APPLIED PHYSICS REVIEWS—FOCUSED REVIEW

Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing

Dieter K. Schroder[a)]
Department of Electrical Engineering and Center for Low Power Electronics, Arizona State University, Tempe, Arizona 85287-5706

Jeff A. Babcock
PDF Solutions, Inc., 101 West Renner Road, Suite 325, Richardson, Texas 75082

(Received 11 July 2002; accepted 17 February 2003)

We present an overview of negative bias temperature instability (NBTI) commonly observed in *p*-channel metal–oxide–semiconductor field-effect transistors when stressed with negative gate voltages at elevated temperatures. We 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 the effects of varying parameters (hydrogen, deuterium, nitrogen, nitride, water, fluorine, boron, gate material, holes, temperature, electric field, and gate length) on NBTI. We conclude with the present understanding of NBTI and its minimization. [DOI: 10.1063/1.1567461]

TABLE OF CONTENTS

[a)]Author to whom correspondence should be addressed; electronic mail: schroder@asu.edu

I. MOTIVATION

For more then 30 years the semiconductor industry has witnessed exceptional growth and achievements in integrated circuit (IC) manufacturing. In the last 10 years, this growth has outpaced Moore's Law,[1] causing numerous modifications of the SIA roadmap to ensure leading edge semiconductor manufacturing remains on target for high performance products and integration of system-on-a-chip (SOC) where both analog and digital signal processing can occur. In spite of some claims that the industry will hit a ''red brick'' wall at the 100 nm technology node just 4 years ago,[2] leading edge research and development is now working on developing 65 nm technology for release to manufacturing within the next 1–3 years. Of course, at some point, semiconductor scaling will approach its final limits[3] and silicon semiconductor manufacturing will become more of a commodity market.[4] The key determining factor in approaching the endpoint for scaling of complementary metal–oxide–semiconductor (CMOS) in semiconductor manufactuarable environments will ultimately be in the economics and not in the physics.[5,6] However, statistical physics and the physics of understanding how complex systems interact will become increasingly important especially as we enter the era of megascale integration and SOC IC production. This is especially true for variance and defect interactions that become magnified when investigating the electrical output characteristics as device geometry shrinks.[7]

Today, we are truly in a global market where many of our day-to-day applications and interactions require interfac-

ing with integrated circuits and the millions of transistors contained in these building blocks. To the average consumer, the dynamics of these ICs are transparent while the expectation for increasing functionality with unquestioned accuracy and quality drive the industry to new heights. This results in a tremendous driving force to continue to scale technologies no matter what roadblocks may be encountered. As semiconductor manufacturing migrates to more advanced deep-submicron technologies,[8] we are facing a new barrage of challenges to overcome for the next generation of SOC and IC products.[9,10] The bulk of these challenges arise due to their impact on product yield, product reliability, chip testability/performance prediction, and understanding how process integration and design layout interact. Some examples of key areas now affecting semiconductor manufacturing include negative bias temperature instability (NBTI), gate oxide leakage current, power consumption, testing of complex product, interaction of layout on chip yield, etc. Based on these issues, within the next 5 years the semiconductor industry is poised for a paradigm shift point in Moore's Law and technology scaling if yield and performance prediction of complex circuits such as SOC are not addressed. The pressure to maintain the roadmaps outlined above will continue to increase, but without addressing these secondary issues, multibillion dollar fabrication facilities may not be able to deliver competitive products. Second designers are now under extensive economic pressure to have their designs work the first time. Several factors are leading to this paradigm shift but the strongest ones include cost of generating mask sets, which can now approach a million dollars and higher, time-to-market pressure, and wafer cost. This is exacerbated by the fact that the cost of fabricating wafers and that IC processing typically requires wafers to be processed in wafer lots containing 25 wafers. State-of-the-art volume manufacturing can achieve 1.5 d per level of processing, which translates to best in the world full flow cycle times of 37.5 d.[11] If failures or yield issues are detected, this typically occurs after wafers have been processed, which means that time delays for implementing process fixes can impact many wafers in the line. Because of these multiplying effects and considering a typical high volume wafer fab with throughput capacity for 25 000 wafers/month,[12] mistakes or low yielding products can result in severe economic consequences to a manufacturer. The ability to address these issues will make the difference between having a highly profitable product versus an uncompetitive high loss product. As pointed out, multi-redesign of product or technology nodes are no longer available options. This pressure will be even worse as we go to 300 mm wafers.

In this article, we address the issue of NBTI with first a review of the present day understanding of the mechanisms governing NBTI, second, the process interactions affecting NBTI, and finally a review of some potential solutions to minimized the impact of NBTI induced from the technology processing.

II. INTRODUCTION

NBTI occurs in p-channel MOS devices stressed with negative gate voltages at elevated temperatures. It manifests

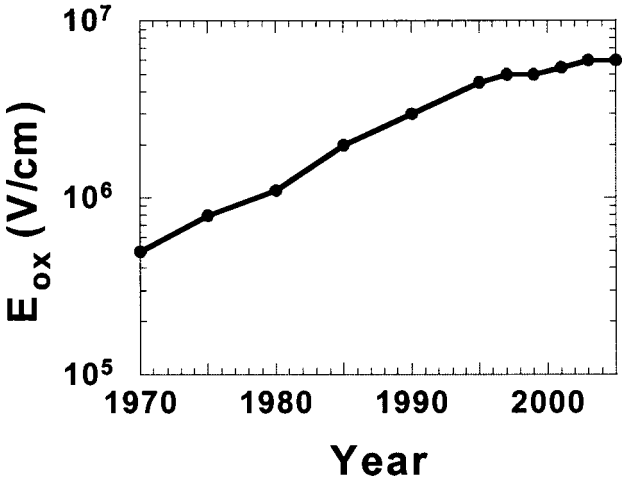

FIG. 1. Oxide electric field as a function of time for CMOS circuits.

itself as absolute drain current I_{Dsat}, and transconductance g_m, decrease and absolute ''off'' current I_{off}, and threshold voltage V_T increase. Typical stress temperatures lie in the 100–250 °C range with oxide electric fields typically below 6 MV/cm, i.e., fields below those that lead to hot carrier degradation. Such fields and temperatures are typically encountered during burn in, but are also approached in high-performance ICs during routine operation. Figure 1 shows the trend in electric fields for CMOS circuits, showing that oxide electric fields of the magnitude to generate NBTI are typical in today's circuits. Either negative gate voltages or elevated temperatures can produce NBTI, but a stronger and faster effect is produced by their combined action. It occurs primarily in p-channel MOSFETs with negative gate voltage bias and appears to be negligible for positive gate voltage and for either positive or negative gate voltages in n-channel MOSFETs.[13] In MOS circuits, it occurs most commonly during the ''high'' state of p-channel MOSFETs inverter operation. It also leads to timing shifts and potential circuit failure due to increased spreads in signal arrival in logic circuits. Asymmetric degradation in timing paths can lead to nonfunctionality of sensitive logic circuits and hence lead to product field failures.

A fraction of NBTI degradation can be recovered by annealing at high temperatures if the NBTI stress voltage is removed. The electric field applied during anneal can play a role in the recovery of NBTI degradation. Positive bias anneals exhibit the largest recovery in device characteristics.[14] Although it has not been reported yet, however, based on the historical results we expect this recovery to be unstable and for the original degradation to reappear soon after reapplication of the stress condition, assuming that hydrogen does not play a reversible role under this condition.

Since NBTI occurs for negative gate voltages, it is particularly detrimental for p-MOSFETs with either p^+ or n^+ poly-Si gates. However, recent data suggest that buried-channel (BC) p-MOSFETs are significantly less susceptible to NBTI.[15] The improved reliability in buried-channel compared to surface-channel devices is attributed to the naturally reduced oxide field for the same gate voltage due to the work

function difference of n^+ gates compared to p^+ gates and no boron diffusion from the n^+ gate. Also, the effective oxide is thicker in buried- than in surface-channel devices.[16] Thus their use can improve NBTI and $1/f$ noise, though suffering from worse short-channel effects and difficulty in manufacturing due to variance control issues. Currently the majority of digital CMOS technology requires surface channel devices, with this trend continuing into the future. Scaling of technology results in a significant increase in the susceptibility to NBTI degradation. Hence it may ultimately limit device lifetime, since NBTI is more severe than hot carrier stress for thin oxides at low electric fields.

Many advanced CMOS technologies now offer dual- or multigate oxide processes, with thicker gate oxides for analog-sensitive circuits, while logic performance optimization requires thin and medium thick gate oxides. Because these thicker gate oxides typically have lower nitrogen incorporation near the Si/SiO_2 interface, these devices tend to be less susceptible to NBTI, but it is still a concern for its impact on analog mixed signal circuit applications, because V_T shifts are a major reliability concern. This is especially a concern in matching applications where circuit operation may force matched transistors into asymmetrical bias conditions resulting in a significant asymmetric stress induced mismatch.[17] If the mismatch exceeds circuit tolerances, which can be as tight as 100 ppm, differences in operating characteristics between devices for some high performance design applications result in a failure or yield loss during burn-in or worse yet in field operation. Analog design techniques can circumvent some of these issues and minimize asymmetric NBTI bias stress conditions, but this comes at the expense of design complexity and potential performance tradeoffs such as power consumption, noise margins, or increased chip area.[18]

Because digital circuits tend to dominate most ICs and most SOCs and the number of digital transistors approach millions of devices on a chip, NBTI-induced shifts have become a critical issue for most manufacturers. NBTI stress-induced variances in digital device saturation drive current (I_{Dsat}), due to degradation in p-channel MOSFET I_{Dsat}, lead to significant timing issues. If digital signals arrive at different times, signal processing becomes corrupted and ultimately results in circuit failure. With V_T degradation, the gate overdrive ($V_G - V_T$) decreases, leading to reduced current and frequency degradation of ring oscillators and reduced standard random access memory noise margin.[19] Design modifications can alleviate some of these problems at the expense of circuit complexity or performance degradation. Of more ominous concern for future SOC designs, is the potential for statistical variation of NBTI tails of the failure time distribution that impact chip failures when millions of devices are present.[20] The reason for concern is the probability of a circuit encountering lethal accelerated NBTI degradation in single devices increase as both devices are scaled and the number of devices on chip increases. Hence the impact of NBTI on yield is expected to increase as SOC complexity and integration increase, even if design modifications are used.

NBTI has been known since the very early days of MOS device development, having been observed as early as 1967.[21] Deal named it Instability Number VI.[22] Goetzberger *et al.* at Bell Labs were one of the first groups to show detailed characterization of negative bias, temperature stress.[23] They used metal gate devices on 100 nm oxides, stressed at -10^6 V/cm at 300 °C and found an interface trap density D_{it} peak in the lower half of the band gap. The higher the starting D_{it}, the higher the final stress-induced D_{it}. For positive gate voltage, they noted a very small D_{it} increase. D_{it} increased with gate voltage and with time with a time dependence of $t^{0.25}$. D_{it} $(T=300\,^{\circ}\mathrm{C}) > D_{\mathrm{it}}$ $(T=250\,^{\circ}\mathrm{C})$ and p-type substrates gave higher D_{it} than n-type substrates. Early MOS devices containing oxides as the gate dielectric exhibited NBTI. Migration to nitrided oxides aggravated NBTI coinciding with a shift from the research stage to the forefront by around 1999 shortly after nitrided oxides became the industry standard in advanced CMOS.

The interface trap density induced by NBTI increases with decreasing oxide thickness, whereas the fixed oxide charge density induced by NBTI appears to have no thickness dependence. This t_{ox}^{-1} dependence of interface-trap generation implies that NBTI becomes more severe for ultrathin oxides. Furthermore, the NBTI-generated interface traps and fixed charges are likely to have an adverse effect on $1/f$ noise, which is believed to be closely related to these charges. NBTI has also been reported for HfO_2 high-k insulators.[24]

Examples of NBTI degradation are shown in Fig. 2. Figure 2(a) shows the change of threshold voltage and charge pumping current as a function of stress time.[25] Charge pumping current is used to measure the increase in interface trap density with degradation. Clearly, both V_T and I_{cp} change by similar amounts, plainly indicating that interface traps are created. Figure 2(b) shows similar behavior for threshold voltage and transconductance change.[26] Transconductance is related to mobility that is degraded during the stress. Although such plots vary from researcher to researcher, the general NBTI trends are embodied in these figures.

