa. A large data buffer of the sampling stage allows to measure at gate voltages farther from $E_T = E_F$, where higher ratios of τ_c/τ_e are observed. An example of a suitable sampling unit are high speed oscilloscopes which are often used in combination with custom-designed circuits [22, 23].

To estimate the capture and emission time of a defect, at least a handful of charge capture and emission events have to be visible in each measurement trace, while the sampling time (T_S) should be around two orders of magnitude [24] below the lowest time constant for reliable results. This requirement effectively reduces the measurement range by about three decades in time. To give an example, consider a measurement system with a sampling buffer size (n) of one million samples and a defect with a charge capture and emission time at the intersection point ($\tau_c = \tau_e$) of around a second. In order to obtain a large characterization range (in bias), one could choose a sampling time of around $t_S = 1$ ms, leading to a total measurement time of $nT_S = 1$ ks. This effectively allows to characterize RTN signals with transition times in the range of $\approx$100 ms to $\approx$100 s.

2.1 Number of Observations

The error made by observing a limited number of events can be predicted using a chi-square distribution, which gives the one-sigma confidence limit for the estimated time constant [25]. For 4, 10, and 100 observations (N), the measured time constant will lie approximately within a 50, 30, and 10% error margin, respectively, 68% of the time. Alternatively, one may use the width of the Gaussian distribution to obtain the relative error as $1/\sqrt{N}$, which gives values close to the errors calculated using a chi-square distribution for all but very small numbers of N. To reflect this in the design of experiment, a minimum cumulative measurement time of $t_m = N(\tau_c + \tau_e)$ should be used.

3 Defect Parameter Extraction

A number of methods have been developed to extract the properties of RTN signals [26, 27]. The method of choice depends on (1) the parameters which should be extracted, for instance, step height and/or (ratio of) transition times, (2) the number of defects visible in the measurements, and (3) their signal-to-noise ratio. In the following, the classical methods such as the histogram and time lag plot methods are discussed briefly, before more advanced methods based on edge detection and hidden Markov models (HMMs) are presented.

3.1 Histogram and Time Lag Plots

The most straightforward method for RTN parameter extraction is the histogram method [28]. For this, a histogram of the samples recorded from the drain current, or equivalent the ΔV_{th} values, is drawn in a first step, as shown in Fig. 7a. Given a large enough SNR, Gaussian peaks, which correspond to the individual charge

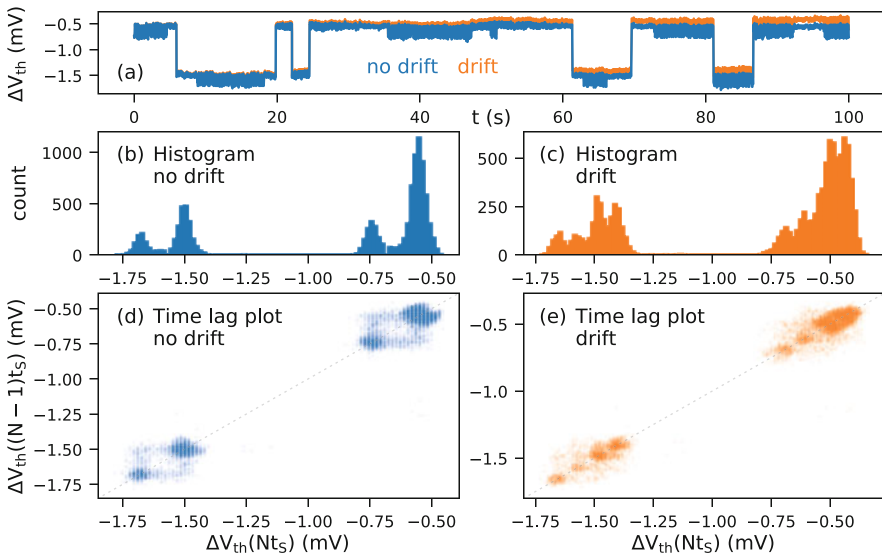


Fig. 7 Histogram (**b**) and time lag plot (**d**) of a measured RTN signal (**a**) (blue figures). The histogram exhibits two defects, one with a smaller and one with a larger step height. A clearer separation between the distinct levels can be observed for the time lag plot. Additionally, the time lag plot shows the transitions in the off-diagonals. To show the influence of drift on these methods, a linear drift of 150 μV was added to the signal (orange plots). It can be seen that both the histogram (**c**) and time lag plot (**e**) can only distinguish the larger defect with the drift present. Note that the smaller of the defects shows abnormal RTN behavior, which these methods do not reflect

states of the single defects, should become visible in the histogram. The step height of the defect can then be extracted directly from the distance between two peaks. To determine the charge transition times, threshold values can be defined to assign regions of drain current to the charge states of the defect(s). The charge transition times can then be extracted by counting and averaging the time spent in each charge state. The ratio of the charge transition times can be obtained directly from the areas of the peaks in the histogram.

An improvement of the histogram method is the time lag plot method [26], where instead of a simple histogram, a scatterplot is created from the samples, with the x- and y-values of the points defined by each two consecutive drain current steps (see Fig. 7b). This leads to a better separation of the (now two-dimensional) peaks, and thus enables a more reliable extraction of the defect parameters, especially in the case of multiple defects. In the time lag plot representation, transitions between the states show as off-diagonal clusters.

The largest drawback of the histogram and time lag plot methods is that a relatively large SNR is required for reliable assignment of drain current values to defect states, as few erroneous assignments may lead to large errors of the extracted charge transition times. This can be mitigated to some extent by the removal of outliers or circumvented by just obtaining the ratio of the transition times which will be much less affected by a small number of wrong assignments. Another disadvantage of these methods is their reliance on the absolute drain current values for any combination of charge states. Any measurement drift, low frequency noise, or slow defects which are not subject of the analysis will widen the extracted step height distributions and make the extraction unreliable. This is a serious challenge especially when long-term RTN measurements in the kilo second range or even larger have to be analyzed. A recent improvement called the weighted time lag method has been proposed by Martin-Martinez et al., it allows the extraction of RTN at higher noise levels [29]. The weighted time lag method is presented in Chapter 14 of this book [30].

3.2 Edge Detection

Especially when analyzing measurement data which is superimposed by a slow drift, it is favorable to consider the first derivative of the dataset. The most straightforward approach seems to be to directly calculate the first derivative; however, this direct approach is disadvantageous for two reasons. First, a step occurring during a sampling period might be distributed over two samples, causing a lower derivative, and second any measurement noise will fully contribute to the result. A more reliable approach is to utilize an edge detection filter such as the Canny filter, which is widely used in digital image processing [31]. This filter allows to find the positions of the discrete switching events in the trace even for signals with low SNR. Once the positions of the discrete steps are identified, the step heights can be obtained by subtracting the I_{D} values before and after the steps from the original data, or

