

Simulating Negative Bias Temperature Instability of p-MOSFETS

Introduction

The degradation of MOSFET devices having relatively thin oxide layers is generally accepted as being mainly associated with the depassivation of silicon dangling bonds at the Si/SiO_2 interface. These dangling bonds are initially passivated during the fabrication process by heating in hydrogen or, more rarely, a deuterium environment. The interface trap density is typically reduced by two orders of magnitude by this passivation process, to around $10^{10}cm^{-2}$ or even less[1].

The mechanisms for depassivation under device operating conditions have been the subject of much investigation. It is found that devices passivated with deuterium show much improved degradation properties [2] (and references therein), suggesting that the Si-D bond is much stronger than the Si-H bond. As pointed out by van de Walle and Tuttle [3], the static electronic bonding is the same for Si-H bonds and Si-D bonds, and therefore the apparent difference in bond strengths must arise from the different dynamic nature of the bonds caused by the larger mass of the deuterium nucleus.

It has been suggested that multiple excitation of transverse vibrational states by inelastic carrier scattering can liberate the hydrogen or deuterium atom from the bond. In the case of deuterium the energy quanta of the excitation are believed to couple strongly to a bulk phonon mode and thus the bending mode excitations have a short relaxation time. In the case of hydrogen, however, the vibrational modes are only weakly coupled to phonon states and a vibrational ladder is climbed until the bond dissociates at a median energy of around 1.5 eV above the ground state [4]. Consequently, charge carriers with relatively low energy may contribute to the depassivation process, and carriers with sufficient energy to be injected into the gate oxide are not primarily responsible for this generation of interface traps [2]. Another suggested reaction for the depassivation of the Si-H bond is that the bond first captures an inversion layer hole which reduces its binding energy, thereby increasing the rate of thermal dissociation [5]. This process is expected to be the dominant one in p-MOSFETS biased into channel inversion.

Silvaco ATLAS is capable of modelling the depassivation process when the reaction driving agent is any combination of hot channel current, tunnelling current, oxide field or inversion layer hole capture. The mode of stressing determines which is the most significant mechanism. A critical part of the simulation is the modelling of transport of the hydrogen liberated as part of the depassiva-

tion process. The model assumes that atomic, neutral hydrogen is created by the depassivation, and that the trap becomes charged to $\pm Q$ *Coulombs* almost instantaneously. The atomic hydrogen can diffuse and also dimerise to H_2 molecules, and the H_2 molecules also diffuse and can be converted back to atomic hydrogen.

The interface charge density N_{it} as a function of time is found by integrating the following equation forward in time

$$\frac{dN_{it}}{dt} = K_f(RD.SIHTOT - N_{it}) - RD.KR0 H N_{it} \quad \text{EQ 1}$$

where K_f is the forward reaction rate, RD.SIHTOT is the total available dangling bond density, RD.KR0 is a repassivation rate and H is the local density of atomic hydrogen. The rate K_f can include contributions to bond-breaking from direct field effects, channel hot carriers and Fowler-Nordheim tunnelling. Because the binding energies of the Si-H bonds are expected to form a continuous distribution, the forward reaction rate is obtained as an integral over energy

$$K_f = \int_0^{+\infty} K_f(E)G(E)dE \quad \text{EQ 2}$$

where

$$K_f(E) = RD.KF0 \exp\left(-\frac{(E+F_{\perp}RD.AESLOPE)}{KT}\right) \times (1.0 + RD.INVHCOEF P)$$

and

$$G(E) = \frac{1}{2RD.AEVAR(1 + \cosh(\frac{(E-RD.AE)}{RD.AEVAR}))} \quad \text{EQ 3}$$

where $G(E)$ is the derivative with respect to energy of the Fermi-Dirac distribution function [4]. The nominal (median) activation energy for the depassivation is given by the parameter RD.AE, where RD.AEVAR is the energy width parameter for the distribution of activation energy. In the expression for $K_f(E)$, F_{\perp} is the field in the oxide perpendicular to the interface and P is the hole density at the interface. The terms dependent on current have been omitted for clarity. We only consider thermal dissociation enhanced by hole capture in this article. The effect of perpendicular field is to lower the activation energy of the depassivation reaction. This is thought to be a linear dependence [4]. The full expression for K_f , along with all the default values of parameters are listed in the ATLAS

manual. In this article we analyse the Reaction-Diffusion equation, and present the results of a full numerical simulation of degradation of a p-MOSFET. The Appendix gives suggestions for ensuring that the numerical solution is of high integrity.

