The "Permanent" Component of NBTI Revisited: Saturation, Degradation-Reversal, and Annealing

T. Grasser, M. Waltl* G. Rzepa*, W. Goes*, Y. Wimmer*,*
A.-M. El-Sayed†,, A.L. Shluger†, H. Reisinger•, and B. Kaczer°*

*Institute for Microelectronics, TU Wien, Austria †University College London, UK
•Infineon, Munich, Germany °imec, Leuven, Belgium

***Abstract*— While the defects constituting the recoverable component R of NBTI have been very well analyzed recently, the slower defects forming the more "permanent" component P are much less understood. Using a pragmatic definition for P, we study the evolution of P at elevated temperatures in the range 200°C to 350°C to accelerate these very slow processes. We demonstrate for the first time that P not only clearly saturates, with the saturation value depending on the gate bias, but also that the degradation at constant gate bias can also slowly reverse. Furthermore, at temperatures higher than about 300°C, a significant amount of additional defects is created, which are primarily uncharged around V_{th} but contribute strongly to P at higher V_{G}. Our new data are consistent with our recently suggested hydrogen release model which will be studied in detail using newly acquired long-term data.**

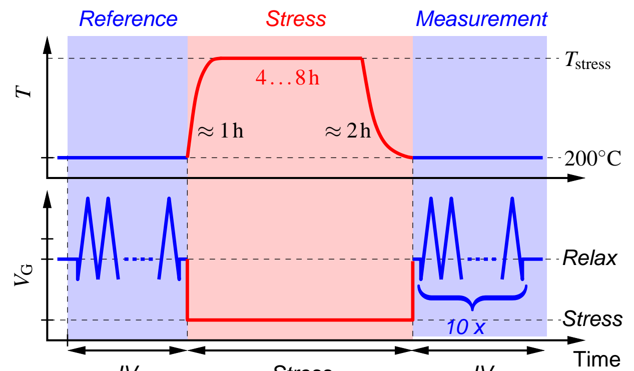

Fig. 1. **The measurement setup:** IV: reference $I_{\mathrm{D}}/V_{\mathrm{G}}$ curve. Then, the device is stressed at a higher temperature (including heating and cooling phases). After the end of the cooling phase, 20 $I_{\mathrm{D}}/V_{\mathrm{G}}$ curves are recorded. From each, $P_i = \Delta V_{\mathrm{th}} @ V_{\mathrm{th}}$ is extracted relative to the initial $I_{\mathrm{D}}/V_{\mathrm{G}}$, yielding $P_{\min} = \min P_i$.

I. Introduction

During the long puzzling history of NBTI the quasi-permanent component P has proven particularly elusive, as it is normally masked by the recoverable component R in typical ΔV_{th} measurements. A number of other experimental techniques, such as electron spin resonance (ESR) [1, 2], spin-dependent recombination (SDR) [2], DCIV [3], charge pumping (CP) [4–6], and CV [7] have been used to extract the increase in the number of slowly recovering defects typically referred to as "interface states". Indeed, all these measurements have revealed defects which are difficult to anneal. Conventionally, these defects are associated with Si dangling bonds, but a conversion of these data to ΔV_{th} and thus to P introduces uncertainties as it requires assumptions regarding the density of states and/or time constants of the responsible defects [8, 9].

In order to avoid these conversion related issues, it is necessary to measure P using the same method as R, which is difficult since ΔV_{th} keeps recovering for very long times, certainly longer than weeks after longer stress times [10]. We have recently suggested and evaluated a pragmatic definition of P using deliberately slow $I_{\mathrm{D}}/V_{\mathrm{G}}$ measurements [11]. During these measurements, V_{G} is ramped into accumulation several times, thereby removing a lot of trapped charge typically associated with R [12]. The remaining degradation is then pragmatically called P, although it is understood that a precise separation of R and P using this approach is impossible [10]. Still, as has been shown [11], the P obtained in this manner is no longer sensitive to any measurement delays and also only recovers weakly. In this context we recall that the degradation accumulated during BTI stress is not really permanent but rather recovers very slowly in a thermally activated manner. In particular, it has been demonstrated that it can be "baked away" at higher temperatures (350–400°C) [13–16]. These temperatures are considerably higher than those typically used during BTI stressing ($\lesssim$ 200°C [17]) and close to those used during the forming gas anneal applied during device fabrication (around 400°C for less 30 minutes [18]). Not surprisingly, we have recently shown that longer bakes (weeks to months) at such high temperatures can lead to additional effects [11], which will also be explored more closely here.

II. Experimental

For our experiments we use packaged devices of a 130nm commercial technology [19] (2.2nm SiON) in computer-controlled furnaces to guarantee flexible and automatic temperature control as well as long-term stability of the experiments. This allows us to acquire long-term data over several months to further elucidate the mechanisms responsible for the creation and annealing of P using the experimental scheme of Fig. 1. Given the thermally activated nature of P and the long time constants involved, degradation is accelerated at higher T (250–350°C). Since the experimental determination of ΔV_{th} at these high temperatures was found to be unreliable, *the temperature was always lowered back to 200°C for the measurement of* ΔV_{th}. Due to the large time constants of the furnaces, these temperature changes took about 1-2 hours. In order to avoid additional effects due to voltage stress, the maximum gate voltage used during stress (−1.5 V) was kept close to the V_{DD} of our technology. At each measurement point, 20 $I_{\mathrm{D}}/V_{\mathrm{G}}$ measurements are taken by slowly sweeping between −1 V and +1 V with a rate of 1 V/s and the remaining ΔV_{th} is extracted at a particular bias, say $V_{\mathrm{G}} = V_{\mathrm{th}}$. In order

to make R and P comparable, this readout voltage should equal the voltage at which R is measured. From these 20 values, the minimum value $P_{\min} = \min_i \Delta V_{\mathrm{th},i}$ is assumed to be representative for P to minimize the measurement noise and remove recovery-related artifacts.

III. Gate-Sided Hydrogen Release Model

We will interpret our experimental results from the perspective of a recently formulated hydrogen-release model in which hydrogen is released from the gate side of the oxide to migrate towards the channel [11]. The release of hydrogen has been the subject of intense discussion during the last decades [18, 20–29], for instance as a consequence of hot electron injection [30, 31]. Since we also see NBTI-like drifts at very small voltages, including 0 V [11], a different mechanism than the injection of hot electrons is required to explain H release. In this context it is worth recalling that the mechanism responsible for the build-up of P cannot rely on proton transport, as has often been speculated in the context of post-irradiation buildup of interface states [18, 20, 26], since the application of a negative voltage at the gate would draw the protons towards the gate rather than the channel. However, just like from the irradiation perspective, where a mechanism based on the release of *neutral* H or H_2 has been suggested [32], a similar mechanism also appears fully consistent with our data as will be shown in the following.

