Multi-Phonon Processes as the Origin of Reliability Issues

W. Goes[a], M. Toledano-Luque[b], F. Schanovsky[a], M. Bina[a], O. Baumgartner[a], B. Kaczer[b], and T. Grasser[a]

[a] Institute for Microelectronics, TU Wien, Vienna, Austria
[b] imec, Leuven, Belgium

Recently, correlated drain and gate current fluctuations have been observed in nano-scaled MOSFETs, indicating that their occurrence is linked to the same defect. One explanation for this observation can be given by the multi-state defect model, which has been developed to describe the hole capture and emission processes involved in the bias temperature instability (BTI). This model relies on a combination of metastable defect states and nonradiative multi-phonon transitions. Our detailed investigations demonstrate that the observed gate current fluctuations and their correlation with the drain noise can be well reproduced using a TAT mechanism based on the multi-state model.

Introduction

Years ago, on-the-fly measurements and the measure-stress-measure techniques were employed to probe the bias temperature instability (BTI), where the latter uses the shape of relaxation curves to characterize the stressed devices. By contrast, ultra-scaled device technologies contain only a handful of oxide defects, whose discharging events appear as abrupt steps in the relaxation curves. Each of these steps is of a certain height, which is characteristic for one oxide defect. Using this fingerprint, these steps can be used to directly study the detrapping behavior of defects and thus the recoverable component of BTI. This idea was implemented in a new measurement technique called time-dependent defect spectroscopy (TDDS), which has led to several crucial findings regarding the hole capture and emission involved in BTI. The more permanent component, which considerably contributes to long-term degradation (1–6), is not discussed here.

Traditionally, BTI has been discussed within the framework of the reaction-diffusion (RD) model, in which the kinetics are governed by hydrogen diffusion from and towards the interface (7–9). However, TDDS studies have demonstrated that the recoverable component of BTI must be ascribed to a first-order reaction (10), which is consistent with NMP hole capture and emission processes but not with the RD model.

Furthermore, the capture times exhibit a strong gate bias and temperature dependence. The latter cannot be reconciled with elastic tunneling, which is only weakly temperature dependent. By contrast, nonradiative multi-phonon (NMP) processes (11, 12), discussed later in detail, offer an explanation since they are thermally activated and also predict the experimentally observed exponential gate bias dependence.

Finally, the TDDS also provides strong experimental evidence that the oxide defect must feature metastable states:

(*i*) In a simple two-state model (see Fig. 1), there must exist a functional correlation of the form $\tau_c(V_g,T) = f(\tau_e(V_g,T))$ between the hole capture (τ_c) and emission

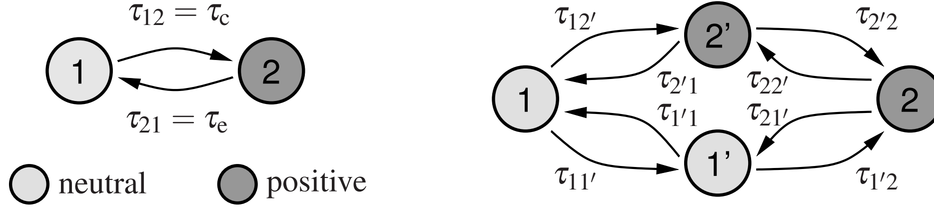

Figure 1: State diagram of an oxide defect with two states (left) and four states (right). The single states are enumerated (1, 2, 1′, 2′) while the arrows indicate transitions between them with the corresponding time constants (τ_{12}, τ_{21}, $\tau_{12'}$, $\tau_{2'1}$, ...). While hole capture and emission in the two-state model must be correlated, this is not the case for the four-state model due to the introduction of intermediate states (1′, 2′). The findings of the points (*i*)-(*iv*) indicate the existence of metastable states as they are present in the four-state model.

(τ_e) time constants if NMP processes are assumed. However, no such a correlation could be confirmed experimentally, indicating the existence of a metastable state (2′) in a four-state model (13).

(*ii*) The gate bias dependence of the capture times (14) shows a curvature that is best described by the introduction of the metastable state 2′.

(*iii*) AC TDDS experiments (15, 16) have revealed a pronounced frequency dependence of the capture times, which has been traced back to the occupancy of the metastable state 2′ (17, 18).

(*iv*) In the TDDS studies two types of defects have been observed — namely, fixed oxide traps, which have temperature independent emission times, and switching oxide traps, which show a pronounced gate bias dependence around the recovery voltage (13, 14). In the four-state model (see Fig. 1), their distinct detrapping behavior may result from the different possible pathways over the two intermediate states 1′ and 2′ during a hole emission event.

(*v*) The metastable configurations can furthermore give an explanation for particular noise phenomena, such as temporary (tRTN) and anomalous (aRTN) random telegraph noise. In such a four-state model, RTN can be understood by temporary transitions to states, in which the defect dwells for a certain time interval but does not produce a noise signal (14).

In summary, the TDDS has produced several crucial hints for modeling and thus understanding of the BTI phenomenon. However, we suspect that also other reliability issues, such as the drain current RTN (20, 21), involve the same type of oxide defects and rely on the same charge capture and emission processes. This hypothesis is supported by a number of studies that link NMP processes to drain noise (22, 23) and the gate leakage (24–26) but disregard the possible existence of metastable states. Interestingly, recent publications (19, 27–29) report that the drain and gate current noise is sometimes correlated (see Fig. 2) in both pFET or nFET devices. Microscopically, this means that the level of the gate current is linked to the charge state of the oxide defects (28, 30). A possible explanation relies on trap-assisted tunneling (TAT), where the bistable defect is only conductive in one of its configurations and the hole capture and emission into and out of

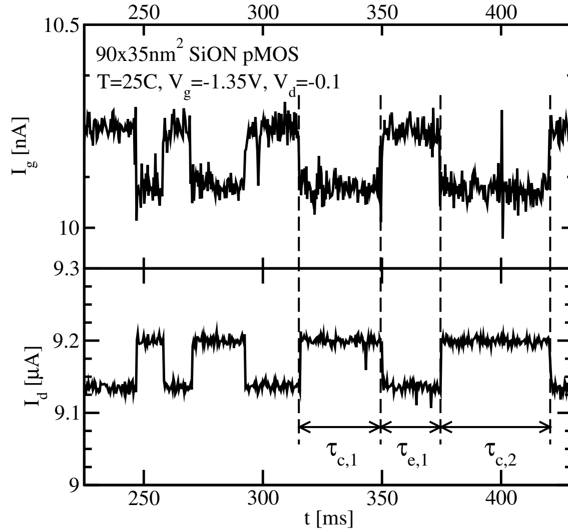

Figure 2: Simultaneously recorded gate (I_g) and drain (I_d) current traces shown in the upper and the lower panel, respectively (19). After a time period of $\tau_{c,i}$ ($i = 1,2,\ldots$), a hole is captured and I_d drops to its lower level while I_g increases significantly at the same time. When the hole is emitted again (after a time period of $\tau_{e,i}$), both currents return to their initial values. Microscopically, this means that the defect provides an additional conductive path through the oxide in its positive charge state (higher level of the gate current) while this path is closed when the defect is neutral (lower level of the gate current).

the oxide defect proceed via NMP transitions. This puzzling correlation is a phenomenon, which is observed for a few defects since it becomes only visible within certain gate bias and temperature ranges. Nevertheless, it reveals much about the nature of defects and thus may help to understand the trapping behavior of oxide defects involved in several reliability issues. In this paper we apply the multi-state defect model, which was developed for the explanation of the BTI, to the correlated drain and gate current noise. Our work provides a physical link between BTI, drain RTN, and the gate current noise.

