Tunneling discharge of positive trapped charge in p-channel field effect transistors

K. E. Kambour[1a], C. Kouhestani[2], D. Nguyen[2], N. Rosen[2] and R. A. B. Devine[2]

[1] SAIC, Kirtland AFB, New Mexico 87117, USA

[2] AFRL/RVSE, 3550 Aberdeen Ave., Kirtland AFB, New Mexico 87117, USA

[a] e-mail - Kenneth.E.Kambour@SAIC.com

ABSTRACT

Negative bias temperature instability has been studied in 130 nm and 90 nm channel length, SiO_2 gate insulator field effect transistors at room temperature. Pulsed voltage stressing and subsequent recovery using times starting in the tens of microsecond regime were employed together with a single point data acquisition time $\sim$ 4 μs. Threshold voltage shifts were characteristic of oxide charge trapping as opposed to interface state generation. Recovery of the threshold voltage was modeled assuming quantum mechanical tunneling of trapped charges from the dielectric to the Si substrate.

Prediction of the extended reliability lifetime of high performance circuits for space applications is a topic of increasing importance. Specifically, the question is how to determine a reliability lifetime ≥ 10 years based upon accelerated testing for a period of time on the scale of hours. For this to be feasible, very accurate measurement and modeling of inherent reliability loss mechanisms must be performed including such factors as duty cycle[1]. Inherent to this is the ability to precisely measure device characteristics, for example metal-oxide-semiconductor field effect transistor (MOSFET) threshold voltages[2], as rapidly as possible. The present work reports the results of a demonstrative study of one such reliability mechanism, negative bias temperature instability[3] (NBTI) in commercially produced p-channel MOSFETs (PMOS).

Devices with130 nm and 90 nm nominal channel lengths and nitrided SiO_2 gate insulators were constructed using a commercially available technology[4].   The gate dielectric thicknesses were 3.2 nm and 2.2 nm respectively and the channel widths were 5 μm. Electrical measurements were carried out using a Keithley 4200 SCS system with rapid data acquisition enabling a single source-drain current (I_{ds}) data point to be acquired in ≤ 1 μs or a 10 point source-drain current versus gate-source voltage, $I_{ds}(V_{gs})$, sweep acquired in ~ 10 μs. To optimize the signal to noise ratio data points were usually acquired in 4 μs. Devices were subjected to constant voltage stressing on the gate with source, drain and body contacts held at zero volts – this was periodically interrupted and the single data point acquired ($I_{ds}(t_{stress})$). For the 130 nm devices we stressed at $V_{gs} = -3.3$ V and measured $I_{ds}(V_{gs} = -0.6V, V_{ds} = -0.05V)$ whilst for the 90 nm devices we stressed at $V_{gs} = -3.0$ V and measured $I_{ds}(V_{gs} = -0.75V, V_{ds} = -0.05V)$, the use of a small source-drain voltage, (V_{ds}) ensured the device operated in the linear regime[2]. Stressing followed by relaxation with all contacts shorted, was carried out for different maximum initial accumulated stress time ranges starting at 100 μs and extending to 50 s.   Identical

measurement times were used for the stressing and relaxation parts of the cycle. The threshold voltage shift, ΔV_{th}, was determined from the single point I_{ds} measurement using the approximation:

$$\Delta V_{th} = (V_{gs} - V_{th}^0)(1 - I_{ds}/I_{ds}^0) \qquad (1)$$

where V_{th}^0 is the device threshold voltage before stressing and I_{ds}^0 is the initial current value. For the relaxation case Ids^0 is taken as the last I_{ds} value of the stressing sequence. The use of this approximate form in the case of small $1 - I_{ds}/I_{ds}^0$ has been justified previously[5]. Variation of the threshold voltage shift, ΔV_{th}, as a function of accumulated stress time in 90 nm devices at room temperature is shown in Fig. 1.

In Fig. 2 we show the variation of $\Delta V_{th}(t)$ in the relaxation phase where the devices were initially stressed for various total times (e.g. 100 µs - filled square symbols). The start of relaxation data point, not shown due to the logarithmic scale, corresponds to the final stress value (e.g. $\Delta V_{th}(t_{stress})$ = 100µs) – see Fig. 1. More detailed relaxation data for one set of devices initially stressed for 50 s is shown in Fig. 3. The "classical" vision of NBTI associates[4] the trapped and relaxing charge with the generation and relaxation of charged interface states – the charging law has been expressed as:

$$\Delta V_{th} = A(T,E) \, t^\beta \qquad (2)$$

where $A(T,E)$ is a temperature and oxide electric field during stressing dependent term and β is a constant whose value 0.16 or 0.25 dependent upon the mechanisms involved in interface state generation[6]. It is straightforward using the data shown in Fig. 1 to demonstrate that Eq. 2 cannot explain the time dependence observed in our experiments. Furthermore, published data on thin

SiO$_2$ gate oxide PMOS devices[4] which has been argued to result from interface state generation, when extrapolated to our operational conditions of T and E suggests that for times ~ 50 seconds ΔV_{th} would be typically a few mV. This value is substantially smaller than the values shown in Fig. 1. Therefore we conclude that in our case we are primarily dealing with charges trapped in the "bulk" of the gate dielectric but presumably close to the dielectric/semiconductor interface.

