Available online at www.sciencedirect.com
ScienceDirect
Microelectronics Reliability 47 (2007) 841–852
MICROELECTRONICS RELIABILITY
www.elsevier.com/locate/microrel

Negative bias temperature instability: What do we understand?

Dieter K. Schroder

Department of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, USA

Received 24 May 2006; received in revised form 31 August 2006
Available online 1 December 2006

Abstract

We present a brief overview of negative bias temperature instability (NBTI) commonly observed for in p-channel metal–oxide–semiconductor field-effect transistors (MOSFETs) when stressed with negative gate voltages at elevated temperatures and discuss the results of such stress on device and circuit performance and review interface traps and oxide charges, their origin, present understanding, and changes due to NBTI. Next we discuss some of the models that have been proposed for both NBTI degradation and recovery and p- versus n-MOSFETs. We also address the time and energy dependence effects of NBTI and crystal orientation. Finally we mention some aspect of circuit degradation. The general conclusion is that although we understand much about NBTI, several aspects are poorly understood. This may be due to a lack of a basic understanding or due to varying experimental data that are likely the result of sample preparation and measurement conditions.

1. Introduction

Negative bias temperature instability has been known since 1966 [1]. It is only during the last few years, however, that it has become a reliability issue in silicon integrated circuits, because the gate electric fields have increased as a result of scaling, increased chip operating temperature, surface p-channel MOSFETs have replaced buried channel devices, and nitrogen is routinely added to thermally grown SiO_2. In 2003, we wrote a review paper on NBTI and tried to include most of the published literature up to that time [2]. Although many papers had been published at that time, there was much discrepancy between the various models and the experimental data. For example, it was poorly understood that the time between NBTI stress and measuring the effect after terminating the stress was important, because the NBTI recovery was just beginning to be understood. Now it is understood that the sooner a degraded device is measured after stress, i.e., within ms or sooner, the more relevant are the data. Since the publication of our paper, many more papers have been published and NBTI is beginning to be understood better. Nevertheless, there are still many gaps and experimental data sometimes contradict each other. Several excellent review papers have since been published. In this paper, I will draw upon the excellent recent reviews by Stathis and Zafar [3], Huard et al. [4], and Alam and Mahapatra [5] as well as many other papers published in the last few years. In particular, the careful, detailed measurements of Huard et al. have confirmed or debunked some earlier assumptions and I draw heavily on their results. I have attempted to highlight those NBTI areas where there is considerable uncertainty and where experimental results contradict each other. To highlight our understanding or misunderstanding I have ended most sections with questions.

2. What is NBTI?

NBTI is an increase in the absolute threshold voltage, a degradation of the mobility, drain current, and transconductance of p-channel MOSFETs. It is almost universally attributed to the creation of interface traps and oxide charge by a negative gate bias at elevated temperature. The oxide electric field is usually, but not always, lower than that leading to hot carrier degradation. The oxide electric field and temperature are similar to those typically

E-mail address: schroder@asu.edu

0026-2714/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.microrel.2006.10.006

encountered during *burn-in* and sometimes encountered during high-performance chip operation. The details of how NBTI occurs are not entirely clear, however. Since much has been written about the origin of NBTI, we will touch on it only briefly. The mechanism is ascribed to breaking of SiH bonds at the SiO_2/Si substrate interface by a combination of electric field, temperature, and holes, resulting in dangling bonds or interface traps at that interface, designated as D_{it} and N_{it}, and positive oxide charge, N_{ot}, that may be due to H^+ or trapped holes. Detailed measurements by various researchers have shown the activation energy and time dependence to differ for D_{it} and N_{ot} generation.

3. NBTI models

Various NBTI models have been proposed, of which the *Reaction–Diffusion* (R–D) model is the most prevalent [5–7]. In this model, interface traps are generated at the SiO_2/Si interface (reaction) with a linear dependence on stress time. Hydrogen is released during this reaction phase. In the subsequent diffusion phase, the hydrogen diffuses from the interface into the oxide (diffusion) with the time dependence t^n, where n for neutral hydrogen species is frequently given as 0.25. It may also be possible for hydrogen to diffuse into the substrate. While the diffusion-limited regime has been verified many times, the reaction-limiting regime has not as frequently been observed, because it occurs during a very short time. Yang et al. extended the stress measurement times from 10^{-3} to 10^2 s and show the $n = 1$ and $n = 0.25$ dependence, with the breakpoint between the two regimes occurring at $t \approx 0.02$–0.03 s, shown in Fig. 1 [8]. A caution about the data in Ref. [8]. The measurements are affected by the stress-to-measure delay and the $n = 1$ value is likely to be lower if this delay is considered. By measuring NBTI degradation as a function of nitrogen density in nitrided oxides, they conclude that D_{it} generation is enhanced by nitrogen and it is mainly due to enhanced hydrogen diffusion in the nitrided oxide. This enhanced diffusion aids D_{it} generation.

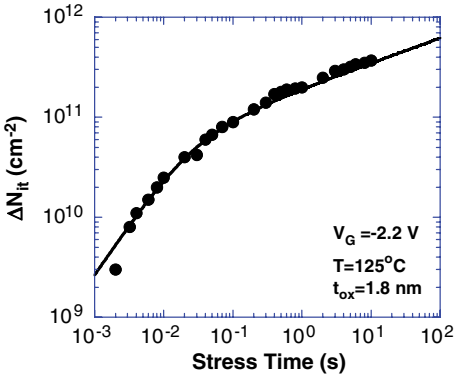

Fig. 1. Measured and simulated interface trap density change ΔN_{it} versus stress time for 15% nitrogen concentration. In Eq. (1): $A = 2.67 \times 10^{12}$ cm^{-2}/s, $B = 6.95 \times 10^{-11}$ cm^2/s$^{1/2}$, $D_H = 8.28 \times 10^{-16}$ cm^2/s, $W/L =$ 20 μm/0.15 μm. After Yang et al. [8].

The generation of interface traps is given by [8]

$$\Delta N_{it}(t) = \frac{2At}{1+\sqrt{1+4ABt^{3/2}}} = \frac{2k_F N_0 t}{1+\sqrt{1+4k_F N_0 k_R t^{3/2}/(0.5\sqrt{D_H})}} \quad (1)$$

where k_F is the forward reaction rate, k_R the reverse reaction rate, N_0 the initial defect density, and D_H the hydrogen diffusion coefficient. According to Eq. (1), for small t during the *reaction* phase of NBTI degradation

$$\Delta N_{it}(t) \approx At = k_F N_0 t \quad (2a)$$

and during the *diffusion* phase at later times

$$\Delta N_{it}(t) \approx \sqrt{\frac{A}{B}}t^{1/4} = \sqrt{\frac{k_F N_0}{2k_R}}(D_H t)^{1/4} \quad (2b)$$

Eq. (1) is plotted in Fig. 1 together with experimental data.

There is little controversy about the applicability of the R–D model to NBTI. What is still debated is the time exponent n, which is predicted to be 0.25 above, but frequently observed to be lower. The true value is important to be able to predict NBTI lifetimes. The generation of interface traps is universally accepted, but, the nature of the positive oxide charge generated during NBTI stress is not clear yet. It has been attributed to hole traps and also to trapped hydrogen.

Question: Is the R&D model the best model? Does reaction dominate at short times and diffusion at longer times? What is the true n value?

4. Interface traps and oxide charges

Silicon is tetrahedrally bonded with each Si atom bonded to four Si atoms in the wafer bulk. When the Si is oxidized, the bonding configuration at the surface is as shown in Fig. 2(a) for (111) and (b) for (100) orientation with most Si atoms bonded to oxygen. Some Si atoms bond to hydrogen. An interface trapped charge, often called interface trap, is an interface trivalent Si atom with an unsaturated (unpaired) valence electron at the SiO_2/Si interface, denoted by

$$Si_3 \equiv Si\bullet \quad (3)$$

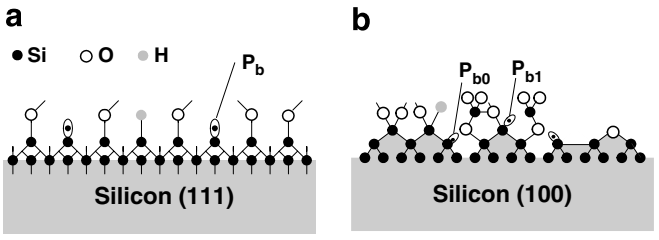

Fig. 2. Structural model of the (a): (111)Si surface and (b): (100)Si surface.

