Charge Trapping
and
the Negative Bias Temperature Instability

W. Goes[*], F. Schanovsky[°], Ph. Hehenberger[°], P.-J. Wagner[*], and T. Grasser[°]

[*] Christian Doppler Laboratory for TCAD at the Institute for Microelectronics, TU Wien, Wien, Austria

[°] Institute for Microelectronics, TU Wien, Wien, Austria

During the last couple of years new measurement techniques have provided insight into the physics behind the negative bias temperature instability (NBTI) and indicate that the recoverable component of NBTI is due to some kind of charge trapping. As a consequence, charge trapping processes have been investigated and modeled in detail. We review the evolution of the latest charge trapping model by focussing on the correct temperature- and field dependence as well as on the quality of agreement with experimental stress/relaxation curves.

Introduction

The discovery of the negative bias temperature instability (NBTI) dates back to the middle of the last century (1–4). The phenomenon attracted little attention for a long period of time but in the meantime has risen to a dominant present-day reliability concern. Since reliable lifetime projection of CMOS devices or even suppression of NBTI is sought for industrial purposes, the physical origins are of great interest. Electrical measurements can only determine variations of charges in the dielectric but do not provide a detailed insight into the underlying physics. As a consequence, modeling plays a central role in our understanding of NBTI.

In the early NBTI literature, the conventional measure/stress/measure (MSM) (5) technique was used to assess NBTI degradation. The focus was on the stress phase whose temporal evolution was fitted to a power law $\Delta V_{\text{th}} = A \times t^n$. These measurements delivered an extracted power law exponent of approximately 0.25, which is consistent with the first and simplest variant of the popular reaction-diffusion (RD) model (6–9). Therein, NBTI is traced back to charge carriers captured in silicon dangling bonds which have been identified with the so-called P_{b} centers in electron spin resonance (ESR) measurements (10). During fabrication, these dangling bonds are saturated by hydrogen atoms which moves the electron level out of the silicon bandgap. In the stress phase, the RD model assumes that a reaction takes place which releases the hydrogen atoms from the P_{b} centers (10) and triggers a hydrogen diffusion away from the interface. This reaction is initiated by holes from the inversion layer but is stopped during relaxation due to the absence of holes. Then these bonds are annealed by hydrogen atoms which diffuse back from deep in the dielectric to the interface and saturate the P_{b} centers again. While the interface reaction is assumed to be in equilibrium, the reaction kinetics are controlled by the diffusion of hydrogen from or to the interface. Interestingly, first principles investigations predict hydrogen to be only stable as a positive atom or a molecule (11, 12).

Furthermore, the estimated barrier for the hole-assisted hydrogen release (13) is too high to be overcome within timescales of interest for NBTI. Even though the RD model is in contradiction to these findings, the concept of the RD model is still in widespread use today.

Detailed MSM studies have revealed that the degradation recovers appreciably during the time between the removal of stress and the actual measurement at the readout voltage. The unrecorded relaxation during this delay cannot be experimentally assessed and is thus not considered in the extraction of the power law exponent n. Improved MSM measurements with a minimum delay in the order of $1\,\mu$s (14) yield exponents in the range $0.1 - 0.15$. The on-the-fly (OTF) method does not suffer from the interruption of stress since it directly monitors the degradation of the drain current I_D in the linear regime and thereby avoids any unrecorded relaxation. The measured degradation is then converted to the threshold voltage using the following SPICE-level-1 conversion scheme:

$$\Delta V_{\mathrm{th}}^{\mathrm{OTF}} \approx \frac{I_D - I_{D,0}}{I_{D,0}}(V_G - \Delta V_{\mathrm{th},0}^{\mathrm{OTF}}), \qquad [1]$$

where $\Delta V_{\mathrm{th},0}$, V_G, and $I_{D,0}$ denote the threshold voltage of the undegraded device, the applied gate bias, and the first recorded drain current, respectively. Although this technique does not suffer from artifacts due to the recovery during the measurement, it is affected by mobility variations and inaccuracies in $I_{D,0}$ (15–17). The power law exponents extracted from both OTF and MSM measurements ranged around 0.15 (18, 19) and were related to a modified RD model predicting $n = 1/6$. In this model, molecular hydrogen is assumed to be the diffusing species reacting with the P_b centers.

In the following, the focus is moved to the NBTI recovery phase which exhibits an interesting universal pattern (14, 20–22) independent of the applied gate bias: The relaxation shows a logarithmic time behavior that appears as straight lines in a time-logarithmic plot and spans over 11 decades, before it levels off. This pattern has been identified with the recoverable component of NBTI (22), however, there also exists an additional component which is permanent or at least slowly recovering. Most importantly, the curve shapes do not disclose the underlying microscopic process but can be used as a benchmark for physics-based models. For instance, the recovery lasting about 11 decades is inconsistent with the 4 decades predicted by the RD model. Therefore the experimental data could not be reconciled with the aforementioned RD models which triggered some new attempts to maintain this model by small refinements or extensions:

- One variant of the RD model rests on the assumption that the dielectric is populated by both atomic and molecular hydrogen and their $H \leftrightarrow H_2$ interconversion not only takes place at the interface but also deep in the dielectric. As found in (14, 23, 24) via rigorous simulations, this model predicts basically the same relaxation behavior as the normal RD model and thus still contradicts the experimental relaxation behavior.

- It has been speculated that an appreciable amount of hydrogen can be trapped at oxide defects, thus is retained from the interface during recovery, and prolongs the recovery in this way. The most sophisticated model, the reaction dispersive
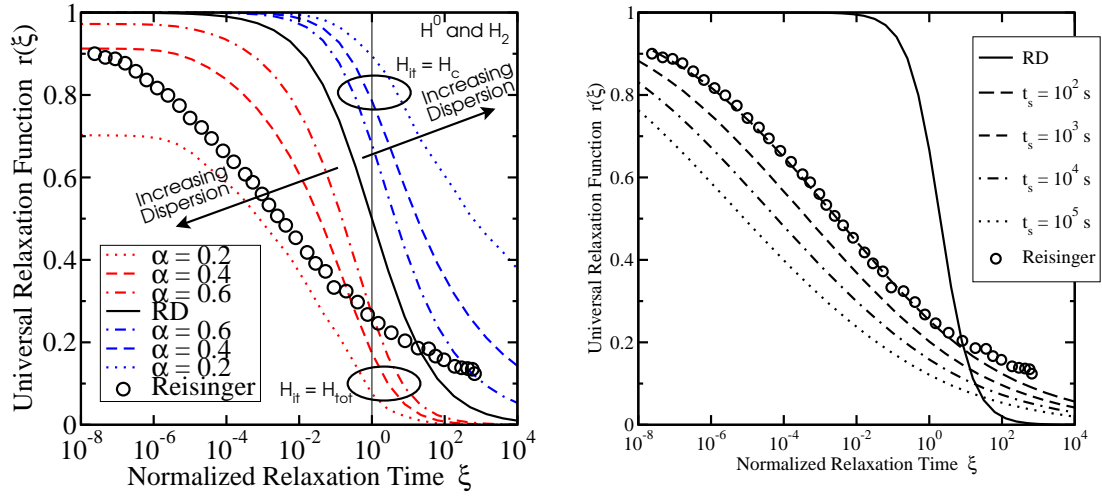
Figure 1: Left: Comparison between the RDD model and experimental recovery data. The solid line stands for the RD model whereas the other lines correspond to the case where only untrapped hydrogen H_c or the the total hydrogen H_{tot} participates in the interface reaction. The circles are measurement data from Reisinger *et al.* (33). Although the slope of the simulation and the experimental data give a similar slope in a certain time range, the overall slopes of both curves do not match. **Right**: The dispersive rate model evaluated against experimental data. A good fit can be obtained for a single recovery curve.

diffusion (RDD) model (5, 25–27), has been developed based on simplified versions of multiple-trapping (MT) theory. Therein, hydrogen can be captured in traps of different energetical depths, leading to a wide distribution of detrapping times. This concept of hydrogen transport reflects spatially and temporally strongly varying effective diffusion constants in the oxide on a microscopic level. As can be seen in Fig. 1, the time range of recovery is significantly extended but still does not follow the experimental relaxation behavior.

- In other variants of the RD model, emphasis is changed to the interface reaction rather than the hydrogen transport. It is argued that due to the amorphous nature of the interfacial layer at the substrate, the dissociation barrier height shows a wide dispersion (28, 29). Although the dispersive rate model (30, 31) already delivers quite a good agreement for the relaxation phase (see Fig. 1), in (14) it is shown that this model yields a reduced power-law exponent of $n \approx 0.03$ during stress. An extended model, the triple well model (32), accounts for a further hydrogen transition, which follows the interface reaction and has been identified as the permanent component of NBTI. It has already been capable of reproducing complex stress/relaxation curves for distinct temperatures, however, no satisfying agreement with experimental data obtained for different gate voltages could be achieved.

None of the above listed RD variants seems to capture the main part of the physics behind NBTI where the universal relaxation is viewed as the most challenging obstacle. Since the RD model appears to be incompatible with the experimental data, a great deal of effort has been put into the search for alternative explanations. All of them rely on
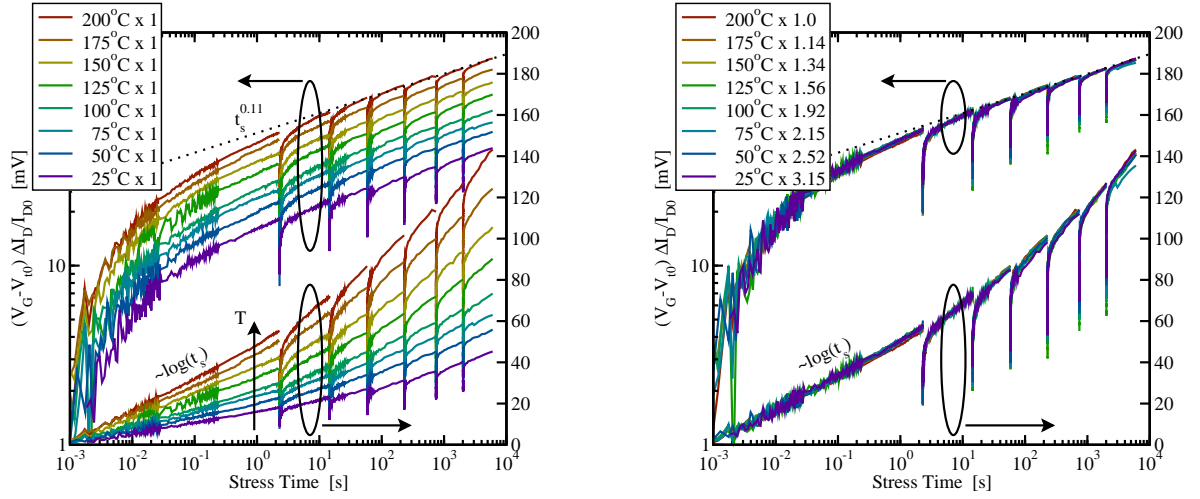
Figure 2: Left: Degradation during subsequent stress/relax cycles measured by the eMSM technique (5, 34). The temperature ranges from 25°C to 200°C for a fixed gate voltage of 2 V. The logarithmic behavior during the initial degradation changes into a power law behavior at long-time stress. **Right**: The same data as on the left-hand side but multiplied with the factors shown in the top left insert. The scaling results in a perfect overlap of the curves and indicates the dominant trapping process that is amplified for higher temperatures.

some kind of charge trapping — with differences in their exact description of the charge trapping process. Charge trapping has always been envisioned as the capture of charge carriers, which stem from the substrate and move into defects located within the dielectric. This mechanism behind charge trapping always involves quantum mechanical tunneling even if it may not be the rate-limiting factor but is expected to make up the core of NBTI degradation. Also, it is held responsible for the recoverable part of NBTI. Nevertheless this does not preclude the existence of a permanent component, most likely the generation of interface states via some rate-limited process.