Multi-State Multi-Phonon Model

As discussed before, the TDDS suggested two decisive ingredients for modeling BTI. These parts are the NMP processes for the charge transfer between the defects and the bulk, on the one hand, and the metastable states, on the other hand. The former can be understood at an atomistic level. Starting from the Huang-Born approximation, the electrons and the nuclei are treated as a common quantum mechanical system. However, the positions of the nuclei enter in this system parametrically and thus determine the energy of the system. This means that the energy and the configuration of the defect are coupled and the defect must move on the so-called adiabatic potential energy surfaces (V_i, V_j). These surfaces exist for different charge states and may intersect (see Fig. 3 left). At these intersection points, the probability of an NMP transition reaches its maximum so that the NMP processes, such as hole capture or emission in BTI, primarily takes place there. A rigorous quantum mechanical derivation of these NMP transitions leads to the so-called lineshape functions (LSF). For device simulation we have chosen their high-temperature limit of the LSFs, which provides a good approximation at practical device operation temperatures (31). In order to incorporate all NMP transitions to and from the conduction and the valence band

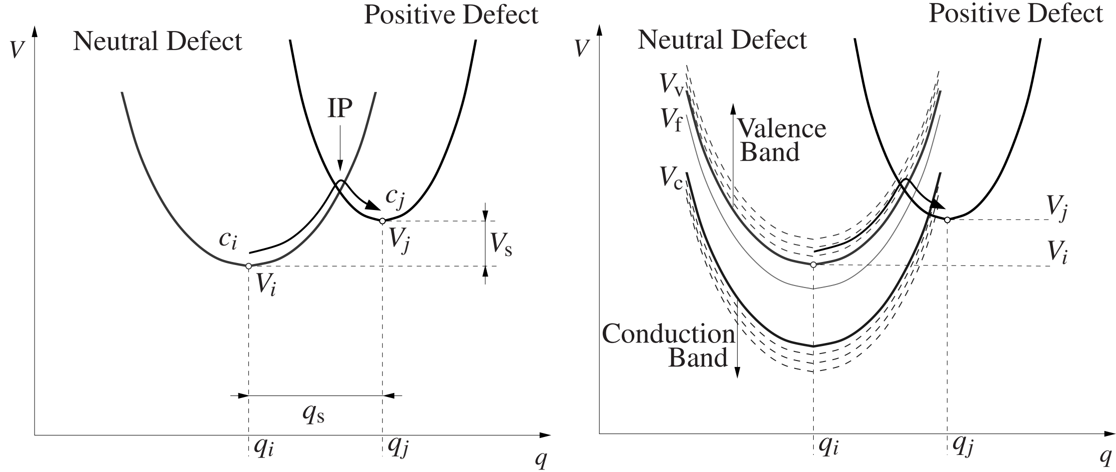

Figure 3: Configuration coordinate diagram for hole capture. **Left:** The defect is present in two charge states, labeled i and j here. Their adiabatic potentials determines their equilibrium configurations q_i and q_j and their corresponding minimum energies V_i and V_j. The arrow indicates an NMP transition for a hole which is emitted from the substrate into the defect. During this process, the defect must be thermally excited from q_i on the left curve (neutral defect) up to the intersection point (IP) with the right curve (positive defect). At this point, the actual hole capture takes place. Subsequently, the defect relaxes to the minimum energy of the right curve at q_j. **Right:** Here, all band states within the substrate conduction (lower parabolas) and valence (upper parabolas) band have been taken into consideration. The parabola of the initial state i shifts with the energy E of electron which is exchanged with the substrate during the NMP process. By contrast, the equilibrium configurations remain unaffected by the electron energy. V_c and V_v correspond to the adiabatic potentials for the case when the hole is captured from the conduction or the valence band edge, respectively.

(see Fig. 3 right), the overall hole capture ($k_{0/+}$) and emission ($k_{+/0}$) rates are obtained by the integrals over the carrier energy (32)

$$k_{0/+} = k_0 \int D_\mathrm{p}(E)\lambda_\mathrm{p}(E,x_\mathrm{t}) f_{0/+}(V_\mathrm{s} = E - E_\mathrm{t}) f_\mathrm{p}(E)\mathrm{d}E \qquad [1]$$

$$k_{+/0} = k_0 \int D_\mathrm{p}(E)\lambda_\mathrm{p}(E,x_\mathrm{t}) f_{+/0}(V_\mathrm{s} = E_\mathrm{t} - E) f_\mathrm{n}(E)\mathrm{d}E \qquad [2]$$

with E and $D_\mathrm{p}(E)$ being the carrier energy and the valence band density of states, respectively. $f_\mathrm{n}(E)$ and $f_\mathrm{p}(E)$ stand for the electron and hole occupancy, respectively. They follow the Fermi-Dirac distribution for typical BTI operation conditions and are therefore related by $f_\mathrm{p}(E) = 1 - f_\mathrm{n}(E)$. $\lambda_\mathrm{p}(E,x_\mathrm{t})$ denotes the electron/hole tunneling factor, derived from Wentzel-Kramers-Brillouin theory. E_t and x_t are the thermodynamic trap level and the trap depth, respectively. $f_{0/+}$ and $f_{+/0}$ denote the LSF for hole capture and emission, respectively, and can be evaluated using

$$f_{i/j}(V_\mathrm{s}) = \frac{1}{2}\sqrt{\frac{c_i\beta}{\pi}} \frac{\mathrm{e}^{-\beta V_{i/j}(V_\mathrm{s})}}{|c_i\Delta q_1 - c_j(\Delta q_1 - q_\mathrm{s})|} \qquad [3]$$

with $i, j = 0, +$. The LSF gives the probability for the thermal excitation up to the intersection point where the charge transfer from the initial charge state i to the final charge state j of the defect (32) can occur. The NMP barrier $V_{i/j}(V_\mathrm{s})$ which must be overcome during this process, is governed by the shape of the corresponding adiabatic potential energy surfaces.

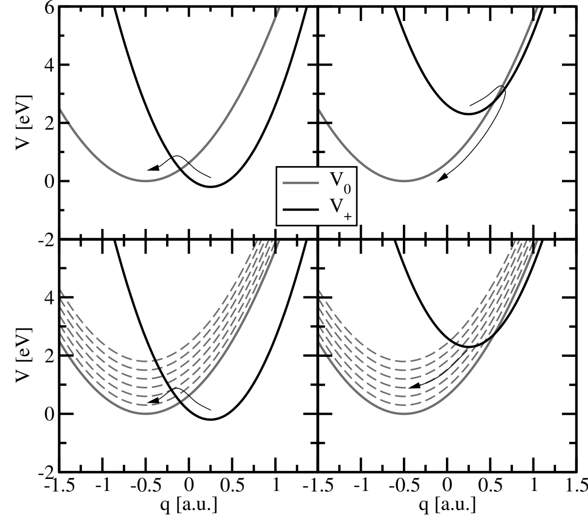

Figure 4: Configuration coordinate diagram for the exemplary case of hole emission, which is depicted for strong (left) and weak (right) electron-phonon coupling considering charge exchange with the whole valence band (bottom) or only the valence band edge (top). As before, the left and the right parabola in each figure shows the adiabatic potential for the case when the hole is located at the valence band edge or in the defect, respectively. The adiabatic potentials of the excited valence band states are represented by the dashed curves. Note that the dominating transition (indicated by the arrows) changes when all band states are considered in the weak electron-phonon coupling regime.

