

# On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?

S. Mahapatra<sup>1#</sup>, K. Ahmed<sup>2</sup>, D. Varghese<sup>1,3</sup>, A. E. Islam<sup>3</sup>, G. Gupta<sup>1</sup>, L. Madhav<sup>1</sup>, D. Saha<sup>1</sup> and M. A. Alam<sup>3</sup>

<sup>1</sup>Department of Electrical Engineering, IIT Bombay, Mumbai 400076, India (#Email: souvik@ee.iitb.ac.in)

<sup>2</sup>Applied Materials, Santa Clara, CA; <sup>3</sup>School of EECS, Purdue University, W. Lafayette, IN

## ABSTRACT

Negative Bias Temperature Instability (NBTI) is studied in plasma (PNO) and thermal (TNO) Si-oxynitride devices having varying EOT. Threshold voltage shift ( $\Delta V_T$ ) and its field ( $E_{OX}$ ), temperature (T) and time (t) dependencies obtained from no-delay on-the-fly linear drain current ( $I_{DLIN}$ ) measurements are carefully compared to that obtained from Charge Pumping (CP). It is shown that thin and thick PNO and thin TNO devices show very similar NBTI behavior, which can primarily be attributed to generation of interface traps ( $\Delta N_{IT}$ ). Thicker TNO devices show different NBTI behavior, and can be attributed to additional contribution from hole trapping ( $\Delta N_h$ ) in pre-existing bulk traps. A physics based model is developed to explain the experimental results. [Keywords: NBTI, plasma and thermal nitridation, interface traps, hole trapping, reaction-diffusion model]

## INTRODUCTION

NBTI is a serious reliability concern for Si-oxynitride (SiON) p-MOSFETs [1]-[5]. Proper stress, measurement and extrapolation are required for reliable estimation of device lifetime. It has been shown that stress  $V_G$  must be carefully chosen to minimize bulk trap generation [6], and the delay between stress and measurement must be reduced or completely eliminated to avoid complications due to recovery effects [3],[7]. However, the extrapolation from stress to operating bias ( $V_G$ ) and from stress time to the end-of-life requires knowledge about NBTI physical mechanism [3]-[5], [8]-[11], which is a subject of significant practical interest and needs careful attention.

It is generally believed that NBTI stress results in donor like  $N_{IT}$  [12] by breaking Si-H bonds at the Si-SiO<sub>2</sub> interface [13]-[16]. It has been shown that with increasing N<sub>2</sub>%, generated  $N_{IT}$  shows higher density near conduction band (CB) edge in the band gap [4],[17] and shows ESR signature different from the usual Pb<sup>0</sup> centers [18]. It is also shown that  $\Delta N_{IT}$  is  $E_{OX}$  driven (not  $V_G$  driven) process, and inversion layer holes in the channel (not electrons tunneling from gate) play a crucial role [6],[19].

However, although NBTI in pure SiO<sub>2</sub> or moderate N<sub>2</sub>% PNO devices is shown to be dominated by  $\Delta N_{IT}$  [3],[4], the importance of  $\Delta N_h$  to overall  $\Delta V_T$  (especially in thicker TNO films) remains unresolved [8]-[11],[20],[21]. Note,  $\Delta N_{IT}$  and  $\Delta N_h$  show different time dependence (power law versus log respectively), different T activation energy and different  $E_{OX}$  dependence. Therefore, any presence of  $\Delta N_h$  would significantly impact the nature (time and  $E_{OX}$  dependence) of the composite degradation. This in turn would impact reliable estimation of the safe operating voltage ( $V_{SAFE}$ ) for a given technology node. Moreover, careful decomposition of the roles of  $\Delta N_{IT}$  and  $\Delta N_h$  are necessary for accurate determination of NBTI-aware process and IC design [22].

In this work, the EOT dependence of NBTI in PNO and TNO devices is studied using on-the-fly  $I_{DLIN}$ . Extracted  $\Delta V_T$  magnitude and its time,  $E_{OX}$  and T dependencies are analyzed and carefully

compared to CP measured in thicker EOT devices. Based on these experiments, we conclude that  $\Delta V_T$  is due to  $\Delta N_{IT}$  for PNO and thin TNO devices. Thicker TNO devices show  $\Delta N_{IT}$  and additional  $\Delta N_h$ , though the latter is only significant at short stress time and high stress  $V_G$ . A physics based model is developed to analyze and explain the EOT dependence of NBTI in PNO and TNO devices.

## EXPERIMENTAL DETAILS

Experiments were performed on PNO and TNO devices having varying EOT (12Å<sup>0</sup> through 22Å<sup>0</sup>) and varying N<sub>2</sub> dose (up to 21% for thin EOT, up to 29% for thick EOT). On-the-fly  $I_{DLIN}$  [7] and delay ( $t=100ms$ )  $I_{DLIN}$  [21] measurements were done on all devices. Separate CP [23] measurements were done on thick (EOT=22Å<sup>0</sup>) devices.  $\Delta V_T$  was calculated using:

- (1) *On-the-fly*  $I_{DLIN}$ :  $\Delta V_T = -\Delta I_D / I_{D0} * (V_{GS} - V_{T0})$ ;  $V_{GS}$  is stress  $V_G$ ,  $V_{T0}$  is pre-stress  $V_T$ , and  $I_{D0}$  is first data point obtained within  $t_0=1ms$  of the application of  $V_{GS}$ .
- (2) *Delay*  $I_{DLIN}$ :  $\Delta V_T = -\Delta I_D / I_{D0} * (V_{GM} - V_{T0})$ ;  $V_G$  is pulsed from  $V_{GS}$  (stress) to  $V_{GM}$  (measurement) and back to  $V_{GS}$  in  $t_{delay}=100ms$ , where  $I_{D0}$  is pre-stress  $I_D$  at  $V_G=V_{GM}$  ( $|V_{T0}| \ll |V_{GM}| \leq |V_{GS}|$ ) obtained from I-V sweep prior to stress.
- (3) *CP*:  $\Delta V_T = q * \Delta N_{IT} / C_{INV}$ ;  $\Delta N_{IT} = \Delta I_{CP} / (q.f.A_G)$ , where  $\Delta I_{CP}$  is increase in CP current measured at frequency  $f$ ,  $A_G$  is gate area and  $q$  is electronic charge;  $C_{INV}$  is gate capacitance in inversion (obtained from CV measurements).

Note that mobility degradation by charges generated due to NBTI stress has been neglected while calculating  $\Delta V_T$  from  $I_{DLIN}$ , due to inversion layer screening effect at high gate overdrive [4]. However, this results in slight overestimation of  $\Delta V_T$  as the entire  $I_{DLIN}$  degradation is attributed to  $V_T$  shift alone. For delay  $I_{DLIN}$  measurements, gate pulse directly switches from  $V_{GS}$  to  $V_{GM}$  and back to  $V_{GS}$  without going through  $V_G=0V$ , which is an important requirement for recovery reduction. To minimize recovery during CP, measurements were performed (after stress is stopped) using gate pulse having fixed amplitude and frequency. Delay due to measurement after each stress step was 500ms typical, unless mentioned otherwise.

## COMPARISON OF $I_{DLIN}$ and CP TECHNIQUES

It is well known that  $I_{DLIN}$  measures total  $\Delta V_T$  due to  $\Delta N_h$  and  $\Delta N_{IT}$ , while CP measures contribution due to  $\Delta N_{IT}$  only. So, a simple comparison of  $I_{DLIN}$  and CP would ideally resolve relative contribution of  $\Delta N_h$  and  $\Delta N_{IT}$  to overall  $\Delta V_T$ . However, such straightforward analysis is impossible as (i)  $I_{DLIN}$  and CP scan different zones in the energy band-gap and (ii) delay inherent in CP measurements allows  $N_{IT}$  relaxation during the measurement time. Therefore, any systematic analysis of the contribution of  $\Delta N_h$  and  $\Delta N_{IT}$  must address the question of measurement at the very onset.