III. INTERFACE TRAPPED CHARGE AND FIXED OXIDE CHARGE

Before discussing the published experimental data and proposed models of NBTI stress, it is useful to review some basic MOSFET concepts such as threshold voltage, interface trapped charge and fixed oxide charge. The p-channel MOSFET threshold voltage is given by

$$V_T = V_{\mathrm{FB}} - 2\phi_F - |Q_B|/C_{\mathrm{ox}}, \tag{1}$$

where $\phi_F = (kT/q)\ln(N_D/n_i)$, $|Q_B| = (4qK_s\varepsilon_o\phi_F N_D)^{1/2}$ and C_{ox} is the oxide capacitance per unit area. The other symbols have their usual meaning. The flatband voltage is given by

$$V_{\mathrm{FB}} = \phi_{\mathrm{MS}} - \frac{Q_f}{C_{\mathrm{ox}}} - \frac{Q_{\mathrm{it}}(\phi_s)}{C_{\mathrm{ox}}}, \tag{2}$$

where Q_f is the fixed charge density and Q_{it} the interface trap density. We assume that neither substrate doping density N_D nor oxide thickness vary with NBTI stress. This may not

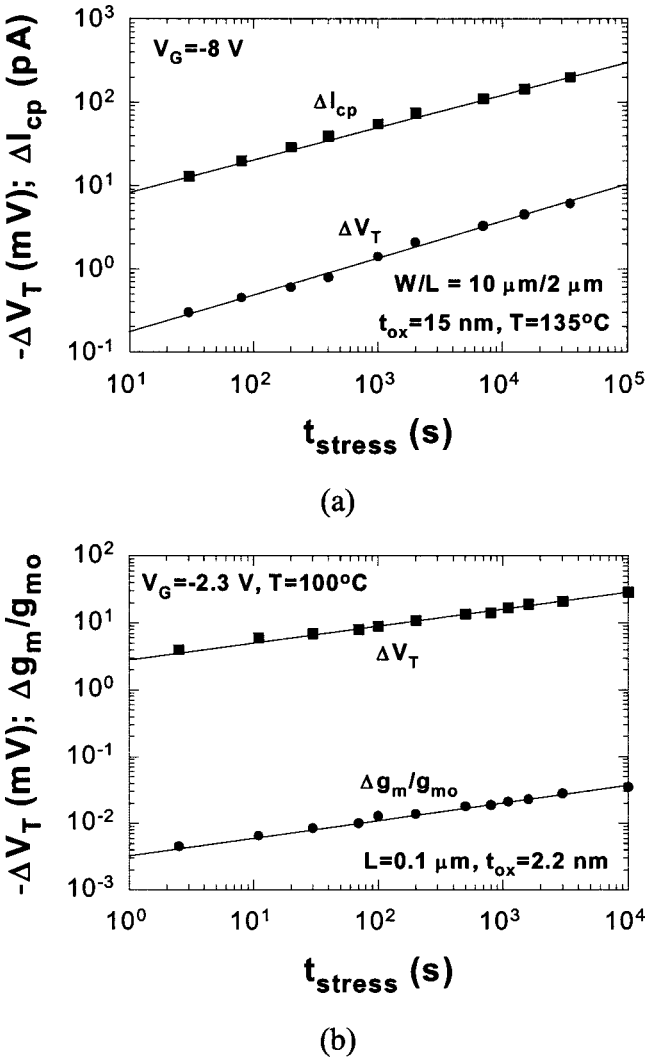

FIG. 2. (a) Time-dependent threshold voltage and stress-induced charge pumping current increase for *p*-MOSFETs. Data after Schlünder *et al.* Ref. 25; (b) time-dependent threshold voltage and stress-induced transconductance increase. Data after Kimizuka *et al.* Ref. 26.

be true if hydrogen in the silicon has deactivated some of the substrate doping atoms.[27] Such activation can, of course also occur in the poly-Si gates. Boron–hydrogen pair formation occurs typically in the $T=90-130\,^{\circ}\mathrm{C}$ temperature range.

The only parameters that can lead to threshold voltage shifts are the fixed charge density Q_f and the interface trapped charge density Q_{it}. Q_{it} in Eq. (2) depends on the surface potential ϕ_s, because the occupancy of the interface trapped charge is surface potential dependent. Positive increases in either of these charge densities, leads to negative threshold voltage shifts. Since NBTI typically leads to negative threshold voltage shifts, either or both of these charge densities are changed during NBTI stress.

The MOSFET saturation drain current and transconductance in its simplest form are given by

$$I_D=(W/2L)\mu_{\mathrm{eff}}C_{\mathrm{ox}}(V_G-V_T)^2;$$
$$g_m=(W/L)\mu_{\mathrm{eff}}C_{\mathrm{ox}}(V_G-V_T). \tag{3}$$

Two parameters leading to I_D and g_m degradation are threshold voltage and mobility μ_{eff} changes. Threshold voltage changes are discussed above. Mobility changes come about through interface trap generation, leading to additional surface-related scattering. For saturation velocity dominated devices, it is the velocity that is degraded.

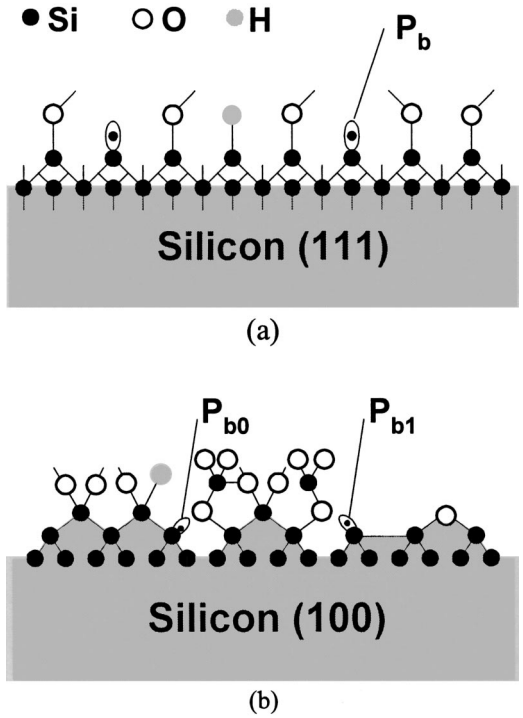

FIG. 3. Structural model of the: (a) (111) Si surface and (b) (100) Si surface.

A. Interface trapped charge

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 Figs. 3(a) and 3(b) with most Si atoms bonded to oxygen at the surface. 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. It is usually denoted by

$$\mathrm{Si_3}\equiv\mathrm{Si}\cdot. \tag{4}$$

The $\equiv$ represents three complete bonds to other Si atoms (the Si_3) and the $\cdot$ represents the fourth, unpaired electron in a dangling orbital (dangling bond). Interface traps are also known as P_b centers.[28] Interface traps are designated as D_{it} ($\mathrm{cm^{-2}\,eV^{-1}}$), Q_{it} ($\mathrm{C/cm^2}$), and N_{it} ($\mathrm{cm^{-2}}$).

On (111)-oriented wafers, the P_b center is a $\mathrm{Si_3}\equiv\mathrm{Si}\cdot$ 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. 3(a). On (100) 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. 3(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 $\mathrm{Si_2O}\equiv\mathrm{Si}\cdot$.[28] This identification was found to be incorrect, as the calculated energy levels for

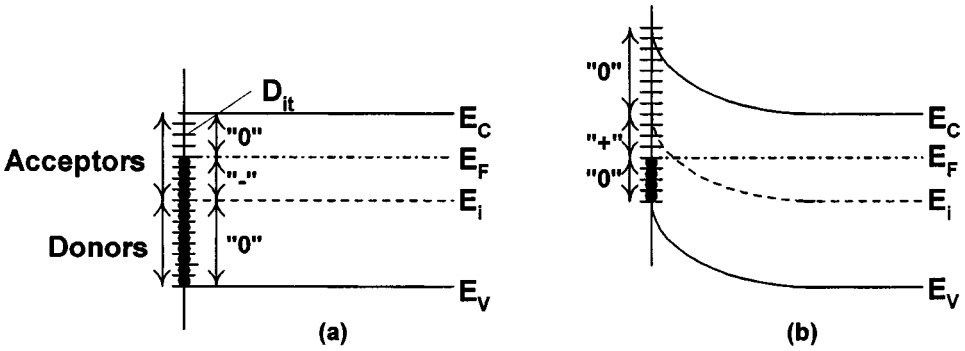

FIG. 4. Band diagrams of the Si substrate of a *p*-channel MOS device showing the occupancy of interface traps and the various charge polarities for a *p*-substrate with (a) negative interface trap charge at flatband and (b) positive interface trap charge at inversion. Each of the small horizontal lines 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.

this defect do not agree with experiment.[29] A recent calculation suggests the P_{b1} center to be an asymmetrically oxidized dimer, with no first neighbor oxygen atoms.[30] By 1999, it was unambiguously established that both P_{b0} and P_{b1} are chemically identical to the P_b center.[31] 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.[32] The two different effects are the result of strain relief in (100) silicon. The defects result from the naturally occurring mismatch induced stress in the SiO_2/Si layer during oxide growth.

P_{b0} centers result when strain relaxation occurs with a defect residing at (111) microfacets at the Si/SiO_2 interface, while P_{b1} centers result when strain relaxation occurs with a defect at (100) 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 (100) 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.[33] 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.[34] 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 also contribute to threshold voltage shifts, given by

$$\Delta V_T = -\frac{\Delta Q_{\rm it}(\phi_s)}{C_{\rm ox}}, \tag{5}$$

where ϕ_s is the surface potential. The surface potential dependence of the occupancy of interface traps is illustrated in Fig. 4.

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.[35] This is in contrast to doping atoms, which are donors in the upper half and acceptors in the lower half of the band gap. Hence, as shown in Fig. 4(a), at flatband, where electrons occupy states below the Fermi energy, the states in the lower half of the band gap are neutral (designated by "0"), being occupied donor states. Those between mid gap and the Fermi energy are negatively charged (designated by "−"), being occupied acceptor states and those above E_F are neutral (unoccupied acceptors). For an inverted *p*-channel MOSFET, shown in Fig. 4(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.

Interface traps, being *acceptors* in the upper half of the band gap and *donors* in the lower half, affect V_T shifts in *n*-channel and *p*-channel MOSFETs differently. Figure 5 shows an *n* channel in (a) and *p* channel in (b). At flatband, the *n* channel has *positive* and the *p* channel has *negative* interface trap charge. At inversion, $\phi_s = |2\phi_F|$, the *n* channel has *negative* and the *p* channel has *positive* interface trap charge. Since the fixed charge is positive, we have at inversion: *n* channel: $Q_f - Q_{\rm it}$, *p* channel: $Q_f + Q_{\rm it}$, hence *p* channel MOSFETs are more severely affected. This was clearly shown by Sinha and Smith where the threshold voltage of MOS capacitors on (111) *n*-Si decreases by 1.5 V while V_T of (111) *p*-Si decreases by only about 0.2 V.[36] Negative bias stress generates donor states in the lower half of the band gap.[19,37]

Both interface trap and fixed charge densities of state-of-the-art devices are in the range of 10^{10} cm^{-2} or lower. For a MOSFET with a 0.1 μm × 1.0 μm gate, i.e., $A = 10^{-9}$ cm^2, and $N_f = N_{\rm it} = 10^{10}$ cm^{-2}, there are only ten interface traps and ten fixed charges at the SiO_2/Si interface under the gate. Twenty charges lead to a threshold voltage shift of

$$\begin{aligned}\Delta V_T &= -\frac{Q_{\rm it}+Q_f}{C_{\rm ox}} \\ &= -\frac{20q}{K_{\rm ox}\varepsilon_o A}t_{\rm ox} \\ &= -\frac{1.6\times10^{-19}\times20}{3.45\times10^{-13}\times10^{-9}}t_{\rm ox} \\ &= -9.2\times10^{3}t_{\rm ox}.\end{aligned} \tag{6}$$

For $t_{\rm ox} = 5$ nm, this gives $\Delta V_T \approx -5$ mV. Device failure is sometimes defined as $\Delta V_T = -50$ mV, corresponding to $\Delta N_{\rm it} = \Delta N_f = 10^{11}$ cm^{-2}, showing that a modest increase in

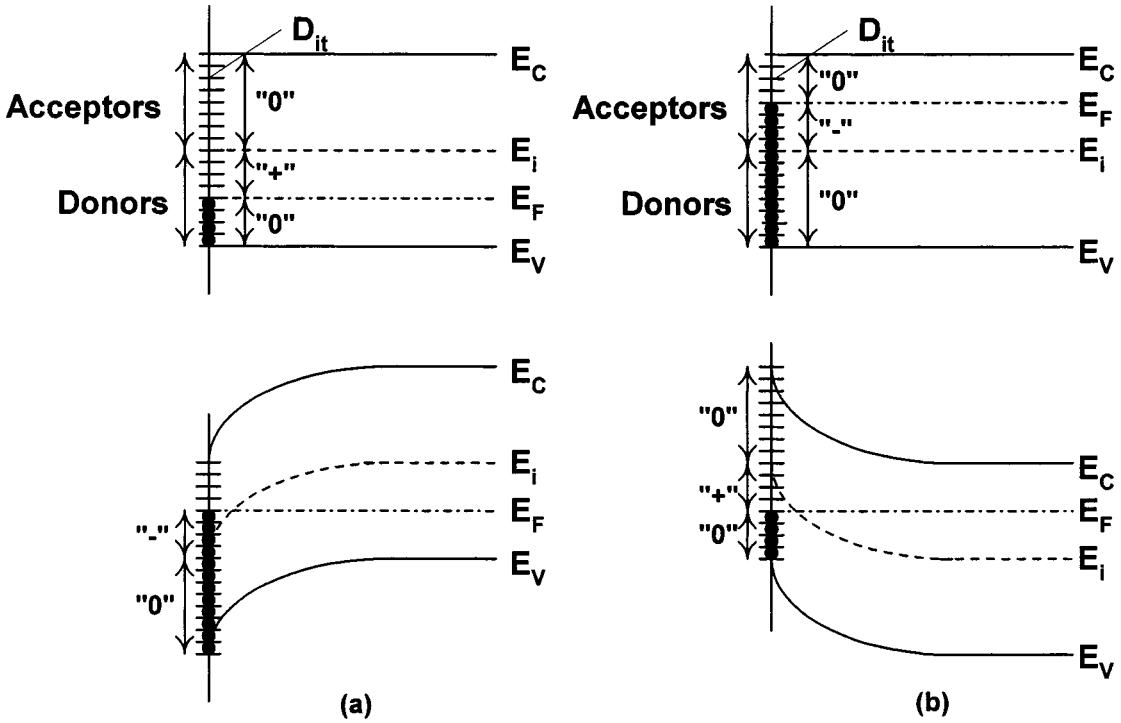

FIG. 5. Band diagrams of the Si substrate showing the occupancy of interface traps and the various charge polarities: (a) *p* substrate with positive interface trap charge at flatband and negative interface trap charge at inversion and (b) *n* substrate with negative interface trap charge at flatband and positive interface trap charge at inversion.

