Intrinsic limitations on ultimate device performance and reliability at (i) semiconductor–dielectric interfaces and (ii) internal interfaces in stacked dielectrics

G. Lucovsky,[a] H. Yang, and H. Niimi
Departments of Physics and Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27695-8202

M. F. Thorpe
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

J. C. Phillips
Lucent Bell Laboratories, Murray Hill, New Jersey 07974

(Received 17 January 2000; accepted 30 May 2000)

The scaling of electrical oxide thickness to 1.0 nm and below for advanced silicon devices requires a change from thermally grown oxides and nitrided oxides to deposited dielectrics which have dielectric constants, k, significantly greater than that of silicon dioxide, $k_0 \sim 3.8$. Implementation of the higher-k dielectrics into field effect transistor devices requires a processing protocol that provides separate and independent control over the properties of the Si–dielectric interface and the bulk dielectric film. Experiments to date have shown that plasma-grown nitrided oxides, $\sim 0.5-0.6$ nm thick, satisfy this requirement. This paper addresses chemical bonding issues at the Si–dielectric interface and at the internal dielectric interface between the plasma-grown nitrided oxides and the high-k alternative dielectrics by applying constraint theory. $Si-SiO_2$ is a prototypical interface between a ''rigid'' Si substrate and a ''floppy'' network dielectric, SiO_2, and the interfacial properties are modified by a monolayer-scale transition region with excess suboxide bonding over what is required for an ideal interface. Additionally, the defect properties at the internal interface between a nitrided SiO_2 interface layer and a bulk dielectric film reflect differences in the average number of bonds/atom, N_{av}, of the dielectrics on either side of that interface. Experimentally determined interfacial defect concentrations are shown to scale quadratically with increasing differences in N_{av} thereby establishing a fundamental basis for limitations on device performance and reliability. © *2000 American Vacuum Society.* [S0734-211X(00)08204-4]

I. INTRODUCTION

The aggressive scaling of Si complementary metal–oxide–semiconductor (CMOS) devices to gate lengths <100 nm requires an increase in gate dielectric capacitance corresponding to an electrical oxide thickness (EOT) referenced to SiO_2, of <1 nm, and eventually extending down to 0.5–0.6 nm. The direct tunneling current density, J, through SiO_2 films with a physical thickness, t_0, of ~ 1 nm, and at a 1 V potential drop across the oxide exceeds 100 A/cm^2. However, to maintain field effect transistor (FET) performance and reliability, the tunneling current density at an operating bias of approximately 1 V over threshold must not exceed about 1 A/cm^2. This means that physically thicker gate insulators with dielectric constants, k, greater than that of SiO_2, $k_0 = 3.8$, must be introduced into the gate stack. This integration of *alternative* deposited gate dielectrics is based on the premise that increased physical thickness, t, of these high-k insulators,

$$t = t_0(k/k_0), \tag{1}$$

which provides the same capacitance as physically thinner SiO_2 dielectrics, will yield a tunneling current density significantly less than 1 A/cm^2. This in turn requires that conduction band offset energies and electron tunneling masses in the high-k dielectric be about the same as they are at $Si-SiO_2$ interfaces and in the SiO_2 dielectric film, respectively.

Introduction of gate dielectrics other than SiO_2 requires that these be prepared by deposition, rather than conventional thermal or rapid thermal growth processes. Since the $Si-SiO_2$ interface is continuously regenerated in a thermal growth process, the electrical properties of that interface are determined by the Si surface preparation/cleaning, the growth conditions, and subsequent thermal annealing/processing after completion of the growth phase. The electrical properties of thermally grown interfaces have been optimized for both performance and reliability, and have set the standards for performance and reliability required from alternative gate dielectrics.

Previous studies have demonstrated that the transition from thermally grown SiO_2 to deposited SiO_2 necessitates separate processing steps for interface formation and dielectric deposition.[1–9] As an example, it has been shown that direct deposition of SiO_2 by 300 °C remote plasma enhanced chemical vapor deposition (RPECVD) onto Si(100) surfaces that have been subjected to an *in situ* hydrogen plasma clean and/or an HF last rinse, resulted in highly defective

[a]Electronic mail: gerry_lucovsky@ncsu.edu

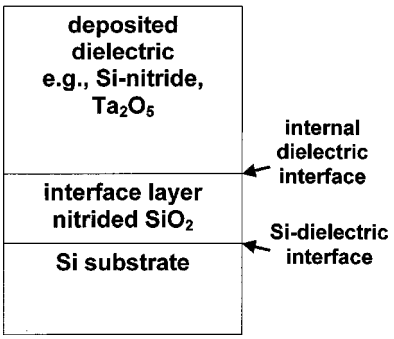

FIG. 1. Schematic representation of a stacked gate dielectric, indicating the Si–dielectric interface and the internal dielectric interface as well.

interfaces.[3] It was subsequently shown that a 300 °C remote plasma-assisted oxidation (RPAO) prior to REPCVD could be used to *in situ* clean and passivate the Si surface by growth of an ultrathin *passivating* oxide layer ~0.5–0.6 nm thick.[3] This approach has been demonstrated to give equally good results when combined with SiO_2 layers deposited by RPECVD at 300 °C, or by rapid thermal chemical vapor deposition (RTCVD), at higher temperatures, 600–750 °C.[3,7] The predeposition oxidation passivated the Si–dielectric interface against chemical attack in the form of a subcutaneous oxidation of the Si buried interface during deposition of the SiO_2 film.[1,2]

Other studies, in which deposited (i) hydrogenated Si nitride, (ii) Si oxynitride pseudobinary alloys, $(SiO_2)_{1-x}(Si_3N_4)_x$, and (iii) high-k dielectrics such as the Zr silicates, $(SiO_2)_{1-x}(ZrO_2)_x$ and Ta_2O_5, were substituted for RPECVD and RTCVD SiO_2, also indicated the necessity for using a predeposition, plasma-assisted oxidation step.[10–12] As shown in Refs. 3 and 7, the RPAO process of Refs. 10 and 11 formed the Si dielectric interface and a thin oxide layer that also prevented chemical attack of the substrate during film deposition. Stacked gate dielectrics formed in this way include two interfaces, the Si–dielectric interface and the internal dielectric interface, in close proximity to the Si substrate (see Fig. 1). Defects at these two interfaces play a significant role in determining device performance and reliability. This paper discusses defects at these interfaces from the perspective of constraint theory,[5] and then identifies fundamental aspects of the chemical bonding that play a role in determining the ultimate limits of device performance and reliability.

