

# Mechanism of Dynamic NBTI of pMOSFETs

B. Zhu<sup>1</sup>, J. S. Suehle<sup>2</sup>, J. B. Bernstein<sup>1</sup> and Y. Chen<sup>3</sup>

<sup>1</sup> University of Maryland, College Park, MD 20742

<sup>2</sup> National Institute of Standards and Technology, Gaithersburg, MD 20899

<sup>3</sup> Jet Propulsion Laboratory, Pasadena, CA 91109

**Abstract:** Negative bias temperature instability of pMOSFETs is studied. It is found that there is strong frequency dependence of the parameters shift. At a certain high frequency, the threshold voltage shift is only about half of that generated by dc stress. The possible sources of this dependence are explored. An empirical model is established based on the reduction of fixed oxide charges. This model is further used to explain some observed phenomena.

**Index Terms:** Negative bias temperature instability, pMOSFETs, threshold voltage, interface traps.

## I. INTRODUCTION

As the trend of scaling of MOSFETs goes further, NBTI becomes a more prominent issue in the current VLSI technology. It may become one of the ultimate limiting factors since NBTI is more severe than hot carrier stress for ultrathin oxides at low electric fields. The degradation of NBTI includes the absolute decrease of drain drive current  $I_{on}$  and transconductance  $g_m$  and increase of threshold voltage  $V_{th}$ . Compared with other type of devices; pMOSFETs are especially vulnerable to NBTI [1]. In MOS circuits, NBTI mainly occurs during the on state of inverter operation. The decrease in drain current will lead to significant timing issues. The degradation of threshold voltage will lead to reduced current and frequency degradation of ring oscillators and reduced standard random access memory noise margin [2]. Recent studies have shown that a fraction of NBTI degradation can be annealed if the applied gate voltage is removed [3-6]. It was speculated that this kind of annealing effect came from two factors, the interface trap and fixed oxide charge reductions. Some models are established to explain this phenomenon [7]. In this work, we conduct detailed experiments and show that this annealing process can be explained by the reduction of fixed oxide charge generated by NBTI stress, and this reduction is the dominating factor.

## II. EXPERIMENT DESCRIPTION

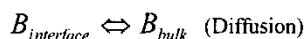
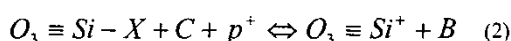
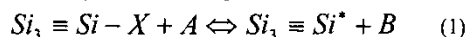
Surface channel dual-gate p-MOSFETs with 2.0 nm and 5.0 nm gate oxide thicknesses are used in this study. The channel width is 10.0  $\mu\text{m}$ , and the length is 0.25  $\mu\text{m}$ .

Constant voltage stresses are applied to the gate under dc and unipolar-pulsed stresses with source, drain, and substrate grounded. For the pulsed stress, the amplitude is identical to

the dc stress voltage during the pulse "on" phase and was zero during the pulse "off" phase. The duty cycle is maintained at 50 %. The stress was periodically interrupted and the drain saturation current  $I_{on}$  was measured immediately.

## III. EXPERIMENTAL RESULTS

During the NBTI stress, the sources that lead to drain drive current,  $I_{on}$ , and threshold voltage,  $V_{th}$ , shifts are the fixed oxide charges and the interface trapped charges [8]. According to different models proposed for the NBTI generation mechanisms, the generation of fixed oxide charges and interface traps could be expressed as [9-11]:



$Si_3 \equiv Si - X$  is a trivalent silicon, a defect near the interface. X, A, B, and C are unknown neutral species.  $Si_3 \equiv Si^*$  is a neutral trivalent silicon backbonded by three other silicon atoms, functioning as an interface trap near the interface. The  $\equiv$  represents three complete bonds to other Si atoms and the  $\bullet$  represents the fourth, unpaired electron in a dangling bond. Interface traps are also known as  $P_b$  centers [12]. Being electrically active defects, interface traps act as generation/recombination centers and contribute to the leakage current and reduced mobility, drain current and transconductance. Electrons or holes can occupy interface traps, thus leads to the threshold voltage shift, given by

$$\Delta V_{th} = -\frac{\Delta Q_{it}(\phi_s)}{C_{OX}}$$

$\phi_s$  is the surface potential,  $C_{OX}$  is the oxide capacitance per unit area.

