High Temperature Annealing of the Interface State Component of Negative-Bias Temperature Instability (NBTI) in MOSFET Devices

D. D. Nguyen[a], K. E. Kambour[b], C. Kouhestani[a], H. P. Hjalmarson[c] and R. A. B. Devine[d]

[a] COSMIAC, Albuquerque, NM 87106
[b] SAIC, Albuquerque, NM 87106
[c] Sandia National Laboratories, Albuquerque, NM 87185-1179, USA
[d] Think Strategically, Albuquerque, NM 87111

We have performed negative bias instability (NBTI) measurements on 130 nm channel length Si MOSFETs with SiON gate dielectrics at different temperatures up to 220 °C. Using a pulsed stressiing and pseudo-DC measurements we have focused on the charging and annealing kinetics of only the interface state related component. At temperatures > 180 °C we find clear evidence for a diminution of the growth kinetics suggesting the onset of an annealing effect during interface state generation. We contrast the NBTI data with interface state growth in devices exposed to ionizing radiation. Possible physical mechanisms contributing to the "annealing" effect are discussed and associated with diffusing H_2 molecules present in the substrate/gate insulator/polycrystalline Si gate structure.

Introduction

Negative bias temperature instability (NBTI) is a degradation mechanism present primarily in p-channel metal-oxide-semiconductor field effect transistors (MOSFETs) in inversion, though it can also be seen in n-channel MOSFETS in accumulation. Under normal PMOS operation [1] a negative potential is applied to the gate contact with respect to the substrate, and a hole inversion layer is induced in the nominally n-type region immediately under the gate insulator between the source and drain contacts. Experimental observation [2] has shown that extended negative biasing in such structures results in degradation of the device working parameters and in particular the threshold voltage, V_{th} [1]. In the extreme, this process ultimately leads to changes in the circuit behaviour and, finally, to failure. At this point it is necessary to underline that "failure" in itself is something requiring definition since each device characteristic (threshold voltage, mobility,…) or chain of devices (pulse delay, ..) may have a different criteria. We limit ourselves here to individual devices and the particular characteristic, threshold voltage. In order to develop predictive models for this failure, a deeper physical understanding of the trapped charges involved and their creation/elimination kinetics is required. Though it is generally recognized that NBTI is associated with positive charging at the semiconductor/insulator interface and in the "bulk" of the insulator, there is significant debate [3-4] as to the physics involved and the mechanism by which charging or discharging of the different trapping sites takes place. Previously [1], we reported on the development of an experimental protocol that, given certain approximations, enables

independent extraction of the various components of NBTI-induced charging. This works by applying pseudo-DC electrical stressing methods together with a technique in which stress is applied in a pulsed manner with a specifically chosen on/off duty cycle. Using this methodology has enabled us to extract the three distinct components of charging-induced threshold voltage variation: $\Delta V_{th(IS)}$ associated with interface states, $\Delta V_{th(RC)}$ associated with insulator charge which spontaneously recovers after removal of the negative bias by charge tunnelling into the substrate and $\Delta V_{th(FRC)}$ associated with charges in the insulator which are neutralized under the application of a positive oxide field but will quickly recharge under a small negative bias. There remains significant debate as to the permanent nature of $\Delta V_{th(IS)}$. Based on our results further work has been performed to separate out and to explore the temperature dependence of only the IS charge component and to evidence the high temperature annealing of such charge if present..

Experimental

Measurements were performed on p-channel MOSFETs produced using a proprietary IBM process on bulk Si wafers with a channel length of 130 nm and channel width of 5 μm. The gate insulator was 3.4 nm thick SiON. The electrical measurements were performed using a Keithley Instruments Inc. 4200 Semiconductor Characterization System with rapid data acquisition facility. Devices were probed using tungsten tips at the wafer level. The measurements were made at multiple temperatures in the range of room temperature to 220 °C.