A. Basic Considerations

Based on numerous literature reports [23, 30, 33–36] and our own theoretical and experimental data [37–40], we have recently suggested a hydrogen release (HR) mechanism to explain the evolution of P [11]. In this model, P results when hydrogen atoms trapped in suitable precursor sites [38, 39] are neutralized during stress and detrap over a thermal barrier. The released H^0 can then quickly migrate towards the channel during stress, become trapped in preexisting H-trapping sites, form a defect, and thereby lead to a ΔV_{th} shift (see Fig. 2). We emphasize that the probability of having 'free' (interstitial) H is very small as the vast majority of H is bonded, consistent with experimental results where the observable concentration of interstitial H is below the experimental resolution for temperatures above 130 K [32]. Also note that a significant amount of H exists even in 'ultra-dry' oxides [41], making H an essential player in reliability considerations.

Such a *gate-sided hydrogen release mechanism* has been repeatedly suggested in a number of different contexts and is consistent with several observations:

- In nuclear reaction analysis (NRA) it has been observed that H is moved from the gate towards the channel during stress and back during recovery [36].
- The radical atomic hydrogen released from the gate side also provides the most plausible explanation for the otherwise puzzling observation [42] that during NBTI stress the strong Si-H bonds ($E_{\mathrm{B}} = 2.5$ eV for the direct removal of H [29, 43]) are broken.

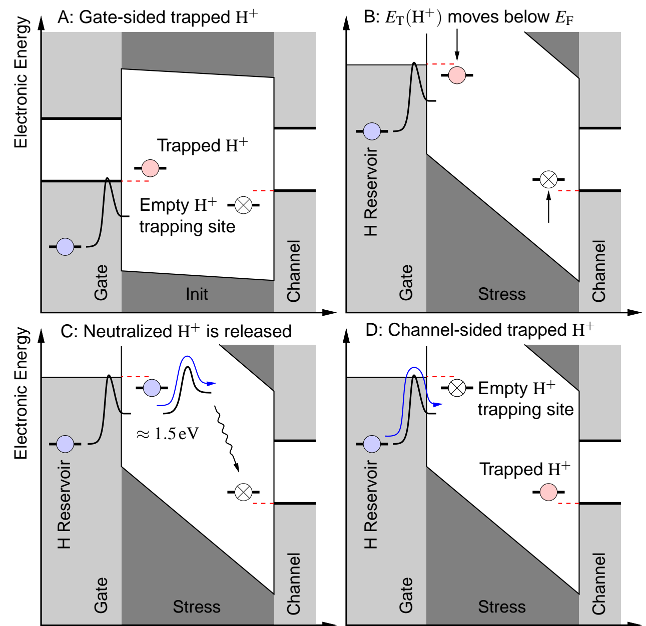

Fig. 2. **Schematic H-release mechanism:** A trapped H^+ at the gate side (A) is neutralized during stress (B/C), released as H^0 over a barrier (C), moves to the channel (C) and becomes trapped in previously inaccessible H^+ trapping sites (D). Eventually, particularly at high T, additional H can be released from the reservoir.

- Along similar lines, it has been found that dopants in the channel can be passivated by H following BTI stress, consistent with the release of H from the oxide [23].
- In an interesting reversal of the situation, in samples with very low H content, BTI stress leads to the *passivation* of Si dangling bonds rather than the typically observed depassivation, also consistent with the release of H inside the oxide and the formation of Si-H bonds [44].
- During BTI stress, an increase in noise has been observed [45, 46], consistent with the creation of H-related defects next to the channel. This noise can be considerably reduced by a bake in accumulation, which moves the H back towards the gate in a reverse reaction [11].

In the next subsection we will develop a mathematical model consistent with the above observations. For this it has to be kept in mind that the configurational flexibility of the SiO_2/Si system is enormous, in particular when hydrogen is added and the various reactions between hydrogen and defect precursor sites are considered [41, 47, 48], not to mention hydrogen-hydrogen reactions and reactions of hydrogen with hydrogen complexes, e.g. defect sites with trapped hydrogen.

In this first modeling attempt, in order to keep the model as simple as possible, all reactions which create H_2 as additional species are neglected. This includes the aforementioned reaction of H with hydrogen-passivated P_{b} centers, P_{b}–H, which creates the notorious P_{b} centers and releases H_2. While it is a well established experimental fact that P_{b} centers *are* created during NBTI stress and *are* the dominant ESR-active defect type in SiO_2 systems [42], it is not clear whether these defects also form the dominant contribution to ΔV_{th}. Based on the observation that the P_{b} center signal saturates while ΔV_{th} keeps increasing, it has been suggested that non-ESR active defects

can dominate ΔV_{th} [30], such as hydrogen related defects. This is consistent with our recent single-trap data obtained on small devices using similar experiments as in the present study, in which P appears do be dominated by oxide traps [11]. So while certainly not all possible reactions are included in this first version of the model explicitly, great care has been taken that the design of the model is compatible with experimental findings and can be extended to include additional reactions to describe the second-order observations listed above if need should arise.

The cornerstone of our model is the recently made demonstration that according to DFT calculations hydrogen can bind to Si–O–Si bridges in amorphous SiO_2 even in defect-free structures [38, 39], as has been suggested before based on experimental data [23]. The only prerequisite for this to happen appears to be the presence of stretched Si–O–Si bonds, which occur in abundance particularly close to the interfaces. Note that this is not the case in crystalline SiO_2, where only the proton can bind in a stable manner to the bridging oxygen [49]. Upon attachment of H to these stretched bonds, the H can give up its electron to one of the Si atoms and bind to the O in a configuration similar to the well known protonic configuration. Alternatively, the extra H can break one of the Si–O bonds and form a hydroxyl group (–OH), which then faces the dangling bond on the other Si, a configuration very similar to the familiar E' center [50]. We have recently shown that the properties of this so-called hydroxyl E' center are in good agreement with the defects responsible for charge trapping in pMOS transistors [37], and thus for the recoverable component of BTI, R. Furthermore, since the H can be removed over a thermal barrier, leaving an electrically inactive defect precursor behind, this defect appears to be consistent with the observed volatility of oxide traps [40] as well. In a next evolutionary step of these findings we suggest that the removal of H from this defect-family (either directly from hydroxyl E' centers or their non-bondbreaking variants) leads to H migration towards the gate and thus to the creation of P [11]. This provides a very promising H release mechanism to explain the experimental observations listed in Section III-A.