$O_3 \equiv Si^+$  is a positively charged species, serving as a positive fixed charge in the oxide. It may take place at the interface or in the oxide close to the interface. Similar to interface traps, fixed oxide charges will also lead to threshold voltage shift by

$$\Delta V_{th} = -\frac{\Delta Q_f}{C_{OX}}$$

$I_{on}$  and  $V_{th}$  have similar shift behavior with the NBTI stress, which can be explained by the simplest relation between them:

$$I_{on} = \frac{W}{2L} \mu C_{OX} (V_G - V_{th})^2.$$

$W$  and  $L$  are the device channel width and length respectively,  $\mu$  is the mobility of minority carriers, and  $V_G$  is the gate voltage. However,  $I_{on}$  and  $V_{th}$  are not simply square rootly related since there is also mobility degradation during the NBTI stress.

Most of the tests in this work are taken by measuring  $I_{on}$ . However, the evaluation of the  $V_{th}$  shift is also performed.  $V_{th}$  is extrapolated from the linear region of the drain current vs. gate voltage ( $I_d$ - $V_g$ ) measurement [13]. The interface traps density change is monitored by  $I_{cp}$ , which is the peak current of the conventional charge pumping method [14].

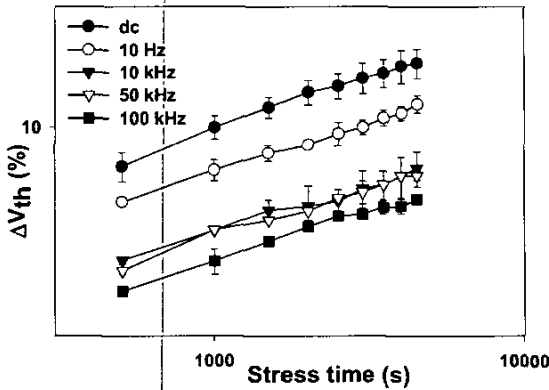


Fig. 1. The frequency dependence of  $\Delta V_{th}$ . Stress condition:  $V_{stress} = -3.0$  V,  $T = 150$  °C.

Fig. 1 shows the  $\Delta V_{th}$  under negative unipolar voltage stress with pulse repetition frequencies up to 100 kHz. According to this figure,  $\Delta V_{th}$  exhibits significant frequency dependence.  $\Delta V_{th}$  decreases as the frequency increases and with a frequency of 500 kHz or higher,  $\Delta V_{th}$  is only about half of that under dc stress condition, which is also observed by the other researchers [3]. Compared with  $V_{th}$ ,  $I_{on}$  shows the similar dependence on frequency, as shown in Fig.2.

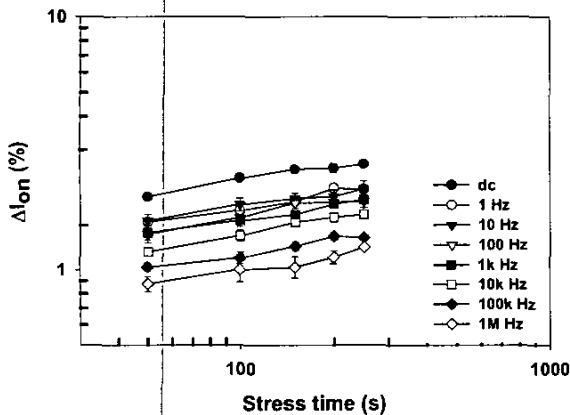


Fig. 2. The frequency dependence of  $\Delta I_{on}$ . Stress condition:  $V_{stress} = -3.0$  V,  $T = 150$  °C.

IV. DISCUSSION

One of the possible explanations for the frequency dependence is that the damage caused by NBTI during the stress on period is reduced or annealed during the stress off period.

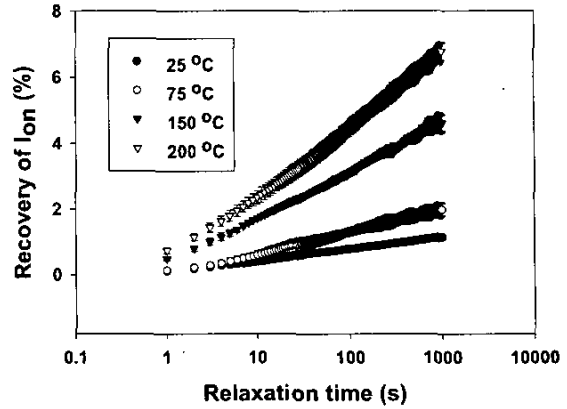


Fig. 3. The recovery of  $I_{on}$  at different temperatures. Stress condition:  $V_g = -2.0$  V,  $V_d = -0.05$  V, total stress time = 1000 s.