The remainder of this review is organized as follows: The next chapter gives an overview of the status quo with respect to the experimental side. This includes the experimentally observed stress/relaxation patterns as well as a recently developed measurement technique that provides insight into the physics behind NBTI. In the following chapters, several different charge trapping models will be presented and their underlying concepts will be outlined. Furthermore, their results will be discussed and compared to the experimental findings used as a benchmark for these models.

Experimental Findings

Extended MSM Technique

A valuable reliable set of measurement data can be obtained from the extended MSM (eMSM) technique (5, 34), which combines the advantages of the OTF method during stress and the conventional MSM method during relaxation. During stress the OTF method is employed while one changes to the conventional MSM approach when proceeding to recovery. Based on the eMSM technique, striking correlations in the behavior

of stress/relaxation curves for various different combinations of stress temperatures and voltages have been observed. The resultant findings, serving as criteria for any NBTI modeling attempt, are delineated in the following:

- The data in Fig. 2 demonstrate that the shape of the stress/relaxation curves remains the same for all recorded temperatures (35, 36) up to medium stress times. This peculiarity is best visualized by the complete overlap of these curves when they are scaled by a suitably chosen factor. As shown in (36), the congruence of the scaled data can also be achieved for different stress voltages. Mathematically the temperature and voltage dependence can thus be expressed by

$$\Delta V_{\text{th}}(t, T_{\text{s}}, V_{\text{s}}) = f(T_{\text{s}}, V_{\text{s}}) \times \Delta \widetilde{V}_{\text{th}}(t), \qquad [2]$$

where $f(T_{\text{s}}, V_{\text{s}})$ represents a stress voltage V_{s} and temperature T_{s} dependent scaling factor. The short-time stress phase can be well fitted by a logarithmic time-dependence (37)

$$\Delta \widetilde{V}_{\text{th,stress}}(t_{\text{s}}) = A_{\text{s}} \log(t_{\text{s}}/t_0), \qquad [3]$$

where A_{s} and t_0 stand for an unspecified prefactor and the delay of the first OTF measurement point. In the long-time part of the stress phase, a slight deviation from this behavior can be recognized, which follows a power-law (37). As pointed out in (38), it shows a strong temperature dependence and most probably originates from a permanent component. The relaxation phase is again split into two parts

$$\Delta \widetilde{V}_{\text{th,relax}}(t_{\text{s}}, t_{\text{r}}) = A_{\text{r}} \log(1 + t_{\text{s}}/t_{\text{r}}) + P(t_{\text{s}}), \qquad [4]$$

where A_{r} represents a prefactor for the relaxation phase and $P(t_{\text{s}})$ denotes the stress time-dependent permanent component. Note that the degradation during stress is produced much faster than it recovers during relaxation. This aspect is reflected in the fact that the time slopes in the logarithmic plots have a certain ratio $A_{\text{s}}/A_{\text{r}} \approx 2.5$ (39).

- The temperature and field dependence of the degradation is an important criterion and has attracted increasing attention during the last few years. Their scaling factor has been empirically determined and approximated as $f(T_{\text{s}}, V_{\text{s}}) \approx A_{\text{s},0} T^2 F^2$ (36, 38, 39). In recently published investigations with emphasis on the long-term degradation of NBTI, a F^4-dependence of the permanent component and a temperature activation in the range of $0.15 - 0.25\,\text{eV}$ has been reported (38).

- During recent years the idea of two processes contributing to the NBTI degradation has been frequently proposed (31). One is attributed to hole trapping into pre-existing traps deep in the dielectric (19, 31, 40, 41), while the other is identified with the creation of interface states (6–9). In this context the question arises whether these two components are acting independently (17, 42) or coupled. The scalability of stress/relaxation curves suggests that the NBTI dynamics is governed by two interacting mechanisms, most probably with one process triggering the other (36, 39). However, due to the lack of concise data, this question is still open for discussions (38).
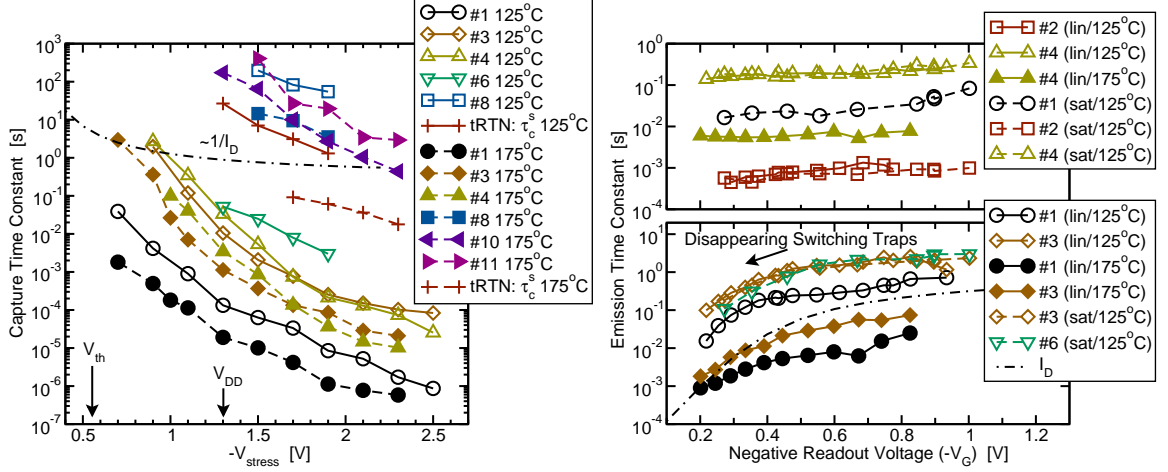
Figure 3: Left: Capture time constants extracted from TDDS experiments. One line corresponds to a single defect whose time constants are collected at one temperature but for various different gate voltages. The curve shapes clearly show that the trapping process are field-accelerated and temperature-activated. **Right**: Distinct behavior of two types of defects in NBTI. While defects in the upper panel exhibit no voltage dependence, defects from the lower panel strongly respond to different gate voltages during relaxation.

Time-Dependent Defect Spectroscopy

The shrinking of device geometries in modern technology nodes has advanced to the point where single charging or discharging events of defects can be recognized as discrete steps in the V_{th} recovery traces. Previously, the capture (τ_c) and emission (τ_e) time constants of these events has experimentally been assessed via random telegraph noise (RTN) experiments. Recently, a new technique, termed time-dependent defect spectroscopy (TDDS) (43), has been developed to extend the typical measurement window for τ_c and τ_e. In these experiments, the nature of the mechanism behind NBTI has been identified as a superposition of independent Poisson processes. As pointed out in (43), this finding is consistent with the concept of hole trapping but not with the RD model. The latter presumes indirect interactions between the individual interface states via hydrogen diffusion and therefore contradict the concept of independent processes.

An investigation on the voltage and temperature behavior of the trapping and detrapping events has been conducted and yielded intriguing results. Both the capture and emission time constants depicted in Fig. 3 show a temperature dependence with a wide spread of activation energies. The field acceleration of τ_c can be fitted to $\exp(c_1 F + c_2 F^2)$, which is in line with the quadratic behavior measured by eMSM. Regarding τ_e, one defect type is not influenced by the magnitude of V_G while the emission time constants for the other significantly decrease as V_G is reduced below V_{th}. No correlation between capture and emission time constants has been identified, which, however, does not preclude the existence of such a correlation.

Atomistic Perspective on Charge Trapping

Defects are the main actors in charge trapping and their properties strongly vary with their chemical nature. Most importantly, they introduce localized states into the oxide that are accessible for charge carriers from the bulk or the gate. A large number of defects has been discovered in various dielectric materials used in MOS structures. Their properties are influenced by the atomic structure of their surrounding. In some cases, the defect properties are strongly affected by the amorphous structure of the host material. Small variations in the bond lengths and angles of the surrounding can then result in a wide distribution of trap levels, which in turn impacts the trapping dynamics. Since these properties are hard to determine by measurements, theoretical methods, such as density function theory, can provide additional input to research in the field of oxide reliability.

Silicon Dioxide

Silicon dioxide has been successfully applied in MOS technologies for decades and still attracts scientific as well as industrial interest because even modern gate stacks have a thin interfacial layer of silicon dioxide. Although it can form several stable crystalline structures, such for instance as quartz and cristobalite, MOS technologies are based on thin layers of an amorphous phase as confirmed by transmission electron microscopy (44).

A series of E' center defects have been detected by ESR measurements (45) in high-quality silicon dioxide films of MOS transistors. The E'_γ center (see Fig. 4) and the oxygen vacancy (see Fig. 4) rank among the most prominent ones and together make up the core of the Harry-Diamond-Laboratories (HDL) model (46, 47) which has been used to explain irradiation response of MOS devices. Lenahan $et\ al.$ quite early established that the E' centers are the dominant defects for charge trapping in silicon dioxide, evidenced by their observed one-to-one correspondence in irradiated oxides (48, 49). The HDL model takes the oxygen vacancy as a starting point. This defect consists of a strained Si−Si in its neutral charge state (dimer configuration) and carries two electrons in the corresponding orbital. Upon hole capture, this bond is weakened, breaks up and leaves behind a Si atom with a singly occupied dangling bond. The other, positively charged Si atom relaxes into the middle of the three neighboring O atoms or, according to first principles calculations, weakly bonds to another adjacent O atom (puckered configuration). In this configuration, the E' center features an energy level in the middle of the silicon dioxide bandgap and can thus easily exchange charge carriers with the substrate via tunneling. Consequently, it is classified as a switching trap, whose occupation may quickly respond to gate bias sweeps. When it is neutral, the defect may return to its initial oxygen vacancy configuration by re-bonding of the two oxygen atoms. This is in line with a first principles investigation which predicts the E' center to be more stable in the positive charge state and the dimer configuration in the neutral charge state (50). Lu $et\ al.$ (51) found two groups of E'_γ centers — one immediately returns back to the dimer configuration when it is positively charged, while the other remains stable in its puckered configuration. Fleetwood $et\ al.$ linked these configurations to defects giving rise to $1/f$ noise (52).