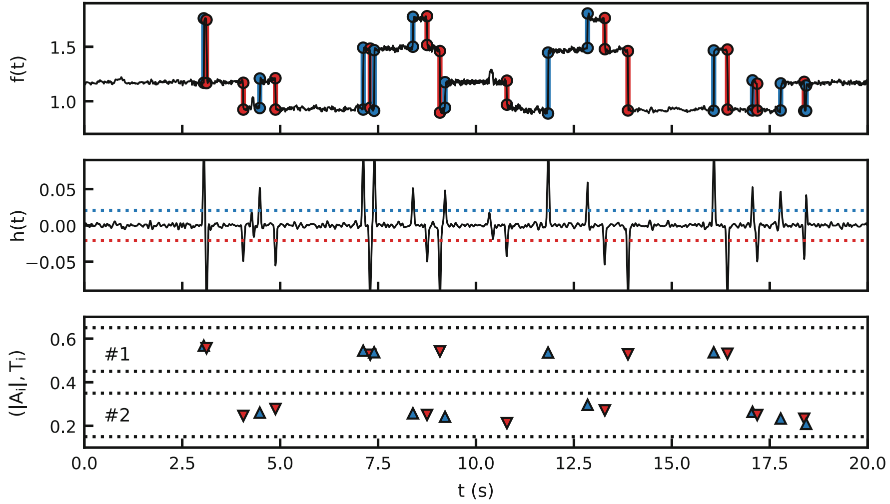


Fig. 8 Working principle of the Canny edge detection algorithm used for RTN parameter extraction, shown on an exemplary measurement trace. The original signal $f(t)$ is convoluted with the first derivative of a Gaussian to obtain the filtered signal $h(t)$. The positions of the local extrema in $h(t)$ above a certain threshold give the positions of the steps. This method is insensitive to most drifts in the measurement trace

alternatively by integrating the derivatives. Afterwards, steps with similar heights can be considered for further evaluation, and the time intervals between the step-up and step-down transitions can be analyzed. An illustration showing a trace analyzed with the Canny edge detection algorithm can be seen in Fig. 8.

To calculate the positions of the edges in the signal $f(t) = \Delta V_{\mathrm{th}}(t)$, we first calculate the derivative $\partial \Delta \mathrm{V}_{\mathrm{th}}(t)/\partial t$ and convolute it with a Gaussian kernel $g(t, \sigma)$, or equivalently, convolute the original signal with the derivate of the Gaussian kernel

$$h(t) = f(t) * \frac{\partial g(t, \sigma)}{\partial t}. \tag{7}$$

The positions of the relative maxima in the signal above a noise threshold m then give the positions of the steps:

$$\mathbf{T} = \mathbf{argrelmax}(|h(t)| > m) \tag{8}$$

With (**argrelmax**(*x*,*y*)) defined as a function yielding the *x* positions of the local maxima in *y* (c.f. *findpeaks* in matlab). The step heights can be obtained by taking the difference of the measurement values before and after the step:

$$\mathbf{A} = f(\mathbf{T} + \delta) - f(\mathbf{T} - \delta) \tag{9}$$

or alternatively by integrating over the segment in $h(t)$. Improved variants of the algorithm exist, i.e., methods which use multiple values of σ [32] or estimation of the threshold value using Otsu's method [33].

After detection, the steps can be binned by step height and finally the time differences between positive and negative steps can be measured. If only a single two-state defect is responsible for the steps in a bin, the mean values of the time differences between the charge transition events directly give the charge capture and charge emission time. If a defect with more than two states is observed, a histogram of the differences between the charge transition events may be used to fit the transition times. Similarly, if two defects fall within a single bin, a state machine might be used to assign individual steps to the defects given additional information, as shown in Fig. 9 for correlated defects. At this point, however, it might already be favorable to use the method based on hidden Markov models which is discussed next.

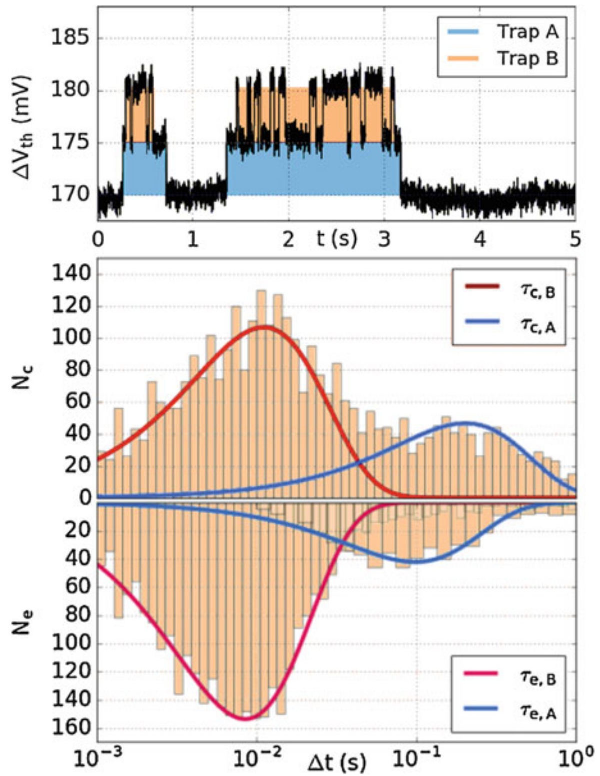


Fig. 9 Two defects with similar step heights, or the charge transition times of a single defect with more than two states, may be distinguished by fitting the exponential distributions to the differences in step-up and step-down times. The example shows steps obtained from a measurement trace with two correlated defects where defect B was only active when defect A captured charge. This feature allowed to clearly separate the extracted steps for each defect. For regular defects with similar step height, extraction using HMMs might be necessary to obtain their parameters. (Picture taken from [34])

3.3 Hidden Markov Models

The last and most sophisticated method which is presented here is based on maximum likelihood estimation using hidden Markov models (HMMs). A Markov model as described in Sect. 4 is constructed from the suspected defects visible in the measured data and its parameters are typically trained to the measured data using the Baum–Welch algorithm [35].

Given the observed data $O(t) = I_D(t)$ (or $\Delta V_{th}(t)$) and the structure of the Markov chain, we want to find the parameters of the model θ, whereby θ consists of:

- the transition matrix $\boldsymbol{K} = \{k_{ij}\}$ with the transition rates between all defect states i and j,
- the observed mean drain current $\boldsymbol{M} = \{\mu_i\}$ for each state i, and
- the standard deviation σ of the drain current noise.

The Baum–Welch algorithm is a variant of an expectation–maximization algorithm and each iteration consists of two steps:

- In the expectation step, the probabilities of being in each state for each point in time, as well as the average number of transitions between the states are calculated given the observations and the model parameters from the previous iteration.
- In the maximization step, new model parameters are calculated from these statistics and the observations to maximize the likelihood of the expectations.

Once the algorithm has converged, the parameters and observations may also be used to calculate a most likely sequence of states using the Viterbi algorithm [36].

The main advantage of the HMM method is that it is suitable for signals exhibiting a lower SNR than required for the previously presented methods. Thus, this method can be used to extract parameters from defects which are barely visible in the measurement [37]. Another advantage of this method is that it is possible to distinguish defects with similar step heights and extract defects with more than two states. The method, however, is negatively affected by any slow drift present in the measurement data, just like the histogram and time lag plot method. To mitigate this, a method for baseline correction can be implemented within the iterative Baum–Welch solver. Suitable solutions comprise spline smoothing, local regression [38], and asymmetric least squares [39]. For a large set of measurement data, this method is less suitable than, for example, the Canny edge detection method as the defects have to be defined prior to extraction, requiring manual interaction. Furthermore, the method has to be used carefully, as it regularly converges to wrong results with good scores (*likelihood*). For instance, overfitting may occur if too large a number of defects is chosen, often with better likelihood than a regular fit. It can also happen that a single real defect is fitted using multiple defects, or measurement noise is fitted with defects. The method is also computationally expensive as the computing time increases exponentially with the number of states in the Markov model. Thus,

an efficient implementation is required if traces with multiple defects should be analyzed. The basic algorithm can be found in [40], and various implementations are available, e.g., for python [41, 42]. To train a HMM to multiple defects, additional pre- and post-processing is necessary to assemble and disassemble the individual Markov chains before and after each iteration as described by Frank et al. in [43]. A refined approach to fitting multiple defects was proposed by Puglisi et al. in [44] where they used a factorial hidden Markov model (FHMM) [45] to fit the individual defects which decreases the computational complexity. This approach, instead of compiling a single Markov chain representing all defects, uses separate Markov chains for each defect and optimizes their state variables in a self-consistent manner.

