On the Differences between Ultra-fast NBTI Measurements and Reaction-Diffusion Theory

*A.E. Islam, S. Mahapatra*, S. Deora*, V. D. Maheta*, and M.A. Alam

*Email: aeilam@ieee.org, Phone: 765-532-6514, Fax: 765-494-6441*

Department of ECE, Purdue University, West Lafayette, IN 47906, USA; *CEN and Department of EE, IIT Bombay, India

**Abstract**

Reaction-Diffusion (R-D) theory, well-known to successfully explain most features of NBTI stress, is perceived to fail in explaining NBTI recovery. Several efforts have been made to understand differences between NBTI relaxation measured using ultra-fast methods and that predicted by R-D theory. Many alternative theories have also been proposed to explain ultra-fast NBTI relaxation, although their ability in predicting features of NBTI stress remains questionable. In this work, a hole-trap/interface-trap (N_{HT}/N_{IT}) separation framework (Fig. 1a) is used to demonstrate that N_{IT} relaxes slower compared to overall NBTI and this N_{IT} relaxation is consistent with R-D theory. The framework also explains, perhaps for the first time, the observed impacts of nitrogen, stress-time, temperature, frequency, duty cycle, etc. on NBTI degradation. In sum, together with N_{HT}, the R-D model governing N_{IT} is shown to explain NBTI stress and recovery features in nitrided gate oxide p-MOSFETS.

1. Introduction

Recent introduction of ultra-fast NBTI measurements [1-3] have inspired a number of studies [1,2,4-6] to understand the gap between NBTI theory and experiment. Among all NBTI theories, the R-D model explains many experimental signatures for p-MOSFETs having lightly nitrided oxides [7-9] including stress-phase time exponent, activation energy, field acceleration, frequency independence, etc. However, recent reports of ultra-fast NBTI relaxation that initiates at ~µs time-scale is inconsistent with N_{IT} dynamics, predicted by the R-D model (Fig. 1b), where H-H₂ conversion takes place at poly-Si/dielectric interface and long-term diffusion of H₂ occurs in poly-Si [10]. This inconsistency has raised questions regarding the general validity of the R-D theory [1,2,4,5]. The purpose of this work is to perform NBTI stress/relaxation experiments using an ultra-fast on-the-fly (UF-OTF) setup [3] on p-MOSFETs with nitrided dielectric, and to show that the theory-experiment gap in explaining ΔV_T relaxation can be bridged, if one accounts for the respective relaxation dynamics of its components: ΔV_{HT} (due to ΔN_{HT}) and ΔV_{IT} (due to ΔN_{IT}). A ΔV_{HT}/ΔV_{IT} separation scheme (Figs. 1-4; which was previously used in NBTI stress phase [6]) can explain the difference between start of overall NBTI relaxation (tREC,start ~µs) and N_{IT} relaxation (tNIT,start ~sec) (Figs. 5-7). This framework not only anticipates the duty-cycle/frequency dependencies of AC NBTI stress (Fig. 8), but also establishes the AC NBTI dependencies on nitrogen content (%N), temperature (T), and stress time (tSTS).

2. Non-universality of NBTI relaxation

Several studies [1,4,5,11] on NBTI have reported the universality of log-t relaxation, with tREC,start (start of ~5% NBTI relaxation) of ~ µs (Fig. 1b). However, our measurements on p-MOSFETs having a variety of nitrided dielectric (Fig. 1c and [12-14]) using UF-OTF demonstrate that fractional NBTI relaxation depends on %N of the dielectric, as well as on the difference between stress and recovery voltages (V_{STS}-V_{REC}). As shown in Fig. 1c, tREC,start is larger (~ms) for low %N and smaller (V_{STS}-V_{REC}), very clearly indicating the non-universal nature of NBTI recovery. This is due to the existence of different (N_{HT} and N_{IP}) species during stress [6, 14], with very different respective relaxation dynamics during recovery. Therefore, isolation of N_{HT} and N_{IT} is essential, before R-D theory (that governs N_{IT}...
Un-optimized post-nitridation anneal (PNA) can show the presence of fast Δ\(V_{HT}\) component, even in our lightly nitrided transistor. (d) Similar T independence was also reported for lightly doped nitrided transistors, using UFV [1].