II. EXPERIMENTAL RESULTS

The devices described in this paper with hydrogenated Si nitride and Si oxynitride alloys include the remote plasma processing steps described in Refs. 9, 13–17. The devices incorporating Zr silicates and Ta_2O_5 were prepared by similar plasma-processing steps as described in Refs. 10 and 11. Interfacial oxide and nitrided oxide layers have been used in all of the devices addressed in this study. Since the focus of this paper is on the Si–dielectric interface formed in this process as well as the internal dielectric interface that is an

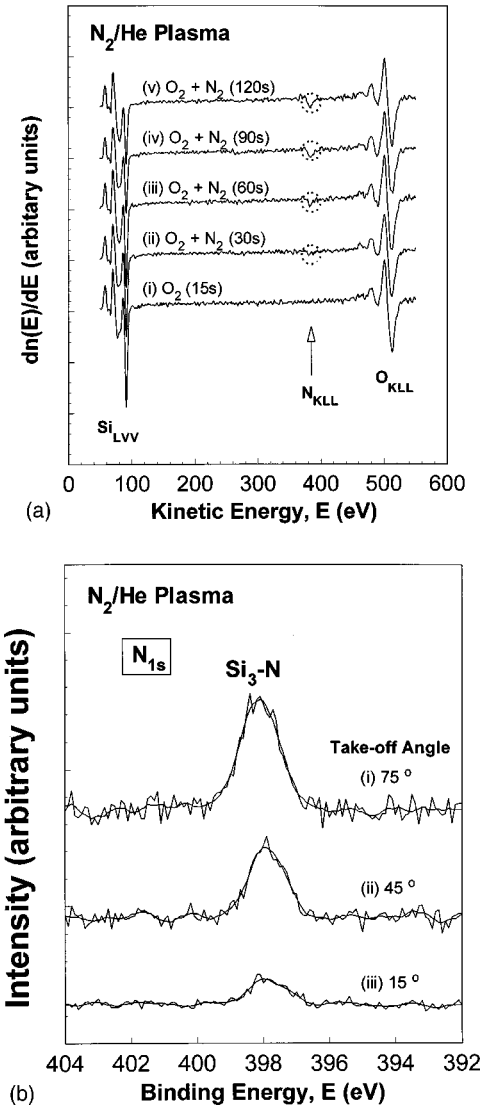

FIG. 2. (a) Differential Auger electron spectroscopy spectra as a function of electron energy for interface nitridation. Traces are included for a 15 s RPAO process, and for nitridation times ranging from 30 to 120 s. (b) Angle-resolved monochromatic N_{1s} XPS for nitrided interfaces (90 s exposure). The decrease of the nitrogen XPS feature with decreasing take off angle (with respect to the substrate) is consistent with nitrogen atoms being localized at, or in the immediate vicinity of, the Si–dielectric interface.

integral component of this interfacial oxide, the processing and characterization of these interfacial layers will be addressed in some detail. The thickness of the plasma-processed interfacial layers has been determined by on-line Auger electron spectroscopy as described in Refs. 3, 13, and 14. The processing conditions and characterizations are described in detail in Refs. 13 and 14 as well as other references cited in that paper. Figures 2(a) and 2(b) present on-line Auger electron spectroscopy (AES), and off-line x-ray photoelectron spectroscopy (XPS) data relevant to the interfacial nitridation. The combination on-line AES of Fig. 2(a), the XPS of Fig. 2(b), with the secondary ion mass spectrometry results in Ref. 14, optical second-harmonic generation (SHG) results in Ref. 17, and nuclear reaction analysis results of Refs. 9 and 14, establish that the nitrogen atoms are

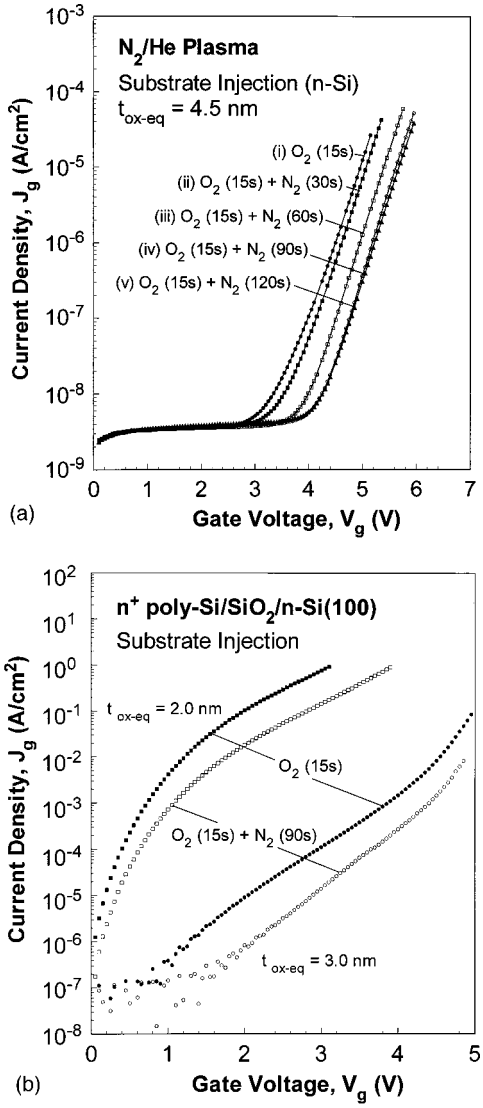

FIG. 3. Current density as function of gate voltage in (a) the Fowler–Nordheim, and (b) the direct tunneling regimes for devices with RPAO interface layers and RPECVD SiO_2 bulk dielectric films. These devices have been subjected to a 900 °C rapid thermal annealing (RTA) in an inert ambient (e.g., He).