4 Oxide Defect Modeling

Before discussing the identification of physical defect properties from the obtained signal parameters, we should take a closer look at the defects themselves. The main culprit giving rise to RTN (and BTI/TDDS) is thought to be defects located within the oxide of the devices. Each defect can capture or emit a charge from the channel or the gate carrier reservoirs. Such a charge transition event changes the Coulomb potential in the vicinity of the defect, and as a consequence affects the surface potential along the semiconductor/insulator interface, and further perturbs the current flux through the transistor [46]. Due to the amorphous nature of the oxide material, a number of atomistic arrangements serving as possible defect candidates exist [47]. The most likely defect candidates for SiO_2 are the hydrogen bridge and the hydroxyl E′ center, which both exhibit a number of different, but stable, atomistic configurations, as shown in Fig. 10. The neutral, puckered, and charged

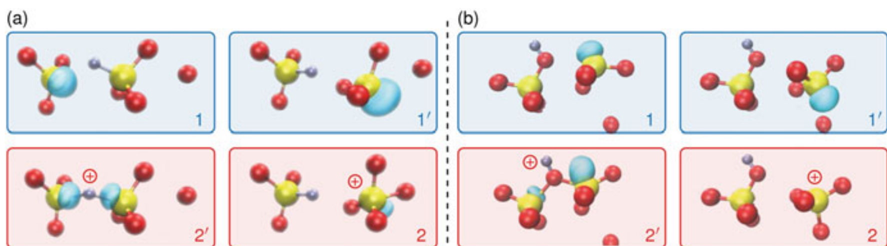


Fig. 10 The hydrogen bridge (**a**) and the hydroxyl E′ center (**b**), two possible defect candidates in SiO_2. Consistent with experimental data, these defects are able to change their atomistic configuration while maintaining the same net charge, leading to multiple states. In this figure, the yellow balls represent silicon atoms, red balls oxygen, gray balls hydrogen, and the blue bubbles represent a negative charge distribution. The numbers 1 and 2 describe the neutral and charged configurations, respectively, and a prime (′) denotes a metastable state. It can be seen that the transition between the stable and metastable states commonly involves the silicon atom transitioning through the plane spanned by three of its oxygen neighbors to enter a so-called puckered configuration [48, 49]. Adapted from [50]

configurations can be determined using computational expensive density functional theory (DFT) calculations. Recent calculations suggest that such defects might be (in)activated by the interaction with the various hydrogenous species always present in any device [50, 51].

The transition between the different states of the defects can be considered memoryless, meaning that the probability of the defect changing state depends only on the current defect state and potential distribution, and not on the previous state of the defect or the device. This is crucial for modeling as so the defects can be modeled as Markov chains where each Markov state represents one of the possible atomistic configurations. In this context, the simplest model for an oxide defect or interface state is given by a two-state Markov chain, representing the defect in its charged and neutral state. To some extent, the two-state description is sufficient to describe the charge trapping behavior observed in many RTN measurements. However, these measurements also reveal defects hibernating in one of the two charge states between phases of RTN, a behavior which is termed anomalous RTN as shown in Fig. 2b. To describe this trapping feature, a Markov chain with at least three states is required. An additional remarkable observation is that defects show a kind of volatility in TDDS measurements, meaning that they can disappear for a number of measurements, and later reappear [52]. This means that the defect becomes stuck in a more stable atomic configuration for a certain amount of time. Furthermore, defects were found to either show strongly bias dependent charge capture and emission times, the so-called switching traps, or bias dependent charge capture times and bias independent emission times, the so-called fixed traps. To explain all these effects, two additional states have been introduced in the model, which leads to the so-called four-state model as shown in Fig. 11 [14, 53].

The probability $P_i(t)$ of the defect being in any state i can be described by the master equation, a set of first-order differential equations

$$\frac{\partial P_i}{\partial t} = \sum_{j \neq i} \left(P_j k_{ji} - P_i k_{ij} \right) \tag{10}$$

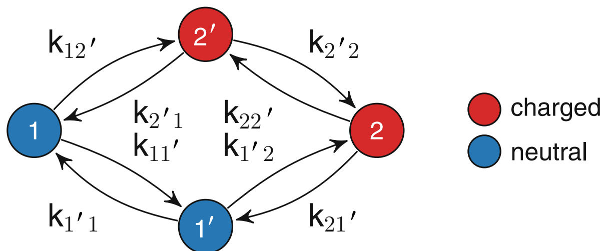


Fig. 11 Markov chain representing an oxide defect modeled using the 4-state model. States 1 and 1′ represent the defect in a neutral state, states 2 and 2′ in a charged state. States marked with ′(prime) represent metastable states

with the sum of all probabilities being 1 ($\sum_i P_i = 1$) and $k = 1/\tau$ denoting the transition rates.

In the four-state model, each of the two charge states is represented by both a "stable" and a "metastable" state. This allows for two types of transitions or reactions to occur: those without carrier exchange, and those with. Transitions without charge transfer can be modeled as thermally activated reactions, where overcoming a thermal barrier allows the defect to switch its configuration. The transition rates are typically modeled in the form of Arrhenius equations:

$$k = k_0 \mathrm{e}^{-\beta \mathcal{E}} \tag{11}$$

with $\beta = (k_\mathrm{B} T)^{-1}$, a pre-factor k_0 and an energetic barrier between the states $\mathcal{E}$.

To describe transitions between different charge states (charge transfer reactions), the tunneling probability for the carrier involved, as well as the availability of carriers or target states for capture and emission, respectively, has to be considered. This process can be described using non-radiative multi-phonon (NMP) theory [54–56] and yields equations in the form of [57–59]:

$$k_{\mathrm{c},12'} = k_0 \int D_\mathrm{p}(E) f_\mathrm{p}(E) \lambda(d, E) \mathrm{e}^{-\beta \mathcal{E}_{12'}(E_\mathrm{T}, E)} \mathrm{d}E \tag{12}$$

$$k_{\mathrm{e},2'1} = k_0 \int D_\mathrm{p}(E) f_\mathrm{n}(E) \lambda(d, E) \mathrm{e}^{-\beta \mathcal{E}_{2'1}(E_\mathrm{T}, E)} \mathrm{d}E, \tag{13}$$

e.g., for the exchange rates of holes between a defect and the valence band for the $1 \leftrightarrow 2'$ transition. Using the assumption that most of the exchange happens close to the band edge, the equations can be simplified to:

$$k_{\mathrm{c},12'} = k_0 p \lambda \mathrm{e}^{-\beta \mathcal{E}_{12'}(E_\mathrm{T}, E_\mathrm{v})} \tag{14}$$

$$k_{\mathrm{e},2'1} = k_0 N_\mathrm{v} \lambda \mathrm{e}^{-\beta \mathcal{E}_{2'1}(E_\mathrm{T}, E_\mathrm{v})}. \tag{15}$$

Here, k_0 denotes a pre-factor, λ a tunneling coefficient (commonly calculated using the WKB approximation), D the density of states, f the Fermi factor, p the density of holes, N_v the effective density of states in the valence band, E_t and E_v the defect and valence band energy level, respectively, and $\mathcal{E}_{ij}$ the energy barrier heights of the reaction. The barrier heights depend on the effective defect energy and the band energy of the exchange point, which strongly determines the bias dependence of the rates. In the simplest and most common approximation, the potential energy surfaces of the individual states are approximated by parabolic functions along the reaction coordinates, and the energy barriers are given by the intersection point between them. A more elaborate choice of the potential for the investigated defect might improve the physical accuracy and one example for such a potential is the Morse potential [60].

