Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs

Souvik Mahapatra, *Senior Member, IEEE*, Vrajesh D. Maheta, *Member, IEEE*, Ahmad Ehteshamul Islam, *Student Member, IEEE*, and Muhammad Ashraful Alam, *Fellow, IEEE*

***Abstract*—In this paper, a simple phenomenological technique is used to isolate the hole-trapping and interface trap generation components during negative bias temperature instability (NBTI) stress in plasma nitrided oxide (PNO) p-MOSFETs. This isolation methodology reconciles the apparent differences between experimentally measured NBTI power-law time exponents obtained by ultrafast on-the-fly I_{DLIN} method, which are the ones obtained using slightly delayed but very long-time measurements, and the corresponding exponents predicted by the reaction–diffusion model. A systematic validation of the isolation technique is provided through degradation data taken over a broad range of operating conditions and a wide variety of PNO processes, to establish the robustness and uniqueness of the separation procedure.**

***Index Terms*—Activation energy, field acceleration, hole trapping, interface traps, negative bias temperature instability (NBTI), plasma oxynitride, p-MOSFET, reaction–diffusion (R–D) model, time exponent.**

I. Introduction

NEGATIVE BIAS temperature instability (NBTI) is a serious reliability concern for Si oxynitride (SiON) p-MOSFETs [1]–[12]. In spite of extensive research during the last few years, there are still debates regarding the nature, magnitude, and physical mechanisms of time evolution, oxide field (E_{OX}) dependence, and temperature (T) activation of NBTI. It is now well accepted by most researchers that NBTI degradation measured at moderate to long stress time[1] follows power-law time dependence (t^n), although there is disagreement regarding the value of the measured time exponent (n). It has been shown that n, obtained using conventional stress–measure–stress (SMS) method (particularly at relatively small or moderate range of stress time), is higher than the "true" value due to recovery related artifacts [4], [5], [13], [14]. To avoid measurement delay and obtain "recovery free" NBTI degradation from very short (approximately microseconds) to moderately long measurement time, an ultrafast on-the-fly (UF-OTF) I_{DLIN} method has been proposed [8] and used [10] to determine NBTI degradation of SiON p-MOSFETs having different nitridation types, thicknesses, and nitrogen (N) densities. It has been shown that long-time n reduces with a reduction in time-zero (t_0) delay[2] of OTF measurement and saturates to its final value for $t_0 \leq 10\ \mu$s. It has been shown that N density at the Si/SiON interface [controlled by nitridation type—plasma or thermal, N dose, starting base oxide thickness, and postnitridation anneal (PNA)] strongly influences n, although a measured n value remains independent of stress E_{OX} and T [10]. However, the value of n [$\sim$0.12 for plasma nitrided oxide (PNO) processes, resulting in low Si/SiON N density that reduces further at higher N density] obtained using UF-OTF is always lower than the value $(n \sim 0.16)$[3] obtained by very long-time measurements (albeit with slightly higher delay, although the recovery is expected to have minimal impact at very long stress time) [11], [12], [15]. The T independence of long-time n suggests Arrhenius-like T activation [1], [5], [7], and an earlier report of n as a linear function of T [17] has been shown to be an artifact of measurement delay [5]. The activation energy (E_A) measured at a fixed stress time also shows strong process dependence, with $E_A \sim 0.08$ eV for PNO processes resulting in low Si/SiON N density that reduces further at higher N density [10]. The E_{OX} independence of long-time n suggests the absence of bulk-trap generation (in the range of stress biases used for acceleration) [9], [18], and an inversion-layer hole initiated E_{OX} assisted model has been proposed to model the E_{OX} dependence of NBTI [11], [19]. The exponential dependence[4] of NBTI degradation on E_{OX} is predicted by the model, and the E_{OX} dependent slope (Γ) shows strong process dependence, with $\Gamma \sim 0.6$ cm/MV for PNO processes resulting in low Si/SiON N density that reduces further at higher N density [10].

There is also an intense debate whether NBTI is due to interface trap generation (ΔN_{IT}) [4], [5], [17], [20] or hole

Manuscript received August 6, 2008; revised October 16, 2008. This work was supported by the Applied Materials, USA. The review of this paper was arranged by Editor J. Suehle.

S. Mahapatra and V. D. Maheta are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India (e-mail: souvik@ee.iitb.ac.in).

A. E. Islam and M. A. Alam are with the School of Electrical Engineering and Computer Science, Purdue University, West Lafayette, IN 47906 USA.

Digital Object Identifier 10.1109/TED.2008.2010569

[1]In a log–log plot, the degradation obtained from UF-OTF measurements asymptotically converges to a "final" value for a stress time of $\sim$1 s and higher [10]. All exponents in this paper are obtained for 10-s time and higher and are denoted as "long-time" exponent. The term "very long time" is used for published data measured in the time range of days to several weeks.

[2]The time delay between gate voltage reaching stress level and the corresponding peak drain current being measured at the initial stage of OTF measurement [8].

[3]Data measured at Texas Instruments [11], TSMC [12], and Freescale Semiconductors [15] show this trend, although data from ST Microelectronics [16] show much higher value of $n \sim 0.25$. This paper would assume $n \sim 0.16$ as the long term value.

[4]Power-law E_{OX} dependence has also been proposed by performing measurements over a wide range of stress biases [12]. The time exponent n has been found to increase particularly at higher stress biases due to the generation of bulk oxide traps [9]. Hence, the entire stress period used in [12] cannot be associated to the acceleration of only the defects responsible for NBTI.