The "≡" represents three complete bonds to other Si atoms (the Si_3) and the "•" represents the fourth, unpaired electron in a dangling orbital (dangling bond). Interface traps are also known as P_b centers [9]. Interface traps are designated as D_{it} ($cm^{-2}\,eV^{-1}$), Q_{it} (C/cm^2), and N_{it} (cm^{-2}).

On (1 1 1)-oriented wafers, the P_b center is a $Si_3 \equiv Si\bullet$ center, situated at the Si/SiO_2 interface with its unbonded central-atom orbital perpendicular to the interface and aimed into a vacancy in the oxide immediately above it, as shown in Fig. 2(a). On (1 0 0)Si, the four tetrahedral Si–Si directions intersect the interface plane at the same angle. Two defects, named P_{b1} and P_{b0}, have been detected by electron spin resonance (ESR), shown in Fig. 2(b). The P_{b1} center was originally thought to be a Si atom backbonded to two substrate Si atoms, with the third saturated bond attached to an oxygen atom, designated as $Si_2O \equiv Si\bullet$. This identification was found to be incorrect, as the calculated energy levels for this defect do not agree with experiment [10]. A recent calculation suggests the P_{b1} center to be an asymmetrically oxidized dimer, with no first neighbor oxygen atoms [11]. By 1999, it was unambiguously established that both P_{b0} and P_{b1} are chemically identical to the P_b center [12]. However, there is a charge state difference between these two centers indicating P_{b0} is electrically active, while some authors believe the P_{b1} to be electrically inactive [13]. The two different effects are the result of strain relief in (1 0 0) silicon. The defects result from the naturally occurring mismatch-induced stress at the SiO_2/Si interface during oxide growth.

P_{b0} centers result when strain relaxation occurs with a defect residing at (1 1 1) microfacets at the Si/SiO_2 interface, while P_{b1} centers result when strain relaxation occurs with a defect at (1 0 0)Si/SiO_2 transition regions. Based on these results and the fact that P_{b1} centers are believed to be electrically inactive, defects resulting from P_{b0} centers are considered the key culprits in creating interface traps in (1 0 0) silicon. It is worth mentioning that recent work indicates P_{b1} centers to be electrically inactive at low temperatures ($T = 77$ K). However, at room temperature and higher these defects contribute to the electrical activity of total interface traps [14]. Recent ESR measurements show the P_{b1} center to be electrically active with two distinct, narrow peaks close to midgap in the silicon band gap [15]. However, P_{b1} centers are typically generated at densities considerably lower than P_{b0} centers, making them potentially less important.

Interface traps are electrically active defects with an energy distribution throughout the Si band gap. They act as generation/recombination centers and contribute to leakage current, low-frequency noise, and reduced mobility, drain current, and transconductance. Since electrons or holes occupy interface traps, they become charged and contribute to threshold voltage shifts. The surface potential dependence of the occupancy of interface traps is illustrated in Fig. 3.

Interface traps at the SiO_2/Si interface are acceptor-like in the upper half and donor-like in the lower half of the band gap [16]. Hence, as shown in Fig. 3(a), at flatband, with electrons occupying states below the Fermi energy, the states in the lower half of the band gap are neutral (occupied donors designated by "0"). Those between midgap and the Fermi energy are negatively charged (occupied acceptors designated by "−"), and those above E_F are neutral (unoccupied acceptors). For an inverted p-MOSFET, shown in Fig. 3(b), the fraction of interface traps between mid gap and the Fermi level is now unoccupied donors, leading to positively charged interface traps (designated by "+"). Hence interface traps in p-channel devices in inversion are *positively charged*, leading to negative threshold voltage shifts. Negative bias stress generates donor states in the lower half of the band gap [17,18].

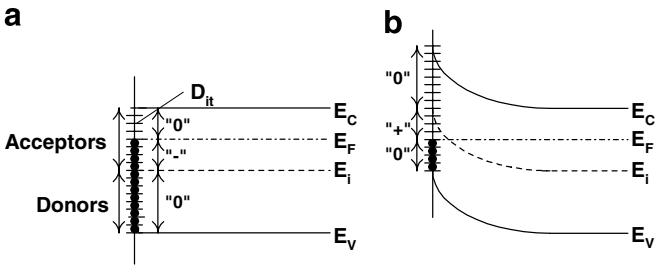

Fig. 3. Band diagrams of the Si substrate of a p-channel MOS device showing the occupancy of interface traps and the various charge polarities for a p-substrate with (a) negative interface trap charge at flatband and (b) positive interface trap charge at inversion. Each of the small horizontal lines represents an interface trap. It is either occupied by an electron (solid circle) or occupied by a hole (unoccupied by an electron), shown by the lines.

The oxide charge that contributes to NBTI is not well understood. Oxide charge can consist of various entities, including mobile charge, e.g., Na, K, Li, oxide trapped charge, e.g., electrons and/or holes, and fixed charge. We exclude mobile charge, because NBTI is observed when there is no mobile charge. Oxide charge is located within the oxide and may communicate with the Si conduction and valence bands and it is positive. Some believe it is H^+ trapped in the oxide near the SiO_2/Si interface. However, it is generally believed that hole trapping is the dominant mechanism and that the hole traps or their precursors may exist in the insulator prior to the stress.

The traps are positively charged when occupied by holes and neutral when unoccupied. It is possible that such positively charged traps can be neutralized by electrons when the n-substrate is at flatband and in accumulation. Similar traps have been proposed to be responsible for low-frequency (lf) noise where electrons or holes tunnel into traps [19]. The trap distance from the SiO_2/Si interface determines the tunnel time and hence the lf noise frequency response. Typical trap distances are 1–2 nm. The range 0.9–1.75 nm covers the frequency range 1–10^4 Hz where 10 Hz corresponds to ~2 nm distance. If carriers can tunnel over this frequency range that would correspond to tunnel times of 10^{-4}–1 s and suggests that the positive charge can be discharged in very short times, as suggested by NBTI experiments. Oxide charge located closer to the oxide/substrate

interface leads to higher threshold voltage shifts than charge near the gate/oxide interface. It is rarely pointed out that charge build-up in the oxide also alters the oxide electric field. Depending on the charge and their location, the oxide electric field may increase or decrease locally, e.g., near the SiO_2/substrate interface.

The NBTI threshold voltage change does not depend on the hole density, determined by changing the channel hole density and measuring the threshold voltage change. The hole density can be changed by changing V_T through fabrication or substrate bias. When this was done, there was very little change in interface trap generation, indicating that hole density is a secondary effect [4,20]. However, it appears that holes are required for NBTI degradation.

The p-MOSFET threshold voltage is

$$V_T = \phi_{MS} - \frac{Q_{ox}}{C_{ox}} - \frac{Q_{it}(2\phi_F)}{C_{ox}} - 2\phi_F - \frac{Q_S}{C_{ox}} \quad (4)$$

where ϕ_{MS} is the work function difference between the gate and substrate, ϕ_F the Fermi potential, Q_{ox} the positive oxide charge density, Q_S the semiconductor charge density, and C_{ox} the oxide capacitance/unit area. Q_{it}, shown to be dependent on the surface potential in Fig. 3, is given by

$$Q_{it} = qD_{it}\Delta E = qN_{it} \quad (5)$$

where ΔE is the energy range over which interface traps are active. Q_{it} may be positive or negative (Fig. 3).

Since neither gate nor substrate doping density nor oxide thickness change during stress, the threshold voltage *change* is due to changes in Q_{ox} and Q_{it} as

$$\Delta V_T = -\frac{\Delta Q_{ox} + \Delta Q_{it}}{C_{ox}} = -\frac{q(\Delta N_{ox} + \Delta N_{it})}{K_{ox}\varepsilon_0} t_{ox}$$
$$= -4.6 \times 10^{-7}(\Delta N_{ox} + \Delta N_{it}) t_{ox} \quad (6)$$

In today's ICs, N_{ox} and N_{it} are approximately $10^{10}\,cm^{-2}$ or slightly less *before* stress. Typical ΔV_T is on the order of -10 mV leading to $\Delta N_{ox} + \Delta N_{it} \approx 10^{11}\,cm^{-2}$ for $t_{ox} = 2$ nm, showing the NBTI-generated charges/states are typically higher than the starting values. How many SiH bonds are there to be dissociated? After Si oxidation, $N_{it} = 10^{12}$–$10^{13}\,cm^{-2}$ which is reduced to approximately $10^{10}\,cm^{-2}$ after low-temperature forming gas anneal. This suggests that ΔV_T may approach -100 mV or higher after prolonged stress before the 10^{12}–$10^{13}\,cm^{-2}$ SiH bonds are broken.

