A Combined Study of p- and n-Channel MOS Devices to Investigate the Energetic Distribution of Oxide Traps After NBTI

Thomas Aichinger, Michael Nelhiebel, and Tibor Grasser, Senior Member, IEEE

Abstract—The aim of this paper is to highlight the effect of gate bias switches and charge pumping (CP) on oxide trap and interface-state recovery. In general, variations in gate bias correspond to shifts of the Fermi level (E_F) across the silicon band gap and trigger carrier exchange between stress-induced oxide traps and the silicon substrate. Our measurements strongly indicate that interface-state recovery is accelerated by the CP measurement itself, whereas oxide-trap occupation can be controlled more efficiently by slow Fermi level switches. Oxide-trap neutralization/ charging by electron/hole capture works similarly to interface states but with larger time constants indicating inelastic carrier tunneling between the silicon substrate and stress-induced donorlike oxide traps. In a combined study, we compare thresholdvoltage shifts and recovery of identically processed NMOS and PMOS devices which allows us to gain access to the full silicon band gap by appropriate gate biasing. Additional CP measurements on identically stressed reference devices allow us furthermore to differentiate quantitatively between interface-state and oxide-trap contributions. Finally, by referring oxide-trapdependent V_{TH} shifts after stress to certain gate voltages during recovery, energetic profiling of oxide traps with respect to the Fermi level position becomes possible.

Index Terms—Charge pumping (CP), interface states, negativebias temperature instability (NBTI), oxide traps.

I. Introduction

THE CURRENT understanding of negative-bias temperature instability (NBTI) in MOS devices suggests interface states and/or oxide traps to be created during stress [1]–[3]. The degradation of an electrically stressed MOS device is usually characterized by evaluating the threshold-voltage shift using a constant drain current criterion. During the measurement of a standard transfer curve, the gate voltage is generally swept from accumulation to inversion while a constant bias is applied between source and drain. From an energetic point of view,

Manuscript received March 17, 2009; revised August 24, 2009. First published October 30, 2009; current version published November 20, 2009. This work was supported in part by the Federal Ministry of Economics and Labour of the Republic of Austria under Contract 98.362/0112-C1/10/2005 and in part by the Carinthian Economic Promotion Fund (KWF) under Contract 98.362/0112-C1/10/2005. The review of this paper was arranged by Editor J. Suehle.

- T. Aichinger is with Kompetenzzentrum Automobil- und Industrielektronik, 9500 Villach, Austria (e-mail: thomas.aichinger@k-ai.at).
- M. Nelhiebel is with Infineon Technologies Austria, 9500 Villach, Austria (e-mail: michael.nelhiebel@infineon.com).
- T. Grasser is with the Christian Doppler Laboratory for TCAD in Microelectronics, Institute for Microelectronics, Technical University of Vienna, 1040 Vienna, Austria (e-mail: grasser@iue.tuwien.ac.at).
- Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2009.2033008

this means that the Fermi level E_F at the gate-oxide-siliconsubstrate interface moves incrementally from one band edge to the other, resulting in an exponential growth of the drain current in the subthreshold region of the device. According to Shockley-Read-Hall (SRH) theory, interface states located energetically within the silicon band gap can follow such changes in E_F very quickly and quasi-instantaneously adjust their charge state (relative to the time scales of our measurements). This means that traps above the Fermi level become uncharged while traps below E_F stay occupied as the Fermi level crosses the silicon band gap during a gate bias ramp. As a result, we observe a variation in the subthreshold slope of the transfer curve which is usually attributed to charging of stressinduced interface states $(N_{\rm it})$ [4]-[9]. Recent research suggests, however, that oxide traps can also change their occupancy with respect to E_F , which results in an additional contribution to the variation in the subthreshold slope [9]–[11].

In this paper, we extend the idea of rechargeable oxide traps with larger carrier exchange time constants than interface states and will prove that such traps have to be considered in order to explain the full differences in the $V_{\rm TH}$ shifts recorded at different gate biases. The position of the Fermi level at which we evaluate the threshold-voltage shift will turn out to be of fundamental importance, since it has a considerable impact on the extracted $V_{\rm TH}$ shifts and the recovery characteristics.

II. PMOS AND NMOS COMBINATION TECHNIQUE

As we drive a device from accumulation to inversion by a gate bias sweep, we cannot reasonably measure an appropriate exponential growth in the drain current before the density of minority carriers does not exceed the density of majority carriers within the channel. Consequently, when evaluating $V_{\rm TH}$ shifts from transfer curves, the energy range between the valenceband edge (E_V) and the intrinsic energy (E_i) is not accessible for the NMOS device. Similarly, the energy range between E_i and the conduction-band edge (E_C) remains hidden for the PMOS device. Only by combining the study of both devices can we maximize the accessible range of $V_{\rm TH}$ shift evaluation versus gate voltage: On a PMOS device, we have the possibility to scan roughly the lower half of the silicon band gap by varying the Fermi level between E_i and E_V . By using an identically processed but oppositely doped NMOS device, which has a completely different Fermi level position for gate biases around its threshold voltage, we can scan the upper half of the silicon

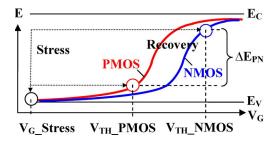


Fig. 1. Schematic illustration of the Fermi level position at the interface during stress and recovery for the PMOS and the NMOS devices. The values at stress level and at $V_{\rm TH}$ are highlighted by circles. During stress, the oxide field is identical for both types of devices. The energy interval $\Delta E_{\rm PN}$ represents the difference in Fermi level position between the NMOS and the PMOS devices during recovery at $V_{\rm TH}$.

band gap by varying the Fermi level between E_i and E_C (Fig. 1).

In order to generate identical stress levels by identical stress biases, we have fabricated our devices with a single-gate polyprocess (n⁺⁺ gate material for both NMOS and PMOS devices). The threshold voltage of the PMOS device is approximately -1.1 V, the threshold voltage of the NMOS device +1.1 V, respectively. We remark that despite of the different substrate and junction doping, during stress, the same gate bias results in the same oxide field and identical carrier concentrations under the gate oxide. This statement holds at least for thick oxide technologies stressed at gate biases far beyond threshold voltage. Due to the considerable overdrive $(V_{\rm GS}-V_{\rm TH})$ of the PMOS device during stress in inversion, the surface potential becomes equal to the drain and source voltage (0.0 V). On the other hand, the surface potential of the NMOS device stressed in accumulation becomes equal to the substrate voltage (0.0 V). We consider this a particular strength of our combined analysis because it allows us to monitor the impact of comparable degradation levels in different parts of the silicon band gap.