4. Separation of \(N_{HT}\)/\(N_{IT}\) from \(\Delta V_T\)

Based on the observations of (i) T independent \(\Delta V_T\) at short \(t_{STS}\) and (ii) saturation of \(\Delta V_{HT}\) within \(t_{STS} \leq 1\)s for the thin oxides under study [9], a \(N_{HT}/N_{IT}\) separation algorithm is presented, which was used for explaining stress phase experiments in [6]. To estimate \(\Delta V_{IT}\) using this algorithm, a constant (saturated) \(\Delta V_{HT}\) is subtracted from \(\Delta V_T\) for \(t_{STS} > 1\)s thus can explain the observed T insensitivity in \(\Delta V_T\). At long \(t_{STS}\), \(\Delta V_{IT}\) starts to take over and results T dependent \(\Delta V_T\).

Moreover, the increase in \(\Delta V_{HT}\) with the increase in %N is also consistent with the decrease in power-law time exponent \((n\) for \(\Delta V_T \sim t^n\); see Fig. 3) for higher %N transistors [8,14]. Note that T-independence at short \(t_{STS}\) was observed in [1], using ultra-fast \(V_T\) (UFV) scheme, even for low %N (Fig. 2e). However, as shown in Fig. 2d, the quality of post-nitridation anneal (PNA) plays significant role for the presence of \(\Delta V_{HT}\), which can explain the observation of T independent \(\Delta V_T\) at short \(t_{STS}\) in [1], even for smaller %N.

4. Separation of \(N_{HT}/N_{IT}\) from \(\Delta V_T\)

Based on the observations of (i) T independent \(\Delta V_T\) at short \(t_{STS}\) and (ii) saturation of \(\Delta V_{HT}\) within \(t_{STS} \leq 1\)s for the thin oxides under study [9], a \(N_{HT}/N_{IT}\) separation algorithm is presented, which was used for explaining stress phase experiments in [6]. To estimate \(\Delta V_{IT}\) using this algorithm, a constant (saturated) \(\Delta V_{HT}\) is subtracted from \(\Delta V_T\) for \(t_{STS} > 1\)s thus can explain the observed T insensitivity in \(\Delta V_T\). At long \(t_{STS}\), \(\Delta V_{IT}\) starts to take over and results T dependent \(\Delta V_T\).

Moreover, the increase in \(\Delta V_{HT}\) with the increase in %N is also consistent with the decrease in power-law time exponent \((n\) for \(\Delta V_T \sim t^n\); see Fig. 3) for higher %N transistors [8,14]. Note that T-independence at short \(t_{STS}\) was observed in [1], using ultra-fast \(V_T\) (UFV) scheme, even for low %N (Fig. 2e). However, as shown in Fig. 2d, the quality of post-nitridation anneal (PNA) plays significant role for the presence of \(\Delta V_{HT}\), which can explain the observation of T independent \(\Delta V_T\) at short \(t_{STS}\) in [1], even for smaller %N.

4. Separation of \(N_{HT}/N_{IT}\) from \(\Delta V_T\)

Based on the observations of (i) T independent \(\Delta V_T\) at short \(t_{STS}\) and (ii) saturation of \(\Delta V_{HT}\) within \(t_{STS} \leq 1\)s for the thin oxides under study [9], a \(N_{HT}/N_{IT}\) separation algorithm is presented, which was used for explaining stress phase experiments in [6]. To estimate \(\Delta V_{IT}\) using this algorithm, a constant (saturated) \(\Delta V_{HT}\) is subtracted from \(\Delta V_T\) for \(t_{STS} > 1\)s thus can explain the observed T insensitivity in \(\Delta V_T\). At long \(t_{STS}\), \(\Delta V_{IT}\) starts to take over and results T dependent \(\Delta V_T\).

