Microscopic Oxide Defects Causing BTI, RTN, and SILC on High-K FinFETs

G. Rzepa*, M. Waltl*, W. Goes*, B. Kaczer†, and T. Grasser*
*Institute for Microelectronics, TU Wien, Vienna, Austria
†imec, Leuven, Belgium

Abstract—Reliability issues of MOSFETs such as bias temperature instability (BTI), random telegraph noise (RTN), and stress-induced leakage current (SILC), are linked to the trapping of charges in oxides. Even though the chemical structure of these oxide defects is still the subject of debate, detailed studies of these reliability phenomena have shown that their physical behavior can be successfully described by non-radiative multi-phonon (NMP) theory. In this work we characterize and study a pMOS high-k FinFET technology starting from degradation measurements up to the simulation of the energy barriers in the framework of NMP theory. This allows to investigate the aforementioned reliability issues all based on their common cause, the microscopic oxide defects.

INTRODUCTION

Recent MOS technologies have become highly susceptible to the impact of single charge trapping events in their oxides. On the one hand, this has resulted in investigations regarding variability aspects in general [1], on the other hand there is a growing need to understand the details behind the various facets of oxide defects such as RTN [2–6], BTI [7–14] and SILC [15]. The first step towards a physical description that can cover all these effects was the application of NMP theory [3, 16, 17]. Based on detailed studies of individual defects, the four-state NMP model has been suggested recently [18] and its validity was confirmed by successful application to charge trapping related issues [13, 14, 18, 19]. In this work we will characterize this model for a state-of-the-art technology, followed by a detailed investigation of RTN, BTI, and SILC at the single defect level.

EXPERIMENTAL

In this work we investigate a pMOS FinFET technology with a 1.6 nm thick HfO₂ layer (HK) and a 0.6 nm thick SiO₂ interface layer (IL). The fin length is 100 nm and the fin width and height are 10 and 30 nm, respectively. In order to characterize BTI for this technology, we have performed measure-stress-measure (MSM) cycles [20] on structures with 22 fins in parallel (1.54 µm effective width). As the gate voltage during recovery phases \( V_{th}^R \) was fixed at time-zero threshold voltage \( V_{th,0} \), the threshold voltage shift \( \Delta V_{th} \) was calculated based on the measured drain current using the initial transfer characteristics. This was done for several stress setups, each on “fresh” devices, with gate voltages during stress varying between \( V_{th}^I \) and \( -0.85 \) and \(-1.45 \) V and temperatures between \( T = 75 \) and \( 175 \) °C. The stress times were varied from 0.1 µs up to 10 ks and recovery was recorded with a short measurement delay of a few µs and recovery times up to 10 ks. The drain voltage during stress and recovery was \(-0.1 \) V. The measurement data are shown in Fig. 1 and Fig. 2 together with the simulation results, which will be discussed in the following.

Fig. 1: Comparison of simulated \( \Delta V_{th} \) (lines) to the experimental data (circles) during stress \( t_s \) and recovery \( t_r \). The stress times of the consecutive stress-recovery cycles of this MSM experiment were stepped between 1 ms for the first cycle (blue) and 1 ks for the last cycle (red), while the recovery time was 1 ks for each cycle in this exemplary setup with \( V_{th}^I = -1.25 \) V and \( T = 175 \) °C. The drain current was recorded during recovery only, hence there is no \( \Delta V_{th} \) data for stress.

Fig. 2: Degradation measured with MSM experiments (circles) and the corresponding simulation results (lines) for setups with different combinations of temperature and stress voltage. Each condition was measured on another fresh device. Only the data of the last stress \( t_s \) and recovery \( t_r \) cycle is shown.