Initially the source-drain current (I_{ds}) was measured as a function of applied gate-source voltage (V_{gs}) with a small applied source-drain voltage (V_{ds}) chosen so the device was operating in the linear regime [5]. V_{ds} was then typically – 50 mV. The absolute threshold voltage, V_{th}^{o}, was then determined from the $I_{ds}(V_{gs})$ plot by locating the V_{gs} value ($V_{gs(max.)}$) for maximum slope of dI_{ds}/dV_{gs} then extrapolating the line of slope $dI_{ds}/dV_{gs}\big|_{max}$. touching the curve at $V_{gs(max.)}$ back to the x axis. In this regime $I_{ds}(V_{gs})$ is approximately expressed as:

$$I_{ds} = (W/L)\mu C_{ox}(V_{gs} - V_{th})\, V_{ds} \qquad [1]$$

where W and L are the width and length of the conducting inversion channel, μ is the carrier mobility and C_{ox} is the gate insulator capacitance. For small changes in V_{th} one can to first order avoid performing the full $I_{ds}(V_{ds})$ measurement and approximate the shift in V_{th}, ΔV_{th}, as:

$$\Delta V_{th} = (1 - I_{ds}/I_{ds}^{o})(V_{gs(m)} - V_{th}^{o}) \qquad [2]$$

where I_{ds}^{o} and V_{th}^{o} are initial time (t = 0) values prior to any NBTI stressing. I_{ds} is the value of the current at a measurement voltage $V_{gs(m)}$ following some electrical stressing sequence. Note that $V_{gs(m)}$ was determined as the value of V_{gs} at which for the unstressed, $I_{ds}(V_{gs})$ curve dI_{ds}/dV_{gs} was maximized. Typically $V_{gs(m)} \sim$ - 0.52 V whilst $V_{th}^{o} \sim$ - 0.40 V.

An alternative approach was also used to determine $\Delta V_{th(IS)}$. In this case the measured $I_{ds}(V_{gs})$ curve was plotted and then replotted as a function of V_{gs}' where this value was simply displaced by an amount ΔV_{th} with respect to the initial V_{gs} value. Repeating this process one generates a family of $I_{ds}(V_{gd})$ curves each separated by an amount ΔV_{th}.

Working at a unique value of $V_{gs(m)}$ one can then generate a polynomial relating ΔV_{th} and ΔI_{ds} so that measured ΔI_{ds} (t_{stress}) values can be used directly to ascertain ΔV_{th} (t_{stress}). We verified that we obtained the same experimental ΔV_{th} (t_{stress}) values as those using Eq. 2.

Pseudo-DC measurements

The first methodology used here to examine the IS components is a single long term stressing at chosen gate bias followed by a long period of recovery similar to the pseudo-DC stressing discussed previously[1]. These measurements were made following the voltage application schematic shown in Figure 1a. At time t = 0, the device gate voltage was ramped to the measurement $V_{gs(m)}$ in approximately 20 ns and I_{ds}^{o} was measured, V_{ds} was -50 mV. The measurement time, t_m, for I_{ds} was ~ 4 μs. Subsequently, V_{gs} was ramped back to the stressing voltage $V_{gs(stress)}$ and V_{ds} set to 0 V – this stressing condition was maintained for a chosen time, t_{stress}. Following the stress sequence, V_{gs} was reduced to $V_{gs(m)}$ and $I_{ds}(t_{stress})$ determined after which V_{gs} was returned to $V_{gs(stress)}$, again in ~ 20 ns. The measurement process was periodically repeated, typically starting at stress times ~ 10 μs out to 1500 s with the time step increasing logarithmically to allow coverage of both very short and long total stress times. Using Equation 2, a stress curve $\Delta V_{th}(t_{stress})$ was calculated. Following a sequence of stress/measurements the devices were allowed to relax with V_{gs} = +1.5 V. Thus allowing nearly complete relaxation of both RC and FRC components of NBTI, leaving only the IS component.