We monitored the recovery of  $I_{on}$  in detail as a means to evaluate the damage caused by NBTI after the stress termination, as shown in Fig.3. Although the monitoring of  $V_{th}$  is more preferable, however we cannot conduct that test as fast as  $I_{on}$ , since  $I_d$ - $V_g$  measurement takes longer time. As pointed by some reports, the time between the measurements is critical for the study of NBTI [15]. The data in this plot are taken every 1 s after the stress terminated.

Apparently, the  $I_{on}$  recovery data at each test could not fit a simple straight line, which implies the existence of multiple mechanisms. This is further confirmed by the extrapolated activation energy using the data in Fig.3 as shown in Fig. 4. A chemical reaction process is related with unique activation energy which is the minimal energy needed.

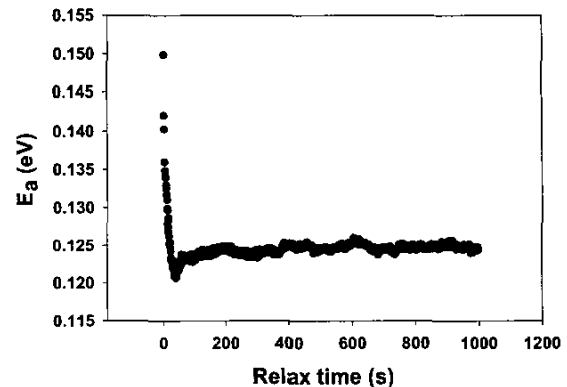


Fig. 4. The activation energies extrapolated from data in Fig. 3. Two stages are observed.

Obviously, the activation energy associated with  $I_{on}$  recovery is much higher right after the NBTI stress terminated and saturates later, which suggests the existence of at least two different processes. The first process has higher activation energy, faster annealing rate and happens earlier. The second process has lower activation energy, slower annealing rate and dominates the later period. The first process may happen so fast that we can't get the exact value due to the equipment time response limit. The observed decreasing period should be the transient period from the first process to the second one.

Since both reactions (1) and (2) can be reversible, the recovery of NBTI could be attributed to the annealing of either fixed oxide charges or interface traps, or a combination of the two.

At first, the annealing of interface traps is discussed. It has been reported that different from NBTI, which includes both interface trap generation and positive charge formation, the PBTI only results in interface trap generation, and the final saturation level is fixed by the number of defects in the device [16]. Thus, the study of the frequency dependence of PBTI will reveal the fact if there is annealing of interface traps, as shown in Fig.5. This shows that there is almost no frequency dependence of  $V_{th}$  with PBTI, at least in the frequency range from dc to 10k Hz. This suggests that the annealing of PBTI, or interface traps during the stress off period within each pulse is much smaller or negligible. The interface traps generated by PBTI and NBTI should be the same as identical devices are tested.

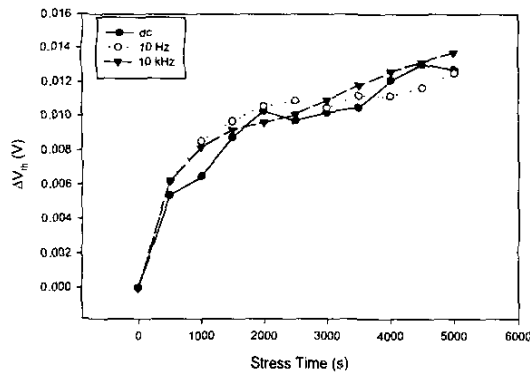


Fig. 5.  $\Delta V_{th}$  under positive unipolar stress conditions with different frequencies,  $V_{stress}=+3.0$  V,  $T=150$  °C.

Additional tests are performed to study the annealing of interface traps during the stress off period as shown by Fig. 6.