$N_{\rm it}$ and N_f suffices to cause failure. Since NBTI-generated $\Delta N_{\rm it}$ and ΔN_f are random in nature, suppose that in a matched analog circuit, one MOSFET experiences $\Delta V_T \approx -10$ mV and the other $\Delta V_T \approx -25$ mV. This 15 mV mismatch in a $V_T = -0.3$ V technology is a 5% mismatch. This is a very significant mismatch, especially if one considers high performance analog transistor pairs that can require mismatch tolerances of 0.1%–0.01%. Of course designers can trade off area to improve mismatch, but at the expense of adding capacitance and increasing chip area. Unfortunately there are no simple design solutions for some circuit building blocks where the existence of asymmetric bias conditions will induce NBTI asymmetric mismatch degradation. These issues are also of importance for sensitive digital circuits where such mismatch affects timing and edge triggering of high-performance digital circuits and adds to variances already present in the process. These undesirable NBTI effects can potentially affect yield seriously.

Rauch shows that NBTI-induced mismatch shifts are uncorrelated to the initial mismatch and that variances of the mismatch shifts add to the initial mismatch variations.[38] Threshold voltage shifts of only 20–40 mV increase the V_T sigma mismatch by $>10\%$. MOSFET β mismatch was found to be relatively unimportant at low gate overdrive, but became important at higher overdrives.

IV. NBTI TRAP GENERATION MODELS

In this section we discuss two basic approaches for modeling interface trap generation during NBTI processes. The first models discuss trap creation via hydrogen interaction dynamics. The second set of models describes more general trap creation via chemical species interaction and diffusion. The exact model describing the NBTI physics remains somewhat elusive at this time. However, these models are consistent with a number of observations in NBTI degradation and are expected to form the building blocks for a more complete understanding of NBTI physics.

A. Hydrogen models

A hydrogen terminated interface trap, shown in Fig. 3 is denoted by

$$\mathrm{Si_3 \equiv SiH}. \tag{7}$$

Since a definitive model of interface trap and fixed charge creation is yet to be developed, we describe the various models that have been put forth. High electric fields can dissociate the silicon–hydrogen bond, according to the model

$$\mathrm{Si_3 \equiv SiH \rightarrow Si_3 \equiv Si^{\bullet} + H^0} \tag{8}$$

where H^0 is a neutral interstitial hydrogen atom or atomic hydrogen. Recent first-principles calculations show that the positively charged hydrogen or proton H^+ is the only stable charge state of hydrogen at the interface and that H^+ reacts directly with the SiH to form an interface trap, according to the reaction[39]

$$\mathrm{Si_3 \equiv SiH + H^+ \rightarrow Si_3 \equiv Si^{\bullet} + H_2}. \tag{9}$$

This model uses the fact that the SiH complex (or passivated dangling bond chemical species) is polarized such that a more positive charge resides near the Si atom and a more negative charge resides near the hydrogen atom. Mobile

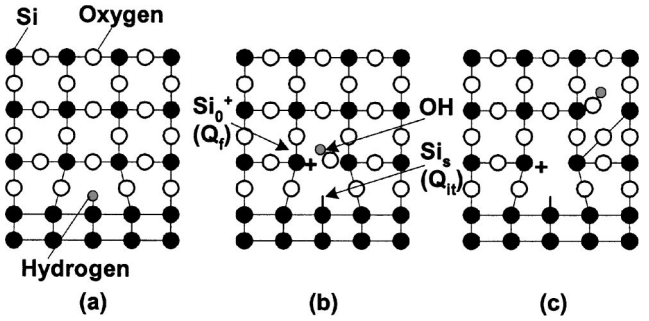

FIG. 6. Schematic two-dimensional representation of the $Si-SiO_2$ interface, showing (a) the $\equiv$SiH defect, (b) how this defect may be electrically activated during NBTI to form an interface trap, a fixed oxide charge, and a hydroxyl group, and (c) the OH diffuses through the oxide. After Jeppson and Svensson Ref. 45.

positive H^+ migrates towards the negatively charged dipole region of the SiH molecule. The H^+ atom then reacts with the H^- to form H_2 leaving behind a positively charged Si dangling bond (or trapping center). This is an appealing mechanism of the model since it is consistent with theoretical models that predict lower activation energies for hydrogen dissociation from SiH when charged states exist.[40] In this model, H_2 can later dissociate to again act as a catalyst to disrupt additional SiH bonds. This process, in theory, can continue so long as hydrogen is available and SiH bonds are available to react.

Since it is the proton that disrupts the SiH complex and forms the interface trap, interface trap formation depends on the oxide electric field that aids in the transport of the proton to the SiO_2/Si interface. Radiation experiments have shown that the electric field must be directed *toward* the Si substrate, for efficient interface trap formation to occur, if H^+ is created or released in the bulk SiO_2.[41] This is the typical condition during radiation experiments, where the electric field is directed from the gate to the substrate. Such radiation-induced degradation affects both *n*- and *p*-channel MOSFETs. This model is also consistent with the prediction that H^0 is unstable in both Si and SiO_2.[42] Although the model has successfully explained radiation experiments, where ionizing radiation creates H^+ in the oxide, it is inconsistent with NBTI experiments. The model is inconsistent with H^+ generated in the bulk oxide since a negative bias on the poly-Si gate causes the charged hydrogen to drift from the Si/SiO_2 interface at the channel. On the other hand, the model is consistent if one assumes H^+ can exist in the silicon below the Si/SiO_2 interface. At present this remains controversial since the literature claims H^- exists in *n*-type and H^+ exists in *p*-type Si. Hence if true, hydrogen would again move in the wrong direction for the applied fields, since *p*-MOSFETs are fabricated in *n* wells. H^- does not make intuitive sense since it would require two electrons circulating the proton. Hence H^+ may still exist in *n*-type Si, but it alters the material characteristics to appear similar to H^- existing in this material. It has been shown that SiH dissociation in *n*-type Si has lower activation energy than in *p*-type Si.[43] As the carrier density increases, the interaction of SiH dissociation mechanisms is expected to increase. This is consistent with *p*-MOSFETs showing more sensitivity to NBTI than *n*-MOSFETs.[44] It is also consistent with the activation energy for hydrogen dissociation in SiH bonds being lower in bulk Si than in SiO_2. Additional research will be required to confirm or dismiss this hypothesis. However, this is the most consistent NBTI model.

A different model that is sometimes used to explain NBTI induced trap formation considers the interaction of SiH with ''hot holes'' or holes near or at the Si/SiO_2 interface. Dissociation involving holes is given by

$$Si_3{\equiv}SiH + h^+ \rightarrow Si_3{\equiv}Si\cdot + H^+. \tag{10}$$

During NBTI stress, the proton is transported *from* the SiO_2/Si interface by the electric field directed from substrate to gate. Although this is still somewhat controversial, it is consistent with recent results indicating reverse substrate bias (V_{BS}) accelerates NBTI mechanisms.[25] This model is also consistent with hydrogen dissociation from SiH having reduced activation energy as the free carrier density increases.[43] It is not clear if it is consistent with dissociation dynamics in *n*-type Si (free carriers being electrons) versus channel inversion (free carriers being holes).

B. NBTI chemical reacting species models

We briefly describe some of the models that have been proposed. One model assumes that species Y diffuses to the interface and creates an interface trap

$$Si_3{\equiv}SiH + Y \rightarrow Si_3{\equiv}Si\cdot + X, \tag{11}$$

where Y is unknown. Jeppson and Svensson were the first to propose a model of NBTI.[45] They stressed MOS devices with Al gates and 95 nm thick oxides at -4 to -7×10^6 V/cm. These devices had been annealed in forming gas at 500 °C for 10 min. They find equal densities of Q_{it} and Q_f formation during NBTI stress. The D_{it} density created during NBTI stress decreases slowly if the device remains at the stress temperature with grounded gate. They also find a $t^{0.25}$ dependence and propose the model in Fig. 6. The H of the SiH bond reacts with the SiO_2 lattice to form an OH group bonded to an oxide atom, leaving a trivalent Si atom (Si_0^+) in the oxide and one trivalent Si_s at the Si surface. The Si_0^+ forms the fixed positive charge and the Si_s forms the interface trap. Their model forms the basis for models subsequently modified by others. They propose

$$\underset{Y}{Si_3{\equiv}SiH} + \underset{}{O_3{\equiv}SiOSi{\equiv}O_3} \rightarrow \underset{D_{it}}{Si_3{\equiv}Si\cdot} + \underset{Q_f}{O_3{\equiv}Si^+} + \underset{X \text{ to Si.}}{O_3{\equiv}SiOH} + e^- \tag{12}$$

For the interstitial hydrogen attack model X is H_2 and Y is H^0.

C. Fixed charge

Fixed charges are designated as Q_f (C/cm^2) and N_f (cm^{-2}).[46] The fixed charge is a charge near the SiO_2/Si interface, contributing mainly to threshold voltage shift. Q_f is also a byproduct trivalent Si defect in the oxide, denoted by

$$O_3 \equiv Si^+. \tag{13}$$

Similar to interface trap creation modeled as

$$Si_3 \equiv SiH + h^+ \rightarrow Si_3 \equiv Si\bullet + H^+ \tag{14}$$

Q_f generation can be modeled as

$$O_3 \equiv SiH + h^+ \rightarrow O_3 \equiv Si^+ + H^0 \tag{15}$$

Both interface traps and fixed charges are generated from the dissociation of SiH bonds. For Q_f this may take place at the interface or in the oxide close to the interface. Ogawa *et al.* from NTT determined fixed oxide charge densities from capacitance–voltage measurements and interface trap densities from conductance measurements of MOS capacitors.[47] Based on these measurements, they formulated expressions for N_{it} and N_f generation

$$\Delta N_{it}(E_{ox}, T, t, t_{ox}) = 9 \times 10^{-4} E_{ox}^{1.5} t^{0.25} \exp(-0.2/kT)/t_{ox} \tag{16}$$

$$\Delta N_f(E_{ox}, T, t) = 490 E_{ox}^{1.5} t^{0.14} \exp(-0.15/kT), \tag{17}$$

where t is the time. They find ΔN_f to be independent of oxide thickness, but ΔN_{it} to be inversely proportional to t_{ox}. This suggests that NBTI is worse for thinner oxides. This is not always observed, however, and is highly dependent on the process conditions. Their model is consistent with a diffusion model for N_{it} formation.

Until this point we have considered that NBTI shifts do not saturate as can be seen in Eqs. (16) and (17). More recent results indicate NBTI shifts tend to saturate over time, indicating a reaction-limited mechanism. Under these conditions, one would expect the total shift in device characteristics, such as $\Delta V_T = f(\Delta N_{it}, \Delta N_{ot})$, to be limited by the total amount of hydrogen available for breaking SiH bonds and the total number of potential trapping sites at the Si/SiO_2 interface and in the oxide. In this case ΔV_T is given by[48]

$$\Delta V_T(\Delta N_{it}, \Delta N_f) = B_1[1 - \exp(-t/\tau_1)] + B_2[1 - \exp(-t/\tau_2)], \tag{18}$$

where B_1 and B_2 can be related to Eqs. (16) and (17) and τ_1 and τ_2 are the reaction limiting time constants, related to the forward and reverse reaction rates for trap formation versus trap passivation.