From the experimental data, we can only distinguish between charged or discharged states, resulting in effective capture and emission times being measured,

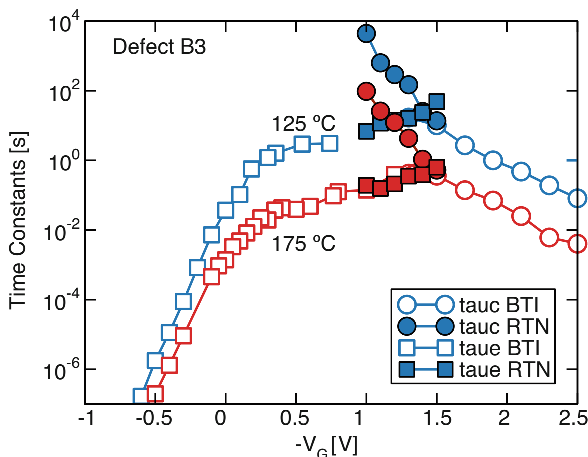


Fig. 12 Gate voltage dependence of effective charge capture and charge emission time characteristics for a defect characterized using TDDS (open symbols) and RTN (filled symbols) measurements, at two temperatures. The voltage ranges for both measurement methods overlap and the obtained data show good agreement. From [16]

as shown in Fig. 12. Depending on the gate voltage, usually one of the two possible transitions with charge transfer will be dominant. The capture and emission times obtained from the experiments will then be either the inverse rates between a stable state of one charge and metastable states of the other charge, or the first passage times between the two stable states, depending on the defect and bias. The expectation value of the first passage time from, e.g., state 2 to state 1 over state 2′ can be calculated as:

$$\tau_{e,2\to 2'\to 1} = \frac{\tau_{22'} + \tau_{2'1} + \tau_{2'2}}{\tau_{22'}\tau_{2'1}} \tag{16}$$

The atomistic aspects of oxide defects can be studied with DFT calculations. A more detailed discussion of this aspect of reliability engineering can be found in Chapter 19 of this book [61].

5 Defect Characterization

With the step height and the charge transition times at a number of bias points, conclusions about a defect parameters such as its position and energy can be drawn using either first-order calculations or TCAD simulations.

5.1 TCAD Simulation

To perform TCAD simulations, an abstracted version of the DUT reflecting the essential geometry and doping details of the real device is designed in the simulation

software, and in a first step calibrated using I_D-V_G and/or C-V measurement data. In a second step, a defect as described in Sect. 4 is placed in the oxide. Its charge transition times can then be simulated at the bias points of the measurement and in that way the model parameters calibrated to reproduce the experimental data.

5.2 First-Order Calculations

Due to the fact that RTN measurements give both charge capture and emission times of the defect close to the intersection point, RTN data allows us to draw conclusions from simple calculations, as we will show below. This proves useful if, for example, large sets of defects have to be analyzed, if TCAD simulation is not yet available for the technology, or as a starting point for TCAD calibration. Starting from the rate equations for charge capture and emission

$$k_{\mathrm{c,e}} = k_{0\mathrm{c,e}} \exp\left(-\frac{\mathcal{E}_{\mathrm{c,e}}}{k_{\mathrm{B}}T}\right), \tag{17}$$

the change in defect energy with gate voltage can be obtained from: [27]

$$\frac{\partial \ln k_{\mathrm{e}}}{\partial V_{\mathrm{G}}} - \frac{\partial \ln k_{\mathrm{c}}}{\partial V_{\mathrm{G}}} = \frac{1}{k_{\mathrm{B}}T}\frac{\partial(\mathcal{E}_{\mathrm{c}} - \mathcal{E}_{\mathrm{e}})}{\partial V_{\mathrm{G}}} = \frac{1}{k_{\mathrm{B}}T}\frac{\partial E_{\mathrm{T}}}{\partial V_{\mathrm{G}}} \tag{18}$$

assuming a constant carrier density in the channel. Further assuming a homogeneous electric field in the oxide gives a link to the position of the defect

$$\partial E_{\mathrm{T}} = -\frac{d}{t_{\mathrm{ox}}}\partial V_{\mathrm{G}}. \tag{19}$$

The vertical position of the defect can then be estimated as

$$d = -t_{\mathrm{ox}}\frac{\partial E_{\mathrm{T}}}{\partial V_{\mathrm{G}}} = -t_{\mathrm{ox}}k_{\mathrm{B}}T\left(\frac{\partial \ln \tau_{\mathrm{c}}/\tau_{\mathrm{e}}}{\partial V_{\mathrm{G}}}\right). \tag{20}$$

Integrating Equation 19 gives the energetic position of the defect with a still unknown integration constant C

$$E_{\mathrm{T}}(V_{\mathrm{G}}) = k_{\mathrm{B}}T\left(\frac{\partial \ln \tau_{\mathrm{c}}/\tau_{\mathrm{e}}}{\partial V_{\mathrm{G}}}\right)V_{\mathrm{G}} + C. \tag{21}$$

At $V_{\mathrm{G}} = V_{\mathrm{G},i}$, the defect's capture and emission time are equal, which puts its energy at the Fermi level, i.e.:

$$E_{\mathrm{T}}(V_{\mathrm{G},i}) = E_{\mathrm{F}} \tag{22}$$

From this, the integration constant can be obtained to yield:

$$E_{\mathrm{T}}(V_{\mathrm{G}}) = k_{\mathrm{B}}T \left(\frac{\partial \ln \tau_{\mathrm{c}}/\tau_{\mathrm{e}}}{\partial V_{\mathrm{G}}} \right) (V_{\mathrm{G}} - V_{\mathrm{G},i}) + E_{\mathrm{F}} \tag{23}$$

The equations can be easily adapted for gate stacks consisting of more than one material as shown in [34].

6 TDDS

In the previous sections, we discussed RTN, which can be observed if a small-area device is operated in equilibrium. Another way of obtaining the capture and emission times of a defect is to measure not at static bias conditions as discussed above but to record the response of the device to transient bias conditions. For this, the gate bias is switched from one value to another, usually chosen in a manner to push the defect far above or far below the Fermi level. This forces defects in the device to charge or discharge starting with the bias change, with a defect capturing or emitting on average after τ_{c} or τ_{e} at the new voltage. The affected region in the oxide is called active energy region, as illustrated in Fig. 13.

This type of measurement is the basis for both the TDDS characterization method in small devices and stress-recovery measurements in large devices. An illustration is given in Fig. 14, where an overview of traces recorded at transient and static bias conditions on large and small devices is shown. The link between the static behavior of small-area and large-area devices was already discussed in the introduction. For the transient case, the link between the responses of large- and small-area devices can easily be seen as the result of a superposition of numerous charge transition events of single defects, thereby smoothing the drain currents.