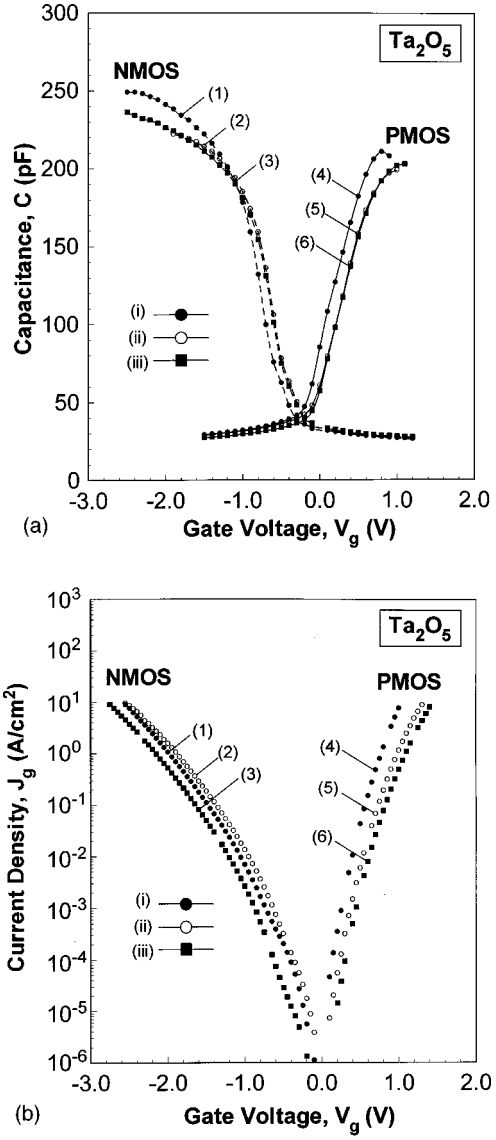

FIG. 4. (a) Capacitance–voltage, *C–V* and (b) current density–voltage, *J–V*, plots for devices with stacked gate dielectrics comprised of RPAO SiO_2 or nitrided RPAO SiO_2 and RPECVD Ta_2O_5. For comparison, these plots also include *C–V* and *J–V* traces for RPECVD Ta_2O_5 directly onto HF-last silicon substrates. These devices have been subjected to a 800 °C RTA in an inert ambient (e.g., He). The closed circles are for deposition on HF-last Si, the open circles for a RPAO SiO_2 layer, and the closed squares are for monolayer nitrided RPAO SiO_2. The EOT values are: (1) 1.09 nm, (2) 1.20 nm, (3) 1.17 nm, (4) 1.19 nm, (5) 1.30 nm, and (6) 1.29 nm.

bonded to the Si atoms at the Si–dielectric interface. In addition for a 90 s plasma-nitridation treatment[9,14] the nitrogen atom concentration of $7.5 \pm 0.5 \times 10^{14}$ cm^{-2} is equivalent to one nitrogen atom per Si(100) surface silicon atom, i.e., monolayer coverage. Figures 3(a) and 3(b) present electrical data that demonstrate the effects of interface nitridation. The data in Fig. 3(a), in the Fowler–Nordheim regime, indicate that tunneling is reduced in proportion to interface nitridation up to the monolayer level (90 s) and then saturates. The data in Fig. 3(b) indicate that reductions in the direct tunneling regime are smaller, $\sim 10\times$, compared to $50-60\times$ in the Fowler–Nordheim regime, and are effectively independent of the effective oxide thickness. The reductions in Figs. 3(a) and 3(b) are not due to shifts in the flat band voltage as determined from capacitance–voltage measurements.[10,11] The data in Figs. 4(a) and 4(b) are for device structures in which RPECVD Ta_2O_5, deposited at 300 °C, and subjected to a rapid thermal anneal at 800 °C replaced the RPECVD SiO_2 layers of the devices in Fig. 3. Monolayer interface nitridation has essentially the same effect on these devices with Ta_2O_5 dielectrics as it does in Fig. 3(b) with SiO_2. Additionally, the flatband voltages for the Ta_2O_5 capacitors on *n*-type and *p*-type substrates are the same to ± 0.005 V as for devices fabricated on substrates from the same lot with RPECVD SiO_2 dielectrics. Initial experiments performed on devices with RPECVD Zr silicate alloy dielectrics demonstrated improved interface properties, e.g., reduced hysteresis in capacitance–voltage, *C–V*, curves, when a RPAO step

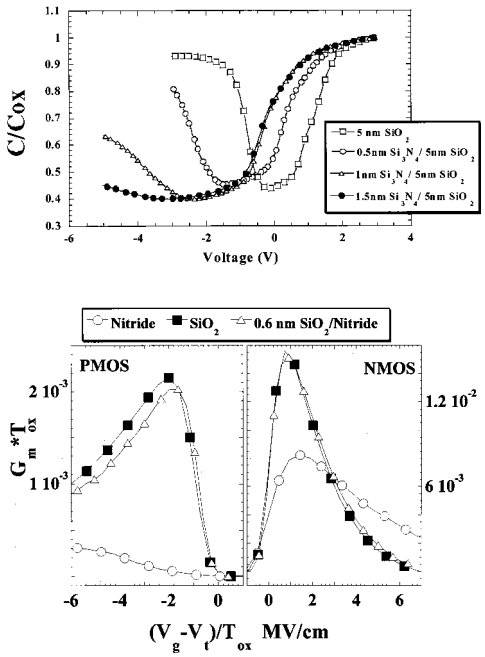

FIG. 5. (a) Negative voltage shifts of capacitance–voltage, *C–V*, curves for direct deposition of Si_3N_4 onto crystalline silicon. (b) $g_m - T_{ox}$ plots for NMOS and PMOS FETs with (i) RPAO and SiO_2 dielectric layers, (ii) Si_3N_4 dielectric layers deposited directly onto HF-last silicon, and with RPAO and Si_3N_4 dielectric layers. These devices have been subjected to a 900 °C RTA in an inert ambient (e.g., He).

