**Advanced Characterization of Oxide Traps: The Dynamic Time-Dependent Defect Spectroscopy**


* Institute for Microelectronics, TU Wien, Austria  
° Infineon, Munich, Germany  
° imec, Leuven, Belgium

**Abstract**—An accurate understanding of oxide traps is essential for a number of reliability issues, including the bias temperature instability, hot carrier degradation, time-dependent dielectric breakdown, random telegraph and 1/f noise. Recent results have demonstrated that hole capture and emission into oxide traps in pMOS transistors are more complicated than the usually assumed Shockley-Read-Hall-like process. In particular, both charging and discharging proceed via a non-radiative multiphonon (NMP) mechanism involving metastable defect states. The existence of these metastable states can be demonstrated by extending the previously introduced time-dependent defect spectroscopy (TDDS) to a more general dynamic case by employing AC stresses and precisely timed discharge pulses during recovery. Application of AC stresses clearly reveals a frequency-dependence of the effective capture time, which confirms the existence of an intermediate metastable state. Application of pulses during recovery, on the other hand, allows extraction of the effective emission time also in depletion as well as in accumulation, thereby clearly revealing a metastable switching state. While all investigated traps show a frequency-dependent capture time constant, suggesting them to be of the same microscopic origin, we find two different kinds of emission behavior, namely fixed positive and switching traps. We finally demonstrate that our multi-state NMP model perfectly captures both cases.

**I. INTRODUCTION**

The effective capture and emission times, $\tau_c$ and $\tau_e$, of oxide defects have conventionally been studied via the analysis of random telegraph noise (RTN). This method analyzes the fluctuations around the equilibrium occupancy of the traps at a certain bias and works best when capture and emission times are of the same order of magnitude. Typically, $\tau_c$ and $\tau_e$ can be resolved over 3 orders of magnitude over a gate voltage range of 200mV (in thin oxides) [1–3]. As we have demonstrated [4], this limited measurement window does not even begin to reveal the rich features of $\tau_c$ and $\tau_e$ over the whole operating regime of the transistor.

By switching the gate voltage, dynamic transitions between the old and the new equilibrium occupancy can be studied. This fact has been exploited in the recently introduced time-dependent defect spectroscopy (TDDS) [4,5], which is an individual-trap variant of the deep-level transient spectroscopy (DLTS) [6,7]. Up to now we have predominantly used constant voltages during the charging and discharging phases [5,8–10]. These experiments have revealed the following schematic model for the dominant defects in the oxide of SiON pMOSFETs (see Fig. 1):

- Defects can be either fixed positive or switching traps, consistent with previous studies on large-area devices [11–13]. This implies and confirms the existence of a neutral metastable state (1’ in our notation).

**Fig. 1.** Left: The four states of oxide defects extracted from DC TDDS experiments [4]. Each defect has two stable states, 1 and 2, and possibly two metastable states 1’ and 2’. The metastable state 2’ seems to be always present, while the existence of the metastable state 1’ depends on whether the trap behaves like a fixed positive or a switching trap [11,12]. Right: An effective two-state approximation of the four-state defect using the first-passage times $\tau_{12}$ and $\tau_{21}$ [4,17].

- The capture times are strongly bias dependent, consistent with previous RTN studies [1,2]. However, for very large oxide fields, a saturation behavior is observed.
- The capture and emission times of individual defects appear de-correlated. This suggests that charging proceeds via a metastable state 2’.

In our initial study [4] we have introduced the metastable state 2’ and described the transitions $1 \rightarrow 2'$ using a conventional non-radiative multiphonon (NMP) model [1,14–16], while the transitions $2' \rightarrow 2$ were modeled using a simple relaxation over a thermal barrier. Using these conventional models, the metastable state 2’ introduces both a saturation behavior over a thermal barrier. Using these models, the metastable state 2’ introduces both a saturation in $\tau_{2'}$ and the existence of a non-puckered vs. puckered configuration. While both experimental observations provide strong evidence for state 2’, they do not rule out alternative explanations via an (hitherto unknown) charging/discharging mechanism different to the assumed non-radiative multiphonon mechanism.