It is frequently claimed that hydrogen undergoes various reactions with defects in silicon dioxide (53, 54). Conley $et\ al.$ proposed hydrogenated variants of the E'_γ cen-
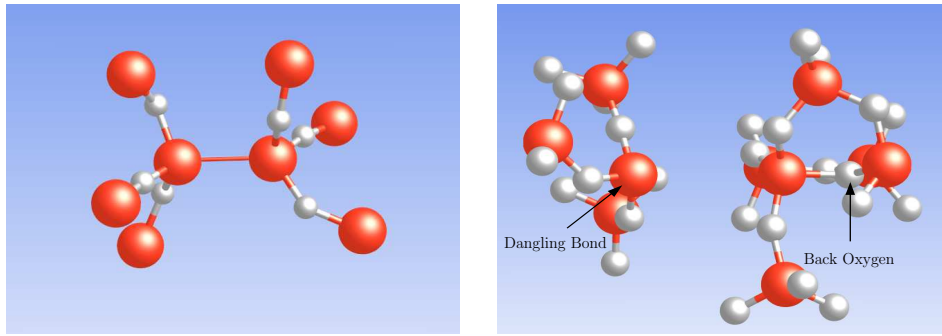
Figure 4: Left: Atomic structure of an oxygen vacancy. The large dark and the small light spheres are silicon and oxygen atoms, respectively. It features a strong Si−Si bond in its neutral charge state and is the precursor defect in the HDL model. **Right**: Representation of an E'_γ center. The silicon atom at the left-hand side carries the dangling bond which can take two electrons. The positively charged silicon atom at the right-hand side forms a bond to a nearby oxygen atom (back oxygen). This special configuration is also referred to as the puckered configuration.
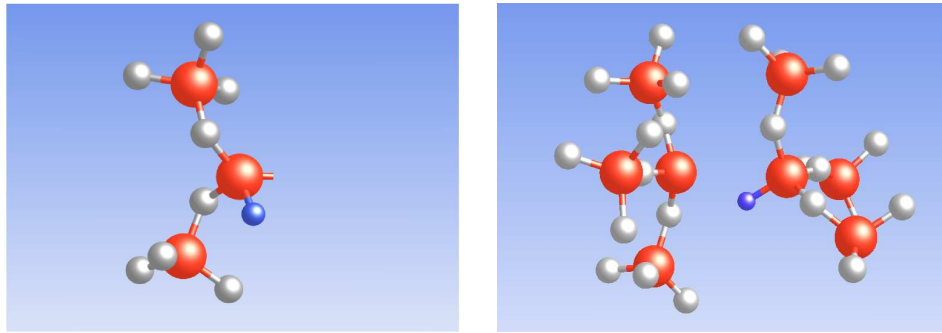

Figure 5: Left: Atomic structure of the 74 Gauss doublet center. The small dark atom represents a hydrogen atom. Just as the E'_γ center, it carries a dangling bond but is bonded to one hydrogen atom. **Right**: Representation of a hydrogen bridge. The hydrogen atom is situated inbetween two silicon atoms but shows a strong bond (indicated by the stick) and a weak bond (not shown here) to the neighboring silicon atoms.

ter (55–58), namely the 74G doublet (see Fig. 5) and the 10.4G doublet center, which both are formed after exposure to a hydrogen ambient. Another hydrogen-related defect in silicon dioxide is the so-called hydrogen bridge which has been suggested to increase the leakage current via charge trapping (59). Several recent investigations (60–62) examined the impact of the amorphous nature of silicon dioxide on the behavior of hydrogen-related defects. Atomic hydrogen is found to introduce several energy levels in silicon dioxide that might be taken into account for charge trapping.

Conventionally, elastic tunneling is regarded as a transition of an electron between a band state in the silicon and a "fixed" defect level in the dielectric. Following theoretical considerations, however, the defect configuration reacts to a change of the charge state, which might be accompanied by breaking or weakening of bonds. This structural relaxation occurs on timescales considerably smaller than that usually observed for tunneling.
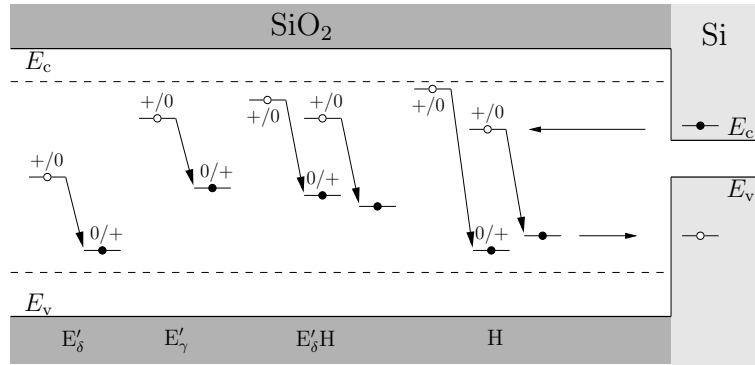
Figure 6: Band scheme including the trap levels calculated by density functional theory (60–62). When an unoccupied defect (open circles) captures an electron, its corresponding energy level vanishes and reappears at another position (filled circles). From there, it is capable of emitting an electron again. Upon electron emission, this energy level rises upwards to its initial position again. In amorphous materials and thus also in interfacial layers, the defect level may be subject to a large spread due to variation in bond angles and lengths.

As a consequence, there exist two defect levels with one for electron capture and one for electron emission (see Fig. 6). In (63, 64) it is shown that the combination of both levels can strongly impact the trapping dynamics and thus the behavior of the defect.

Silicon Oxynitrides

The beneficial role of nitrogen as a blocking barrier against impurity penetration has initially driven the interest in nitrogen incorporation into dielectrics. Besides this, nitrogen increases the effective oxide thickness, thus reduces the leakage current due to direct tunneling and in a further consequence prolongs the lifetime of silicon dioxide-based dielectrics. Despite its great success, only limited experimental insight into the microscopic structure of silicon oxynitride has been gained so far.

From the point of view of ESR measurements, the most likely candidate for charge trapping in silicon oxynitrides is the so-called K_n center (65, 66). It has has been identified as a silicon dangling bond backbonded to three nitrogen atoms and is found to be located close to the interface. Angular dependent ESR measurements indicate that the K_n centers do not have a preferred orientation and therefore must be ascribed to bulk rather than to interface defects.

Earlier investigations based on infrared absorption spectra (67) and ESR measurements (68) have also taken the N_4 complex and the nitrogen bridge (N_2) into account. In first principles calculations, the latter one is found to feature a bistability similar to that of the E' center in pure silicon dioxide (69, 70).

High-κ Dielectrics

As the miniaturization of semiconductor devices continues, silicon dioxide as well as silicon oxynitride are about to be replaced by high-κ dielectrics in order to reduce undesirable

leakage and dielectric breakdown. Among these new materials, hafnium dioxide appears to be the most promising candidate. Depending on its treatment, it either crystallizes with an additional nonstoichiometric amorphous silicon dioxide interlayer (71, 72) or forms a hafnium silicate, an amorphous mixture of hafnium and silicon dioxide. In addition to the defects known for silicon dioxide, the oxygen vacancy and the hafnium interstitial are found to introduce several electron levels within the hafnium dioxide bandgap (73–75), which may act as charge traps.

Reaction Rate Models

Since the discovery of the NBTI phenomenon, a number of microscopic interpretations based on diverse experimental and theoretical findings have been proposed. Their evaluation against experimental data requires the development of reaction based models.

Yang Model

In the early stages of research, only a few authors consider trapped charges as at least partially responsible for the NBTI phenomenon. During stress conditions, the charge carriers in the inversion layer have been envisioned to move into oxide defects via quantum mechanical tunneling. Each single charge carrier trapped therein then contributes to the shift in the threshold voltage. When the stress is removed, the charge carriers return back to the substrate so that the ΔV_{th} will decrease again. A phenomenological hole trapping model, based on broad distributions of trapping times, was put forward by Yang *et al.* (76). Its trapping dynamics are governed by a first-order rate equation:

$$\dot{f}_{\mathrm{T}}(t, \tau_{\mathrm{c}}, \tau_{\mathrm{e}}) = \frac{1}{\tau_{\mathrm{c}}}(1 - f_{\mathrm{T}}) - \frac{1}{\tau_{\mathrm{e}}} f_{\mathrm{T}} \qquad [5]$$

with f_{T} being the occupancy of a single trap. Each of these traps is characterized by its own capture (τ_{c}) and emission (τ_{e}) time constant. A dispersion of τ_{c} and τ_{e} has been assumed in order to explain the logarithmic time behavior of stress and recovery over several decades. In (76) the time constants are presumed to follow a log-normal distribution:

$$g(\tau) = \frac{1}{\sqrt{2\pi}\tau\sigma} \exp\left(-\frac{1}{2}\left(\frac{\log(\tau) - \mu}{\sigma}\right)^2\right) \qquad [6]$$

The time evolution of trapped charges $p(t)$ can be expressed as

$$p(t) = \int d\tau \, f_{\mathrm{T}}(t, \tau) \, g(\tau). \qquad [7]$$

Using this model, reasonable fits can be achieved for a single relaxation curve but too small slopes during the stress phase have been obtained (23). Nevertheless, the above formulation already suggests a suitable mathematical structure of charge capture in NBTI: Trapping and detrapping are reflected in the forward and the reverse rate of equation [5], where each of them depend on the trap occupancy and their inverse time constants τ_{c} and τ_{e}. Since the variations of the defect properties have a strong impact on the trapping process, the time constants are subject to a dispersion as realized in equation [6]. However, the model does not contain any information about the physics behind this process, in particular the capture and emission time constants are missing in this model. Therefore, the goal is to find the right distribution of τ_{c} and τ_{e} as well as their field and temperature dependences.
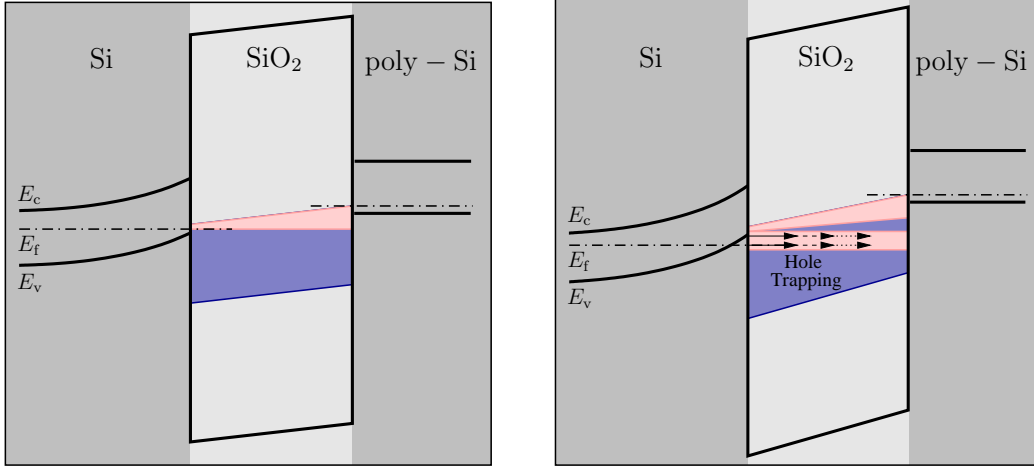
Figure 7: Schematics of the gate bias dependence of charge trapping. No charge trapping from the gate is considered for simplicity. **Left**: The band edge diagram including a broad band of traps at a low negative gate bias. Traps below the E_f (dark grey area within the dielectric) are occupied by electrons (neutral), while traps located above E_f (bright grey area) are empty (positive). **Right**: For a larger gate bias, a part of the traps in the dielectric is shifted above the Fermi level and will capture holes from the substrate as indicated by the arrows. The area available for hole capture and thus the possible amount of captured holes depends on the shift of the Fermi level relative to the energetical position of the traps. However, this correlation does not show the required quadratically field dependence.