0018-9383/$25.00 © 2008 IEEE

trapping (ΔN_h) in preexisting bulk SiON traps [2], [21], [22]. A recent "consensus" among different researchers appears that both $\Delta N_{\rm IT}$ and ΔN_h take place during NBTI stress; while the latter "dominates" at very early stress time, shows weak T dependence, and saturates quickly, the former is a relatively slower but strongly T activated process that builds up in time and "rides above" ΔN_h to eventually cause device failure [3], [6]–[10], [16]–[23]. Once again, the relative contribution of $\Delta N_{\rm IT}$ and ΔN_h to overall NBTI is qualitatively shown to be strongly dependent on SiON process [7]–[10], and ΔN_h effect is mostly observed when NBTI is measured by UF-SMS [2] or UF-OTF [8], [10], or to some extent by conventional OTF techniques [7], [9]. The presence of ΔN_h (more for films with high Si/SiON N density) in addition to $\Delta N_{\rm IT}$ reduces the time and T dependence of overall NBTI [3], [6], [7], [10]. The separation of $\Delta N_{\rm IT}$ and ΔN_h (see [3], [6], [16], [24], and [25] for recent attempts) is important for the extrapolation of stress data to end of life for the estimation of device lifetime.[5]

In this paper, a simple method is proposed to isolate the $\Delta N_{\rm IT}$ and ΔN_h component of overall NBTI degradation in PNO p-MOSFETs treated with proper two-step PNA.[6] It is shown that, once ΔN_h has been subtracted out from the overall NBTI degradation, the time and T dependence of the extracted $\Delta N_{\rm IT}$ part (for longer stress time) are fully consistent with the simple version of reaction–diffusion (R–D) model with nondispersive molecular H_2 diffusion [19], [29]. This is verified in devices having a wide range of thicknesses (EOT of 1.2–2.2 nm) and N content (N atomic of 16%–42%) which justifies the robustness of the isolation scheme and underlying physical mechanism.

II. Results and Discussion

NBTI is measured on p-MOSFETs ($W/L = 15/0.16\ \mu$m) using UF-OTF $I_{\rm DLIN}$ technique with 1-μs resolution [8]. The measured $I_{\rm DLIN}$ degradation is expressed as $\Delta V_T = -(\Delta I_{\rm DLIN}/I_{\rm DLIN0})^*(V_{\rm GSTRESS} - V_{T0})$ [4], where $I_{\rm DLIN0}$ is the drain current measured within 1 μs after the application of stress, $V_{\rm GSTRESS}$ is the applied stress bias, and V_{T0} is the prestress V_T. Note that ΔV_T obtained by the aforementioned equation is proportional to but different from ΔV_T that is typically obtained using transfer I–V sweeps for SMS method, as mobility correction[7] is not considered [30]. CV measurements followed by quantum mechanical corrections and XPS were used to estimate EOT and atomic N% of these devices, respectively.

Fig. 1 shows the measured ΔV_T transients from short to moderately long stress time at two different stress T's, but identical stress $E_{\rm OX}$'s, on devices having identical starting

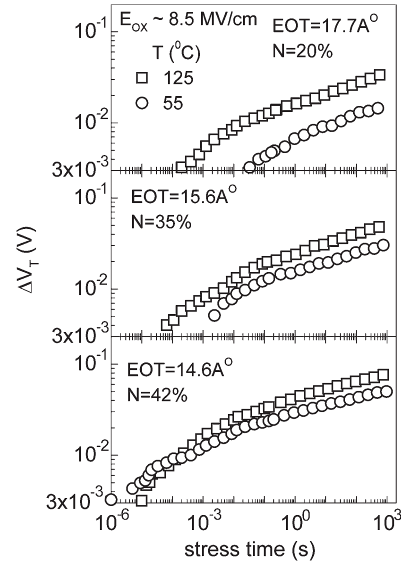

Fig. 1. Time evolution of degradation under NBT stress at identical $E_{\rm OX}$'s but different T's, for PNO devices having identical starting base oxide thicknesses but different N doses, resulting in different atomic N% values and EOTs.

base oxide thicknesses (20 Å) but different N% values and EOTs due to the difference in PNO N doses. Although not observed for low N%, devices having higher N% show very large short-time (in submilliseconds) degradation; the amount of this rapid short-time degradation increases with an increase in N%. The overall degradation also increases with an increase in N% as expected. Furthermore, while the low N% device shows clear T dependence for the entire stress duration, higher N% devices show lower or negligible T dependence at shorter stress duration and lower T activation at longer stress duration. Note that the increase in the overall degradation as N% is increased can be largely attributed to the appearance of this short-time weak T dependent degradation. Fig. 2 shows the measured power-law exponents for long stress time ($t \geq 10$ s) as a function of stress $E_{\rm OX}$ and T for the devices used in Fig. 1. Note that n remains constant (within experimental scatter due to noise in $I_{\rm DLIN0}$) in the range of $E_{\rm OX}$ and T used, although it reduces with an increase in N%. Similar results were presented in much greater detail in [10], where the observed increase in short-time weak T dependent degradation (for certain SiON processes) is always associated with reduction in long-time n. It was concluded that an increase in N density at Si/SiON interface (at higher PNO dose[8]) is responsible for the observed behavior. It is proposed that the short-time very high, but weak, T dependent degradation, observed for devices having high N%, is due to ΔN_h [2], [3], [6]–[10], while degradation at longer stress time is due to $\Delta N_{\rm IT}$, as ΔN_h is likely to saturate at a time of much less than 1 s for the thickness of films used in this study [3], [6], [7]. The $\Delta N_{\rm IT}$ and ΔN_h components will be separated using the method described as follows.

