Analysis of the Threshold Voltage Turn-Around Effect in High-Voltage n-MOSFETs Due to Hot-Carrier Stress

I. Starkov*, H. Enichlmair*, S. Tyaginov*, and T. Grasser*
*Institute for Microelectronics, Vienna University of Technology, Gußhausstraße 27-29, A-1040 Wien, Austria
*Process Development and Implementation Department, Austriamicrosystems AG, Tobelbader Straße 30, A-8141 Unterpremstaetten, Austria
Email: starkov@iue.tuwien.ac.at

Abstract—The turn-around effect of the threshold voltage shift during hot-carrier stress has been investigated. Such a phenomenon is explained by the interplay between interface states and oxide traps, i.e. by the compensation of the rapidly created oxide charges by the more slowly created interface states. To prove this idea, a refined extraction scheme for the defect distribution from charge-pumping measurements has been employed. The obtained results are in a good agreement with the findings of our physics-based model of hot-carrier degradation. This approach considers not only damage produced by channel electrons but also by secondary generated channel holes. Although the contribution of the holes to the total defect creation is smaller compared to that of electrons, their impact on the threshold voltage shift is comparable with the electronic contribution. The reason behind this trend is that hole-induced traps are shifted towards the source, thereby more severely affecting the device behavior.

Index Terms—threshold voltage, turn-around effect, hot-carrier degradation, worst-case conditions, Monte-Carlo, interface states, MOSFET, TCAD

I. INTRODUCTION

It is commonly accepted that processes related to trapping/detrapping in the oxide bulk play a crucial role in the bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since BTI and hot-carrier degradation (HCD) are believed to have a similar bias temperature instability (BTI) [1, 2]. Since B...
Fig. 1. The topology of the n-MOSFET with the net doping profile highlighted. Inset: the dependence of the oxide thickness vs. the lateral coordinate.

when $V_{gs} = (0.4 - 0.5)V_{ds}$ [5] and, hence, it is assumed that the most severe degradation corresponds to these stress conditions. This idea is supported by Fig. 2 demonstrating both experimental and simulated (with our HCD model [10, 14]) dependences of worst case conditions (WCC) with respect to $V_{gs}$ and $V_{ds}$. As we have discussed in [9, 15], for long-channel and/or high-voltage device the values of the acceleration integrals $I^{(MC)}$ are rather high, thereby the dose of damage provided by the MC-component is homogeneously distributed over the lateral coordinate. In other words, the density of interface states generated by “colder” carriers does not vary with the lateral coordinate. Hence, the strong localization of HCD is related to the single-carrier process and the position of the $N_{it}$ peak is defined by the acceleration integral maximum. Very often other criteria are being used in the literature and among them are the electric field, carrier averaged energy, dynamic energy, and the most prolonged tails of the carrier energy distribution function. However, as we showed in [10], the $N_{it}$ peak only coincides with the AI peak and is shifted with respect to maxima of other quantities.

We have intentionally chosen stress voltages close to the WCC (see checkpoints in Fig. 2) in order to maximize the contribution of all the degradation effects. The threshold voltage was monitored up to $10^5$s. Fig. 3 demonstrates that first $V_{th}$ decreases due to hole trapping in the oxide bulk while after 10ks $V_{th}$ increases due to trapping of electrons by interface traps. To check these speculations, the CP technique with varying amplitude of the gate pulse [16, 17] has been employed to investigate the $N_{it} (x)$ and $N_{ot} (x)$ distributions.

In our work we use a standard experimental scheme where the gate of the transistor is connected to a pulse generator and a small constant reverse bias is applied to the source and drain. We measured the substrate current at a gate pulse frequency of 100kHz. We use $V_{gl} = 5V$ and increase $V_{gh}$ from -4 to 2V (varying high level technique) and $V_{gh} = 3V$ with decreasing $V_{gh}$ from 2 to -5V (varying base level technique) in 0.02V increments, where $V_{gl}$ and $V_{gh}$ are the base and the high levels of the gate pulse, respectively. Such a small voltage step is required in order to obtain sufficient spatial resolution.