Elastic Tunneling

Conventionally, charge trapping is envisioned as an elastic tunneling process (64, 77–80), where no energy is dissipated and the initial and the final energy of the tunneling charge carrier coincide. Such transitions are calculated based on Fermi's golden rule whose matrix element contains a time-independent Hamiltonian. This leads to the same expressions for the forward r_in and the reverse rate r_out of the tunneling transitions. According to the approximate derivation in (78), the most sensitive factor of the matrix element is the WKB factor, which shows an exponential dependence on the trap depth. The obtained rates enter a first order differential equation:

$$\dot{f}_\mathrm{T}(E_\mathrm{T}, x) = \underbrace{n(E_\mathrm{T})\ r_\mathrm{in}(E_\mathrm{T}, x)}_{1/\tau_\mathrm{c}(E_\mathrm{T}, x)}\ (1 - f_\mathrm{T}(E_\mathrm{T}, x)) - \underbrace{p(E_\mathrm{T})\ r_\mathrm{out}(E_\mathrm{T}, x)}_{1/\tau_\mathrm{e}(E_\mathrm{T}, x)}\ f_\mathrm{T}(E_\mathrm{T}, x)\ , \quad [8]$$

where $n(E_\mathrm{T})$ and $p(E_\mathrm{T})$ denote the density of occupied or empty bulk states with an energy of E_T at the substrate interface. The first term at the right hand side of equation [8] corresponds to electron trapping, while the second term represents hole trapping, respectively. Due to the strong exponential decay of the Fermi Dirac distribution, $n(E_\mathrm{T})$ strongly outnumbers $p(E_\mathrm{T})$ in the energy region below the Fermi level E_f. This favors electron injection into traps below E_f, which are therefore statistically occupied by electrons. Note that the opposite holds true for traps in the region above the the Fermi level. With the onset of stress, the shift of the Fermi level gives rise to a recharging of traps. The magnitude of this shift correlates with the amount of trapped charges as illustrated in Fig. 7. Since it does not depend quadratically on the oxide field, this model predicts the
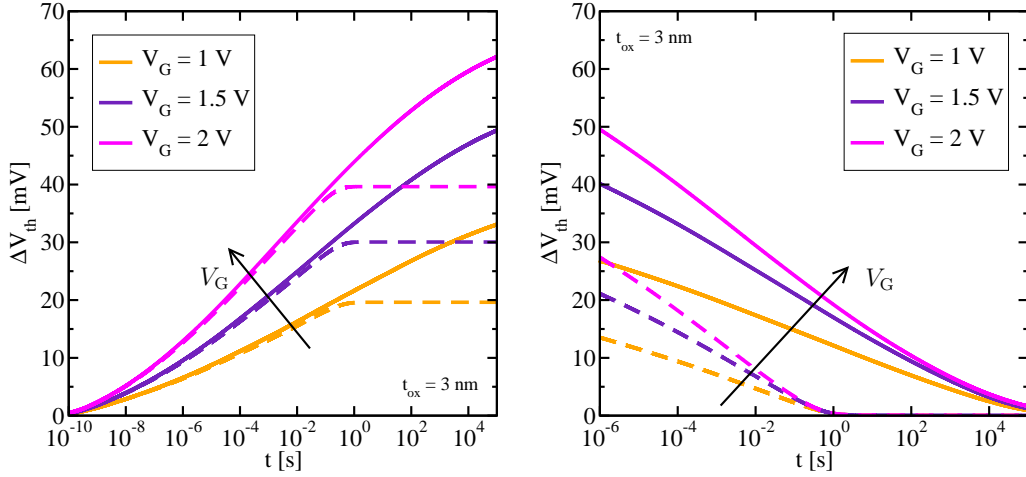
Figure 8: Left: Numerical simulations of the time evolution of V_{th} during the stress phase for different gate voltages. The solid lines denote V_{th} curves neglecting the gate contact, while the dashed lines belong to the case with consideration of the poly-gate interface. In the latter, only traps near the substrate interface (with short tunneling time constants) are capable of capturing holes. **Right**: The same as on the left-hand side but for the recovery phase. Due to the impact of the gate contact, only traps with small tunneling times participate in charge trapping during stress which is reflected in an early loss of trapped charges. In the case of the no charge injection from the gate, also traps with long tunneling times are involved in detrapping again and results in slow decay of V_{th}.

wrong field acceleration. Futhermore, the Fermi level remains nearly at the same position relative to the bandgap so that this model does not show any appreciable temperature dependence. The temporal filling of traps is thus dominated by the WKB coefficient and thus the tunneling time constants increase exponentially with the trap distance to the interface. Therefore, charge trapping starts from the interface and proceeds deep into the dielectric. For a spatially uniform trap distribution (see Fig. 8), this behavior yields a logarithmic time dependence consistent with experimental observations of the short time stress behavior (37). As pointed out in (63, 64), this model predicts the same time slopes during both phases (see Fig. 8) since the same defects with the same time constants are involved during stress and relaxation. This is in strong contrast to experimental observations, in which the recovery occurs much slower than the degradation during stress.
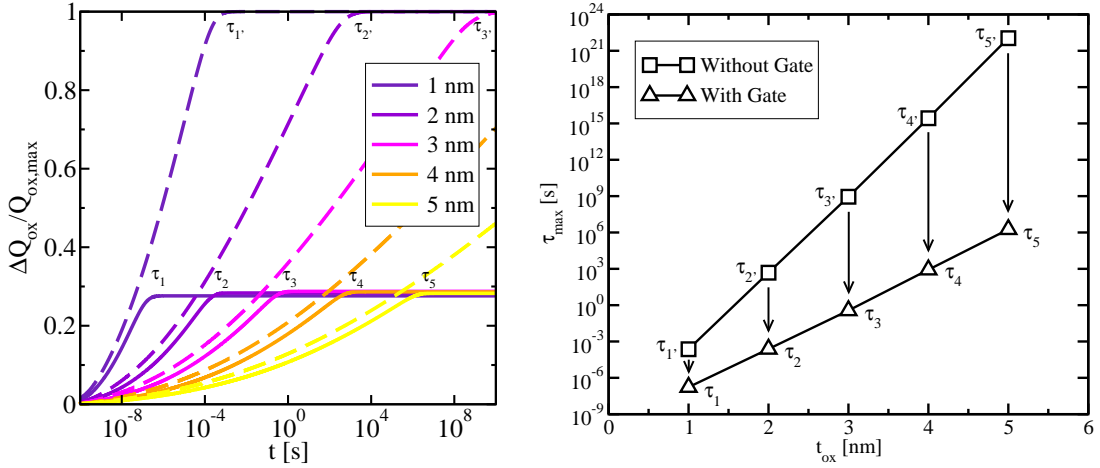
Figure 9: Left: Charge trapping for various gate thicknesses. The dashed lines represent the simple model without consideration of the gate contact, while the solid lines refer to the extended model including charge injection from the gate. This figure proves the importance of the gate contact when thin gate dielectrics are considered. **Right**: Dependence of the saturation of charge trapping on the oxide thickness. The upper and the lower line give the saturation of charge trapping for elastic tunneling with and without the consideration of the gate contact, respectively. For thicker gate dielectrics, the impact of the gate becomes relevant at later times.

Charge trapping is often restricted to charge injection from the substrate and the poly-gate interface is neglected as a source or a sink of charge carriers. This assumption is not justified when the thickness of the gate dielectric enters the nanometer range (80). The model is extended by introducing additional rates on the right-hand-side of equation [8].

$$\dot{f}_{\mathrm{T}}(E_{\mathrm{t}}, x) = \quad n_{\mathrm{s}}(E_{\mathrm{T}})\, r_{\mathrm{in,s}}(E_{\mathrm{T}}, x)\, (1 - f_{\mathrm{T}}(E_{\mathrm{T}}, x)) \quad - \quad p_{\mathrm{s}}(E_{\mathrm{T}})\, r_{\mathrm{out,s}}(E_{\mathrm{T}}, x)\, f_{\mathrm{T}}(E_{\mathrm{T}}, x)$$
$$+ \, n_{\mathrm{g}}(E_{\mathrm{T}})\, r_{\mathrm{in,g}}(E_{\mathrm{T}}, x)\, (1 - f_{\mathrm{T}}(E_{\mathrm{T}}, x)) \quad - \quad p_{\mathrm{g}}(E_{\mathrm{T}})\, r_{\mathrm{out,g}}(E_{\mathrm{T}}, x)\, f_{\mathrm{T}}(E_{\mathrm{T}}, x) \quad [9]$$

The subscripts s and g refer to substrate and gate quantities, respectively. Here, the occupancy of defects located closer to the poly interface than to the substrate interface is determined by the charge injection from the gate contact. These defects are not recharged during stress or relaxation and thus do not participate in charge trapping from the gate. As pointed out in (64, 80), this can be recognized as a saturation in the ΔV_{th} curves of Fig. 8. For thin gate dielectrics ($< 3\,\mathrm{nm}$), the time of saturation is moved to timescales smaller than typical for NBTI (see Fig. 9) so that elastic tunneling cannot explain the recoverable component of NBTI (81) also from this point of view.

Note that this model requires a very broad energetic distribution of traps, however, it is more natural to assume a narrow distribution. For the latter (64), one observes a quick recharging of traps which occurs at earlier times for higher gate biases (see Fig. 10).