Equation (18) is based on the electrochemical reactions[48]

$$Si_3 \equiv SiH + A + h^+ \leftrightarrow Si_3 \equiv Si\bullet + H^+, \tag{19}$$

$$O_3 \equiv SiH + A + h^+ \leftrightarrow O_3 \equiv Si\bullet + H^+, \tag{20}$$

where A is a neutral water-related species at the SiO_2/Si interface and h^+ is a hole at the silicon surface. During NBTI stress, positive hydrogen ions are released from hydrogen-terminated silicon bonds. Some of the hydrogen ions diffuse from the interface into the oxide bulk where some are trapped, causing a threshold voltage shift. During the early stress stage, reactions (19) and (20) favor the generation of interface states and positive hydrogen ions at the interface. The process is limited by the dissociation rate of hydrogen terminated silicon bonds. However, after some stress time, the transport of hydrogen ions from the interface into the oxide limits the process and the diffusion rate is controlled by a gradually decreasing electric field at the SiO_2/Si interface due to positive charge trapping in the oxide and increasing interface state density. As a result, further diffusion of hydrogen ions is reduced and ΔV_T is reduced and finally saturates.

Although mobile oxide charge usually results from ionic contamination due to sodium ions or other ions such as Li^+, K^+, Ca^{++}, and Mg^{++}, there is strong evidence that H^+ can also readily exist for prolonged periods in oxides as mobile charge that can be cycled in the gate oxide.[48,49] Although NBTI has been concerned primarily with unidirectional degradation shifts in V_T and I_{Dsat}, the interaction of mobile H^+ remains a concern that should be de-embedded from nonreversible stress induced effects. This is especially true in the case where mobile charge induced shifts in p-MOS devices are in the range of expected NBTI induced V_T shifts of 20–30 mV.[50]

Before leaving this topic, we would like to point out that there are some additional comments and concerns with regard to the NBTI degradation models as predicted by Eqs. (16) and (17). Although these equations remain as a fundamental building block for modeling and understanding NBTI in p-MOSFETs, deep submicron technologies may deviate somewhat from these predictions.[51] Deviations have been observed in the electric field dependence, the activation energies, and in the time evolution of interface and fixed trap creation. Depending on the gate oxide process and whether or not the oxide has received plasma-induced damage, the NBTI oxide electric field exponent m (E_{ox}^m) lies in the range of $1.5 \leqslant m \leqslant 3.0$. The activation energies (E_a) appear to be in the range $0.15\ eV \leqslant E_A \leqslant 0.325\ eV$ and may even be modeled with some stretched exponents and distributions in $E_A \pm \sigma E_A$, depending on what dominates the reaction kinetics and species involved, for example H_2O, H^+, SiH, SiN, SiO, SiF, etc.[52] Regarding the time dependence, more recent data suggest that interface and fixed charge are time evolving with similar dependences of $t^{0.25}$, although some deviations have been observed it appears that the time dependence factor lies in the range of 0.2–0.3.

Because of the variance in activation energies and mechanisms governing the reaction kinetics of NBTI degradation, it is critical to use statistical reliability models for predicting IC performance, yield, and reliability.[9] The importance of this increases directly as the level of complexity and the number of devices on a chip increase. Accuracy and understanding the deviation in the statistical distribution of fast shifting devices versus slow shifting devices can mean the difference between successful product introduction and economic disaster. Currently, the statistically small sample sizes

used in many investigations do not address these issues adequately.

V. EFFECTS OF VARIOUS PROCESS/DEVICE PARAMETERS ON NBTI

In this section we review process conditions that impact NBTI sensitivity in advanced CMOS processing. Wherever possible we provide references and interpret results and recommendations for improving NBTI performance.

A. Hydrogen

Hydrogen is a most common impurity in MOS integrated circuits oxides, being incorporated into the oxide during various phases of IC fabrication, e.g., nitride deposition and forming gas anneal. Hydrogen has been found at concentrations of 10^{19} cm^{-3} in dry oxides and concentrations of 10^{20} cm^{-3} in steam oxides.[53] The distribution is nonuniform with a substantial pileup near the SiO_2/Si interface.[54] Higher concentrations can exist depending on processing and anneal conditions. Hydrogen can exist in its atomic state H^0, as molecular hydrogen H_2, as positively charged hydrogen or proton H^+, as part of the hydroxyl group OH, as hydronium, H_3O^+, or as hydroxide ions OH^-. More recent literature suggests that H^0 is unstable in both silicon and SiO_2 and is not expected to exist.[42] As discussed earlier, hydrogen is believed to be the main passivating species for Si dangling bonds and plays a major role during NBTI stress, when SiH bonds are depassivated forming interface traps. Hydrogen and its interaction with nitrogen in nitrided gate oxides are also a concern. During NBTI stress, because of lower activation energy, mobile hydrogen ions are more likely to combine with nitrogen in Si–N bonds rather then with Si–O bonds and hence form positive fixed charge in nitrided oxides.[55]

B. Deuterium

Deuterium has been shown to reduce NBTI. Deuterium (D)—a ''heavy'' variant of hydrogen—is a stable isotope of hydrogen with a natural abundance of 0.015% containing a proton as well as a neutron in its nucleus. Due to its heavier mass or giant isotope effect, SiD bonds are more resistant than SiH bonds to hot carrier stress as well as NBTI. NBTI degradation with deuterium passivation generally exhibits a parallel shift compared to hydrogen passivation with a similar $t^{0.25}$ dependence. This is likely due to its slower diffusion coefficient. Deuterium can be introduced early or late into the device fabrication process. Its incorporation at the oxide/Si interface by thermal annealing involves three processes: (i) deuterium must diffuse through the insulating layer(s); (ii) it must passivate unpassivated interface traps; and (iii) it must replace the existing hydrogen in SiH bonds. Mechanism (iii) is believed to be the rate limiting process.[56]

In addition to the differences in the expected arrival rate of H and D, recent results indicate that deuterium desorption is not only governed by the isotope effect (a difference in vibrational excitation of SiH versus SiD), but also by the probability of D recombination enhancement due to the slower movement of D. This effect has been observed as a factor of 10 increase in the capture cross section of D compared to H.[57] Of course, this effect may be somewhat detrimental if deuterium is breaking the SiH bond since it is possible that deuterium may remain longer at a SiH bond and hence enhance the interface trap formation once near this molecular complex. However, it is expected that this effect is more significant for desorption then for absorption. Based on these results of slower deuterium diffusion and some more recent work,[58] it is also expected that D should help reduce NBTI even further at lower electric fields. This potential beneficial effect is expected to be more important for normal circuit operation compared to the high fields used for accelerated NBTI studies. Regardless of the final mechanisms governing D_2 physics, it is clear that D_2 can significantly reduce NBTI sensitivity. It is not clear yet, whether this improvement is enhanced at the low fields that are more typical of actual circuits.

Deuterium passivation can be significantly improved by depassivating SiH bonds *before* deuterium anneal or with deuterium-containing processes, e.g., D_2 instead of H_2 in forming gas, deuterated silane SiD_4 instead of SiH_4, and deuterated ammonia ND_3 instead of NH_3. By annealing the devices in a 10% deuterium ambient forming gas and providing the devices with a deuterated barrier nitride film, it is possible to retain the SiD bonds even if later stages of the process use hydrogen-containing forming gas.[59] The deuterated silicon nitride provides a deuterium reservoir and a barrier to subsequent in-diffusion of hydrogen and outdiffusion of deuterium. Clark *et al.* at IBM verified the efficacy of this process with hot carrier stress measurements.[60] Liu *et al.* from United Microelectronics/Infineon/IBM[61] and Kimizuka *et al.* from NEC/Bell Labs[26] find deuterium to reduce NBTI. However, there are some concerns when using early deposition of deuterium rich layers in process integration, especially if the layer is subjected to high energy implants at later steps in the process. Higher energy implants through deuterium rich layers risk activation and gamma ray production increases during processing.[60,62]

C. Nitrogen and nitrides

Nitrogen generally enhances NBTI. Nitrogen is commonly incorporated into gate oxides to reduce boron diffusion in *p*-channel MOSFETs, improve hot carrier resistance, and increase the dielectric constant. However, nitrogen tends to degrade NBTI. Chaparala *et al.*[63] from National Semiconductor used $L=0.4$ μm MOSFETs with $t_{ox}=6.8$ nm at $T=85$ and 150 °C. They measured hot carrier instability and NBTI and find that higher nitrogen concentration makes NBTI worse with an activation energy of $E_A=0.84$ eV. Kimizuka *et al.*[64] from NEC used 0.18 μm MOSFETs with $t_{ox}=2-4$ nm. They measured V_T, g_m, and I_G and find more severe NBTI for *p*-channel than for *n*-channel MOSFETs. Nitrided oxide made NBTI worse. Liu *et al.* found N_2O nitrided oxides, rapid thermal anneal nitric oxides, and remote plasma nitrided (RPN) oxides, where the gate oxide is exposed to a high density of remote nitrogen discharge, and all had lower NBTI degradation than SiO_2.[61] However, RPN oxides with higher nitrogen density exhibited worse NBTI

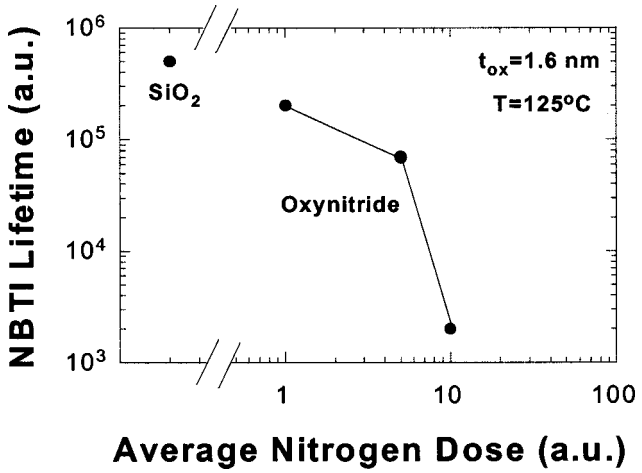

FIG. 7. NBTI lifetime as a function of nitrogen concentration in the oxynitride film. After Imai and Ono, Ref. 69.

than those oxides with lower nitrogen. Kimizuka *et al.* find nitrogen to degrade NBTI, with higher nitrogen content leading to more severe degradation.[64] Ichinose *et al.* used Si_3N_4 deposition to form sidewalls and find the NBTI lifetime to depend on the SiH concentration in the nitride film.[65] Lower SiH concentrations lead to longer NBTI lifetimes. Liu *et al.* from United Microelectronics compared NO-nitrided oxides, RPN, $N_2O+RPN+NO$, and reoxidized RPN.[66] They find the $N_2O+RPN+NO$ to give the best NBTI behavior in terms of interface roughness, effective mobility, gate current, NBTI V_T shift, and hot carrier degradation. They attribute this to low SiH bond densities due to the smooth SiO_2/Si interface as a result of the NO anneal. Ono *et al.* also attributed the better NBTI behavior to a smoother interface.[67] They achieved this through a hydrogen pretreatment of the wafer prior to oxidation, removing any chemical oxide. Recent work from Chartered Semiconductor provides additional support that decoupled plasma-nitrided oxides (DPNOs), where a pure thermally grown oxide is exposed to a high density decoupled nitrogen plasma source and has significant NBTI improvement compared to rapid thermal nitrided oxides (RTNOs).[68] Indeed, their results indicate the NBTI activation energy of DPNO oxides is 0.325 eV compared with 0.25 eV for RTNO oxides. This increase in activation energy translates directly to significantly improved NBTI lifetimes in DPNO oxides although the time dependent degradation remains proportional to $t^{0.25}$ for both DPNO and RTNO oxides, which indicates diffusion of a reaction species such as positively charged hydrogen is still a dominating mechanism for NBTI in these structures.

Imai and Ono used radical nitridation from electron cyclotron resonance to increase the nitrogen concentration near the gate/oxide interface in a 1.6 nm gate oxide, poly-SiGe gate technology.[69] As shown in Fig. 7, the nitrogen must be carefully controlled or it leads to excessive NBTI degradation. That paper clearly illustrates the tradeoff between optimizing for good NBTI lifetime and optimal device performance (gate leakage current, gate insulator boron penetration etc.).

D. Water

Water in the oxide enhances NBTI. Sasada *et al.*[70] from Sanyo stressed 11 nm oxide devices at $T=200\,^{\circ}C$, using charge pumping to determine D_{it}. By covering various regions of the device with nitride, a barrier to water diffusion, they found that water vapor is the dominant degrader. The activation energy of $E_A=1$ eV is consistent with water diffusion through oxide. Blat *et al.*[71] carried out a series of experiments growing 56 nm oxides on (111) oriented Si in "dry," "damp," and "wet" environments. Dry oxides are grown in dry oxygen. Damp oxides are formed by exposing dry oxides to a postmetal anneal at 450 °C, driving the thin water layer that forms on the oxide surface when the wafer is withdrawn from the oxidation furnace, into the oxide. Oxidizing in the presence of water vapor forms wet oxides. D_{it} and Q_f increases are observed in damp and wet oxides and the authors conclude that the diffusion species is water. Kimizuka *et al.* find wet H_2-O_2 grown oxide to exhibit worse NBTI than dry O_2 grown oxides.[26]

Helms and Poindexter in their review paper on the Si/SiO_2 system, propose water as the NBTI culprit.[72] Surveying the pre-1994 literature, they conclude that H_2O is the most likely—if unproven—depassivating reactant. H appears to be the less likely attacking reactant. The H_2O model is shown in Fig. 8. In Fig. 8(a), the prospective reactant H_2O is near the passivated P_b center and in (b) the electric field has oriented it into its attack position. In (c), the proton has been dislodged from the $\equiv$SiH site and combined with the H_2O molecule to yield H_3O^+. Finally, in (d), the now positively charged H_3O^+ is pulled away by the electric field preventing any reverse reaction.