Analysis of Negative Bias Temperature Instability Degradation

This type of degradation occurs in p-MOSFETs and impacts on digital and analog CMOS circuits, and affects parameters such as threshold voltage, transconductance, saturation current etc. It is thought to be due to both interface charge generation, and the creation of defects in the SiO_2 itself. One experimentally observed feature of NBTI is partial degradation recovery after stressing has finished. The interface charge creation is thought to be at least partly reversible, while permanent NBTI degradation effects are thought to be associated with creation of trapping sites in the bulk oxide, possibly by structural relaxation [6]. For devices with thin oxides the interface states are dominant[7], with bulk states making a contribution with thicker gate oxide thicknesses. The reversibility of the degradation can be attributed to repassivation by any lingering hydrogen. Another possibility is that holes tunnel into states which extend into the bulk, and have a large range of characteristic re-emission rates. This can be modelled using the Heiman model available in SILVACO ATLAS[8].

The passivated dangling bonds at the interface can be weakened by the capture of a hole, which removes a bonding electron from the system. Thus the forward reaction rate will depend on inversion layer hole density. The hydrogen removed during the depassivation reaction can diffuse and also repassivate the dangling bonds. This repassivation process is described by the last term in equation (1). Equation (1) is known as the Reaction-Diffusion equation, and initially depassivation dominates until enough hydrogen has been liberated to result in a significant repassivation rate. Therefore the Reaction term controlled by K_f dominates at first and the diffusion term, proportional to $RD.KR0$, becomes important at a later time. The diffusion of hydrogen is numerically solved for throughout the device, with material dependent diffusion parameters. Some models restrict diffusion only to the insulator side of the interface [9], but diffusion can also occur on the Silicon side of the interface [3]. The transport of hydrogen in silicon is more complicated than the straightforward diffusion considered here. However, it is important to include the material dependence of the diffusion rate [10],[11], and the diffusion coefficients can be used as fitting parameters.

A simplified equation can be derived in order to analyze the time behavior of the interface charge density. Assuming that the interface trap density is initially zero, the total hydrogen liberated after a certain time is equal to the

total number of interface traps. In terms of densities, H can be related to N_{it} by the introduction of a characteristic length, the most obvious length to use being \sqrt{Dt} , where D is the diffusion coefficient and t the simulation time. Thus $H \approx N_{it}/\sqrt{Dt}$, and by omitting field and hole concentration dependence, equation (1) may be rewritten

$$\frac{dN_{it}}{dt} = K_f(RD.SIHTOT - N_{it}) - K_r \frac{N_{it}^2}{\sqrt{t}D_{eff}} \quad \text{EQ 4}$$

where D_{eff} is an effective diffusion coefficient. This is a non-linear first order differential equation and will have more than one solution in general. There is a solution which is asymptotic to $\frac{\sqrt{D_{eff}}}{2K_r} t^{-0.5}$ as $t \rightarrow 0+$. This solution is unbounded at $t = 0$, and is therefore unphysical. With the initial condition that $N_{it} = 0.0$ there is a solution which is asymptotic to $N_{it} = K_f RD.SIHTOT t$ for small t, and sublinear in time as the solution develops. At large time, then any bounded solution must tend asymptotically to $N_{it} = RD.SIHTOT$, in other words all the dangling bonds become depassivated. It is also possible to show that this solution is monotonically increasing in time. If the repassivation rate K_r is sufficiently large, then for part of the solution the reaction term and the diffusion term will both be much larger than the nett depassivation rate. If it is further assumed that $N_{it} \ll RD.SIHTOT$ it is deduced that