III. EXPERIMENTAL DETAILS

In order to minimize tunneling currents in thin oxides and the strongly process-dependent impact of nitridation, we use MOS transistors with identical thermally grown SiO_2 gate oxides $(t_{\mathrm{ox}}=30~\mathrm{nm})$. It has recently been shown that the basic mechanisms behind NBTI are essentially the same in thin and thick SiO_2 and SiON technologies [14]. In particular, it has been shown that in pure SiO_2 technologies, the oxide-charge contribution relative to the interface-state contribution is smaller [11], [13], [14].

In order to extract stress-induced $V_{\rm TH}$ shifts for different Fermi level positions, we measure the drain current (I_d) at different gate biases before stress and during recovery. In the analysis, we then convert changes in the drain current into equivalent threshold-voltage shifts [15]. In order to separate between interface states $(\Delta V_{\rm TH}^{\rm it})$ and oxide-trap $(\Delta V_{\rm TH}^{\rm ox})$ contributions to the total $V_{\rm TH}$ shift, we use the charge-pumping (CP) technique on reference devices [16]–[18]. To record the maximum CP current, we pulsed the gate symmetrically at a frequency of 500 kHz between deep accumulation and deep inversion, with rise and fall times of 375 ns, respectively. The frequency of 500 kHz implies that all traps with carrier

exchange time constants shorter than 1 μ s are, by definition, interface states. For the conversion of the maximum CP current into an effective interface-state-dependent $V_{\rm TH}$ shift, it is essentially important to consider that the scanned energy interval during CP does not generally coincide with the energy interval of electrical active (charged) interface traps at a certain constant gate bias. During CP, the scanned energy interval $(\Delta E_{\rm CP})$ only depends on the pulse slopes and on temperature [17]. In particular, it is independent of the charge state of the interface traps and contains information about parts of both the upper and lower half of the silicon band gap.

By contrast, if we measure the drain current at a constant gate bias, the Fermi level position and the nature of traps alone define the electrically active energy interval of charged interface states $(\Delta E_{\rm VG})$ that cause a contribution to the thresholdvoltage shift. In the analysis of our data, we make use of the fact that interface states are amphoteric, i.e., traps in the upper half of the silicon band gap are negatively charged upon capture of an electron and neutral when empty, and traps in the lower half of the band gap are positively charged upon hole capture or neutral when empty [19]-[21]. This fact is not in contradiction with the observation that NBTI on PMOS devices predominantly creates donorlike defects [22]–[24]. This is also observed in our data; however, it will be shown that these donorlike defects are predominantly oxide charges. The only small uncertainty related to the conversion of ΔICP_{max} to ΔV_{TH}^{it} is that we assume the amphoteric transition at midgap and a flat density of states (DOS) within the silicon band gap. We note that this assumption is probably not totally correct, since it has been reported for P_b centers that the $D_{it}(E)$ profile might have a minimum around midgap and two symmetric maxima at roughly $E_i \pm 280$ meV in the upper and lower half of the silicon band gap [25]. The extracted $D_{\rm it}$ from CP measurements is generally an average value over $\Delta E_{\rm CP}$. Using our pulse setup, this energy interval is estimated to be $E_i \pm 180 \text{ meV}$ at 125 °C and therefore does not cover the proposed maxima completely. Consequently, the flat DOS approach might slightly underestimate the interface-state-dependent $V_{\rm TH}$ shift $(\Delta V_{\rm TH}^{\rm it})$ at a constant gate voltage V_G (ΔE_{VG}), particularly for gate biases close to threshold voltage. However, this has no impact on our general conclusions.

All our devices discussed in the following are stressed identically for several thousand seconds at a temperature of 125 °C and a constant oxide field of 5.5 MV/cm. During recovery, we vary the gate bias and study the response of the interface states and the total threshold-voltage shift. For our investigations, it is necessary to compare different sets of identically processed devices with the same geometry on the same wafer by subjecting them serially to NBTI stress. Care has been taken in the designing of the layout that NMOS and PMOS devices are electrically isolated and can be stressed and characterized separately. In order to assure that these devices actually degrade identically, we select devices with comparable unstressed transfer and CP characteristics. This was done because preliminary measurements have indicated a good correlation between the degradation levels observed in equally fabricated devices provided they had identical virgin CP characteristics.

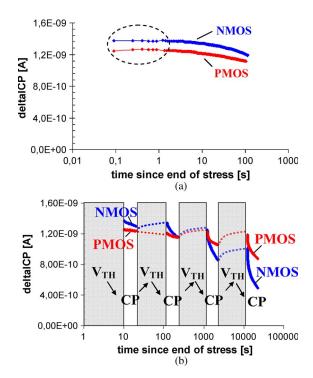


Fig. 2. (a) Recovery of ΔICP_{max} for a PMOS and an NMOS device after stressing both devices under the same bias and temperature conditions. Both devices show similar degradation of ICP_{max} , indicating a comparable D_{it} after stress. Due to continuous gate pulsing, moderate interface-state recovery can be observed. However, within the first second, the reduction of ICP_{max} is insignificant and independent of the bias conditions. (b) Recovery of ICP_{max} for the PMOS and the NMOS device after stressing both devices at the same bias and temperature. Within the gray-shaded areas, the gate voltage is constant at $V_G = -1.1$ V (PMOS), respectively, +1.1 V (NMOS). Within the white areas, we pulse the gate in order to record ICP_{max} . During the gate-pulsing periods, we observe interface-state recovery. During the constant-bias periods, some of the previously recovered interface states can be restored again (indicated by dotted lines).

IV. BIAS DEPENDENCE OF CP AND CONSTANT BIAS RECOVERY

The first aim of this paper is to differentiate between permanent and recoverable components in the $V_{\rm TH}$ shift. However, before we dare to separate them, we have to check how oxide traps and interface states react to the gate bias. This is essentially important since we must use different measurement setups to quantify both contributions independently. By gate pulsing between accumulation and inversion (CP), we may record interface-state degradation separately. On the other hand, both contributions superimpose in the total $V_{\rm TH}$ shift recorded at a constant gate bias. By combining both techniques, it is possible to separate interface from oxidecharge contributions by subtracting the $\Delta V_{\mathrm{TH}}^{\mathrm{it}}$ component from the total $\Delta V_{\rm TH}$ shift. The challenge here is to determine the contribution and recovery of charged interface states at a constant gate bias by means of a measurement technique that requires gate pulsing from inversion to accumulation (CP). This is not straightforward at all if gate pulsing turns out to have an impact on either oxide-trap or interface-state recovery. To check this, we perform two key experiments, cf. Figs. 2 and 3.