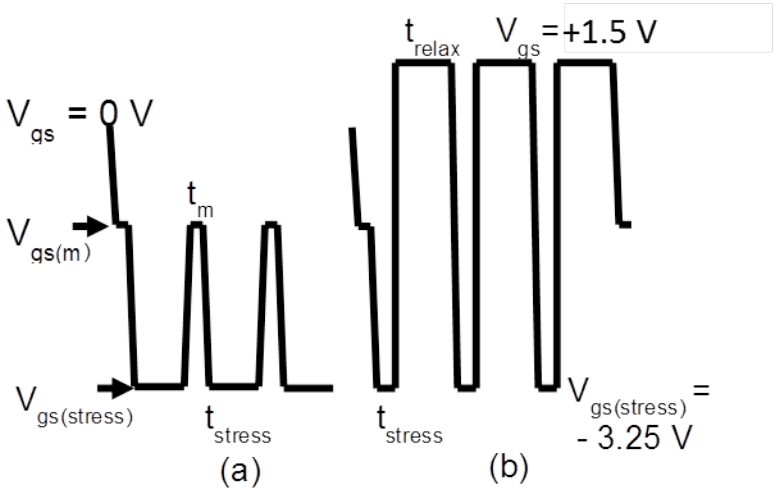

Figure 1. a) the pseudo-DC measurement sequence for V_{gs}. b) the pulsed stress/measurement sequence for V_{gs}.

Pulsed measurements

The pulsed measurement was made similar to that reported earlier [1]. The pulsed stress/measurement methodology enables determination of permanent component of NBTI. The sequence is shown in Figure 1b wherein a pulse is employed with a duty cycle, D ($0 \leq D \leq 1$). The pulse repetition frequency is ν. At time t = 0, it was analogous to the pseudo-DC case and we measured I_{ds}^{o}. The first stressing pulse was then applied with $V_{gs} = V_{gs(stress)}$ for a time $t_{stress} = D/\nu$, after which $V_{gs(stress)}$ changed to either 0 V or + 1.5 V for a time $t_{relax} = (1\text{-}D)/\nu$. This allowed complete relaxation of recoverable charge to occur during this time. Note that, $t_{stress} + t_{relax} = 1/\nu$ during each pulse. The total relaxation time per pulse sequence is $(1\text{-}D)/\nu$. Then for a total experimental time, t_{real},

(t_{real}) ν pulses are applied and, from Figure 1b, a measurement of $I_{ds}(t_{stress})$ is made such that $t_{stress} = D\ t_{real}$. The unique character of the measurement performed is that $I_{ds}(t_{stress})$ is determined immediately following the relaxation period at $V_{gs(stress)} = +1.5$ V of the final pulse i.e. after the (t_{real}ν)'th pulse. With this measurement sequence, during a pulse one measures only that charge which is permanent (non-spontaneously recoverable).

Results

Pseudo-DC measurements

Psuedo-DC measurements of NBTI have been made at various temperatures up to 220° C. Measurements of the source – drain current, I_{ds}, were made in the linear regime with a source-gate voltage, $V_{gs(m)}$, ~ -0.51 V. Bias stressing was carried out at a gate bias of -3.25 V with the source, drain and body contacts shorted at 0 V for a total stressing time of 1500 s. This was followed by 3000 s of recovery at a gate bias of $V_{gs}(rec) = +1.5$ V. The measured I_{ds} variations with stress were converted into threshold voltage shifts, ΔV_{th} using Eq. 2.

The experimental results of this set of measurements are shown in Fig.2. The charging data shows a similar shape and consistent growth for all temperatures below 180° C. However, the results for higher temperatures show the growth slowing and eventually turning around, so after 1500 s of stress there is less threshold voltage shift at 220° C than at 180° C. This can be explained by the presence of a recovery mechanism in the interface states which is only active at high temperatures.