Fig.1 plots time evolution of  $\Delta V_T$  generation from on-the-fly  $I_{DLIN}$ , delay  $I_{DLIN}$ , and CP for thick ( $EOT=22\text{\AA}$ ) PNO and TNO devices and the following observations are made: (i)  $\Delta V_T$  obtained from  $I_{DLIN}$  is larger than  $\Delta V_T$  obtained from CP, (ii) delay  $I_{DLIN}$  shows lower  $\Delta V_T$  than on-the-fly  $I_{DLIN}$  only at lower stress time, the difference becomes negligible at longer stress time, (iii)  $\Delta V_T$  from CP is slightly larger for TNO w.r.t PNO, while (iv)  $\Delta V_T$  from  $I_{DLIN}$  is much larger for TNO w.r.t PNO.

As mentioned before,  $\Delta V_T$  from  $I_{DLIN}$  is a sum of  $\Delta N_{IT}$  and  $\Delta N_h$  (if any), while  $\Delta V_T$  calculated from CP is a measure of  $\Delta N_{IT}$  only. As  $\Delta V_T$  measured from on-the-fly  $I_{DLIN}$  is much larger than that from CP, one might led to believe (incorrectly as we show later) that  $\Delta N_h$  is the primary cause of NBTI for PNO and TNO devices. In addition, Fig.2 shows time evolution of  $\Delta V_T$  generation and recovery from delay  $I_{DLIN}$  and CP measurements. It can be clearly seen that the recovery of  $\Delta V_T$  obtained from  $I_{DLIN}$  is much larger than that from CP. Similar results were also obtained by others and NBTI recovery was attributed to detrapping of trapped holes [8],[20]. This again leads one to believe that  $\Delta N_h$  dominates  $\Delta V_T$  during NBTI stress [8],[9]-[11],[20].

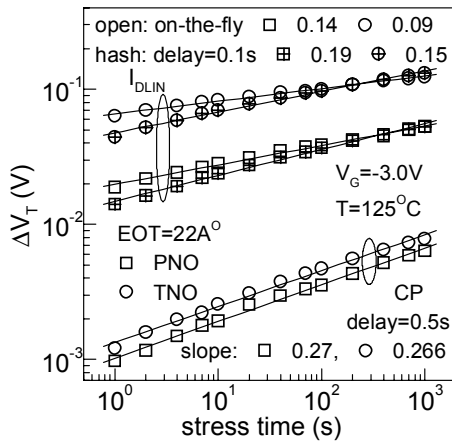


FIG.1. TIME EVOLUTION OF  $\Delta V_T$  GENERATION FROM ON-THE-FLY  $I_{DLIN}$ , DELAY  $I_{DLIN}$  AND CP ON THICK PNO AND TNO DEVICES.

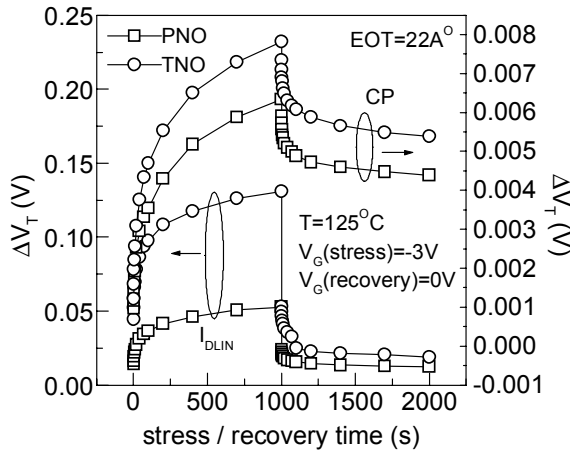


FIG.2. GENERATION AND RECOVERY OF  $\Delta V_T$  FROM DELAY  $I_{DLIN}$  AND CP ON THICK PNO AND TNO DEVICES.

To interpret the results from CP and  $I_{DLIN}$  properly, first note that the use of delay  $I_{DLIN}$  prevents unintentional stress during recovery measurements by making  $V_{GM}$  less than  $V_{GS}$ . Note (see

Fig.1) that the difference in  $\Delta V_T$  obtained by on-the-fly  $I_{DLIN}$  and delay  $I_{DLIN}$  is negligible at long stress time ( $t=1000s$ ) used before the onset of the recovery phase. Hence, the magnitude of  $\Delta V_T$  recovery determined by delay (small)  $I_{DLIN}$  method does not suffer from errors due to un-recorded recovery between end of stress and onset of measurement at the beginning of recovery phase [24]. However CP suffers from larger recovery than  $I_{DLIN}$  [3], due to larger inherent delay and application of zero  $V_G$  during switching from stress to measurement. Hence, large un-recorded recovery would take place between the end of stress and start of first CP measurement in the recovery phase, which would result in lower magnitude of recordable recovery once the stress is removed. Since actual magnitude of recovery cannot be determined from CP due to associated delay, the magnitude of  $\Delta V_T$  recovery from CP and  $I_{DLIN}$  cannot be compared, as done in [8], to comment on the presence (or the lack of)  $\Delta N_h$  during NBTI stress.

Similarly, inherent delay and hence large recovery associated with CP would cause lower  $\Delta V_T$  magnitude during NBTI stress. Moreover, the energy zone in the band gap scanned by CP (near midgap) is quite different from that scanned by  $I_{DLIN}$  (full band gap,  $E_G$ ). Therefore, generation of  $\Delta V_T$  as directly measured from CP and  $I_{DLIN}$  cannot be compared to each other, as done in [9], without correcting for differences in scanned energy zones in the band gap and measurement delays.

**Energy Zone Correction:** Fig.3 shows time evolution of  $\Delta V_T$  for thick PNO device obtained from CP measurements with different rise and fall times of the gate pulse. Energy zone in the band gap scanned by CP increases with reduction in rise-fall time, resulting in higher  $\Delta I_{CP}$  and calculated  $\Delta V_T$  magnitude as per the following relations (see [23] for details):

$$\Delta I_{CP} = q \cdot f \cdot A_G \cdot \Delta N_{IT} \quad (1a)$$

$$\Delta N_{IT} = \langle \Delta D_{IT} \rangle \cdot \Delta E \quad (1b)$$

$$\Delta E = -(2kT) \cdot \ln [n_i \cdot v_{th} \cdot \sigma_0 \cdot (t_R \cdot t_F)^{0.5} \{ (V_T - V_{FB}) / \Delta V \}] \quad (1c)$$

where  $D_{IT}$  is interface trap density distribution in the band gap ( $eV^{-1} \cdot cm^{-2}$ ),  $\Delta E$  is energy zone in band gap scanned by CP,  $\langle \Delta D_{IT} \rangle$  is average  $D_{IT}$  generation in  $\Delta E$ ,  $kT$  is thermal energy,  $n_i$  is intrinsic carrier density,  $v_{th}$  is thermal velocity,  $\sigma_0$  is mean capture cross section of electrons and holes,  $t_R$  and  $t_F$  are rise and fall time of the gate pulse having amplitude  $\Delta V$ ,  $V_T$  and  $V_{FB}$  are threshold and flatband voltages as per CP definition.