Question: While it is generally accepted that interface traps are one of the NBTI culprits, the oxide charge is less well understood. Is the positive charge due to hydrogen or trapped holes? Are the hole traps induced by NBTI stress or are they pre-existent in the oxide?

5. NBTI recovery

Until recently, most NBTI measurements were made by stressing the device, then measuring the threshold voltage, interface trap density, drain current, transconductance or other device parameter. But, there is usually a time delay between stress and characterization and that delay time was frequently not given in the published papers and it may have varied widely. More recently, it was found that the time delay is very important, because the stress damage recovers very rapidly. It may take many seconds to generate NBTI damage, but it recovers within a few seconds! Does the NBTI damage recover completely? That is not clear and the reported values depend on the authors. Ershov et al. propose that there are two degradation components: a permanent component which remains after stress removal and a reversible component that recovers [21]. Rangan et al., on the other hand, show complete recovery with the device being reset to its original state after the stress is removed [22]. Huard et al. have done extensive measurements and find partial recovery [4]. They attribute NBTI degradation to interface trap generation and hole trapping, N_{ot} and claim the recovery to be due to N_{ot} recovery with D_{it} remaining largely unchanged during the recovery phase. Fig. 4 shows the threshold voltage and interface trap density change during stress and recovery periods. Both ΔV_T and ΔN_{it} change proportionally during stress, but only ΔV_T shows a significant drop during the recovery phase with ΔN_{it} remaining almost constant. In contrast, Yang et al. show the threshold voltage recovery after stress cessation to be primarily due to interface trap passivation, not due

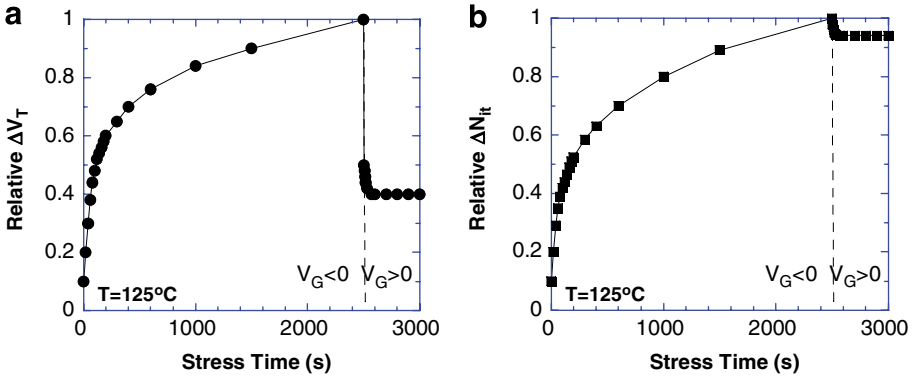

Fig. 4. Relative shifts for (a) ΔV_T and (b) ΔN_{it} versus stress time for negative and positive gate voltages [4].

to hole de-trapping [23]. This passivation effect is enhanced by positive gate bias during the recovery phase.

Why does the N_{ot}-controlled portion of V_T recovery occur? If N_{ot} is due to trapped holes, then the holes can be emitted during recovery. For $+V_G$ recovery bias, it is possible that accumulation electrons near the Si surface are injected or tunnel into the oxide to neutralize the trapped holes or that holes are "pushed" out of the traps. How do the hole traps originate? The trap generation is possibly related to hydrogen released from the SiH bonds at the interface and more traps are generated in nitrided oxides than in SiO_2. But what is the mechanism?

Tsujikawa et al. investigated NBTI recovery and found that V_T recovered more than D_{it} with zero or positive gate voltage during recovery [24]. They attributed N_{ot} to hole trapping induced by hydrogen and these positive charges are neutralized by electrons in the accumulation layer when the MOSFET is biased to $+V_G$. They correlated NBTI damage to stress-induced leakage current (SILC) and time-dependent dielectric breakdown (TDDB) through oxide traps generated during NBTI stress. Ang, during NBTI recovery measurements, found that while V_T recovers for $V_G = 0$ and $V_G > 0$, D_{it} hardly recovers at all for $V_G > 0$ [25] in agreement with Huard et al., showing that D_{it} recovery is suppressed under $+V_G$ recovery bias. However, the mechanism is not yet understood. If D_{it} hardly recovers, does that imply that hydrogen does not diffuse back to the SiO_2/Si substrate interface during recovery?

NBTI recovery leads to less severe device/circuit degradation under ac operation compared to dc and depends on the gate voltage duty cycle [21]. Most NBTI degradation measurements, however, are made with dc bias. The lifetime enhancement can be significant. The accelerated lifetime factor for ac operation is

$$F = \left(\frac{t_2}{t_1}\right)^n \Rightarrow t_2 = t_1 F^{1/n} \tag{7}$$

Eq. (7) predicts lifetime enhancements of $t_2/t_1 = 16$ for $n = 0.25$ and 76 for $n = 0.16$ for 50% ac NBTI degradation, pointing out the importance of an accurate knowledge of the n-factor. Since, as we show later, n depends on the time lag between stress and recovery, it then also depends on the frequency or duty cycle of the ac signal.

Question: What fraction of recovery is due to interface traps and oxide charge? Are oxide traps generated during NBTI stress? How does N_{ot} recovery work?

6. Why are p-MOSFETs different from n-MOSFETs?

NBTI is observed in both p-MOSFETs and n-MOSFETs. However, the effect is more extreme in p-MOSFETs. What is the reason for this? Fig. 5 shows the threshold voltage shift from two different research groups when p- and n-MOSFETs are biased with positive and negative gate voltages [17,4]. Clearly p-MOSFETs under negative gate bias are the most severely affected. It is commonly assumed that holes are necessary for NBTI degradation and since n-MOSFETs biased into accumulation also have holes at the surface, they should show similar V_T degradation. But they do not. We will put forward several explanations for this.

Tsetseris et al. propose the following NBTI model [26]. Their first-principles calculations show that positively charged hydrogen or protons, H^+, react directly with the SiH to form interface traps, according to the reaction

$$Si_3 \equiv SiH + H^+ \rightarrow Si_3 \equiv Si\bullet + H_2 \tag{8}$$

where $Si_3 \equiv SiH$ is a hydrogen-terminated interface trap and $Si_3 \equiv Si\bullet$ an interface trap with the dot representing the dangling bond. As shown in Fig. 6, the hydrogen is assumed to originate from phosphorus-hydrogen bonds in the n-Si substrate. The P–H bonds dissociate and the hydrogen on the way to the SiO_2/Si interface "picks up" a hole to become H^+, then reacts with the H from the SiH bond to form H_2 leaving behind a positively charged

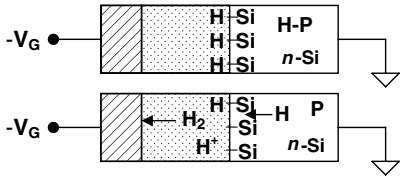

Fig. 6. Possible interface trap creation by hydrogen.

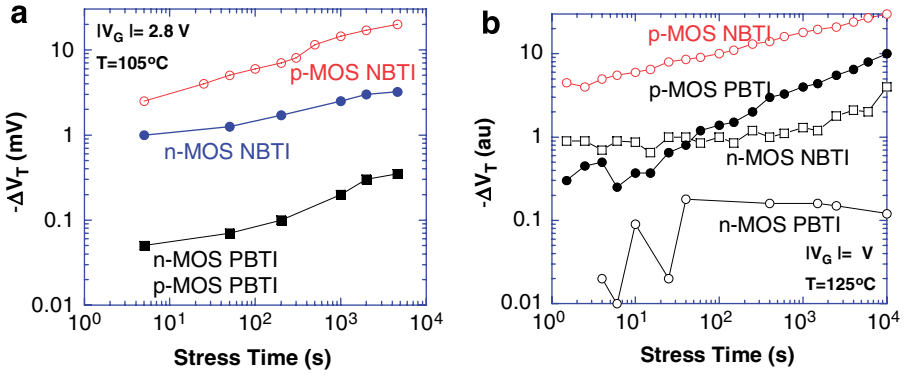

Fig. 5. Threshold voltage shifts for p- and n-MOSFETs for positive and negative gate bias [17,4].

Si dangling bond (or trapping center). The H_2 diffuses from the interface into the oxide or poly-Si gate. It can later passivate a dangling bond by diffusing back to the interface when the stress voltage is interrupted. The reason for the reduced NBTI activity in n-MOSFETs in this model is that it is more difficult for boron-hydrogen bonds to be broken in p-Si substrates. So the different behavior of p- versus n-MOSFETs in this model is due to the ease or difficulty of breaking P–H and B-H bonds in the Si substrate. Threshold voltage shifts in this case are due to interface traps, oxide charge, and a change in the substrate doping density after P–H or B-H depassivation.