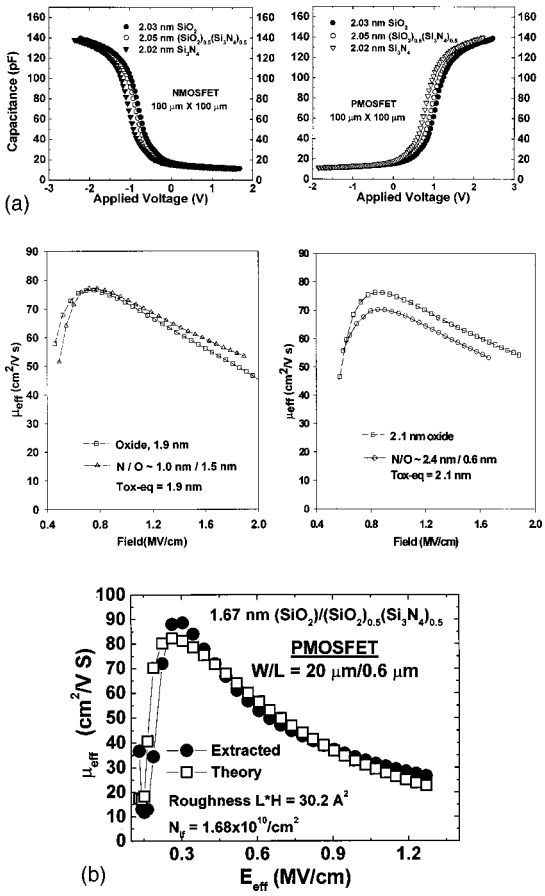

FIG. 6. (a) Capacitance–voltage, *C–V*, curves for NMOS and PMOS devices (EOT ~2 nm) with RPAO interfaces (~0.6 nm) and RPECVD oxide, oxynitride, and nitride bulk dielectric layers. (b) Mobility data for PMOS devices. These devices have been subjected to a 900 °C RTA in an inert ambient (e.g., He).

was used prior to deposition of the Zr silicate layer.[10] High resolution transmission electron microscopy micrographs indicated subcutaneous oxidation and/or silicate formation at an HF-last Si substrate surface.

The data in Figs. 5(a) and 5(b) present another aspect of interfacial oxidation. Direct deposition of nitride layers on Si(100), or the plasma nitridation of Si(100), each prior to oxide deposition, results in highly defective interfaces as indicated by flatband voltage shifts in Fig. 5(a) and the transconductance-EOT, $G_m - T_{ox}$, plots of Fig. 5(b).[5,6,13] The insertion of a RPAO step to form approximately 0.6 nm of interfacial SiO_2 restores the $g_m - T_{ox}$ product (which is proportional to channel mobility) to values essentially equal to those in devices with $Si–SiO_2$ interfaces [see Fig. 5(b)].

The data in Figs. 6(a) and 6(b) demonstrated flatband voltage shifts that are attributed to fixed positive charge at the internal dielectric interface.[18,19] Each of the devices in Fig. 6(a) have RPAO non-nitrided interface layers approximately 0.6 nm thick, and then RPECVD layers of (i) SiO_2, (ii) a Si oxynitride alloy with equal formula weights of SiO_2 and Si_3N_4, or (iii) hydrogenated (~15%–20% H) Si nitride. The capacitance–voltage, *C–V*, curves of Fig. 6(a) with Si oxynitride and nitride dielectrics, display systematic shifts to negative voltages independent of the substrate conductivity type. The shifts are larger for the devices with nitride dielectrics, ~0.15 V, than for the devices with the oxynitride dielectrics, ~0.4 eV. These shifts are consistent with fixed positive charge levels of about 7.5 and 2×10^{11} cm^{-2} at the respective internal dielectric interfaces. This interpretation is supported by plots of hole channel mobility versus surface field in Fig. 6(b). The decrease in peak channel mobility is larger for the nitride device with the thinner interfacial layer. In addition, the magnitude of the peak channel mobilities in the devices with Si nitrides is consistent with interfacial defect levels in the low to mid 10^{10} cm^{-2} range, supporting the assignment of the flatband voltage shirts to charge at the internal dielectric interface.

Finally, Fig. 7(a) displays current-density–voltage, *J–V*, plots and Fig. 7(b) displays time to breakdown for devices with gate dielectrics similar to those in Fig. 6(a).[18,19] Consistent with a tunneling model discussed in Ref. 20, the tunneling current at a bias voltage of 1 V across the dielectric in substrate inversion decreases by about one order of magnitude in going from the device with oxide dielectric to the device with nitride dielectric and an additional order of magnitude in going to the device with the oxynitride dielectric. As shown in Fig. 7(b), the values of time to breakdown at a fixed bias voltage across the dielectric increase inversely as the square of the current density across the stacked gate dielectric are not correlated with flatband voltage shifts and fixed positive charge discussed previously. This reliability scaling behavior is similar to what is discussed in Ref. 21 for stacked devices with Ta_2O_5 layers and $Si–SiO_2$ interfaces,

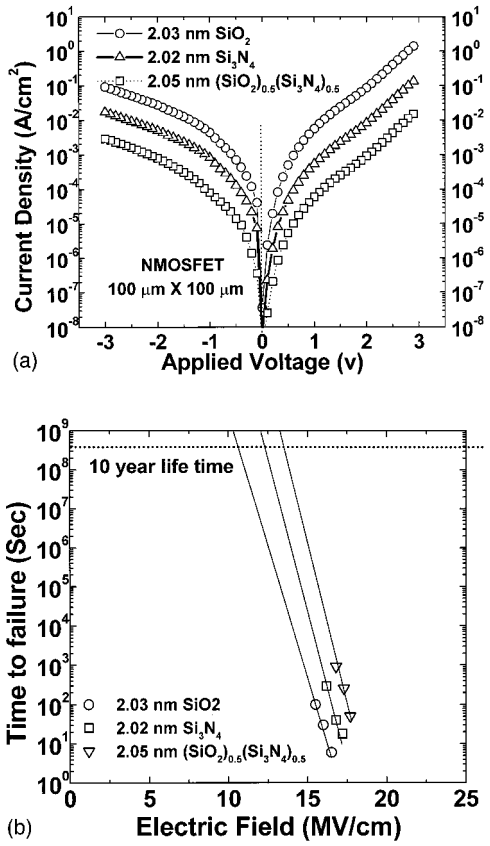

FIG. 7. (a) Current-density–voltage, *J–V*, plots for PMOS and NMOS FETs for the devices of Fig. 6(a). (b) Time to breakdown for NMOS devices of Figs. (a) and 6(a). These devices have been subjected to a 900 °C RTA in an inert ambient (e.g., He).

and has been attributed to a $Si–SiO_2$ interface controlled dielectric breakdown.