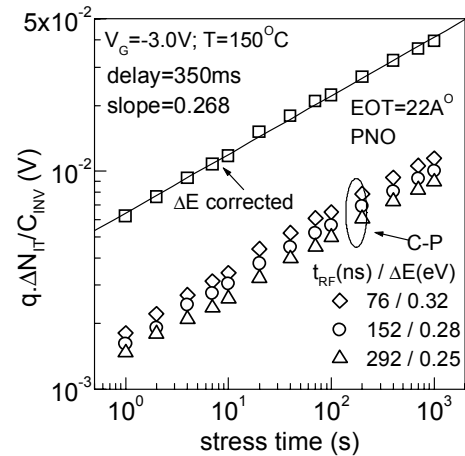


FIG.3. TIME EVOLUTION OF  $\Delta V_T$  FROM CP FOR VARIOUS RISE/FALL TIME OF GATE PULSE, AND CORRECTED FOR DIFFERENCE IN  $\Delta E$  W.R.T  $I_{DLIN}$ .

Note that extracted CP energy zones are much smaller than  $E_G$ . Hence, the difference in CP and  $I_{DLIN}$  energy zones must be accounted for by multiplying the CP data with the  $E_G/\Delta E$  ratio as a first step towards direct comparison (as shown)<sup>1</sup>. However, any localized (near CB edge in the energy band gap [17]) peak in  $D_{IT}$  generation while detected by  $I_{DLIN}$  cannot be accounted for in CP, even after  $\Delta E$  correction.

**Delay Correction:** Fig.4 shows time evolution of  $\Delta V_T$  for thick PNO device from CP measurements, after  $\Delta E$  correction (as described above) but with different measurement delay.  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  is also shown. Reduced delay reduces  $n$  and increases  $\Delta V_T$  magnitude. Therefore, zero delay data must be calculated as a second step for direct comparison of CP and  $I_{DLIN}$  results. When corrected based on R-D model with molecular  $H_2$  diffusion (as justified later), the difference between  $\Delta V_T$  obtained from  $I_{DLIN}$  and CP reduces to within 20% of each other (without accounting for any localized peak interface trap generation near CB edge in the band gap).

Fig.5 shows the stress  $E_{OX}$  dependence of generated  $\Delta V_T$  (after  $t=1000s$  stress) and post-stress  $E_{OX}$  dependence of recovered  $\Delta V_T$  (in  $t=1s$ , after  $t=1000s$  stress), obtained from CP and  $I_{DLIN}$  for thick PNO and TNO devices.  $\Delta V_T$  from CP is calculated directly without energy zone and delay corrections (as described in Fig.3 and Fig.4). Identical field dependence of  $\Delta V_T$  (versus stress  $E_{OX}$  during generation, versus post-stress  $E_{OX}$  during recovery) is obtained for CP and  $I_{DLIN}$  measurements for PNO device. Due to identical  $E_{OX}$  dependent slopes, the ratio (correction factor) of  $\Delta V_T$  obtained from  $I_{DLIN}$  and CP measurements remains fixed at all stress  $E_{OX}$  during generation (post-stress  $E_{OX}$  during recovery). However for TNO,  $\Delta V_T$  obtained from  $I_{DLIN}$  shows stronger field dependence than that of CP, and the ratio of  $\Delta V_T$  calculated from  $I_{DLIN}$  and CP increases at higher stress  $E_{OX}$  (more positive post-stress  $E_{OX}$ ).

### ANALYSIS OF $I_{DLIN}$ AND CP RESULTS

The preceding  $I_{DLIN}$  and CP results are explained as follows. When carefully analyzed, the magnitude of  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  in thick PNO device is found to be similar to that from CP (see Fig.4). As  $\Delta V_T$  is dominated by  $\Delta N_{IT}$ ,  $I_{DLIN}$  and CP results for PNO show similar  $E_{OX}$  dependence of generation and post-stress  $E_{OX}$  dependence of recovery (see Fig.5). Therefore, the observed power law time dependence with  $n \sim 0.14$  using on-the-fly  $I_{DLIN}$  measurement in PNO (see Fig.1) can be explained using the Reaction-Diffusion (R-D) model with molecular  $H_2$  diffusion<sup>2</sup> that results in theoretical slope<sup>3</sup> of  $n=1/6$  [13]-[16],[25]. Note that our analysis is different from [5],[10], where *significant* hole trapping (log time dependence) is used together with R-D model (atomic H diffusion, power law time dependence with  $n=0.25$ ) to explain the observed power law time dependence with  $n \sim 1/6$ . Also note that  $H_2$  is more stable than  $H^0$  [26] and is likely to be the long-time

<sup>1</sup> Donor like  $N_{IT}$  throughout the band gap is assumed for this calculation. This model can successfully explain the skew in CV curve after NBTI stress [12], without assuming any unrealistic 1:1 correlation of interface traps and fixed charges.

<sup>2</sup> Post-stress  $E_{OX}$  dependence of  $N_{IT}$  recovery (Fig.5) does not necessarily indicates drift/diffusion of  $H^+$  in the R-D model framework [20]. As  $N_{IT}$  generation depends on holes and is higher for more negative  $E_{OX}$ , recovery depends on electrons and is higher for more positive  $E_{OX}$ .

<sup>3</sup> Small difference of experimental  $n$  ( $\sim 0.14$ ) to that predicted from theory ( $n=1/6$ ) can be due to small dispersion correction for  $H_2$  diffusion [15] and also due to  $E_{OX}$  reduction as device is stressed [27].

diffusion species (see [27] for details of  $H^0$  and  $H_2$  diffusion during NBTI).

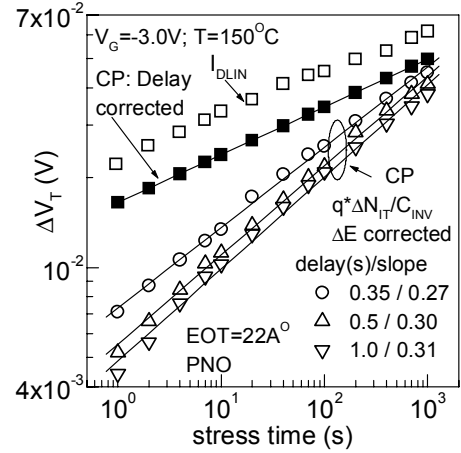


FIG.4. TIME EVOLUTION OF  $\Delta V_T$  FROM ON-THE-FLY  $I_{DLIN}$  AND FROM CP FOR VARIOUS MEASUREMENT DELAY (BUT CORRECTED FOR  $\Delta E$ ). ADDITIONAL CORRECTION FOR DELAY IS DONE ON CP DATA USING R-D MODEL, TO COMPARE WITH  $I_{DLIN}$  DATA.

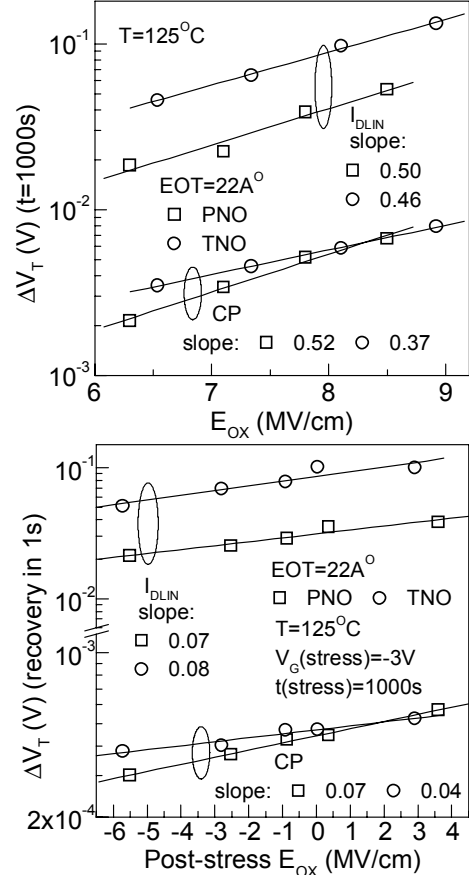


FIG.5.  $E_{OX}$  DEPENDENCE OF  $\Delta V_T$  GENERATION AND RECOVERY FROM  $I_{DLIN}$  AND CP ON THICK PNO AND TNO DEVICES.