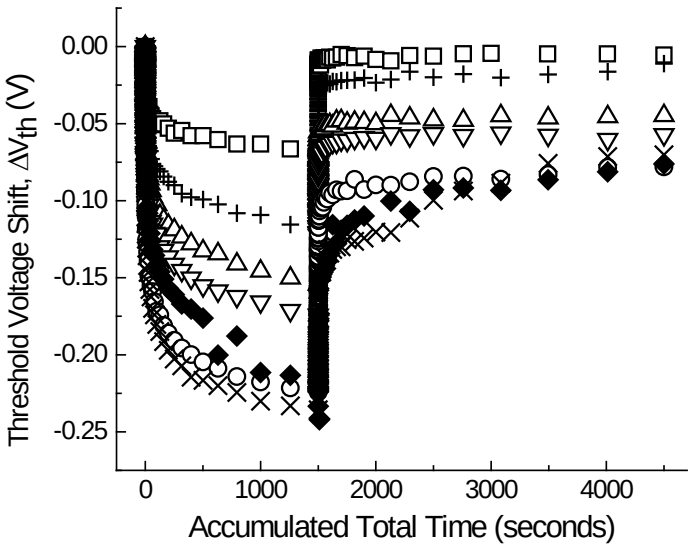

Figure 2: The measured threshold voltage shift due to NBTI with $V_{gs}(stress) = -3.25$ V for 1500s and $V_{gs}(rec) = +1.5$ V for multiple temperatures 50° C (□), 90° C (+), 120° C (△), 150° C (▽), 180° C (○), 200° C (×), and 220° C (◆).

There is a corresponding change in the shape of the recovery data. At temperatures below 150° C, there is a rapid recovery followed by a flat plateau representing the charged interface states because the device has discharged both the recoverable and field recoverable charge. The shape of the curve at higher temperatures shows a second recovery slope during what at lower temperatures was a plateau, showing recovery of

previously permanent charge and supporting the idea that the interface states are annealing at high temperatures.

Pulsed measurements

In order to measure the dynamic growth of the IS charge, while allowing the RC and FRC to fully recover, pulsed measurements with a frequency of 10 kHz over a range of temperatures from 50° to 220° C were made for a duty cycle of 0.1. Measurements of the source – drain current, I_{ds}, were again made in the linear regime with a source-gate voltage, V_{gs}, ~ -0.51 V. Bias stressing was carried out at a gate bias of -3.25 V with the source, drain and body contacts shorted at 0 V for a total stressing time of 200 s. This corresponds to a total time of 2000 s with a duty cycle of 0.1. The resulting growth curve is shown in Fig. 3.

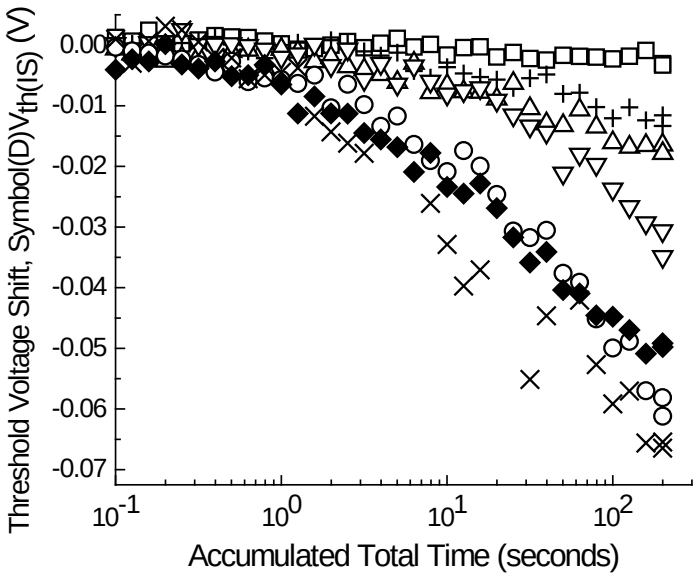

Figure 3: The measured growth of the interface component of the threshold voltage shift ΔV_{th}(IS) as a function of total stress time of 200 seconds is shown for multiple temperatures: 50° C (□), 90° C (+), 120° C (△), 150° C (▽), 180° C (○), 200° C (×), and 220° C (◆) .

The most striking result here is the steady increase in magnitude of the IS charge with increasing temperature for temperatures below 150° C, while above 180° C the increase with temperature begins to slow and by 220° C the IS charge is actually less than the result at 200° C.