III. DISCUSSION

Several studies have addressed chemical bonding arrangements at thermally annealed $Si\text{-}SiO_2$ interfaces and have shown that *optimized* interfaces display transition regions $\sim$0.3 nm thick with excess suboxide bonding arrangements different from those expected at abrupt $Si–SiO_2$ metallurgical interfaces. These studies include: (i) XPS on ultrathin $Si–SiO_2$ interfaces using monochromatic synchrotron radiation,[22] (ii) *in situ* AES,[23] and (iii) *in situ* Fourier transform infrared (FTIR).[24] From the perspective of constraint theory, these transition regions are at an interface between a ''rigid'' crystalline Si material and an effectively ''floppy'' amorphous material, SiO_2, with a random covalent network structure.[25]

Applications of constraint theory to noncrystalline solids have focused primarily on bulk glasses,[26] thin films,[27] and $Si–SiO_2/Si_3N_4$ interfaces.[5] Recently, both theory[25] and experiment[28] have identified a new aspect of constraint theory by demonstrating that ''rigid'' or ''overconstrained'' and ''floppy'' or ''understrained'' microstructural regions in an alloy glass are separated by monolayer scale, self-organized interface layers that are ''overconstrained'' with

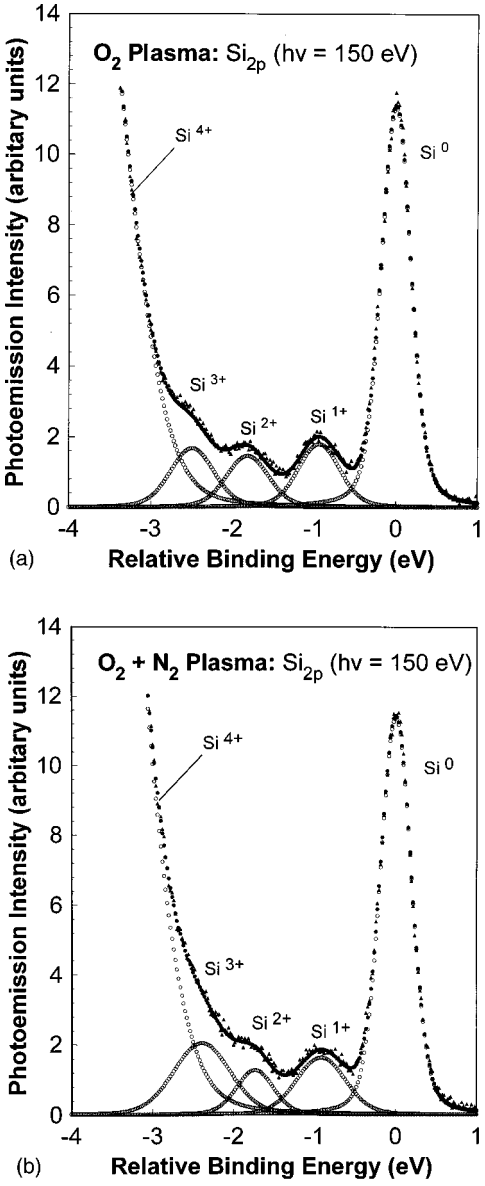

FIG. 8. (a) Synchrotron XPS data for RPAO interfaces without nitridation and (b) with monolayer interface nitridation. These devices have been subjected to a 900 °C RTA in an inert ambient (e.g., He).

respect to bonding coordination, yet not ''mechanically strained.'' Experiments cited in Refs. 22–24, combined with the theory of Refs. 25 and 29, suggest that the interfacial transition regions in advanced Si FET gate stacks are *intrinsic*, and result in part from entropy effects. These transition regions constitute a basic limitation for the aggressive scaling of CMOS silicon devices, as well as other semiconductor devices with similar ''steps'' in interfacial bonding coordination.

For example, the data in Figs. 3(a) and 3(b) have been explained in terms of a model based on XPS data of Ref. 21 [see Figs. 8(a) and 8(b)]. Analysis of the interfacial features in XPS data of Figs. 8(a) and 8(b) labeled I_1, I_2, and I_3, yields concentrations of silicon atoms in suboxide bonding environments in excess of those that are required for an ideal and abrupt interface. The excess silicon atom concentrations

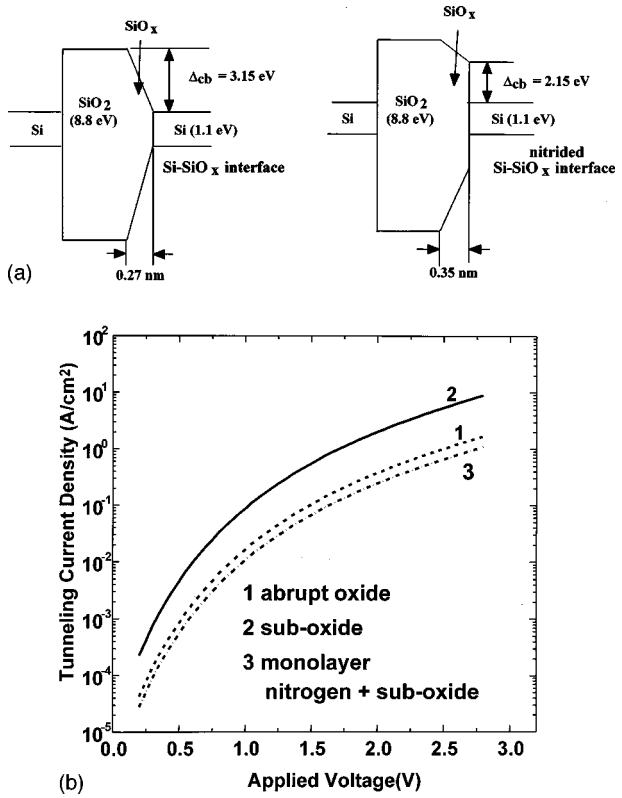

FIG. 9. (a) Interfacial band structure for tunneling calculations for RPAO and nitrided RPAO interfaces as derived from the XPS data of Fig. 8. (b) Calculated direct tunneling for devices with EOT=2 nm, for the interfaces of (a) and ideal $Si–SiO_2$ interfaces with no suboxide bonding.

