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

*A.E. Islam, S. Mahapatra[1], S. Deora[1], V. D. Maheta[1], and M.A. Alam

*Email: aeislam@ieee.org, Phone: 765-532-6514, Fax: 765-494-6441
Department of ECE, Purdue University, West Lafayette, IN 47906, USA; [1]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-H_2 conversion takes place at poly-Si/dielectric interface and long-term diffusion of H_2 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 $\Delta V_{HT}/\Delta V_{IT}$ separation scheme (Figs. 1-4; which was previously used in NBTI stress phase [6]) can explain the difference between start of overall NBTI relaxation ($t_{REC,start}$ ~μs) and N_{IT} relaxation ($t_{NIT,start}$ ~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 (t_{STS}).

2. Non-universality of NBTI relaxation

Several studies [1,4,5,11] on NBTI have reported the universality of *log-t relaxation*, with $t_{REC,start}$ (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, $t_{REC,start}$ 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_{IT}) species during stress [6, 14], with very different respective relaxation dynamics during recovery. Therefore, isolation of N_{IT} and N_{HT} is essential, before R-D theory (that governs N_{IT}

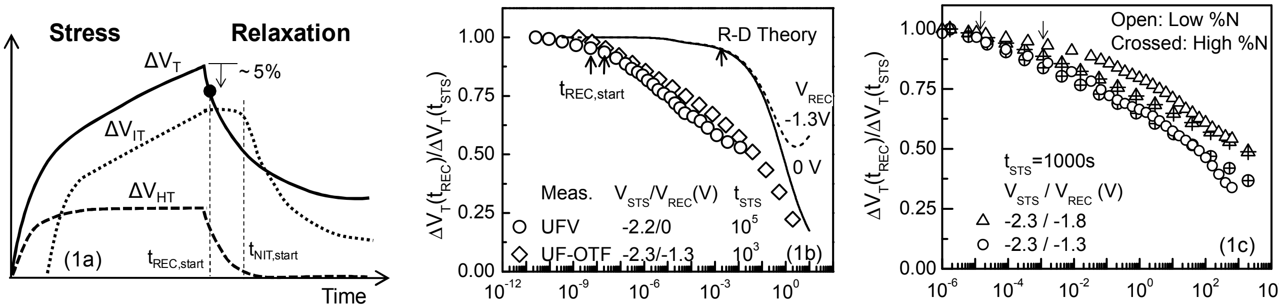

Fig. 1: (a) ΔV_T due to NBTI stress recovers once the stress is removed (solid line). On nitrided transistors, ΔV_T can be attributed to: interfacial traps (ΔV_{IT}, dotted line) and pre-existing hole traps (ΔV_{HT}, dashed line) with respective time-dynamics. (b) $t_{REC,start}$ in ultra-fast V_T (UFV) measurement [1], at V_{REC} ~ 0, commences ~5 decades earlier in time compared to the prediction of H-H_2 R-D theory. Our UF-OTF measurement at V_{REC} ~ -1.3V shows similar relaxation trend. (c) Though $t_{REC,start}$ is considered to be ~ μs in several reports [4,5,11], UF-OTF measurements on nitrided transistors, at different V_{REC}, demonstrate that $t_{REC,start}$ depends on %N and (V_{STS} – V_{REC}). In general, $t_{REC,start}$ ranges from μs to ms and is smaller in the presence of higher %N and for larger (V_{STS} – V_{REC}).