Ushio *et al.*[73] from Hitachi did a first-principles molecular calculation to investigate hole-trapping reactions. They determined the reaction energies of hole trapping by subtracting the total molecular energy before hole trapping to hole reaction energy from that after hole trapping. The bonding configurations are shown in Fig. 9. Figure 9(a) shows the bonding configurations for an interface trap creation by H in SiO_2. Figure 9(b) shows the bonding configurations for an interface trap creation by H_2O in SiO_2 and Fig. 9(c) for an interface trap creation by H_2O in SiO_xN_y. Water has a lower reaction energy than hydrogen. H_2O reacts with O or N vacancies by inserting OH into the vacancy and generating a H atom to stabilize the hole-trapped state. The H atom combines with the H of a nearby Si–H bond, creating an interface trap. The final products are D_{it}, Q_f, and H_2. Their conclusions: water-originated reaction has lower energy at the Si/SiO_xN_y interface than at the Si/SiO_2 interface, i.e., NBTI is enhanced by water and by nitrogen incorporation in the oxide.

Water is often present on wafers from the contact and via formation. When the wafers are etched and cleaned, there is no problem getting water into the small holes mainly through capillary action, the problem is getting water out of the holes. An N_2 bake at 200 °C for >24 h is often used for good contact and via resistances. Water and moisture mostly travel along interfaces. This potentially makes water a key issue in pattern or layout dependent NBTI shifts. As pointed

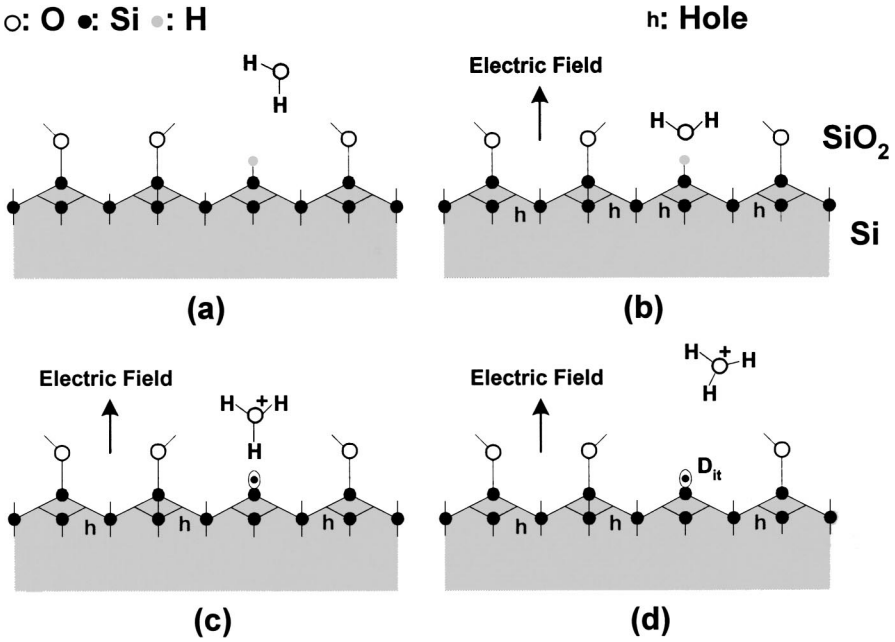

FIG. 8. Depassivation of P_b center by attack of H_2O molecule under electric field stress: (a) passivated, unstressed condition; (b) orienting of H_2O molecule and attraction of holes (h) to the interface by the electric field; (c) capture of a hole and H atom from the P_b center, forming H_3O^+; and (d) H_3O^+ removed from the reaction site by the electric field, leaving an active interface trap. After Helms and Poindexter, Ref. 72.

out several times throughout this article, this can lead to fast NBTI tails in the distribution of lifetimes if just a small amount of water is present in some pattern dependent locations within the die. Failure of a critical path device can cause yield loss during burn-in or field operation. Hence care must be taken to minimize the presence of water during processing.

E. Fluorine

Fluorine reduces NBTI. Fluorine is known to have many beneficial effects on MOS devices, improving hot carrier immunity, dielectric integrity, and NBTI. Hook *et al.* from IBM show that fluorine reduces NBTI and the reduction increases with increasing implanted fluorine dose, as shown in Fig. 10.[74] Liu *et al.* also observe an NBTI improvement.[61] However, these beneficial effects need to be weighed against detrimental effects such as enhanced boron diffusion in the gate oxide and higher junction leakage current in some devices.

F. Boron

Boron enhances NBTI. Boron diffuses into the gate oxide from the boron-doped gate and from the source/drain implants. NBTI should be gate-length dependent, if diffusion from the source/drain regions is important. A significant lifetime improvement is observed, as shown in Fig. 11, if boron is kept out of the gate oxide. Boron penetration has been observed to enhance fixed charge generation but to suppress interface trap generation. Reduced interface trap formation is attributed to the formation of Si–F bonds from the BF_2 boron implant. Enhanced Q_f has been attributed to increased oxide defects due to boron in the oxide.[75] As long as boron penetration into the gate oxide can be prevented, the use of BF_2 implants compared to standard B implants can gain almost a 1 order of magnitude improvement in NBTI lifetimes.[76]

G. Oxide damage

Oxide damage enhances NBTI. Damage in the oxide, particularly at the SiO_2/Si interface is detrimental for NBTI. For example, ion implantation creates defects at the interface that subsequently lead to more severe NBTI. Several articles report higher NBTI degradation for higher initial interface trap density. In fact, NBTI has been used to characterize plasma-induced damage and was shown to be a good predictor of oxide susceptibility to plasma charging.[77]

H. Gate material

Blat *et al.* use evaporated Al gates.[71] Gerardi *et al.*[78] do not use gates at all, but use negative corona charge and observe P_b formation on (111) Si. Most others use poly-Si gates. This implies that gate metal or poly-Si is not required and NBTI appears to be independent of gate material.

I. Gate precleans

Although there are few publications on the interaction of gate precleans with NBTI sensitivity, it is obvious that this step can potentially impact NBTI performance. Gate precleans are known to impact the gate oxide quality. Improper cleans can leave contaminants or damage nucleation sites that affect NBTI and $1/f$ noise. One of the more interesting reports for improving NBTI is to use hydrogen pretreatment of the silicon surface before gate oxide growth.[67] Unfortunately, the authors did not give details of their hydrogen pretreatment technique. However, it is well know that hydrogen binds to exposed silicon protecting it from oxidation and from contaminants such as water and carbon.[79] Hydrogen ambient at high temperatures results in breaking up of native oxides and removal of contaminants and is used extensively in SiGe epitaxial technologies.[80]

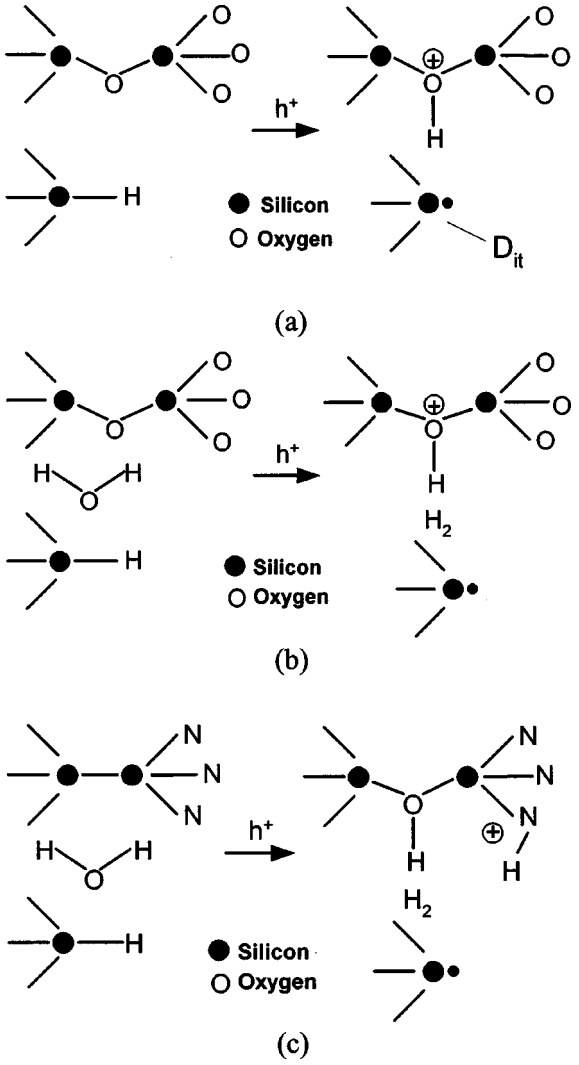

FIG. 9. (a) Interface trap formation in Si/SiO_2 with H (9.19 eV); (b) interface trap formation in Si/SiO_2 with H_2O (7.43 eV); and (c) interface trap formation in Si/SiO_xN_y with H_2O (5.82 eV). After Ushio *et al.*, in Ref. 73.

J. Silicon starting material

As we discussed earlier in the article, the type of silicon starting material orientation can have a direct impact on

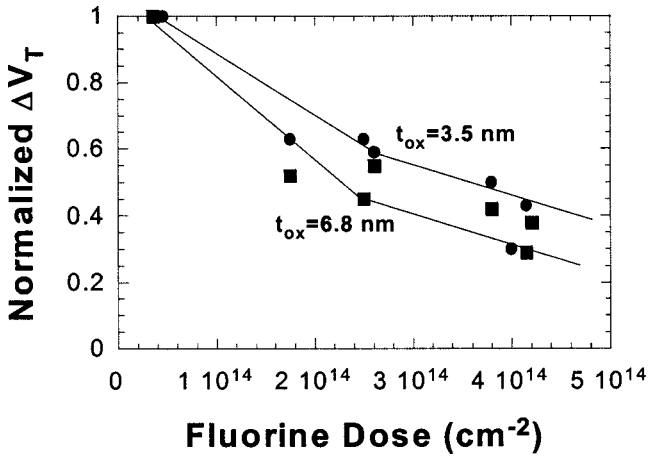

FIG. 10. Normalized NBTI V_T shift as a function of fluorine dose. After Hook *et al.*, Ref. 74.

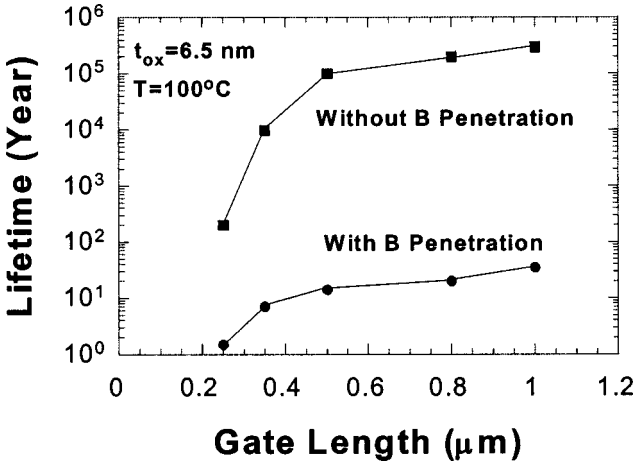

FIG. 11. Device lifetime vs gate length with and without boron penetration. The lifetime is defined as a $V_T = -20$ mV shift. The lifetimes were extrapolated from the temperature stress dependence. After Yamamoto *et al.*, Ref. 75.

NBTI sensitivity. Typically (100) silicon has been preferred over (111) silicon due to fewer P_b centers. Interestingly, however, recent results indicate improved device performance of *p*-MOSFETs fabricated on (110) silicon wafers. Momose *et al.* have observed a factor of 1.5–1.9 improvement in low field mobility and transconductance and similar improvements in drive current of these devices (though *n*-MOSFETs showed less improvements).[81] Unfortunately, *p*-MOSFET NBTI sensitivity was much worse in (110) silicon compared to (100). This increased sensitivity was found to correlate with significantly increased $1/f$ noise. Transmission electron microscope cross section indicated that surface roughness may have played a role in the NBTI sensitivity and hence it is speculated that it may be possible to overcome these issues with more careful gate oxidation methods. More research will be required in order to quantify if (110) silicon can maintain better *p*-MOSFET device characteristics such as I_{Dsat} while minimizing NBTI and $1/f$ noise sensitivity.