The following are assumed: 1) the measured ΔV_T for all PNO samples consists of $\Delta V_{\rm IT}$ ($= q.\Delta N_{\rm IT}/CET$, CET being

[5] As ΔN_h de traps fast, its contribution to overall NBTI should be isolated to find the contribution of $\Delta N_{\rm IT}$ alone, and then the DC/AC de rating can be applied on $\Delta N_{\rm IT}$ data to determine ac NBTI lifetime. Capture of ΔN_h can be avoided by using delayed measurements, but would also bring in $\Delta N_{\rm IT}$ recovery related artifacts unless very long-time measurement is done.

[6] Bulk of the measurement data is presented in [10], PNO process details are in [26], and PNA details are described in [27].

[7] As shown in [30], mobility correction on OTF data can only be done on devices having low N%. Mobility correction cannot be done in this work as devices having a wide N% range have been used. Note that mobility correction impacts NBTI magnitude and not time exponent n, and therefore, it would not affect the main conclusion of this paper.

[8] N density at the SiON/poly-Si interface increases much larger than that at Si/SiON interface for higher PNO dose [31], but as shown in [10] using different types of nitridation processes, it is the Si/SiON N density that impacts NBTI most.

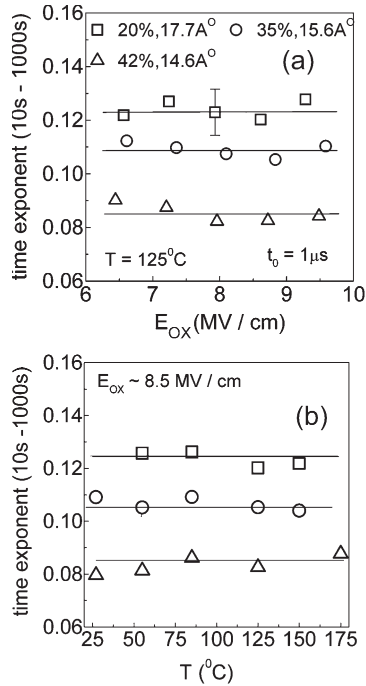

Fig. 2. Measured long-time power-law exponents as a function of (a) stress $E_{\rm OX}$ and (b) stress T for PNO devices having identical starting base oxide thicknesses but different N doses, resulting in different atomic N% values and EOTs.

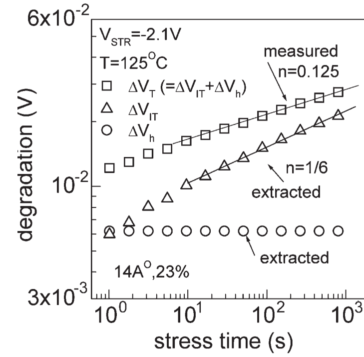

Fig. 3. Measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components according to the separation method proposed in this paper.

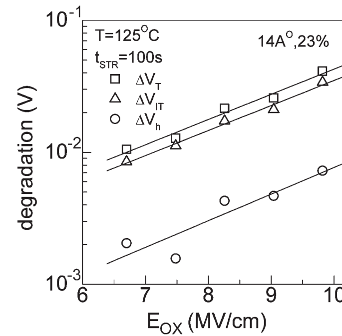

Fig. 4. $E_{\rm OX}$ dependence of measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components. Identical Γ's imply identical fractional $\Delta V_{\rm IT}$ and ΔV_h contributions as $E_{\rm OX}$ is varied.

capacitance equivalent thickness) and ΔV_h $(= q.\Delta N_h/CET)$, although their magnitude and relative dominance are determined by N density and thickness of the gate insulator; 2) the ΔV_h part saturates in less than 1 s for such thin insulators [3], [6], [7], and hence, its contribution to overall ΔV_T is constant for longer $(t > 1$ s) stress time; and 3) the $\Delta V_{\rm IT}$ part at longer stress time (equals the overall measured ΔV_T less constant ΔV_h component) follows power-law dependence with $n = 0.16$, as observed by very long-time measurements [11], [12], [15] and also as predicted by a simple solution of R–D model with H_2 diffusion [29].[9] Note that the procedure is different from [3], [6], and [16], where such separation was performed by assuming a long-time exponent of $n = 0.25$ (which is not consistent with E_A of the extracted $\Delta V_{\rm IT}$ and ΔV_h components, as discussed later). This is also different from [24], where the degradation obtained at a stress time of 1 s (assumed on an *ad hoc* basis to be due to ΔV_h only) is subtracted from the overall ΔV_T to find $\Delta V_{\rm IT}$. Finally, the method is also different from [25], where degradations at different measurement delays were subtracted from each other to find out ΔV_h, with the assumption of no $\Delta V_{\rm IT}$ recovery in the range of delay studied which cannot be justified.

Fig. 3 shows the time evolution of the measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components (using the aforementioned method) for a device having relatively lower N% stressed at a particular $E_{\rm OX}$ and T. Note that the measured ΔV_T shows a long-time $(t \geq 10$ s) exponent of $n \sim 0.12$ (which is independent of stress $E_{\rm OX}$ and T [10] that validates the robustness of the use of power law at long stress time), and a constant ΔV_h component is subtracted (note that ΔV_h is assumed to saturate at time less than 1 s) such that the resultant $\Delta V_{\rm IT}$ component shows long-time $n = 0.16$. Identical procedure is followed to isolate $\Delta V_{\rm IT}$ and ΔV_h for a wide range of stress $E_{\rm OX}$ and T, for all other devices (having different N% values and EOTs) used in this work.