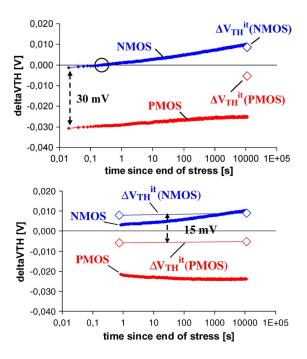


Fig. 3. (a) Constant bias recovery at $V_{\rm TH}$ for the PMOS and the NMOS device. After 10 000-s recovery, we have performed a short CP measurement and converted the ICP_{max} current into an appropriate interface-state-dependent $V_{\rm TH}$ shift $\Delta V_{\rm TH}^{\rm it}$ (open diamonds). (b) Constant bias recovery at $V_{\rm TH}$ for the PMOS and the NMOS devices, including two short CP cycles right after stress and after 10 000-s recovery. The corresponding $\Delta V_{\rm TH}^{\rm it}$ shifts are shown by open diamonds. The CP measurement has a dramatic impact on the recovery of $V_{\rm TH}$. There is no interface-state recovery as long as the gate bias is kept constant (no repetitive CP measurements).

A. Interface-State Recovery Versus Gate Bias

In Fig. 2(a), the recovery of ΔICP_{max} is shown for NMOS and PMOS devices stressed under the same bias and temperature conditions. During CP, we have continuously pulsed the gate of the NMOS device between $-1~{\rm V}~(V_{\rm GL})$ and $+2~{\rm V}$ $(V_{\rm GH})$, the PMOS device between -2 V $(V_{\rm GL})$ and +1 V $(V_{\rm GH})$, respectively. We remark that the NMOS device has a slightly higher (~10%) CP signal than the PMOS device after stress. Within 100 s of uninterrupted gate pulsing, the CP currents of both devices recover in a similar way by about 15%, indicating time-dependent repassivation of dangling bonds at the interface [3], [24], [26], [27]. However, within the first second of gate pulsing, there is no significant decrease in Δ ICP_{max}, neither for the PMOS nor for the NMOS device. This finding is in contradiction to results obtained by on-the-fly CP (OTFIT) measurements reported in [28] and subject of a separate study [29].

In the experiment shown in Fig. 2(b), we have replaced continuous gate pulsing (500 kHz) by *sequences* of constant gate biasing (-1.1-V PMOS and +1.1-V NMOS) followed by gate pulsing (-1/+2-V NMOS and -2/+1-V PMOS). Both sequences were repeated four times with increasing cycle durations (10/100/1000/10000 s).

During the gate-pulsing periods, we observe continuous interface-state recovery similar to those reported in [24] and [30]. We also note that the recovery of the NMOS device is faster than the recovery of the PMOS device which may be due to the different biasing conditions during gate pulsing

[31], [32]. Conversely, between two pulsing periods where the gate bias is constant, there is rather an increase than a decrease in $\Delta ICP_{\rm max}$ which indicates that previously recovered interface states may be restored again. This kind of "degradation" of the CP current is estimated by dotted lines in Fig. 2(b). Remarkably, the increase is similar for the PMOS and NMOS devices despite of the different biasing conditions ($\pm 1.1 \text{ V}$). This indicates that the gate voltage and the carrier concentration at the interface are not crucially determining this phenomenon. In contrast to Yang et al. [31], who concluded that the passivation of interface states is accelerated by an applied positive gate bias, and Ang [32], who suggested that the CP current recovery is suppressed by an applied positive gate bias, our experiments show unambiguously that it is the high-frequency pulsing, i.e., the CP measurement, itself which accelerates interface-state recovery and makes it look like the "classic" recovery curves shown in Fig. 2(a). The pulsing levels and the sign of the constant gate voltage between the CP samples seem to be just of second-order importance as long as the gate voltage does not subject the device to stress like in conventional OTF or OTFIT measurements. The results obtained by Yang et al. [31] and Ang [32] might also be misleading due to the long-lasting CP measurement periods (e.g., 120 s in [32]) and the large measurement delay used to determine the $\Delta N_{\rm it}$ contributions.

Although it is not subject of this paper to find a model for interface degradation and repassivation, we explicitly remark that these observations challenge $V_{\rm TH}$ recovery theories based on interface-state passivation [3], [24], [26], [27] and indicate that continuous gate pulsing is not appropriate to investigate $D_{\rm it}$ recovery during constant gate biasing.

In the following section, we will demonstrate that the interface-state contribution stays quasi-permanent, consistent with [2], if we keep our CP samples shorter than 1 s and abstain from long continuous gate pulsing periods. This finding indicates that the observed increase of the CP current during constant gate biasing [cf. Fig. 2(b)] is rather a restoration of previously recovered interface traps than a true degradation.

B. Threshold-Voltage Recovery Versus Gate Bias

In this section, we further investigate the influence of the gate bias on the total $V_{\rm TH}$ shift. In particular, we expect that CP has a significant impact on the carrier exchange properties and, therefore, on the $V_{\rm TH}$ and oxide-trap recovery, since gate pulsing during CP corresponds to periodic Fermi level switches across the entire silicon band gap.

Standard ($V_{\rm TH^-}$) recovery curves observed after NBTI stress of p- and n-channel MOSFETS are shown in Fig. 3(a). During recovery, the gate-source bias ($V_{\rm GS}$) was kept at a value close to the threshold voltage of the PMOS (-1.1 V) and NMOS device (+1.1 V), respectively. According to Fig. 3(a), the first important finding is that both p- and n-channel devices recover in a comparable way, yet starting from completely different $\Delta V_{\rm TH}$ values visible immediately after the end of stress. We further remark that the $V_{\rm TH}$ shift of the NMOS device is negative within the first 200-ms poststress. Only after 200 ms the shift becomes positive, indicating that the balance between

negative and positive contributions to $V_{\rm TH}$ shift has turned in favor of the negative charge.