In addition to the starting material orientation, some processing steps can induce changes in the orientation of exposed surfaces, which may become more sensitive to NBTI. Local oxidation of silicon (LOCOS) oxides with too much stress can change (100) Si to (111) Si at the bird's beak edge, and shallow trench isolation (STI) can change (100) Si to (110) Si at the trench edge. Both effects influence gate oxide thickness, stress, and oxide charge. For curved surfaces at bird's beak edges in LOCOS, or trench edges in STI, there can be increased oxide charges. With shrinking device dimensions, these curved surfaces take up a higher percentage of the transistor active area. Furthermore the local strain in these regions can alter the activation energy/dynamics for NBTI induced trap creation. We expect this to lead to pattern dependent variation in circuit NBTI sensitivity and hence care must be taken for design and process interactions.

One intriguing method for getting around many of these issues is expected through the use of atomic oxidation methods. Saito *et al.*, have unambiguously demonstrated that atomic oxidation methods significantly improve interface

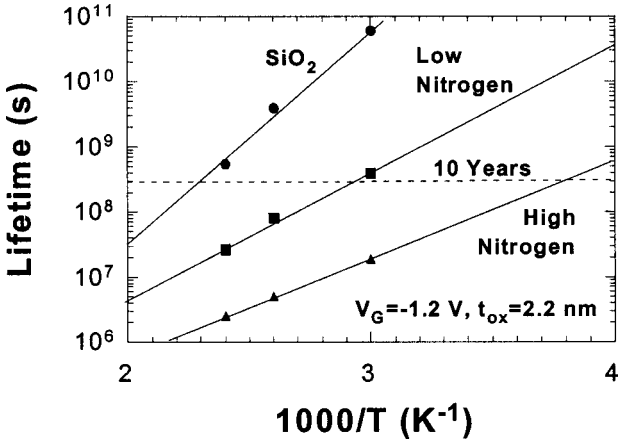

FIG. 12. Lifetime as a function of temperature. Nitridation done in NO. Data after Kimizuka *et al.*, Ref. 26.

trap densities on (111) silicon, yielding an order of magnitude reduction in N_{it} compared to conventional thermal oxide growth.[82] They also demonstrated highly uniform oxide thickness including regions of oxide thickness growth at STI edges where (100) silicon surfaces change to (110) and (111).

K. Holes

Why are holes required in NBTI? Blat *et al.* find no NBTI for *n*-channel MOSFETs and claim it is because there are no holes near the Si surface.[71] This suggests that holes are required, but there is no clear equation that shows what holes do in the thermo–chemical reaction. Figures 8 and 9 show the role of holes qualitatively. This effect becomes stronger for higher reverse body bias, which accelerates thermally generated holes towards the Si/SiO_2 interface,[25] thus increasing NBTI sensitivity. On the other hand, circuits may not see this accelerated condition since *n*-well supply voltages are limited and hence if V_{BS} is increased, the V_G potential to the channel is decreased in the circuit, reducing the oxide electric field. This inherent circuit bias constraint can potentially offset the accelerated V_{BS} NBTI mechanisms. Additional work will be required to quantify the exact effects with regard to realistic bias conditions encountered by *p*-MOSFETs, especially in dual- or multigate and multiple V_T devices.

L. Temperature

Higher temperature stress enhances NBTI. NBTI degradation is thermally activated [see Eqs. (16) and(17)] and, therefore, is sensitive to temperature. NBTI degrades more severely the higher the temperature, as illustrated in Fig. 12. As pointed out earlier, the activation energy of NBTI processes appears to be highly sensitive to the types of potential reacting species and to the type of oxidation methods used. This sensitivity in activation energy is highly sensitive to the temperature and variance in it will led to variance in the predicted NBTI lifetime of different processes. High performance microprocessors and SOCs can have hot spots in a circuit design leading to large temperature gradients across a chip. The net result is pattern dependent dispersion in the activation of NBTI processes and NBTI drift. If care is not taken to understand these issues, timing degradation dependent paths can lead to accelerated circuit failures during burn-in or field operations. Detection of these failures may become difficult due to circuit complexity and hence lead to erroneous data or output conditions.

Temperatures encountered during processing also play a role in the NBTI sensitivity. For example, it has been found that RTA processing and the ambient anneal environment can have a significant effect on NBTI sensitivity, where NBTI depends strongly on the processing temperature and annealing ambient.[49] Oxygen ambient was found to worsen NBTI, while annealing in argon led to a reduction in NBTI sensitivity. It is believed that the oxide growth temperature and annealing ambient can lead to changes in the oxide itself, including the mechanical stress of the oxide, the composition of the oxide (silicon rich versus oxygen rich, the water content in the oxide, etc) and the total number of trapping centers and potential trapping centers that are passivated with SiH or SiD (lower number of total traps leads to improved NBTI).

Process temperatures should be limited to temperature less then 1100 °C.[83] Although the effects of very high process temperatures have not been reported for nitrided gate oxides, it has been shown that temperatures at or above this value can lead to a significant generation of ''nonreactive'' mobile hydrogen trapped in the gate oxide.[83,84] These results are also consistent with results that indicate N_2O-grown oxides show significant reduction in charge-to-breakdown (Q_{BD}) when grown at higher temperature.[85] This is in contrast to pure oxides, which show increased Q_{DB} with increasing oxide growth temperature up to approximately 1050 °C.

Another area effecting NBTI appears to be postmetallization anneal temperature and the forming gas anneal temperature. Resent results indicate that postmetallization anneals temperatures and FG anneals should be minimized to temperatures at or below 370 C to improve NBTI and improve time-dependent dielectric breakdown (TDDB) in narrow devices.[86,87] Finally, we point out that other anneals such as the postsilicide anneal temperature, can impact the NBTI performance and care must be taken to optimize all key anneals and the ambient of the anneals in a process.[88]

It was shown by Roy *et al.* that it is beneficial to grow the gate oxide at a temperature above the SiO_2 viscoelastic temperature.[89] Growth stress incorporation in SiO_2 is the result of SiO_2 viscosity decrease at high growth temperatures and its subsequent increase during cooling. A graded structure aids in stress relaxation, where a pre-grown oxide layer provides grading for the subsequent high-temperature oxide layer to grow below the pregrown seed SiO_2 layer. This seed layer acts as a sink for stress relaxation. With this process design, high-quality SiO_2 interfacial layers, free from localized strain gradients, can be grown. They grew graded oxides at typically 940–1050 °C in a diluted oxidizing ambient (0.1% O) on a pregrown SiO_2 layer, grown at 750–800 °C. This pregrown oxide layer provides grading and stress relief during the cooling phase. Better than 3× reduction in NBTI was observed for the graded oxide when compared to con-

ventional oxides, attributed to the Si/SiO_2 interfacial substructure and a reduced number of weak silicon–oxygen bond within the oxide layer.

Liu *et al.* measured the hydrogen redistribution in steam-grown oxides by resonant nuclear reaction analysis.[90] They did this for 25 nm thick oxides grown at 850 °C and also for oxides postoxidation annealed (POA) at $T>1000$ °C. The measurements showed that the POA sample had significantly lower hydrogen density, lower D_{it}, and lower Q_f. After NBTI stress, the hydrogen density accumulated in an approximately 8 nm wide region near the SiO_2/Si interface, was significantly higher than the NBTI-induced D_{it}. POA clearly improved NBTI.

Bunyan *et al.* show that self heating in silicon on insulator (SOI) MOSFETs can aggravate NBTI-induced flatband voltage shifts and interface state density increases.[91]

M. Interconnects

Processing during backend-of-line (BEOL) metallization appears to have a significant influence on NBTI sensitivity. Data presented in joint work by Sony and Fujitsu[87] indicate copper metalization with typical dual damascene processes degrades NBTI. The increased sensitivity of NBTI to copper metalization was attributed to increased hydrogen present in both the copper metalization and especially in the barrier metal. Using TaN as the barrier metal compared to TiN/Ti improved NBTI. These results are in contrast to earlier results, which indicated similar NBTI sensitivity in antenna structure obtained from Al metalization and from Cu metalization.[92] The differences in these results may be due to differences in the hydrogen or water content present in BEOL processing between the two fabrication facilities.

Clearly another area of significant importance with regard to NBTI and BEOL processing is the intermetal dielectric (IMD).[87,93] Low-*k* IMD materials can introduce significant concentrations of water and hydrogen into the insulator. Materials of concern include spin on glass, SOD, plasma enhanced chemical vapor deposition dielectrics, etc. As discussed earlier, NBTI is strongly sensitive to H_2O that can penetrate the active *p*-MOSFET area in an IC. It has been found that these low-*k* IMD materials may impose significant NBTI risk if hydrogen or water is present along with catalytic anneals that appear to release and/or diffuse hydrogen/water to active regions of *p*-MOSFETs causing significant NBTI degradation. Nitride liners may minimize the impact of some of these issues but it is still possible for H/H_2O to penetrate vias to the active devices and hence care must be taken to minimize the IMD from high levels of H/H_2O.

The final area in BEOL processing to impact NBTI performance is device antenna ratio and the potential for plasma damage.[93] Although forming gas anneals have been shown to remove damage from unstressed devices, when stress conditions are applied, plasma damaged devices become significantly more sensitive to NBTI degradation with highly accelerated shifts in comparison to gate-protected structures or structures without antennas connected. Enhanced NBTI sensitivity in *p*-MOSFET antenna connected structures has been found for both positive and negative plasma potentials.[94] Other than minimizing plasma-induced damage (PID) in processing equipment and through special gate protection structures, a simple method for immediate improvement is through the use of conductive top films (CTFs) prior to using high-density plasma for dielectric depositions. This has been recently demonstrated in a joint effort by Motorola and AMD which clearly indicates that use of a thin undoped amorphous silicon CTF on top of the contact etch stop layer over silicon results in significantly reduced PID and significant improvement in NBTI lifetimes and improved electric field dependences.[95] The improvement in PID and NBTI performance was attributed not so much to the conducting ability of the CTF but more to the photon absorption of this layer from high energy photons generated in the plasma that are suspected of damaging the gate oxide. The sensitivity of this effect to antenna size remains somewhat unclear to us at this point, but is expected to be related to asymmetric charge generation, electron/hole recombination times, and differences in the barrier of the gate dielectric to electron and holes generated via the high energy PID photons.

N. Mechanical stress

Mechanical stress induced by either encapsulating films or by shallow trench proximity can have significant effects on the NBTI sensitivity of a device. Mechanical stress appears to interact with the amount of hydrogen and/or water present near the active regions of PMOS devices to cause increased NBTI sensitivity and even time-dependent dielectric breakdown degradation.[64,86,96] Based on these results it is important to minimize stress in IMD films and STI processing while also minimizing H/H_2O content for improving NBTI. Finally, although it has not been reported, we expect that IC packaging may become increasingly sensitive to NBTI degradation if the packaging material induces mechanical stress in encapsulated silicon chips, if stress is indeed a factor influencing NBTI. If true, this could become an additional concern for a SOC product that is destined for packaged part burn-in, especially if significant amounts of H/H_2O remain in the IMDs of the IC chip. However, we again point out this is speculation on our part and will require additional studies to verify if this is an issue.

O. Oxide electric field

Oxide electric field enhances NBTI. NBTI is very sensitive to electric field as shown in Fig. 13 in terms of the applied gate voltage for two different oxide thicknesses. According to Ogawa *et al.* the interface trap and fixed charge generation show an $E_{ox}^{1.5}$ dependence.[47] To determine the lifetime dependence on E_{ox}, we write the oxide voltage V_{ox} as

$$V_G = V_{FB} + \phi_s + V_{ox} \Rightarrow V_{ox} = V_G - V_{FB} - \phi_s . \tag{21}$$

For the p^+ poly-Si gate/*n*-substrate device, $V_{FB} \approx 0.8$ V and $\phi_s \approx 2\phi_F \approx -0.6$ V. The oxide electric field then becomes

$$E_{ox} = \frac{V_{ox}}{t_{ox}} = \frac{V_G - 0.2}{t_{ox}}, \tag{22}$$

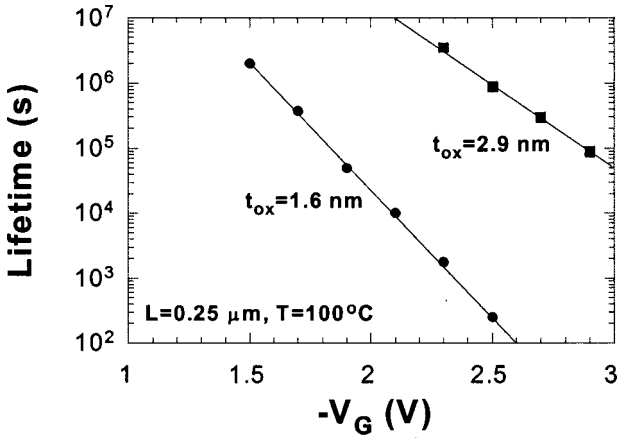

FIG. 13. Lifetime as a function of gate voltage. Data after Kimizuka *et al.*, Ref. 64.

where $V_G<0$. Replotting the data of Fig. 13 in terms of E_{ox}, gives the plot in Fig. 14. The linear data on a semilog plot suggest the lifetime expression

$$\tau = K \exp(-E_{ox}/E_1). \tag{23}$$

Fitting the data yields $K=2.3\times10^{11}$ s and $E_1=7\times10^5$ V/cm for $t_{ox}=1.6$ nm and $K=1.07\times10^{11}$ s and $E_1=5.7\times10^5$ V/cm for $t_{ox}=2.9$ nm.