Fig. 4 shows the $E_{\rm OX}$ dependence (at constant T) of the measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components. Note that the $E_{\rm OX}$ dependent slope (Γ) is similar (~ 0.46) for all three sets of data, implying that both $\Delta V_{\rm IT}$ and ΔV_h (and, hence, overall ΔV_T) are similarly influenced by $E_{\rm OX}$. This is an intuitive result because both components of NBTI degradation are linked to hole tunneling through the barrier at the Si/SiON interface,[10] as discussed in [8], [10], [11], and [19]. While Γ varies with Si/SiON N density [10], it always remains identical for ΔV_T and the extracted $\Delta V_{\rm IT}$ and ΔV_h components, which is true for all the devices (having different N% values and

[9]R–D model suggests the breaking of interfacial Si–H bonds and the subsequent diffusion of released H as responsible for $\Delta N_{\rm IT}$. A simple nondispersive solution of R–D model suggests $n = 0.5$ for H^+ drift/diffusion, $n = 0.25$ for H^0 diffusion, and $n = 0.16$ for H_2 diffusion. At long time and the range of T used, H_2 is the most likely diffusion species [32]–[34]. The nondispersive diffusion is also likely, as the bulk of the diffusion is expected to happen in the poly-Si for such ultrathin oxides [4], [35].

[10]Tunneling of inversion layer holes to interfacial Si–H bonds and bulk (N related) traps are responsible for $\Delta N_{\rm IT}$ and ΔN_h, respectively [7], [8], [10]. Inversion hole density and the tunneling barrier are affected by $E_{\rm OX}$, and hence, both $\Delta N_{\rm IT}$ and ΔN_h show identical $E_{\rm OX}$ dependence. The field acceleration of both $\Delta N_{\rm IT}$ and ΔN_h can be modeled similarly using the method described in [11] and [19].

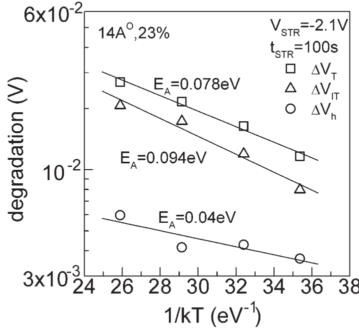

Fig. 5. T dependence of measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components. Higher E_A for $\Delta V_{\rm IT}$ compared to ΔV_h implies higher fractional $\Delta V_{\rm IT}$ contribution than ΔV_h as T is increased.

EOTs) used in this work and are not explicitly shown in this paper. Furthermore, since $\Delta V_{\rm IT}$ and ΔV_h show identical field dependence, both would reduce by the same amount, as $E_{\rm OX}$ is scaled from stress to operating condition. It can be shown that a finite ΔV_h when added to $\Delta V_{\rm IT}$ (ΔV_h and $\Delta V_{\rm IT}$ are scaled by the same ratio) would always result in lower (than 0.16) n of overall ΔV_T irrespective of stress $E_{\rm OX}$. This explains the $E_{\rm OX}$ independence of measured n, as shown in Fig. 2(a), which is expected to hold true for lower $E_{\rm OX}$ values close to operating condition.

Since $\Delta V_{\rm IT}$ depends strongly on T [5], [7]–[9] while ΔV_h does not [3], [16], exploring the T dependence of degradation provides an additional probe to explore the validity of the decomposition technique discussed earlier. Fig. 5 shows the T activation[11] for the measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components for a PNO device at long ($t \geq 10$ s) stress time under constant stress $E_{\rm OX}$. Note that E_A (at fixed stress time) for ΔV_T is found to be 0.078 eV. Once $\Delta V_{\rm IT}$ and ΔV_h components are separated by the algorithm discussed earlier, $\Delta V_{\rm IT}$ shows $E_A = 0.094$ eV (more on this later), while ΔV_h shows $E_A = 0.04$ eV (~3/2 KT and negligible). In other words, (essentially) T-insensitive ΔV_h contribution during initial stress phase ($t < 1$ s) lowers the E_A for overall ΔV_T at long stress time [7]. As mentioned before, the constant ΔV_h contribution at long stress time when added to $\Delta V_{\rm IT}(n \sim 0.16)$ reduces the n of overall ΔV_T. Unlike $E_{\rm OX}$ dependent slope Γ, which is identical for ΔV_h and $\Delta V_{\rm IT}$ and, hence, ΔV_T, the T dependent slope (E_A) for $\Delta V_{\rm IT}$ is much larger than ΔV_h. Therefore, it is expected that ΔV_h would dominate for very low stress T, and resultant n would become much lower than the one shown in Fig. 2(b). On the other hand, $\Delta V_{\rm IT}$ would dominate for very high stress T, and it is expected that long-time n would show a value close to $n \sim 0.16$. In other words, the observed T independence of n as shown in Fig. 2(b) may not hold if a significantly wider T range (than that used in this paper) is studied. Finally, note that similar separation based on a long-time exponent of $n = 0.25$ [3], [6], [16] shows $E_A \sim 0.06{-}0.07$ eV for ΔV_h (not explicitly shown in this paper), which cannot be physically justified without assuming a strong phonon-assisted hole-trapping process and also without assuming H^0 diffusion[12] that is unlikely at long stress time and at the range of T used in this paper [32]–[34].

[11] Measured time exponent n at long stress time is independent of T for all PNO devices [10], and hence Arrhenius T activation is used.