In the NMP theory (11, 12), they are frequently assumed to be harmonic (see Fig. 3) so that they assume a parabolic shape defined by their curvatures c_i and c_j. The relative position of these parabolas are determined by their spatial displacement q_s and their energy separation V_s, which is given by the energy difference between the charge carrier energy and the trap level ($V_\mathrm{s} = E - E_\mathrm{t}$ for hole capture and $V_\mathrm{s} = E_\mathrm{t} - E$ for hole emission). The NMP barrier can be calculated as the energy difference from the energy minimum (V_i, V_j) up to the intersection point (IP) and is given by

$$V_{i/j}(V_\mathrm{s}) = \frac{c_i q_\mathrm{s}^2}{(\frac{c_i}{c_j} - 1)^2} \big(1 \pm \sqrt{\frac{c_i}{c_j} + \frac{V_\mathrm{s}(\frac{c_i}{c_j} - 1)}{c_j q_\mathrm{s}^2}}\big)^2 \, . \qquad [4]$$

Up to now, the multi-state defect model has been used to describe hole capture and emission in BTI, where the trapping dynamics are found to be determined by the regime of strong electron-phonon coupling (see Fig. 4). In this regime, the hole capture and emission (compare to Fig. 4) are dominated by the charge exchange with the valence band edge while the other band states give negligible contributions. Then, the NMP transition rates can be simplified to the analytical expressions published in (14). However, one may also encounter the weak electron-phonon regime when defect behavior in other reliability issues is addressed. In this regime, the dominant transition is hole emission into one of the band states and the integration over the whole valence band as in the equations [1] and [2] must be carried out.

Besides the NMP transitions, metastable states are the other integral part of the multi-state defect model as discussed before. The concept of defects with multiple stable configurations have first been proposed in the Harry Diamond Laboratory (HDL) model (33–36),

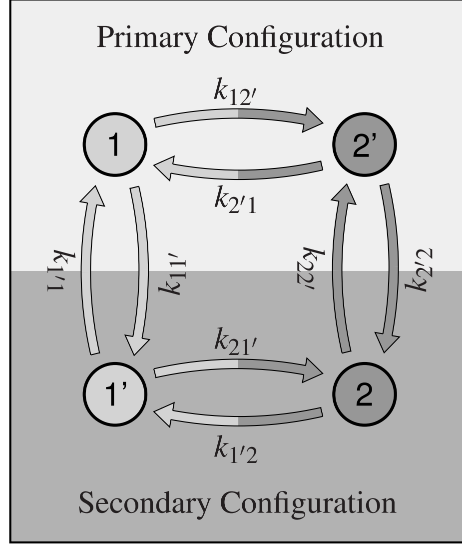

Figure 5: State diagram of the multi-state defect model. The individual states are represented by the circles, where the bright and the dark circles indicate the positive and the neutral charge state of the defect, respectively. The transition rates between two states are denoted by the arrows. The vertical transitions correspond to pure thermal transitions while horizontal ones represent NMP transitions. The latter occur in either the primary (1, $2'$) or the secondary ($1'$, 2) configuration and are associated with the corresponding trap levels E_t and E'_t in the band energy diagram.

which is based on the oxygen vacancy. In the multi-state defect model, this idea has been adopted and generalized to the four-state model for a bistable defect, whose single states are not assigned to the configurations of the oxygen vacancy any more since this defect was found to have a too low trap level (37). This bistable defect is assumed to feature two neutral (1, $1'$) and two positive (2, $2'$) states, where the metastable states are marked by a prime (see Fig. 5). The transitions between different charge states ($1 \leftrightarrow 2'$ and $1' \leftrightarrow 2$) involve a charge transfer and are described by the NMP theory. The transitions $1 \leftrightarrow 1'$ and $2 \leftrightarrow 2'$ are associated with a structural rearrangement of the oxide defect and are modeled using transition state theory (38). Such defects show complex trapping dynamics due to the four states involved and are treated using Markov theory (39), where future transitions depend only on the current state of the investigated system. The time evolution of such a system is described by a Master equation (40), which involves four states for the case of the multi-state defect model. However, the defects dwell in the stable states most of their time while the metastable states are only occupied temporarily. As a result, the four-state model can be simplified to an effective two-state model with the capture and emission times calculated using the concept of first-passage times (40) (see Fig. 6).

$$\tau_\mathrm{c}^{2'} = \frac{k_{12'} + k_{2'1} + k_{2'2}}{k_{12'}k_{2'2}} \quad [5]$$

$$\tau_\mathrm{c}^{1'} = \frac{k_{11'} + k_{1'1} + k_{1'2}}{k_{11'}k_{1'2}} \quad [6]$$

$$\tau_\mathrm{e}^{1'} = \frac{k_{21'} + k_{1'2} + k_{1'1}}{k_{21'}k_{1'1}} \quad [7]$$

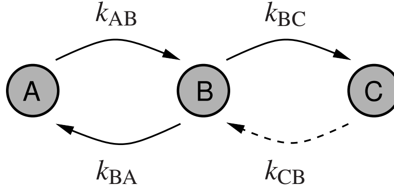

Figure 6: State diagram for a two-step process from the state A to C over the state B. The first-passage time corresponds to the mean time, it takes the defect to arrive at the state C provided that it was initially in the state A and can be calculated as $\tau^{\mathrm{B}} = \tau_{\mathrm{AB}}\tau_{\mathrm{BC}}/(\tau_{\mathrm{AB}} + \tau_{\mathrm{BA}} + \tau_{\mathrm{BC}})$.

$$\tau_{\mathrm{e}}^{2'} = \frac{k_{22'} + k_{2'2} + k_{2'1}}{k_{22'}k_{2'1}} \qquad [8]$$

Here, k_{ij} denotes a transition rate from state i to j and the superscript of $\tau_{\mathrm{c/e}}^{k}$ denotes the metastable state k, which has been passed during the corresponding capture or emission event,. The inverse of the first-passage times are only partial rates of the overall capture and emission rates and must be combined using the equations

$$\frac{1}{\tau_{\mathrm{c}}} = \frac{1}{\tau_{\mathrm{c}}^{1'}} + \frac{1}{\tau_{\mathrm{c}}^{2'}} \qquad [9]$$

$$\frac{1}{\tau_{\mathrm{e}}} = \frac{1}{\tau_{\mathrm{e}}^{1'}} + \frac{1}{\tau_{\mathrm{e}}^{2'}} \qquad [10]$$

to obtain the sought overall hole capture and emission times.

Both parts — the NMP transitions and the bistability of the defect — have been motivated by the findings of TDDS experiments and have therefore been incorporated in the multi-state defect model. In order to evaluate this model, it has been calibrated against the data extracted from TDDS measurements. As it has been demonstrated in Fig. 7, the model yields the correct gate bias and temperature dependence and can reproduce the curvature in the capture times. Furthermore, it also gives an explanation for the two distinct types of defects, the fixed and the switching oxide traps. They differ in the position of their trap level E_{t}', which is associated with the NMP transition $1' \leftrightarrow 2$ in the secondary configuration (see Fig. 5). If this level is far above the substrate Fermi level for a gate bias around the recovery bias, the transition $2 \rightarrow 1'$ is suppressed and the defect follows the pathways $2 \leftrightarrow 2' \rightarrow 1$ in the state diagram of Fig. 5. Then the hole emission process is governed by the thermal transition $2 \rightarrow 2'$, which exhibits no gate bias dependence. The corresponding emission time is independent of the gate bias as it has been observed for the fixed oxide hole trap in the TDDS experiments. In the special case that the trap level E_{t}' lies around the Fermi level, the transition $2 \rightarrow 1'$ is enabled. Provided the barrier $1' \rightarrow 1$ is not too high, the pathway $2 \leftrightarrow 1' \rightarrow 1$ is opened in the state diagram of Fig. 5. Then the defect reacts sensitively to variations of the gate bias, just as the switching oxide hole trap in the TDDS experiments. As discussed in (17), the model also captures the frequency dependence observed in the AC variant of the TDDS as being due to the existence of state $2'$.