have been equated to an average composition of SiO, thereby yielding converted Si areal densities into equivalent transition region widths of 0.27 nm for the non-nitrided interface and 0.35 nm for the nitrided interface.[20] Combining these XPS data with differences in optical SHG for interfaces with and without interfacial nitridation,[17] band models have been generated for the interfacial transition regions [see Fig. 9(a)]. After completing the stacked dielectrics with films of SiO_2 subject to the constraint of maintaining the same values of EOT, calculations of direct tunneling for substrate accumulation in Fig. 9(b) have yielded the reductions in the tunneling current density for devices with nitrided interfaces essentially the same as shown in Fig. 3(b).[20]

Constraint theory has also been invoked at semiconductor dielectric interfaces to explain differences in the defect densities between $Si–Si_3N_4$ and $Si–SiO_2$ interfaces.[5,6] In this application, mechanical bonding constraints at the interface have been characterized in terms of the average number of bonds per atom in the interfacial region. Following Ref. 5, the interface bonding structure is *defined* by 0.5 molecular layers of Si (0.5 atoms and two bonds), and 1.5 molecular layers of the dielectric film (SiO_2 or Si_3N_4). Interface nitridation has been taken into account by inserting one atomic layer of nitrogen between the Si substrate and SiO_2 layer. The $Si–SiO_2$ interface is used as a reference interface and is characterized by an average number of interfacial bonds/atom, $N_{av}^*=2.86$. Based on the above-described model interface bonding model, N_{av} equals 3.47 for a $Si–Si_3N_4$ interface, and 2.89 for a monolayer nitrided $Si–SiO_2$ interface. As discussed in Ref. 5, the concentration of defects relative to the $Si–SiO_2$ interface scales as $(N_{av}–N_{av}^*)^2$. Based on this scaling, the defect concentration at a $Si–Si_3N_4$ interface is expected to be about three orders of magnitude higher than at a monolayer nitrided $Si–SiO_2$ interface, consistent with experimental results.[5,6] Constraint theory and the associated scaling neither predict defect concentrations, nor identify the defect bonding arrangements. Instead, they yield scaling relationships that provide a useful guideline for comparisons of the type discussed previously. In the spirit of the scaling arguments, a value of interfacial $N_{av}\sim 3$ has been proposed in Ref. 5 as a demarcation between device quality and increasing defective interfaces, paralleling a similar criterion applied to defects in homogeneous amorphous thin films.[27]

TABLE I. Application of constraint theory to fixed charge at internal dielectric interfaces.

(i) Average number of bonds per atom in dielectric films, N_{av}			
SiO_2	Si_3N_4	Oxynitride alloy, $x=0.5$	Ta_2O_5
2.67	3.43	3.05	2.86[a]
(ii) Step in number of bonds/atom at internal interface with SiO_2, ΔN_{av}			
	Si_3N_4	Oxynitride alloy, $x=0.5$	Ta_2O_5
	0.76	0.38	0.19
(a) Scaling factor for defects/cm^2 scale as $[\Delta N_{av}]^2$			
	0.58	0.14	0.04
(b) Normalized ratio of defects/cm^2			
	16	4	1
(c) Fixed charge at internal dielectric interface, Q_f (cm^{-2})			
	Si_3N_4	Oxynitride alloy, $x=0.5$	Ta_2O_5
	7.5×10^{11}	2×10^{11}	$<0.6\times10^{11}$ [b]

[a]Based on an extension of the resonating bond model proposed by Phillips (Ref. 34).
[b]Estimated from *C–V* data.

Scaling theory also provides an explanation for differences in fixed charge at internal dielectric interfaces. In this application, the appropriate scaling variable is the difference in the average number of bonds per atom, ΔN_{av}, on either side of the internal dielectric interface. This model will now be applied to the *C–V* data of Figs. 4(a) and 6(a). The interfaces being addressed are between the interfacial SiO_2 and Si nitride and oxynitride in Fig. 5(a), and SiO_2 and Ta_2O_5 in Fig. 4(a). This application of constraint theory is summarized in Table I, which includes: (i) the average number of bonds/atom, N_{av}, for SiO_2, Si_3N_4, an oxynitride alloy with a composition $(SiO_2)_{0.5}(Si_3N_4)_{0.5}$, and Ta_2O_5, (ii) the step in the average number of bonds/atom, ΔN_{av}, for internal interfaces between SiO_2 and Si_3N_4, $(SiO_2)_{0.5}(Si_3N_4)_{0.5}$, and Ta_2O_5, (iii) computed values of $[\Delta N_{av}]^2$ and normalized values of $[\Delta N_{av}]^2$, and (iv) values of fixed charge, Q_f, obtained from an analysis of *C–V* data. The values of Q_f are obtained from the relative shifts in *C–V* curves, as for example in Fig. 5(a). The limiting value of Q_f for the internal $SiO_2–Ta_2O_5$ interface is estimated from the accuracy of the voltage determination in the *C–V* traces in Fig. 3(a). These data are consistent with the scaling expected from constraint theory. As noted previously and in Refs. 5 and 6, constraint theory can

be used to predict relative defect concentrations, but cannot predict the actual defect concentrations, the local bonding arrangements at the defect sites, or the energies of the defects relative to band edges.

We now turn to the reliability data in Fig. 7(b). A comparison between these data and those in Fig. 7(a) indicate that the time to breakdown at a given applied voltage increases as the square of the tunneling current decreases. Additionally, the time to breakdown does not scale inversely as the relative concentrations of Q_f at the internal dielectric interface, so that there is no experimental evidence that these internal interfaces play a significant role in intrinsic device breakdown. On the other hand, the data in Fig. 7(b) are consistent with the time to breakdown being determined predominantly by the interfacial SiO_2 layer and its interface with the crystalline Si substrate, as has been suggested in Ref. 21, which focused on devices with SiO_2 interface layers and Ta_2O_5 bulk dielectrics. The devices of Fig. 7(b) do not have nitrided Si–SiO_2 interfaces; however, as has been demonstrated in Fig. 2, interfacial nitridation reduces both Fowler–Nordheim and direct tunneling so that additional increases in time to breakdown are anticipated in devices with nitrided interfaces.