Model for Thermally Activated Hole Trapping

Since elastic tunneling apparently cannot explain the experimentally observed temperature dependence, alternative charge trapping processes have been applied. In an alternative model, tunneling is only considered as one part of a complex trapping mecha-
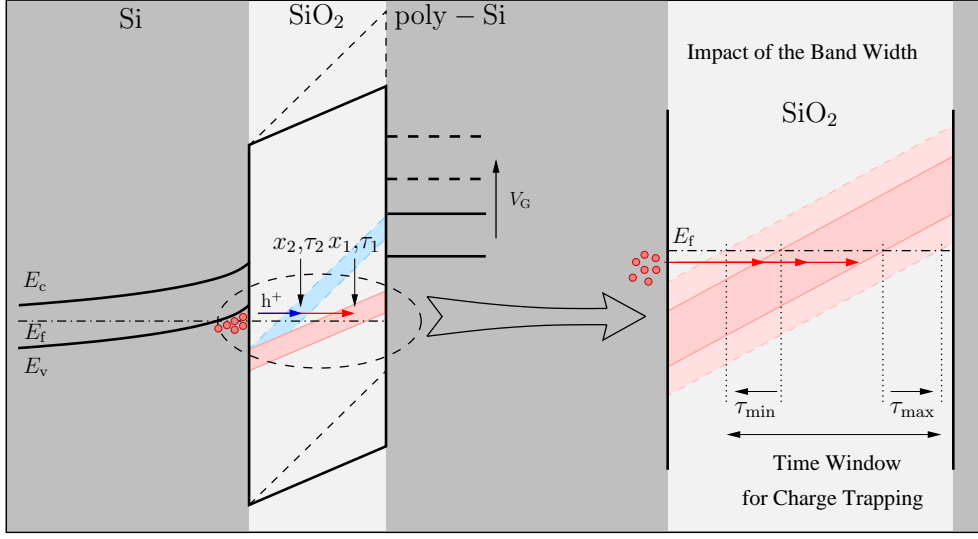
Figure 10: Bandstructure for a narrow distribution of trap levels. The narrow band of trap levels indicated by the grey areas has only a small intersection region with the Fermi level. Since charge trapping only occurs in there, the trapping time constants show a small distribution. Note that for increasing gate bias the intersection region shifts closer to the interface ($x_1 > x_2$) which goes hand in hand with an exponential decrease in the trapping time constants ($\tau_1 > \tau_2$) and an earlier onset of charge trapping. On the right hand side of this figure, the band of traps is zoomed out. The intersections with the band of trap levels determines the minimum ($\tau_{\min}$) and the maximum ($\tau_{\max}$) trapping time constants. For a broadened band of traps, $\tau_{\min}$ moves to shorter times while $\tau_{\max}$ shifted to larger times. This means that with increasing width of the trap band, the time windows for charge trapping is enlarged.

nism and is not necessarily the rate-limiting process. In an early modeling attempt based on such a mechanism, the creation of defects is initiated by thermally activated hole trapping, represented by the left double-well in Fig. 11. It is assumed that a hole located in the substrate tunnels into a defect in the dielectric and thereby initiates the creation of an "electrical" trap, such as a P_b and K_n center for instance. Potential candidates would be the hydrogen bridge or one of the various variants of the oxygen vacancy, both frequently mentioned in the context of oxide reliability issues. The above hole capture mechanism favors a defect transformation, which can only be reversed slowly and is therefore more permanent. This second step is described by a simplified formalism of transition state theory (TST) and is represented by the right double-well in Fig. 11. Hydrogen is suspected to interact with various sorts of defects in the dielectric and thus has been proposed as an actor in the second step. For instance, it could be released during a bond rupture, subsequently migrate away, and would not be available for an immediate inverse reaction.

The defect dynamics are expressed by the following rate equations:

$$
\begin{aligned}
\dot{f}_1 &= -\,\nu_1 \mathrm{e}^{-\beta(V_2-V_1)}\mathrm{e}^{\Delta_1} \times f_1 &&+\; \nu_2 \mathrm{e}^{-\beta(V_5-V_4-\Delta_2)} \times (1-f_1) \\
\dot{f}_2 &= -\,\mathrm{e}^{-\beta(V_2-V_3)} &&\times (1-f_1)f_2 \;+\; \nu_2 \mathrm{e}^{-\beta(V_5-V_6+\Delta_2)} \times (1-f_2)\,, &&\qquad [10]
\end{aligned}
$$

with β being $1/k_{\mathrm{B}}T$. f_1 and f_2 stands for the occupancy of the states 1 and 4, respectively. ν_1 and ν_2 is referred to the attempt frequencies of the first and second step of the model.
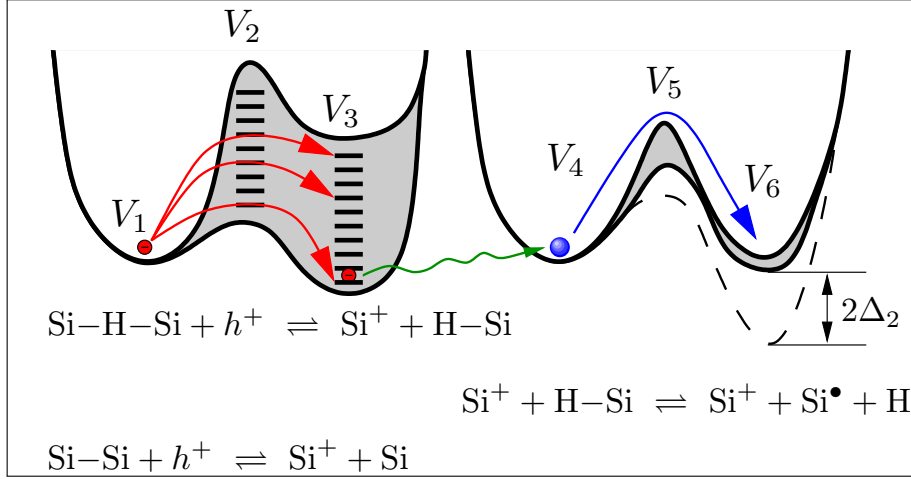
Figure 11: Schematic illustration of the thermally activated hole trapping model: The considered defect exists in four different states denoted by i, where each of them has an energy V_i. The left double-well represents the defect creation upon hole capture in a configuration coordinate diagram. Therein, the total energy of the defect is plotted as a function of the reaction coordinate which summarizes the atomic rearrangement caused by the defect creation. When the defect is in state 1, it has to overcome a thermal barrier of $V_2 - V_1$ to proceed to state 3. The energy barriers are assumed to be uniformly distributed to allow for a wide range of trapping times. The stress parameter Δ_1 used in equation [10] strongly enhances the above transition and in consequence accounts for the field acceleration of the thermally activated hole trapping process. In a second stage (right well), the defect may undergo a bond breakage reaction, described analogously to the interface reaction in the triple-well model (32). The defect in state 4 encounters a barrier of $V_5 - V_4$ whose height follows a Gaussian distribution. The parameter Δ_2 reflects the lowering of the energies V_5 and V_6 due to the electric field.

Δ_1 denotes the stress parameter, which corresponds to an enhancement factor for defect creation. Δ_2 accounts for the field-dependence of the second transition over the barrier in Fig. 11.

As can be seen in Fig. 12, this model can perfectly reproduce complicated stress and relaxation patterns as well as the quadratic field- and temperature dependence. However, the stress parameter is phenomenologically introduced and needs to be physically motivated as done in the two stage model discussed in the next chapter.

Two Stage Model

For elastic tunneling, it has been postulated that the initial electronic energy in the substrate and the final electronic defect energy coincide. Recall that this concept shows only a weak temperature dependence, which is not compatible with the experimental findings. However, only the overall energy of the system, including the lattice and the electronic energy, must be conserved according to Fermi's Golden rule. Thus, the difference in electronic levels can be compensated by exchanging energy with the system of lattice vibrations. Such a mechanism has already been proposed for random telegraph noise (RTN) and $1/f$ noise (82, 83) in microelectronic devices and dates back to the pioneering work of Dutta *et al.* (84, 85). In there, a simple model based on multi-
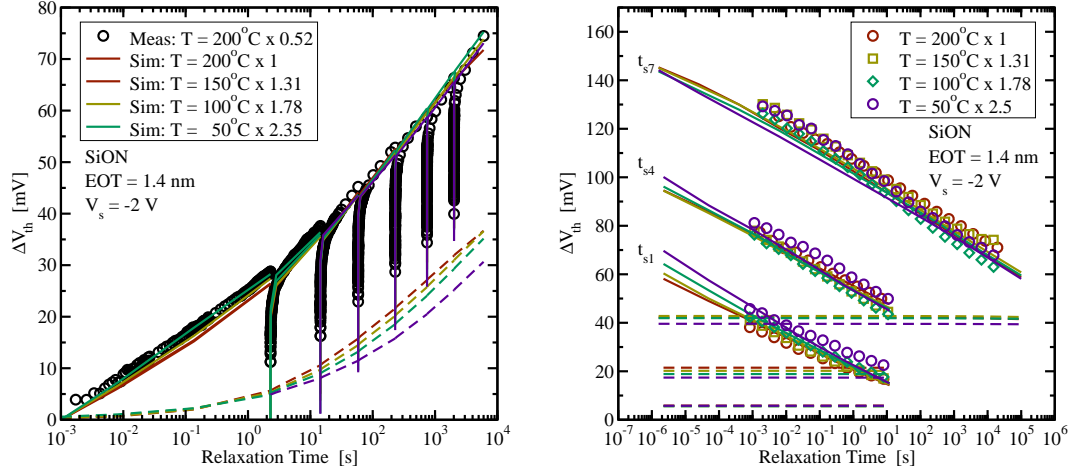
Figure 12: Left: Verification of scalability for the thermally hole trapping model. The simulated stress/relax cycles can be overlapped for different temperatures by using the scaling factors given in the inset. Moreover, the resulting curve shape matches perfectly with the measurement data obtained by the eMSM technique at 200 °C. **Right**: Scaled relaxation data of the stress interruptions visible in the curves on the left-hand side. Recorded V_{th} for different stress temperatures also perfectly line up for the relaxation phase.

phonon emission (MPE), has been employed. However, the concept of this mechanism has also been encountered in the context of phonon-assisted tunneling ionization of deep centers (86–90).

Due to the importance of the MPE process, an introduction to the basic concept will be given in the following. The potential energy curves $U_1(q)$ and $U_2(q)$ (see Fig. 13) give the total energy of a defect as a function of its configuration coordinate q. Depending on whether the charge resides in the channel or the defect, the entire system is described by the curve $U_1(q)$ or $U_2(q)$, respectively. The parabolic shape of these curves arises from interactions between defects and lattice phonons, frequently described by harmonic oscillators. When the system is thermally raised from the initial configuration $U_1(q_1)$ to the intersection point of the two parabolas IP, a charge transfer reaction can take place. Exactly at this point, the potential energy curves $U_1(q)$ and $U_2(q)$ are equal and allow a pure tunneling process. From there, the system relaxes to the equilibrium configuration q_2 with the energy $U_2(q_2)$. In (88), it was found that tunneling is considerably enhanced in bulk defects upon application of an electric field due to the lowering of the tunneling barrier. Theoretical calculations deliver a factor of $\exp(F^2/F_c^2)$ (88), which accounts for this effect and is able to reproduce the required field dependence. However, this process has been derived for bulk defects, where the thermal excitation of trapped charges is given by the Fermi level within the bulk. We remark that in the simple MPE theory used in the two stage model (TSM), the barrier heights are subject to a dispersion, which is explained by the variations in the local bonding structure in an amorphous material. This yields a wide distribution of trapping times analog to Yang's model. The TSM is therefore also consistent with a logarithmic time behavior during stress and relaxation over a wide time range.
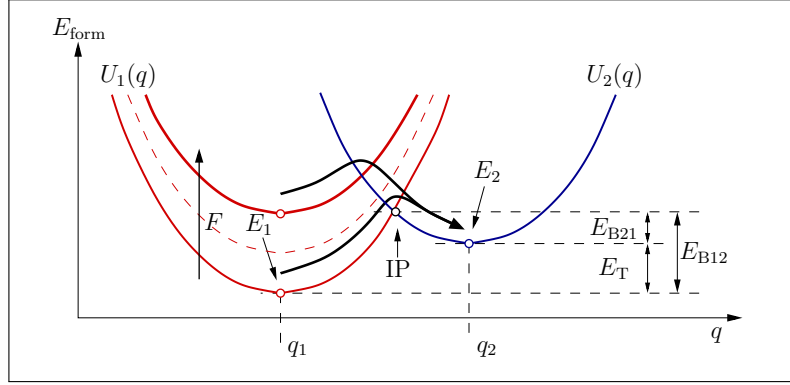
Figure 13: Schematic illustration of the MPE process: Two total energy curves are drawn for the case that the charge carrier exchanged between the substrate and the dielectric is either still located in the substrate or already caught in the defect. Although, in principle, the curves can take an arbitrary shape, they are assumed to be parabolic within the harmonic approximation. E_1 at q_1 and E_2 at q_2 denote the minima of the curves, while the connecting barriers are referred to as E_{B12} and E_{B21}. The energy difference E_T corresponds to the position of the trap level of the defect. However, in more sophisticated models they are calculated as the intersection point of the two parabolas according to non-radiative multiphonon (NMP) theory. Note that the shift of the red curve due to a higher gate voltage results in a lower barrier leading to a higher transition rate for charge transfer reaction.