Discussion

The IS trapped charge generated by irradiation is tentatively seen [6] as being produced through a complex of sequence of defect reactions. In the first case electron-hole pairs are generated in a radiation cascade in the gate dielectric "bulk". Trapping of the slow moving holes at neutral oxygen vacancies leads to positively charged defects. These defects are assumed to interact with always present, diffusing H_2 molecules thereby

generating protons which can diffuse to the Si/SiO_2 interface. At this interface the proton can interact with a passivated Si bond, Si-H to form a charged interface trap:

$$H^+ + Si_3{\equiv}Si\text{-}H \rightarrow H_2 + Si_3{\equiv}Si^+ \qquad [3]$$

We underline that this may be an extremely simplified version of reality. However, we start by assuming a similar mechanism is active for the creation of IS by NBTI. In this case the source of initial holes which trap at neutral oxygen vacancies is the p-channel inversion layer. Subsequently one can invoke a mechanism identical to that for radiation induced IS state buildup: diffusing H_2 cracks at positively trapped charged defects $\rightarrow$ generation of protons $\rightarrow$ diffusion to interface $\rightarrow$ release of passivating H leaving charged IS.

The presence of annealing of **radiation** induced IS at high temperatures through a recovery mechanism which is not simply the reverse of the generation mechanism has also been seen [7]. We can now safely say that such a mechanism also appears to be active for NBTI induced IS states.

It is possible, by choosing a total NBTI stress time, to plot the magnitude of the IS charging as a function of temperature as shown in Fig. 4. in which it is clear that there is at high temperatures a recovery term which becomes active and slows the increase in the growth of the curve. It is important to note that we observed considerable difficulty in

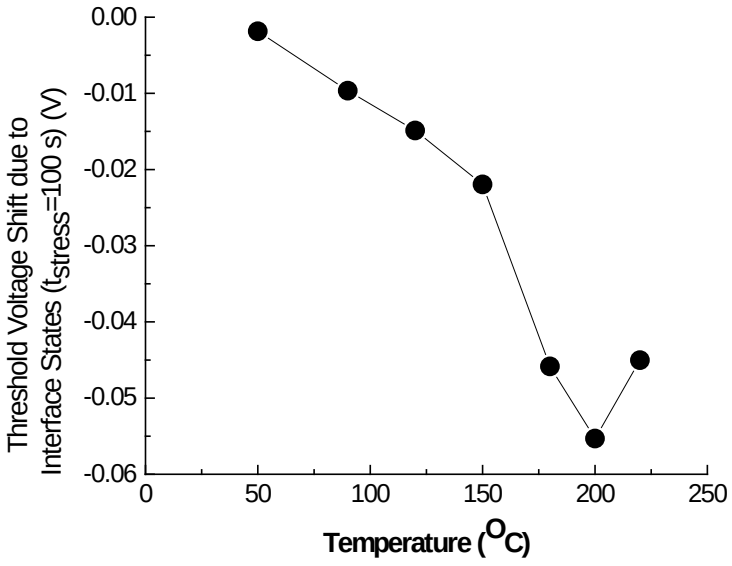

Figure 4: The interface state component of the threshold voltage shift after 100 s of stress as a function of device temperature.

making the measurements at high temperatures because of frequent device breakdown, therefore greater error in the higher temperature data is expected. Even so, it is clear that the measured value at 220° C is much lower than would have been expected based on simple extrapolation of the trend at lower temperatures. Clearly, some form of annealing mechanism is active competing with the basic generation term.

It is of interest to confront the results for the IS produced by radiation with those created by NBTI. This is done in Fig. 5 using the radiation results reported in [6,7]. We observe that the interface state density generated by a constant total dose of radiation

increases monotonically with increasing temperature for T < 150° C . Beyond this limit, the radiation induced interface state density is seen to decrease. This behavior is also observed for the NBTI induced interface state density except for the fact that the curve appears sharper and shifted by ~ 75° C to higher temperatures. This would be consistent with an activation energy of the generation process is different. Note, however, that there is evidence in the radiation induced interface state curve that it too varies with T and may also depend upon the concentration of H_2 in the gate dielectric [6,7].