In this figure, we compare the annealed fraction of  $\Delta I_{cp}$  and  $\Delta V_{th}$ . They are defined as:

$$\Delta V_{th(t)} = \frac{V_{th(t)} - V_{th(1000)}}{V_{th(1000)} - V_{th(0)}} \times 100\%$$

$$\Delta I_{cp(t)} = \frac{I_{cp(t)} - I_{cp(1000)}}{I_{cp(1000)} - I_{cp(0)}} \times 100\%$$

Here,  $V_{th(1000)}$  is the threshold voltage at 1000 s, at which the stress is turned off,  $V_{th(t)}$  is the threshold voltage measured at

time  $t$ ,  $V_{th(0)}$  is the initial threshold voltage for a fresh device.  $I_{cp}$  is defined in a similar way.

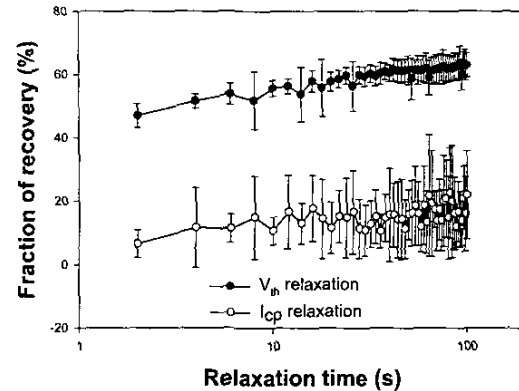


Fig. 6. The annealing of  $\Delta I_{cp}$  and  $\Delta V_{th}$ .

Stress condition:  $V_g=-3$  V,  $T=75$  °C, total stress time =1000 s.

If the annealing is mainly associated with interface traps, the annealing of  $\Delta I_{cp}$  will be comparable to that of  $\Delta V_{th}$ . However, we see the annealing of  $\Delta V_{th}$  is much larger, thus the annealing of fixed oxide charges is dominating. We believe the earlier recovery shown in Fig. 1 is related with the annealing of fixed oxide charge, which is the product mainly from a chemical reaction at the interface, while the later process is because of the annealing of interface traps, as it is associated with diffusion process, which happens at a lower rate.

It has been reported that fixed oxide charges could be reduced or detrapped with a positive gate bias following the NBTI stress [17]. Based on reaction (1), one would argue interface traps would show similar behavior. One provable test is the bipolar stress test in which the device is stressed with negative gate bias and positive gate bias one after another. If both fixed oxide charges and interface traps will reduce with an alternative gate bias stress, the resulted  $\Delta V_{th}$  will reduce consequently. However, Fig.7 shows that  $\Delta V_{th}$  under bipolar stress is larger. This means that there is not only no reduction but also the amount of increase of  $V_{th}$  induced by increased interface traps during the positive bias period could be larger than the reduced amount caused by reduction of fixed oxide charges. This may further suggest that alternative gate bias stress will probably not reduce the degradation caused by NBTI.

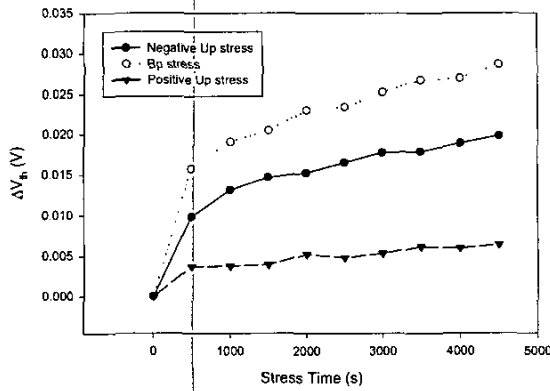


Fig. 7.  $\Delta V_{th}$  under different bias stress conditions with a 10K Hz frequency,  $V_{stress}=+3.0/-3.0$  V,  $T=150$  °C.

As discussed earlier,  $\Delta V_{th}$  consists the change of the fixed oxide charge and interface traps. Thus the reduction of  $\Delta V_{th}$  could only come from the reduction or annealing of fixed oxide charge. Recent studies have also confirmed that during the NBTI stress, the interface traps density is almost independent of the gate voltage frequency [17].

Following this argument, we can establish a model to explain the frequency dependence of NBTI. The parameter shift, such as  $V_{th}$  shift, could be expressed as:

$$\Delta V_{th(f)} = \frac{1}{C_{OX}} [\Delta Q_{it(dc)} + \Delta Q_{f(f)}] \quad (3).$$

$\Delta V_{th(f)}$  is  $V_{th}$  shift with unipolar stress,  $\Delta Q_{it(dc)}$  is the interface traps change with dc stress,  $\Delta Q_{f(f)}$  is the fixed oxide charge changes with unipolar stress and is a function of gate voltage frequency.