The multi-state defect model can not only describe the time constant plots of Fig. 7 but also offers an explanation for temporary and anomalous RTN. The generated noise signal stems from an oxide defect which switches back and forth between states 2 and $1'$ in the secondary configuration. This implies that, first, the defect features a trap level E_{t}' close to the substrate Fermi level when the recovery voltage is applied to the gate, and, second,

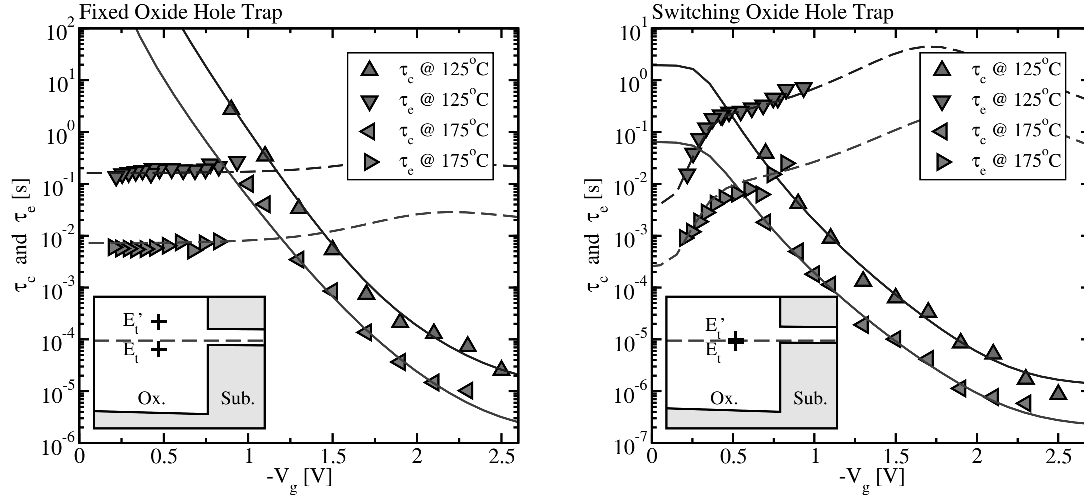

Figure 7: **Left**: Comparison of the simulated capture (solid lines) and emission (dashed lines) times with the measurement data (symbols) for a fixed oxide hole trap. The simulation results are shown to be in remarkable agreement with the experimental data of (14). The trap level E_t', depicted in the inset, is found to be distant from the substrate Fermi level when no voltage stress is applied. This fact characterizes the fixed oxide traps since their emission times become gate bias insensitive under these conditions. **Right**: The same but for a switching oxide hole trap. Compared to the fixed oxide hole trap, emission time constant shows a strong gate bias dependence around the recovery voltage. This is due to the fact that the substrate Fermi level passes the trap level E_t' and thereby enables the pathway $2 \to 1' \to 1$ in the state diagram.

the transitions $2 \leftrightarrow 1'$ must have sufficiently high NMP barriers. If both requirements are met, the noise signal persists as long as the defect dwells in the secondary configuration but vanishes when the defect undergoes a transition into one of two states 1 and $2'$ in the primary configuration. In the case of anomalous RTN, the defect repeatedly changes from the primary configuration to the state $2'$. These transitions are followed by the recurring pauses in the noise signal as depicted in Fig. 8 (left). In the case of temporary RTN, the defect is first forced into the secondary configuration during the stress phase and produces noise at the beginning of the relaxation phase. This noise signal ends when the defect returns to state 1. This qualitative agreement further confirms the validity of the multi-state defect model. It is emphasized here that noise phenomena have increasingly gained importance in the last device technologies as the charging of a single defect can already seriously affect the device characteristics, therefore influencing the device lifetime.

Trap-Assisted Tunneling

The gate leakage current is often ascribed to direct tunneling, which weakly depends on the temperature. Regarding correlated drain and gate current noise (see Fig. 1) discussed before, this behavior is reminiscent of the gate current fluctuations as they are also found to be temperature insensitive. As such, it appears reasonable to also link these fluctuations to direct tunneling for the nFET (41). It may be argued that the charge captured by the defect locally repels the inversion charge. Such a capture event would therefore shift the wavefunction away from the interface and in consequence reduce the tunneling gate current, visible as a step in the leakage current. We remark here that this concept only applies to devices where a charge capture event yields a reduction instead of an increase

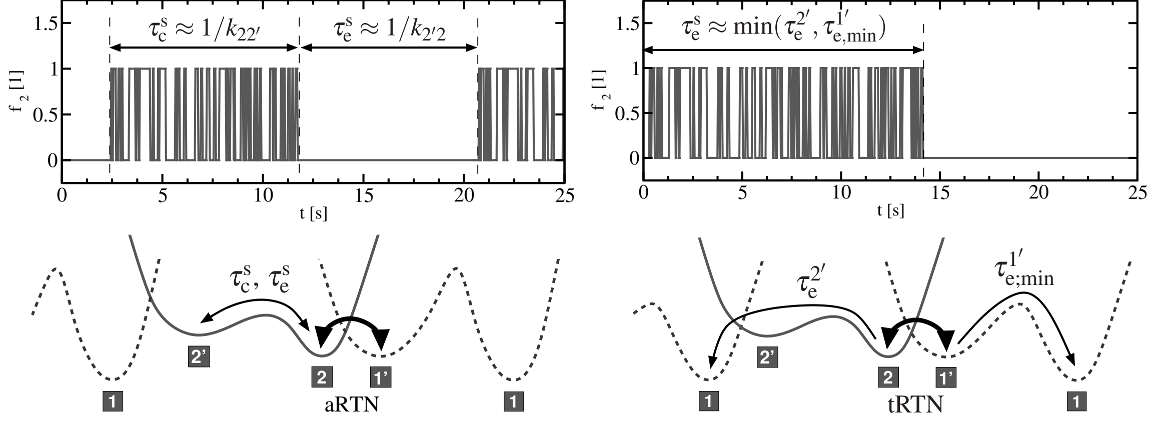

Figure 8: Occupation of the state 2 (top) and the configuration coordinate diagram (bottom) for aRTN (left) and tRTN (right). The RTN noise in the drain current is also reflected in the fluctuations of the trap occupation and is generated by transitions between the states 2 and $1'$. These fluctuations vanish when the defect undergoes a transition into the state 1 or $2'$ of the primary configuration. This usually occurs after a time interval of τ_e^s. While the defect returns back to its secondary configuration (state 2 or $1'$) and produces noise again in the case of aRTN , this does not occur for tRTN, where the defect remains in state 1 and the noise signal has disappeared permanently.

of the gate current. To the best of our knowledge, this is only the case for the nFET studied in (19). Nevertheless, this effect has been investigated on the basis of state-of-the-art non-equilibrium Green's function (NEGF) calculations (41), which is viewed as the most accurate description for direct tunneling at the present. These simulations were carried out for one- and two-dimensional cuts through an nFET assuming different positions of the defect and using a large number of different random dopant configurations. The analysis of these large statistics yielded maximum fluctuations of less than 1% of the total gate current while values of about 75% are seen in experiments. This rules out direct tunneling as an explanation for the gate gate current fluctuations (see Fig. 9).