$$N_{it} \approx \sqrt{\frac{K_f RD.SIHTOT}{K_r}} (D_{eff} t)^{0.25} \quad \text{EQ 5}$$

as various authors have shown [5],[7]. This simple relationship is not valid at all timescales. A comparison between a numerical solution of equation (4) and approximation (5) is shown in Figure 1. The parameter values were $K_f = 100s^{-1}$, $K_r = 10^{-6}cm^3s^{-1}$, $D_{eff} = 10^{-4}cm^2s^{-1}$ and $RD.SIHTOT = 10^{11}cm^{-2}$. The approximate solution is of high quality until the depassivated bond density is a significant fraction of overall dangling bond density. The solution of the full numerical model will show more detail, but the analysis of the simplified case allows some conclusions to be drawn.

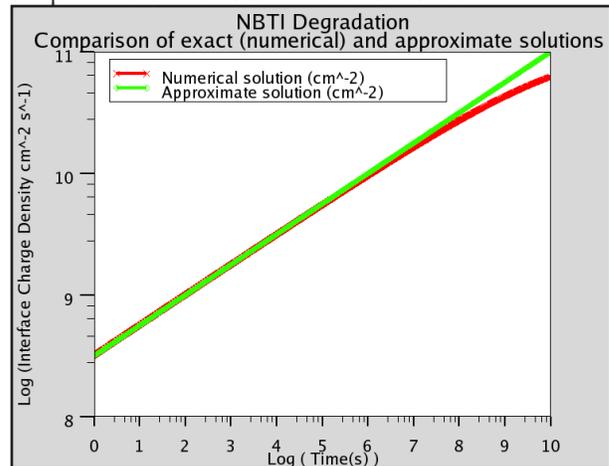


Figure 1. Approximate versus numerical solution of simplified Reaction-Diffusion equation.

When using the TR-BDF2 integration scheme to solve equation (4) it was found that fully implicit handling of the non-linear terms was required, otherwise non-physical solutions were obtained. In the full model this implies that the diffusion equations must be fully coupled to the depassivation equation, which is the default behaviour in ATLAS. In addition, solutions for carrier density and potential are also fully coupled into the equation system. One further conclusion drawn from the analysis arises from the fact that the nett depassivation rate is a difference of two larger numbers. Thus the solutions must be obtained to a sufficiently high relative precision in order to avoid numerical noise dominating the nett depassivation rate.

Application to a p-MOSFET with 9 nm Gate Oxide Thickness

The Silvaco ATHENA package was used to create a sub-micron device structure, with a minimum gate thickness of 9 nm. It is shown in Figure 2.

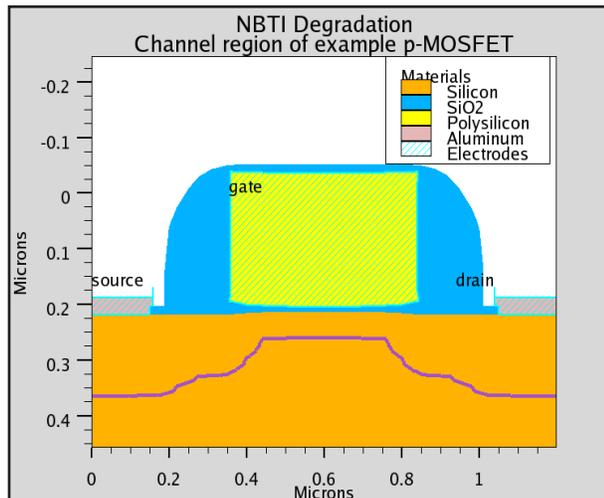


Figure 2. Device structure of active part of simulated p-MOSFET. The position of the p-n junction is also indicated.