B. Model Equations

A simple model can be formulated in which the amorphous oxide is assumed to consist of discrete sites i at which hydrogen can either occur in a neutral interstitial position (H_i), in a trapped neutral configuration (H_i^0), or in a trapped positive configuration (H_i^+) as a "proton". Note that this protonic configuration exists in many more variants than the conventional proton bound to a bridging oxygen, as it can also break the bond and lead to the formation of hydroxyl (–OH) groups and the puckering of the now detached Si atom [38, 39]. These variants are required to explain the recoverable component R [51] but are for now not further considered here.

Since only a small number of H atoms will be allowed to trap at a particular site i, the model will be expressed in absolute numbers rather than in concentrations, so H_i^+ gives the absolute number of trapped and positively charged H atoms at site i, or – to be more precise – its expectation value [52].

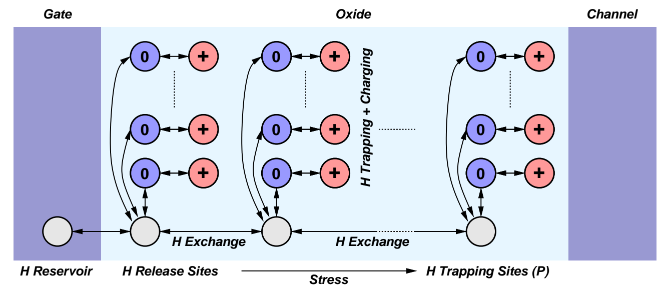

Fig. 3. A schematic view of the simple H-release model (in 1D). H can get trapped and the trapping sites are connected via barriers to the interstitial configurations. Redistribution of H in these configurations (diffusion) is very fast and not rate-limiting [32]. At $T \gtrsim 300°\mathrm{C}$, additional H can be exchanged with the reservoir, separated by a 2.5 eV barrier (e.g. Si-H bonds [29, 43]).

Hydrogen exchange between different sites is modeled as a thermally activated hopping process with hopping rate $k_{\mathrm{H}} = k_{ij} = k_0 \exp(-\mathrm{q}E_{\mathrm{H}}/\mathrm{k_B}T)$. According to literature values for H diffusion in SiO_2, which give a prefactor of $D_{\mathrm{H}} = 10^{-4}\,\mathrm{cm^2/s}$ [32] and a typical hopping distance of $a = 5\,\mathrm{nm}$ [53], we obtain for $k_0 = D_{\mathrm{H}}/a^2 = 4\times10^8\,\mathrm{s^{-1}}$. Together with the very small activation energy for diffusion ($E_{\mathrm{H}} = 0.18\,\mathrm{eV}$) [32] this results in a very large value for k_{H}, meaning that redistribution of the detrapped H will be very fast and the precise value of the rate immaterial. In particular, the model results stay practically unchanged if a uniform concentration of *free interstitial* H is used in the model. Overall, the response of the system is thus determined by the thermal release of the trapped H over barriers with widely distributed barrier heights, and thus *reaction- rather than diffusion-limited.*

In the following, the rate equations for the three different H species (interstitial, trapped neutral, trapped positive) are summarized. The temporal change of the interstitial H at site i is given by

$$\frac{\partial H_i}{\partial t} = -\sum_j k_{\mathrm{H}}\big(H_i - H_j\big) - \sum_n T_{i,n}\ , \qquad (1)$$

where the first sum runs over all neighboring sites j. In order to reduce the computational effort, one spatial interstitial site i is allowed to interact with several trapping sites n, see Fig. 3. This is justified as the free hydrogen H_i is in quasi-equilibrium across the oxide anyway. The corresponding trapping rates $T_{i,n}$ are given by

$$T_{i,n} = k_{01} H_i \big(H_{\mathrm{max}}^{\mathrm{T}} - (H_{i,n}^0 + H_{i,n}^+)\big) - k_{10} H_{i,n}^0\ , \qquad (2)$$

which describes the transition from the interstitial to the neutral trapped configuration via the rate k_{01} and the transition back via the rate k_{10} for each trapping site (i,n). These rates are modeled using simple Arrhenius laws with widely distributed barriers E_{01} and E_{10}. Also, each interstitial position is allowed to harbor an infinite number of hydrogen atoms, a modeling simplification justified by the very low concentration of free atomic hydrogen and the fact that two hydrogen atoms in the same interstitial position would dimerize to H_2 anyway, an aspect currently not considered in the model. At each trapping site (i,n), however, only a certain number of trapped hydrogen is allowed, $H_{i,n}^0 + H_{i,n}^+ \le H_{\mathrm{max}}^{\mathrm{T}}$, expressed by the

$H_{\max}^{T} - (H_{i,n}^{0} + H_{i,n}^{+})$ term in the model.

The temporal change in the number of neutral and positive trapped hydrogen atoms is given by (index n omitted for readability reasons)

$$\frac{\partial H_i^0}{\partial t} = -k_{12} H_i^0 + k_{21} H_i^+ + T_i \ , \quad (3)$$

$$\frac{\partial H_i^+}{\partial t} = \ k_{12} H_i^0 - k_{21} H_i^+ \ . \quad (4)$$

The transition rates k_{12} from the neutral to the positively trapped state and the backward transition k_{21} are modeled using standard non-radiative multiphonon theory [51] again with widely distributed relaxation energies.

Finally, in order to explain data acquired at $T > 300\,°\mathrm{C}$, we assume that in close proximity of the gate stack there is additional hydrogen available bound in a reservoir, H_R. It has been suggested many times that hydrogen trapped in higher layers of the gate stack (poly-silicon and metalization) can have a significant impact on gate stack reliability [35, 36, 54]. In our model we assume that this hydrogen is allowed to enter the oxide at the gate side over a thermal barrier according to

$$\frac{\partial H_\mathrm{R}}{\partial t} = -k_{\mathrm{R}0} H_\mathrm{R} + k_{0\mathrm{R}} H_i \quad (5)$$

for selected sites i at the gate/oxide interface. Since H redistribution is very fast, the precise choice of interaction sites does not matter as the reaction is dominated by the large thermal barrier (2.5 eV) and triggered whenever H_i at the interface is lower than the equilibrium value. This will happen during stress, when H at the interface is neutralized ($H_i^+ \Rightarrow H_i^0$), released over the barrier ($H_i^0 \Rightarrow H_i$), and transported from site i to another site j closer to the channel side ($H_i \Rightarrow H_j$), where it now can become trapped at the changed bias conditions ($H_j \Rightarrow H_j^0 \Rightarrow H_j^+$)

At the beginning of the simulation, a certain amount of hydrogen is placed into the oxide and reservoir (on the order of $10^{19}\,\mathrm{cm}^{-3}$ [32]) and allowed to redistribute by determining the stationary solution of the problem. Note that the initial H distribution between the reservoir and the oxide as well as the spatial distribution in the oxide will depend on the applied gate bias. While this could be used to emulate the initial H distribution after processing, the precise processing conditions are not too well defined, so currently a gate voltage of 0 V is applied.