III. INTERFACE AND OXIDE TRAP PROFILES EVOLUTION

A. Local Oxide Capacitance

Most methods for the extraction of the lateral $N_{it} (x)$ and $N_{ot} (x)$ profiles from CP data employ the oxide capacitance $C_{ox}$ [16–18] as a crucial parameter. Usually $C_{ox}$ is treated as a constant parameter of the device [17]:

$$C_{ox} = \varepsilon_{ox}/t_{ox},$$

where $t_{ox}$ is the oxide thickness at the device center and $\varepsilon_{ox}$ the dielectric permittivity.

This approach is applicable only for transistors with a fixed oxide thickness and leads to erroneous results when $t_{ox}$ varies along the interface. In fact, in such an approach the capacitor is considered ideal or, in other words, the electric field is assumed uniform. An extension of the technique to non-uniform $t_{ox} (x)$ has been suggested in [16]. In practice, however, a substantial distortion occurs near the source/drain ends of the gate. Moreover, even in the case of a uniform thickness distribution such a scheme leads to spurious results due to fringing capacitance. The aforementioned simplification does not strongly affect the transistor characteristics. However, the electric field non-uniformity is of particular importance for the extraction of the interface state density profile after HC stress because $N_{it} (x)$ peaks near the drain end of the gate where the capacitor non-ideality is most pronounced [4]. Additionally, Lee et al. [18] have already demonstrated that...
Fig. 3. The shift of \( V_{th} \) as a function of stress time at various voltages. Inset: \( \Delta V_{th} \) measured for devices of the same architecture but different only in channel lengths stressed at \( V_{gs} = 6.75 \text{V} \) and \( V_{ds} = 2.0 \text{V} \).

Fig. 4. The conformal transformation used to solve the gate/drain fringing problem. The complicated case of the corner gate form (red dashed line) was reduced just to a ray [19].

The coordinate dependence of the capacitance \( C_{ox}(x) \) due to the fringing effect is essential.

The conformal-mapping method is most helpful for fringing electric fields in simple two-dimensional boundary problems (which is just our case) by transforming the boundary to a soluble form [20]. A series of simulations (described below) allow us to conclude that for the \( C_{ox}(x) \) consideration a simplified structure can be used. Namely, the edge of the gate contact can be interpreted geometrically as a ray instead of more complicated corner variant. The problem with the coordinate system is shown in Fig. 4. The \( z \) plane is mapped into the \( w \) plane with the functional relationship between \( z \) and \( w \) described by

\[
z = \frac{t_{ox}}{\pi} (\omega + \exp(\omega))
\]

The suggested conformal transformation is shown in Fig. 4 and reduces the original problem to the Laplace problem between two parallel infinitely long metallic plates at different potentials [20]. The local oxide capacitance is then defined as the ratio between the interface charge density and the interface potential and can be written as [19]

\[
x = \frac{t_{ox}(x)}{\pi} \left( \varphi + \exp(\varphi) \right)
\]

(4)

For \( x \to -\infty \) (or \( \varphi \to -\infty \)) equation (4) asymptotically turns into the well known expression for the parallel-plate capacitance (2).

For evaluating the local oxide capacitance using device modeling we employ the method developed by Lee et al. [18]. To determine \( C_{ox}(x) \) by means of simulations, a careful calculation of the local flatband \( V_{fb}(x) \) and threshold \( V_{th}(x) \) voltage distributions as crucial quantities is performed. Due to the symmetry of the source and drain for the fresh device, we present results only for the drain half of the device [18]. For an unstressed transistor \( V_{th}(x) \) and \( V_{fb}(x) \) can be obtained with a widely-adopted routine [7, 21] employing our device simulator MiniMOS-NT [22]. The carrier concentrations used as a criterion for local threshold and flatband voltages are \( 2 \times 10^{13} \text{cm}^{-3} \) and \( 8 \times 10^{14} \text{cm}^{-3} \), respectively. The validity of the obtained curves has been verified using the Monte-Carlo device simulator MONJU [23, 24].

For any point at the device interface the post-stress local threshold \( V_{th,x}(x) \) and local flatband \( V_{fb,x}(x) \) values deviate from the pre-stressed by [16, 25]:

\[
V_{th,x}(x) - V_{th}(x) = \frac{\varrho \Delta N_{it}(x)}{C_{ox}(x)} + \frac{q \Delta N_{ox}(x)}{2C_{ox}(x)}
\]

(5)

\[
V_{fb,x}(x) - V_{fb}(x) = \frac{\varrho \Delta N_{it}(x)}{C_{ox}(x)} - \frac{q \Delta N_{it}(x)}{2C_{ox}(x)}
\]

where \( \Delta N_{it}(x) \) and \( \Delta N_{ox}(x) \) is the change between pre- and post-stress concentrations of interface traps and bulk oxide.
Fig. 6. The lateral distributions of \( N_{it}(x) \) and \( N_{ot}(x) \) extracted using different \( C_{ox}(x) \) approaches for \( 10^5 \) s at \( V_{th}=6.5 \text{V} \) and \( V_{gs}=2.6 \text{V} \). Peaks of \( N_{it}(x) \) correspond to the maxima of electron and hole acceleration integrals [8].