Another explanation depends on the N_{it} and N_{ot} charge states. Interface traps, being *acceptors* in the upper half of the band gap and *donors* in the lower half, affect V_T shifts in n- and p-MOSFETs differently. Fig. 7 shows the band diagram of n-channel in (a) and p-channel devices in (b). At flatband, the n-channel has *positive* and the p-channel has *negative* interface trap charge. At inversion, $\phi_s \approx |2\phi_F|$, the n-channel has *negative* and the p-channel has *positive* interface trap charge. Since the oxide charge is positive in both cases, we have at inversion: n-channel: $Q_{ox} - Q_{it}$, p-channel: $Q_{ox} + Q_{it}$, hence p-channel MOSFETs are more severely affected. This was clearly shown by Sinha and Smith where the threshold voltage of MOS capacitors on (1 1 1)n-Si decreases by 1.5 V while V_T of (1 1 1)p-Si decreases by only about 0.2 V [27].

Let us put this into quantitative terms. For $\Delta N_{ox} = 10^{11}\,\mathrm{cm}^{-2}$ and $\Delta D_{it} = 10^{11}\,\mathrm{cm}^{-2}\,\mathrm{eV}^{-1}$, we have

$$\Delta V_T\ (\text{p-MOSFET}) = -4.6 \times 10^{-7}(\Delta N_{ox} + \Delta D_{it} E_G/2) t_{ox} = -14.3\ \text{mV} \quad (9a)$$

$$\Delta V_T\ (\text{n-MOSFET}) = -4.6 \times 10^{-7}(\Delta N_{ox} - \Delta D_{it} E_G/2) t_{ox} = -4.3\ \text{mV} \quad (9b)$$

assuming that under strong inversion, the Fermi level at the surface coincides with the valence band in p-MOSFETs and with the conduction band in n-MOSFETs, giving $\Delta E \approx E_G/2$. So the different behavior of p- versus n-MOSFETs in this model is due to the opposing sign of the interface traps.

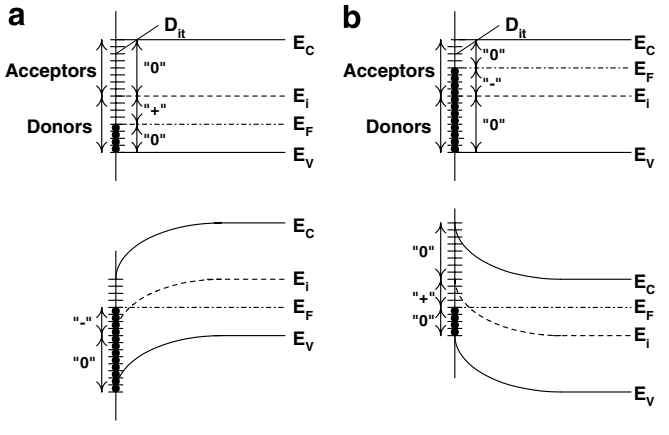

Fig. 7. Band diagrams of the Si substrate showing the occupancy of interface traps and the various charge polarities. (a) p-Substrate with positive interface trap charge at flat band and negative interface trap charge at inversion. (b) n-Substrate with negative interface trap charge at flat band and positive interface trap charge at inversion.

Another reason for the n- versus p-channel behavior has to do with the surface potential–gate voltage behavior, i.e., the gate voltages on n- and p-channel devices are not the same for a given oxide electric field [5]. Since this has not been treated in detail, I give the relevant equations. Consider the band diagrams in Fig. 8, drawn for uniformly doped substrates without threshold voltage adjust implants for simplicity. I use the band diagrams to explain the effect and then use more exact equations for numerical values. The *flatband* voltages differ by about one band gap, assuming a p^+ poly-Si gate with $E_F = E_V$ for the p-channel and $E_F = E_C$ for the n^+ gate of the n-channel devices. For NBTI, the p-channel device is biased into *inversion* and the n-channel device into *accumulation*, causing the p^+ gate to be depleted and the n^+ gate to be accumulated. Clearly, the band diagrams for negative NBTI gate bias are not symmetrical, although they are symmetrical at flatband.

With threshold voltage adjust implant, the gate voltage is given by

$$V_G = V_{FB} + \phi_{s,S} + \phi_{s,G} + V_{ox} + V_i \quad (10)$$

where V_{FB} is the flatband voltage, $\phi_{s,S}$ and $\phi_{s,G}$ the substrate and gate surface potentials, $V_{ox} = Q_S/C_{ox}$ the oxide voltage, and $V_i = Q_i/C_{ox}$ ($Q_i = qN_i$) the threshold voltage adjust implant voltage with N_i the implant dose. I will use: $N_{A,D} = 10^{17}\,\mathrm{cm}^{-3}$, $N_G = 5 \times 10^{19}\,\mathrm{cm}^{-3}$, $n_i = 10^{10}\,\mathrm{cm}^{-3}$, $t_{ox} = 2$ nm, $T = 300$ K, and $E_G/q = 1.12$ V. The flatband voltages for the p- and n-channel devices, assumed to be due only to work function differences, are

$$V_{FB,p} = \frac{E_G}{2q} + \phi_{F,p} = \frac{E_G}{2q} + \frac{kT}{q}\ln\left(\frac{N_D}{n_i}\right) = 0.56 + 0.42 = 0.98\ \text{V} \quad (11a)$$

$$V_{FB,n} = -\frac{E_G}{2q} - \phi_{F,n} = -\frac{E_G}{2q} - \frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right) = -0.56 - 0.42 = -0.98\ \text{V} \quad (11b)$$

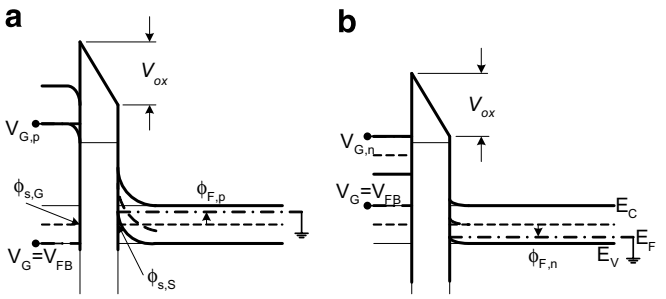

Fig. 8. Band diagrams for (a) p-channel and (b) n-channel MOSFETs. The poly-Si gate Fermi level is taken to be $E_{V,G}$ in (a) and $E_{C,G}$ in (b). The light lines are for $V_G = V_{FB}$ and the heavy lines for strong inversion/accumulation for the same V_{ox}.

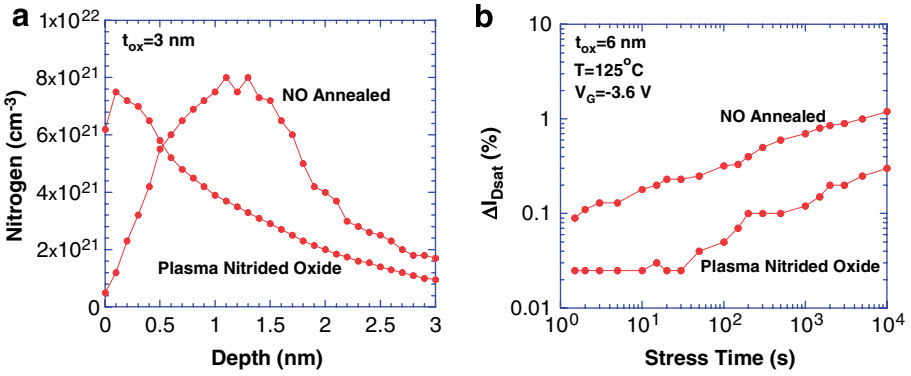

Fig. 9. (a) Nitrogen concentration versus depth and (b) drain current change versus stress time for nitrous oxide and plasma nitrided oxides showing the beneficial effects of plasma nitridation [29].

where ϕ_F is the Fermi potential. Assuming for the substrate $E_V = E_F$ at the surface for both devices for NBTI negative gate voltage bias gives

$$\phi_{s,Sp}\ (\text{inversion}) = -0.98\ \text{V},$$
$$\phi_{s,Sn}\ (\text{accumulation}) = -0.14\ \text{V} \qquad (12)$$

To determine V_i, the two threshold voltages need to be considered. V_T is

$$V_T = V_{FB} + 2\phi_F + \phi_{s,G} + Q_S/C_{ox} + V_i \qquad (13)$$

Substituting numerical values with $Q_S = \sqrt{2qK_s\varepsilon_0 N_S 2\phi_F} = \pm 0.14$ V, $\phi_F = \pm 0.42$ V, and using

$V_{T,p} = -0.4$ V, $V_{T,n} = 0.4$ V gives

$$V_{T,p} = -0.4 = 0.98 - 0.84 - 0.14 + V_i \Rightarrow V_i = -0.4\ \text{V} \qquad (14a)$$
$$V_{T,n} = 0.4 = -0.98 + 0.84 + 0.14 + V_i \Rightarrow V_i = 0.4\ \text{V} \qquad (14b)$$

assuming the gate surface potential is around zero at threshold.