For a unipolar stress condition, the fixed oxide charge density will increase during the stress on period ( $t_{on}$ ), and decrease during the stress off period ( $t_{off}$ ), as the generation of fixed oxide charge can be reversible. By assuming the first order kinetics, we can create the following equations.

During the  $n^{th}$  stress off period:

$$Q_{min}^n = Q_{max}^n e^{-\frac{t_{off}}{\tau_{off}}}, \quad t_{on} + t_{off} = \frac{1}{f}.$$

During the  $n+1^{th}$  stress on period:

$$Q_{max}^{n+1} = Q_{min}^n + (Q_{f(f)} - Q_{min}^n)(1 - e^{-\frac{t_{on}}{\tau_{on}}}).$$

Here  $Q_{max}$  and  $Q_{min}$  are the upper trace and lower trace of  $Q_{f(f)}$  respectively,  $f$  is the frequency of the gate voltage,  $\tau_{on}$  and  $\tau_{off}$  are time constants for the stress on and off period respectively.

After a certain time, the system will be in a quasi-equilibrium condition, that is  $Q_{max}^n = Q_{max}^{n+1}$ . Assuming  $Q_{max}$  to be  $Q_{f(f)}$ , we have

$$Q_{f(f)} = \frac{Q_{f(dc)} \left( 1 - e^{-\frac{t_{on}}{\tau_{on}}} \right)}{1 - e^{-\frac{t_{off}}{\tau_{off}}} e^{-\frac{t_{on}}{\tau_{on}}}}.$$

To simplify the problem by assuming there is no mobility degradation during the NBTI stress, we can directly use this model to fit the  $\Delta I_{on}$  data we get at different frequencies, as in Fig.8. All the data in this figure are normalized to the dc stress data.

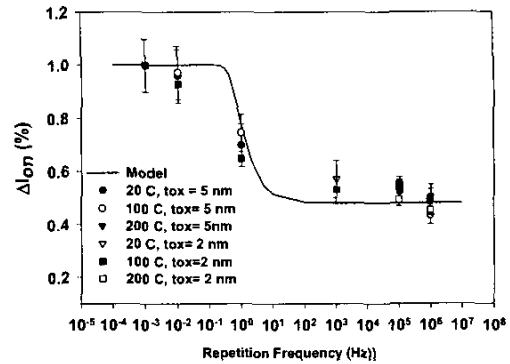


Fig. 8. The model fitting of  $\Delta I_{on}$  with different stress conditions and different oxide thickness.

By this model, we find the parameter shift caused by NBTI will be largely reduced in the unipolar stress conditions. And this kind of reduction will almost saturate at higher repetition frequencies. This can also explain the reason some reports didn't observe the independence at higher frequencies [3]. The fit parameters are 0.4 s and 1 s for  $\tau_{on}$  and  $\tau_{off}$  respectively.

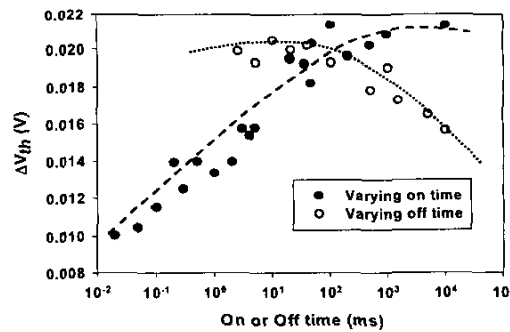


Fig. 9. The measured time constants related with the stress on and stress off period.

In order to experimentally determine the time constants  $\tau_{on}$  and  $\tau_{off}$ , we conduct two tests. In these tests, the pulse on (off) period is set to be 5 ms and the pulse off (on) period varies from 2  $\mu$ s up to 10<sup>4</sup> ms. The total effective (DC) stress time is 1000 s for both tests. The  $\Delta V_{th}$  is plotted against the pulse frequency as in Fig. 9. Based on these tests,  $\tau_{on}$  will be at the order of 100 ms, and  $\tau_{off}$  will be at least one order higher than  $\tau_{on}$ , which is consistent with our model.

## V. CONCLUSION

The root course of the NBTI reduction with unipolar stress was discussed. It was found the annealing of NBTI under unipolar stress is mainly due to the annealing of fixed oxide charge. The annealing of interface traps is a much slower process compared with that of fixed oxide charges. A model based on the fixed oxide charge generation and annealing is established to explain the reduced NBTI under the unipolar stress conditions.