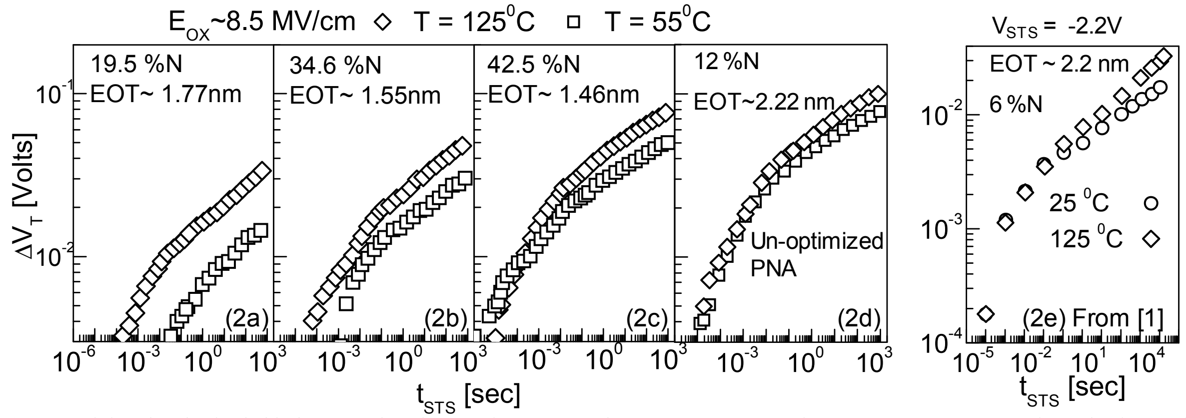

Fig. 2: (a-c) Ascertaining the physics behind NBTI relaxation requires us to analyze NBTI stress experiments at t_{STS} < ms on transistors having a wide variety of plasma nitrided dielectric, measured by UF-OTF [3]. The experiments show T independent ΔV_{HT} dominates over temperature dependent ΔV_{IT} at t_{STS} < ms in high %N transistors. Increase in %N within the dielectric reduces the T dependency at t_{STS} < ms, thus indicating an increase in ΔV_{HT} with %N [6,14]. (d) Un-optimized post-nitridation anneal (PNA) can show the presence of fast ΔV_{HT} component, even in our lightly nitrided transistor. (d) Similar T independence was also reported for lightly dosed nitrided transistors, using UFV [1].

dynamics) is compared to ultra-fast NBTI measurements.

3. Existence of N_{HT}/N_{IT}

As recovery follows stress, it is necessary to establish the existence of N_{HT}, along with N_{IT}, during stress in p-MOSFETs having different gate insulators. Indeed, stress-phase UF-OTF measurements (where, time-zero delay, t_0 is 1μs [3] and ΔV_T is estimated ignoring the mobility correction [15]; *i.e.*, $\Delta V_T \sim \Delta I_D/I_{D0}(V_G-V_{T0})$, Fig.2a-d) reveal the existence of a short-time, T independent fast component (that saturates within ~ms) in transistors having high %N plasma nitrided oxide (Fig. 2c). Decrease in %N within the dielectric reduces the contribution from this fast component (Fig. 2a) and shows the existence of T activated slow component, even at short t_{STS}. Moreover, at long t_{STS} (>1s), all devices show some degree of T activation (less for high %N). Since ΔV_{HT} is fast and T independent compared to slow and T dependent ΔV_{IT} [9], the existence of both ΔV_{HT} and ΔV_{IT} are identified in nitrided transistors (see schematic in Fig. 1a). At short t_{STS} (<ms), ΔV_{HT} dominates over ΔV_{IT} (more for high %N) and thus can explain the observed T insensitivity in ΔV_T. At long t_{STS}, ΔV_{IT} starts to take over and results T dependent ΔV_T.

Moreover, the increase in ΔV_{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 ΔV_{HT}, which can explain the observation of T independent ΔV_T at short t_{STS} in [1], even for smaller %N.