For strong inversion

$$\begin{aligned} V_{G,p} &= V_{FB} + \phi_{s,Sinv} + \phi_{s,G} + V_{ox} + V_i \\ &\approx 0.98 - 0.94 - 0.56 + V_{ox} - 0.4 \\ &= V_{ox} - 0.92\ \text{V} \end{aligned} \qquad (15a)$$

assuming $\phi_{s,Sinv} \approx -(2\phi_F + 0.1) = -0.94$ V and for the depleted gate $\phi_{s,G} \approx -E_G/2q = -0.56$ V. Similarly for n-channel in accumulation

$$\begin{aligned} V_{G,n} &= V_{FB} + \phi_{s,Sacc} + \phi_{s,G} + V_{ox} + V_i \\ &\approx -0.98 - 0.14 + V_{ox} + 0.4 = V_{ox} - 0.72\ \text{V} \end{aligned} \qquad (15b)$$

assuming $\phi_{s,Sacc} \approx -0.14$ V and $\phi_{s,G} \approx 0$. Eqs. (15) show that for a given oxide voltage or given oxide electric field the gate voltages on the p- and the n-channel devices differ by 0.2 V in this fairly realistic example, i.e., the voltage difference for the same oxide electric field does not differ by the band gap, as sometimes inferred. Hence, measurements on p- and n-channel devices should take into account these considerations for equivalent oxide electric field.

Question: Is the p-channel/n-channel NBTI difference due to interface trap charge dependence on surface potential, hydrogen diffusing from the substrate or work function difference?

7. Nitrogen and fluorine

Nitrogen, water and hydrogen in the oxide degrade NBTI. Nitrogen is especially important as it is incorporated into many oxides today and its role in NBTI degradation is poorly understood. It has been suggested that nitrogen creates hole traps [28]. Plasma nitridation leads to reduced NBTI degradation compared to nitrous oxide illustrated in Fig. 9 [29]. During plasma nitridation the nitrogen diffuses from the top into the oxide with the nitrogen concentration skewed more heavily toward the gate/oxide part of the oxide. Excess nitrogen near the oxide/substrate interface leads not only to enhanced NBTI degradation but also to higher fixed oxide charge density [30]. Nitrogen should be located near the gate/oxide interface rather than the oxide/substrate interface for minimum NBTI degradation. Fluorine, on the other hand generally helps. Fluorine atoms release the distortion at the SiO_2/Si interface. Interface-state generation for fluorinated oxides under Fowler–Nordheim stress is suppressed and one might suspect that this also applies to NBTI [31]. NBTI has also been observed in high-K dielectrics.

8. Activation energy and time dependence

NBTI depends on temperature, oxide electric field, and time and the threshold voltage shift can be expressed by

$$\Delta V_T = A \exp(\gamma\varepsilon_{ox}) \exp(-E_A/kT) t^n \qquad (16)$$

where A is a constant, γ the electric field factor, E_A the activation energy and n the time exponent. It is obviously important to determine the various parameters in this equation to predict NBTI degradation for any time, temperature and oxide electric field. The degradation mechanism of hole trapping and hydrogen diffusion from the SiO_2/Si interface into the oxide is usually described by an activation energy and a time dependence, shown in Eq. (16). Neutral H_2 diffusion is described with $E_A \approx 0.2$–0.3 eV and the time

exponent $n \approx 0.25$. Frequently, however, very different values are measured for these parameters and various explanations have been forwarded. Recent measurements of nitrided oxides gave $0.019\ \text{eV} \leqslant E_A \leqslant 0.24\ \text{eV}$ and $0.095 \leqslant n \leqslant 0.158$ over the 183–513 K temperature range, in contrast to $E_A \approx 0.27$ eV and $n \approx 0.27$ for pure SiO_2 [32]. A more commonly-accepted value for dry SiO_2 is 0.1–0.15 eV. The value of 0.27 eV is likely due to partial wet oxide. Ang and Wang attribute the non-Arrhenius behavior to the superposition of two defect generation mechanisms with differing n and E_A: hole trapping with $n \approx 0.1$ and $E_A \approx 0.02$ eV and hydrogen diffusion with $n \approx 0.25$ and $E_A \approx 0.25$ eV [32]. At low temperatures the V_T shift is determined by hole trapping at nitrogen defect precursors. At the higher temperatures, classical hydrogen diffusion takes over. They suggest that previously published reports of temperature-dependent n values in thin SiO_2 [33] may be due to trace amounts of nitrogen in those oxides. This mechanism may account for the V_T recovery being due to "oxide charges" and not interface traps, observed by Huard et al. [4]. Why is NBTI oxide electric field dependent? This implies either a charged oxide species or the field aids the electrochemical reaction.

The threshold voltage degradation has also been modeled by the concept of disorder by [34].

$$|\Delta V_T| = \Delta V_{T\max}\left(1 - \exp\left(-\frac{t}{\tau}\right)^{\beta}\right) \tag{17}$$

where

$$\tau = \left(\frac{N_i}{kT\beta} D_0 \varepsilon_{ox} \exp\left(-\frac{\Delta E}{kT}\right)\right)^{-1/\beta};$$
$$D_H = D_0 t^{-\alpha} = D_0 t^{-(1-\beta)} \tag{18}$$

where N_i is the interstitial site density at the SiO_2/Si interface, β the dispersion coefficient, D_H the dispersive hydrogen diffusion coefficient, α the dispersion coefficient, ε_{ox} the oxide electric field, and ΔE the hydrogen dissociation energy. For short times, $|\Delta V_T| \approx \Delta V_{T\max}(t/\tau)^{\beta}$.

Huard et al. find n to increase with temperature and E_A to vary with stress time [4]. They explain this behavior with the SiH dispersed dissociation energy having various E_A. Each SiH bond follows its own Arrhenius behavior, but the multitude of bonds exhibits a non-Arrhenius behavior. Kaczer et al. use a disorder-controlled kinetics model in which the disorder due to energy distribution of deep localized hydrogen states in the SiO_2 leads to a wide distribution of hydrogen hopping times [35]. This predicts a temperature-dependent n factor. Nitrogen introduces deeper hydrogen states leading to lower n. The n exponent appears to depend on the measurement conditions. When there is a delay between NBTI stress and its characterization, $n \approx 0.25$. However, when the delay is zero, n decreases to ~0.15.

Krishnan et al. show the n exponent for interface trap generation to be 0.16 under the assumption that neutral H_2 diffuses from the SiO_2/Si interface into the oxide and diffusion, not interfacial reaction, is the rate-limiting process [36]. Their data are shown in Fig. 10(a). They posit that the $n = 0.25$ value frequently quoted, is due to the time delay between stress and measurement. In their measurements, they obtain $n = 0.3$ with 1 s delay between stress and N_{it} measurement. Varghese et al. show $n = 0.14$ for "no delay" measurements and $n > 0.14$ when there is a delay between stress and measurement, shown in Fig. 10(b) [37].