As the fluctuations in the gate leakage current cannot be explained by direct tunneling (41), they may originate from TAT through one defect. In order to describe these fluctuations in the multi-state defect model, the defect can only carry a TAT current when it assumes one of its secondary configurations $1'$ or 2. Then TAT proceeds as a two-step process based on two consecutive NMP transitions, for instance hole capture from the substrate ($1' \to 2$) followed by hole emission into the poly-gate ($2 \to 1'$). In this case the hole is transferred from the substrate into the poly-gate, however, this fact remains unconsidered in state diagram of the multi-state defect model so far. In the refined state diagram of Fig. 10, the state $1'$ is split into the states $1'_\mathrm{s}$ and $1'_\mathrm{p}$ where subscripts s and p indicate whether the hole is located in the substrate or in the poly-gate, respectively. In previous works (24, 25, 44), the TAT current is usually modeled as

$$I_\mathrm{g} = q_0 \left(\frac{k_{1'_\mathrm{s}2} k_{21'_\mathrm{p}}}{k_{1'_\mathrm{s}2} + k_{21'_\mathrm{p}}} - \frac{k_{1'_\mathrm{p}2} k_{21'_\mathrm{s}}}{k_{1'_\mathrm{p}2} + k_{21'_\mathrm{s}}} \right), \qquad [11]$$

where the two terms represent the opposite directions of the current. It is emphasized that each term is derived from a two-step process which only involves the forward rates of the overall process as illustrated in Fig. 11. For a pFET this means that a positive gate current constitutes of two consecutive NMP transitions, namely hole capture $1'_\mathrm{s} \to 2$

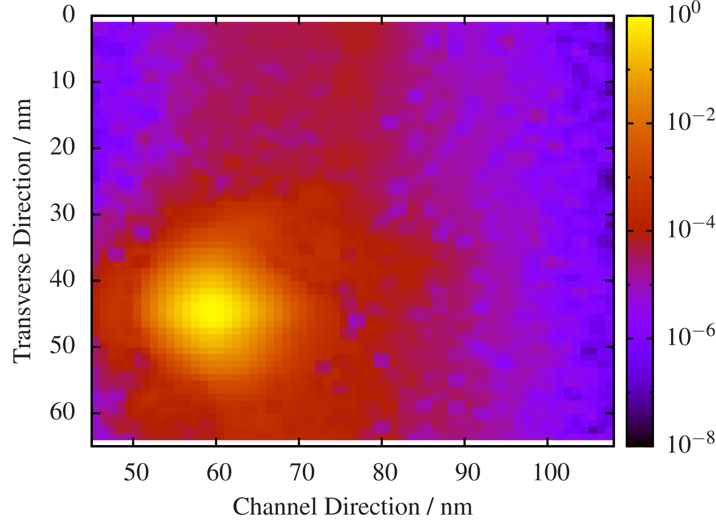

Figure 9: Relative reduction (in percent) of the gate current density due to the charging of one defect. NEGF calculations were employed to calculate the gate tunneling current for several random dopants configurations (42, 43) — once when the defect is charged and once when it is uncharged. For the worst-case dopant configuration, the obtained change in the gate current amounts to a reduction of less than 1% while a value around 75% is found experimentally (19). This indicates that the simple electrostatic picture can not explain the observed gate current fluctuations.

directly followed by hole emission $2 \rightarrow 1'_{\mathrm{p}}$. An improved description of TAT is based on the use of first-passage times, which account for the emission of the captured charge carrier into the substrate as well as into the poly-gate.

$$I_{\mathrm{g}} = q_0 \left(\frac{k_{1'_{\mathrm{s}}2} k_{21'_{\mathrm{p}}}}{k_{1'_{\mathrm{s}}2} + k_{21'_{\mathrm{s}}} + k_{21'_{\mathrm{p}}}} - \frac{k_{1'_{\mathrm{p}}2} k_{21'_{\mathrm{s}}}}{k_{1'_{\mathrm{p}}2} + k_{21'_{\mathrm{p}}} + k_{21'_{\mathrm{s}}}} \right) \quad [12]$$

In the above example, the hole originating from the substrate is emitted back into the substrate before it continues to the poly-gate, thereby possibly reducing the overall TAT current by many orders of magnitude (see Fig. 11). However, a proper derivation of TAT starts from the rate equation

$$\partial_t f_{\mathrm{t}} = (k_{21'_{\mathrm{s}}} + k_{21'_{\mathrm{p}}})(1 - f_{\mathrm{t}}) - (k_{1'_{\mathrm{s}}2} + k_{1'_{\mathrm{p}}2}) f_{\mathrm{t}} \ , \quad [13]$$

where f_{t} denotes the electron occupancy of the trap. For steady state conditions ($\partial f_{\mathrm{t}}/\partial t = 0$), the trap occupancy reads

$$f_{\mathrm{t}} = \frac{k_{21'_{\mathrm{s}}} + k_{21'_{\mathrm{p}}}}{k_{1'_{\mathrm{s}}2} + k_{21'_{\mathrm{s}}} + k_{1'_{\mathrm{p}}2} + k_{21'_{\mathrm{p}}}} \ , \quad [14]$$

which is inserted into

$$I_{\mathrm{g}} = q_0 \left(k_{21'_{\mathrm{p}}} (1 - f_{\mathrm{t}}) - k_{1'_{\mathrm{p}}2} f_{\mathrm{t}} \right) \quad [15]$$

resulting in

$$I_{\mathrm{g}} = q_0 \frac{k_{1'_{\mathrm{s}}2} k_{21'_{\mathrm{p}}} - k_{1'_{\mathrm{p}}2} k_{21'_{\mathrm{s}}}}{k_{1'_{\mathrm{s}}2} + k_{21'_{\mathrm{s}}} + k_{1'_{\mathrm{p}}2} + k_{21'_{\mathrm{p}}}} \ . \quad [16]$$

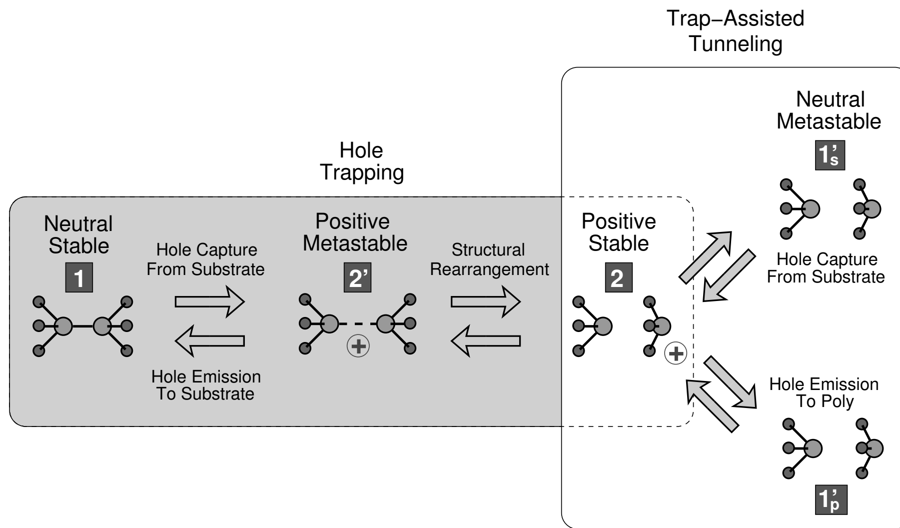

Figure 10: Refined state diagram of the multi-state defect model. Hole capture and emission still occur as transitions between the states 1 and 2 over the metastable state $2'$. However, the state $1'$ has to be split in order to differentiate whether the hole is located in the substrate (s) or in the poly-gate (p). With this modification in the state diagram, one can define a TAT process in which the hole is transferred from the substrate into the poly-gate and can therefore contribute to a gate current.