In order to understand these results, we apply a short CP measurement right after the constant-bias recovery period and convert changes in ICP $_{\rm max}$ with respect to their virgin values into appropriate $D_{\rm it}$ -dependent $V_{\rm TH}$ shifts ($\Delta V_{\rm TH}^{\rm it}$). Assuming a flat DOS, this can be done by considering a weight factor [33] that accounts for the mismatch between the charged energy interval at a constant gate bias V_G ($\Delta E_{\rm VG}$) and the profiled energy range during CP ($\Delta E_{\rm CP}$)

$$\Delta V_{\rm TH}^{\rm it} = -\frac{\Delta Q_{\rm it}}{C_{\rm OX}} \approx -\frac{\Delta E_{\rm VG}}{\Delta E_{\rm CP}} \cdot \frac{\Delta ICP_{\rm max}}{f \cdot A \cdot C_{\rm OX}} \tag{1}$$

where $\Delta Q_{\rm it}$ is the increase in net interface charge at V_G due to stress, C_{OX} is the oxide capacitance, f is the measurement frequency, and A is the gate area. We remark that due to the amphoteric nature of interface traps, $\Delta Q_{\rm it}$ can be either positive or negative depending on the current Fermi level position during read-out at V_G . This implies that although the net charge build up after NBTI is positive in most reported cases (resulting in a negative threshold-voltage shift after NBTI), individual contributions to the $V_{\rm TH}$ shift can interfere either constructively or destructively with respect to their present charge state. To put it in a nutshell, in a PMOS device at $V_{\rm TH}$ (-1.1 V), both interface traps and oxide traps are positively charged, accumulating in a high negative threshold-voltage shift. Conversely, in an NMOS device at $V_{\rm TH}$ (+1.1 V), oxide traps are positively charged while interface states are negatively charged, resulting in either a net-positive or net-negative threshold-voltage shift depending on the contribution that dominates.

In Fig. 3(a), the $D_{\rm it}$ -equivalent $\Delta V_{\rm TH}$ shifts for both devices are shown by large open diamonds. Clearly, after a long recovery period of 10 000 s, the remaining threshold-voltage degradation is essentially due to interface states for the NMOS (\sim 8 mV), whereas for the PMOS, the major part (\sim 25 mV) of the visible shift must be attributed to positive oxide charge, only 5 mV being due to interface states.

In order to elaborate the role of interface states and gate pulsing on the $V_{\rm TH}$ recovery, we perform another key experiment using a second set of p- and n-channel devices, cf. Fig. 3(b). On this second set, we start immediately with CP after the end of stress, i.e., before recording the $V_{\rm TH}$ recovery curve at a constant gate bias. The CP periods were kept as short as possible (< 1 s) in order to prevent the previously discussed $D_{\rm it}$ recovery induced by gate pulsing. A comparison of the $V_{\rm TH}$ recovery characteristics shown in Fig. 3(a) and (b) shows that CP right after stress influences the subsequent $V_{\rm TH}$ shift considerably [31][34]. The $\Delta V_{\rm TH}$ levels of both NMOS and PMOS devices shift in the positive direction, indicating a reduction in net positive charge. We remark that the recovery curve of the PMOS device even changes its direction after the CP measurement.

When comparing the $\Delta V_{\mathrm{TH}}^{\mathrm{it}}$ contributions of Fig. 3(a) and (b), we realize that they are nearly identical right after stress and after 10 000-s recovery at constant gate bias. This indicates that interface states remain permanent if we keep our CP measurement cycles short.

V. SUMMARY OF EXPERIMENTAL FINDINGS

- 1) We must use separate sets of devices and short CP measurements for reliable $\Delta V_{\rm TH}$ and $\Delta V_{\rm TH}^{\rm it}$ investigations during constant bias recovery. These conclusions are drawn from the findings that interface state and $V_{\rm TH}$ recovery are influenced considerably by gate pulsing.
- 2) Interface states are permanent as long as the gate bias is constant between two *short* (\sim 0.5 s) CP measurements, cf. Fig. 3(b). Consequently, they cannot account for the threshold-voltage recovery observed at constant gate bias.
- 3) During $V_{\rm TH}$ recovery, PMOS and NMOS devices shift in the same direction, cf. Fig. 3(a). This indicates that, in both cases, net positive charge is lost. In addition, the recovery rate is larger in the NMOS device biased at a positive gate voltage. This is consistent with previous observations and with the idea of oxide-trap recovery via electron capture, the latter being much more efficient at positive gate biases.
- 4) The net charge at the end of a long recovery phase is positive for the PMOS and negative for the NMOS device. Indeed, the V_{TH} shift of the NMOS device after 10 000-s recovery at +1.1 V equals approximately its interface-state contribution. Conversely, 10 000-s recovery at -1.1 V (PMOS) leaves a considerable amount of additional positive oxide charge [2], [15]. The explanation consistent with the bulk of literature on oxide and interface traps is as follows: Oxide traps (E' centers) are donorlike while interface states are amphoteric [25]. Consequently, the positive V_{TH} shift of the NMOS device after recovery must originate from a dominance of the negative interface-state charge.
- 5) The observed difference of ${\sim}30$ mV between the total $V_{\rm TH}$ shift of the NMOS and PMOS device shown in Fig. 3(a) cannot be explained by interface charge alone, cf. Fig. 3(b). The $D_{\rm it}$ -related fraction (${\Delta}V_{\rm TH}^{\rm it}$) of ${\sim}15$ mV is only about half of the total $V_{\rm TH}$ shift difference.
- 6) After NBTI, significantly more (equivalent another ~15 mV) positive oxide charge is visible on PMOS than on NMOS devices, although the stress field and temperature have been identical for both devices.

VI. RECHARGEABLE OXIDE TRAPS AND THEIR ELECTRICAL RESPONSE

The key results of our experiments describe the nature and dynamics of NBTI stress-induced traps as follows: Interface states are amphoteric, permanent, and cannot account for the whole visible $V_{\rm TH}$ shift, consistent with [2]; additionally, there is stress-induced donorlike oxide charge which is recoverable and predominantly *positive* in PMOS and *neutral* in NMOS devices at their particular $V_{\rm TH}$. This is fully consistent with our results discussed earlier and the energy range accessible by our gate bias switches. We note that it has been suggested in literature [35], [36] that oxide traps can become negatively charged as well; however, since this is not usually observed for NBTI-induced oxide traps, our results suggest that the negative

energy level of NBTI-induced oxide traps is too high to be observed under conventional bias conditions.