NBTI shows a tendency to saturate for long stress times, as shown in Fig. 4. This has been attributed to various causes. A recent paper by Alam and Kufluoglu lists four possible causes [38]. One suggests that a change of hydrogen diffusion from H to H_2 diffusion changes n from 0.25 to 0.16 leading to quasi-saturation of the ΔV_T vs. t characteristics. Another cause may be H_2 pile up at the gate/oxide interface if hydrogen diffuses slower in the poly-Si gate than in the oxide, leading to hydrogen back diffusion and subsequent D_{it} passivation. On the other hand, if hydrogen diffuses faster in the poly-Si gate, then hydrogen "disappears" from the oxide and saturation occurs when all SiH bonds are broken. The dispersion of the bond-breaking dissociation energies leads to low E_A bonds breaking first making it progressively more difficult for stronger bonds to dissociate. Finally, saturation may be a measurement artifact. Delays between stress and NBTI damage measurement lead to repassivation of interface traps. In the model in which NBTI is due to hydrogen

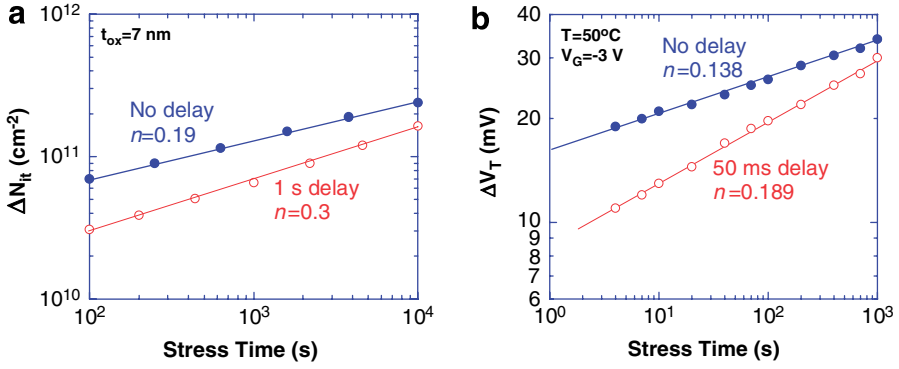

Fig. 10. (a) ΔN_{it}[36] and (b) ΔV_T[37] versus stress time with and without stress/measurement delay.

diffusion from the substrate, is it possible for the P–H bonds to deplete thereby leading to saturation?

Question: What is the true time dependence? Is it $n \approx 0.16$? What is the activation energy? Is the disorder-controlled diffusion mechanism correct? Which mechanism(s) is (are) responsible for the quasi-saturation?

9. Wafer orientation

Most ICs are fabricated on (100) oriented Si wafers. However, it has been known since 1968 that hole mobility is higher for p-MOSFETs on (110) oriented wafers with the channel in the ⟨110⟩ direction [39,40]. Along with the higher hole mobility, however, is a higher interface trap density. D_{it} has been reported as $D_{it}(100) = 7.7 \times 10^{10}$ and $D_{it}(110) = 1.4 \times 10^{11}\ \mathrm{cm^{-2}\,eV^{-1}}$ [41]. Hence one would expect more severe NBTI degradation on (110) oriented wafers, as indeed has been observed, as shown in Fig. 11 [42,43]. This is a potential problem if (110) oriented wafers become important. For certain three-dimensional devices, e.g., FinFETs, when fabricated on (100) wafers with the channel in the conventional ⟨110⟩ direction, the vertical sidewalls are (110) oriented, leading to NBTI problems [42]. However, forming FinFETs on (100) wafers with the channel in the ⟨100⟩ direction, leads to (100) vertical sidewalls, illustrated in Fig. 12. Three-dimensional structures lead to various difficulties, e.g., gate length critical dimension control and the nature of three-dimensional structures. It is possible, however, to form planar (100) and (110) oriented surfaces on one wafer using SOI layer transfer technology [44].

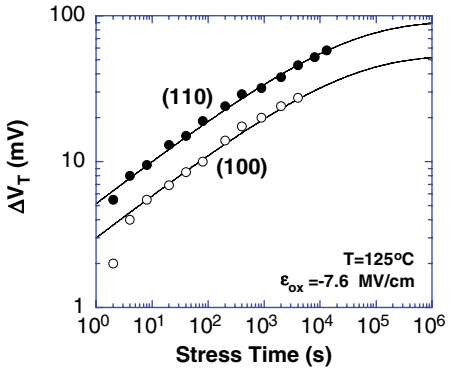

Fig. 11. ΔV_T versus stress time for SiON (2%) with (110) and (100) orientation. Points: experiment, lines: Eq. (17) with $\tau = 1.3 \times 10^4$ s, $\beta = 0.3$, $\Delta V_{Tmax}(110) = 91$ mV and $\Delta V_{Tmax}(100) = 53$ mV [43].

10. Measurements

Traditionally NBTI has been characterized by stressing a device, then interrupting the stress to measure a device parameter, e.g., V_T, I_D, μ_{eff}, D_{it}, N_{ot}, stress again, measure, etc. V_T, I_D, and μ_{eff} are determined from I_D–V_D and I_D–V_G measurements. The interface trap density D_{it} is usually measured with charge pumping [45] or log(I_D)–V_G subthreshold swing measurements. Usually source, drain and substrate are grounded during the stress period. The "dead" time between "stress" and "measure" varied from researcher to researcher and was frequently not given in publications. When it was discovered that that time was quite important, techniques were developed to minimize and eliminate the "dead" time. Rangan et al. use $V_D = 50$ mV and monitor the drain current during stress and recovery [22]. Drain current depends on effective mobility and threshold voltage and in this method neither is measured explicitly, but the drain current degradation is determined.

Huard et al. in their "on the fly" method, apply a low drain voltage pulse of ~50 mV and record the drain current at the end of the stress period, without interrupting the stress, illustrated in Fig. 13(a) [4]. This was repeated and in this way drain current degradation was recorded with no "dead" time. Since the current depends on threshold voltage, mobility and other device parameters, it is not

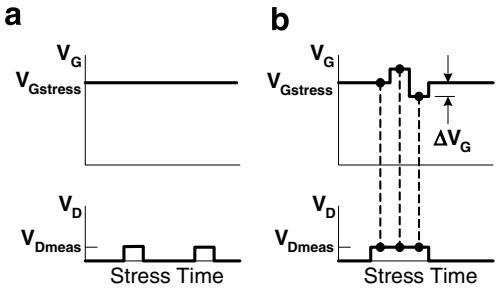

Fig. 13. Measurement schematics for (a) drain current and (b) drain current and transconductance without interrupting stress.

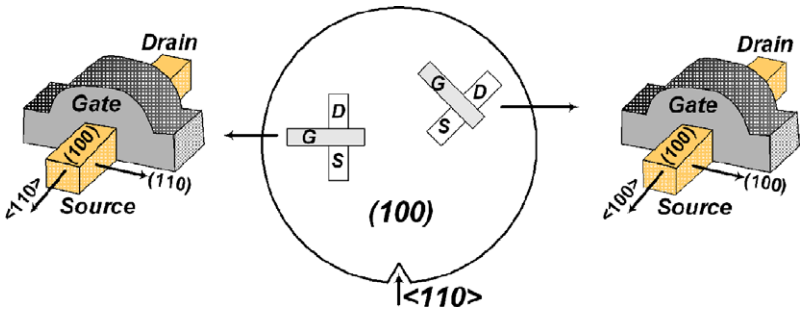

Fig. 12. Top view of (100) oriented Si wafer with FinFET-like transistors with channels in the ⟨110⟩ and ⟨100⟩ directions.

possible to extract ΔV_T. In a modification of this method, shown in Fig. 13(b), small gate voltage pulses are superimposed on the dc gate stress voltage. The drain current is also pulsed and the drain current is measured at three different gate voltages, allowing the transconductance to be determined.

In yet another method, the I_D–V_G characteristic of an unstressed device is measured. Then the gate voltage is lowered briefly from its stress value to a voltage near V_T and the drain current is measured [35]. The gate voltage shift of the stressed drain current is converted to ΔV_T.

From the linear drain current a number of parameters can be derived. The linear drain currents I_{Dlin} for p-MOSFETs in the mobility-dominated and velocity saturation-limited regimes are [46]

$$I_{Dlin} \approx k\mu_{eff}(V_{GS} - V_T)V_{DS};$$
$$I_{Dlin} \approx \frac{k\mu_{eff}}{1 + \mu_{eff}V_{DS}/v_{sat}L}(V_{GS} - V_T)V_{DS}; \quad (19)$$
$$\mu_{eff} = \frac{\mu_0}{1 + \theta(V_{GS} - V_T)}; \quad k = \frac{WC_{ox}}{L}$$

where μ_0 is the low-field mobility, θ the mobility degradation factor, and v_{sat} the saturation velocity. All voltages are negative for p-channel MOSFETs and given as absolute values in Eq. (19). If the change in drain current is compared to the unstressed, I_{Dlin0}, the change in I_{Dlin} is [47]

$$\frac{\Delta I_{Dlin}}{I_{Dlin0}} = -\frac{\Delta V_T}{V_{GS} - V_{T0}} \Rightarrow \Delta V_T = -(V_{GS} - V_{T0})\frac{\Delta I_{Dlin}}{I_{Dlin0}} \quad (20)$$

where V_{T0} is the unstressed threshold voltage and any change in mobility in deriving Eq. (19) is neglected. This, of course, is an approximation since we know that interface traps generated during NBTI stress degrade the mobility by introducing scattering centers. However, the mobility dependence on gate voltage decreases at higher gate voltages.