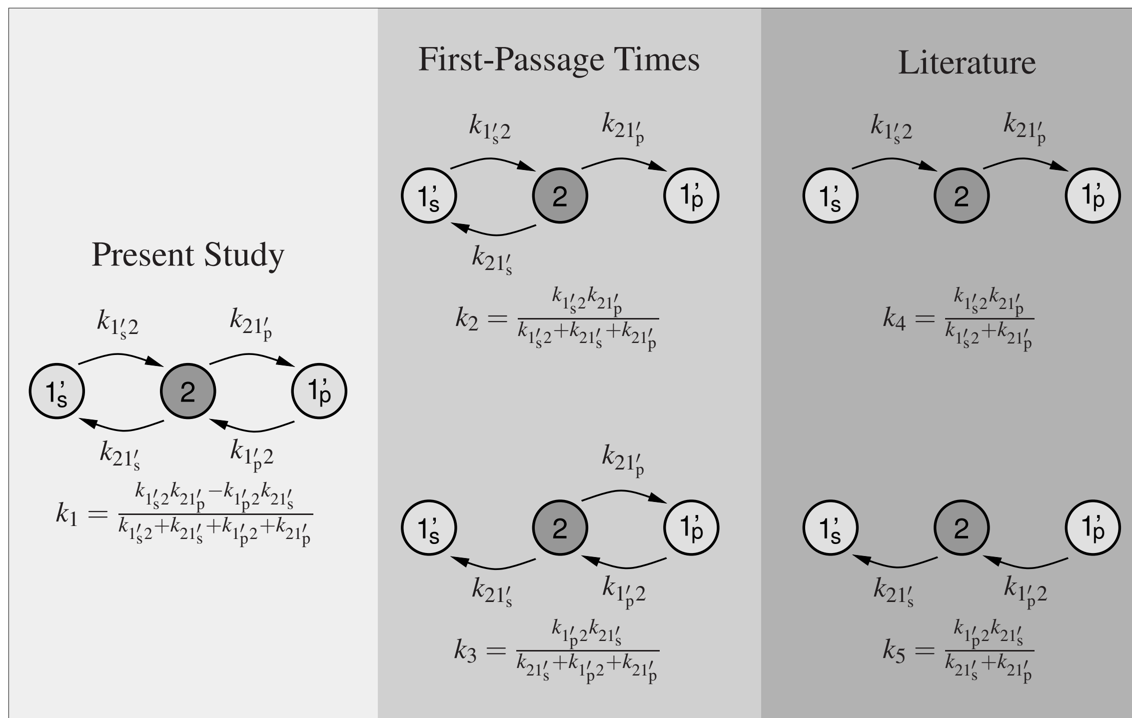

Figure 11: State diagrams and their corresponding rate expressions for the conventional model (right) used in previous attempts (24, 25, 44), the approximation based on first-passage times (middle), and the complete description of TAT used in this paper (left).

The expression above covers all four possible transitions ($1'_\mathrm{s} \to 2$, $2 \to 1'_\mathrm{s}$, $1'_\mathrm{p} \to 2$, $2 \to 1'_\mathrm{p}$) and is therefore a complete description of TAT — apart from approximations employed in the used NMP formulation. The importance to use the accurate expression [16] is pointed out in Fig. 12. The conventional description based on the expression [11] strongly overestimates the TAT current if the trap level lies below the substrate valence band edge. By contrast, the first-passage times still provide a good approximation in this case. When the trap level comes into the energy range of the substrate band gap, also the first-passage times fail since they yield a non-vanishing TAT current for a zero gate bias. As a consequence, the complete description of equation [16] was employed to calculate the TAT current in this work.

In order to obtain appreciable TAT currents, the NMP transition rates $1'_\mathrm{s} \leftrightarrow 2$ and $1'_\mathrm{p} \leftrightarrow 2$ must be large compared to the capture ($1/\tau_\mathrm{c}$) and emission ($1/\tau_\mathrm{e}$) rates. Therefore, these NMP transition rates govern the occupancy of the state 2 from which hole emission starts. Since the concept of first-passage times assumes that the defect must initially be in state 2 for hole emission, the corresponding emission rate $1/\tau_\mathrm{e}^{2'}$ must be multiplied by the hole occupation probability $1 - f_\mathrm{t}$, resulting in the correction

$$\tau_\mathrm{e}^{2'} \to \tau_\mathrm{e}^{2'}/(1 - f_\mathrm{t}) \qquad [17]$$

for the equation [8].

The multi-state defect model was tested for its ability to correctly predict the gate leakage fluctuations. For this purpose, the model was evaluated against the experimental data for the nanoscaled pFET device investigated in (19). The model was found to correctly reproduce the gate bias and temperature dependence observed for the magnitude of the gate current fluctuations (see Fig. 13). This may appear surprising since the data exhibit no temperature activation, which is, at a first glance, in contradiction to the temperature sensitivity of the NMP processes. The defect in these fits is located close to the substrate interface ($x_\mathrm{t} = 4\,\mathrm{\AA}$ measured from the substrate interface) so that the hole capture and emission with the substrate occurs on much shorter timescales than the hole emission into the poly-gate. As a consequence, the gate leakage is controlled by the NMP transition rate $2 \to 1'_\mathrm{p}$, whose configuration coordinate diagram is depicted in Fig. 14. This process is found to be in the weak electron-phonon coupling regime. As already explained, all band states must be considered in this case and the integral in equation [2] has its dominant contributions from the band states that intersect the parabola of state 2 around its minimum. These transitions feature negligible NMP barriers so that the overall transition shows no temperature activation. It is noted that the multi-state defect model also yields a good agreement for the capture and emission times (see Fig. 15) extracted from the noise measurement using the same set of parameters (listed in Tab. 1). The simulations were carried out for the same set of parameters in order to ensure that the model captures the temperature and gate bias trends in both the drain and the gate current noise at the same time. As such, the multi-state defect model has been shown to also capture the correlated drain and gate current noise seen in experiments. This is an important finding as these results require and therefore also confirm the bistability of the oxide defects, already seen in BTI and RTN.

Conclusions

The multi-state defect model relies on two main assumptions, motivated by the findings from the TDDS experiments: First, the NMP transitions describe the charge transfer to