P. Gate length

NBTI does not depend on lateral electric fields, unlike hot carrier degradation, and should therefore not exhibit any gate length dependence. Nevertheless, NBTI is sometimes enhanced with reduced gate length. It is not well understood why that is the case. It may have to do with the closeness of the source and drain and the dielectric spacers to the active channel. As shown in Fig. 15, localized damage near the ends of the channel and possible boron diffusion from source and drain into the gate oxide may play a role. Furthermore, lateral water diffusion into the gate oxide may be enhanced for shorter gate lengths. The threshold voltage shift dependence on gate length is shown in Fig. 16. For these data, the MOSFET was totally or partially covered with silicon nitride to study the effect of water penetration. Clearly, the "no SiN" case allows the most water penetration and exhibits the highest V_T shifts. On the other hand, if fluorine passivation of interface traps is important, then shorter gate lengths may lead to reduced NBTI with reduced gate length, since F finds it easier to diffuse to the SiO_2/Si interface. Although worse NBTI is most commonly associated with shorter channel devices, there are examples that show the opposite trend.[97]

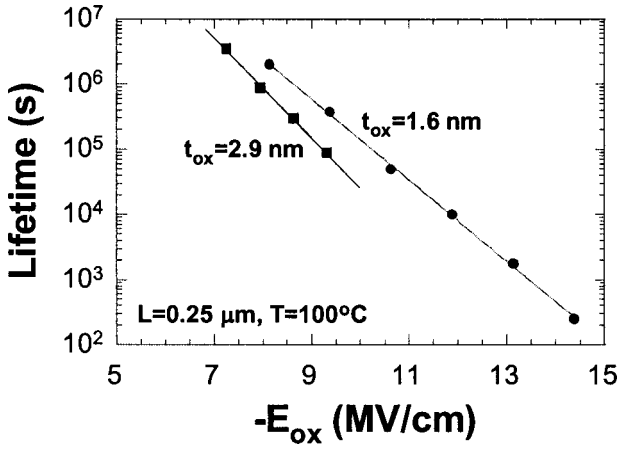

FIG. 14. Lifetime as a function of oxide electric field.

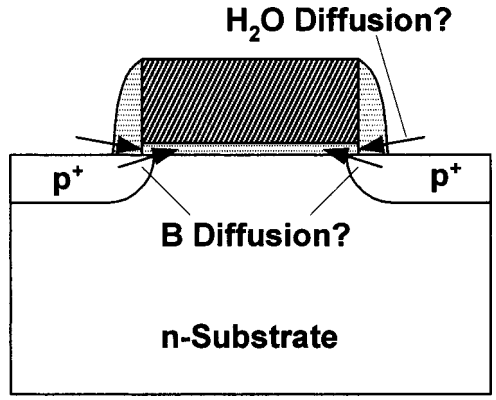

FIG. 15. MOSFET cross section showing possible "leakage" paths for boron and water penetration into the gate oxide.

VI. NBTI MINIMIZATION

To minimize NBTI, it appears to be necessary to have initially low densities of electrically active defects at the SiO_2/Si interface and keep water out of the oxide. During poly-Si deposition, the water on the wafer surface is driven off, leading to low water-containing oxides. The use of a silicon nitride encapsulation layer has been found effective in keeping water away from the active CMOS devices and hence improving NBTI performance.[70] However, it is critical to pattern the nitride film and optimize the geometry in order to assure that hydrogen passivation of dangling bonds can occur while keeping the distance from the active region large enough to ensure that water cannot diffuse to the gate region. Additional studies have found that it is also important to minimize stress[96] and hydrogen content,[65] in these liner nitrides covering the active PMOS devices. It appears to be

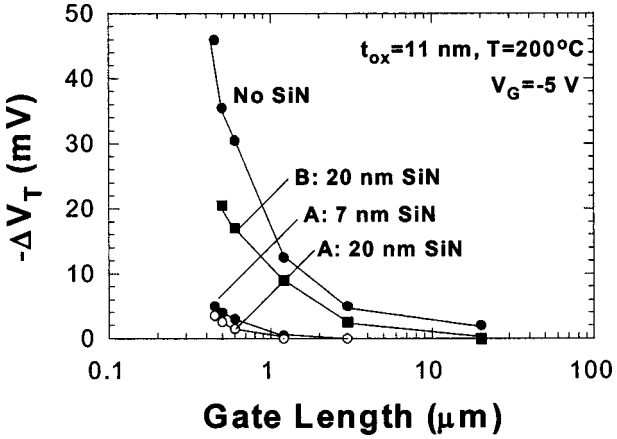

FIG. 16. Threshold voltage change as a function of gate length: (A) SiN all over the device and (B) SiN over gate only. Data from Sasada *et al.*, Ref. 70.

important to keep damage at the SiO_2/Si interface to a minimum during processing. The higher the initial interface trap density, the worse is the subsequent NBTI. For example, nitrogen introduction by ion implantation exhibits higher NBTI than nitrogen introduction by nonimplanted means.[61] Plasma charging damage also degrades NBTI as shown by Krishnan *et al.* from Texas Instruments.[92] This NBTI enhancement with antenna structure was independent of metalization, being observed in both dual damascene copper metalization as well as traditional plasma etched aluminum metalization. The reason for this similarity is not understood, but it is clear that antenna-charging effects are important to both copper and aluminum technologies. Pagaduan *et al.* from Xilinx also show that plasma damage degrades NBTI in *p*-MOSFETs, but *n*-MOSFETs are not degraded.[98] The threshold voltage degradation is consistent with their observation that the rise time of the CMOS output signal, controlled by the *p*-MOSFET, is degraded. The fall time, controlled by the *n*-MOSFET, is unchanged. It is believed that charging damage results in higher interface trap densities. Although these are passivated during subsequent low-temperature sintering, nevertheless, these higher initial Si–H densities, lead to higher NBTI. Hence, this suggests that initial damage should be minimized, even if that damage is annealed by postmetal annealing.

Deuterium is an effective way to improve both hot carrier stress and NBTI. However, it is not trivial to get the deuterium to the SiO_2/Si interface to form SiD bonds. If the MOSFET spacers consist of silicon nitride, deposited in the presence of hydrogen, most of the dangling bonds are already saturated with hydrogen and it is difficult to replace them with deuterium. Hence, the process needs to be altered to ensure deuterium can get to the SiO_2/Si interface and passivate dangling bonds or replace the hydrogen in existing SiH bonds with deuterium.

Nitrogen incorporation has given conflicting results. Some authors claim an NBTI improvement, while others observe degradation. Degradation is more commonly observed. Nitrogen concentration plays a key role in NBTI sensitivity especially if it is located near the Si/SiO_2 interface. Optimization of the nitrogen profile in the gate oxide can significantly improve NBTI sensitivity. Another key method to improve NBTI sensitivity is through the use of remote plasma nitridation of N_2O-grown oxides[61,66] and DPNO oxides. Nitrogen at the oxide/Si interface reduces the activation energy.[99] The higher the nitrogen concentration, the lower the activation energy and both Q_f and Q_{it} exhibit the same activation energy. The location of the nitrogen is also important. The closer the nitrogen is to the oxide/Si interface, the worse is NBTI.[100]

The method and chemistry of oxide growth appears to have significant effects on NBTI. It has been shown that the oxidation atmosphere has significant influence on NBTI, with wet oxides showing worse NBTI degradation then dry oxides. Fluorine, however, does lead to an improvement. It has also been clearly demonstrated that F implants or F concentration in the gate oxides can significantly improve NBTI and $1/f$ noise performance.[74] However, care must be taken when using blanket F implants since detrimental effects can be enhanced such as boron penetration or *n*-MOS performance degradation. Boron degrades NBTI.

As described earlier, the oxide electric field plays a significant role in NBTI sensitivity. Buried channel devices have been shown to reduce NBTI sensitivity, but they may not be suitable for most advanced CMOS technologies. We also expect NBTI sensitivity to be reduced by mid-gap work function gate materials, since the oxide electric field will be reduced due to the work function difference and due to the flatband voltage difference (lower doped channel region). Based on these concepts, it is expected that fully-depleted SOI should also offer improved NBTI immunity, because it uses lower-doped channel regions resulting in lower gate oxide electric fields, which should lead to improved NBTI.

VII. CONCLUSIONS

NBTI is potentially a significant reliability issue in *p*-channel MOSFETs. Although it is not a ''show stopper,'' because it is possible to design around it, it does place additional onus on design and process engineers to take this degradation mode into account. While, the microscopic details of NBTI are not completely understood with most of our present knowledge based on empirical results, NBTI physics and dynamics is gradually becoming clearer. From the literature, it is clear there are many effects that interact with NBTI sensitivity in a process. Most of these effects are second order and they modify the reaction dynamics or influence and shift the rate at which NBTI occurs within a given process and for a given device geometry. Probably the most important issue is the quality of the gate oxide and the spacer regions sounding the edge of the gate. The number of traps or weak spots in the oxide must be minimized to reduce NBTI sensitivity. Passivation of these traps with hydrogen is a partial solution since bonds can be broken and processes that yield very low trap densities (less then 10^{10} cm^{-2}) are critical.[101] Minimizing this high-quality oxide to effects such as antenna charging or other bond breaking process is likely the next most important issue. Passivation of remaining dangling bonds remains critical, but to minimize the sensitivity of these bonds to NBTI induced degradation should be the next level of focus.

Materials or process steps that reduce NBTI activation energy should be minimized. Typical processing conditions that affect the NBTI activation energy have been outlined in this article, but in general the worst culprits for reducing SiH and H interaction dynamics are nitrogen, water, and mechanical stress. Nitrogen, water and other chemical species that lower NBTI activation energies can be minimized in a process or one can tailor the profile in the oxides to lessen these materials at or near the Si/SiO_2 interface. Mechanical stress can be minimized by modifying the temperatures and growth conditions and strain in the oxide, STI or other dielectric layers. Alternatively one may be able to relieve the stress in oxide by methods such as ion implantation in the nitride liner layers or stress compensating layers.[102] Substituting deuterium for hydrogen is the next best approach since D_2 has high binding energies and diffuses slower than hydrogen.

Finally, to understand NBTI's impact on SOC circuits, it is critical to understand the distribution of NBTI tails inherent in advanced CMOS processes. Careful understanding of these issues along with close interaction with the design community on the impact of NBTI can lead to high yielding products. To cross the road of NBTI to deep submicron SOC manufacturing is the task of physicists, engineers, and the process development and manufacturing communities.

ACKNOWLEDGMENTS

This work was partially carried out at the National Science Foundation's State/Industry/University Cooperative Research Centers' (NSF-S/IUCRC) Center for Low Power Electronics (CLPE). CLPE is supported by the NSF (Grant No. EEC-9523338), the State of Arizona, and the following companies: Conexant, Gain Technology, Intel Corporation, Microchip Technology, Motorola, Inc., ON Semiconductor, Philips Semiconductors, Raytheon, Syncron Technologies, LLT, Texas Instruments and Western Design Center. The authors thank B. D. Choi from Arizona State University for assistance with some of the references, L. Anderson from Texas Instruments for illuminating discussions and insight, and C. Bulucea from National Semiconductor for reviewing the paper and motivating us to enumerate NBTI processes in a clearer manner.