In the following experiments, we will use the acquired knowledge on interface and oxide traps to differentiate correctly between various contributions to the total V_{TH} shift. By separating interface from oxide charges, we can study the exchange properties, quantity, and energetic location of NBTI-induced oxide traps by attributing their present charge state to the current position of the Fermi level (E_F) at the interface. We remark that these reference energies within the silicon band gap need not necessarily correspond to the trap energies of the oxide defects within the silicon dioxide band gap, since the gradient of the chemical potential within the insulator is not well defined.

In Fig. 4, we schematically show the band diagrams of a PMOS and an NMOS device during stress and recovery (with the gate bias close to the respective threshold voltage, cf. [37]). Positive and negative charges optionally contributing to the $V_{\rm TH}$ shift are indicated by hatched areas (the latter occurring only for the NMOS due to charging of acceptorlike interface traps). The defects we focus on in this section are switching oxide charges energetically located within the silicon band gap, as shown in the inset of Fig. 4(b). Since these kinds of traps are too slow to contribute to the conventional CP signal, their carrier-exchange time constants must be greater than 1 μ s (half of the gate-pulsing period during CP). According to the simulations of Heh et al. [38], a high/low time of 1 μ s would correspond to a spatial probing depth of approximately 5-6 Å in pure amorphous SiO2. This is an important conclusion consistent with some recent reports [10] but in contradiction to the assumption that charge trapping into the oxide occurs into preexisting traps [2].

Due to different Fermi level positions of the inverted PMOS and NMOS transistor, these traps will be filled (and hence neutral) in the NMOS device and empty (hence positive) in the PMOS device. The critical energy range between the Fermi level of the PMOS $(E_{\rm FP})$ and NMOS $(E_{\rm FN})$ transistors biased in (weak) inversion will be denoted as $\Delta E_{\rm PN}$. By changing the gate bias in a defined way during recovery, we can expand or shrink $\Delta E_{\rm PN}$, which enables us to energetically profile defects to a certain degree.

VII. TIME CONSTANTS AND SPECTROSCOPIC TRAP PROFILING

The occupancy of interface states can follow changes in E_F quasi-instantaneously (SRH theory), with the fastest response times being in the pico- and nanosecond range and the slowest response times for traps located at midgap in the millisecond regime [39], both outside our measurement window. Oxide traps, on the other hand, are expected to require a longer time interval to restore equilibrium with the silicon substrate, since a large quantum–mechanical barrier has to be overcome by elastic or inelastic tunneling. In addition, charging and discharging is likely to be accompanied by structural relaxation [40]–[42]. Generally, this means that if the Fermi level is changed abruptly (for example, when we switch the gate bias from stress level to $V_{\rm TH}$), a new equilibrium condition is generated. Assuming an elastic tunneling process, stress-induced oxide traps within

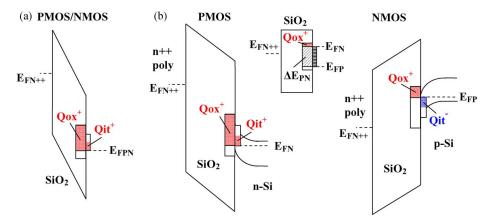


Fig. 4. Schematic band diagram of PMOS and NMOS devices during (a) NBTI stress and (b) recovery at $V_{\rm TH}$. During stress, both the oxide field and the Fermi level position are identical for NMOS and PMOS devices when stressed at a sufficiently large bias. During recovery, the net oxide and interface charge is different for the PMOS and the NMOS devices due to the different position of the Fermi level at $V_{\rm TH}$. The hatched area in the inset of (b) illustrates the difference $\Delta E_{\rm PN}$. Oxide traps and interface states located energetically within this interval will account for additional positive charge in the PMOS device.

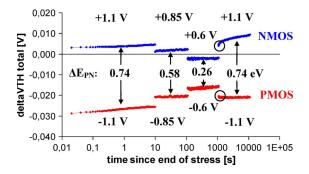


Fig. 5. Total $V_{\rm TH}$ shift during recovery at different gate biases for the PMOS and the NMOS devices. As the gate voltage drives the Fermi levels closer to midgap, the difference between the PMOS and the NMOS devices decreases. As we switch the bias from depletion back to inversion, transients emerge as tails in the recovery curve (circles).

the energy range $E_{\rm Fnew}-E_{\rm Fold}$ would tend to become or stay positively charged while traps below $E_{\rm Fnew}$ would rather be neutralized. Such a "switching" behavior of oxide traps was already reported for E_{γ} centers in MOS structures after irradiation [43], [44].

In Fig. 5, we generate such new equilibrium conditions by switching the gate bias during recovery of the NMOS and PMOS devices in a defined way between weak and deep inversion: after 10 s of recovery at ± 1.1 V (NMOS/PMOS), we change the gate bias for another 100 s toward weaker inversion (± 0.85 V). This causes the Fermi level to move immediately from the band edges toward midgap, cf. Fig. 1. A second gate bias switch from ± 0.85 to ± 0.6 V for another 1000 s moves the Fermi level even closer to midgap. The last recovery cycle was again at ± 1.1 V and lasted for $10\,000$ s.

According to the data shown in Fig. 5, a Fermi level variation toward midgap lowers the total $V_{\rm TH}$ shift for the NMOS and PMOS transistor as well. Consequently, the difference in net positive charge between NMOS and PMOS devices is reduced. Following our model, this is because the energy interval $\Delta E_{\rm PN} = E_{\rm FN} - E_{\rm FP}$ has reduced itself as we bias both devices closer to accumulation. After the last bias switch from ± 0.6 V (weak inversion) back to ± 1.1 V (deep inversion), we observe a huge step in the $V_{\rm TH}$ shift that indicates that

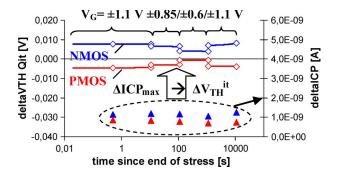


Fig. 6. Interface-state-dependent V_{TH} shift during recovery at different gate biases for the PMOS and the NMOS devices. Before every constant-bias-recovery phase, we have performed a short CP measurement for 0.5 s (full triangles). In the analysis, we have then converted $\Delta \mathrm{ICP_{max}}$ into $\Delta V_{\mathrm{TH}}^{\mathrm{it}}$ (open diamonds) and interpolated linearly between the measurement points. We also accounted for different active energy intervals of charged interface states (ΔE_{VG}) at different gate biases.

previously neutralized defects can be recharged again. This behavior is typically attributed to interface states and not fully appreciated for oxide defects. We remark furthermore that the charges in the $V_{\rm TH}$ shift do not react instantaneously to the bias switch. Instead, we observe moderately fast transients (tails in the recovery curve) after the bias switch, which indicate a relatively slow (i.e., time constants in the range of seconds) dynamic carrier exchange process which may be due to inelastic tunneling between oxide defects and the silicon substrate. SRH recombination of interface states or elastic tunneling would be definitely too fast to produce recovery tails which are within our measurement resolution.