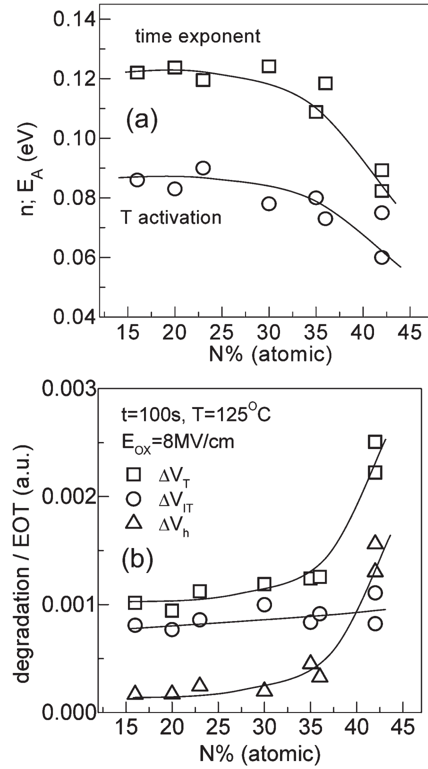

Fig. 6. Atomic N% dependence of (a) n and E_A of measured ΔV_T and (b) measured ΔV_T and extracted $\Delta V_{\rm IT}$ and ΔV_h components for a particular stress time T and $E_{\rm OX}$. Starting base oxide thickness and PNO N dose were varied to vary N%. Lines are guides to the eye.

The robustness of the separation method described earlier needs to be verified by using it over a wide range of samples having different N% values and EOTs. Fig. 6(a) shows the measured n and E_A for long-time ($t > 10$ s) stress ΔV_T data on devices having different N% values (obtained by varying the starting base oxide thickness from 15 to 25 Å and the PNO N dose from 2.8×10^{15} to 6.8×10^{15} cm^{-2}; see [10] for details). Note that n is independent of stress $E_{\rm OX}$ and T for all devices used in this study, and E_A was extracted from varying T but fixed $E_{\rm OX}$ (~8.5 MV/cm) stress. Both n and E_A reduce in a similar fashion with an increase in N% (with very rapid reduction seen for $N > 30\%$), which is due to an increase in N density at the Si/SiON interface [8], [10]. When considered together with N% dependence of ΔV_T transients in Fig. 1 (very large T independent degradation at short stress time for higher N%), it can be argued that the reduction in n and E_A of overall ΔV_T is due to relatively higher contribution of ΔV_h (large short-time degradation and weak T activation) as N% is increased. While $\Delta V_{\rm IT}$ also increases

[12] H^0 is released after the breaking of Si–H bonds. Dimerization can occur by one of the following reactions: $2H^0 = H_2$ (in the oxide bulk [19]), Si-H + H^0 = Si- + H_2 (in the Si/SiON interface or, most likely, in the SiON/poly-Si interface [4], [35]). The short-time diffusion of H^0 is indeed possible (mostly in the thin SiON bulk), but at longer stress time, the diffusing species is very likely to be H_2, and the bulk of the diffusion is likely to be in the poly-Si [4], [35].

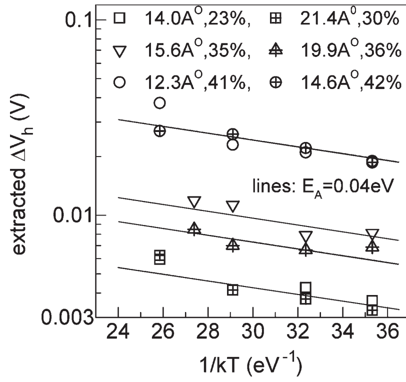

Fig. 7. T dependence of extracted ΔV_h component for devices having different N% values, obtained by varying starting base oxide thickness and PNO N dose. Lines are calculated using $E_A = 0.04$ eV.

with an increase in N% [36], the relative increase in ΔV_h must be higher for the aforementioned effect to hold true. To verify this, Fig. 6(b) shows the measured ΔV_T and extracted ΔV_{IT} and ΔV_h for identical stress time, E_{OX} and T for devices having different N% values. Indeed, ΔV_h increases at a much larger proportion than ΔV_{IT} and significantly accounts for the increase in magnitude and reduction in n and E_A of overall ΔV_T as N% becomes higher. It is important to note that, as Γ is identical for ΔV_{IT}, ΔV_h, and overall ΔV_T (see Fig. 4), the extracted ΔV_h fraction is independent of stress E_{OX}. However, as T activation is different for ΔV_{IT}, ΔV_h, and overall ΔV_T (see Fig. 5), fractional ΔV_h contribution is higher at lower T and vice versa, i.e., for a given device, ΔV_h contribution is more dominating at lower stress T, while relative domination of ΔV_{IT} is more at higher stress T.

Fig. 7 shows the extracted ΔV_h component at different stresses T for devices having different N% values (obtained by using different starting base oxide thicknesses and PNO N doses, as mentioned earlier). Although the magnitude of ΔV_h depends on EOT and N%,[13] it always shows very weak (and identical) T activation for all devices, which is consistent with the fact that hole trapping is a weak T activated process. Identical E_A's for the extracted ΔV_h component across such a wide range of SiON PNO processes point to the robustness of the extraction algorithm.

The T dependence of the extracted ΔV_{IT} contribution for different SiON processes is considered next. The extracted ΔV_{IT} data at long stress time at different T are parallel to each other on a log-log plot (identical time exponents $n = 0.16$) and hence can be scaled along X-axis (at constant ΔV_{IT}) and also along Y-axis (at constant stress time) to unique relations [7], [9], [18]. The X- and Y-axis scaling factors for ΔV_{IT} data on devices having different EOTs and slightly different N% values (similar PNO doses but different starting base oxide thicknesses) as functions of stress T are shown in Fig. 8.

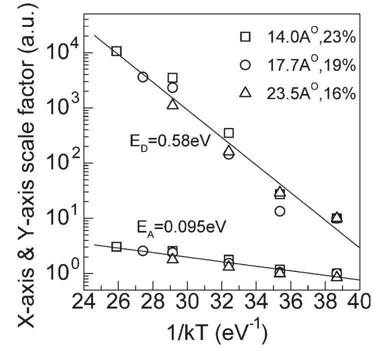

Fig. 8. Multiplicative factors obtained by X-axis scaling (at constant degradation) and Y-axis scaling (at constant time) of the extracted ΔV_{IT} components at different T's, for devices having different starting base oxide thicknesses and EOTs but similar PNO N doses.