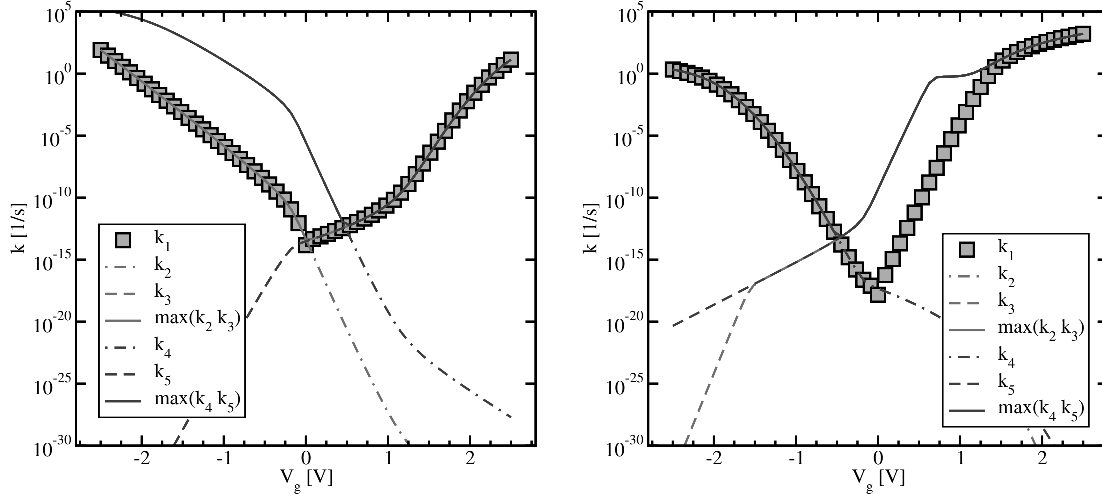

Figure 12: The absolute values of the simulated TAT current rates, evaluated using the conventional model (dark lines, k_4 and k_5), first-passage times (bright lines, k_2 and k_3), and the complete description (symbols, k_1) for the cases when the trap level lies below (left) or inside the substrate band gap (right). The left figure illustrates the weakness of the conventional model, which strongly overestimates the hole current from the substrate to the gate. This is because this model ignores that holes, which are captured by the defect ($1'_\mathrm{s} \to 2$), can be emitted back into the substrate ($2 \to 1'_\mathrm{s}$) instead of going into the poly-gate ($2 \to 1'_\mathrm{p}$) and therefore do not contribute to the gate current. The right figure depicts a case where the first-passage times fail to predict the correct gate current. The first-passage times do not account for the possibility that a hole can be injected from the substrate ($1'_\mathrm{s} \to 2$) and thus occupies the state 2, thereby blocking a hole current from the poly-gate to the substrate ($1'_\mathrm{p} \leftrightarrow 2 \to 1'_\mathrm{s}$) over this state. An analogous consideration also holds true for electrons and explains the overestimated TAT rates based on first-passage times in the above simulation. To conclude, only the complete description predicts the minimum TAT current for a zero gate bias in both cases and is therefore used in this work.

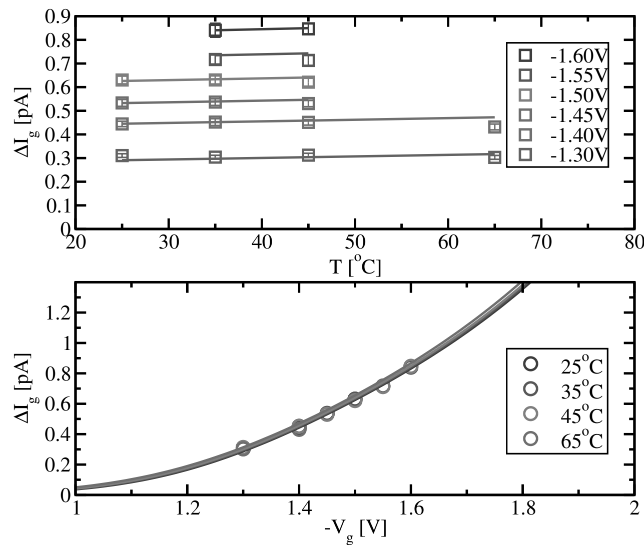

Figure 13: Step heights of the gate current fluctuations plotted as a function of temperature (upper panel) and as a function of the applied gate bias (lower panel). It is shown that the multi-state defect model (lines) captures the weak temperature dependence seen in the experimental data (symbols).

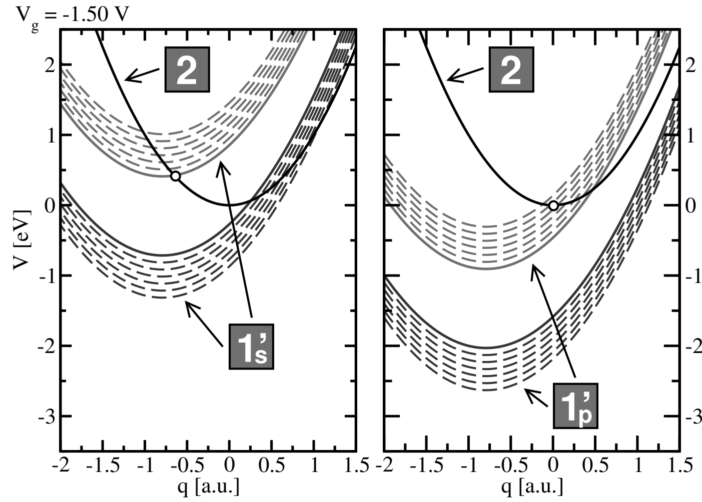

Figure 14: Adiabatic potential energy surfaces extracted from the fits to the experimental data for the pFET in (19). Our simulations consider NMP transitions from all conduction (dashed lower parabolas) and valence (dashed upper parabolas) band states involving charge carriers from the substrate (left) and the poly-gate (right). Note that the hole exchange with the substrate valence band is in the strong electron-phonon coupling regime as usually assumed for hole capture and emission in BTI. However, weak electron-phonon coupling is found for hole exchange with the poly-gate. Since the full width of the valence band states is taken into account, there always exists one parabola that cuts the right parabola in its minimum. Then hole emission to the poly-gate proceeds without a barrier and therefore shows no temperature dependence. As this transition dominates the TAT current, the multi-state defect model can explain the temperature insensitive leakage currents.

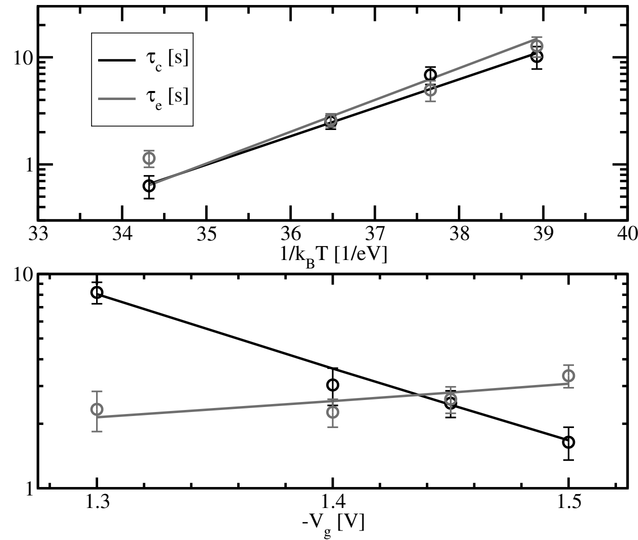

Figure 15: Comparison of measured (symbols) and simulated (lines) for τ_c and τ_e. Good agreement with the experimental data is obtained for the temperature (upper panel) and the gate bias (lower panel) dependence.

Transition $1 \leftrightarrow 2'$		Transition $1' \leftrightarrow 2$	
E_t	-0.27 eV	E'_t	0.15 eV
c_1	1.00 eV/Å^2	$c_{1'}$	0.71 eV/Å^2
$c_{2'}$	1.64 eV/Å^2	c_2	1.00 eV/Å^2
$q_{12'}$	0.90 Å	$q_{1'2}$	1.39 Å
Other Parameters			
$\varepsilon_{\mathrm{T}2}$	0.65 eV	$\varepsilon_{2'2}$	0.10 eV
m_t	0.95	$\varepsilon_{11'}$	10.0 eV

Table 1: Model parameters used in Fig. 13 and 14. The shapes of the parabolic adiabatic potentials can be also given by the quantities $S\hbar\omega = 0.81$ eV, $R = 0.61$, $S'\hbar\omega = 1.37$ eV, and $R' = 0.71$.

and from the oxide defect and, second, the defects have a bistable configuration in two charge states. Both assumptions are required to reproduce the temperature and the gate bias dependence of the hole capture and emission times extracted from the TDDS and they give an explanation for the occurrence of fixed and switching oxide hole traps as well as for anomalous and temporary RTN. As has been demonstrated, this model also explains the correlated drain and gate current noise observed in the newest device technologies. The gate current fluctuations have been found to originate from TAT, where the single steps of the overall process are due to NMP transitions. At a first glance, the temperature insensitivity of the gate current fluctuations appear to be in stark contrast to the expected temperature dependence of the NMP processes in the multi-state defect model. However, this temperature dependence is merely an artifact of neglecting the NMP transitions to and from the excited conduction and valence band states. Given this fact, we suspect not only the gate current fluctuations but also the gate leakage current itself to be caused by this type of oxide defects. From a wider perspective, the success of the multi-state defect model indicates that the bistable defect configurations are not a feature solely seen in BTI. Rather, these oxide defects are likely to also play a dominant role in several other reliability issues apart from BTI.