As shown in Fig.1, on-the-fly  $I_{DLIN}$  measurement in TNO shows power law dependence with much lower  $n$  ( $\sim 0.09$ ). Unlike PNO,  $\Delta V_T$  obtained from  $I_{DLIN}$  for TNO is much bigger than that from CP, even after correction for energy zone and delay. This

larger difference can be attributed to *additional* hole trapping, whose log time dependence [9],[10] when added to the R-D model solution with molecular H<sub>2</sub> diffusion can explain much lower  $n$  observed from on-the-fly  $I_{DLIN}$  measurements. Higher stress  $E_{OX}$  (post-stress  $E_{OX}$ ) dependent slope of  $\Delta V_T$  generation (recovery) for  $\Delta V_T$  from  $I_{DLIN}$  compared to that from CP (see Fig.5) can also be attributed to hole trapping (detrapping), more at higher  $E_{OX}$  during stress (more positive post-stress  $E_{OX}$  during recovery).

Also note (see Fig.1), delay  $I_{DLIN}$  shows higher  $n$  compared to on-the-fly  $I_{DLIN}$  due to recovery [3],[4],[7]. The difference in  $n$  is higher for TNO compared to PNO, which would imply higher recovery for TNO, consistent with Fig.2. Higher recovery for TNO can be due to detrapping of additional trapped holes. Finally, obtained  $n$  for CP is similar for PNO and TNO devices, which however is much larger than that obtained from  $I_{DLIN}$  due to higher inherent delay and recovery effect in CP measurements.

Alternatively, larger non-uniform enhanced  $\Delta N_{IT}$  near CB [17] (beyond energy zone scanned by CP) for TNO compared to PNO can also explain larger difference of  $\Delta V_T$  obtained from  $I_{DLIN}$  and CP for TNO device. If this non-uniformity arises from a different (than Pb<sup>0</sup>) nature of interface trap [18], it may be possible to have different  $E_{OX}$  dependence of  $\Delta V_T$  generation and recovery for CP and  $I_{DLIN}$  measurements. Therefore based on these results alone, it is not possible to conclusively confirm (or refute) the presence of  $\Delta N_h$  in thick TNO device.

### IMPACT OF EOT, STRESS BIAS, TEMPERATURE

We now proceed to identify and isolate the  $\Delta N_h$  and  $\Delta N_{IT}$  contributions to NBTI based on additional measurements. Note, these two contributions have different functional dependencies on oxide thickness (for constant  $E_{OX}$ ,  $\Delta N_h$  reduces at lower EOT while  $\Delta N_{IT}$  does not), stress bias, temperature ( $\Delta N_{IT}$  is activated, while  $\Delta N_h$  is not). Hence, a systematic study of NBTI in PNO and TNO devices as a function of these parameters would help determine relative contributions of  $\Delta N_{IT}$  and  $\Delta N_h$  on overall  $\Delta V_T$ .

#### A. Impact of EOT and stress bias:

Fig.6 shows time evolution of  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  on thin and thick PNO and TNO devices at different stress biases. Thin and thick PNO and thin TNO devices show similar power law time exponents ( $n=0.14 - 0.15$ ) for a wide range of stress biases. However, thick TNO device shows bias dependent  $n$  that reduce at higher stress  $V_G$ . Obtained  $n$  for thick TNO is much lower than other devices.

Fig.7 plots the time evolution of  $\Delta V_T$  from on-the-fly  $I_{DLIN}$  measurements on thick TNO device and plotted on a lin-log scale. A log time dependence can be observed especially at higher stress  $V_G$ . Similar log time dependence was also observed in [9] for thick TNO devices, and was used to postulate hole trapping for those films. However, thin and thick PNO and thin TNO devices never show such log time dependence (when plotted in a lin-log scale [3], not explicitly shown in this paper).

Fig.8 shows the average power law time exponent (obtained from log-log plot) as a function of EOT, measured using on-the-fly  $I_{DLIN}$  for PNO and TNO devices. Averaging takes care of experimental scatter for thin and thick PNO and thin TNO devices, and  $V_G$  dependence as well as experimental scatter for thick TNO device. Obtained  $n$  for PNO devices is independent of EOT in the range studied. However for TNO device,  $n$  increase with reduction in EOT and the smallest EOT TNO shows identical time exponent as PNO devices.

The power law time dependence with  $n=0.14 - 0.15$ , observed for both thin and thick PNO as well as thin TNO devices can be explained by R-D model with molecular H<sub>2</sub> diffusing species as discussed above (also see the section on Theoretical Interpretation

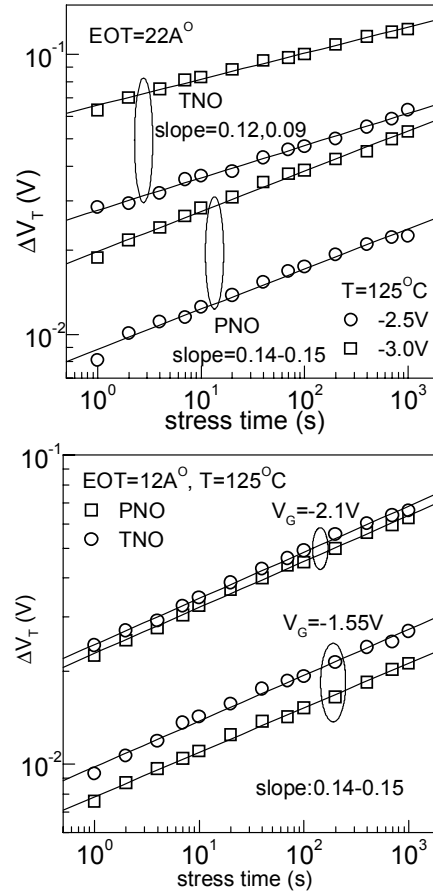


FIG.6. TIME EVOLUTION OF  $\Delta V_T$  MEASURED ON THIN AND THICK PNO AND TNO DEVICES FOR DIFFERENT STRESS BIASES

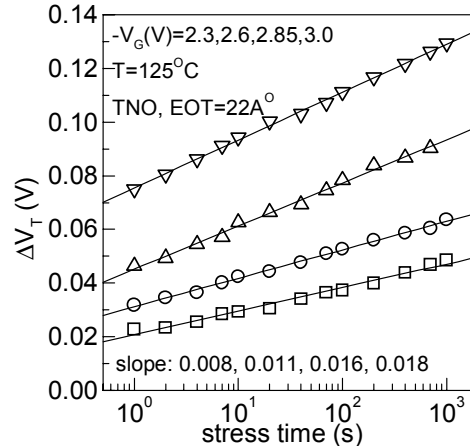


FIG.7. TIME EVOLUTION OF  $\Delta V_T$  AT DIFFERENT STRESS BIAS FOR THICK TNO DEVICE PLOTTED IN A LIN-LOG SCALE (TO HIGHLIGHT LOGARITHMIC DEPENDENCE).

later for quantitative models to support these observations). It is however difficult to explain the EOT dependence of  $n$  for TNO devices in the R-D framework for following reasons. First, the EOT dependence of enhanced  $\Delta N_{IT}$  near the CB edge is difficult

to conceive. Moreover, majority of  $H_2$  diffusion (for stress time window considered here) is expected to take place in poly-Si [4]. Therefore, it is not possible to explain the EOT dependence of  $n$  by invoking EOT dependent dispersion in  $H_2$  diffusion. Further, such a hypothesis does not justify the bias dependence of  $n$  (see Fig.6), and difference in  $E_{OX}$  dependence of  $\Delta V_T$  generation and recovery for  $I_{DLIN}$  and CP data (see Fig.5) as diffusion species is neutral. So, additional hole trapping seems to be the only plausible mechanism that supports EOT dependence of  $n$  for TNO devices, as hole trap volume would decrease and hence  $n$  would increase with reduction in EOT. A mathematical formulation of hole trapping is done later in this paper.