Note that in Fig. 5, the observed total $V_{\rm TH}$ shifts are still the sum of interface and oxide-trap contributions. In order to decouple them, we have performed measurements on additional NMOS and PMOS devices where we introduced short CP measurements before and after every recovery cycle in a similar way, as shown in Fig. 3(b).

In Fig. 6, the extracted CP currents for the PMOS and the NMOS devices are illustrated by triangles. Consistent with the results of the previous section, the CP current does neither recover nor degrade significantly if the pulsing time is kept short and the gate bias is constant between the CP measurements.

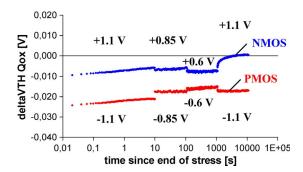


Fig. 7. Oxide-charge-dependent $V_{\rm TH}$ shift during recovery at different gate biases for the PMOS and the NMOS devices. The oxide-charge contribution results from subtracting the interface contribution by the total $V_{\rm TH}$ shift. After 10 000-s recovery, the oxide-charge contribution of the NMOS device is approximately zero.

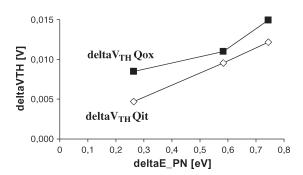


Fig. 8. Extracted (diamonds) $Q_{\rm it}$ and (squares) $Q_{\rm ox}$ energy profiles. The interface-state-dependent part increases linearly with $\Delta E_{\rm PN}$, since we have assumed a constant DOS within the silicon band gap. The oxide-charge-dependent contribution increases in a similar way, indicating also a nearly homogeneous density of oxide traps within $\Delta E_{\rm PN}$.

Following (1), we have converted the extracted $\Delta ICP_{\rm max}$ values into appropriate threshold-voltage shifts (symbolized by open diamonds, values in between are linear interpolations).

Having extracted the interface-state-dependent ΔV_{TH} component $(\Delta V_{\mathrm{TH}}^{\mathrm{it}})$, the oxide-trap-dependent contribution $(\Delta V_{\mathrm{TH}}^{\mathrm{ox}})$ can directly be obtained by subtracting $\Delta V_{\mathrm{TH}}^{\mathrm{it}}$ by the total V_{TH} shift shown in Fig. 5. $\Delta V_{\mathrm{TH}}^{\mathrm{ox}}$ is shown in Fig. 7 and can be attributed to switching oxide charge located energetically within ΔE_{PN} .

The medium concentration of interface states and oxide traps within $\Delta E_{\rm PN}$ can be studied when we average the differences between NMOS and PMOS in $\Delta V_{\rm TH}^{\rm it}$ (cf. Fig. 6) and $\Delta V_{\rm TH}^{\rm ox}$ (cf. Fig. 7) over the recovery time. Both contributions are shown separately as a function of $\Delta E_{\rm PN}$ in Fig. 8. As the interval $\Delta E_{\rm PN}$ shrinks, the oxide-charge contribution decreases quasi-linearly. This indicates a homogeneous distribution of oxide traps within the scanned energetic area, consistent with [45]. The interface-state contribution decreases linearly as well; however, this is a natural consequence of the assumed flat DOS.

VIII. CONCLUSION

We have analyzed interface states and carrier-exchange processes between the silicon substrate and the SiO₂ gate oxide. The impact of gate bias on the recovery of oxide and interface states was studied using NMOS and PMOS devices to make a larger portion of the silicon band gap visible. The occupancy

(charge state) of both types of traps was found to be bias dependent during recovery. However, as opposed to oxide charges, interface-state recovery was found to be negligible when the gate-pulsing time is kept short, and the gate bias is constant between the intermediate pulsing periods. Continuous gate pulsing during recovery showed a significant recovery acceleration of D_{it} . Concerning trapping and detrapping characteristics of defects, we have shown that oxide traps can exchange carriers with the silicon substrate, however, with larger time constants than expected for SRH-controlled interface states and elastic tunneling. The larger time constants originate most likely from an inelastic tunneling process between oxide defects and the silicon substrate and appear in our measurement signal as transients in the $V_{\rm TH}$ recovery curves when we change the gate biasing abruptly. By evaluating the difference in the V_{TH} shift between NMOS and PMOS devices at certain Fermi level positions (gate voltages) and afterward removing the interfacestate contribution from the total V_{TH} shift by an appropriate CP reference measurement, we can finally estimate an energetic profile of oxide defects within a defined energy interval $\Delta E_{\rm PN}$ around midgap. The profile suggests that oxide traps are continuously distributed with respect to the Fermi level position within the silicon substrate, indicating a homogeneous density of oxide traps located close to the gate-oxide-substrate interface.