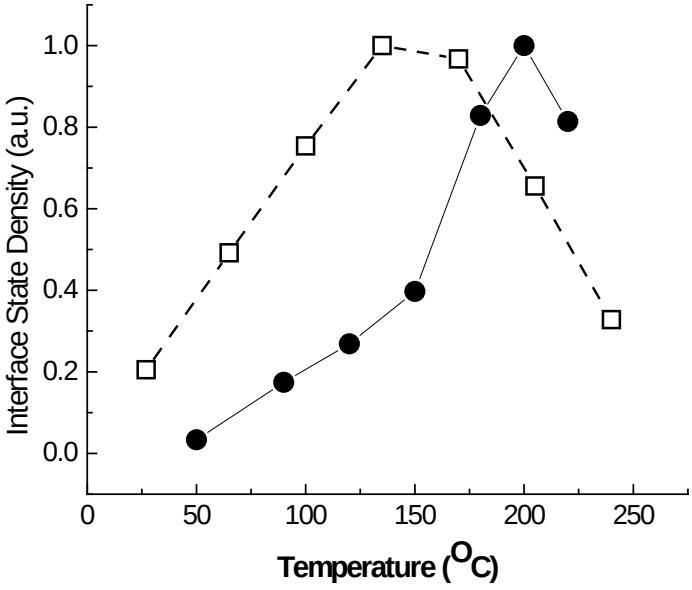

Figure 5: The interface state density in arbitrary units as a function of temperature showing apparent recovery for interface states created by NBTI (●) and irradiation (□) [6,7].

We have demonstrated that NBTI induced interface states in p-channel MOSFETscan be generated in increasingly large numbers by performing the NBTI at temperatures ≤ 200° C. Beyond this temperature the generation efficiency decreases suggesting the presence of an annealing mechanism. This conclusion has been reached using two different experimental measurement methodologies. These observations are consistent with those found in experiments where ionizing radiation is the primary source of creation of interface states, The mechanisms responsible for interface state generation both in NBTI and radiation experiments remain unclear and open for further studies. One can use these conclusions to reflect upon the question of synergy in the event that interface state generation was occurring during simultaneous irradiation exposure and NBTI. It would appear to us that both mechanisms of charged interface state trap generation are competing for the same neutral interface state precursors, and that therefore the interface state generation rate may be enhanced but in an additive fashion.

Acknowledgements

D. D. Nguyen and C. Kouhestani are with COSMIAC Kirtland, AFB, New Mexico USA 87117. This material is based on research sponsored by Air Force Research Laboratory (AFRL) under agreement number FA9453-08-2-0259. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

K. Kambour was supported by the US Air Force under a contract sponsored, monitored, and managed by: United States Air Force Air Force Materiel Command, Air Force Research Laboratory, RVSE, Kirtland AFB, NM 87117-5776.

Sandia National Labs is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

The work performed by R. A. B. D. was supported under a contract sponsored by the Air Force Research Laboratory, Space Vehicles Directorate, Kirtland AFB, NM 87117-5776.

References

[1] C. Mayberry, D. D. Nguyen, C. Kouhestani, K. E. Kambour, H. P. Hjalmarson and R. A. B. Devine, ECS Trans., **50**, 223-232 (2012).
[2] see, for example A. Islam, Ph. D. thesis, Purdue University 2010
[3] M. A. Alam, H. Kufluoglu, D. Varghese and S. Mahapatra, Microelectron. Reliability, 47, 853 (2007)
[4] T. Grasser, Microelectron. Reliability, 52, 39 (2012)
[5] S. M. Sze, The physics of semiconductor devices, (Wiley, N.Y. 1981) Chapt. 8
[6] D.R. Hughart, R. D. Schrimpf, D. M. Fleetwood, N. L. Rowsey, M. E. Law, B. R. Tuttle and S. T. Pantelides IEEE Trans. Nucl. Sci., **59**,3087-3092 (2012).
[7] S. C. Witzak, R. D. Schrimpf, K. F. Galloway, D. M. Fleetwood, R. L. Pease, J. M. Puhl, D. M. Schmidt, W. E. Combs and J. S. Suehle, IEEE Trans. Nucl. Sci. **43**, 3151-3160 (1996)