Acknowledgements

This work has received funding from the Austrian Science Fund (FWF) project n°23390-N24 and the European Community's FP7 n°261868 (MORDRED).

References

1. D. Ang, S. Wang, and C. Ling, *IEEE Elect.Dev.Let.* **26**, 906–908 (2005).
2. V. Huard, M. Denais, and C. Parthasarathy, *Microelectron.Reliab.* **46**, 1–23 (2006).
3. T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, *Microelectron.Eng.* **86**, 1876–1882 (2009).
4. V. Huard, *Proc.IRPS*, 2010, pp. 33 –42.
5. D. Ang, Z. Teo, T. Ho, and C. Ng, *IEEE Trans.Dev.Mater.Rel.* **11**, 19–34 (2011).
6. T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer, *Proc.IRPS*, 2011, pp. 6A.2.1–6A.2.9.
7. K. Jeppson, and C. Svensson, *J.Appl.Phys.* **48**, 2004–2014 (1977).

8. M. Alam, *Proc.IEDM*, 2003, pp. 345–348.
9. T. Grasser, W. Goes, and B. Kaczer, "Towards Engineering Modeling of Negative Bias Temperature Instability," *Defects in Microelectronic Materials and Devices*, edited by D. Fleetwood, R. Schrimpf, and S. Pantelides, Taylor and Francis/CRC Press, 2008, pp. 1–30.
10. T. Grasser, B. Kaczer, W. Goes, H. Reisinger, T. Aichinger, P. Hehenberger, P. Wagner, F. Schanovsky, J. Franco, M. Luque, and M. Nelhiebel, *IEEE Trans.Elect.Dev.* **58**, 3652–3666 (2011).
11. K. Huang, and A. Rhys, *Proc. Royal Society London. Ser.A* **204**, 406–423 (1950).
12. A. Palma, A. Godoy, J. A. Jimenez-Tejada, J. E. Carceller, and J. A. Lopez-Villanueva, *Phys.Rev.B* **56**, 9565–9574 (1997).
13. T. Grasser, K. Rott, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, M. Waltl, M. Toledano-Luque, and B. Kaczer, *Proc.IRPS*, 2013, pp. 2D.2.1–2D.2.7.
14. T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, *Proc.IRPS*, 2010, pp. 16 –25.
15. H. Reisinger, T. Grasser, K. Ermisch, H. Nielen, W. Gustin, and C. Schlunder, *Proc.IRPS*, 2011, pp. 6A.1.1–6A.1.8.
16. M. Toledano-Luque, B. Kaczer, P. Roussel, T. Grasser, G. Wirth, J. Franco, C. Vrancken, N. Horiguchi, and G. Groeseneken, *Proc.IRPS*, 2011, pp. 4A.2.1–4A.2.8.
17. T. Grasser, B. Kaczer, H. Reisinger, P.-J. Wagner, and M. Toledano-Luque, *Proc.IRPS*, 2012, pp. XT.8.1–XT.8.7.
18. T. Grasser, H. Reisinger, K. Rott, M. Toledano-Luque, and B. Kaczer, *Proc.IEDM*, 2012, pp. 19.6.1–19.6.4.
19. M. Toledano-Luque, B. Kaczer, E. Simoen, R. Degraeve, J. Franco, P. Roussel, T. Grasser, and G. Groeseneken, *Proc.IRPS*, 2012.
20. L. K. J. Vandamme, and F. N. Hooge, *IEEE Trans.Elect.Dev.* **55**, 3070–3085 (2008).
21. M. Kirton, and M. Uren, *Adv.Phys.* **38**, 367–486 (1989).
22. A. Avellan, D. Schroeder, and W. Krautschneider, *J.Appl.Phys.* **94**, 703–708 (2003).
23. N. Zanolla, D. Siprak, P. Baumgartner, E. Sangiorgi, and C. Fiegna, *Ultimate Integration of Silicon*, 2008, pp. 137–140.
24. F. Jimenez-Molinos, A. Palma, F. Gamiz, J. Banqueri, and J. A. Lopez-Villanueva, *J.Appl.Phys.* **90**, 3396–3404 (2001).
25. L. Vandelli, A. Padovani, L. Larcher, R. Southwick, W. Knowlton, and G. Bersuker, *Proc. European Solid-State Device Research Conference (ESSDERC)* , 2010, pp. 388–391.
26. L. Vandelli, A. Padovani, L. Larcher, R. Southwick, W. Knowlton, and G. Bersuker, *IEEE Trans.Elect.Dev.* **58**, 2878–2887 (2011).
27. C.-Y. Chen, Q. Ran, H.-J. Cho, A. Kerber, Y. Liu, M.-R. Lin, and R. Dutton, *Proc.IRPS*, 2011.
28. B. Kaczer, M. Toledano-Luque, W. Goes, T. Grasser, and G. Groeseneken, *Microelectron.Eng.* **109**, 123–125 (2013).
29. X. Ji, Y. Liao, C. Zhu, J. Chang, F.Yan, Y. Shi, and Q. Guo, *IPRS*, 2013, pp. XT.7.1–XT.7.5.
30. M. Andersson, Z. Xiao, S. Norrman, and O. Engström, *Phys.Rev.B* **41**, 9836–9842 (1990).

31. F. Schanovsky, O. Baumgartner, V. Sverdlov, and T. Grasser, *Journ. of Computational Electronics* **11**, 218–224 (2012).
32. W. Goes, F. Schanovsky, and T. Grasser, "Advanced Modeling of Oxide Defects," *The Bias Temperature Instability: Experiment, Theory, and Modeling for Devices and Circuits*, T. Grasser (Ed.), Springer-Verlag, 2013.
33. A. Lelis, and T. Oldham, *IEEE Trans.Nucl.Sci.* **41**, 1835–1843 (1994).
34. Conley, and Lenahan, *IEEE Trans.Nucl.Sci.* **42**, 1744–1749 (1995).
35. J. Campbell, and P. Lenahan, *Proc.IIRW*, 2007, pp. 12–17.
36. J. Ryan, P. Lenahan, T. Grasser, and H.Enichlmair, *Proc.IRPS*, 2010, pp. 43–49.
37. F. Schanovsky, W. Goes, and T. Grasser, *Journ. of Computational Electronics* **9**, 135–140 (2010).
38. A. Nitzan, *Chemical Dynamics in Condensed Phases*, Oxford University Press, 2006.
39. O. Ibe, *Markov Processes for Stochastic Modeling*, Academic Press, 2009.
40. T. Grasser, *Microelectron.Reliab.* **52**, 39–70 (2012).
41. O. Baumgartner, M. Bina, M. Toledano-Luque, W. Goes, F. Schanovsky, B. Kaczer, and T. Grasser, Direct Tunneling and Gate Current Fluctuations, accepted for SISPAD 2013.
42. M. Karner, A. Gehring, S. Holzer, M. Pourfath, M. Wagner, W. Goes, M. Vasicek, O. Baumgartner, C. Kernstock, K. Schnass, G. Zeiler, T. Grasser, H. Kosina, and S. Selberherr, *Journ. of Computational Electronics* **6**, 179–182 (2007).
43. O. Baumgartner, M. M. Karner, and H. Kosina, *Proc.SISPAD*, 2008.
44. S. Takagi, N. Yasuda, and A. Toriumi, *IEEE Trans.Elect.Dev.* **46**, 348–354 (1999).