A negative gate bias of -4.5 V was applied to the device, and then a transient simulation was performed, up to a simulated time of 1×10^6 seconds. The Reaction-Diffusion device degradation model was enabled by using the DEVDEG.RD flag on the MODELS statement. The values used for RD.AE and RD.AEVAR were 1.5 eV and 0.05 eV respectively. RD.KF0 was $1 \times 10^{-5} \text{ s}^{-1}$ and RD.AESLOPE was $-5.6 \times 10^{-8} \text{ cm}$. A total dangling bond density of $1 \times 10^{12} \text{ cm}^{-2}$ was assumed, initially passivated 100%. Results were obtained at simulation temperatures of 300 K and 400 K, with the ATLAS default set of models for MOSFET simulation. The value of RD.INVHCOEF was $1 \times 10^{-5} \text{ cm}^3$ and the repassivation coefficient RD.KR0 was $1 \times 10^{-8} \text{ cm}^3 \text{ s}^{-1}$. Only Atomic hydrogen was modelled by setting the dimerisation rate to zero.

The diffusion co-efficients D are calculated according to the Arrhenius formula $D = D_0 \exp(-\frac{E_A H}{KT})$. Consequently, the diffusion rate in all materials is higher at 400 K than at 300 K. The forward reaction rate is also larger at 400K than at 300K. Taken together this implies that nett depassivation will be higher at 400 K, resulting in faster degradation. In Figure 3a the time evolution of the total interface trapped charge is shown, in units of Coulombs per cm of device width. In Figure 3b the shift in threshold voltage as a function of time is shown. These Figures clearly demonstrate that the effect of the stressing is greater at the higher temperature. In Figures 4a and 4b the concentration of atomic hydrogen is shown after 1 s and after 1000 s for the simulation carried out at 300 K. The position of the p-n junction in the device is also shown. Hydrogen can only leave the device at its boundary, and so the hydrogen diffusion through the gate stack, electrodes and silicon is simulated.

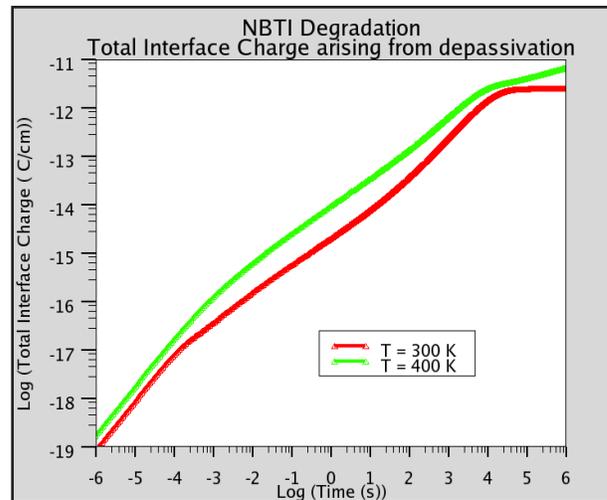


Figure 3a. Log-Log plot of the total positive interface charge created by the degradation model as a function of time and at two temperatures.

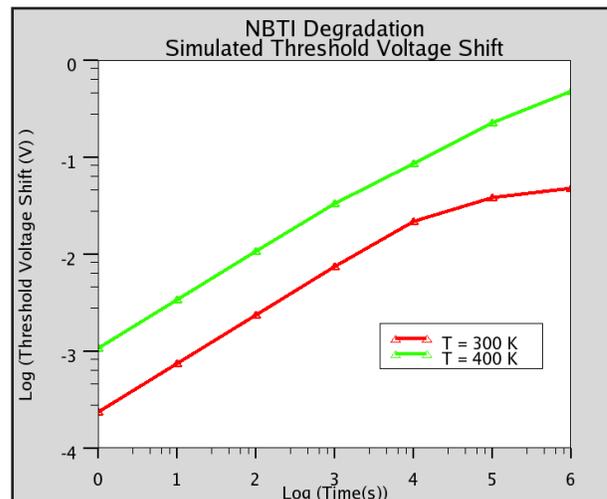


Figure 3b. Log-Log plot of the consequent threshold Voltage shifts as a function of time.