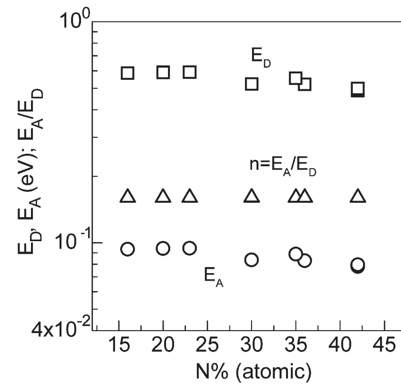

Fig. 9. Atomic N% dependence of E_A, E_D, and ratio of E_A to E_D of the extracted ΔV_{IT} component. See text for details.

According to simple nondispersive solution of the R–D model, the T dependence of X-axis scale factor (at constant ΔV_{IT}) denotes the T activation of diffusion (E_D) of the released H species,[14] while Y-axis scale factor (at constant stress time) is the usual T activation (E_A) of ΔV_{IT}. Note that the obtained E_D value is consistent with molecular H_2 diffusion [33], while the ratio of $E_A/E_D(= n)$ is also consistent with the prediction of R–D model. It is important to note that the obtained E_D from UF-OTF is identical to that obtained from charge pumping measurements [9], the later being a direct estimator of ΔN_{IT}[15] and therefore justifies the correctness of the separation procedure. To prove that the aforementioned observation is universal, Fig. 9 shows the obtained E_D, E_A, and the ratio $n = E_A/E_D$ for the extracted ΔV_{IT} component as a function of atomic N% (once again, obtained by varying the starting base oxide thickness from 15 to 25 Å and PNO N dose from 2.8×10^{15} to 6.8×10^{15} cm^{-2}) for all PNO devices. The consistency of the value of E_D (in the range of 0.5–0.6 eV) with H_2 diffusion [33] and the validity of the $n = \mathrm{E}_A/\mathrm{E}_D$ relation for PNO

[13] ΔV_h magnitude is determined by N density close to the Si/SiON interface, which is controlled by the overall N% and EOT. Interfacial N density is higher for higher N% (fixed EOT) and smaller EOT (fixed N%). Moreover, for a given Si/SiON N density, ΔV_h is likely to be higher for thicker EOT due to lower detrapping probability [7]. Although N% is a good measure of increased ΔV_h, an example of exception in terms of identical N% ($\sim$35%) but different EOTs due to difference in starting base oxide can be observed.

[14] R–D model solution [19]: $\Delta V_{\mathrm{IT}} = [k_F.N_0/k_R]^m[D_0.\exp(-E_D/kT)t]^n$, with $m = 0.67$ and $n = 0.16$ for molecular H_2 diffusion; k_F and k_R are forward and reverse (Si–H bond breaking and annealing) reaction rates; N_0 is the total Si–H bond density; D_0 is the T independent diffusion prefactor; and E_D is the activation of diffusion. Assuming similar T activation for k_F and k_R, time scaling at identical ΔV_{IT} can be used to extract E_D, while $E_A = E_D * n$.

[15] E_A for charge pumping is higher compared to OTF due to higher n as a result of measurement delay.

samples over a wide range of EOTs and N densities once again justify the robustness and validity of the decomposition scheme (the consistency of the obtained E_A for the ΔV_h component over a wide range of N% values has been shown earlier in Fig. 7).

Finally, to ensure that the decomposition is unique, the analysis has been repeated with the assumption that ΔV_{IT} component follows the power-law time dependence with $n \sim 0.25$, as predicted by H^0-only R–D model [3], [28], [29]. It has been found (not explicitly shown in this paper) that the E_D values obtained (in the range of 0.3–0.4 eV) as N% is varied are not consistent with the expected value (in the range of 0.1–0.2 eV) for H^0 [34]. Furthermore, E_D (for such extracted ΔV_{IT} component from UF-OTF) is not consistent with that obtained from charge pumping as well. Once again, this clearly establishes that the decomposition based on $n = 0.16$, i.e., H_2 diffusion at longer stress time, is unique and physically justified. Therefore, even though n and E_A of ΔV_T measured using UF-OTF for PNO devices have often been found to be lower compared to that predicted by a simple R–D model solution, once the component due to ΔV_h is accounted for, the n and E_A of the extracted ΔV_{IT} component at long stress time are broadly consistent with those predicted by H_2-based R–D model [19], [29], while n also remains consistent with very long time measurements [11], [12], [15], where the contribution due to ΔV_h would be negligible.

III. Conclusion

To summarize, a simple phenomenological scheme to isolate hole-trapping (ΔV_h) and interface trap generation (ΔV_{IT}) components from measured ΔV_T during NBTI is proposed. The robustness of the methodology is validated in PNO samples having a wide range of N% values, obtained by using different base oxide thicknesses and PNO N doses. A fast but quickly saturated ΔV_h component with negligible T activation gets added to ΔV_{IT} having power-law time dependence and strong T activation as N% is increased, which lowers the long-time n and E_A of overall ΔV_T measured using UF-OTF technique. However, once the fast ΔV_h component is subtracted from overall measured degradation, the resulting component (ΔV_{IT}), which is characterized in this paper by n, E_D, E_A, and ratio of E_A to E_D are fully consistent with those predicted by classical nondispersive H_2-based R–D model solution, with n being consistent with very long time measurements published in the literature. Moreover, since ΔV_h contribution saturates relatively quickly, the long term degradation (which ultimately leads to failure) should be dominated by ΔV_{IT} degradation, whose dynamics is captured within the existing theoretical framework of interface trap generation.