4. Separation of N_{HT}/N_{IT} from ΔV_T

Based on the observations of (i) T independent ΔV_T at short t_{STS} and (ii) saturation of ΔV_{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 ΔV_{IT} using this algorithm, a constant (saturated) ΔV_{HT} is subtracted from ΔV_T for t_{STS}>1s

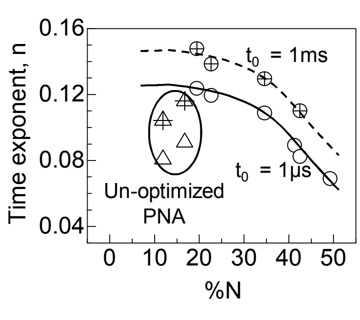

Fig. 3: Power-law NBTI time exponent (n) variation with %N is consistent with the observations in Fig. 2. Here, a reduction in n with %N indicates an increase in ΔV_{HT} [14]. In addition, n shows an increment with increase in t_0-delay [14]. Moreover, the effect of un-optimized PNA (leading to significant ΔV_{HT}, even at low %N) is also evident here.

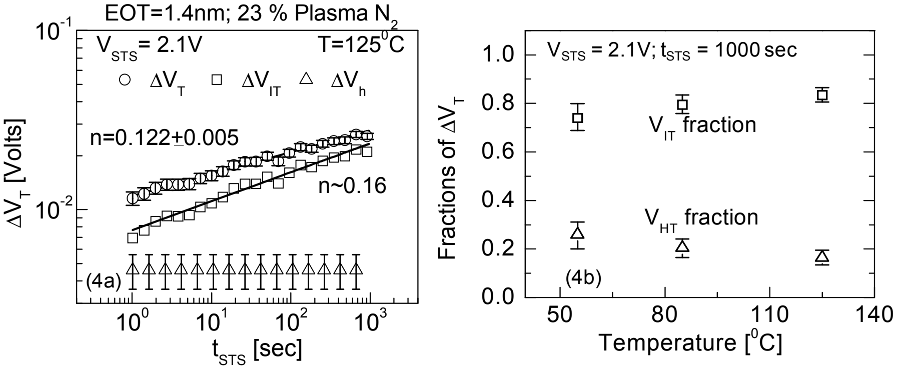

Fig. 4: (a) Figs. 2-3 confirm the existence of ΔV_{HT} in nitrided transistors and presumably ΔV_{HT} saturates within 1s [9]. Thus, it is possible to subtract constant ΔV_{HT} from ΔV_T in such a way [6] that it provides n ~1/6 for ΔV_{IT}. Here, the effect of *mobility* [15] *and electric field-reduction* [9] *is taken into account* in estimating ΔV_T. (b) Estimated $\Delta V_{HT}/\Delta V_T$ and $\Delta V_{IT}/\Delta V_T$ at different temperature indicate an increase (decrease) in $\Delta V_{IT}/\Delta V_T$ ($\Delta V_{HT}/\Delta V_T$) at higher temperature, which is primarily due to the T independence of the ΔV_{HT} component.

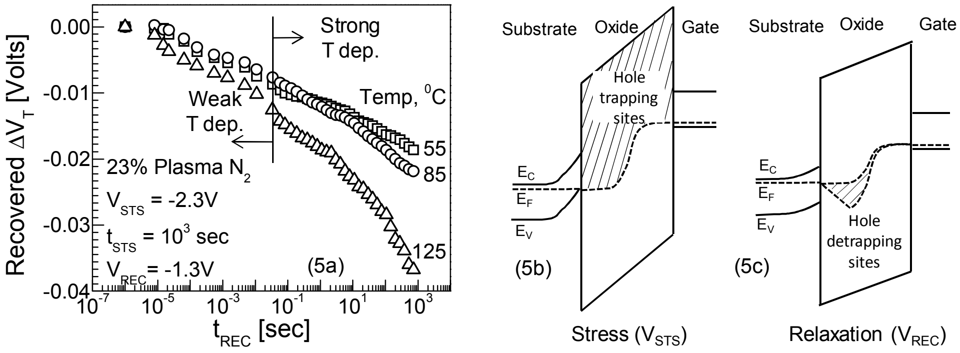

Fig. 5: (a) Similar to NBTI stress phase (Fig. 2), early relaxation phase (up to ~ms) is approximately T independent. Thus hole detrapping from pre-existing oxide traps is the predominant mechanism at $t_{REC} \leq$ ~ms for V_{REC} = -1.3V and start of significant amount of T dependent N_{IT} relaxation ($t_{NIT,start}$) will also be ~ ms. (b) Schematic (based on simulation within a Shockley-Read-Hall trapping-detrapping framework [9]) for expected hole trapping sites (hatched region) at V_{STS}. (c) When gate bias is switched from V_{STS} to V_{REC}, the hatched region will detrap the captured hole in a temperature independent manner.