The MPE process allows to explain trapping processes where the defect level is situated far below the substrate valence band as supposed in (45). Hence, it can be linked to the HDL model (see Fig. 14), suggested in several ESR studies. Therein, the defect precursor, the oxygen vacancy, features a wide distribution of donor levels below the substrate valence band and is capable of trapping substrate holes via the MPFAT process. After the defect transformation to an E' center, visible in ESR measurements (91), the defect has an energy level close or within the substrate bandgap in accordance with (92). In this state, it can be repeatedly charged and discharged by adding or removing one electron from its dangling bond. This switching behavior is in agreement with the experimental observations made in (39). The E' center in its neutral charge state can be fully annealed, becoming an oxygen vacancy again. However, amphoteric traps, unquestionably detected as P_b centers in ESR measurements (91), are involved in the second stage via a complex reaction. Therein, hydrogen originating from the interface is assumed to leave behind the P_b centers and saturates the dangling bond of the E' center. This stage of the entire NBTI process fixes the positive charge at the defect and therefore corresponds to the permanent or slowly recoverable NBTI component.

Mathematically, the dynamics of this complex mechanism are described by a set of rate equations:

$$\begin{aligned}
\dot{f}_1 &= -f_1 k_{12} + f_3 k_{31}, \\
\dot{f}_2 &= +f_1 k_{12} - f_2 k_{23} + f_3 k_{32} - f_2 k_{24} + f_4 k_{42}, \\
\dot{f}_3 &= +f_2 k_{23} - f_3 k_{32} - f_3 k_{31}, \\
\dot{f}_4 &= +f_2 k_{24} - f_4 k_{42},
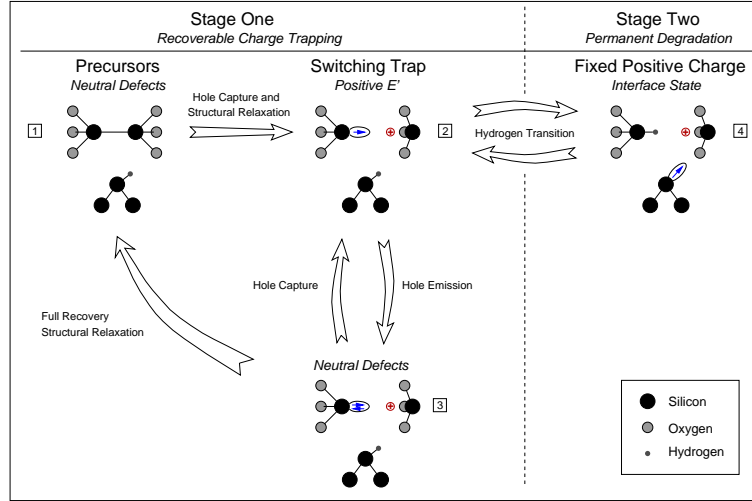\end{aligned} \qquad [11]$$

Figure 14: Trapping dynamics of the two stage model (TSM). On the left hand side, the recoverable component of NBTI, explained by the HDL model, is dealt with. The precursor (state 1) is transformed to a switching trap (state 2) via a MPFAT process which cannot be reversed. The fast response to bias sweeps proceed through hole capture and emission, where the defect changes between the states 2 and 3. From the neutral charge state, the defect can undergo structural relaxation after overcoming a thermal barrier and arrive at its initial configuration. In the second stage, the permanent component of NBTI, attributed to fixed oxide charges and interface states, is addressed. Here, a hydrogen transition from state 2 to state 4 fixes the trapped charge in the defect and creates new interface states.

where the subscripts denote the states as numbered in Fig. 14 and the k_{ij} stand for the rates from state i to j. The electron and hole capture rates $(1 \rightarrow 2$ and $2 \leftrightarrow 3)$ have been derived from the Shockley-Read-Hall-like equations, where the MPE process is incorporated in an empirical manner. The relaxation of the defect $(3 \rightarrow 1)$ is interpreted by TST in which the rate depends exponentially on the height of the corresponding thermal barriers. The hydrogen transition $(2 \leftrightarrow 4)$ has been modeled assuming a thermal, field-dependent barrier for simplicity.

In (39), the TSM has been validated for a large set of measurement data, including various combinations of stress voltages and temperatures. For illustration, a fit to eMSM data is shown in Fig. 15. The agreement with experimental data implies that this model reproduces the field and temperature dependence as required by criteria established at the beginning of this paper. Furthermore, the comparison of the averaged occupancies during stress and relaxation in Fig. 15 also explains the observed asymmetry of the prefactors A_s and A_r. Another feature of the TSM is the so-called occupancy effect: Some defects termed switching traps are quite sensitive to small changes in the gate bias, which can cause quick responses in their occupancy as illustrated in Fig. 16. Since only neutralized defects can be annealed, the recovery is strongly affected by the magnitude of the relaxation voltage as shown in Fig. 16. In conclusion, the TSM meets the established criteria and can also explain the impact of the relaxation voltage on the recovery. Although this model performs well in the above benchmark, the rates of the charge capture process were not rigorously derived from a concise microscopic picture. The required large time
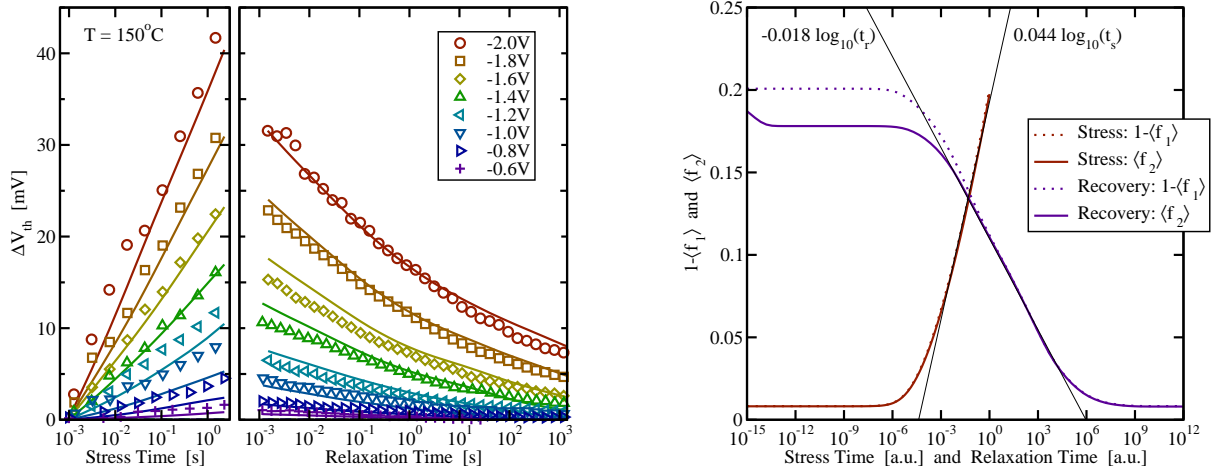
Figure 15: Left: Evaluation of measurement data of a thin SiON device against simulation data for 8 different stress voltages at a temperature of 150°C. The field acceleration and the asymmetry between stress and relaxation have been nicely reproduced. **Right**: Exemplaric occupancy for stress and recovery averaged over all simulated traps. Neglecting the defects in state 4, the dotted lines gives the number of oxide defects (state 2 and 3), while the solid line corresponds to the positively charged E' centers (state 3) only. The ratio between the slope of the stress (A_s) and the relaxation (A_r) curve yields $A_s/A_r \approx 2.5$, as observed experimentally.

constants for trapping and detrapping are achieved by strongly dispersed barrier heights in the MPE process. They are assumed to follow a random distribution instead of being calculated due to variations in the shape of the potential energy curves. Furthermore, the field dependence has been empirically introduced by the factor $\exp(F^2/F_c^2)$ which originates from a derivation of multiphonon-assisted tunneling ionization of bulk traps in the presence of an electric field (88). This description is based on traps located within bulk material, where the tunneling charge carriers face a triangular energy barrier. For increasing electric fields the tunneling length is decreased leading to the aforementioned factor. However, charge carriers from the substrate encounter a trapezoidal energy barrier, where the field dependence is much weaker.