Acknowledgment

The authors would like to thank E. N. Kumar and S. Purawat for the development of UF-OTF setup; C. Olsen and K. Ahmed of Applied Materials, USA, for the useful discussions; and Applied Materials, USA, for the devices used in this work.

References

[1] Y. Mitani, “Influence of nitrogen in ultra-thin SiON on negative bias temperature instability under AC stress,” in *IEDM Tech. Dig.*, 2004, pp. 117–120.

[2] C. Shen, M. F. Li, C. E. Foo, T. Yang, D. M. Huang, A. Yap, G. S. Samudra, and Y. C. Yeo, “Characterization and physical origin of fast Vth transient in NBTI of pMOSFETs with SiON dielectric,” in *IEDM Tech. Dig.*, 2006, pp. 333–336.

[3] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, “Analysis of NBTI degradation- and recovery-behavior based on ultra fast VT measurements,” in *Proc. Int. Reliab. Phys. Symp.*, 2006, pp. 448–453.

[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,” in *IEDM Tech. Dig.*, 2005, pp. 688–691.

[5] 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,” in *IEDM Tech. Dig.*, 2005, pp. 684–687.

[6] K. Sakuma, D. Matsushita, K. Muraoka, and Y. Mitani, “Investigation of nitrogen-originated NBTI mechanism in SiON with high-nitrogen concentration,” in *Proc. Int. Reliab. Phys. Symp.*, 2006, pp. 454–460.

[7] S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, and M. A. Alam, “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?” in *Proc. Int. Reliab. Phys. Symp.*, 2007, pp. 1–9.

[8] E. N. Kumar, V. D. Maheta, S. Purawat, A. E. Islam, C. Olsen, K. Ahmed, M. Alam, and S. Mahapatra, “Material dependence of NBTI physical mechanism in silicon oxynitride (SiON) p-MOSFETs: A comprehensive study by ultra-fast on-the-fly (UF-OTF) I_{DLIN} technique,” in *IEDM Tech. Dig.*, 2007, pp. 809–812.

[9] S. Mahapatra and M. A. Alam, “Defect generation in p-MOSFETs under negative-bias stress: An experimental perspective,” *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 35–46, Mar. 2008.

[10] V. D. Maheta, C. Olsen, K. Ahmed, and S. Mahapatra, “The impact of nitride engineering on silicon oxynitride gate dielectric on negative bias temperature instability of p-MOSFETs: A study by using ultra-fast on-the-fly I_{DLIN} technique,” *IEEE Trans. Electron Devices*, vol. 55, no. 7, pp. 1630–1638, Jul. 2008.

[11] 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,” in *IEDM Tech. Dig.*, 2006, pp. 329–332.

[12] C. L. Chen, Y. M. Lin, C. J. Wang, and K. Wu, “A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2 nm ultra thin oxide,” in *Proc. Int. Reliab. Phys. Symp.*, 2005, pp. 704–705.

[13] S. Rangan, N. Mielke, and E. C. C. Yeh, “Universal recovery behavior of negative bias temperature instability,” in *IEDM Tech. Dig.*, 2003, pp. 341–344.

[14] M. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, J. Babcock, R. Lindley, P. Clifton, M. Redford, and A. Shibkov, “Dynamic recovery of negative bias temperature instability in p-type metal–oxide–semiconductor field-effect transistors,” *Appl. Phys. Lett.*, vol. 83, no. 8, pp. 1647–1649, Aug. 2003.

[15] A. Haggag, G. Anderson, S. Parihar, D. Burnett, G. Abeln, J. Higman, and M. Moosa, “Understanding SRAM high-temperature-operating-life NBTI: Statistics and permanent vs. recoverable damage,” in *Proc. Int. Reliab. Phys. Symp.*, 2007, pp. 452–456.

[16] V. Huard, C. Parthasarathy, T. Guerin, T. Valentin, E. Pion, M. Mammasse, N. Planes, and L. Camus, “NBTI degradation: From transistor to SRAM arrays,” in *Proc. Int. Reliab. Phys. Symp.*, 2008, pp. 289–300.

[17] B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, “Disorder controlled kinetics model for negative bias temperature instability and its experimental verification,” in *Proc. Int. Reliab. Phys. Symp.*, 2005, pp. 381–387.

[18] 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 Trans. Electron Devices*, vol. 51, no. 9, pp. 1371–1379, Sep. 2004.

[19] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, “Recent issues in negative-bias temperature instability: Initial degradation, field dependence of interface trap generation, hole trapping effects,

and relaxation," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2143–2154, Sep. 2007.

[20] T. Grasser, B. Kaczer, and W. Goes, "An energy-level perspective of bias temperature instability," in *Proc. Int. Reliab. Phys. Symp.*, 2008, pp. 28–38.

[21] V. Huard, M. Denais, F. Perrier, N. Revil, C. Parthasarathy, A. Bravaix, and E. Vincent, "A thorough investigation of MOSFETs NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 83–98, Jan. 2005.

[22] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative bias temperature instability," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 1, pp. 22–34, Mar. 2008.

[23] T. L. Yang, M. F. Li, C. Shen, C. H. Ang, Z. Chunxiang, Y. C. Yeo, G. Samudra, S. C. Rustagi, and 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," in *VLSI Symp. Tech. Dig.*, 2005, pp. 92–93.

[24] A. Neugroschel, G. Bersuker, R. Choi, C. Cochrane, P. Lenahan, D. Heh, C. Young, C. Y. Kang, B. H. Lee, and R. Jammy, "An accurate lifetime analysis methodology incorporating governing NBTI mechanisms in high-*k*/SiO_2 gate stacks," in *IEDM Tech. Dig.*, 2006, pp. 1–4.