IV. Results and Discussion

At first, the surprising ΔV_th shift at $V_\mathrm{G} = 0\,\mathrm{V}$ is studied in more detail, see Fig. 4. While at 250 and 300°C the shift of about 5 mV is consistent with a collection of normally distributed first-order processes [51], at 350°C an about 2.4 times higher drift is observed, which we interpret as the creation of additional defects. This implies that the previously suggested method [13, 14] to restore a device to its pristine state using a bake at 350/400°C has to be used with care.

Next, a device is cycled between 0 and −1.5 V at various T, see Fig. 5. In this first experiment, the device is initially and between the stress phases annealed at 0 V/350°C. As already anticipated in Fig. 4, even during this anneal additional

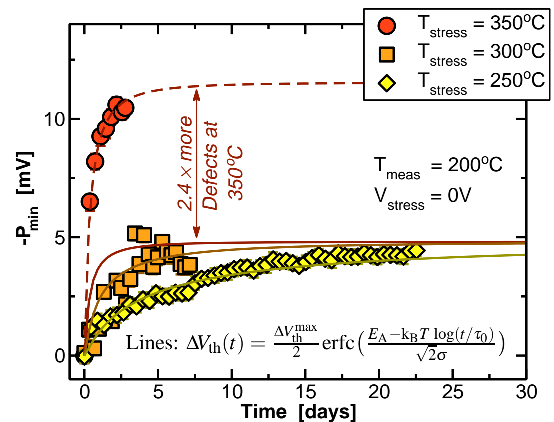

Fig. 4. Devices drift slowly even at 0 V “stress”. This effect can be more easily characterized at high T. The degradation typically saturates at a few millivolts, with the saturation level depending on V_G and T. At $T > 300\,°\mathrm{C}$ new defects are created and saturation occurs at a higher level. Lines: simple model based on normally distributed “effective” activation energies [55].

defects are created. During stress at 300°C, the degradation saturates after 3 days. However, in the subsequent stress phase at −1.5 V/350°C, a massive amount of new defects is created, which are consistently modeled by a release of H from the reservoir.

The HR model can explain all these observations very well, see lines in Fig. 5 (shown with and without coupling to the reservoir). In particular, the H released from the reservoir can explain the additionally created defects at higher T. This experiment is analyzed in more detail in Fig. 6: in the first degradation phase trapped protons at the gate side are very slowly neutralized following a nonradiative multiphonon process [51, 56, 57] and eventually released over a distributed barrier with mean 1.5 eV [38, 39]. The released H^0 can now find previously unavailable trapping sites at the channel side and become charged. With increasing stress time, the H centroid thus moves from the gate to the channel side of the oxide, in agreement with experimental NRA data [36]. Fig. 7 shows the increase in trapped H due to the release from the reservoir. These newly created H-related defects appear identical to the preexisting ones and are more likely to be positively charged at −1.5 V than 0 V.

In the next long-term experiment, see Fig. 8, bias stress at 350°C is avoided to clarify whether new defects are created predominantly during stress or recovery. As can be seen during the 300°C stress in Phase H, defect creation was also very efficient during the preceding recovery in Phase G. The simulated H profiles are shown in Fig. 9, clearly highlighting how H is moved back and forth inside the oxide. Also shown is the increase of the H concentration by the interaction with the reservoir even during recovery at 0 V/350°C. Interestingly, in Phases D and F (−1.5 V/300°C), *the degradation is partially reversed*, an effect not yet covered by the model.

In the final measurement of Fig. 10, 350°C is avoided altogether. The saturation during the −1.5 V/250°C stress (Phase B) is well observable together with a reversal of degradation at higher readout voltages. This reversal is not yet fully understood but appears to be due to defects with higher

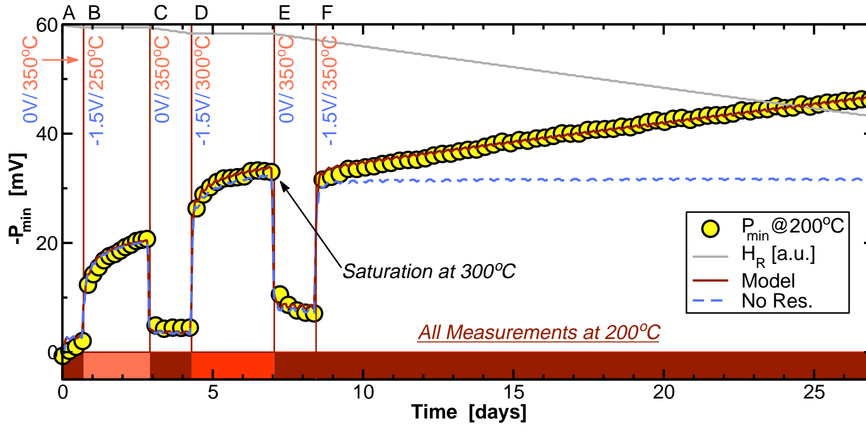

Fig. 5. Long-term degradation experiment where V_G is cycled between 0 V and $V_{DD} = -1.5$ V at various T. At the end of Phase D (7 days), degradation at 300°C saturates at about 33 mV. In the subsequent recovery in Phase E, a slightly higher P_{min}@200°C is already observed. During the next stress at 350°C, a large number of additional defects are created, modeled by the release of H from a reservoir with concentration H_R separated by a 2.5 eV barrier. In the next anneal step (Phase G) at 350°C, most of these defects are neutralized but can be easily recharged in Phase H even at 300°C. Only few additional defects are created at 300°C. Also shown is the model without a reservoir where the maximum amount of H in the oxide is fixed.