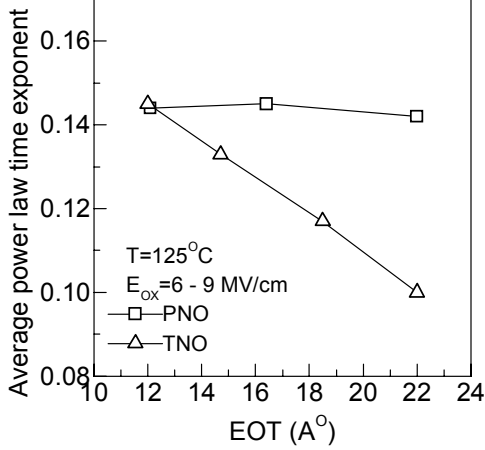


FIG.8. EOT DEPENDENCE OF AVERAGE POWER LAW TIME EXPONENTS OBTAINED FROM ON-THE-FLY  $I_{DLIN}$  IN PNO AND TNO DEVICES.

As a further proof of EOT dependence of  $\Delta N_h$ , Fig.9 shows  $E_{OX}$  dependence of  $\Delta V_T$  (normalized to  $T_{INV}$ ) for thin and thick PNO and TNO devices obtained from on-the-fly  $I_{DLIN}$ . Thick PNO shows lower normalized  $\Delta V_T$  and similar  $E_{OX}$  dependent slope when compared to thin PNO device. For thick PNO,  $I_{DLIN}$  and CP show similar slopes as well (see Fig.5). This suggests absence of  $\Delta N_h$  in PNO, and slightly lower  $N_2$  density at the Si-SiO<sub>2</sub> interface for thick compared to thin PNO (irrespective of higher total dose used for thick PNO).  $E_{OX}$  dependent slope obtained from  $I_{DLIN}$  for thin TNO matches well with that obtained from CP (in thick TNO, see Fig.5). This also suggests  $\Delta N_{IT}$  dominated NBTI for thin TNO and similar  $N_2$  density at the Si-SiO<sub>2</sub> interface for thick and thin TNO. Larger  $E_{OX}$  dependent slope for thin PNO compared to thin TNO suggests lower  $N_2$  density at the Si-SiO<sub>2</sub> interface for thin PNO, irrespective of slightly higher total dose for PNO [19],[28].

However, thick TNO device shows higher normalized  $\Delta V_T$  as well as higher  $E_{OX}$  dependent slope compared to thin TNO (from  $I_{DLIN}$  measurements). Moreover,  $I_{DLIN}$  for thick TNO shows larger  $E_{OX}$  dependent slope than CP (see Fig.5). This is fully consistent with additional  $\Delta N_h$  for thick TNO.

#### B. Impact of EOT and temperature:

In addition to the bias dependence, the T dependence of NBTI provides a mechanism of isolating the  $\Delta N_{IT}$  and  $\Delta N_h$  components of overall  $\Delta V_T$ . This is because  $\Delta N_h$  is insensitive to T but  $\Delta N_{IT}$  is activated (activation energy  $E_A \sim 0.1eV$  [3],[4]). The presence of additional  $\Delta N_h$  would reflect in the overall reduction in the T activation, one can document the relative magnitude  $\Delta N_{IT}$  and  $\Delta N_h$  by studying  $E_A$  as EOT (and hole trapping volume) is varied.

Fig.10 shows T activation of  $\Delta V_T$  at fixed stress time obtained from on-the-fly  $I_{DLIN}$  measurements on PNO and TNO devices. Thin and thick PNO and thin TNO devices show very similar  $E_A$ , which is higher than that of thick TNO device. Fig.11 shows EOT

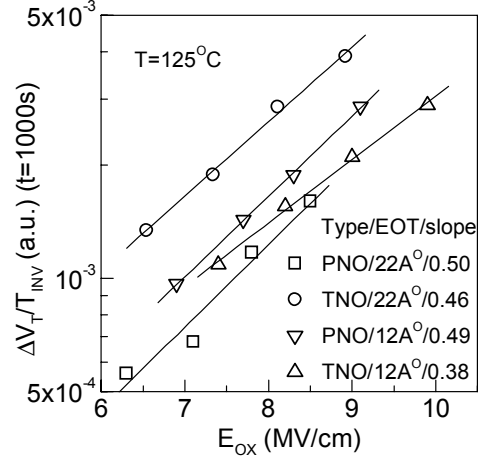


FIG.9.  $E_{OX}$  DEPENDENCE OF NORMALIZED  $\Delta V_T$  FOR THIN AND THICK PNO AND TNO DEVICES.

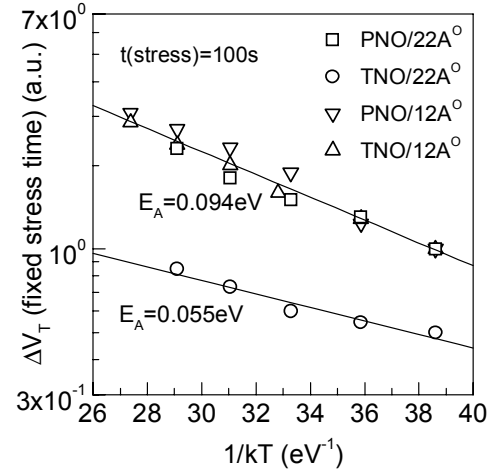


FIG.10. T DEPENDENCE OF  $\Delta V_T$  FOR THIN AND THICK PNO AND TNO DEVICES.

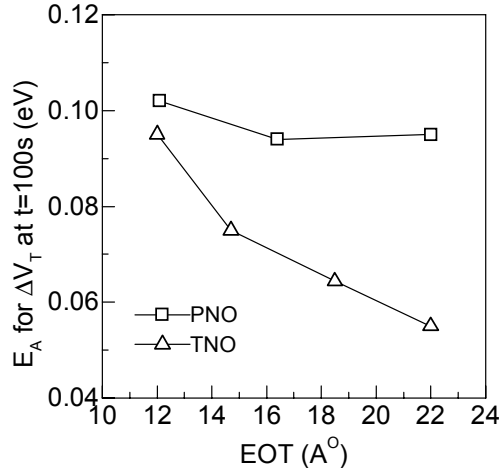


FIG.11. EOT DEPENDENCE OF  $E_A$  OBTAINED FROM ON-THE-FLY  $I_{DLIN}$  IN PNO AND TNO DEVICES.

dependence of  $E_A$  for PNO and TNO devices.  $E_A$  does not change with EOT in the range studied for PNO. However for TNO,  $E_A$

increases with reduction in EOT and smallest EOT TNO device shows identical  $E_A$  as PNO devices. Note that identical  $n$  and  $E_A$  for all PNO devices is consistent with  $\Delta N_{IT}$  driven  $\Delta V_T$ . For TNO devices, the reduction of  $\Delta N_h$  causes both  $n$  and  $E_A$  to increase as EOT is decreased. Hence,  $\Delta N_{IT}$  driven  $\Delta V_T$  becomes true for thin TNO device as well.