[25] J. H. Lee, W. H. Wu, A. E. Islam, M. A. Alam, and A. S. Oates, "Separation method of hole trapping and interface trap generation and their roles in NBTI reaction–diffusion model," in *Proc. Int. Reliab. Phys. Symp.*, 2008, pp. 745–746.

[26] P. A. Kraus, K. Z. Ahmed, C. S. Olsen, and F. Nouri, "Physical models for predicting plasma nitrided SiON gate dielectric properties from physical metrology," *IEEE Electron Device Lett.*, vol. 24, no. 9, pp. 559–561, Sep. 2003.

[27] C. Olsen, "Two-step post nitridation annealing for lower EOT plasma nitrided gate dielectrics," U.S. Patent 017 596 1A1, Sep. 9, 2004.

[28] M. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 71–81, Jan. 2005.

[29] S. Chakravarthi, A. T. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability," in *Proc. Int. Reliab. Phys. Symp.*, 2004, pp. 273–282.

[30] A. E. Islam, V. D. Maheta, H. Das, S. Mahapatra, and M. A. Alam, "Mobility degradation due to interface traps in plasma oxynitride PMOS devices," in *Proc. Int. Reliab. Phys. Symp.*, 2008, pp. 87–96.

[31] S. Rauf, S. Lim, and P. L. G. Ventzek, "Model for nitridation of nanoscale SiO_2 thin films in pulsed inductively coupled N_2 plasma," *J. Appl. Phys.*, vol. 98, no. 2, pp. 024 305-1–024 305-10, Jul. 2005.

[32] D. Griscom, "Diffusion of radiolytic molecular hydrogen as a mechanism for the post-irradiation buildup of interface states in SiO_2-on-Si structures," *J. Appl. Phys.*, vol. 58, no. 7, pp. 2524–2533, Oct. 1985.

[33] M. L. Reed and J. D. Plummer, "Chemistry of Si–SiO_2 interface trap annealing," *J. Appl. Phys.*, vol. 63, no. 12, pp. 5776–5793, Jun. 1988.

[34] B. Tuttle, "Energetics and diffusion of hydrogen in SiO_2," *Phys. Rev. B, Condens. Matter*, vol. 61, no. 7, pp. 4417–4420, Feb. 2000.

[35] S. Chakravarthi, A. T. Krishnan, V. Reddy, and S. Krishnan, "Probing negative bias temperature instability using a continuum numerical framework: Physics to real world operation," *Microelectron. Reliab.*, vol. 47, no. 6, pp. 863–872, Jun. 2007.

[36] 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," in *Proc. Solid State Devices Mater.*, 2003, pp. 70–71.

Souvik Mahapatra (M'02–SM'07) received the Ph.D. degree in electrical engineering from the Indian Institute of Technology Bombay (IIT Bombay), Mumbai, India, in 1999.

From 2000 to 2001, he was with Bell Laboratories, Murray Hill, NJ. Since 2002, he has been with the Department of Electrical Engineering, IIT Bombay, where he is currently an Associate Professor. He has published more than 70 papers in peer-reviewed international journals and conferences, delivered invited talks at major international conferences including the IEDM, given tutorials at IRPS, and served as a Reviewer for several international journals and conferences. His research interests are CMOS device and Flash memory reliability.

Vrajesh D. Maheta (M'06) received the B.E. degree in electronics from Sardar Patel University, Gujarat, India, in 1993 and the M.E. degree in microelectronics from Birla Institute of Technology and Science, Pilani, India, in 2002. He is currently working toward the Ph.D. degree in electrical engineering in the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India.

His research interests are semiconductor device physics and simulation, modeling, and characterization of CMOS silicon devices. He is working on the gate insulator reliability issues, particularly bias temperature instability in advanced CMOS silicon devices.

Ahmad Ehteshamul Islam (S'02) received the B.S. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2004. He is currently working toward the Ph.D. degree in the School of Electrical Engineering and Computer Science, Purdue University, West Lafayette, IN.

During 2004–2005, he was a Lecturer with the Department of EEE, BUET. His research mainly focuses on studying the impact of variabilities in semiconductor devices. He is currently working with BTI instabilities in nanoscale transistors. He has authored and coauthored more than 15 journals and conference papers.

Mr. Islam has been a Student Member of the IEEE Electron Devices Society (since 2002) and American Physical Society (since 2008) and also serves as a Reviewer for several IEEE and Elsevier journals. He is the recipient of Kintar-Ul-Haque Gold Medal for his undergraduate result in 2005 and IEEE EDS Ph.D. Fellowship for his work on transistor reliability in 2008.

Muhammad Ashraful Alam (F'05) received the B.S.E.E. degree from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 1988, the M.S. degree in electrical engineering from Clarkson University, Potsdam, NY, in 1991, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1994.

From 1995 to 2001, he was with Bell Laboratories, Lucent Technologies, Murray Hill, NJ, as a Member of the Technical Staff in the Silicon ULSI Research Department. From 2001 to 2003, he was a Distinguished Member of the Technical Staff with Agere Systems, Murray Hill. Since 2004, he has been with Purdue University, where he is currently a Professor of electrical and computer engineering in the School of Electrical Engineering and Computer Science, with his research and teaching focusing on physics, simulation, characterization, and technology of classical and emerging electronic devices, including theory of oxide reliability, nanonet thin-film transistors, and nanobiosensors. He also holds the position of distinguished University Scholar at Purdue University. He has published over 90 papers in international journals and has presented many invited and contributed talks at international conferences.

Dr. Alam is the recipient of the 2006 IEEE Kiyo Tomiyasu Award for his contributions to device technology for communication systems.