Moreover, the increase in \(\Delta V_{HT}\) with the increase in %N is also consistent with the decrease in power-law time exponent \((n\) for \(\Delta V_T \sim t^n\); see Fig. 3) for higher %N transistors [8,14]. Note that T-independence at short \(t_{STS}\) was observed in [1], using ultra-fast \(V_T\) (UFV) scheme, even for low %N (Fig. 2e). However, as shown in Fig. 2d, the quality of post-nitridation anneal (PNA) plays significant role for the presence of \(\Delta V_{HT}\), which can explain the observation of T independent \(\Delta V_T\) at short \(t_{STS}\) in [1], even for smaller %N.

4. Separation of \(N_{HT}/N_{IT}\) from \(\Delta V_T\)

Based on the observations of (i) T independent \(\Delta V_T\) at short \(t_{STS}\) and (ii) saturation of \(\Delta V_{HT}\) within \(t_{STS} \leq 1\)s for the thin oxides under study [9], a \(N_{HT}/N_{IT}\) separation algorithm is presented, which was used for explaining stress phase experiments in [6]. To estimate \(\Delta V_{IT}\) using this algorithm, a constant (saturated) \(\Delta V_{HT}\) is subtracted from \(\Delta V_T\) for \(t_{STS} > 1\)s thus can explain the observed T insensitivity in \(\Delta V_T\). At long \(t_{STS}\), \(\Delta V_{IT}\) starts to take over and results T dependent \(\Delta V_T\).

Moreover, the increase in \(\Delta V_{HT}\) with the increase in %N is also consistent with the decrease in power-law time exponent \((n\) for \(\Delta V_T \sim t^n\); see Fig. 3) for higher %N transistors [8,14]. Note that T-independence at short \(t_{STS}\) was observed in [1], using ultra-fast \(V_T\) (UFV) scheme, even for low %N (Fig. 2e). However, as shown in Fig. 2d, the quality of post-nitridation anneal (PNA) plays significant role for the presence of \(\Delta V_{HT}\), which can explain the observation of T independent \(\Delta V_T\) at short \(t_{STS}\) in [1], even for smaller %N.
in such a way that it provides time exponent n=1/6 for ΔVT (Fig. 4a). Since ΔVT has higher T activation compared to ΔVIT, extracted ΔVIT/ΔVT increases (i.e. ΔVIT/ΔVT decreases) with increase in T at fixed tSTS (Fig. 4b). The error bar in Fig. 4 results from noise in Id,lin measurement for UF-OTF [3], which causes ±0.005% error in ΔVT and a ±1mV error in estimated ΔVIT.

Next, we use the same NIT/NIT separation technique to explore the inconsistency between R-D theory and NBTI relaxation experiments (see Fig. 1b). Note that, UF-OTF relaxation measurements clearly indicate a weak T dependence up to tREC ~ ms (Fig. 5a and [13, 14]). Thus hole detrapping from pre-existing oxide traps is the predominant mechanism at tREC = ~ms for VREC = -1.3V and start of significant amount of T dependent NIT relaxation (tNIT,start) is also ~ ms. Moreover, UF-OTF at VREC = -1.8V shows insignificant relaxation up to ~ms (Fig. 6a), which leads to the conclusion that hole detrapping occurs predominantly for VREC = -1.8V, and that the amount of hole detrapping is similar from VREC = -1.3V to -1.6V (Fig. 6a). Thus trapping sites within the quasi-Fermi levels at VREC = -1.8V and -1.6V (shown schematically by the hatched region in Fig. 5c) will detrapping all the holes that were captured during stress. Considering such total hole detrapping at VREC = -1.3V, tNIT,start = ms is obtained at all VREC, which is comparable with the relaxation predicted by R-D theory (Fig. 6b). Remaining theory-experiment gap (x10 in tSTS, or 10% in NIT relaxation) is reduced (Fig. 6c), if it is realized that NBTI is also TDDB stress on p-MOSFETs, which causes hole trapping/detrapping from generated oxide traps NOT at VSTR at VREC (resultant ~3mV ΔVOF in ΔVf for Fig. 6c). Indeed, Figs. 7 confirms the existence of universal ΔVOF, even at NBTI stress conditions.