We remark that the multistate model of Fig. 1 agrees in many ways with properties previously ascribed to $E'$ centers [11,12,18], in particular the metastable switching configuration 1’ and the existence of a non-pucker vs. puckered configuration (possibly 2’ and 2) [19]. However, density functional theory calculations in bulk SiO$_2$ [20,21] do not give thermodynamic energy levels and barriers fully consistent with our electrical data [21,22]. Whether this is due to the influence of additional strain or the amorphous nature at the interface, or other limiting assumptions in the model calculations is unclear at the moment and the identification of the microscopic nature of these oxide defects requires further work [23].
applied for the duration increase, visible by their clusters in the spectral map becoming fainter. As defects A1 and A2 have capture times shorter than $10\mu s$ [24, 25], hole capture is delayed under AC conditions compared to DC (Fig. 4). With increasing frequency, the capture time constants of defects A3 and A4 increase, visible by their clusters in the spectral map becoming fainter. As defects A1 and A2 have capture times shorter than $10\mu s$ at the two temperatures considered here, any potential frequency dependence could not be resolved.

In the following the existence of both metastable states 1' and 2' as well as the physical processes responsible for the various transitions between them will be studied using an extended TDDS setup, see Fig. 2. A considerable amount of revealing data can be produced by replacing the DC stress period of the conventional TDDS setup by an AC signal switching between $V_1$ and $V_2$, followed by a discharge period at $V_r$. This sequence is typically repeated 100 times to allow for a statistical analysis of the discrete emission events. Middle: In the dynamic AC TDDS measurement, defects are subjected to an AC signal switching between $V_1$ and $V_2$, followed by a discharge period at $V_r$. Right: In the dynamic AC TDDS, a pulse $V_p$ is applied for the duration $t_p$ between the charging and discharging biases. Depending on the value of the emission time constant at $V_r$, the occupancy can be significantly reduced compared to what is obtained after the pure DC pulse. In the simplest case, the occupancy follows $\exp(-t_p/\tau)$, which can be used to extract $\tau$ over a wide range of $V_p$.

II. THE AC TDDS

In a decisive experiment to confirm the existence of state 2', the DC stress phase of the TDDS setup was replaced by an AC stress, see Fig. 2 [25,26]. Although the two-state approximation is designed to have the same expectation values for the capture and emission times under DC conditions [17], it is fundamentally different under AC conditions because it has a frequency-independent effective capture time $\tau_e$ [26,30,31]. The full four-state model, on the other hand, where only the pathway $1' \rightleftharpoons 2'$ is relevant during stress $22'$, can result in a frequency-dependent effective $\tau_e$ [24]. This overall frequency-dependence can be weaker or stronger, depending on the values of the four time constants $\tau_{12'}, \tau_{21}, \tau_{22'}$, and $\tau_{22'}$. Since these time constants depend on bias and temperature, the
previously charged defect is neutralized (2 in [4]) moves below the Fermi-level of the substrate. Then, the frequency-dependence as a function of bias and temperature demonstrates that the defect model accurately captures this map decreases with increasing frequency. Furthermore, Fig. 4 shows how the occupancy of the traps in the spectral have confirmed that this is indeed the case [25]. For example, by the same parameter set. Dynamic AC TDDS experiments must be able to describe words, since no new parameters are introduced in the four-conditions as a function of temperature and bias. In other conditions as a function of temperature and bias. In other words, since no new parameters are introduced in the four-state model, both DC and AC conditions must be captured by the same parameter set. Dynamic AC TDDS experiments have confirmed that this is indeed the case [25]. For example, Fig. 3 shows how the occupancy of the traps in the spectral map decreases with increasing frequency. Furthermore, Fig. 4 demonstrates that the defect model accurately captures this frequency-dependence as a function of bias and temperature.

III. THE RECOVERY-PULSE TDDS

While charging follows the pathway 1 $\leftrightarrow$ 2 $\rightarrow$ 2, discharge may proceed either via 2’ or 1’. The latter route becomes relevant when the defect level $E_2'$ (the switching level, details in [4]) moves below the Fermi-level of the substrate. Then, the previously charged defect is neutralized (2 $\rightarrow$ 1’). The duration of this transition is determined by $\tau_{21'}$, which will have become smaller than the bias-independent $\tau_{22'}$ under these conditions.

Regarding the discharge of the defects it has been observed that some defects, the so-called switching traps, are sensitive to the gate bias [11, 12], while others are not. From a reliability standpoint, it has been demonstrated that the recoverable component of NBTI, which has been linked to such oxide traps [32–36] is highly sensitive to switches into the depletion regime [37–39]. However, the DC TDDS requires a certain current flow through the transistors and $\tau_e$ cannot be completely studied in depletion to better understand and link the behavior of individual traps to NBTI, see Fig. 5. By introducing a well controlled discharge pulse right after stress, we extend the TDDS to allow for the extraction of $\tau_e$ throughout the whole depletion and even into the accumulation regimes of the transistor.