Fig. 12 shows the T activation of lifetime (at fixed  $\Delta V_T$ ) from on-the-fly  $I_{DLIN}$  (thin and thick PNO, thin PNO) and CP (thick PNO and thick TNO) measurements, obtained by lateral scaling of  $\Delta V_T$  versus  $t$  data (plotted in a log-log scale) [6],[15]. Identical T activation of lifetimes are obtained from CP for thick TNO and PNO devices, which are similar to those obtained from  $I_{DLIN}$  for thin and thick PNO and thin TNO devices. This is consistent with  $\Delta N_{IT}$  driven  $\Delta V_T$  for all these films. Note, lateral scaling cannot be performed on  $I_{DLIN}$  measurement data in thick TNO device, since it shows very low  $n$  (varies with T) or  $\log(t)$  dependence due to additional  $\Delta N_h$  contribution.

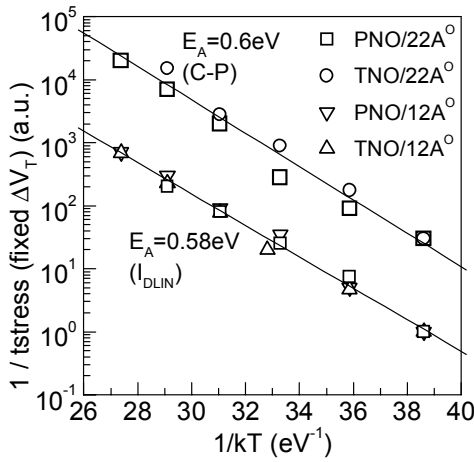


FIG.12. T DEPENDENCE OF LIFETIME FOR THIN AND THICK PNO AND TNO DEVICES.

R-D model suggests that activation energy obtained by lateral scaling relates to T activation of diffusion,  $E_A(D)$ , which is linked to the overall NBTI activation,  $E_A(\Delta V_T)$  (see Fig.10) by  $E_A(\Delta V_T) = E_A(D) * n$  [6],[15]. It can be clearly seen that the above relation readily holds for thin and thick PNO and thin TNO devices, and is consistent with NBTI being dominated by  $\Delta N_{IT}$  and governed by R-D model for these devices. Moreover, obtained  $E_A(D)$  value is consistent with diffusion of neutral  $H_2$  [29], and is also consistent with values of  $n$  ( $\sim 1/6$ ) observed and predicted by R-D model for  $H_2$  diffusion [16],[25]. This reconfirms that for thin and thick PNO and thin TNO devices,  $\Delta V_T$  during NBTI stress is dominated by  $\Delta N_{IT}$ . Finally, we wish to add that due to delay induced increase in  $n$ , larger  $E_A$  (for overall  $\Delta V_T$ ) is obtained from CP than  $I_{DLIN}$  and therefore cannot be compared against each other to conclude about the presence or absence of hole trapping.

## THEORETICAL INTERPRETATION

### A. Separation of interface trap generation and hole trapping:

It is clear from the above discussion that  $\Delta V_T$  in PNO devices is dominated by  $\Delta N_{IT}$ . On the other hand, while thin TNO shows only  $\Delta N_{IT}$  induced  $\Delta V_T$ , thick TNO shows contribution from both  $\Delta N_{IT}$  and  $\Delta N_h$ . As CP data for thick TNO (see Fig.5) shows very similar  $E_{OX}$  dependence as  $I_{DLIN}$  data on thin TNO (see Fig.9), it

can be assumed [19] that the Si-SiO<sub>2</sub> interfacial N<sub>2</sub> density is very similar for thin and thick TNO devices. Hence for identical  $E_{OX}$ , the  $\Delta N_{IT}$  contribution (to a first order) to  $\Delta V_T$  in thick TNO can be obtained by multiplying  $\Delta V_T$  for thin TNO by the thickness (EOT + poly-Si depletion) ratio of the two devices.

Fig.13(a) shows time evolution of  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  in thin TNO device stressed at fixed  $E_{OX}$  but different T. Lines are  $\Delta V_T$  contribution from  $\Delta N_{IT}$ , calculated using following R-D model solution [6],[15],[19]:

$$\Delta V_T = A * \exp(-E_A/kT) * t^n \quad (2)$$

where 'A' depends on  $E_{OX}$  and N<sub>2</sub>% [19], and  $n$  and  $E_A$  values are obtained from Figs. 8 and 11. A good match with experimental data is consistent with  $\Delta N_{IT}$  driven  $\Delta V_T$  for thin TNO devices.

Fig.13(b) shows time evolution of  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  in thick TNO device stressed at fixed  $E_{OX}$  but different T. As  $\Delta N_h$  is T independent,  $\Delta V_T$  increases due to activation of  $\Delta N_{IT}$  with T. So,  $\Delta V_T$  data (symbols) is estimated (lines) by summing up T activated  $\Delta N_{IT}$  (obtained by scaling thin TNO data) and T independent  $\Delta N_h$  as shown. Excellent agreement for a wide range of T validates extraction of  $\Delta N_{IT}$  and  $\Delta N_h$  components.

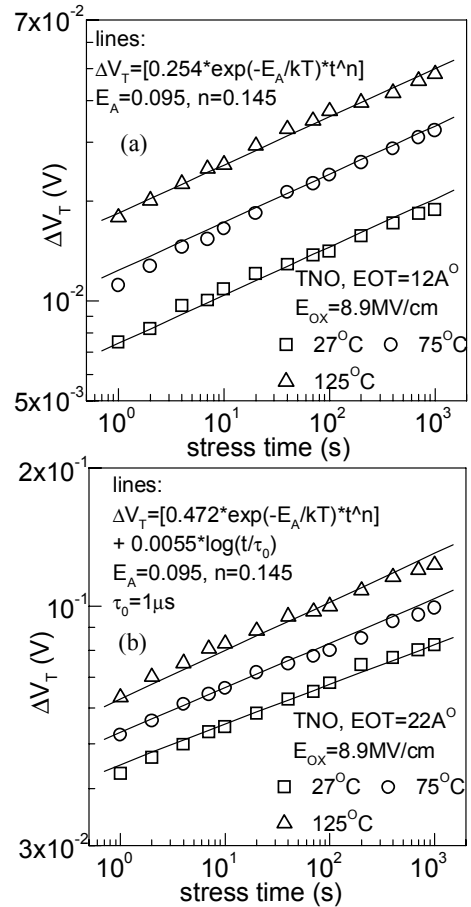


FIG.13. TIME EVOLUTION OF  $\Delta V_T$  FOR THIN AND THICK TNO DEVICE AT DIFFERENT T (CONSTANT  $E_{OX}$ ). LINES ARE MODEL CALCULATION.

Fig.14 shows the  $E_{OX}$  dependence of  $\Delta V_T$  and extracted  $\Delta N_{IT}$  and  $\Delta N_h$  contribution for thick TNO device extracted at short and long stress time. Once again,  $\Delta N_{IT}$  is computed by scaling thin EOT data as described above. Note,  $\Delta N_h$  shows stronger  $E_{OX}$  dependence than  $\Delta N_{IT}$ , and the contribution of  $\Delta N_h$  to overall  $\Delta V_T$

becomes less significant at longer stress time, as well as when  $E_{OX}$  is scaled to operating conditions [20].