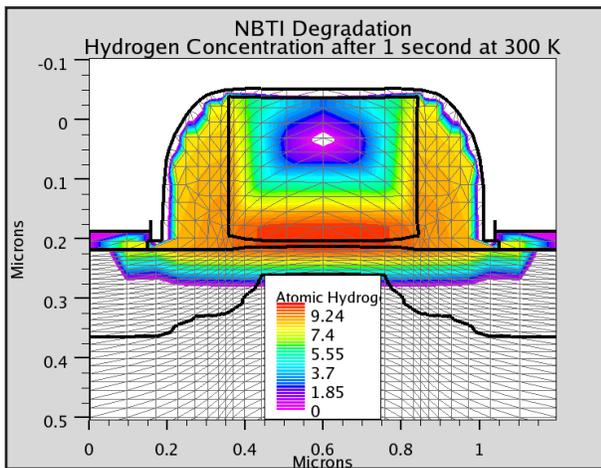


Figure 4a. Spatial hydrogen concentration at a simulation time of 1 s for the simulation at 300 K. Also shows simulation mesh.

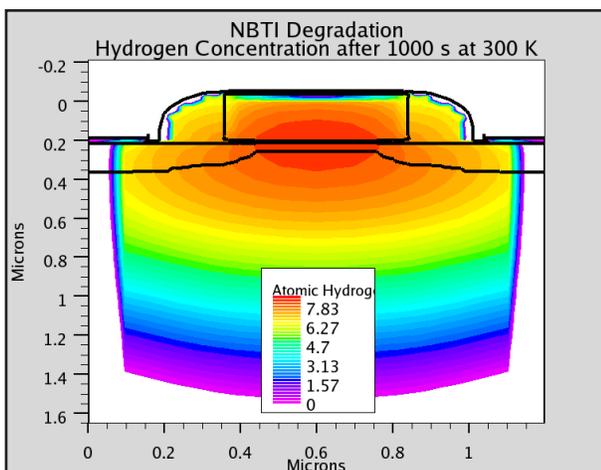


Figure 4b. Spatial hydrogen concentration at a simulation time of 1000 s for the simulation at 300 K.

Discussion

Experimental curves of V_t shift versus time of stressing at negative gate Voltage can show quite different behavior. For example in Figure 1. of Mahapatra et al [7] a power law relationship with a single exponent is evident for lower values of gate bias. At higher values of gate bias the curves switch to a power law with a larger exponent after a certain time, t_{break} . This is explained in terms of the creation of trapping states in the bulk of the oxide. The experimental results given in Figure 10 of Ielmini et al [6] show a curve which has a more logarithmic dependence on simulation time. These can be described by a hole tunnelling and trapping model, enhanced by thermal activation probabilities for the trapping and emission processes. In reality, both hole tunnelling and depassivation are probably present in both data sets. A critical parameter is the total number of available dangling bonds at the interface. When a significant fraction of them have been depassivated there is no longer a simple power law dependence. This could partly explain the results of Ielmini et al [6].

The curves in Figure 3b of this article are initially linear, but become sublinear. This is thought to be due to a combination of the increasing fraction of depassivated bonds, and details of the hydrogen diffusion process. For the result at 400 K, where hydrogen diffuses much more rapidly, the effect is much less pronounced. This can be clearly seen from the hydrogen concentration as a function of time at three positions. In Figure 5a is the hydrogen concentration at the center of the Si/SiO_2 interface. Except at the very smallest timescale, the simulation done at 300 K shows the higher hydrogen concentration, despite the higher depassivation rate for the simulation at 400 K. This feature is a consequence of the hydrogen diffusing away from the interface much more quickly at 400 K. This is clearer in Figures 5b and 5c, which are the hydrogen concentrations at a point in the Poly-Si contact and at a point deep in the silicon respectively. The diffusion front arrives earlier, and has a larger peak at 400 K. At longer timescales the concentration is larger for the 300 K simulation. This means that the repassivation rate is larger here, and this would explain the behaviour of the threshold voltage shift at longer times for the 300 K simulation. With an assumed activation energy of 1 eV for the diffusion of Hydrogen in Silicon, the diffusion co-efficient will be approximately 2×10^4 larger at 400 K than at 300 K.