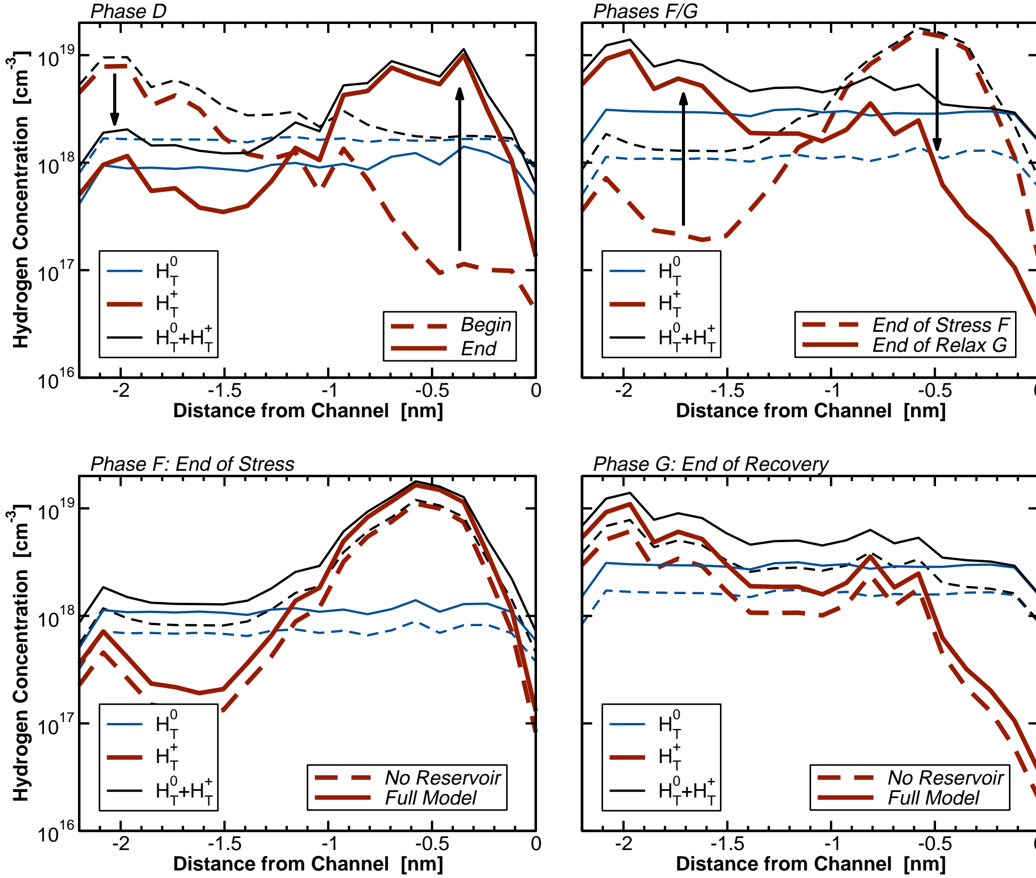

Fig. 6. **H profiles to Fig. 5:** Shown are the trapped neutral hydrogen (H_T^0), the trapped protons (H_T^+), as well as their sum at the beginning and end of Phase D (left) and the end of the stress and recovery in Phases F and G (right). The noise in the simulation data is due to the stochastic simulation algorithm employed.

Fig. 7. Comparison of the H profiles to Fig. 5 for the model with and without the reservoir. The H is more likely to be positively charged and thus impact ΔV_{th} during stress (left) than during recovery (right).

energy levels which are annealed during stress. As can be seen in the subsequent stress phase D, once defects are lost during this degradation reversal, they can no longer be charged.

Summarizing these observations, our model suggests that at temperatures lower than 300°C the build-up of P is the consequence of H being moved towards the channel, while the recovery of P would be due to the inverse process. The creation of additional defects at $T > 300$°C, on the other hand, is modeled by H released from a reservoir over a 2.5 eV barrier (close to the Si-H binding energy [43]), see Fig. 2(D). The effect of the additional hydrogen reservoir is determined by the relatively high barrier of 2.5 eV, which essentially "turns on" H release when the temperature is changed e.g. from 300°C to 350°C: assuming the process to follow an Arrhenius law with an attempt frequency of $\nu = 10^{13}$ Hz, we obtain for the mean time constant $\tau = \nu^{-1}\exp(-qE_A/k_BT)$ about 190 days at 350°C, which is within reach of our experimental efforts, and about 30 years at 300°C. Naturally, under regular operating conditions this process is completely irrelevant, as $\tau(25°C) \approx 6 \times 10^{21}$ years, but it is fundamental for the correct interpretation of accelerated high temperature degradation data. This newly released H can now go into previously inaccessible precursor sites at different bias conditions. Once the H is in the oxide, the newly created H-defects are difficult to anneal and act as regular defects also at lower T.

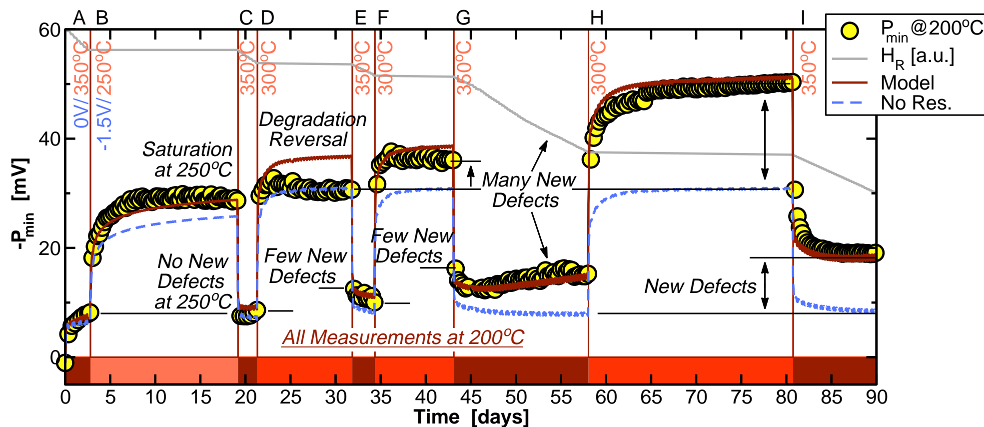

Fig. 8. Long-term experiment with bias stress (−1.5 V) at a maximum of 300°C in order to minimize defect creation during stress at 350°C. Still, in the long recovery phase at 0 V/350°C, a large number of defects are obviously created. While the *measurable* degradation at 0 V/350°C is small, since most H released from the reservoir remains neutral, all these newly created H-related defects can be charged at the subsequent stress phase, resulting in nearly a doubling of P_{min}. Note the small reversal of the degradation during Phases D and F.

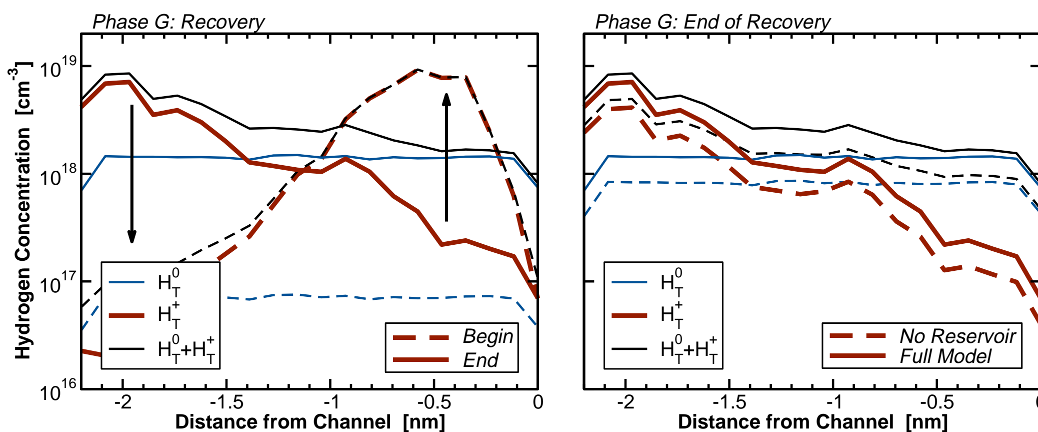

Fig. 9. **H profiles to Fig. 8:** During the long recovery in Phase G at 350°C, the amount of H in the oxide is considerably increased. At the beginning of recovery (left), most trapped H-traps are positively charged and at the channel side of the oxide. During recovery, the traps are neutralized, H^0 is released over a barrier, and eventually migrates back to the gate side of the oxide where it gets trapped in unoccupied defect precursors, thereby activating the defect. Right: comparison between the full model and the model without a reservoir reveals the considerably increased H (and thus defect) concentration in the oxide.