Nonradiative Multi-Phonon Model

The recently developed TDDS is capable of directly measuring the field and temperature dependence of the charge capture process for single defects and thus provides the necessary information to model the charge capture process more accurately. Furthermore, it has revealed that some traps can show a switching trap behavior. To be able to deal with the stochastic nature of the charge capture process, a Master equation has been set up with the rates given by a homogeneous continuous-time jump Markov process (43). For larger devices the stochastic behavior averages out and the process can be described by simple rate equations. This stochastic approach has been employed in a recently published model where the charge transfer reaction is based on a nonradiative multiphonon (NMP) process (43). Therein, the shapes of the potential energy curves in Fig. 13 are assumed to be parabolic according to the harmonic approximation. As a result, the transition probability for a charge transfer reaction is calculated by the intersection point of
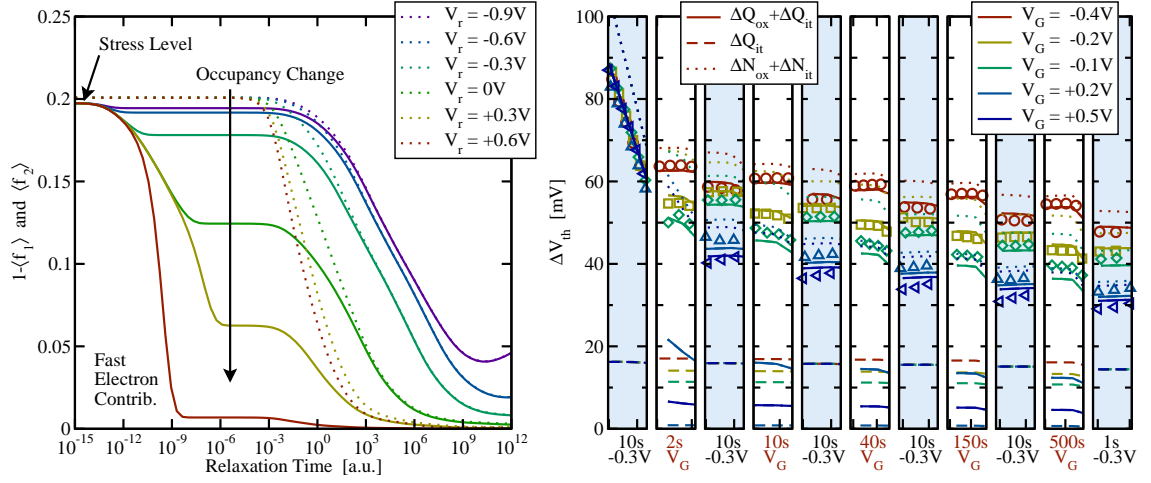
Figure 16: Left: The bias-dependent occupancy of the E' center. A high positive gate bias during recovery neutralizes the defect in its switching trap configuration (state 3) and thus increases the recovery of defects over the rate $(3 \rightarrow 1)$. **Right**: Gate bias dependence of recovery. Five devices were stressed at $V_s = -2$ V and thus subjected to the same level of degradation. Their V_{th} degradation was monitored during relaxation which was interrupted for some seconds at a different V_r. The splitting of the V_{th} curves after the first interruption gives the accelerated recovery for higher positive gate voltages.
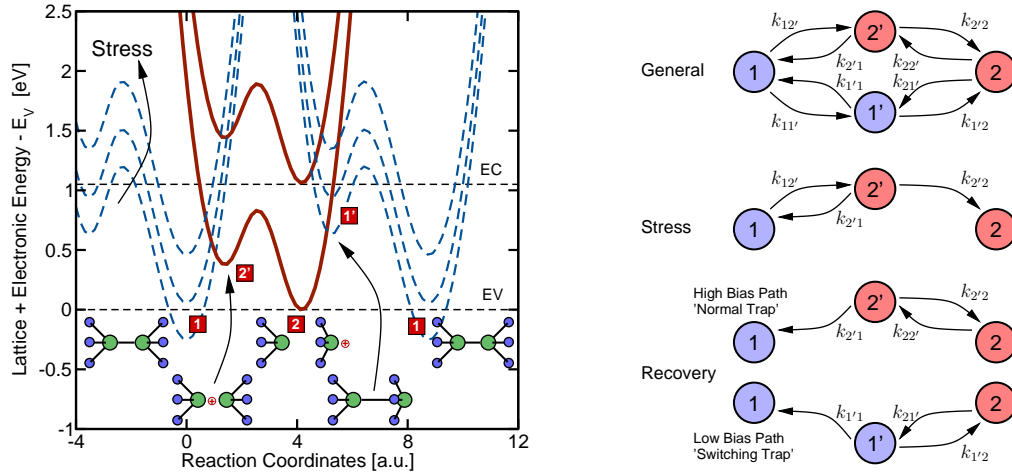

Figure 17: Left: Configuration coordinate diagram for the NMP model. The solid curve corresponds to the case, where the hole is trapped by the defect. Note that the energetical distance of state 2 and 1' gives the position of the effective trap level for hole capture. The neutral defect is represented by the dashed curve, which is shifted vertically with a change in the gate bias. This affects the intersection point between the solid and the dashed curve and in consequence the transition rate for hole capture. **Right**: State diagram for the general case (top), stress case (middle), and two alternative relaxation pathways (bottom).

the parabola for two different charge states of the defect. When a bias is applied to the gate, the relative position of the potential energy curves is changed. This affects the barrier height and as a further consequence influences the probability for a charge transfer
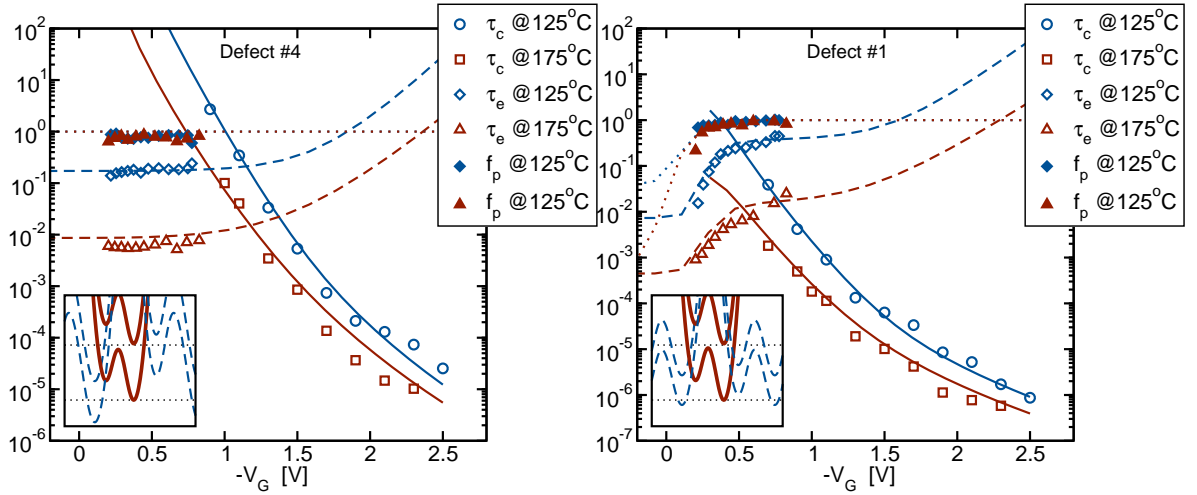
Figure 18: Left: Comparison of simulated and measured capture and emission time constants at different gate voltages in the case of normal kinetics. In the left bottom insert the corresponding configuration coordinate diagram is depicted. **Right**: The same as left but for switching trap kinetics.

reaction. In this way the required field dependence of NBTI is incorporated in the basic concept of this model. Therein the defect consists of a stable and a metastable state for the positively charged defect (state $2'$ in Fig. 17). This is, for instance, in line with first principles calculations of the E' center (50) where the Si−Si bond is weakened but not ruptured upon hole capture. Therefore, this metastable configuration is maintained for a short time before the defect surmounts a thermal barrier and relaxes to the new equilibrium configuration (state 2). The corresponding configuration coordinate diagram is composed of two double-wells, where the dashed and solid curve in Fig. 17 belong to the neutral and the positive charge state of the defect, respectively. Each of these double-wells has a metastable state and a equilibrium configuration with a thermal barrier inbetween. The charge transfer reactions, however, take place at the intersection points between the two double-wells in Fig. 17 and are described by the NMP theory as mentioned before. As illustrated in Fig. 17, one possible candidate for such a defect can be the E' center along with its neutral counterpart, the neutral oxygen vacancy. The latter features two different equilibrium configurations deviating in their Si−Si bond lengths for the positive and the neutral charge state. First principles calculations (50) predict a considerable fraction of these E' centers to remain stable at both charge states, which makes this defect a potential condidate for the NMP model.

TDDS experiments have revealed that the capture time constants of some traps are strongly affected by the gate bias while others are not (see Fig. 3). Their distinct behavior can be traced back to the recovery kinetics explained by the state diagram of Fig. 18. The pathway during stress proceeds from state 1 over state $2'$ to state 2 (see Fig. 18). Here, the field dependence is governed by the first transition $1 \leftrightarrow 2'$ and is reflected in the capture time constants shown in Fig. 18. The NMP theory used for the transition probabilities of the charge transfer reactions yields a nearly perfect agreement with the experimental bias-dependent capture time constants for different temperatures. This directly shows that the charge transfer reaction describes the experimentally observed field acceleration

along with the temperature activation. For relaxation, the relative positions of the minima 2 and $1'$ are crucial for the chosen reaction pathway, which determines whether the defect acts as a switching trap or a normal trap. When the minima 2 and $1'$ are energetical far separated, the rate-limiting step is given by a purely thermal barrier $2 \leftrightarrow 2'$ whose transition rate is voltage independent. This is in line with the measured emission time constants of defect #4 in Fig. 18). Switching traps, by contrast, feature a strong sensitivity on small variations in the gate bias as shown for defect #1 in Fig. 18. In this case, the defects chose the reaction path $2 \leftrightarrow 1' \rightarrow 1$, where the minima 2 and $1'$ are only separated by small barriers. Then the corresponding capture and emission time constants are short so that the occupancy of the traps can quickly respond to small shifts of the Fermi level. In conclusion, the NMP model can accurately describe single charge trapping events, however, the exact agreement between simulations and experimental data of the full ΔV_{th} curve is still pending. Nevertheless, this model appears to be the most promising one since its field dependence is inherent to the approach of NMP theory and does not require the introduction of a field dependent tunneling process as in the case of the TSM.

Conclusions

In this paper, the findings of experimental NBTI studies have been presented and used as criterions for the discussed charge trapping models. Most notably, the concept of elastic tunneling has been shown to have the wrong temperature and field dependences and cannot provide the required range of timescales. Therefore, this process must be ruled out as a possible NBTI model. More complex models, such as the TSM and NMP model, are already capable of reproducing the shape of stress/relaxation curves and the experimentally observed field- and temperature dependences. At the moment, the NMP model appears to be the most promising one since it is based on the experimental findings of TDDS and has a sound theoretical basis.

Acknowledgment

References

1. A. Goetzberger, and H. Nigh, *IEEE Proc.* **54**, 1454–1454 (1966).
2. B. Deal, M. Sklar, A. S. Grove, and E. H. Snow, *J.Electrochem.Soc.* **114**, 266 (1967).
3. D. Schroder, and J. Babcock, *J.Appl.Phys.* **94**, 1–18 (2003).
4. D. Schroder, *Microelectron.Reliab.* **47**, 841–852 (2007).
5. B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, *Proc.IRPS*, 2005, pp. 381–387.
6. K. Jeppson, and C. Svensson, *J.Appl.Phys.* **48**, 2004–2014 (1977).
7. M. Alam, *Proc.IEDM*, 2003, pp. 345–348.
8. M. Alam, *Proc.IRPS*, 2005, (Tutorial).
9. M. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, *Microelectron.Reliab.* **47**, 853–862 (2007).
10. J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, *Appl.Phys.Lett.* **87**, 204106 (2005).