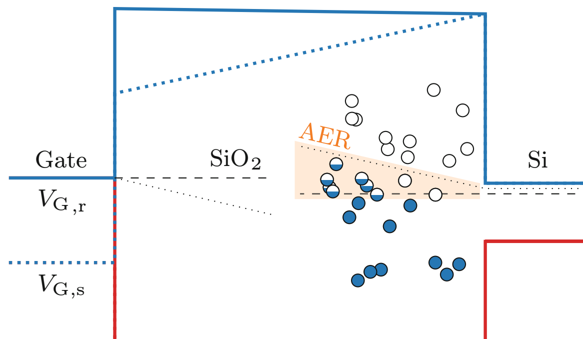


Fig. 13 Band diagram illustrating the scanning window of TDDS—the active energy region—for a specific set of stress and recovery biases. Defects located, e.g., below the Fermi level during stress conditions and located above the Fermi level during recovery conditions change their occupation after switching the bias, enabling their characterization

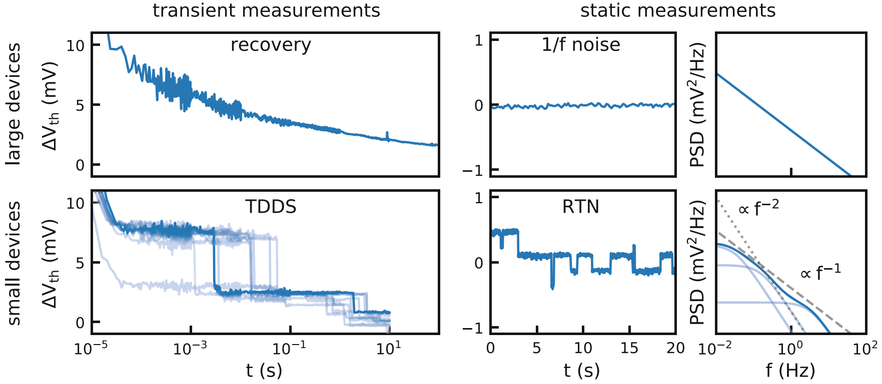


Fig. 14 Exemplary transient measurements recorded after the devices have been stressed (left) and noise measurements (static, right) shown for large-area (top) and small-area (bottom) devices. It can be consistently observed that for large-area devices continuous trends are observed, while a discrete behavior is obtained for the small-area counterparts. Both techniques can be used to characterize individual defects in small-area devices. For the static measurements, illustrations of the frequency behavior are given to show the link between small- and large-area devices. In large-area devices, many Lorentzian power spectral densities (PSDs) produced by individual defects add up to 1/f noise

6.1 Measurement

For TDDS measurements, the same setup as shown in Fig. 6 for RTN measurements can be used, but now the measurement sequence consists of two separate phases. The first phase is the stress or charging phase where the device is usually operated in strong inversion for a defined stress (or charging) time. In this phase, no sampling is typically done, as no drain current is available due to $V_{\mathrm{D,str}} = 0\,\mathrm{V}$. After the stress phase, the gate bias is switched to a voltage in the sub-threshold regime, and a small drain bias of typically $V_{\mathrm{D,rec}} = 50\,\mathrm{mV}$ is applied and the data sampling started. The emission time is then considered the time interval between the bias switch and the occurrence of the charge transition event. In contrast, the capture time cannot be measured directly but is deducted indirectly by doing measurements at the same bias conditions but with varying stress times [14]. Compared to RTN measurements where linear sampling is required, the application of a logarithmic sampling scheme with, for instance, 200 points per decade is preferred for TDDS. This lowers the requirement on the sampling buffer size but now an exact timing of the bias switches and the subsequent start of sampling is inevitable to achieve accurate charge emission times. As previously mentioned, the change in the occupancy of a defect between stress and relaxation conditions determines the probability of an emission event to occur during the measurement sequence. This entails that the stress time should be chosen such that the occupancy of the defect is as close as possible to unity after stress. This can also be achieved by the modification of the stress bias; however, the bias is limited by the oxide breakdown voltage. Regarding

the error due to a limited number (N) of observations, the same considerations as for RTN discussed in Sect. 2 apply. To achieve a given amount of observations, a total number of measurements $N_\mathrm{m} = \frac{N}{o_\mathrm{str}}$, with the occupancy after stress, o_str, have to be recorded.

6.2 Measurement Limitations

RTN and TDDS measurements differ in the gate voltage range where defects can be characterized. An overview of the limits for both RTN and TDDS measurements shown for an exemplary defect and device is given in Fig. 15. In the figure, it can be seen that TDDS allows to measure the charge capture time (red area) in a voltage range limited by the breakdown voltage for high fields and limited by low defect occupancy for low fields. The minimum and maximum time constants which can be measured depend on the minimum time a stress bias can be applied and the maximum time one is willing to spend. The range where the charge emission time (blue area) may be measured is limited by the defect occupancy and measurement resolution of larger drain currents, i.e., the defect visibility in the current data for high fields, and measurement sensitivity and noise in the depletion region. The time constants which can be measured depend on the minimum sampling time and the time required for the bias switch, and again the maximum measurement time one is willing to use determines the upper limit.

RTN (orange area) is limited to the voltage range where both capture and emission times are in the range between the sampling time and the total recording time of the measurement traces. Note that by modifying the device temperature, the defect's charge transition times move to shorter or longer times. Thus, the temperature can be used to cleverly adjust the measurement window.

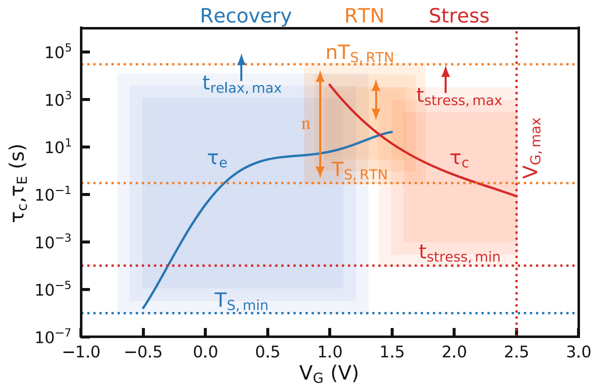


Fig. 15 Limits for RTN and TDDS measurements shown for an exemplary defect and device. $T_{\mathrm{S,min}}$, $t_{\mathrm{stress,min}}$, and n are limitations of the measurement setup, while the upper limit of stress bias $V_{\mathrm{G,max}}$ is usually given by the device. Notice that the position of the RTN window can be shifted by the choice of $T_{\mathrm{S,RTN}}$

6.3 Defect Parameter Extraction

For TDDS measurements, the recovery traces are analyzed using a step-detection algorithm such as the Canny algorithm discussed in Sect. 3. The extracted charge emission times and step heights are then plotted in the charge emission time versus step height plane, which is called spectral map, as shown in Fig. 16. As can be seen, the single charge emission events tend to form a cluster in the spectral map for each identified defect. The emission time and step height of a defect can then be calculated by taking the average values of the points belonging to one cluster. The capture events can usually not be observed directly as there is no current during stress. To indirectly obtain the capture time, the measurement has to be repeated at varying stress times. One then plots the probability of a charge emission event to occur (i.e., the number of points in each cluster divided by the number of measurements) over the stress times, which can be described by an exponential CDF which then finally gives the charge capture time.

This technique can also be used for "mixed" measurement traces recorded at a recovery bias where the defects also produce RTN. In this case, only the first charge emission event obtained for each defect should be used for the construction of the spectral map. The remaining points may then be used for RTN extraction. Comparing parameter extraction from RTN and TDDS data, TDDS has the advantage that

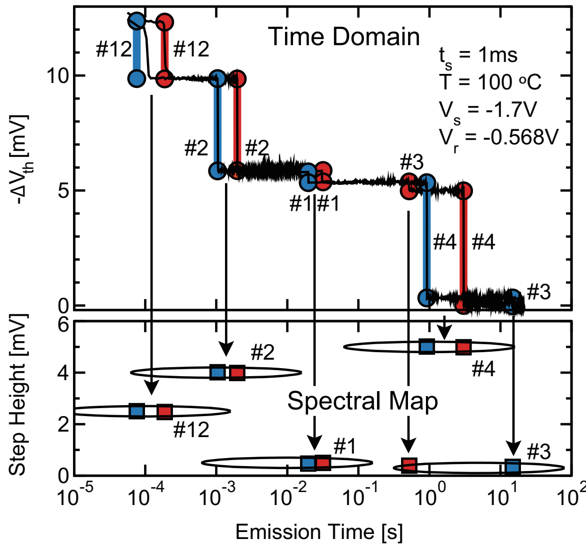


Fig. 16 Working principle of TDDS. Steps in ΔV_{th} (top) are marked in a scatterplot showing step height over emission time (bottom). Multiple measurements produce clusters in the scatterplot with exponentially distributed emission times and step heights normally distributed around a mean, which itself is exponentially distributed. Each cluster corresponds to an individual defect. The emission time of the defect (at $V_{\text{G,rec}}$) can be obtained directly from the measurement. To obtain the capture time (at $V_{\text{G,str}}$) multiple such measurements at varying stress times have to be conducted. From [14]

due to the bias switch, which enforces the occupational change of the defects, both step height and emission time can be directly obtained from the measurement. This allows to easily distinguish between defects even with similar step height as they will form separate clusters, provided their emission times are sufficiently different.