in such a way that it provides time exponent n~1/6 for ΔV_{IT} (Fig. 4a). Since ΔV_{IT} has higher T activation compared to ΔV_{HT}, extracted $\Delta V_{IT}/\ \Delta V_T$ increases (*i.e.* $\Delta V_{HT}/\Delta V_T$ decreases) with increase in T at fixed t_{STS} (Fig. 4b). The error bar in Fig. 4 results from noise in $I_{D,lin0}$ measurement for UF-OTF [3], which causes a ±0.005 error in *n* for ΔV_T and a ±1mV error in estimated ΔV_{HT}.

Next, we use the same N_{HT}/N_{IT} 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 t_{REC} ~ ms (Fig. 5a and [13, 14]). Thus hole detrapping from pre-existing oxide traps is the predominant mechanism at $t_{REC} \leq$ ~ms for V_{REC} = -1.3V and start of significant amount of temperature dependent N_{IT} relaxation ($t_{NIT,start}$) is also ~ ms. Moreover, UF-OTF at $V_{REC} \leq -1.8$V shows insignificant relaxation up to ~ms (Fig. 6a), which leads to the conclusion that hole detrapping occurs predominantly for $V_{REC} > -1.8$V, and that the amount of hole detrapping is similar from V_{REC}= −1.3V to −1.6V (Fig.6a). Thus trapping sites within the quasi-Fermi levels at V_{REC} = −1.8V and −1.6V (shown schematically by the hatched region in Fig. 5c) will detrap all the holes that were captured during stress. Considering such total hole detrapping at V_{REC}= −1.3V, $t_{NIT,start}$ ~ ms is obtained at all V_{REC}, which is comparable with the relaxation predicted by R-D theory (Fig. 6b). Remaining theory-experiment gap (x10 in t_{STS}, or 10% in N_{IT} 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 N_{OT} at V_{STS} (resultant ~3mV ΔV_{OT} in ΔV_T for Fig. 6c). Indeed, Fig. 7 confirms the existence of universal ΔV_{OT}, 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 dependent 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 ΔV_T is measured at the end of AC cycles) for ≤50% duty cycle will measure $\Delta V_{IT}(AC)/[\Delta V_{HT}(DC) + \Delta V_{IT}(DC)]$ and hence will always be less than the contribution from N_{IT}'s component, $AC/DC(N_{IT}) = \Delta V_{IT}(AC)\ /\Delta V_{IT}(DC)$, predicted by R-D theory [7] (Fig. 8a). Moreover, as ΔV_{HT} decreases for smaller %N, AC/DC ratio in low %N transistors will only