11. P. Bunson, M. D. Ventra, S. Pantelides, R. Schrimpf, and K. Galloway, *IEEE Trans.Nucl.Sci.* **46**, 1568–1573 (1999).
12. J. Godet, and A. Pasquarello, *Microelectron.Eng.* **80**, 288–291 (2005), ISSN 0167-9317.
13. L. Tsetseris, X. J. Zhou, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, *Appl.Phys.Lett.* **86**, 142103 (2005).
14. T. Grasser, W. Goes, V. Sverdlov, and B. Kaczer, *Proc.IRPS*, 2007, pp. 268–280.
15. T. Grasser, P.-J. Wagner, P. Hehenberger, W. Goes, and B. Kaczer, *IEEE Trans.Dev.Mater.Rel.* **8**, 526 – 535 (2008).
16. H. Reisinger, U. Brunner, W. Heinrigs, W. Gustin, and C. Schlünder, *IEEE Trans.Dev.Mater.Rel.* **7**, 531–539 (2007).
17. A. Islam, E. N. Kumar, H. Das, S. Purawat, V. Maheta, H. Aono, E. Murakami, S. Mahapatra, and M. Alam, *Proc.IEDM*, 2007, pp. 1–4.
18. C. Shen, M.-F. Li, X. Wang, Y.-C. Yeo, and D.-L. Kwong, *IEEE Elect.Dev.Let.* **27**, 55–57 (2006).
19. C. Shen, M.-F. Li, C. E. Foo, T. Yang, D. Huang, A. Yap, G. Samudra, and Y.-C. Yeo, *Proc.IEDM*, 2006, pp. 333–336.
20. T. Grasser, B. Kaczer, P. Hehenberger, W. Goes, R. O'Connor, H. Reisinger, W. Gustin, and C. Schlünder, *Proc.IEDM*, 2007, pp. 801–804.
21. M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, C. Guerin, G. Ribes, F. Perrier, M. Mairy, and D. Roy, *Proc.IRPS*, 2006, pp. 735–736.
22. T. Grasser, and B. Kaczer, *Proc.ESSDERC*, 2007, pp. 127–130.
23. 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.
24. T. Grasser, W. Goes, and B. Kaczer, *ECS Trans.* **19**, 265.
25. B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, *Appl.Phys.Lett.* **86**, 1–3 (2005).
26. T. Grasser, W. Goes, and B. Kaczer, *Proc.IIRW*, 2006, pp. 5–10.
27. T. Grasser, W. Goes, and B. Kaczer, *IEEE Trans.Dev.Mater.Rel.* **8**, 79–97 (2008).
28. A. Stesmans, *Appl.Phys.Lett.* **68**, 2076–2078 (1996).
29. A. Stesmans, *Phys.Rev.B* **61**, 8393–8403 (2000).
30. V. Huard, M. Denais, F. Perrier, N. Revil, C. Parthasarathy, A. Bravaix, and E. Vincent, *Microelectron.Reliab.* **45**, 83–98 (2005).
31. V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammase, D. Barge, and C. Ouvrard, *Proc.IEDM*, 2007, pp. 797–800.
32. T. Grasser, W. Goes, and B. Kaczer, *Proc.SISPAD*, 2008, pp. 65–68.
33. H. Reisinger, O. Blank, W. Heinrigs, A. Mühlhoff, W. Gustin, and C. Schlünder, *Proc.IRPS*, 2006, pp. 448–453.
34. B. Kaczer, T. Grasser, P. Roussel, J. Martin-Martinez, R. O'Connor, B. O'Sullivan, and G. Groeseneken, *Proc.IRPS*, 2008, pp. 20–27.
35. T. Grasser, B. Kaczer, and W. Goes, *Proc.IRPS*, 2008, pp. 28–38.
36. T. Grasser, B. Kaczer, T. Aichinger, W. Goes, and M. Nelhiebel, *Proc.IIRW*, 2008, pp. 91–95.
37. P. Hehenberger, P.-J. Wagner, H. Reisinger, and T. Grasser, *Proc.ESSDERC*, 2009, pp. 311–314.

38. V. Huard, *Proc.IRPS*, 2010, pp. 33 –42.
39. T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, *Proc.IRPS*, 2009, pp. 33–44.
40. V. Huard, M. Denais, and C. Parthasarathy, *Microelectron.Reliab.* **46**, 1–23 (2006).
41. S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, and M. A. Alam, *Proc.IRPS*, 2007, pp. 1–9.
42. A. Haggag, W. McMahon, K. Hess, K. Cheng, J. Lee, and J. Lyding, *Proc.IRPS*, 2001, pp. 271–279.
43. T. Grasser, H. Reisinger, P.-J. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, *Proc.IRPS*, 2010, pp. 16 –25.
44. A. Diebold, D. Venables, Y. Chabal, D. Muller, M. Weldon, and E. Garfunkel, *Mat.Sci. in Semicond.Process.* **2**, 103–147 (1999).
45. E. Poindexter, and W. Warren, *J.Electrochem.Soc.* **142**, 2508–2516 (1995).
46. A. Lelis, T. Oldham, H. Boesch, and F. McLean, *IEEE Trans.Nucl.Sci.* **36**, 1808–1815 (1989).
47. A. Lelis, and T. Oldham, *IEEE Trans.Nucl.Sci.* **41**, 1835–1843 (1994).
48. P. Lenahan, and P. Dressendorfer, *IEEE Trans.Nucl.Sci.* **30**, 4602 –4604 (1983).
49. P. Lenahan, and P. Dressendorfer, *J.Appl.Phys.* **55**, 3495–3499 (1984).
50. M. Boero, A. Pasquarello, J. Sarnthein, and R. Car, *Phys.Rev.Lett.* **78**, 887–890 (1997).
51. Z.-Y. Lu, C. Nicklaw, D. Fleetwood, R. Schrimpf, and S. Pantelides, *Phys.Rev.Lett.* **89**, 285505 (2002).
52. D. Fleetwood, H. Xiong, Z.-Y. Lu, C. Nicklaw, J. Felix, R. Schrimpf, and S. Pantelides, *IEEE Trans.Nucl.Sci.* **49**, 2674–2683 (2002).
53. P. Bunson, M. D. Ventra, S. Pantelides, D. Fleetwood, and R. Schrimpf, *IEEE Trans.Nucl.Sci.* **47**, 2289–2296 (2000).
54. R. V. Ginhoven, H. Hjalmarson, A. Edwards, and B. Tuttle, *Nucl.Instr.&Meth.Phys.Res.Sect.B* **250**, 274–278 (2006).
55. V. Afanas'ev, and A. Stesmans, *EuroPhys.Lett.* **53**, 233–239 (2001).
56. J. Conley, and P. Lenahan, *IEEE Trans.Nucl.Sci.* **39**, 2186–2191 (1992).
57. J. Conley, and P. Lenahan, *Appl.Phys.Lett.* **62**, 40–42 (1993).
58. P. Lenahan, and J. Conley, *J.Vac.Sci.Technol.B* **16**, 2134–2153 (1998).
59. P. Blöchl, and J. Stathis, *Phys.Rev.Lett.* **83**, 372–375 (1999).
60. A. Alkauskas, and A. Pasquarello, *Phys.B Condens.Matter* **401** (2007).
61. W. Goes, and T. Grasser, *Proc.SISPAD*, 2007, pp. 157–160.
62. W. Goes, and T. Grasser, *Proc.IIRW*, 2007, pp. 27–32.
63. W. Goes, M. Karner, V. Sverdlov, and T. Grasser, *IEEE Trans.Dev.Mater.Rel.* **8**, 491–500 (2008).
64. W. Goes, M. Karner, V. Sverdlov, and T. Grasser, *Proc.IPFA*, 2008, pp. 249–254.
65. J. Campbell, and P. Lenahan, *Proc.IRPS*, 2006, pp. 442–447.
66. J. Campbell, and P. Lenahan, *Proc.IRPS*, 2007, pp. 503–510.
67. H. Ono, T. Ikarashi, Y. Miura, E. Hasegawa, K. Ando, and T. Kitano, *Appl.Phys.Lett.* **74**, 203–205 (1999).
68. J. H. Stathis, J. Chapple-Sokol, E. Tierney, and J. Batey, *Appl.Phys.Lett.* **56**, 2111–2113 (1990).
69. S. Joeng, and A. Oshiyama, *Phys.Rev.B* **86**, 3574 (2001).
70. E.-C. Lee, and K. Chang, *Phys.Rev.B* **66**, 233205 (2002).

71. H. Baik, M.K., G.-S. Park, S. Song, M.Varela, A. Franceschetti, S. Pantelides, and S. Pennycook, *Appl.Phys.Lett.* **85**, 672–674 (2004).

72. K. van Benthem, A. Lupini, M. Kim, H. Baik, S. Doh, J.-H. Lee, M. Oxley, S. Findlay, L. Allen, J. Luck, and S. Pennycook, *Appl.Phys.Lett.* **87**, 034104 (2005).

73. P. Broqvist, and A. Pasquarello, *Appl.Phys.Lett.* **89**, 262904 (2006).

74. P. Broqvist, A. Alkauskas, and A. Pasquarello, *Appl.Phys.Lett.* **92**, 132911 (2008).

75. J. Gavartin, D. Mũnoz Ramo, A. Shluger, G. Bersuker, and B. Lee, *Appl.Phys.Lett.* **89**, 082908 (2006).

76. T. Yang, C. Shen, M.-F. Li, C. Ang, C. Zhu, Y.-C. Yeo, G. Samudra, S. Rustagi, M. Yu, and D.-L. Kwong, *IEEE Elect.Dev.Let.* **26**, 826–828 (2005).

77. A. McWhorter, *Sem.Surf.Phys* pp. 207–228 (1957).

78. T. Tewksbury, , *Relaxation Effects in MOS Devices due to Tunnel Exchange with Near-Interface Oxide Traps*, Ph.D. Thesis, MIT (1992).

79. V. Huard, C. Parthasarathy, C. Guerin, and M. Denais, *Proc.IRPS*, 2006, pp. 733–734.

80. W. Goes, M. Karner, S. Tyaginov, P. Hehenberger, and T. Grasser, *Proc.SISPAD*, 2008, pp. 69–72.

81. J. Campbell, J. Qin, K. Cheung, L. Yu, J. Suehle, A. Oates, and K. Sheng, *Proc.IRPS*, 2009, pp. 382–388.

82. M. Kirton, and M. Uren, *Adv.Phys.* **38**, 367–486 (1989).

83. K. Ralls, W. Skocpol, L. Jackel, R. Howard, L. Fetter, R. Epworth, and D. Tennant, *Phys.Rev.Lett.* **52**, 228–231 (1984).

84. P. Dutta, P. Dimon, and P. Horn, *Phys.Rev.Lett.* **43**, 646–649 (1979).

85. M. Weissman, *Rev.Mod.Phys.* **60**, 537–571 (1988).

86. S. Makram-Ebeid, and M. Lannoo, *Phys.Rev.B* **25**, 6406–6424 (1982).

87. C. Henry, and D. Lang, *Phys.Rev.B* **15**, 989–1016 (1977).

88. S. Ganichev, W. Prettl, and I. Yassievich, *Phys.Solid State* **39**, 1703–1726 (1997).

89. A. Palma, A. Godoy, J. A. Jimenez-Tejada, J. E. Carceller, and J. A. Lopez-Villanueva, *Phys.Rev.B* **56**, 9565–9574 (1997).

90. N. Zanolla, D. Siprak, P. Baumgartner, E. Sangiorgi, and C. Fiegna, *Ultimate Integration of Silicon*, 2008, pp. 137–140.

91. P. Lenahan, *Microelectron.Eng.* **69**, 173–181 (2003).

92. C. Nicklaw, D. Fleetwood, R. Schrimpf, and S. Pantelides, *IEEE Trans.Nucl.Sci.* **49**, 2667–2673 (2002).