IV. CONCLUSIONS

The application of constraint theory to silicon dielectric and internal dielectric interfaces has identified fundamental limitations on the performance and reliability of FET devices that include stacked gate dielectrics. To date the only interfacial layers that have been demonstrated not to degrade FET current drive, or equivalently reduce the effective channel mobilities of both holes and electrons, are those which are either SiO_2 or nitrided SiO_2. Based on these observations, this paper has focused on stacked gate dielectrics which are comprised of a Si substrate, an ultrathin SiO_2 or nitrided SiO_2 layer, and higher-k bulk dielectric films such as Si_3N_4, a Si-oxynitride alloy [e.g., $(SiO_2)_{0.5}(Si_3N_4)_{0.5}$], or Ta_2O_5. Other research results for devices with stacked gate dielectrics have demonstrated excellent performance and reliability when the devices included an interfacial layer of SiO_2, either grown prior to deposition, or formed during a postdeposition anneal.[30,31] Therefore, the conclusions drawn in this paper extend to other stacked gate dielectric structures as well, including the devices of Refs. 30 and 31 where the interfacial layers were not prepared by RPAO.

The paper has addressed three different aspects of interface bonding and defect structure that limit device performance and reliability, (i) interfacial transition regions between the Si substrate and a dielectric film, (ii) differences in defect concentrations at interfaces between Si and different dielectric films, and (iii) differences in defect concentrations at internal interfaces between different dielectrics.

It has been shown that the interfacial transition regions with suboxide bonding in excess of what is required at an abrupt interface between Si and SiO_2 are typically about 0.3 nm thick.[22–24] These regions are formed during oxidation of the silicon substrate and minimization of their spatial extent requires a postoxidation anneal, e.g., 30 s to 1 min at 900 °C. It is particularly noteworthy that interfaces prepared by thermal oxidation at 900 °C also require an anneal at the same temperature to optimize interfacial smoothness or equivalently minimize suboxide bonding.[32] The existence of an interfacial transition region between crystalline silicon and SiO_2 is anticipated on the basis of constraint theory as has been discussed in Ref. 25. Constraint theory predicts that a transition region of the order of one molecular layer must be present at the interface between two materials in which the number of bonds/atom is different, or equivalently at which the number of bonding constraints per atom is different. The resulting interfacial region is overconstrained with respect to the lower constraint partner of the interface, and is self-organized in a way that minimizes the development of mechanical strain. Interface layers of this type have been reported within glassy alloys with microstructure composed of a floppy or underconstrained constituent, and an overconstrained bonding partner.[28] ''Floppy'' is defined in terms of a maximum average bonding coordination of 2.4,[26] except for SiO_2, where 2.67 is more appropriate due to unusually small bonding forces at the oxygen atom sites.[27] Values of $N_{av} < 2.4$ (or 2.67) define a regime in which the average number of bond constraints per atom is lower than the network dimensionality.[27] These concepts, originally applied to bulk glasses, have been extended to thin film amorphous materials[27] and more recently to internal interfaces between crystalline and amorphous materials.[5] In this paper, it has been shown that control of interfacial bonding structure by interfacial nitridation can modify interfacial transition regions between Si and SiO_2 and produce significant reductions in tunneling currents, thereby improving device performance. The experiments of Refs. 22–24, combined with the theory of Refs. 25 and 29, suggest that the interfacial transition regions in advanced Si FET gate stacks are *intrinsic*, and result from entropy effects. Entropy is not a factor in the crystalline substrate, but configurational entropy is one of the most important factors in allowing for the formation of amorphous covalent random networks. Therefore, it is not surprising that it plays a role at the interface between a crystalline solid and an ideal amorphous covalent random network solid.

A second application of constraint theory as applied to Si–dielectric interfaces has been discussed in detail in Refs. 5 and 6, and deals with interfaces between Si and other dielectrics such as Si nitride. These papers demonstrate that differences in performance, e.g., channel mobilities of FETs and threshold and flatband voltages, can be understood in terms of the average number of interfacial bonds/atom. The bonding model proposed in Ref. 5 gives an average interfacial N_{av} of 2.86 for Si–SiO_2, which increases to (i) 2.89 for a Si–SiO_2 interface with monolayer nitridation, and (ii) 3.47 for Si–Si_3N_4. Based on an empirical criterion for quantification of defect formation, mostly dangling bonds in amorphous thin films, a N_{av} value of about 3 marks a boundary between device quality and defective interfaces. This is consistent with the experimental results presented in Ref. 5,

which show that interfacial defects increase as N_{av} increases following the quadratic scaling predicted by constraint theory.

Finally, constraint theory has accounted for differences in the fixed charge at internal interfaces of stacked gate dielectrics. In particular, the quadratic scaling with ΔN_{av} has accounted for differences in fixed charge at $SiO_2-Si_3N_4$, $SiO_2-(SiO_2)_{0.5}(Si_3N_4)_{0.5}$, and $SiO_2-Ta_2O_5$ internal interfaces. This interfacial charge shifts flatband voltages and therefore will also change threshold voltages in FETs, however, it does not appear to degrade reliability. If the charge is high enough, and if the interfacial layer is sufficiently thin, it can also reduce effective channel mobilities as in *p*-channel FETs with bulk Si_3N_4 dielectrics.

In conclusion, future generations of silicon-based FET devices will have to incorporate deposited dielectric layers with dielectric constants greater than that of SiO_2 in order to reduce direct tunneling currents which can be present at the scaled down operating voltages. Experiments to date indicate that these dielectrics will have to have a stacked structure comprised of an ultrathin SiO_2 interfacial layer, $\sim 0.5-0.6$ nm thick, in order to ensure device performance and reliability that is comparable to devices with thermally grown oxides of past generations of devices. This paper has shown that there are fundamental physical effects that limit the bonding structure of these interfacial oxides, both the silicon substrate and at the internal interface with high-*k* dielectric. Monolayer scale transition regions at the $Si-SiO_2$ interface are common for both homogeneous and stacked structures, and increases in direct tunneling associated with the ''average SiO'' interfacial composition can be reduced by incorporation of about one monolayer of nitrogen atoms bonded directly at the $Si-SiO_2$ interface [see Fig. 8(b)]. This interfacial bonding arrangement is consistent with optical SHG studies,[17] and has been confirmed by recent experiments using synchrotron XPS.[33]