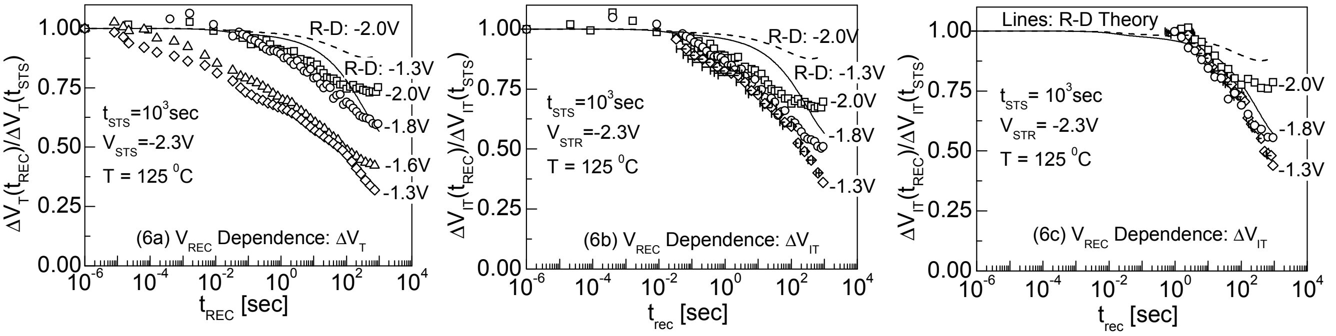

Fig. 6: (a) NBTI relaxation at different V_{REC} show significantly different $t_{REC,start}$. Since relaxation for $V_{REC} \leq$ -1.8V is very close the R-D theory, there is an additional relaxation mechanism for V_{REC} > -1.8V. The observation in Fig. 2-5 suggests hole detrapping to be the additional mechanism. (b) Here, we consider total hole detrapping at V_{REC} = -1.3V in the early phase of NBTI relaxation. Resultant $t_{NIT,start}$ at V_{REC} = -1.3V is similar with the one at $V_{REC} \leq$ -1.8V. This confirms the presence of negligible hole detrapping at $V_{REC} \leq$ -1.8V for this transistor. The error bar along the x-axis for V_{REC} = -1.3V is due to the $I_{D,lin0}$ error in calculating $\Delta V_{IT}/\Delta V_T$ (see Fig. 4). (c) Though Figs. 4-6b assume hole trapping (detrapping) into (out of) pre-existing oxide traps, *i.e.*, ΔV_{HT} saturates before t_{STS} ~ sec [9]; possibility of some amount oxide trap (N_{OT}) generation (resultant ΔV_{OT} in ΔV_T) at V_{STS} can still be present (see Fig. 7). Assuming hole detrapping from N_{OT} (with ΔV_{OT} ~ 3mV at t_{STS} =10^3 sec, having the time dependence similar to Fig. 7c), $t_{NIT,start}$ comes closer to the R-D prediction.