REFERENCES

- D. K. Schroder, "Negative bias temperature instability: What do we understand?" Microelectron. Reliab., vol. 47, no. 6, pp. 841–852, Jun. 2007.
- [2] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanism to modeling," *Microelectron. Reliab.*, vol. 46, no. 1, pp. 1–23, Jan. 2006.
- [3] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 71–81, Jan. 2005.
- [4] J.-S. Lyu, K.-S. Nam, and C. Lee, "Determination of the interface traps density in metal oxide semiconductor field-effect transistor through subthreshold slope measurement," *Jpn. J. Appl. Phys.*, vol. 32, no. 10, pp. 4393–4397, Oct. 1993.
- [5] N. S. Saks and M. G. Ancona, "Generation of interface states by ionizing radiation at 80 K measured by charge pumping and subthreshold slope techniques," *IEEE Trans. Nucl. Sci.*, vol. NS-34, no. 6, pp. 1347–1354, Dec. 1987
- [6] S. C. Witczak, J. S. Suehle, and M. Gaitan, "An experimental comparison of measurement techniques to extract Si–SiO2, interface trap density," *Solid State Electron.*, vol. 35, no. 3, pp. 345–355, Mar. 1992.
- [7] C. Schlünder, M. Hoffmann, R.-P. Vollertsen, G. Schindler, W. Heinrigs, W. Gustin, and H. Reisinger, "A novel multi-point NBTI characterization methodology using smart intermediate stress (SIS)," in *Proc. IRPS*, 2008, pp. 79–86.
- [8] D. Brisbin and P. Chaparala, "The effect of the subthreshold slope degradation on NBTI device characterization," in *Proc. IIRW*, 2008, pp. 96–99.
- [9] J. F. Zhang, Z. Ji, M. H. Chang, B. Kaczer, and G. Groeseneken, "Real Vth instability of pMOSFETs under practical operation conditions," in *IEDM Tech. Dig.*, 2007, pp. 817–820.
- [10] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative bias temperature instability," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 22–34, Mar. 2008.
- [11] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "A two-stage model for negative bias temperature instability," in *Proc. IRPS*, 2009, pp. 33–44.
- [12] T. Grasser, W. Goes, and B. Kaczer, "An energy-level perspective of bias temperature instability," in *Proc. IRPS*, 2008, pp. 28–38.
- [13] T. Grasser, B. Kaczer, T. Aichinger, W. Goes, and M. Nelhiebel, "Defect creation stimulated by thermally activated hole trapping as the driving force behind negative bias temperature instability in SiO2, SiON, and high-k gate stacks," in *Proc. IIRW*, 2008, pp. 91–95.

- [14] H. Reisinger, R. Vollertsen, P.-J. Wagner, T. Huttner, A. Martin, S. Aresu, W. Gustin, T. Grasser, and C. Schlünder, "The effect of recovery on NBTI characterization of thick non-nitrided oxides," in *Proc. IIRW*, 2008, pp. 1–6.
- [15] B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, "Disorder-controlled-kinetics model for Negative Bias Temperature Instability and its experimental verification," in *Proc. IRPS*, 2005, pp. 381–387.
- [16] J. S. Brugler and P. G. A. Jespers, "Charge pumping in MOS devices," *IEEE Trans. Electron Devices*, vol. ED-16, no. 3, pp. 297–302, Mar. 1969.
- [17] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. De Keersmaecker, "A reliable approach to charge-pumping measurements in MOS transistors," *IEEE Trans. Electron Devices*, vol. ED-31, no. 1, pp. 42–53, Jan. 1984.
- [18] G. Van Den Bosch, G. Groeseneken, P. Heremens, and H. E. Maes, "Spectroscopic charge pumping: A new procedure for measuring interface trap distribution on MOS transistors," *IEEE Trans. Electron Devices*, vol. 38, no. 8, pp. 1820–1831, Aug. 1991.
- [19] J. Bardeen, "Surface states and rectification at a metal semiconductor contact," *Phys. Rev.*, vol. 71, no. 10, pp. 717–727, May 1947.
- [20] Z. Chen, B. B. Jie, and C.-T. Sah, "Microscopic theory of energy distribution of SiO2/Si interface traps: A survey of history and some new results," in *Proc. ICSICT*, 2006, pp. 1227–1233.
- [21] D. M. Fleetwood, M. R. Shaneyfelt, W. L. Warren, J. R. Schwank, T. L. Meisenheimer, and P. S. Winokur, "Border traps: Issues for MOS radiation response and long-term reliability," *Microelectron. Reliab.*, vol. 35, no. 3, pp. 403–428, Mar. 1995.
- [22] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability," in *Proc. IRPS*, 2002, pp. 248–254.
- [23] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, A. Varghese, R. B. Khamankar, and S. Krishnan, "Material dependence of hydrogen diffusion: Implications for NBTI degradation," in *IEDM Tech. Dig.*, 2005, pp. 688–691.
- [24] S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, M. A. Alam, "On the physical mechanism of NBTI in silicon oxynitride p-MOSFETs: Can differences in insulator processing conditions resolve the interface trap generation versus hole trapping controversy?" IRPS, 2007, pp. 1–9.
- [25] E. H. Poindexter, G. J. Gerardi, M.-E. Rueckel, P. J. Caplan, N. M. Johnson, and D. K. Biegelsen, "Electronic traps and Pb centers at the Si/SiO2 interface: Band-gap energy distribution," *J. Appl. Phys.*, vol. 56, no. 10, pp. 2844–2849, Nov. 1984.
- [26] M. A. Alam, "A critical examination of the mechanics of dynamic NBTI for PMOSFETs," in *IEDM Tech. Dig.*, 2003, pp. 14.4.1–14.4.4.
- [27] Y. Wang, "On the recovery of interface state in pMOSFETs subjected to NBTI and SHI stress," *Solid State Electron.*, vol. 52, no. 2, pp. 264–268, Feb. 2008
- [28] M.-F. Li, D. Huang, C. Shen, T. Yang, W. J. Liu, and Z. Liu, "Understand NBTI mechanism by developing novel measurement techniques," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 62–71, Mar. 2008.
- [29] P. Hehenberger, T. Aichinger, T. Grasser, W. Goes, O. Triebl, B. Kaczer, and M. Nelhiebel, "Do NBTI-induced interface states show fast recovery? A study using a corrected on-the-fly charge-pumping measurement technique," in *Proc. IRPS*, 2009, pp. 1033–1038.
- [30] S. Mahapatra, D. Saha, D. Varghese, and P. B. Kumar, "On the generation and recovery of interface traps in MOSFETs subjected to NBTI, FN, and HCI stress," *IEEE Trans. Electron Devices*, vol. 53, no. 7, pp. 1583–1592, Jul. 2006.
- [31] T. Yang, C. Shen, M. F. Li, C. H. Ang, C. X. Zhu, Y.-C. Yeo, G. Samudra, and D.-L. Kwong, "Interface trap passivation effect in NBTI measurement for p-MOSFET with SiON gate dielectric," *IEEE Electron Device Lett.*, vol. 26, no. 10, pp. 758–760, Oct. 2005.
- [32] D. S. Ang, "Observation of suppressed interface state relaxation under positive gate biasing of ultrathin oxynitride gate p-MOSFET subjected to negative-bias temperature stressing," *IEEE Electron Device Lett.*, vol. 27, no. 5, pp. 412–415, May 2006.
- [33] T. Aichinger, M. Nelhiebel, and T. Grasser, "Unambiguous identification of the NBTI recovery mechanism using ultra-fast temperature changes," in *Proc. IRPS*, 2009, pp. 2–7.
- [34] B. Kaczer, T. Grasser, R. Fernandez, and G. Goeseneken, "Toward understanding of the wide distribution of time scales in negative bias temperature instability," *ECS Trans.*, vol. 6, no. 3, pp. 265–281, 2007.
- [35] C. J. Nicklaw, Z.-Y. Lu, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "The structure, properties, and dynamics of oxygen vacancies in amorphous SiO2," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 6, pp. 2667–2673, Dec. 2002.