7 Link Between RTN and TDDS

While we presented both RTN and TDDS as methods to characterize oxide defects, the question whether both methods characterize the same sets of (oxide) defects has been left open. Grasser et al. reported in [16] a number of defects initially identified using TDDS measurements, all of which could also be observed using RTN measurements (see Fig. 17). Based on observations on around a hundred nanoscale devices, they concluded that both RTN and the recoverable component of BTI are due to the same defects.

Considering this from a theoretical perspective, oxide defects may be separated in three groups. The first group comprises defects whose charge capture time decreases relative to the charge emission time towards stress conditions. Their equilibrium occupancy is higher at stress bias than at recovery bias. This is the kind of defect which is most commonly found, and due to its decrease in occupancy readily emits during the recovery phase in TDDS measurements. The second group consists of defects whose charge capture times increase relative to their charge emission times towards stress bias. This can be defects close to the gate, positioned below the Fermi level during stress and above during recovery (for NBTI, and vice versa for PBTI). After switching to recovery conditions, they capture a charge instead of emitting it and are thus visible as inverse steps in TDDS measurements. Both

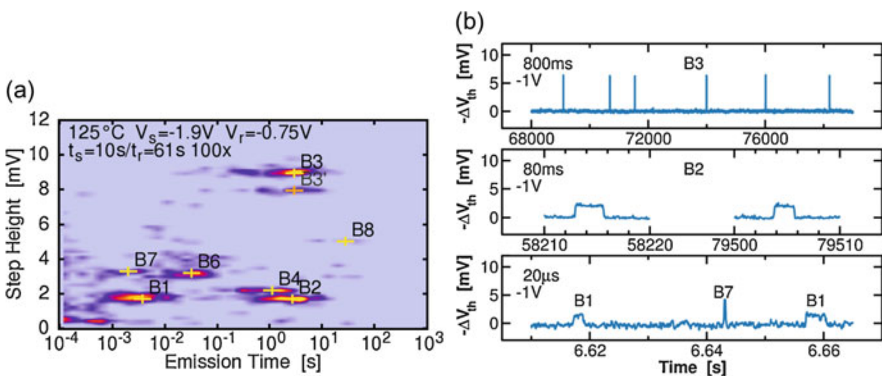


Fig. 17 Spectral map from TDDS measurements (**a**) showing a number of defects in a device and RTN traces (**b**) recorded on the same device, showing some of the same defects. All defects found in (**a**) could also be measured using RTN measurements and vice versa, indicating that RTN and TDDS identify the same type of defects. From [16]

groups of defects can also be found in RTN measurements, if the experiments are performed at suitable conditions in between stress and recovery conditions, as their thermodynamic trap level has to cross the Fermi level between stress and recovery conditions. The last group of defects are the ones whose charge capture to charge emission time ratio does not change significantly with gate bias. Their occupancies thus stay relatively constant, usually close to zero or close to one both during stress and recovery. This might be defects located energetically outside the TDDS measurement range or very close to the interface. They do not show up in TDDS measurements but are still able to sporadically emit or capture a charge for a short period of time. While these defects still produce RTN, most will not be visible in regular RTN measurements due to the large factor between their charge transition times and the limited time span of the RTN measurement window. Simulated band diagrams illustrating the suspected defect bands in SiO_2 together with measurement windows for TDDS and RTN are given in Fig. 18 for *P*- and NMOS devices (or NBTI- and PBTI-like measurements). Note that while parts of the defect bands

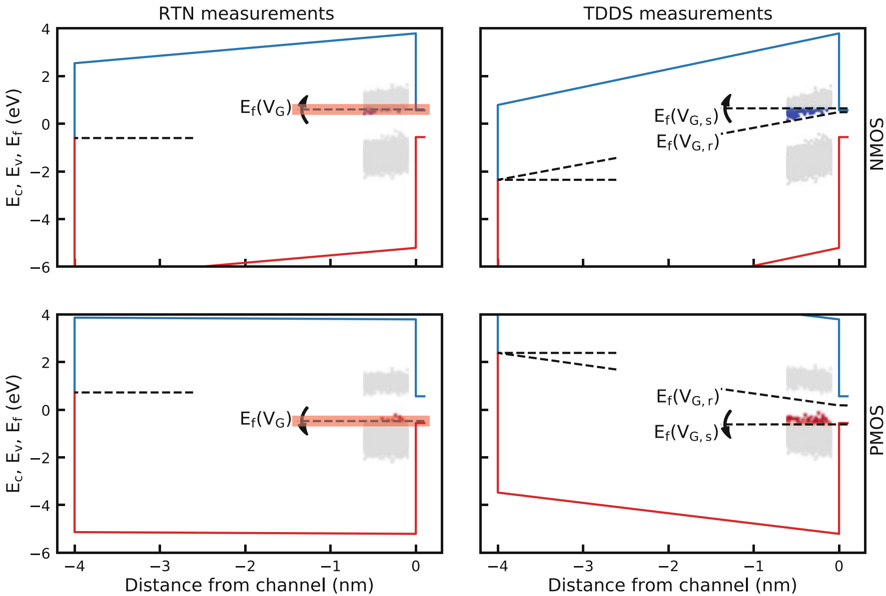


Fig. 18 Band diagrams with the effective scanning ranges for RTN and TDDS characterization for positive and negative bias conditions with exemplary defect distributions plotted close to the channel interface. TDDS measurements are sensitive over a trapezoidal area spanned by the interface Fermi levels at $V_{G,str}$ and $V_{G,rec}$ called the active energy region, while RTN measurements are sensitive to defects with an effective energy close to the channel and gate Fermi levels (marked as red bars), with the width depending on the limitations of the measurement setup. For both methods, modification of the applied bias changes the scanned area, and thus allows to pinpoint the located defects. As TDDS measurements are usually performed at bias conditions far beyond operating specifications, defects located outside the scanning ranges are unlikely to affect device performance at operating conditions. Values for the locations of the defect bands are taken from [62]

are inaccessible to the electrical measurements presented here, those are also very unlikely to affect device performance during operating conditions.

8 Summary

Non-ideal device behavior during operation is mainly determined by defects located in the oxide or at the oxide/semiconductor interface. In order to identify these defects, RTN or TDDS measurements can be performed. From these measurements, the charge transition times, i.e., charge capture time and charge emission time, and the contribution of the defect to the overall ΔV_{th}, i.e., the defect's step height, can be extracted. Afterwards the charge transition times can be analyzed using advanced computer simulations and a possible link to defect candidates using DFT simulations can be established.