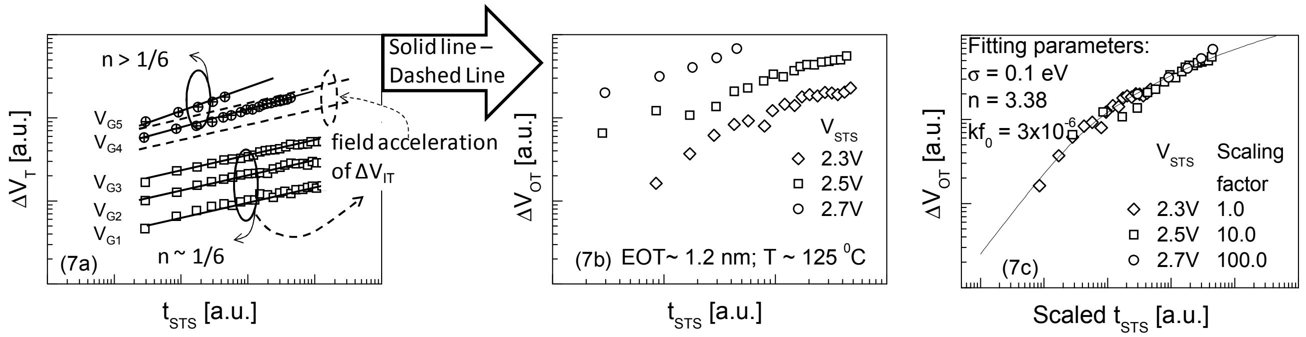

Fig. 7: (a) Time exponent of NBTI degradation (fitted solid lines) increases at $V_{STS} \geq V_{G4}$ [16], thus indicating clear presence of N_{OT} (or ΔV_{OT}) at these V_{STS}. Considering negligible ΔV_{OT} at $V_{STS} \leq V_{G3}$ and calculating field acceleration for ΔV_{IT}, we can estimate ΔV_{IT} at $V_{STS} \geq V_{G4}$ (dashed lines) and hence ΔV_{OT} (= $\Delta V_T - \Delta V_{IT}$) at $V_{STS} \geq V_{G4}$. (b) Similar calculation of ΔV_{OT} for the measurements in [15] shows a decrease in time exponent for ΔV_{OT} at higher V_{STS}. (c) Estimated ΔV_{OT} at different V_{STS} scales universally (similar to the observation in [17]) and can be fitted using $N_{OT} = \int n_{OT}(E)dE$, $n_{OT}(E) = g(E)[1-\exp(-kf(E)*t)]$, $kf = kf_0\exp(-(E-E_0)/kT)$ and $g(E) \sim 1/\sigma \exp((E-E_0)/\sigma)/[1+\exp((E-E_0)/\sigma)]^2$; where, E_0 is the average bond-dissociation energy and σ is its standard deviation.

have AC/DC(N_{IT}), 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_{IT}) 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_{IT}). As a result, AC/DC ratio (similar to AC/DC(N_{IT})) is always frequency independent, irrespective of %N (Fig. 8b). However, due to the presence of ΔV_{HT} in DC stress, the magnitude of AC/DC ratio is less than AC/DC(N_{IT}) 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_{IT}) 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 ~ 1ms [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

[1] Reisinger *et al.*, IRPS '06, p.448. [2] Shen *et al.*, IEDM '06, p.333. [3] Maheta *et al.*, TED '08, p.2614. [4] Grasser *et al.*, IEDM '07, p.801. [5] Grasser *et al.*, IRPS '09, p.33. [6] Mahapatra *et al.*, TED '09, p.236. [7] Alam, IEDM '03, p.345. [8] Mahapatra *et al.*, IRPS '07, p.1. [9] Islam *et al.*, TED '07, p.2143. [10] Chakravarthi *et al.*, IRPS '04, p.273. [11] Kaczer *et al.*, IRPS '08, p. 20. [12] Kapila *et al.*, IEDM '08, p.103. [13] Deora *et al.*, EDL '09, p.978. [14] Mahapatra *et al.*, ECS '09, p.243. [15] Islam *et al.*, IRPS '08, p.87. [16] Chen *et al.*, IRPS '05, p. 704. [17] Varghese *et al.*, IEDM '06, p.751. [18] Fernandez *et al.*, IEDM '06, p.337. [19] Chen *et al.*, EDL '02, p. 734.

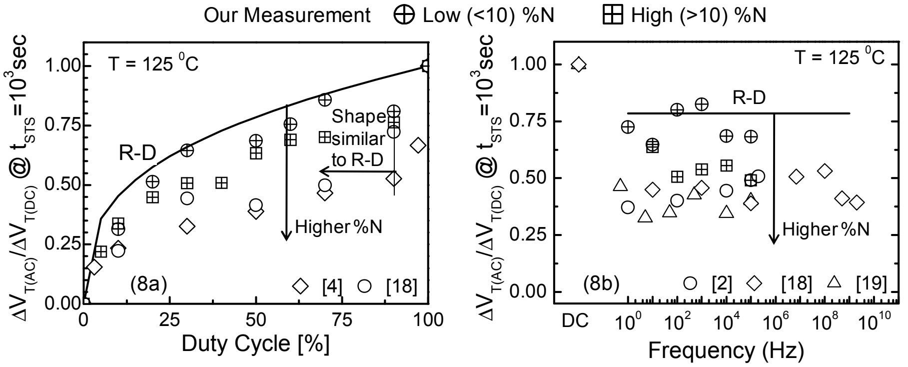

Fig. 8: (a) AC/DC ratio (when ΔV_T is measured at the end of AC cycles) vs. duty cycle plot for different nitrided transistors. The experiments show significant consistency with the prediction of R-D theory (solid line) for low %N, where $\Delta V_T \sim \Delta V_{IT}$. Consideration of ΔV_{HT}, in addition 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.