Charges. In other words, the presence of a uniform probe oxide charge \( Q_{ot, \text{uniform}} \) leads to a local threshold voltage shift \( \Delta V_{th, \text{uniform}}(x) \) [18]. The local oxide capacitance is thus found as

\[
C_{ox}(x) = -Q_{ot, \text{uniform}}/\Delta V_{th, \text{uniform}}(x).
\]

A typical example of \( \Delta V_{th, \text{uniform}}(x) \) induced by a uniform oxide charge density of \( 10^{12} \text{cm}^{-2} \) is shown in Fig. 5, inset.

A comparison of the simulations and developed analytical approach for the local capacitance profile shows good agreement, see Fig. 5. At the same time, the expression \( C_{ox} = \varepsilon_{ox}/t_{ox} \) – even when corrected for \( t_{ox} = t_{ox}(x) \) (Fig. 1, inset) – leads to substantially different results. The value of \( C_{ox}(x) \) has its maximum with \( C_{ox} = \varepsilon_{ox}/t_{ox} \) in the middle of the gate and decreases gradually toward a much lower value outside the gate edge due to the fringing effect. As expected, the most pronounced peculiarity is observed at the drain side of the gate where the abrupt change in the oxide thickness occurs, Fig. 1, inset. From a detailed analysis using the suggested conformal transformation we conclude that under the gate electrode the fringing effect can be neglected for distances larger than \( 1.14t_{ox}(x = l_g) \) from the gate edge (this is reflected in Fig. 5). Note that an abrupt reduction in \( C_{ox}(x) \) at \( x = 1.0 \mu\text{m} \) is unphysical and such an approach should not be used.

B. Extraction of \( N_{ot}(x) \) and \( N_{it}(x) \) Profiles

For extraction of the evolution of \( N_{it}(x) \) and \( N_{ot}(x) \) with the stress time we employed the analytical \( C_{ox}(x) \) distribution incorporated into the scheme described in [16] based on system (5). The results obtained under different assumptions on the \( C_{ox}(x) \) distribution are presented in Fig. 6. The extracted \( N_{it}(x) \) and \( N_{ot}(x) \) profiles were subjected to further validation as input parameters to simulate the transfer characteristics of the degraded device. Comparison of simulated and experimental curves once again confirms applicability of the developed model (see Fig. 7). The role of the fringing effect at large stress times is obvious.

IV. RESULTS AND DISCUSSION

The extracted \( N_{it}(x) \) profile features two peaks, see Fig. 6. This finding is in good agreement with the results of our HCD model [8] which shows that these peaks are related to the contributions induced by primary channel electrons and secondary generated holes (note that the latter one is shifted towards the source). Fig. 6 demonstrates that these peaks just correspond to the maxima of the electron and hole acceleration integrals. In our previous works [8, 10] we have already demonstrated that in long-channel devices the multiple-carrier-process of Si-H bond-breakage leads to \( N_{it} \) homogeneously distributed over \( x \) (see Figs. 8,9). At the same time, this is just the component which gives the main contribution to \( \Delta V_{th} \) because the device is most sensitive to \( N_{it} \) created closer to the channel middle where the SC-related contribution is negligible. Hence, notwithstanding the fact that the maximum value of \( N_{ot} \) is higher than the \( N_{th} \) peak value already from the beginning (Figs. 8,9), \( N_{th} \) starts to prevail over \( N_{ot} \) at longer times (\( \geq 10^5 \text{s} \) in this particular case). This is because at these times only the peak \( N_{ot} \) value is higher than the \( N_{it} \) maximum while within the device center (which defines the \( V_{th} \) behavior) \( N_{it} \) is higher. Strong localization of the \( N_{it} \) SC-component near the drain edge of the gate and, as a consequence, the secondary influence on \( \Delta V_{th} \) can be confirmed by saturation of \( V_{th} \) and a turn-around effect. This trend is depicted in Fig. 3, where the inset also shows for four MOSFETs of the same architecture with different gate lengths.