MODELING

BTI-related degradation is modeled within the framework of the NMP theory with microscopic oxide and interface defects. The various transitions between the four states of the oxide defects are described using the energy potential surfaces of the defects along selected configuration coordinates (CC) [18]. In addition to oxide defects, the generation and the recovery of interface defects are described by a double-well model [21] while their charge is determined using an amphoteric Shockley-Read-Hall model [22]. The parameters of the microscopic defects were sampled from distributions which were calibrated to measurement data [19, 23]. Simulations of \( \Delta V_{th} \) based on these defects match the measurement data for various stress setups (Fig. 1 and Fig. 2), hence a representative set of microscopic defects has been obtained.
**Discussion**

In the following, reliability mechanisms will be discussed with the help of a single four-state NMP defect, denoted as “S”. This defect is shown as a star in the scatter plots. Fig. 3 shows a randomly selected subset of interface and oxide defects plotted according to their capture ($\tau_c$) and emission time constants ($\tau_e$) while their color indicates their contribution to $\Delta V_{th}$. These capture/emission time (CET) map plots refer to DC stress with stress voltage $V_G^H$ and recovery voltage $V_G^C$ [24, 25]. For such stress conditions, the active regions [18] can be specified in the band diagram (Fig. 4). Only traps within these regions have a trap level “passing” the Fermi level of the gate or the channel for the given stress voltages, hence can capture or emit a hole from the gate or the channel, respectively, and thus contribute to a change in the measured $V_{th}$.

While an increase in the stress voltage increases the defect’s probability to capture a hole from the channel, it decreases the probability to capture a hole from the gate. These interactions always appear concurrently and result in effective capture and emission times measurable with time-dependent defect spectroscopy (TDDS) [26]. Simulations allow for separate analysis of these effects as shown for defect “S” in Fig. 5. In accordance with the simulated occupancy of this defect (Fig. 6) $\tau_e$ is huge around the threshold voltage ($V_G = V_{th,0} = -0.35$ V) and, therefore, the defect is neutral and inactive while for moderate stress voltages ($V_G = -1.0$ V) $\tau_e$ is of the same order of magnitude as $\tau_c$ and typical RTN steps can be observed. For high stress ($V_G = -1.45$ V) capture times via the channel ($\tau_c,C$) and emission times via the gate ($\tau_e,G$) decrease and the defect now contributes to SILC. These three stress conditions are further analyzed in Fig. 7 where each row shows the same plots only for different $V_G$. In the 1st column the shift of the trap levels is depicted in the band diagram. The electric field is smaller in the HK layer, hence the shift of the trap levels is less position dependent compared to the IL. This is in agreement with the density of states (DOS) of oxide defects shown in the 2nd column. The occupancy probability $f$, which is indicated by the color of the circles in the 1st column, depends on the Fermi levels of both, the channel and the gate, as well as the energy barriers and the tunneling coefficients between the respective carrier reservoirs. The energy levels and barriers of defect “S” are investigated in the CC diagrams for gate interaction (3rd column) and channel interaction (4th column) which helps to explain the bias dependent contribution of this defect to BTI, RTN, and SILC in agreement with the observations made in Fig. 5. The fact that $\tau_e$ of an oxide defect can be determined by a channel interaction while $\tau_c$ is determined by a gate interaction and vice versa is not only a necessary finding to understand SILC [27], but it also explains the broad distribution of couplings of time constants found by RTN based measurements [28]. The coupling factors of the simulated oxide defects are shown in Fig. 8.

Fig. 5: Simulated bias dependence of $\tau_e$ (red) and $\tau_c$ (blue) of defect “S” for gate only (diamonds), channel only (circles), and the combination of both interactions (lines). A discharge of this defect can be triggered by applying gate voltages around 0 V. In this regime the emission into the channel can become very fast which is why this type of defects is termed “switching oxide trap” [18]. Note that the combined interactions yield larger emission times than sole gate interaction. This is because of the interplay of the transition rates via the metastable states.

Fig. 6: Band diagram of the high-k pMOS FinFET. Active regions are depicted with respect to $V_G^C = -0.35$ V for gate (MG) interaction (blue area) and channel (SI) interaction (red area). The Fermi levels are indicated by dashed black lines.

Fig. 3: Capture time constants at $V_G^H$ and emission time constants at $V_G^C$ of a randomly selected subset of simulated oxide and interface defects. The color of the defects represent their impact on $\Delta V_{th}$ for the given setup. In accordance with the measurements, a broad distribution of time constants can be observed. Defect “S” is depicted by a star.

Fig. 4: Band diagram of the high-k pMOS FinFET. Active regions are depicted for gate only (diamonds), channel only (circles), and the combination of both interactions (lines). A discharge of this defect can be triggered by applying gate voltages around 0 V. In this regime the emission into the channel can become very fast which is why this type of defects is termed “switching oxide trap” [18]. Note that the combined interactions yield larger emission times than sole gate interaction. This is because of the interplay of the transition rates via the metastable states.