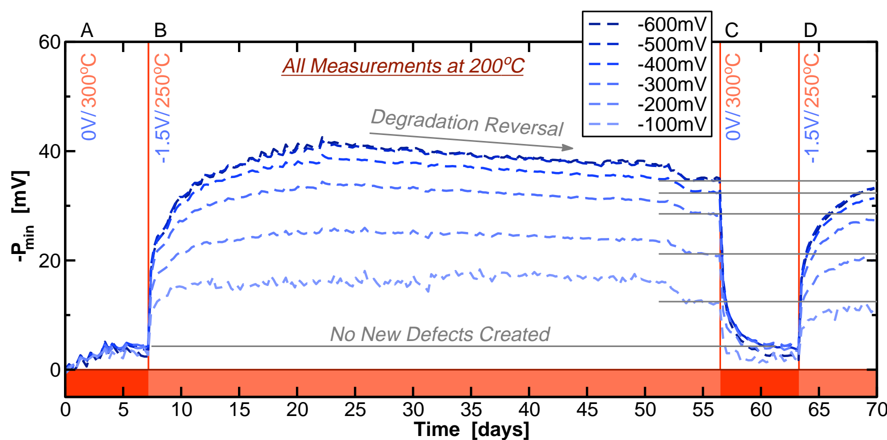

Fig. 10. Long-term experiment without 350°C phases. Shown is $P_{min}(V_G)$ at 200°C extracted from the same set of I_D/V_G curves at different V_G. At $V_G = -100$ mV the degradation saturates after a couple of days while at higher V_G the degradation saturates at a much later time (≈15 days) and then begins to *decrease* again, similarly to Phases D/F in Fig. 8. This decrease indicates that defects with an energy level closer to E_V are permanently annealed. The defects annealed in Phase B cannot be recharged again in Phase D and appear to be lost.

V. Link to Previous Models

During standard BTI degradation at $T < 300$°C, the HR model will behave similarly to the first-order reaction-limited model proposed by Huard *et al.* [17]. The most important difference will be in the interpretation of the effective activation energies and the bias-dependence of the model: while the empirical model of Huard is based on a dipole-field-interaction of the barriers, the bias-dependence of the HR model will be essentially determined by the active areas of the oxide where defects can change their charge state [51]. While for higher stress fields these models will be very similar, the HR model is more accurate at lower voltages (use-conditions). We have previously extended Huard's model to also consider recovery of the "permanent" component [55, 58]. The main difference in the HR model is that the effective forward and backward rates are *uncorrelated*, since different precursor sites will be responsible for degradation and recovery, consistent with experimental data.

VI. Conclusions

Using new long-term data we have carefully validated our previously suggested hydrogen release model which also tries to reconcile a number of puzzling observations inconsistent with available models. In this model, H is slowly released at

the gate side to hop quickly towards the channel side of the oxide, thereby increasing ΔV_{th}. As a side effect, this model naturally explains the depassivation of Si-H bonds as well as the passivation of channel dopants, the increase/decrease of noise during stress/recovery, as well as the sensitivity of the permanent component to the H concentration introduced during fabrication. Previously suggested reaction-limited models will be special cases of the HR model valid at large fields but lower temperatures. Finally, the HR model can correctly describe high-temperature data through coupling to a H reservoir. While this coupling will be unimportant at use-conditions, the HR model thereby allows for a correct extrapolation from high temperature data.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Austrian Science Fund (FWF) project n°26382-N30, the European Community's FP7 project n°619234 (MoRV), as well as the Intel Sponsored Research Project n°2013111914.

REFERENCES

[1] P. Lenahan and J. Conley Jr., "What Can Electron Paramagnetic Resonance Tell Us about the Si/SiO$_2$ System?" *J.Vac.Sci.Technol.B*, vol. 16, no. 4, pp. 2134–2153, 1998.

[2] J. Campbell, P. Lenahan, A. Krishnan, and S. Krishnan, "Identification of the Atomic-Scale Defects Involved in the Negative Bias Temperature Instability in Plasma-Nitrided p-Channel Metal-Oxide-Silicon Field-Effect Transistors," *J.Appl.Phys.*, vol. 103, no. 4, p. 044505, 2008.

[3] A. Neugroschel, G. Bersuker, R. Choi, C. Cochrane, P. Lenahan, D. Heh, C. Young, C. Kang, B. Lee, and R. Jammy, "An Accurate Lifetime Analysis Methodology Incorporating Governing NBTI Mechanisms in High-k/SiO2 Gate Stacks," in *Proc. Intl.Electron Devices Meeting (IEDM)*, 2006, pp. 1–4.

[4] D. Ang, S. Wang, and C. Ling, "Evidence of Two Distinct Degradation Mechanisms from Temperature Dependence of Negative Bias Stressing of the Ultrathin Gate p-MOSFET," *IEEE Electron Device Lett.*, vol. 26, no. 12, pp. 906–908, 2005.

[5] T. Aichinger, M. Nelhiebel, and T. Grasser, "A Combined Study of p- and n-Channel MOS Devices to Investigate the Energetic Distribution of Oxide Traps after NBTI," *IEEE Trans.Electron Devices*, vol. 56, no. 12, pp. 3018–3026, 2009.

[6] Z. Teo, D. Ang, and K. See, "Can the Reaction-Diffusion Model Explain Generation and Recovery of Interface States Contributing to NBTI?" in *Proc. Intl.Electron Devices Meeting (IEDM)*, 2009, pp. 737–740.

[7] L. Ragnarsson and P. Lundgren, "Electrical Characterization of P_b Centers in (100)Si/SiO$_2$ Structures: The Influence of Surface Potential on Passivation During Post Metallization Anneal," *J.Appl.Phys.*, vol. 88, no. 2, pp. 938–942, 2000.