For highest accuracy of the extracted data, elaborate experiments have to be conducted. This can be achieved by either using general purpose instruments or by using custom-design hardware based on the circuits shown initially. The advantage of custom solutions is that they can be optimized for highest SNR and performance. For the case of custom solutions data acquisition is typically performed using high speed digital storage oscilloscopes. To maximize the measurement ranges of RTN and TDDS measurements, the sampling tool should on the one hand be able to allow for fast and precise sampling and on the other hand be able to record a large number of samples for a long time period. To keep the number of samples in a feasible range, logarithmic time intervals for data sampling are typically used for TDDS. The next step is to analyze the measurement data. For this, several methods like the histogram and time lag plot methods, parameter extraction using the Canny edge detection algorithm, and maximum likelihood estimation using hidden Markov models have been presented. The method of choice depends among others on the amount of measurement sets, the signal-to-noise ratio, and the complexity of the recorded data. Finally, the extracted charge transition time characteristics can be used to calibrate physics-based models like the four-state NMP model, which is implemented in modern microelectronic device simulators.

In summary, in this chapter we have briefly discussed the characterization as well as the evaluation of the measurement data and the modeling of individual defects. The main focus is put on defect characterization by means of noise measurements, i.e., RTN measurements, and TDDS measurements. The careful execution of the experiments, whether with custom-build circuit or general purpose instruments, allows a deep insight into the physical mechanisms of charge trapping in MOS transistors, which is responsible for device altering even in extremely scaled technologies.

Acknowledgements This work was supported in part by the Austrian Science Fund (FWF) under Project P 26382-N30, Project P 23958-N24, and Project I2606-N30, in part by the European Union FP7 Project ATHENIS 3-D under Grant 619246, and in part by the Austrian Research Promotion Agency (FFG, Take-Off Program) under Project 861022 and Project 867414.

References

1. L. Terman, An investigation of surface states at a silicon/silicon oxide interface employing metal-oxide-silicon diodes. Solid State Electron. **5**(5), 285–299 (1962)
2. J.S. Brugler, P.G.A. Jespers, Charge pumping in MOS devices. IEEE Trans. Electron Devices **16**, 297–302 (1969)
3. G. Groeseneken, H.E. Maes, N. Beltran, R.F. De Keersmaecker, A reliable approach to charge-pumping measurements in MOS transistors. IEEE Trans. Electron Devices **31**, 42–53 (1984)
4. D.V. Lang, Deep-level transient spectroscopy: a new method to characterize traps in semiconductors. J. Appl. Phys. **45**(7), 3023–3032 (1974)
5. A. Neugroschel, C.-T. Sah, K.M. Han, M.S. Carroll, T. Nishida, J.T. Kavalieros, Y. Lu, Direct-current measurements of oxide and interface traps on oxidized silicon. IEEE Trans. Electron Devices **42**(9), 1657–1662 (1995)
6. C.-T. Sah, A. Neugroschel, K. M. Han, and J. T. Kavalieros, Profiling interface traps in MOS transistors by the DC current-voltage method. IEEE Electron Device Lett. **17**(2), 72–74 (1996)
7. B. Kaczer, T. Grasser, J. Roussel, J. Martin-Martinez, R. O'Connor, B. O'sullivan, G. Groeseneken, Ubiquitous relaxation in BTI stressing-new evaluation and insights, in *2008 IEEE International Reliability Physics Symposium*, pp. 20–27 (IEEE, Piscataway, 2008)
8. A.L. McWhorter et al., *1/f Noise and Related Surface Effects in Germanium* (MIT Lincoln Laboratory, Cambridge, 1955)
9. H. Mikoshiba, 1/f noise in n-channel silicon-gate MOS transistors. IEEE Trans. Electron Devices **29**, 965–970 (1982)
10. K.S. Ralls, W.J. Skocpol, L.D. Jackel, R.E. Howard, L.A. Fetter, R.W. Epworth, D.M. Tennant, Discrete resistance switching in submicrometer silicon inversion layers: individual interface traps and low-frequency ($\frac{1}{f}$) noise. Phys. Rev. Lett. **52**, 228–231 (1984)
11. B. Stampfer, F. Zhang, Y.Y. Illarionov, T. Knobloch, P. Wu, M. Waltl, A. Grill, J. Appenzeller, T. Grasser, Characterization of single defects in ultrascaled MoS2 field-effect transistors. ACS Nano **12**(6), 5368–5375 (2018). PMID: 29878746
12. M. Bina, O. Triebl, B. Schwarz, M. Karner, B. Kaczer, T. Grasser, Simulation of reliability on nanoscale devices, in *Proceedings of the 16th International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 109–112 (2012)
13. A. Ghetti, C. Compagnoni, A. Spinelli, A. Visconti, Comprehensive analysis of random telegraph noise instability and its scaling in deca–nanometer flash memories. IEEE Trans. Electron Devices **56**(8), 1746–1752 (2009)
14. T. Grasser, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, B. Kaczer, The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability, in *IEEE International Reliability Physics Symposium*, pp. 16–25 (2010)
15. S. Machlup, Noise in semiconductors: spectrum of a two-parameter random signal. J. Appl. Phys. **25**(3), 341–343 (1954)
16. T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, B. Kaczer, A unified perspective of RTN and BTI, in *2014 IEEE International Reliability Physics Symposium*, pp. 4A.5.1–4A.5.7 (2014)
17. C. Marquez, O. Huerta, A.I. Tec-Chim, F. Guarin, E.A. Gutierrez-D, F. Gamiz, Systematic characterization of random telegraph noise and its dependence with magnetic fields in MOSFET devices, in *Noise in Nanoscale Semiconductor Devices*, ed. by T. Grasser (Springer, Cham, 2019). https://doi.org/10.1007/978-3-030-37500-3_4
18. B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, M. Goodwin, Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification, in *IEEE International Reliability Physics Symposium, 2005. Proceedings.*, pp. 381–387 (2005)

19. A. Kerber, K. Maitra, A. Majumdar, M. Hargrove, R. J. Carter, E. A. Cartier, Characterization of fast relaxation during BTI stress in conventional and advanced CMOS devices with HfO_2/TiN gate stacks, IEEE Trans. Electron Devices **55**(11), 3175–3183 (2008)
20. H. Reisinger, O. Blank, W. Heinrigs, A. Mühlhoff, W. Gustin, C. Schlünder, Analysis of NBTI degradation- and recovery-behavior based on ultra fast V_{th}-measurements, in *IEEE International Reliability Physics Symposium Proceedings*, pp. 448–453 (2006)
21. B. Ullmann, K. Puschkarsky, M. Waltl, H. Reisinger, T. Grasser, Evaluation of advanced MOSFET threshold voltage drift measurement techniques. IEEE Trans. Device Mater. Reliab. **19**(2), 358–362 (2019)
22. G.A. Du, D.S. Ang, Z.Q. Teo, Y.Z. Hu, Ultrafast measurement on NBTI. IEEE Electron Device Lett. **30**, 275–277 (2009)
23. M. Maestro, J. Diaz, A. Crespo-Yepes, M. Gonzalez, J. Martin-Martinez, R. Rodriguez, M. Nafria, F. Campabadal, X. Aymerich, New high resolution random telegraph noise (RTN) characterization method for resistive RAM. Solid State Electron. **115**, 140–145 (2016). Selected papers from the EUROSOI-ULIS conference
24. 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 Mater. Reliab. **14**(2), 616–622 (2014)
25. H. Reisinger, The time-dependent defect spectroscopy, in *Bias Temperature Instability for Devices and Circuits* (Springer, Berlin, 2014), pp. 75–109
26. T. Nagumo, K. Takeuchi, S. Yokogawa, K. Imai, Y. Hayashi, New analysis methods for comprehensive understanding of random telegraph noise, in *2009 IEEE International Electron Devices Meeting (IEDM)*, pp. 1–4 (2009)
27. T. Nagumo, K. Takeuchi, T. Hase, Y. Hayashi, Statistical characterization of trap position, energy, amplitude and time constants by RTN measurement of multiple individual traps, in *2010 IEEE International Electron Devices Meeting (IEDM)*, pp. 28–3 (IEEE, Piscataway, 2010)
28. Y. Yuzhelevski, M. Yuzhelevski, G. Jung, Random telegraph noise analysis in time domain. Rev. Sci. Instrum. **71**(4), 1681–1688 (2000)
29. J. Martin-Martinez, J. Diaz, R. Rodriguez, M. Nafria, X. Aymerich, New weighted time lag method for the analysis of random telegraph signals. IEEE Electron Device Lett. **35**, 479–481 (2014)
30. J. Martin-Martinez, R. Rodriguez, M. Nafria, Advanced characterization and analysis of random telegraph noise in CMOS devices, in *Noise in Nanoscale Semiconductor Devices*, ed. by T. Grasser (Springer, Cham, 2019). https://doi.org/10.1007/978-3-030-37500-3_14
31. J. Canny, A computational approach to edge detection, in *Readings in Computer Vision* (Elsevier, Amsterdam, 1987), pp. 184–203
32. P. Bao, L. Zhang, X. Wu, Canny edge detection enhancement by scale multiplication, IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1485–1490 (2005)
33. N. Otsu, A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**, 62–66 (1979)
34. A. Grill, B. Stampfer, M. Waltl, K.-S. Im, J.-H. Lee, C. Ostermaier, H. Ceric, T. Grasser, Characterization and modeling of single defects in GaN/AlGaN fin-MIS-HEMTs, in *IEEE International Reliability Physics Symposium Proceedings* (2017), pp. 3B–5
35. L.E. Baum, T. Petrie, G. Soules, N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. Ann. Math. Stat. **41**(1), 164–171 (1970)
36. A. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, IEEE Trans. Inf. Theory **13**(2), 260–269 (1967)