Compared to measurement data of typical high-k devices, the technology investigated in this work shows a significant fraction of defects with large emission time constants. This pronounced contribution of “permanent” defects could be due to the higher interface trap density at the side walls of the fin structures [29], but the clarification of this issue will require additional investigations.
Fig. 6: Transient simulation of the occupancy of the neutral state “1” \(f_1\) \cite{18} of defect “S”, which strongly depends on \(V_G\). For \(V_G = V_{th} = -0.35\, V\) (top) the capture time constant is huge, hence the defect is basically not active. With increasing stress voltage the frequency of capture events via the channel increase significantly. Around \(V_G = -1.0\, V\) (middle) the capture and emission time constants are on the same order of magnitude which results in typical RTN traces. For a further rise of the stress voltage this trend continues. As the frequency of emission events into the gate increase too, this defect contributes to SILC. Additionally, at \(V_G = -1.45\, V\) (bottom) holes capture much faster than they emit, so the average of \(f_1\) is close to zero. Therefore, the defect is positively charged most of the time and will cause BTI.

![Diagram](image1)

Fig. 7: Each row shows the same plots only for different stress voltage: \(1^{\text{st}}\) row: \(V_G = -0.35\, V\), \(2^{\text{nd}}\) row: \(V_G = -1.0\, V\), \(3^{\text{rd}}\) row: \(V_G = -1.45\, V\). \(1^{\text{st}}\) column: Band diagrams with oxide defects (circles, defect “S”: star), their color indicates their occupancy probability at equilibrium. For \(V_G = -0.35\, V\) most defects are below the Fermi levels of both, the gate and the channel and are thus neutral. With increasing stress voltage the trap levels of more and more defects are moved above the Fermi level of the channel, thus can capture holes from the channel. Still, the trap levels are below the Fermi level of the gate to which they can emit holes and, therefore, the occupancy probability results from the balance of these carrier exchanges. \(2^{\text{nd}}\) column: The DOS of defects in the interface layer (green) and in the high-k (blue). The trap levels of defects in layers with higher permittivity are less position dependent. Still, they see a larger shift with respect to the band edges of the channel while their distribution stays narrow compared to the trap levels of the defects in the IL. \(3^{\text{rd}}\) column: CC diagrams for gate interaction with defect “S”. \(4^{\text{th}}\) column: CC diagrams for channel interaction with defect “S”. In accordance with the pronounced decrease of \(\tau_e\) around \(V_G = 0\, V\) (compare Fig. 5) this defect acts as a switching oxide trap and favors emission via the smaller barrier to 1’ for these gate voltages. The green arrows highlight hole capture and emission between defect “S” and the gate (solid arrows) or the channel (dashed arrows) while thicker arrows represent larger rates, hence shorter time constants. For \(V_G = -0.35\, V\) (\(1^{\text{st}}\) row) holes are emitted to both, the gate and the channel but capture events are very unlikely and the defect is thus inactive (compare Fig. 6). An increase of \(V_G\) (\(2^{\text{nd}}\) row) shifts the trap level of defect “S” above the Fermi level of the channel and, therefore, the defect tends to capture holes from the channel. At the same time, holes are emitted to the gate on similar time scales which causes RTN. A further increase of \(V_G\) (\(3^{\text{rd}}\) row) increases both, \(\tau_c\) and \(\tau_e\), hence defect “S” contributes to SILC. As this defect now captures much faster than it emits, its occupancy probability is close to 1 and it will give rise to BTI.
Based on degradation measurements we have successfully characterized a pMOS high-k FinFET technology. Close analysis of the extracted microscopic oxide defects, described by the four-state NMP model, has revealed the details of the mechanisms which let them cause BTI, RTN, and SILC. The distribution of trap levels was investigated and the interplay of channel carrier exchange with the gate and with the channel was studied in detail as it explains the variety of experimental observations related to charge trapping of oxides.

Acknowledgements
This work has received funding from the the European Community’s FP7 n°619234 (MoRV) and the Intel Sponsored Research Project n°2013111914.

References