[8] T. Aichinger, M. Nelhiebel, and T. Grasser, "On the Temperature Dependence of NBTI Recovery," *Microelectronics Reliability*, vol. 48, no. 3, pp. 1178–1184, 2008.

[9] ——, "Refined NBTI Characterization of Arbitrarily Stressed PMOS Devices at Ultra-Low and Unique Temperatures," *Microelectronics Reliability*, vol. 53, no. 7, pp. 937–946, 2013.

[10] T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P.-J. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer, "The 'Permanent' Component of NBTI: Composition and Annealing," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, Apr. 2011, pp. 605–613.

[11] T. Grasser, M. Waltl, Y. Wimmer, W. Goes, R. Kosik, G. Rzepa, H. Reisinger, G. Pobegen, A. El-Sayed, A. Shluger, and B. Kaczer, "Gate-Sided Hydrogen Release as the Origin of "Permanent" NBTI Degradation: From Single Defects to Lifetimes," in *Proc. Intl.Electron Devices Meeting (IEDM)*, Dec. 2015.

[12] T. Grasser, K. Rott, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, M. Waltl, M. Toledano-Luque, and B. Kaczer, "Advanced Characterization of Oxide Traps: The Dynamic Time-Dependent Defect Spectroscopy," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, Apr. 2013, pp. 2D.2.1–2D.2.7.

[13] A. Katsetos, "Negative Bias Temperature Instability (NBTI) Recovery with Bake," *Microelectronics Reliability*, vol. 48, no. 10, pp. 1655–1659, 2008.

[14] G. Pobegen, T. Aichinger, M. Nelhiebel, and T. Grasser, "Understanding Temperature Acceleration for NBTI," in *Proc. Intl.Electron Devices Meeting (IEDM)*, Dec. 2011, pp. 27.3.1–27.3.4.

[15] Y. Yonamoto, "Generation/Recovery Mechanism of Defects Responsible for the Permanent Component in Negative Bias Temperature Instability," *J.Appl.Phys.*, vol. 113, p. 154501, 2013.

[16] G. Pobegen and T. Grasser, "On the Distribution of NBTI Time Constants on a Long, Temperature-Accelerated Time Scale," *IEEE Trans.Electron Devices*, vol. 60, no. 7, pp. 2148–2155, 2013.

[17] V. Huard, M. Denais, and C. Parthasarathy, "NBTI Degradation: From Physical Mechanisms to Modelling," *Microelectronics Reliability*, vol. 46, no. 1, pp. 1–23, 2006.

[18] D. Brown and N. Saks, "Time Dependence of Radiation-Induced Trap Formation in Metal-Oxide-Semiconductor Devices as a Function of Oxide Thickness and Applied Field," *J.Appl.Phys.*, vol. 70, no. 7, pp. 3734–3747, 1991.

[19] H. Reisinger, O. Blank, W. Heinrigs, A. Mühlhoff, W. Gustin, and C. Schlünder, "Analysis of NBTI Degradation- and Recovery-Behavior Based on Ultra Fast V_{th}-Measurements," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, 2006, pp. 448–453.

[20] F. McLean, "A Framework for Understanding Radiation-Induced Interface States in SiO$_2$ Structures," *IEEE Trans.Nucl.Sci.*, vol. 27, no. 6, pp. 1651–1657, Dec 1980.

[21] J. Conley Jr. and P. Lenahan, "Molecular Hydrogen, E$'$ Center Hole Traps, and Radiation Induced Interface Traps in MOS Devices," *IEEE Trans.Nucl.Sci.*, vol. 40, no. 6, pp. 1335–1340, 1993.

[22] E. Cartier, J. Stathis, and D. Buchanan, "Passivation and Depassivation of Silicon Dangling Bonds at the Si(111)/SiO$_2$ Interface by Atomic Hydrogen," *Appl.Phys.Lett.*, vol. 63, no. 11, pp. 1510–1512, 1993.

[23] J. de Nijs, K. Druijf, V. Afanas'ev, E. van der Drift, and P. Balk, "Hydrogen Induced Donor-Type Si/SiO$_2$ Interface States," *Appl.Phys.Lett.*, vol. 65, no. 19, pp. 2428–2430, 1994.

[24] D. DiMaria and J. Stathis, "Non-Arrhenius Temperature Dependence of Reliability in Ultrathin Silicon Dioxide Films," *Appl.Phys.Lett.*, vol. 74, no. 12, pp. 1752–1754, 1999.

[25] J. Zhang, C. Zhao, H. Sii, G. Groeseneken, R. Degraeve, J. Ellis, and C. Beech, "Relation between Hole Traps and Hydrogenous Species in Silicon Dioxides," *Solid-State Electron.*, vol. 46, pp. 1839–1847, 2002.

[26] S. Rashkeev, D. Fleetwood, D. Schrimpf, and S. Pantelides, "Dual Behavior of H$^+$ at Si-SiO$_2$ Interfaces: Mobility Versus Trapping," *Appl.Phys.Lett.*, vol. 81, no. 10, pp. 1839–1841, 2002.

[27] J. Sune and Y. Wu, "Hydrogen-Release Mechanisms in the Breakdown of Thin SiO_2 Films," *Physical Review Letters*, vol. 92, no. 8, pp. 087 601 (1–4), 2004.

[28] D. Fleetwood, X. Zhou, L. Tsetseris, S. Pantelides, and R. Schrimpf, "Hydrogen Model for Negative-Bias Temperature Instabilities in MOS Gate Insulators," in *Silicon Nitride and Silicon Dioxide Thin Insulating Films and Other Emerging Dielectrics VIII*, R. Sah, M. Deen, J. Zhang, J. Yota, and Y. Kamakura, Eds., 2005, pp. 267–278.

[29] L. Tsetseris, X. Zhou, D. Fleetwood, R. Schrimpf, and S. Pantelides, "Hydrogen-Related Instabilities in MOS Devices Under Bias Temperature Stress," *IEEE Trans.Dev.Mat.Rel.*, vol. 7, no. 4, pp. 502–508, 2007.

[30] E. Cartier, "Characterization of the Hot-Electron-Induced Degradation in Thin SiO$_2$ Gate Oxides," *Microelectronics Reliability*, vol. 38, no. 2, pp. 201–211, 1998.

[31] J. Zhang, H. Sii, R. Degraeve, and G. Groeseneken, "Mechanism for the Generation of Interface State Precursors," *J.Appl.Phys.*, vol. 87, no. 6, pp. 2967–2977, 2000.

[32] D. Griscom, "Diffusion of Radiolytic Molecular Hydrogen as a Mechanism for the Post-Irradiation Buildup of Interface States in SiO$_2$-on-SiStructures," *J.Appl.Phys.*, vol. 58, no. 7, pp. 2524–2533, 1985.