## ACKNOWLEDGMENT

The author would like to thank Dr. Eric Vogel and Dr. Jinping Han for helpful discussions.

## REFERENCES

- [1] M. Makabe, T. Kubota, and T. Kitano, "Bias-temperature degradation of pMOSFETs: mechanism and suppression," in IEEE Int. Reliability Phys. Symp. 2000, pp. 205-209.
- [2] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability," in IEEE Int. Reliability Phys. Symp. 2002, pp. 248-254.
- [3] G. Chen, et al. "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling," IEEE EDL, Vol. 23 p. 734 (2002).
- [4] B. Abadeer, et al, "Behavior of NBTI under AC dynamic circuit conditions," Proc. IRPS, p. 17 (2003).
- [5] S. Rangan, "Universal recovery behavior of negative bias temperature instability [PMOSFETs]," IEDM 2003
- [6] B. S. Doyle, B. J. Fishbein, and K. R. Mistry, "NBTI-enhanced hot carrier damage in p-channel MOSFETs," in IEEE IEDM Tech. Dig., 1991, pp. 529 - 532A.
- [7] M. Alam, "A critical examination of the mechanics of dynamic NBTI for PMOSFETs," IEDM 2003.
- [8] D. K. Schroder, J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," J. of applied physics. vol. 94, 2003, pp.1-18.
- [9] C. E. Blat, E. H. Nicollian, and E. H. Poindexter, "Mechanism of negative-bias-temperature instability," J. Appl. Phys. vol. 69, 1991, pp. 1712.
- [10] K. O. Jeppson and C. M. Svensson, "Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices," J. Appl. Phys., vol. 48, pp. 2004, 1977.
- [11] S. Ogawa, M. Shimaya, and N. Shiono, "Interface-trap generation at ultrathin SiO<sub>2</sub> (4-6nm)-Si interface during negative-bias temperature aging," J. Appl. Phys., 77, 1137, (1995).
- [12] E. H. Poindexter and P. J. Caplan, Prog. Surf. Sci. 14, 201-294, 1983.
- [13] Sze. Physics of Semiconductor Devices, Wiley-Interscience, New York, 1981.
- [14] S. C. Wiczak, J. S. Suehle, and M. Gaitan, "An experimental comparison of measurement techniques to extract Si-SiO<sub>2</sub> interface trap density," Solid-State Electron. 35, 345, (1992).
- [15] G. La Rosa, IRPS 2003 tutorial.
- [16] J. F. Zhang and W. Eccleston, "Positive Bias Temperature Instability in MOSFET's," IEEE Trans. on Elec. Devices, vol. 45, no. 1. (1998)
- [17] V. Huard, and M. Denais, "Hole trapping effect on methodology for dc and ac negative bias temperature instability measurements in pMOS transistors," in Proc. of IEEE IRPS, pp. 40-45, 2004.

## QUESTIONS AND ANSWERS

Q: The slope appears dependent on temperature (reference page NBT2-6, graph 11) which is inconsistent with a diffusion-limited process, is something else going on?

Q: Were plots (reference page NBT2-6, graph 11) also made in log-log scale? In log-log scale, are different slopes seen? if log-log plots are parallel, could this indicate diffusion?

A: I will answer these two questions here since they are closely related. The reason for the different slopes at different temperature is that we plotted the data in the linear-logarithm scale for in the clear view. If we change the scale into the logarithm-logarithm scale, the slopes are quite similar after the initial first step, which is consistent with the diffusion-limited process. However, as the paper suggested, we think the first step is more important for the model, and the first step is based on the reaction process not the diffusion process.

Q: Is the relaxation measurement (reference page NBT2-6, graph 11) taken at stress temperature?

A: Yes. Since we couldn't cool the device in a very short time from high temperature to room temperature, those test are taken at the stress temperature.

Q: Is Ion temperature-dependent?

A: Yes, Ion is dependent on temperature as the mobility, threshold voltage, etc. However, when we compare the degradation of Ion, we compare them with the relative percentage change, which has eliminate temperature factor.

Q: If you add the two stress component plots (negative, positive) do they sum to the total stress plot (bipolar, reference NBT2-4, graph 8)?

A: The sum is quite close to the bipolar stress.