The idea of this recovery-pulse TDDS is simple: suppose that we apply a certain stress/charging pulse $V_s$ of duration $t_s$ followed by a discharge period at $V_r$. Assuming for simplicity that in equilibrium at $V_r$ the defect becomes fully charged ($f_2 = 1$), while during relaxation at $V_r$ it becomes completely discharged ($f_2 = 0$), the expectation value of the occupancy is given by [17]

$$f_2 = (1 - \exp(-t_s/\tau_c)) \exp(-t_s/\tau_e). \quad (1)$$

State 2 is here the stable positive state of the multi-state defect model, see Fig. 1. If we now insert a pulse of duration $t_p$ at a voltage $V_p$, we have with $\tau_p = \tau_e(V_p)$ and for $|V_p| < |V_r|$

$$f_2 = (1 - \exp(-t_s/\tau_c)) \exp(-t_p/\tau_p) \exp(-t_s/\tau_e). \quad (2)$$

Impact of Pulse
Naturally, if \( t_p \gg \tau_p \), the defect will completely disappear from the map. On the other hand, for \( t_p \ll \tau_p \), the occupancy will not be affected. However, by choosing \( t_p \approx \tau_p \), the differences in \( f_2 \) between eq. (2) and eq. (1) allow us to extract \( \tau_p \). Since extraction of \( f_2 \) via the spectral map of the TDDS requires many repetitions of the basic charging/discharging sequence shown in Fig. 2 for accurate statistics and since \( \tau_p \) is not known \textit{a priori}, it appears more reliable to choose a few values of \( t_p \) around the expected \( \tau_p \) and fit the data to eq. (2), which is essentially the exp\((-t_p/\tau_p)\) term.

\section{A. Proof of Concept}

The basic features of the dynamic TDDS are demonstrated in Fig. 6. In our previous TDDS studies we had been using a single device for a duration of about three years [4, 5, 25], which unfortunately was destroyed by a handling accident unrelated to the actual experiments during the course of this study. In the following, defects belonging to this initial device have the prefix A (see Fig. 3 and Fig. 4), while the recovery-pulse TDDS experiments were performed on a new device, B. Based on the same setup as used in the DC TDDS of the leftmost panel in Fig. 5, we introduce a pulse from \( V_l \) down to \( V_p = 0 \) \( V \) for \( t_p = 100 \) \( \mu s \), 1ms, 10ms, and 100ms. Clearly, the occupancy of defect B3 is reduced with increasing \( t_p \). When the occupancies extracted from the spectral maps are plotted as a function of \( t_p \), good agreement with the theoretical \( \exp(-t_p/\tau_p) \) dependence is obtained, allowing for the extraction of \( \tau_p(0\,V) = 40\,\text{ms} \).

The recovery-pulse TDDS is now applied to study the bias dependence of \( \tau_p \) for the defects B1 and B3. Unfortunately, B2 disappeared after two weeks of measurements, similarly to A6 in our previous study [4], implying another possible defect transformation [13, 36, 40–42]. The dependence of the occupancy \( f_2 \) for a few selected values of \( V_p \) is shown in Fig. 7 and Fig. 8. As can be seen, \( \tau_p \) of defect B1 remains at a bias-independent value of about 0.4ms throughout depletion and into accumulation. On the other hand, \( \tau_p \) of B3 depends exponentially on the bias in depletion, continuing the trend already observable in Fig. 5.

A comparison of \( \tau_p \) obtained by the DC and recovery-pulse TDDS is given in Fig. 9. Both methods deliver the same result in the region where the DC TDDS can be applied, while the recovery-pulse TDDS works well into the depletion regime and even into accumulation. The exceptional resolution of the recovery-pulse TDDS is clearly seen, allowing us in this case to extract \( \tau_p \) over 6 orders of magnitude for defect B3. From a practical viewpoint, the DC TDDS has already been used to extract values of \( \tau_p \) of up to 1ks, giving the recovery-pulse TDDS an unparalleled resolution of 10 orders of magnitude. As such, the whole operating regime of the transistor can be covered, allowing us to extract values of \( \tau_p \) and \( \tau_e \) from accumulation into strong inversion.
Here it has to be noted that, roughly speaking, the TDDS gives best results for the shorter of the two time constants, as this time constant dominates the transition between the equilibrium states [17]. To be more precise on this somewhat subtle issue consider for example a two-state defect. The expectation value of the defect occupancy during a transition between states 1 and 2 is given by