[33] E. Cartier and J. Stathis, "Atomic Hydrogen-Induced Degradation of the Si/SiO$_2$ Structure," *Microelectronic Engineering*, vol. 28, no. 1-4, pp. 3–10, 1995.

[34] V. Afanas'ev and A. Stesmans, "Hydrogen-Induced Valence Alternation State at SiO$_2$ Interfaces," *Physical Review Letters*, vol. 80, pp. 5176–5179, 6 1998.

[35] M. Nelhiebel, J. Wissenwasser, T. Detzel, A. Timmerer, and E. Bertagnolli, "Hydrogen-Related Influence of the Metallization Stack on Characteristics and Reliability of a Trench Gate Oxide," *Microelectronics Reliability*, vol. 45, pp. 1355–1359, 2005.

[36] Z. Liu, S. Fujieda, H. Ishigaki, M. Wilde, and K. Fukutani, "Current Understanding of the Transport Behavior of Hydrogen Species in MOS

Stacks and Their Relation to Reliability Degradation," *ECS Trans.*, vol. 35, no. 4, pp. 55 – 72, 2011.

[37] T. Grasser, W. Goes, Y. Wimmer, F. Schanovsky, G. Rzepa, M. Waltl, K. Rott, H. Reisinger, V. Afanas'ev, A. Stesmans, A. El-Sayed, and A. Shluger, "On the Microscopic Structure of Hole Traps in pMOSFETs," in *Proc. Intl.Electron Devices Meeting (IEDM)*, Dec. 2014.

[38] A. El-Sayed, M. Watkins, T. Grasser, V. Afanas'ev, and A. Shluger, "Hydrogen induced rupture of strained Si-O bonds in amorphous silicon dioxide," *Physical Review Letters*, vol. 114, no. 11, p. 115503, 2015.

[39] A. El-Sayed, Y. Wimmer, W. Goes, T. Grasser, V. Afanas'ev, and A. Shluger, "Theoretical Models of Hydrogen-Induced Defects in Amorphous Silicon Dioxide," *Physical Review B*, vol. 92, no. 11, p. 014107, 2015.

[40] T. Grasser, M. Waltl, W. Goes, A. El-Sayed, A. Shluger, and B. Kaczer, "On the Volatility of Oxide Defects: Activation, Deactivation, and Transformation," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, June 2015.

[41] E. Poindexter, "Chemical Reactions of Hydrogeneous Species in the Si/SiO_2 System," *J. Noncryst. Solids.*, vol. 187, pp. 257–263, 1995.

[42] J. Campbell, P. Lenahan, A. Krishnan, and S. Krishnan, "Observations of NBTI-Induced Atomic-Scale Defects," *IEEE Trans.Dev.Mat.Rel.*, vol. 6, no. 2, pp. 117–122, 2006.

[43] A. Stesmans, "Dissociation Kinetics of Hydrogen-Passivated P_b Defects at the (111)Si/SiO_2 Interface," *Physical Review B*, vol. 61, no. 12, pp. 8393–8403, 2000.

[44] M. Houssa, V. Afanas'ev, A. Stesmans, M. Aoulaiche, G. Groeseneken, and M. Heyns, "Insights on the Physical Mechanism behind Negative Bias Temperature Instabilities," *Appl.Phys.Lett.*, vol. 90, no. 4, p. 043505, 2007.

[45] G. Kapila, N. Goyal, V. Maheta, C. Olsen, K. Ahmed, and S. Mahapatra, "A Comprehensive Study of Flicker Noise in Plasma Nitrided SiON p-MOSFETs: Process Dependence of Pre-Existing and NBTI Stress Generated Trap Distribution Profiles," in *Proc. Intl.Electron Devices Meeting (IEDM)*, 2008, pp. 103–106.

[46] B. Kaczer, T. Grasser, J. Martin-Martinez, E. Simoen, M. Aoulaiche, P. Roussel, and G. Groeseneken, "NBTI from the Perspective of Defect States with Widely Distributed Time Scales," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, 2009, pp. 55–60.

[47] C. Van de Walle and B. Tuttle, "Microscopic Theory of Hydrogen in Silicon Devices," *IEEE Trans.Electron Devices*, vol. 47, no. 10, pp. 1779–1786, 2000.

[48] P. Blöchl, "First-Principles Calculations of Defects in Oxygen-Deficient Silica Exposed to Hydrogen," *Physical Review B*, vol. 62, no. 10, pp. 6158–6179, 2000.

[49] J. Godet, F. Giustino, and A. Pasquarello, "Proton-Induced Fixed Positive Charge at the SiO_2 Interface," *Physical Review Letters*, vol. 99, no. 12, pp. 126 102–1–126 102–4, 2007.

[50] P. Lenahan, "Atomic Scale Defects Involved in MOS Reliability Problems," *Microelectronic Engineering*, vol. 69, pp. 173–181, 2003.

[51] T. Grasser, "Stochastic Charge Trapping in Oxides: From Random Telegraph Noise to Bias Temperature Instabilities," *Microelectronics Reliability*, vol. 52, pp. 39–70, 2012.

[52] D. Gillespie, *Markov Processes: An Introduction for Physical Scientists*. Academic Press, 1992.

[53] G. Malavasi, M. Menziani, A. Pedone, and U. Segre, "Void Size Distribution in MD-Modelled Silica Glass Structures," *J.Non-Cryst.Solids*, vol. 352, no. 3, pp. 285–296, 2006.

[54] T. Aichinger, S. Puchner, M. Nelhiebel, T. Grasser, and H. Hutter, "Impact of Hydrogen on Recoverable and Permanent Damage following Negative Bias Temperature Stress," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, 2010, pp. 1063–1068.

[55] T. Grasser, P.-J. Wagner, H. Reisinger, T. Aichinger, G. Pobegen, M. Nelhiebel, and B. Kaczer, "Analytic Modeling of the Bias Temperature Instability Using Capture/Emission Time Maps," in *Proc. Intl.Electron Devices Meeting (IEDM)*, Dec. 2011, pp. 27.4.1–27.4.4.

[56] C. Henry and D. Lang, "Nonradiative Capture and Recombination by Multiphonon Emission in GaAs and GaP," *Physical Review B*, vol. 15, no. 2, pp. 989–1016, 1977.

[57] M. Kirton and M. Uren, "Noise in Solid-State Microstructures: A New Perspective on Individual Defects, Interface States and Low-Frequency (1/f) Noise," *Adv.Phys.*, vol. 38, no. 4, pp. 367–486, 1989.

[58] T. Grasser, B. Kaczer, and W. Goes, "An Energy-Level Perspective of Bias Temperature Instability," in *Proc. Intl.Rel.Phys.Symp. (IRPS)*, 2008, pp. 28–38.