37. H. Miki, N. Tega, M. Yamaoka, D.J. Frank, A. Bansal, M. Kobayashi, K. Cheng, C.P. D'Emic, Z. Ren, S. Wu, J. Yau, Y. Zhu, M.A. Guillorn, D.-G. Park, W. Haensch, E. Leobandung, K. Torii, Statistical measurement of random telegraph noise and its impact in scaled-down high-k/metal-gate MOSFETs, in *2012 International Electron Devices Meeting*, pp. 19.1.1–19.1.4 (2012)
38. W.S. Cleveland, Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. **74**(368), 829–836 (1979)
39. P.H. Eilers, A perfect smoother. Anal. Chem. **75**(14), 3631–3636 (2003)
40. L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE **77**(2), 257–286 (1989)
41. R. Weiss, S. Du, J. Grobler, S. Lebedev, and G. Varoquaux, hmmlearn 0.2.2 (2017)
42. J. Schreiber, Pomegranate: fast and flexible probabilistic modeling in python. J. Mach. Learn. Res. **18**(1), 5992–5997 (2017)
43. D.J. Frank, H. Miki, Analysis of oxide traps in nanoscale MOSFETs using random telegraph noise, in *Bias Temperature Instability for Devices and Circuits* (Springer, Berlin, 2014), pp. 111–134
44. F.M. Puglisi, P. Pavan, Factorial hidden Markov model analysis of random telegraph noise in resistive random access memories. ECTI Trans. Electr. Eng. Electron. Commun. **12**(1), 24–29 (2014)
45. Z. Ghahramani, M.I. Jordan, Factorial hidden Markov models, in *Advances in Neural Information Processing Systems* (1996), pp. 472–478
46. E. Simoen, B. Dierickx, C.L. Claeys, G.J. Declerck, Explaining the amplitude of RTS noise in submicrometer MOSFETs. IEEE Trans. Electron Devices **39**, 422–429 (1992)
47. A.-M. El-Sayed, Y. Wimmer, W. Goes, T. Grasser, V.V. Afanas' ev, A.L. Shluger, Theoretical models of hydrogen-induced defects in amorphous silicon dioxide. Phys. Rev. B **92**(1), 014107 (2015)
48. A.L. Shluger, K.P. McKenna, Models of oxygen vacancy defects involved in degradation of gate dielectrics, in *2013 IEEE International Reliability Physics Symposium (IRPS)* (2013), pp. 5A.1.1–5A.1.9
49. S. Pantelides, Z.-Y. Lu, C. Nicklaw, T. Bakos, S. Rashkeev, D. Fleetwood, R. Schrimpf, The E' center and oxygen vacancies in SiO_2. J. Non-Cryst. Solids **354**(2–9), 217–223 (2008)
50. Y. Wimmer, A.-M. El-Sayed, W. Goes, T. Grasser, A.L. Shluger, Role of hydrogen in volatile behaviour of defects in SiO_2-based electronic devices, in *Proceedings of the Royal Society A*, vol. 472 (The Royal Society, London, 2016), p. 20160009
51. T. Grasser, M. Waltl, Y. Wimmer, W. Goes, R. Kosik, G. Rzepa, H. Reisinger, G. Pobegen, A. El-Sayed, A. Shluger, et al., Gate-sided hydrogen release as the origin of "permanent" NBTI degradation: from single defects to lifetimes, in *IEEE International Electron Devices Meeting (IEDM)* (IEEE, Piscataway, 2015), pp. 20–21
52. T. Grasser, M. Waltl, W. Goes, Y. Wimmer, A.-M. El-Sayed, A. Shluger, B. Kaczer, On the volatility of oxide defects: activation, deactivation and transformation, in *IEEE International Reliability Physics Symposium Proceedings* (2015), pp. 5A.3.1–5A.3.8
53. M.J. Uren, M.J. Kirton, S. Collins, Anomalous telegraph noise in small-area silicon metal-oxide-semiconductor field-effect transistors. Phys. Rev. B **37**, 8346–8350 (1988)
54. K. Huang, A. Rhys, Theory of light absorption and non-radiative transitions in f-centres, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 204 (1950), pp. 406–423
55. D. Lang, C. Henry, Nonradiative recombination at deep levels in GaAs and GaP by lattice-relaxation multiphonon emission. Phys. Rev. Lett. **35**(22), 1525 (1975)
56. M. Kirton, M. Uren, Noise in solid-state microstructures: a new perspective on individual defects, interface states and low-frequency (1/f) noise. Adv. Phys. **38**(4), 367–468 (1989)
57. T. Grasser, Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities. Microelectron. Reliab. **52**(1), 39–70 (2012)

58. W. Goes, F. Schanovsky, T. Grasser, Advanced modeling of oxide defects, in *Bias Temperature Instability for Devices and Circuits* (Springer, Berlin, 2014), pp. 409–446
59. W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. Shluger, T. Grasser, Identification of oxide defects in semiconductor devices: a systematic approach linking DFT to rate equations and experimental evidence. Microelectron. Reliab. **87**, 286–320 (2018)
60. Y. Wimmer, Hydrogen Related Defects in Amorphous SiO_2 and the Negative Bias Temperature Instability. Dissertation, Technische Universität Wien (2017)
61. D. Waldhoer, A.-M. El-Sayed, Y. Wimmer, M. Waltl, T. Grasser, Atomistic modeling of oxide defects, in *Noise in Nanoscale Semiconductor Devices*, ed. by T. Grasser (Springer, Cham, 2019). https://doi.org/10.1007/978-3-030-37500-3_18
62. G. Rzepa, J. Franco, B. O'Sullivan, A. Subirats, M. Simicic, G. Hellings, P. Weckx, M. Jech, T. Knobloch, M. Waltl, P. Roussel, D. Linten, B. Kaczer, T. Grasser, Comphy – a compact-physics framework for unified modeling of BTI. Microelectron. Reliab. **85**, 49–65 (2018)