$$f(t) = f^H + (f^L - f^H) e^{-t/\tau}$$

with

$$f^L = \frac{\tau^L_e}{\tau^L_e + \tau^L_c}, \quad f^H = \frac{\tau^H_e}{\tau^H_e + \tau^H_c},$$

and

$$\frac{1}{\tau} = \frac{1}{\tau^L_c} + \frac{1}{\tau^H_c},$$

where the superscripts L and H denote the values at the recovery and stress voltage, respectively. Usually, as also assumed in eq. (1), $f^L = 0, f^H = 1$, and $\tau = \tau^H_c$. For low stress voltages or high recovery voltages, however, this is no longer the case. Since the TDDS uses the occupancy of the defect during stress to calculate the capture time constant, it actually determines the decorrelation time $\tau$ rather than $\tau_c$. During emission, on the other hand, the statistics of the first emission event are recorded, which corresponds to $\tau_e$. Finally, the recovery-pulse TDDS also monitors the occupancy in the spectral map to calculate the emission time at a different bias, so it also determines $\tau$ rather than $\tau_e$. However, for typical bias conditions such as those shown in Fig. 9 we have $\tau \approx \tau_e$. As such, the TDDS is considerably simpler when one of the time constants dominates.

For the three-state defect (assuming the transition via 2' to be dominant), the TDDS extracts the first passage time for the transition 2 → 1 (the average emission time)

$$\tau_e = \frac{k_{2'2} + k_{2'1} + k_{2'1}}{k_{2'2}}$$

since it again analyzes the first step in the emission transient. Extraction of the capture time proceeds via the occupancy in the spectral map, which is given by the effective time constant of the system (the decorrelation time)

$$\tau = \frac{k_{1'2'2} + k_{1'2'1} + k_{1'2'1}}{k_{2'2} k_{2'1} + k_{1'2} (k_{2'2'} + k_{2'2})}$$

assuming that the pathway 2' is dominant. For sufficiently large bias, the time constant corresponds to the first passage time for the transition 1 → 2 (the average capture time)

$$\tau_c = \frac{k_{1'2'} + k_{1'2}}{k_{1'2'} k_{1'2}}.$$

Note that eq. (4) simplifies to eq. (3) for lower biases. Finally, since the recovery-pulse TDDS also monitors the occupancy, it in principle also extracts $\tau$, which, however, is equivalent to $\tau_e$ for low enough biases.

B. Theory

Using the recovery-pulse TDDS, the time constants for defect B1 and B3 have been extracted from accumulation to strong inversion at 125°C and 175°C and fitted by the multistate NMP model. The highly satisfactory results for defect B1 are shown in Fig. 10. Higher stress voltages than 2.5 V were omitted for this 2.2 nm oxide [44] in order to avoid oxide breakdown, but pose otherwise no limitation to the TDDS. As noted before, B1 has a bias-independent $\tau_c$, meaning that the neutral metastable switching state 1' is not accessible, as shown in the configuration coordinate and state diagrams of Fig. 11. As such, recovery proceeds from the stable positive state 2 to the neutral stable state 1 via the pathway 2 → 2' → 1. In other words, discharging follows the opposite route as charging.

On the other hand, in the switching trap B3, the metastable state 1' is accessible, leading to an exponential bias dependence of $\tau_c$ below the threshold voltage well into depletion, see Fig. 12, closely tracking the surface hole concentration. In this regime, discharge takes the alternative pathway 2 → 1' → 1, while (roughly) above the threshold voltage the conventional pathway 2 → 2' → 1 is followed, see Fig. 13. For low enough biases, recovery is then accelerated by the large surface
strongly supporting its validity. The model gives a perfect match to the experimental data, thereby over 8 orders of magnitude.

![Fig. 12. The capture and emission times of switching trap B3 at two temperatures.](image)

...the metastable state $1'$, but for the switching trap B3. Contrary to defect B1, the metastable state $1'$ seems to be always present during capture, two types of the charging sequence and introduces a strong frequency dependence. In particular, AC stresses and controlled pulses right after the charging reveal an intermediate metastable defect state, which is part of the multi-state NMP model accurately describes and explains the experimental behavior.

**V. ACKNOWLEDGMENTS**

The research leading to these results has received funding from the FWF project no. 23390-M24 and the European Community’s FP7 project no. 261868 (MORDRED)

**REFERENCES**