The main conclusion drawn from these results is that the hole contribution is considerably shifted towards the source. The single-electron component generates traps situated outside...
the channel which explains why the hole-induced traps have a stronger relative impact on the threshold voltage turn-around effect. The contribution of channel holes to the total defect concentration is much less than the corresponding fraction of the $V_{th}$ change. This trend becomes more pronounced for longer devices (Fig. 3, inset), see also [8]. Another tendency supporting these conclusions is that the $\Delta V_{th}$ turn-around effect is less pronounced at room temperature (Fig. 3 for the stress regime with $V_{ds} = 6.25\text{V}$ and $V_{gs} = 2.0\text{V}$). This is because trapping/detrapping in the oxide bulk is triggered by non-elastic trap-assisted tunneling which is accelerated at higher temperatures [26, 27].

In Fig. 10 the $N_{ot}(x)$ profiles are resolved for short stress times. The tendencies typical for $N_{it}$ are also pronounced in the case of bulk charge density: at short stress times $N_{ot}$ demonstrates two peaks related to channel electrons and to secondary generated holes. Initially, these peaks are differently located in space while after $\sim 5\text{s}$ they unite into a single common hump. This behavior can be explained assuming that trap-assisted tunneling responsible for charging bulk oxide traps occurs not necessarily perpendicular to the Si/SiO$_2$ interface. As a result, the bulk states situated (over the $x$ coordinate) between two initial peaks are also being filled resulting in a widened hump on the $N_{ot}$ profile.

It is worth emphasizing that the position of the $N_{it}(x)$ and $N_{ot}(x)$ peaks does not change with time (Figs. 8,9). This is in a good agreement with the model predictions where the position of the carrier acceleration integral maximum just coincides with the $N_{it}(x)$ peak position. To additionally check this we self-consistently calculated the acceleration integrals considering $N_{it}(x)$. Fig. 11 shows the AI vs. $x$ and confirms the previous tendency. Note finally that if during simulations the oxide thickness is assumed uniform ($t_{ox}(x) = t_{ox}$), both AI peaks are shifted (relatively to the AI peaks typical for the real device geometry), which implicitly confirms that the peculiarities of the device architecture have to be taken into account.

V. CONCLUSION

We have examined the turn-around effect of the threshold voltage shift under the worst-case hot-carrier stress conditions for long-channel 5V n-MOSFETs. For this purpose, we have improved the conventional defect density profile extraction techniques from charge-pumping measurements. We have demonstrated that by ignoring the spatial variation of $C_{ox}(x)$, a spurious result is produced, leading to an ambiguous picture of HCD with the model ignoring the $C_{ox}(x)$ distribution. The turn-around of the $V_{th}$ shift is explained to be due to the interplay between $N_{it}$ and $N_{ot}$ induced contributions. This idea is supported by our model of hot-carrier degradation which considers contribution of channel electrons as well as secondary holes generated by impact ionization. We have shown that in spite of a less pronounced hole contribution to the total interface states density, the device behavior is more sensitive to the hole-induced trap generation as compared to the electron contribution. This is related to the fact that the hole-induced portion of $N_{it}$ is situated closer to the channel, thereby affecting the $V_{th}$ in a stronger fashion.

REFERENCES


[22] Institute for Microelectronics, TU Wien, MiniMOS-NT Device and Circuit Simulator.


Fig. 10. Filling of the oxide traps revealed for small stress times at the stress regime with $V_{ds}=6.25V$ and $V_{gs}=2.6V$. The position of the $N_{x}(x)$ and $N_{ox}(x)$ peaks does not change with the stress time. The assumption of a constant oxide thickness results in a significant shift of the AI peak for holes and electrons ($A_{i_{x}}(h) \rightarrow A_{i_{x}}(n)$).

Fig. 11. The evolution of $A_{i_{x}}$ with the stress time for $V_{ds}=6.5V$ and $V_{gs}=2.6V$. The position of the $N_{x}(x)$ and $N_{ox}(x)$ peaks does not change with the stress time. The assumption of a constant oxide thickness results in a significant shift of the AI peak for holes and electrons ($A_{i_{x}}(h) \rightarrow A_{i_{x}}(n)$).