Recent insights into bonding counting for silicate glasses,[34] can be extended to other non-SiO_2-based dielectrics as well, and have shown that most of the important high-*k* candidates, such as Ta_2O_5, ZrO_2, HfO_2, and Zr, and Hf silicates have values of N_{av} that, when combined with a crystalline silicon substrate, will yield interfacial values of $N_{av}<3$. This means that high defect concentrations will not result from overconstraint; however, other aspects of bonding, e.g., interfacial Si–metal atom (silicide) bonds,[35] as well as Si–O–metal atom groups, are expected to degrade performance and/or reliability. For example, fixed positive charge associated with silicide bonds, as well as second-neighbor effects on interfacial silicon atoms from different local bonding arrangements in the dielectric at the silicon buried interface, can enhance scattering of charged carriers in the channel region and decrease channel mobilities.

ACKNOWLEDGMENTS

This research has been supported in part by the ONR, SRC, and SRC/SEMATECH FEP Center.

[1]G. Lucovsky, S. S. Kim, D. V. Tsu, G. G. Fountain, and R. J. Markunas, J. Vac. Sci. Technol. B **7**, 861 (1989).
[2]S. S. Kim and G. Lucovsky, J. Vac. Sci. Technol. A **8**, 2039 (1990).
[3]T. Yasuda, Y. Ma, S. Habermehl, and G. Lucovsky, Appl. Phys. Lett. **60**, 434 (1992).
[4]S. V. Hattangady, H. Niimi, and G. Lucovsky, J. Vac. Sci. Technol. A **14**, 3017 (1996).
[5]G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. C. Phillips, Appl. Phys. Lett. **74**, 2005 (1999).
[6]V. Misra, Z. Wang, Y. Wu, H. Niimi, G. Lucovsky, J. J. Wortman, and J. R. Hauser, J. Vac. Sci. Technol. B **17**, 1836 (1999).
[7]G. Lucovsky, J. J. Wortman, T. Yasuda, X.-L. Xu, V. Misra, S. V. Hattangady, Y. Ma, and B. Hornung, J. Vac. Sci. Technol. B **12**, 2839 (1994).
[8]G. Lucovsky, IBM J. Res. Dev. **43**, 301 (1999), and references therein.
[9]H. Niimi, H. Y. Yang, and G. Lucovsky, in *Characterization and Metrology for ULSI Technology: 1998 International Conference*, edited by D. G. Seiler, A. C. Diebold, W. M. Bullis, T. J. Shaffner, R. McDonald, and E. J. Walters (AIP, Woodbury, NY, 1998), p. 273.
[10]G. B. Rayner, R. Therrien, and G. Lucovsky (unpublished).
[11]P. K. Roy and I. C. Kizikyalli, Appl. Phys. Lett. **72**, 2835 (1998); I. C. Kizikyalli, R. Y. S. Huang, and P. K. Roy, IEEE Electron Device Lett. **19**, 423 (1998).
[12]H. Niimi and G. Lucovsky (unpublished).
[13]H. Niimi and G. Lucovsky, J. Vac. Sci. Technol. A **17**, 3185 (1999).
[14]H. Niimi and G. Lucovsky, J. Vac. Sci. Technol. B **17**, 2610 (1999).
[15]Y. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998).
[16]C. R. Parker, G. Lucovsky, and J. R. Hauser, IEEE Electron Device Lett. **19**, 106 (1998).
[17]G. Lucovsky, H, Niimi, K. Koh, D. R. Lee, and Z. Jing, in *The Physics of SiO_2 and Si–SiO_2 Interfaces-3*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (The Electrochemical Society, Pennington, NJ, 1996), p. 441.
[18]H. Yang and G. Lucovsky, Tech. Dig. Int. Electron. Devices Meet. 245, (1999).
[19]G. Lucovsky, H. Yang, Y. Wu, and H. Niimi, Banpis Proc. (in press).
[20]H. Yang, H. Niimi, J. W. Keister, G. Lucovsky, and Jack E. Rowe, IEEE Electron Device Lett. **21**, 76 (2000).
[21]R. Degraeve, B. Kaczer, M. Houssa, G. Groeseneken, M. Hyens, J. S. Jeon, and A. Halliyal, Tech. Dig. Int. Electron. Devices Meet. 327 (1999).
[22]J. W. Keister, J. E. Rowe, J. J. Kolodzie, H. Niimi, N.-S. Tao, T. E. Madey, and G. Lucovsky, J. Vac. Sci. Technol. A **17**, 1340 (1999).
[23]M. Weldon *et al.*, J. Vac. Sci. Technol. B **17**, 1795 (1999).
[24]G. Lucovsky, A. Banerjee, B. Hinds, B. Claflin, K. Koh, and H. Yang, J. Vac. Sci. Technol. B **15**, 1074 (1997).
[25]M. F. Thorpe, J. Non-Cryst. Solids (in press).
[26]J. C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979); **43**, 37 (1981).
[27]G. Lucovsky and J. C. Phillips, J. Non-Cryst. Solids **227**, 1221 (1998).
[28]P. Boolchand *et al.*, J. Non-Cryst Solids (in press).
[29]Y. Yu and J. Tersoff (unpublished).
[30]S. C. Song, H. F. Luan, Y. Y. Chen, M. Gardner, J. Fulford, M. Allen, and D. L. Kwong, Tech. Dig. Int. Electron. Devices Meet. 373 (1998).
[31]T. P. Ma, IEEE Trans. Electron Devices **45**, 680 (1998); X. Guo and T. P. Ma, IEEE Electron Device Lett. **19**, 207 (1998).
[32]X. Chen and J. M. Gibson, Appl. Phys. Lett. **70**, 1462 (1997).
[33]J. W. Keister, J. E. Rowe, Y.-M. Lee, H. Niimi, G. Lucovsky, and G. J. Lapeyre (unpublished).
[34]J. C. Phillips (unpublished).
[35]A. Y. Mao, K. A. Son, J. M. White, D. L. Kwong, D. A. Roberts, and R. N. Vrtis, Mater. Res. Soc. Symp. Proc. **567**, 473 (1999).