[1] G. E. Moore, IEEE IEDM, 1 (1975).
[2] S. Thompson, P. Packan, and M. Bohr, Intel Technol. J. **Q3**, 1 (1998).
[3] E. J. Nowak, IBM J. Res. Dev. **46**, 169 (2002).
[4] T. H. Ning, IEEE CICC , 49 (2000).
[5] R. R. Schaller, IEEE Spectrum **34**, 53 (1997).
[6] G. D. Hutcheson and J. D. Hutcheson, Sci. Am. **274**, 54 (1996).
[7] G. Kamarinos and P. Felix, J. Phys. D **29**, 487 (1996).
[8] H. Iwai, IEEE J. Solid-State Circuits **34**, 357 (1999).
[9] K. Hess, A. Haggag, W. McMahon, K. Cheng, J. Lee, and J. Lyding, IEEE Circuits Devices Mag. **17**, 33 (2001).
[10] P. L. Levin and R. Ludwig, IEEE Spectrum **39**, 38 (2002).
[11] R. C. Leachman and D. A. Hodges, IEEE Trans. Semicond. Manuf. **9**, 158 (1996).
[12] R. Doering and Y. Nishi, Proc. IEEE **89**, 375 (2001).
[13] M. Makabe, T. Kubota, and T. Kitano, IEEE Int. Reliability Phys. Symp. **38**, 205 (2000).
[14] B. S. Doyle, B. J. Fishbein, and K. R. Mistry, IEEE IEDM , 529 (1991).
[15] Y. Nishida, H. Sayama, K. Oda, M. Katayama, Y. Inoue, H. Morimoto, and M. Inuishi, IEEE IEDM , 869 (2001).
[16] C. Bulucea and D. Kerr, Solid-State Electron. **41**, 1345 (1997).
[17] Y. Chen, J. Zhou, S. Tedja, F. Hui, and A. S. Oates, IRW Final Rep. No. 41, 2001.
[18] R. Thewes, R. Brederlow, C. Schlünder, P. Wieczorek, A. Hesener, B. Ankele, P. Klein, S. Kessel, and W. Weber, IEEE IEDM , 81 (1999).
[19] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, IEEE Int. Rel. Symp. **40**, 248 (2002).
[20] A. Haggag, W. McMahon, K. Hess, K. Cheng, J. Lee, and J. Lyding, IRW Final Rep. No. 179, 2000.
[21] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, J. Electrochem. Soc. **114**, 266 (1967).
[22] B. E. Deal, J. Electrochem. Soc. **121**, 198C (1974).
[23] A. Goetzberger, A. D. Lopez, and R. J. Strain, J. Electrochem. Soc. **120**, 90 (1973).
[24] K. Onishi, C. S. Kang, R. Choi, H. J. Cho, S. Gopalan, R. Nieh, E. Dharmarajan, and J. C. Lee, IEEE IEDM , 659 (2001).
[25] C. Schlünder *et al.*, Microelectron. Reliab. **39**, 821 (1999).
[26] N. Kimizuka, K. Yamaguchi, K. Imai, T. Iizuka, C. T. Liu, R. C. Keller, and T. Horiuchi, IEEE VLSI Symposium 2000, p. 92.
[27] C. T. Sah, J. Y. C. Chen, and J. J. T. Tzou, Appl. Phys. Lett. **43**, 204 (1983); S. T. Pantelides, *ibid.* **50**, 995 (1987).
[28] E. H. Poindexter and P. J. Caplan, Prog. Surf. Sci. **14**, 201–294 (1983).
[29] M. L. Reed and J. D. Plummer, J. Appl. Phys. **63**, 5776 (1988).
[30] A. Stirling, A. Pasquarello, J. C. Charlier, and R. Car, Phys. Rev. Lett. **85**, 2773 (2000).
[31] A. Stesmans and V. V. Afanas'ev, Microelectron. Eng. **48**, 113 (1999).
[32] A. Stesmans and V. V. Afanas'ev, Phys. Rev. **57**, 10030 (1998).
[33] P. Lenahan, IEEE 2002 Rel. Phys. Tutorial Notes, Adv. Rel. Topics, IEEE IRPS , section 223 (2002).
[34] T. D. Mishima and P. M. Lenahan, Appl. Phys. Lett. **76**, 3771 (2000); J. P. Campbell and P. M. Lenahan, *ibid.* **80**, 1945 (2002).
[35] P. V. Gray and D. M. Brown, Appl. Phys. Lett. **8**, 31 (1966).
[36] A. K. Sinha and T. E. Smith, J. Electrochem. Soc. **125**, 743 (1978).
[37] K. Saminadayar and J. C. Pfister, Solid-State Electron. **20**, 891 (1977).
[38] S. E. Rauch, III, IEEE Trans. Devices Mat. Rel. **2**, 89 (2002).
[39] S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, IEEE Trans. Nucl. Sci. **48**, 2086 (2001).
[40] A. Stesmans, Phys. Rev. B **61**, 8393 (2000).
[41] D. B. Brown and N. S. Saks, J. Appl. Phys. **70**, 3734 (1991).
[42] C. G. Van de Walle and B. R. Tuttle, IEEE Trans. Electron Devices **47**, 1779 (2000).
[43] B. Tuttle and C. G. Van de Walle, Phys. Rev. B **59**, 12884 (1999).
[44] H. Katto IRW Final Rep. No. 54, 2001.
[45] K. O. Jeppson and C. M. Svensson, J. Appl. Phys. **48**, 2004 (1977).
[46] B. E. Deal, IEEE Trans. Electron Devices **ED-27**, 606 (1980).
[47] S. Ogawa, M. Shimaya, and N. Shiono, J. Appl. Phys. **77**, 1137 (1995).
[48] C.-H. Liu *et al.*, Jpn. J. Appl. Phys., Part 1 **41**, 2423 (2002).
[49] D. Lu, G. A. Ruggles, and J. J. Wortman, Appl. Phys. Lett. **52**, 1344 (1988).
[50] W. W. Abadeer, IEEE Trans. Dev. Mat. Rel. **1**, 60 (2001).
[51] G. La Rosa, S. Rauch, and F. Guarin, IEEE 2002 Rel. Phys. Tutorial Notes, Rel. Fundamentals, , Section 133 (2002).
[52] M. S. Krishnan and V. Kol'dyaev, IEEE Int. Rel. Symp. **40**, 421 (2002).
[53] A. G. Revesz, J. Electrochem. Soc. **126**, 122 (1979).
[54] J. Krauser, F. Wulf, M. A. Briere, J. Steiger, and D. Bräunig, Microelectron. Eng. **22**, 65 (1993).
[55] E. Takeda, E. Murakami, K. Torii, Y. Okuyama, E. Ebe, K. Hinode, and S. Kimura, Microelectron. Reliab. **42**, 493 (2002).
[56] K. Cheng, K. Hess, and J. W. Lyding, IEEE Electron Device Lett. **22**, 441 (2001); K. Cheng, K. Hess, and J. W. Lyding, *ibid.* **22**, 203 (2001).
[57] M. M. Albert and N. H. Tolk, Phys. Rev. B **63**, 035308 (2001).
[58] C. H. Lin, M. H. Lee, and C. W. Liu, Appl. Phys. Lett. **78**, 637 (2001).
[59] T. G. Ference, J. S. Burnham, W. F. Clark, T. B. Hook, S. W. Mittl, K. M. Watson, and L. K. K. Han, IEEE Trans. Electron Devices **46**, 747 (1999).
[60] W. F. Clark, T. G. Ference, S. W. Mittl, J. S. Burnham, and E. D. Adams, IEEE Electron Device Lett. **20**, 501 (1999).
[61] C. H. Liu *et al.*, IEEE IEDM , 861 (2001).
[62] K. Saadatmand, E. McIntyre, S. Roberge, Z. Wan, K. Wenzel, R. Rathmell, and J. Dykstra, IEEE International Conference on Ion Implant. Technology Proceedings, edited by J. Matsuo, G. Takaoka, and I. Yamada, 1999, Vol. 1, p. 292.
[63] P. Chaparala *et al.*, IRW Final Rep. No. 95, 2000.
[64] N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai, and T. Horiuchi, IEEE VLSI Symposium, 1999, p. 73.
[65] K. Ichinose, T. Saito, Y. Tanagida, Y. Nonaka, K. Torri, H. Sato, N. Saito, S. Wada, K. Mori, and S. Mitani, IEEE VLSI Symposium, 2001, p. 103.
[66] C. H. Liu *et al.*, IEEE Int. Rel. Symp. **40**, 268 (2002).
[67] A. Ono, K. Fukasaku, T. Hirai, M. Makabe, S. Koyama, N. Ikezawa, K. Ando, T. Suzuki, K. Imai, and N. Nakamura, IEEE VLSI Symposium, 2001, p. 79.
[68] S.-S. Tan, T. Chen, C.-H. Ang, C.-M. Lek, W. Lin, J. Z. Zheng, A. See and L. Chan, IEEE International Symposium on Plasma/Process-Induced Damage, 2002, p. 146.
[69] K. Imai and A. Ono, IEICE Trans. Electron. **E85-C**, 1057 (2002).
[70] K. Sasada, M. Arimoto, H. Nagasawa, A. Nishida, H. Aoe, T. Dan, S. Fujiwara, Y. Matsushita, and K. Yodoshi, Microelectron. Test Struct. , 207 (1998).
[71] C. E. Blat, E. H. Nicollian, and E. H. Poindexter, J. Appl. Phys. **69**, 1712 (1991).
[72] C. R. Helms and E. H. Poindexter, Rep. Prog. Phys. **57**, 791 (1994).
[73] J. Ushio, K. Kushida-Abdelghafar, and T. Maruizumi, IEEE Semiconductor Device Res. Symposium 2001, p. 158; K. Kushida-Abdelghafar, K. Watanabe, J. Ushio, and E. Murakami, Appl. Phys. Lett. **81**, 4362 (2002).
[74] T. B. Hook, E. Adler, F. Guarin, J. Lukaitis, N. Rovedo, and K. Schruefer, IEEE Trans. Electron Devices **48**, 1346 (2001).

[75] T. Yamamoto, K. Uwasawa, and T. Mogami, IEEE Trans. Electron Devices **46**, 921 (1999).
[76] D.-Y. Lee, H.-C. Lin, W.-J. Chaing, W.-T. Lu, G.-W. Huang, T.-Y. Huang, and T. Wang, IEEE International Symposium on Plasma/Process-Induced Damage, 2002, p. 150.
[77] S. Fang and J. P. McVittie, IEEE Electron Device Lett. **13**, 288 (1992).
[78] G. J. Gerardi, E. H. Poindexter, P. J. Caplan, M. Harmatz, and W. R. Buchwald, J. Electrochem. Soc. **136**, 2609 (1989).
[79] B. S. Meyerson, F. J. Himpsel, and K. J. Uram, Appl. Phys. Lett. **57**, 1034 (1990).
[80] T. O. Sedgwick and P. D. Agnello, J. Vac. Sci. Technol. A **10**, 1913 (1992).
[81] H. S. Momose, T. Ohguro, K. Kojima, S. Nakamura, and Y. Toyoshima, IEEE VLSI Symposium, 2002, p. 156.
[82] Y. Saito, K. Sekine, N. Ueda, M. Hirayama, S. Sugawa, and T. Ohmi, IEEE VLSI Symposium 2000, p. 176.
[83] J. F. Zhang, C. Z. Zhao, G. Groeseneken, R. Degreave, J. N. Ellis, and C. D. Beech, J. Appl. Phys. **90**, 1911 (2001).
[84] K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur, and Z. J. Lemnios, Nature (London) **386**, 587 (1997).
[85] G. W. Yoon, A. B. Joshi, J. Kim, G. Q. Lo, and D.-L. Kwong, IEEE Electron Device Lett. **13**, 606 (1992).
[86] E. Morifuji *et al.*, IEEE VLSI Symposium, 2002, p. 218.
[87] A. Suzuki *et al.*, IEEE VLSI Symposium, 2002, p. 216.
[88] M. Mehrotra *et al.*, IEEE VLSI Symposium, 2002, p. 124.
[89] P. K. Roy, Y. Chen, and S. Chetlur, IEEE Trans. Electron Devices **48**, 2016 (2001).
[90] Z. Liu, S. Fujieda, K. Terashima, M. Wilde, and K. Fukutani, Appl. Phys. Lett. **81**, 2397 (2002).
[91] R. J. T. Bunyan, M. J. Uren, J. C. Alderman, and W. Eccleston, IEEE SOI Conference 1992, p. 130.
[92] A. T. Krishnan, V. Reddy, and S. Krishnan, IEEE IEDM , 865 (2001).
[93] N. Matsunaga, H. Yoshinari, and H. Shibata IEEE International Symposium on Plasma/Process-Induced Damage, 2002, p. 142.
[94] D.-Y. Lee, H.-C. Lin, M.-F. Wang, M.-Y. Tsai, T.-Y. Huang, and T. Wang, Jpn. J. Appl. Phys., Part 1 **41**, 2419 (2002).
[95] S.-C. Song *et al.*, IEEE VLSI Symposium 2002, p. 72.
[96] E. Morifuji *et al.*, IEEE VLSI Symposium, 2001, p. 117.
[97] Y. H. Lee, K. Wu, T. Linton, N. Mielke, S. Hu, and B. Wallace, IEEE Int. Rel. Symp. **38**, 77 (2000).
[98] F. E. Pagaduan, J. K. J. Lee, V. Vedagarbha, K. Lui, M. J. Hart, D. Gitlin, T. Takaso, S. Kamiyama, and K. Nakayama, IEEE Int. Rel. Symp. **39**, 315 (2001).
[99] S. S. Tan, T. P. Chen, C. H. Ang, and L. Chan, Appl. Phys. Lett. **82**, 269 (2003).
[100] S. S. Tan, T. P. Chen, C. H. Ang, Y. L. Tan, and L. Chan, Jpn. J. Appl. Phys. **41**, L1031 (2002).
[101] T. Ohmi, S. Sugawa, K. Kotani, K. Hirayama, and A. Morimoto, Proc. IEEE **89**, 394 (2001) (see Fig. 20 in the article).
[102] A. Shimizu, K. Hachimine, N. Ohki, H. Ohta, M. Koguchi, Y. Nonaka, H. Sato, and F. Ootsuka, IEEE IEDM , 433 (2001).