5. Implications for AC/DC Ratio

Our consideration of significant hole detrapping during recovery (Fig. 7) can explain the duty cycle and frequency dependence NBTI measurements (Fig. 8). Since hole trapping (Fig. 2) and detrapping (Fig. 5) occur at similar time-scales, total hole detrapping is expected for ≤50% duty cycle. In other words, in high %N transistors AC/DC ratio (when ΔVT is measured at the end of AC cycles) for ≤50% duty cycle will measure ΔVT(AC)/[ΔVIT(DC) + ΔVIT(DC)] and hence will always be less than the contribution from NIT’s component, AC/DC(NIT) = ΔVT(AC)/ΔVIT(DC), predicted by R-D theory [7] (Fig. 8a). Moreover, as ΔVIT decreases for smaller %N, AC/DC ratio in low %N transistors will only
have AC/DC(N_{HT}), and thus show remarkable consistency with R-D theory (Fig. 8a). Increase in %N decreases the AC/DC ratio at a particular duty cycle; however, the shape of AC/DC ratio vs. duty cycle plot at ≤50% duty cycle is mainly governed by AC/DC(N_{HT}) and is accurately anticipated by R-D theory, even up to ~80% duty cycle (Fig. 8a).

Since total hole detrapping happens at 50% duty cycle, frequency dependence of AC/DC ratio will follow that of AC/DC(N_{HT}). As a result, AC/DC ratio (similar to AC/DC(N_{HT})) is always frequency independent, irrespective of %N (Fig. 8b). However, due to the presence of ΔV_{HT} in AC/DC(N_{HT}) for higher %N. Thus the co-existence of N_{HT} and N_{IT} can explain both the duty cycle and frequency dependent NBTI experiments on nitrided transistors. Moreover, as the effect of ΔV_{HT} is higher at low T (Fig. 4) or at low t_{STS} (Fig. 2), AC/DC ratio for low T or low t_{STS} will always be less than AC/DC(N_{HT}) in high %N transistors.

In general, therefore, Fig. 8 can be used to roughly estimate ΔV_{HT}, which appears to be significant for the transistor reported in [4, 5, 11]. This is again consistent with n ~ 0.11 for t_{0} ~ 1 ms [5] in this transistor, which we expect for high %N or un-optimized PNA dielectrics (Fig. 3). Thus R-D theory – appropriately augmented with the effects of hole trapping/detrapping – can explain both the duty-cycle and the frequency dependent NBTI degradation in any nitrided oxide.

3. Conclusion

We use a consistent N_{HT}/N_{IT} decomposition framework to bridge the difference between R-D theory and ultra-fast NBTI measurements. Thus we demonstrate that the start of N_{IT} relaxation is slow and consistent with the predictions of R-D theory. Our work highlights the importance of careful interpretation of relaxation experiments considering the properties of gate oxide. Consequently, this analysis, for the first time, explains %N, T, t_{STS}, duty cycle dependencies of AC/DC ratio for nitrided transistors, which should be extremely important for NBTI AC analysis.

Acknowledgement: AMAT, TSMC, SRC, MCIT (GoI) for financial support and NCN for computational resources.

References


Fig. 8: (a) AC/DC ratio (when ΔV_{T} is measured at the end of AC cycles) vs. duty cycle plot for different nitrided transistors. The experiments show significant consistency with the prediction of R-D theory (solid line) for low %N, where ΔV_{T}~ ΔV_{HT}. Consideration of ΔV_{HT}, in addition to ΔV_{IT}, can also explain the cases for high %N transistors. The shape of AC/DC ratio vs. duty cycle is similar to R-D’s prediction for high %N transistors, even up to ~80% duty cycle. (b) Though AC/DC ratio for any %N is always frequency independent, there is significant %N dependency due to the presence of ΔV_{HT} in DC stress.