The mobility dependence can be eliminated by measuring the transconductance in Fig. 13(b) according to

$$g_m \approx \frac{k\mu_0 V_{DS}}{[1 + \theta(V_{GS} - V_T)]^2};$$
$$g_m \approx \frac{k\mu_0 V_{DS}(1 + \mu_0 V_{DS}/v_{sat}L)}{[1 + \theta(V_{GS} - V_T) + \mu_0 V_{DS}/v_{sat}L]^2} \quad (21)$$

and

$$\frac{I_{Dlin}}{\sqrt{g_m}} = \sqrt{k\mu_0 V_{DS}}(V_{GS} - V_T);$$
$$\frac{I_{Dlin}}{\sqrt{g_m}} = \frac{\sqrt{k\mu_0 V_{DS}}(V_{GS} - V_T)}{\sqrt{1 + \mu_0 V_{DS}/v_{sat}L}} \quad (22)$$

Eq. (19) neglects the source/drain parasitic resistance. Including it adds another term, but does not alter Eq. (22) [48]. The mobility degradation factor θ has been eliminated in Eq. (22), but it still contains the low field mobility. The determination of μ_0 is discussed in detail by Huard et al. [4].

Question: How much time can be allowed between stress and measurement? Must it be zero?

11. Effect on circuits

The drain or *on* current is important in analog and digital circuits. In digital circuits, with MOSFETs being switches, charging and discharging capacitors, higher drain current leads to faster capacitor charging and higher frequency operation. The delay time is

$$t_d = \frac{C|V_{DD}|}{I_D} = \frac{2LC}{W\mu_{eff}C_{ox}(V_{DD} - V_T)^2} \quad (23)$$

where C is the capacitance and V_{DD} the supply voltage. NBTI stress leads to μ_{eff} reduction and V_T increase, both giving delay time increases – clearly undesirable. In analog circuits the drain current is also important. An interesting effect that occurs due to NBTI degradation is an increased gate-to-drain overlap capacitance increase due to generated interface traps. Such increased C_{GD} degrades analog circuits due to the Miller effect. For digital circuits V_T and C_{GD} degradation contribute about equally to circuit degradation while for analog circuits, C_{GD} is the major circuit degrader, e.g., frequency response of operational amplifiers.

The fractional change of drain current due to threshold voltage change is

$$I_D \approx \frac{W\mu_{eff}C_{ox}}{L}(V_G - V_T)V_D \Rightarrow \frac{1}{I_D}\frac{dI_D}{dV_T} = -\frac{\Delta V_T}{V_G - V_T} \quad (24)$$

for MOSFETs in the linear region, and

$$I_{Dsat} \approx \frac{W\mu_{eff}C_{ox}}{2L}(V_G - V_T)^2 \Rightarrow \frac{1}{I_D}\frac{dI_D}{dV_T} = -\frac{2\Delta V_T}{V_G - V_T} \quad (25)$$

for the saturation region, showing twice the degradation in saturation, as observed experimentally [47]. With scaling, device dimensions and voltages shrink. However, the gate voltage typically shrinks more than V_T, hence the headroom (V_G–V_T) is smaller for shorter channel devices resulting in more degradation. For example, with $V_G = -1.5$ V, $V_T = -0.4$ V and $\Delta V_T = -0.05$ V, the fractional drain current degradation from Eqs. (24) and (25) becomes 4.5% and 9%, respectively. The fractional change due to mobility change is

$$\frac{1}{I_D}\frac{dI_D}{d\mu_{eff}} = \frac{\Delta\mu_{eff}}{\mu_{eff}} \quad (26)$$

The mobility degrades due to interface trap generation, since μ_{eff} depends inversely on D_{it}.

The one parameter that appears to be positively influenced by NBTI is the *off* current, i.e., the drain current for $V_G = 0$. The *off* current for p-MOSFETs is given by

$$I_{\text{off}} = I_{\text{T}} \exp\left(\frac{qV_{\text{T}}}{nkT}\right); \quad n = 1 + \frac{C_{\text{b}} + C_{\text{it}}}{C_{\text{ox}}} = 1 + \frac{K\sqrt{N_{\text{D}}} + qD_{\text{it}}}{C_{\text{ox}}} \tag{27}$$

where to first order I_{T} and K are constants. As V_{T} becomes more negative, I_{off} decreases. However, D_{it} increases resulting in higher n and reduced slope of the $\log(I_{\text{D}})$–V_{G} characteristics. This leads to a slight I_{off} increase, but the V_{T} change generally dominates. The n-factor also becomes less important as the oxide capacitance increases for thinner oxides.

For SRAM cells the static noise margin (SNM), which is a measure of the read stability of the 6-T SRAM cell, is degraded. A simple solution to recover the SNM uses a data flipping technique [49]. Knowing the threshold voltage degradation of a single transistor due to NBTI, one can predict the performance degradation of a digital circuit with a reasonable degree of accuracy.

Digital circuits are less sensitive (approximately 9.2% performance degradation in ten years for 70 nm technology) to NBTI degradation than previously anticipated [50].

Question: How severe a problem is NBTI for digital and analog circuits? Which circuits are most severely affected? How much threshold voltage shift can a circuit tolerate?

12. Conclusions

A brief review of NBTI is followed by a summary of recent publications elucidating various NBTI mechanisms. I have tried to highlight those aspects of NBTI that are still not fully understood – at least not by me. I have used the most relevant equations to bring out particular points. At the end of most sections I have posed some questions that I believe are still unanswered. Although NBTI is much better understood today than a few years ago, the very mechanism causing NBTI is not completely understood. For example, I do not believe that we really know what causes the positive oxide charge, although there is little doubt about what causes interface trap changes. There is a fair amount of discrepancy on the time dependence of NBTI degradation; is it $n = 0.15$ or 0.25? While that may appear marginally important, it becomes very important in predicting long-term NBTI degradation. The reported activation energy also covers a wide range. What exactly does nitrogen in nitrided oxides do? It appears that the amount and location of nitrogen are important. The further the nitrogen is from the oxide/substrate interface, the less is the NBTI degradation. Most NBTI measurements are made at elevated temperatures and some high-performance ICs operate at or near those temperatures. However, many ICs operate at lower temperature. Is the degradation mechanism the same at various temperatures? This, of course, is always a problem when devices are stressed during reliability measurements. The stress (electromigration, gate oxide integrity, hot carriers) is usually carried out at higher current, voltage, temperature and the degradation is extrapolated to normal operating conditions to very long times, e.g., 10 years. This relies on the degradation mechanisms during stress and operation being the same. Several important NBTI degradation questions remain unanswered. However, I am confident that the world-wide attention it has garnered will lead to a better understanding. Also, very little has been published on NBTI reduction/elimination, perhaps because much of this is proprietary information.

Acknowledgements

I thank the reviewers for constructive and thoughtful comments and suggestions, many of which were incorporated into the paper.

References

[1] Miura Y, Matukura Y. Investigation of silicon–silicon dioxide interface using MOS structure. Jpn J Appl Phys 1966;5:180;
Goetzberger A, Nigh HE. Surface charge after annealing of Al–SiO_2–Si structures under bias. Proc IEEE 1966;54:1454.
[2] Schroder DK, Babcock JA. Negative bias temperature instability: a road to cross in deep submicron silicon semiconductor manufacturing. J Appl Phys 2003;94:1–18.
[3] Stathis JH, Zafar S. The negative bias temperature instability in MOS devices: a review. Microelectron Rel 2006;46:270–86.
[4] Huard V, Denais M, Parthasarathy C. NBTI degradation: from physical mechanisms to modeling. Microelectron Rel 2006;46:1–23.
[5] Alam MA, Mahapatra S. A comprehensive model of PMOS NBTI degradation. Microelectron Rel 2005;45:71–81.
[6] Jeppson KO, Svensson CM. Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices. J Appl Phys 1977;48:2004–14.
[7] Ogawa S, Shiono N. Generalized diffusion–reaction model for the low-field charge build up instability at the Si–SiO_2 interface. Phys Rev 1995;51:4218–30.
[8] Yang JB, Chen TP, Tan SS, Chan L. Analytical reaction–diffusion model and the modeling of nitrogen-enhanced negative bias temperature instability. Appl Phys Lett 2006;88:172109-1–9-3.
[9] Poindexter EH, Caplan PJ. Characterization of Si/SiO_2 interface defects by electron spin resonance. Progr Surf Sci 1983;14:201–94.
[10] Reed ML, Plummer JD. Chemistry of Si–SiO_2 interface trap annealing. J Appl Phys 1988;63:5776–93.
[11] Stirling A, Pasquarello A, Charlier JC, Car R. Dangling bond defects at SiO_2/Si interfaces: atomic structure of the P_{b1} center. Phys Rev Lett 2000;85:2773–6.
[12] Stesmans A, Afanasév VV. Nature of the P_{b1} defect in (100)Si/SiO_2 as revealed by electron spin resonance ^{29}Si hyperfine structure. Microelectron Eng 1999;48:113–6.
[13] Stesmans A, Afanasév VV. Electrical activity of interfacial paramagnetic defects in thermal (100)Si/SiO_2. Phys Rev 1998;57:10030–4.
[14] Lenahan P, IEEE Rel. Phys. Tutorial Notes, Adv. Rel. Topics, IEEE IRPS, 2002; section 223.
[15] Mishima TD, Lenahan PM, Weber W. Do P_{b1} centers have levels in the Si band gap? Spin-dependent recombination study of the P_{b1} "hyperfine spectrum". Appl Phys Lett 2000;76:3771–3;
Campbell JC, Lenahan PM. Density of states of P_{b1} Si/SiO_2 interface trap centers. Appl Phys Lett 2002;80:1945–7.
[16] Gray PV, Brown DM. Density of SiO_2–Si interface states. Appl Phys Lett 1966;8:31–3;
Fleetwood DM. Long-term annealing study of midgap interface-trap charge neutrality. Appl Phys Lett 1992;60:2883–5.
[17] Reddy V, Krishnan AT, Marshall A, Rodriguez J, Natarajan S, Rost Y, et al. Impact of negative bias temperature instability on digital circuit reliability. IEEE Ann Int Rel Symp 2002;40:248–54.