- [36] A. V. Kimmel, P. Sushko, A. Shluger, and G. Bersuker, "Positive and negative oxygen vacancies in amorphous silica," ECS Trans., vol. 19, no. 2, pp. 3–17, May 2009.
- [37] Y.-R. Cao, S.-G. Hu, X.-H. Ma, and Y. Hao, "Mechanism of NBTI under negative voltage stress," *Chin. Phys. Lett.*, vol. 25, no. 9, pp. 3393–3396, Sep. 2008.
- [38] D. Heh, G. A. Brown, A. Diebold, and J. B. Bernstein, "Spatial distributions of trapping centers in HfO2/SiO2 gate stack," *IEEE Trans. Electron Devices*, vol. 54, no. 6, pp. 1338–1345, Jun. 2007.
- [39] T. L. Tewksbury, "Relaxation effects in MOS devices due to tunnel exchange with near-interface oxide traps," Ph.D. dissertation, MIT Press, Cambridge, MA, 1992.
- [40] P. V. Sushko, S. Mukhopadhyay, A. S. Mysovsky, V. B. Sulimov, A. Taga, and A. L. Shluger, "Structure and properties of defects in amorphous silica: New insights from embedded cluster calculations," *J. Phys., Condens. Matter*, vol. 17, no. 21, pp. 2115–2140, Jun. 2005.
- [41] A. J. Lelis and T. R. Oldham, "Time dependence of switching oxide traps," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 1835–1843, Dec. 1994.
- [42] J. F. Conley, Jr., P. M. Lenahan, A. J. Lelis, and T. R. Oldham, "Electron spin resonance evidence for the structure of a switching oxide trap: Long term structural change at silicon dangling bond sites in SiO2," *Appl. Phys. Lett.*, vol. 67, no. 15, pp. 2179–2181, Oct. 1995.
- [43] V. S. Pershenkov, S. V. Cherepko, A. V. Sogoyan, V. V. Belyakov, V. N. Ulimov, V. V. Abramov, A. V. Shalnov, and V. I. Rusanovsky, "Proposed two-level acceptor-donor (AD) center and the nature of switching traps in irradiated MOS structures," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 6, pp. 2579–2586, Dec. 1996.
- [44] J. F. Conley, Jr. and P. M. Lenahan, "Electron spin resonance evidence that E' γ centers can behave as switching oxide traps," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 6, pp. 1744–1749, Dec. 1995.
- [45] T. Sakurai and T. Sugano, "Theory of continuously distributed trap states at the Si–SiO₂ interfaces," *J. Appl. Phys.*, vol. 52, no. 4, pp. 2889–2896, Apr. 1981.



Thomas Aichinger was born in Villach, Austria, in 1983. He received the Mag.Rer.Nat. degree in physics from Karl-Franzens-Universität Graz, Graz, Austria, in 2007. He is currently working toward the Ph.D. degree at the Kompetenzzentrum Automobilund Industrielektronik, Villach, in cooperation with "Christian Doppler Laboratory" group of the Institute for Microelectronics, Technical University of Vienna, Vienna.

During his Diploma degree thesis, concerning the implementation of charge pumping in the laboratory

software and hardware environment, he was integrated in the Quality and Reliability Group of Infineon Technologies Austria, Villach.

Mr. Aichinger was the recipient of the Best Paper Award at ESREF 2008.



Michael Nelhiebel was born in Vienna, Austria, in 1970. He received the M.S. degree in physics from Vienna University of Technology, Vienna, in 1995 and the Ph.D. degree in solid state physics from the Ecole Centrale Paris, Paris, France, in 1999. The focus of his Ph.D. degree thesis was on interferometric effects in the chemical microanalysis of crystals using electron energy loss spectrometry in the transmission electron microscope.

In 1999, he joined as a Reliability Engineer with Infineon Technologies Austria, Villach, where he

worked in the silicon wafer production and is currently a Senior Staff Engineer with the Quality Department, Infineon Automotive Business Division, responsible for wafer- and package-technology-related product reliability in the development phase. He has coordinated the qualification necessary for the delivery release of major automotive technology platforms and participates in research activities of Infineon Technologies Austria targeting technology reliability.



Tibor Grasser (M'05–SM'05) was born in Vienna, Austria, in 1970. He received the Dipl.Ing. degree in communications engineering, the Ph.D. degree in technical sciences, and the Venia Docendi in microelectronics from the Technische Universität Wien, Wien, Austria, in 1995, 1999, and 2002, respectively.

He was a Visiting Research Engineer at Hitachi Ltd., Tokyo, Japan, and in the Alpha Development Group, Compaq Computer Corporation, Shrewsbury, MA. Since 1997, he has been heading the Minimos-NT development group, working on the successor

of the highly successful MiniMOS program. Since 2003, he has been the Head of the Christian Doppler Laboratory for TCAD in Microelectronics, an industry-funded research group embedded in the Institute for Microelectronics,

Technical University of Vienna, Vienna, where he is currently also an Associate Professor. He is the coauthor or author of over 250 articles in scientific books, journals, and conference proceedings and is the Editor of a book on advanced device simulation. His current scientific interests include circuit and device simulation, device modeling, and reliability issues.

Dr. Grasser has been involved in the program committees of conferences such as the International Conference on Simulation of Semiconductor Devices and Processes (SISPAD), the International Workshop on Computational Electronics, the European Solid-State Device Research Conference, the IEEE International Reliability Physics Symposium (IRPS), the International Integrated Reliability Workshop, and the International Semiconductor Device Research Symposium. He was the recipient of the Best Paper Awards at the IRPS and the European Symposium on Reliability of Electron Devices, Failure Physics and Analysis. He was also the Chairman of SISPAD 2007.