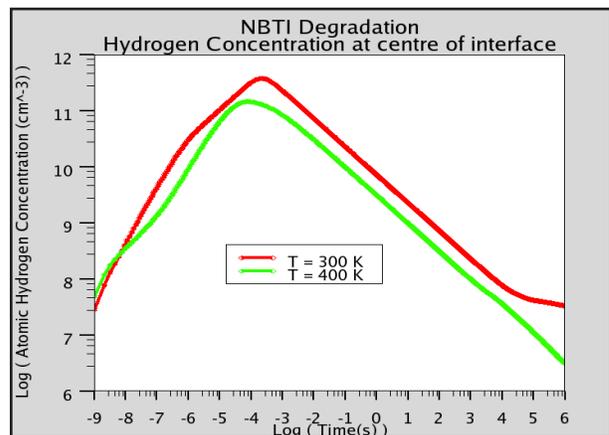


Figure 5a. Log-Log plot of the time evolution of the hydrogen concentration in the centre of the interface.

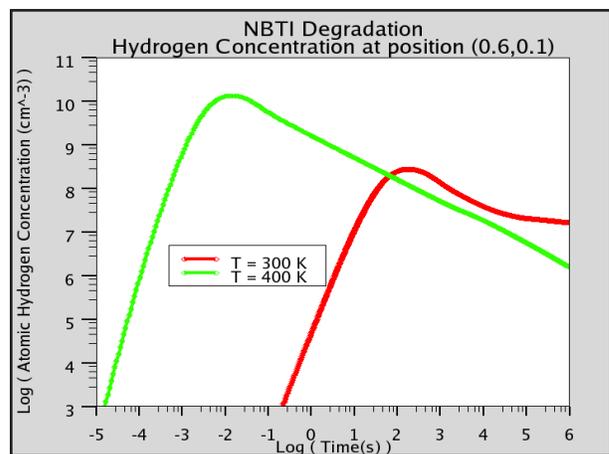


Figure 5b. Log-Log plot of the time evolution of the hydrogen concentration at a point in Poly-Si gate contact.

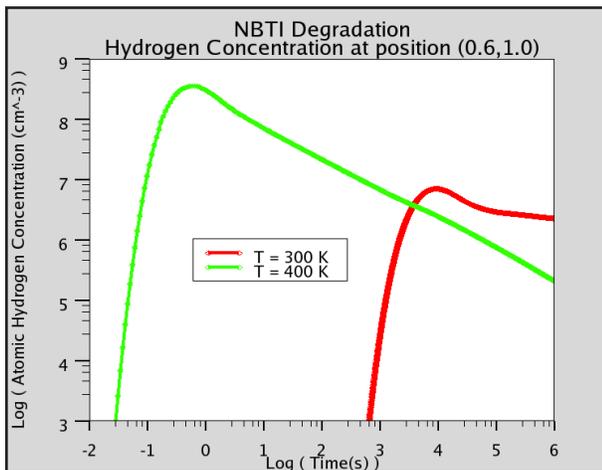


Figure 5c. Log-Log plot of the time evolution of the hydrogen concentration in the Silicon channel.

In summary, Silvaco ATLAS has both a sophisticated Reaction-Diffusion model and the HEIMAN model available for studying NBTI degradation. The HEIMAN model considers tunnelling into traps in the oxide, and has recently been enhanced to allow the effect of structural relaxation of the amorphous oxide, as suggested by Ielmini et al. The Reaction-Diffusion model can simulate many stressing modes, and also the diffusion and inter-conversion of H and H_2 throughout the device.