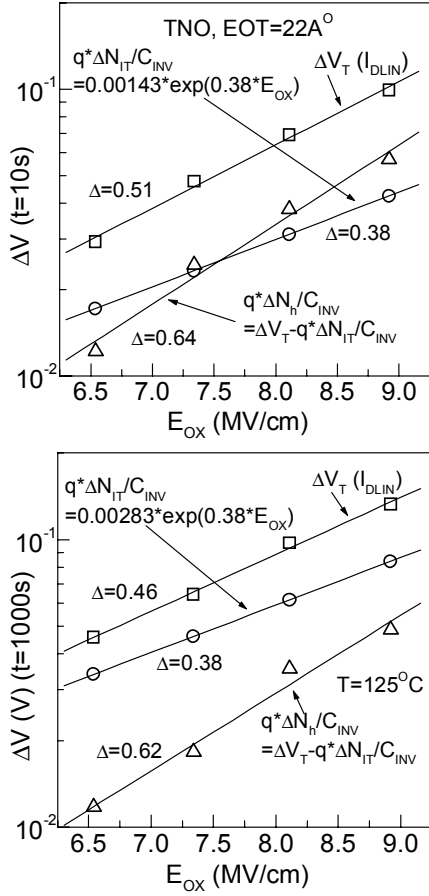


FIG.14.  $E_{OX}$  DEPENDENCE OF  $\Delta V_T$ ,  $\Delta N_{IT}$  AND  $\Delta N_H$  FOR THICK TNO DEVICE AT SHORT AND LONG STRESS TIME.

#### B. Development of hole trapping-detrapping model:

A qualitative interpretation of the results in Figs.8 and 11 can be obtained by considering a simple trapping/de-trapping model, in which holes are trapped from the inversion layer into traps (PNO: peak at  $SiO_2$ -poly, decreasing towards  $Si-SiO_2$ ; TNO: peak at  $Si-SiO_2$ , decreasing towards  $SiO_2$ -poly [30],[31]) in the oxide. This is followed by detrapping of trapped holes towards the empty states in both poly-Si and channel. Trap filling (probability,  $f_T$ ) evolves with time according to the equation:

$$df_T/dt = \sigma \cdot v_{th} * [p \cdot T_1 \cdot (1-f_T) - n_c \cdot T_1 \cdot f_T - n_G \cdot T_2 \cdot f_T] \quad (3)$$

where  $\sigma$  is capture cross section,  $v_{th}$  is thermal velocity,  $p$  is inversion layer hole density,  $n_c$  is concentration of detrapping states at channel ( $\sim 0$ ; filled with holes);  $n_G$  is concentration of detrapping states at poly ( $\sim 10^{22} \text{ cm}^{-3}$ ),  $T_1$  and  $T_2$  are tunneling probabilities (obtained by WKB) of holes from channel to trap and trap to poly. The density of occupied traps are given by  $n_T = N_T f_T$ , where  $N_T$  is the bulk trap density, assumed constant in time<sup>4</sup> and uniformly distributed in energy in the  $SiON$  band gap.

<sup>4</sup> Presently, only single energy hole traps are assumed at the valence band level and can explain the small difference in saturation behavior between theory (Fig.15) and experiment (Fig.13).

Fig.15 shows time evolution of overall  $\Delta V_T$  and contribution by  $\Delta N_h$  and  $\Delta N_{IT}$  for thick TNO device, obtained by numerically solving (3) and the R-D model [16],[25],[27], respectively. R-D model solution for  $\Delta N_{IT}$  shows power law time dependence with  $n \sim 1/3$  at early time and  $n \sim 1/6$  at long time respectively due to conversion of H to  $H_2$  and diffusion of  $H_2$  [25],[27].  $\Delta N_h$  contribution dominates at short time but saturates at long time (time constant consistent with [32]). Overall  $\Delta V_T$  shows lower  $n$  due to additional contribution from  $\Delta N_h$  at intermediate time, but is expected to show  $n \sim 1/6$  when  $\Delta N_{IT}$  completely dominates  $\Delta N_h$  at longer stress time (and lower stress  $E_{OX}$ ).

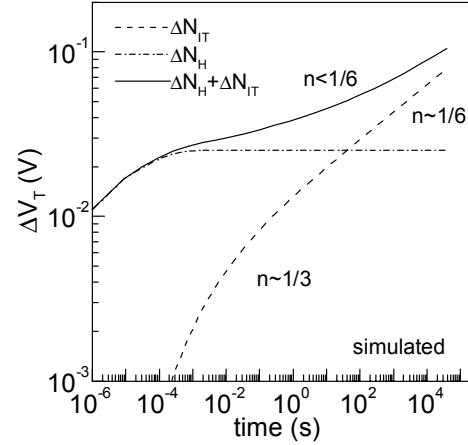


FIG.15. SIMULATED TIME EVOLUTION OF  $\Delta N_{IT}$ ,  $\Delta N_H$  AND TOTAL  $\Delta V_T$  FOR THICK TNO DEVICE.

Similar simulation in thin TNO samples shows that  $\Delta N_h$  does not have significant effect on total  $\Delta V_T$ . This result is intuitive, as thinner oxides allow faster detrapping of holes through the gate [2]. Furthermore,  $\Delta N_h$  contribution is found to be insignificant in both thin and thick PNO devices, as traps are close to the  $SiO_2$ -poly interface, which results in higher trap to poly (compared to channel to trap) hole tunneling probability. The overall effects of hole trapping for different EOT devices are summarized in Fig.16, which shows that both time exponent and activation energy for total  $\Delta V_T$  changes significantly for TNO, while it remains constant for PNO devices. Good qualitative agreement with observations in Fig.8 and Fig.11 clearly suggests that the conclusions based on experimental observations are indeed reasonable.

## CONCLUSION

To summarize, the dependence of NBTI on EOT is studied in PNO and TNO devices under a wide range of stress conditions and the following key features are observed:

- (1) On-the-fly  $I_{DLIN}$  measurements in all PNO devices result in very similar  $E_{OX}$  dependence, power law time exponent ( $n$ ) and  $T$  activation energy ( $E_A$ ) in the range of EOT ( $12A^0$ - $22A^0$ ) and  $N_2\%$  (up to 21% for thinner and up to 29% for thicker EOT).
- (2) PNO devices show identical  $E_{OX}$  dependence and  $E_A$  when on-the-fly  $I_{DLIN}$  results are compared to CP measurements.
- (3) For thicker PNO device,  $\Delta V_T$  calculated from on-the-fly  $I_{DLIN}$  and CP measurements (after proper correction) are close to each other.
- (4) For on-the-fly  $I_{DLIN}$  measurements in TNO devices,  $n$  and  $E_A$  increases and  $E_{OX}$  dependent slope decreases as EOT is scaled.

- (5) For TNO devices, on-the-fly  $I_{DLIN}$  results only from thinner EOT show identical  $E_{OX}$  dependent slope and  $E_A$  compared to that measured from CP.
- (6) For thicker TNO device, on-the-fly  $I_{DLIN}$  measurements show stronger  $E_{OX}$  dependence and lower  $E_A$  compared to CP results.
- (7) For thicker TNO device,  $\Delta V_T$  obtained from on-the-fly  $I_{DLIN}$  is much larger than that obtained from CP (even after proper correction).

It is shown that NBTI in PNO (for the range of EOT and  $N_2\%$  studied) is dominated by interface trap generation. For TNO, hole trapping is significant for thicker films, but reduces at lower EOT, lower  $E_{OX}$  and higher stress time. Experimental observations are analyzed and explained using a physics based theoretical model.