We assume that the observed relaxation under zero voltage bias on the gate, source, drain and body contacts proceeds via quantum mechanical tunneling. A similar approach has been adopted[6] for the relaxation of electrons trapped in SiO$_2$ following injection. The standard expression for ΔV_{th} is:

$$\Delta V_{th}(t) = \frac{1}{Ct_{ox}} \int_0^{t_{ox}} (t_{ox} - x)\rho(x,t)\,dx \tag{3}$$

in terms of a charge distribution $\rho(x,t)$ inside the gate oxide dielectric of thickness t_{ox}. x is measured from the semiconductor/dielectric interface and C is the gate capacitance. To treat the case of trapped charge loss via tunneling from the dielectric into the Si substrate Eq. 3 may be modified to include a term allowing for the spatially dependent tunneling effect. The final form is shown in Eq. 4 where τ_0 is a tunneling time constant and α is related to the natural logarithm of the probability of the charge tunneling through the dielectric. N_0 is the density of trapped charge initially present.

$$\Delta V_{th}(t) = \frac{1}{Ct_{ox}} \int_0^{t_{ox}} (t_{ox} - x)N_0 e^{\left(\frac{-t}{\tau_0 e^{\alpha x}}\right)}\,dx. \tag{4}$$

We will assume that only some oxide thickness δ close to the semiconductor/dielectric interface with uniform volume density of charge N_0 is actually relevant. Under these approximations we can express $\Delta V_{th}(t)$ in terms of the initial value $\Delta V_{th}(0)$:

$$\frac{\Delta V_{th}(t)}{\Delta V_{th}(0)} = \left[\frac{1}{t_{ox}\delta - \dfrac{\delta^2}{2}} \right] \int_{o}^{\delta} (t_{ox} - x) e^{\left(\dfrac{-t}{\tau_0 e^{\alpha x}} \right)} dx \tag{5}$$

Integrating Eq. 5 numerically we have fitted to the data shown in Fig.2 for the 90 nm devices and for 130 nm devices (data not shown). A typical fit for a 90 nm device initially stressed for 50 seconds at $V_{gs} = -3.0$ V then relaxed is shown in Fig. 3.

The fitting procedure involves three unknowns, α, δ and τ_0 with the product $\alpha\delta$ turning out to be the most relevant term. In Fig. 4 we plot the fit values of $\alpha\delta$ versus total stressing time obtained from the 90 nm data shown in Fig. 2 and similar data for the 130 nm devices alluded to earlier. To demonstrate the sensitivity of the fitting procedure to τ_0 we show fit data for $\alpha\delta$ for 130 nm devices using two different values: $\tau_0 = 1 \times 10^{-6}$s and 1×10^{-7} s. The fit is clearly not very sensitive to the choice of τ_0. Though we cannot uniquely determine α and δ, we do observe that α should not be a variable of the experiment so variations in the $\alpha\delta$ product shown in Fig. 4 must result from changes in δ. This enables us to conclude that the primary result of extended electrical stressing is to drive trapped charge further into the dielectric layer and not to simply increase the charge density in a thin defect filled layer at the semiconductor/dielectric interface. A very crude estimate of α can be made assuming $\delta \leq t_{ox}$, if $t_{ox} \sim 3.2$ nm and $\alpha\delta$ is 24 then $\alpha \geq 7.5 \times 10^9$ m^{-1}.

Room temperature NBTI stressing in SiO_2 gate PMOS devices shows evidence for positive charge trapping at timescales ~ tens of μs. Relaxation via tunneling is demonstrated to explain the experimental data.

REFERENCES

1. W. Wang, S. Yang, S. Bhardwaj, S. Vrudhula, F. Liu, Y. Cao, IEEE Trans. on VLSI Systems **18**, 173 (2010)
2. S. M. Sze, **Physics of semiconductor devices** (Wiley, N. Y. 1981) Chapt. 8
3. M.A. Alam, H. Kufluoglu, D. Varghese and S. Mahapatra, Microelect. Reliability **47** 853 (2007)
4. See for example: http://www.mosis.com/ibm/
5. J.K. Mee, R.A.B. Devine, L. Trombetta, R.J. Kaplar and P. Gouker, Microelect. Reliability. **51**, 1113 (2011)
6. K. Yamabe and Y. Miura, J. Appl. Phys. **51**, 6258 (1980)

Figure Captions

Figure 1: The threshold voltage shift during the stress phase as a function of total stressing time.

Figure 2: The threshold voltage shift during recovery as a function of the time since the stress was discontinued.

Figure 3: Fitted recovery curve for the case of a 50 second stress duration then relaxation. For simplicity the threshold voltage shift is normalized to unity for the final stress value ($\Delta V_{th}(t_{stress} = 50 \text{ s}) = 1$). The fit parameters were $\tau_0 = 1 \times 10^{-6}$ s and $\alpha\delta = 24$.

Figure 4: The product $\alpha\delta$ as a function of initial stress time deduced from fits to the data for 90 nm devices such as in Fig. 2($\square$) and 130 nm devices (not included in the text). Fits to the 130 nm data were performed for ($\bigcirc$) $\tau_0 = 1 \times 10^{-6}$ s and for ($\triangle$) $\tau_0 = 1 \times 10^{-7}$ s to test the sensitivity to this parameter.

α*δ (y-axis)

Total Recovery Time (seconds) (x-axis)

□ 90 nm Devices
○ 130 nm Devices ($\tau_p = 1*10^{-8}$ s)
△ 130 nm Devices ($\tau_p = 1*10^{-7}$ s)