Appendix : Time Step Control

The default method of time-step control used with the TR-BDF2 time integration scheme can be prone to numerical noise when simulating MOSFET inversion layer channels. One alternative is to use a higher precision version of ATLAS. The alternative used here is the LT-E2STEP method. This uses a differencing algorithm applied to values of the solution variables from the previous two timesteps to estimate the third derivatives of the transient solution variables. These are then combined in the usual way to form an estimate of Local Truncation Error (LTE). The LTE is then used to obtain the next timestep which will conform to the specified LTE upper limit. To use this method you specify LTE2STEP on the METHOD statement, and also a value of TOL.TIME which is smaller than the default value of 5.0×10^{-3} . It is recommended to reduce the value of TSTEP.INCR from its default of 2 when using the DEVDEG.RD model. Because of the numerical difficulties with this equation, it is recommended to obtain solutions with a small enough timestep, and also with a small enough value of CX.TOL. Specifying XNORM ensures that equation convergence is only accepted when relative errors are small.

References

- [1] M. L. Reed and J. D. Plummer, "Chemistry of Si-SiO₂ interface trap annealing", *J. Appl. Phys.*, Vol. 63, No. 12, (June 1988), pp. 5776-5793.
- [2] Z. Chen, K. Hess, J. Lee, J. W. Lyding, E. Rosenbaum, I. Kizilyalli, S. Chetlur and R. Huang, 'On the mechanism for interface trap generation in MOS transistors due to Channel Hot carrier stressing', *IEEE Electron Device Letters*, Vol. 21, No. 1, (Jan 2000), pp. 24-26.
- [3] C. G. Van de Walle and B. R. Tuttle, 'Microscopic Theory of Hydrogen in Silicon Devices', *IEEE Trans. Electron Devices*, Vol. 47, No. 10, (Oct 2000), pp. 1779-1785.
- [4] C. Guerin, V. Huard and A. Bravaix, 'General framework about defect creation at the Si/SiO₂ interface', *J. Appl. Phys.*, Vol. 105, (2009), 114513.
- [5] H. Kuffuoglu and M. A. Alam, 'A generalized Reaction-Diffusion Model with explicit H-H₂ Dynamics for NBTI Degradation', *IEEE Trans. Electron Devices*, Vol. 54, No. 5, (2007), pp. 1101-1107.
- [6] D. Ielmini, M. Manigrasso, F. Gattel and M. G. Valentini, 'A new NBTI Model Based on Hole Trapping and Structural Relaxation in MOS Dielectrics', *IEEE Trans. on Electron Devices*, Vol. 56, No.9 (2009), pp. 1943-1952.
- [7] S. Mahapatra, P. Bharath Kumar and M. A. Alam, 'Investigation and Modeling of Interface and Bul Trap Generation During Negative Bias Temperature Instability of p-MOSFETS', *IEEE Transactions on Electron Devices*, Vol. 51, No. 9, (2004), pp. 1371-1379.
- [8] 'Simulating the Hysteresis effects of Si/SiO₂ traps', *Silvaco Simulation Standard*, Vol. 20, No. 2, April-May-June 2010, www.silvaco.com.
- [9] K. L. Brower, 'Dissociation kinetics of hydrogen-passivated (111) Si-SiO₂ interface defects', *Phys. Rev. B*, Vol. 42, (1990), pp. 3444-3453.
- [10] S.F. Wan, M. Hatta, N. Soin, J.F. Zhang, 'The effect of gate Oxide Thickness and Drain Bias on NBTI Degradation in 45 nm PMOS', *Proc. ICSE2010* (2010), pp. 210-213.
- [11] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, A. Varghese, R. B. Khamankar, S. Krishnan, 'Material Dependence of Hydrogen Diffusion: Implications for NBTI Degradation', *IEDM Technical Digest*, (5 December 2005), pp. 691-695.