[18] Saminadayar K, Pfister JC. Evolution of surface-states density of Si/wet thermal SiO_2 interface during bias-temperature treatment. Solid-State Electron 1977;20:891–6.
[19] Jayaraman R, Sodini CG. A 1/f noise technique to extract the oxide-trap density near the conduction band edge of silicon. IEEE Trans Electron Dev 1989;36:1773–82.
[20] Mitani Y, Nagamine M, Satake H, Toriumi A. NBTI mechanism in ultra-thin gate dielectric-nitrogen-originated mechanism in SiON. IEEE Int Electron Dev Meet 2002:509–12.
[21] Ershov M, Saxena S, Karbasi H, Winters S, Minehane S, Babcock J, et al. Dynamic recovery of negative bias temperature instability in p-type metal–oxide–semiconductor field-effect transistors. Appl Phys Lett 2003;83:1647–9.
[22] Rangan S, Mielke N, Yeh ECC. Universal recovery behavior of negative bias temperature instability. IEEE Int Electron Dev Meet 2003:341–4.
[23] Yang T, Shen C, Li MF, Ang CH, Zhu CX, Yeo YC, et al. Interface trap passivation effect in NBTI measurement for p-MOSFET with SiON gate dielectric. IEEE Electron Dev Lett 2006;26:758–60.
[24] Tsujikawa S, Watanabe K, Tsuchiya R, Ohnishi K, Yugami J. Experimental evidence for the generation of bulk traps by negative bias temperature stress and their impact on the integrity of direct-tunneling gate dielectrics. IEEE VLSI Technol 2003:139–40.
[25] Ang DS. Observation of suppressed interface state relaxation under positive gate biasing of the ultrathin oxynitride gate p-MOSFET subjected to negative-bias temperature stressing. IEEE Electron Dev Lett 2006;27:412–5.
[26] Tsetseris L, Zhou XJ, Fleetwood DM, Schrimpf RD, Pantelides ST. Physical mechanism of negative-bias temperature instability. Appl Phys Lett 2005;86:142103-1–3-3.
[27] Sinha AK, Smith TE. Kinetics of the slow-trapping instability at the Si/SiO_2 interface. J Electrochem Soc 1978;125:743–6.
[28] Kushida-Abdelghafar K, Watanabe K, Ushio J, Murakami E. Effect of nitrogen at SiO_2/Si interface on reliability issues – negative-bias-temperature instability and Fowler–Nordheim-stress degradation. Appl Phys Lett 2002;81:4362–4.
[29] Tavel B, Bidaud M, Emonet N, Barge D, Planes N, Brut H, et al. Thin oxynitride solution for digital and mixed-signal 65 nm CMOS platform. IEEE Int Electron Dev Meet 2003:643–6.
[30] Ishimaru K, Takayangi M, Watanabe T. Scaled CMOS with SiON and high *k*. In: Huff H, Iwai H, Richter H, editors. Silicon materials science and technology. Pennington, NJ: Electrochem. Soc.; 2006. p. 317–27.
[31] Mitani Y, Satake H, Nakasaki Y, Toriumi A. Improvement of charge-to-breakdown distribution by fluorine incorporation into thin gate oxides. IEEE Trans Electron Dev 2003;50:2221–6.
[32] Ang DS, Wang S. On the non-Arrhenius behavior of negative-bias temperature instability. Appl Phys Lett 2006;88:093506-1–6-3.
[33] Kaczer B, Arkhipov V, Jurczak M, Groeseneken G. Negative bias temperature instability (NBTI) in SiO_2 and SiON gate dielectrics understood through disorder-controlled kinetics. Microelectron Eng 2005;80:122–5.
[34] Zafar S, Lee BH, Stathis J, Callegari A, Ning T. A model for negative bias temperature instability (NBTI) in oxide and high *K* pFETs. IEEE VLSI Technol 2004:208–9.
[35] Kaczer B, Arkhipov V, Degraeve R, Collard N, Groeseneken G, Goodwin M. Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification. IEEE Ann Int Rel Symp 2005;43:381–7.
[36] Krishnan AT, Chakravarthi S, Nicollian P, Reddy V, Krishnan S. Negative bias temperature instability mechanism: the role of molecular hydrogen. Appl Phys Lett 2006;88:153518-1–8-3.
[37] Varghese D, Saha D, Mahapatra S, Ahmed K, Nouri F, Alam A. On the dispersive versus Arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: measurements, theory, and implications. IEEE Int Electron Dev Meet 2005:684–7.
[38] Alam M, Kufluoglu H. On quasi-saturation of negative bias temperature degradation. In: Massoud HZ, Stathis JH, Hattori T, Misra D, Baumvol I, editors. Physics and chemistry of SiO_2 and the $Si–SiO_2$ interface – 5. Pennington, NJ: Electrochem. Soc.; 2005. p. 139–46.
[39] Colman D, Bate RT, Mize JP. Mobility anisotropy and piezoresistance in silicon p-type inversion layers. J Appl Phys 1968;39:1923–31.
[40] Sato T, Takeishi Y, Hara H. Effect of crystallographic orientation on mobility, surface state density, and noise in p-type inversion layers on oxidized silicon surfaces. Jpn J Appl Phys 1969;8:588–98.
[41] Sze SM, editor. VLSI Technology. 2nd ed. NY: McGraw-Hill; 1988. p. 110.
[42] Maeda S, Choi JA, Yang JH, Jin YS, Bae SK, Kim YW, et al. Negative bias temperature instability in triple gate transistors. IEEE Int Reliab Phys Symp 2004;42:8–12.
[43] Zafar S, Yang M, Gusev E, Callegari A, Stathis J, Ning T, et al. A comparative study of NBTI as a function of Si substrate orientation and gate dielectrics (SiON) and $SiON/HfO_2$. IEEE VLSI Technol 2005:128–9.
[44] Yang M, Chan VWC, Chan KK, Shi L, Fried DM, Stathis JH, et al. Hybrid-orientation technology (HOT): opportunities and challenges. IEEE Trans Electron Dev 2006;53:965–78.
[45] Schlünder C, Brederlow R, Wieczorek P, Dahl C, Holz J, Röhner M, et al. Trapping mechanisms in negative bias temperature stressed p-MOSFETs. Microelectron Rel 1999;39:821–6.
[46] Taur Y, Ning TH. Fundamentals of Modern VLSI Devices. Cambridge University Press; 1998. p. 150–1.
[47] Krishnan AT, Reddy V, Chakravarthi S, Rodriguez J, John S, Krishnan S. NBTI impact on transistor and circuit: models, mechanisms, and scaling effects. IEEE Int Electron Dev Meet 2003:349–52.
[48] Ghibaudo G. New method for the extraction of MOSFET parameters. Electron Lett 1988;24:543–5.
[49] <http://doi.ieeecomputersociety.org/10.1109/ISQED.2006.73>.
[50] Paul BC, Kang K, Kufluoglu H, Alam MA, Roy K. Impact of NBTI on the temporal performance degradation of digital circuits. IEEE Electron Dev Lett 2005;26:560–2.