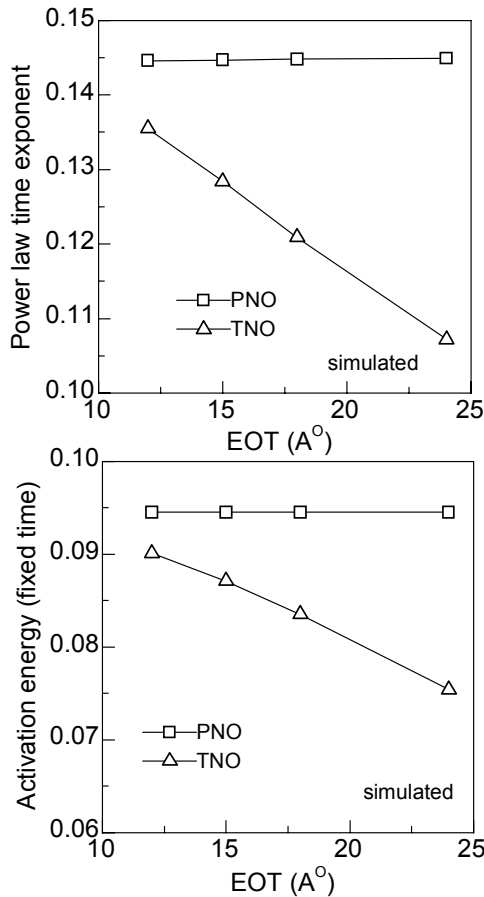


FIG.16. SIMULATED EOT DEPENDENCE OF POWER LAW SLOPE AND T ACTIVATION FOR PNO AND TNO DEVICES.

#### ACKNOWLEDGEMENT

We wish to thank AMAT, Renesas, TI and TSMC for research funding to IIT-B and Purdue.

#### REFERENCES:

- [1] S.S. Tan, T.P. Chen, J.M. Soon, K.P. Loh, C.H. Ang, W.Y. Teo and L. Chan, "Neighboring effect in nitrogen-enhanced negative bias temperature instability", SSDM, p.70, 2003.
- [2] Y. Mitani, "Influence of nitrogen in ultra-thin SiON on negative bias temperature instability under AC stress", IEDM, p.117, 2004.
- [3] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam, "On the dispersive versus arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: measurements, theory, and implications", IEDM, p.684, 2005.
- [4] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, A. Varghese, R. B. Khamankar, and S. Krishnan, "Material dependence of hydrogen diffusion: implications for NBTI degradation", IEDM, p.688, 2005.
- [5] K. Sakuma, D. Matsushita, K. Muraoka and Y. Mitani, "Investigation of nitrogen originated NBTI mechanism in SiON with high nitrogen concentration", IRPS, p.454, 2006.
- [6] S. Mahapatra, P. Bharath Kumar and M. A. Alam, "Investigation and modeling of interface and bulk trap generation during Negative Bias Temperature Instability of p-MOSFETs", IEEE TED, v.51, p.1371, 2004.
- [7] S. Rangan, N. Mielke, and E. C. C. Yeh, "Universal recovery behavior of negative bias temperature instability", IEDM, p.341, 2003.
- [8] V. Huard and M. Denais, "Hole trapping effect on methodology for DC and AC negative bias temperature instability measurement in PMOS transistors", IRPS, p.40, 2004.
- [9] M. Denais, A. Bravaix, V. Huard, C. Parthasarathy G. Ribes, F. Perrier, Y. Rey-Tauriac, and N. Revil, "On-the-fly characterization of NBTI in ultra-thin gate-oxide pMOSFETs", IEDM, p.109, 2004.
- [10] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI degradation- and recovery- behaviour based on ultra fast  $V_T$  measurements", IRPS, p.448, 2006.
- [11] C. R. Parthasarathy, M. Denais, V. Huard, G. Ribes, E. Vincent, and A. Bravaix, "New insights into recovery characteristics post NBTI stress", IRPS, p.471, 2006.
- [12] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability", IRPS p. 248, 2002.
- [13] K. O. Jeppson and C. M. Svensson, "Negative Bias Stress of MOS Devices at High Electric-Fields and Degradation of MNOS Devices", JAP, vol. 48, p. 2004, 1977.
- [14] M. A. Alam, "A Critical Examination of the Mechanics of Dynamic NBTI for p-MOSFETs" IEDM, p.345, 2003.
- [15] M. A. Alam and S. Mahapatra, "A Comprehensive Model for PMOS NBTI Degradation", Microelectronics Reliability, v.45 (1), p.71, 2005.
- [16] S. Chakravarthi, A. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability", IRPS, p.273, 2004.
- [17] J. H. Stathis, G. LaRosa and A. Chou, "Broad energy distribution of NBTI-induced interface states in p-MOSFETs with ultra-thin nitrided oxide", IRPS, p.25, 2004.
- [18] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "NBTI: An atomic-scale defect perspective," IRPS, p.442, 2006.
- [19] A. E. Islam, G. Gupta, S. Mahapatra, A. Krishnan, K. Ahmed, F. Nouri, A. Oates, and M. A. Alam, "Gate leakage vs. NBTI in Plasma Nitrided Oxides: Characterization, Physical Principles, and Optimization", IEDM, p.329, 2006.
- [20] T. L. Yang, M.F. Li, C. Shen, C.H. Ang, Z. Chunxiang, Y.C. Yeo, G. Samudra, S.C. Rustagi, M.B. Yu, "Fast and slow dynamic NBTI components in p-MOSFET with SiON dielectric and their impact on device life-time and circuit application", VLSI Tech., p.92, 2005.

- [21] B. A. Kaczer, R. Degraeve, N. Collaert, G. Groeseneken, M. Goodwin, "Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification", IRPS, p.381, 2005.
- [22] B. Paul, K. Kunhyuk, H. Kufluoglu, M. A. Alam, and K. Roy, "Impact of NBTI on the temporal performance degradation of digital circuits", IEEE EDL, v.26, p.560, 2005.
- [23] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. Dekeersmaecker, "A Reliable Approach to Charge-Pumping Measurements in Mos-Transistors", IEEE T-ED, v.31, p.42, 1984.
- [24] T. Yang, C. Shen, M. F. Li, C. H. Ang, C. X. Xue, Y. C. Yeo, G. Samudra, D. L. Kwong, "Interface trap passivation effect in NBTI measurement for p-MOSFET with SiON gate dielectric", IEEE EDL, p.758, 2005.
- [25] M. A. Alam, NBTI Tutorial (part-1), 2006.
- [26] A. T. Krishnan, S. Chakravarthi, P. Nicollian, V. Reddy, and S. Krishnan, "Negative bias temperature instability mechanism: The role of molecular hydrogen", APL, v.88, 153518, 2006.
- [27] A. E. Islam, D. Varghese, H. Kufluoglu, and M. A. Alam "A Critical Analysis of Short-term Negative Bias Temperature Instability Measurements: Explaining the effect of time-zero delay for On-the-fly Measurements", to appear in APL, 2007.
- [28] G. Gupta, S. Mahapatra, L. Madhav, D. Varghese, K. Ahmed and F. Nouri, "Interface-trap driven NBTI for ultrathin (EOT~12) plasma and thermal nitrided oxynitrides", IRPS, p.731, 2006.
- [29] M. L. Reed and J. D. Plummer, "Chemistry of Si-SiO<sub>2</sub> interface trap annealing", JAP, v.63, p.5776, 1988.
- [30] J. R. Shallenberger, D. A. Cole, and S. W. Novak, "Characterization of silicon oxynitride thin films by X-ray photoelectron spectroscopy", JVST-A, v.17, p.1086, 1999.
- [31] S. Rauf, S. Lim, and P. L. G. Ventzek, "Model for nitridation of nanoscale SiO<sub>2</sub> thin films in pulsed inductively coupled N-2 plasma," JAP, v.98, 024305, 2005.
- [32] L. Lundkvist, L. Lundstrom, and C. Svensson, "Discharge of MNOS Structures", SSE, v.16, p.811, 1973.