JOURNAL OF APPLIED PHYSICS VOLUME 89, NUMBER 10 15 MAY 2001

APPLIED PHYSICS REVIEW

High-κ gate dielectrics: Current status and materials properties considerations

G. D. Wilk[a]
Agere Systems, Electronic Device Research Laboratory, Murray Hill, New Jersey 07974

R. M. Wallace[b]
University of North Texas, Department of Materials Science, Denton, Texas 76203

J. M. Anthony
University of South Florida, Center for Microelectronics Research, Tampa, Florida 33620

(Received 9 November 2000; accepted for publication 19 January 2001)

Many materials systems are currently under consideration as potential replacements for SiO_2 as the gate dielectric material for sub-0.1 μm complementary metal–oxide–semiconductor (CMOS) technology. A systematic consideration of the required properties of gate dielectrics indicates that the key guidelines for selecting an alternative gate dielectric are (a) permittivity, band gap, and band alignment to silicon, (b) thermodynamic stability, (c) film morphology, (d) interface quality, (e) compatibility with the current or expected materials to be used in processing for CMOS devices, (f) process compatibility, and (g) reliability. Many dielectrics appear favorable in some of these areas, but very few materials are promising with respect to all of these guidelines. A review of current work and literature in the area of alternate gate dielectrics is given. Based on reported results and fundamental considerations, the pseudobinary materials systems offer large flexibility and show the most promise toward successful integration into the expected processing conditions for future CMOS technologies, especially due to their tendency to form at interfaces with Si (e.g. silicates). These pseudobinary systems also thereby enable the use of other high-κ materials by serving as an interfacial high-κ layer. While work is ongoing, much research is still required, as it is clear that any material which is to replace SiO_2 as the gate dielectric faces a formidable challenge. The requirements for process integration compatibility are remarkably demanding, and any serious candidates will emerge only through continued, intensive investigation. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1361065]

I. INTRODUCTION. 5243
II. SCALING AND IMPROVED PERFORMANCE. 5244
III. METAL-INSULATOR-SEMICONDUCTOR (MIS) GATE STACK STRUCTURES. 5245
IV. SCALING LIMITS FOR CURRENT GATE DIELECTRICS. 5247
  A. Ultrathin SiO_2 properties. 5247
  B. Ultrathin SiO_2 reliability. 5248
  C. Boron penetration and surface preparation. 5249
  D. SiO_xN_y and Si–N/SiO_2 dielectrics. 5249
  E. Fundamental limitations. 5249
  F. Device structures. 5250
V. ALTERNATIVE HIGH-κ GATE DIELECTRICS. 5250
  A. High-κ candidates from memory applications. . 5251
  B. Issues for interface engineering. 5252
  C. Recent high-κ results. 5253
    1. Group IIIA and IIIB metal oxides. 5254
    2. Group IVB metal oxides. 5256
    3. Pseudobinary alloys. 5262
    4. High-κ device modeling and transport. 5265
VI. MATERIALS PROPERTIES CONSIDERATIONS. 5266
  A. Permittivity and barrier height. 5266
  B. Thermodynamic stability on Si. 5268
  C. Interface quality. 5269
  D. Film morphology. 5270
  E. Gate compatibility. 5271
  F. Process compatibility. 5272
  G. Reliability. 5272
VII. CONCLUSIONS. 5273

I. INTRODUCTION

The rapid progress of complementary metal–oxide–semiconductor (CMOS) integrated circuit technology since the late 1980's has enabled the Si-based microelectronics industry to simultaneously meet several technological requirements to fuel market expansion. These requirements in-

---

[a] G. D. Wilk is formerly of Bell Laboratories, Lucent Technologies; electronic mail: gwilk@agere.com
[b] Electronic mail: rwallace@unt.edu

clude performance (speed), low static (off-state) power, and a wide range of power supply and output voltages.[1] This has been accomplished by developing the ability to perform a calculated reduction of the dimensions of the fundamental active device in the circuit: the field effect transistor (FET)—a practice termed ''scaling.''[2–4] The result has been a dramatic expansion in technology and communications markets including the market associated with high-performance microprocessors as well as low static-power applications, such as wireless systems.[5]

It can be argued that the key element enabling the scaling of the Si-based metal–oxide–semiconductor field effect transistor (MOSFET) is the materials (and resultant electrical) properties associated with the dielectric employed to isolate the transistor gate from the Si channel in CMOS devices for decades: silicon dioxide. The use of amorphous, thermally grown SiO_2 as a gate dielectric offers several key advantages in CMOS processing including a stable (thermodynamically and electrically), high-quality Si–SiO_2 interface as well as superior electrical isolation properties. In modern CMOS processing, defect charge densities are on the order of $10^{10}/\mathrm{cm}^2$, midgap interface state densities are $\sim 10^{10}/\mathrm{cm}^2\,\mathrm{eV}$, and hard breakdown fields of 15 MV/cm are routinely obtained and are therefore expected regardless of the device dimensions. These outstanding electrical properties clearly present a significant challenge for any alternative gate dielectric candidate.

II. SCALING AND IMPROVED PERFORMANCE

The industry's demand for greater integrated circuit functionality and performance at lower cost requires an increased circuit density, which has translated into a higher density of transistors on a wafer.[3] This rapid shrinking of the transistor feature size has forced the channel length and gate dielectric thickness to also decrease rapidly. As will be discussed in the next few sections, the current CMOS gate dielectric SiO_2 thickness can scale to at least 13 Å, but there are several critical device parameters that must be balanced during this process.

The improved performance associated with the scaling of logic device dimensions can be seen by considering a simple model for the drive current associated with a FET.[1] The drive current can be written (using the gradual channel approximation) as

$$I_D = \frac{W}{L}\mu C_{\mathrm{inv}}\left(V_G - V_T - \frac{V_D}{2}\right)V_D, \tag{1}$$

where W is the width of the transistor channel, L is the channel length, μ is the channel carrier mobility (assumed constant here), C_{inv} is the capacitance density associated with the gate dielectric when the underlying channel is in the inverted state, V_G and V_D are the voltages applied to the transistor gate and drain, respectively, and the threshold voltage is given by V_T. It can be seen that in this approximation the drain current is proportional to the average charge across the channel (with a potential $V_D/2$) and the average electric field (V_D/L) along the channel direction. Initially, I_D increases linearly with V_D and then eventually saturates to a maximum when $V_{D,\mathrm{sat}} = V_G - V_T$ to yield

$$I_{D,\mathrm{sat}} = \frac{W}{L}\mu C_{\mathrm{inv}}\frac{(V_G - V_T)^2}{2}. \tag{2}$$

The term $(V_G - V_T)$ is limited in range due to reliability and room temperature operation constraints, since too large a V_G would create an undesirable, high electric field across the oxide. Furthermore, V_T cannot easily be reduced below about 200 mV, because $kT \sim 25\,\mathrm{mV}$ at room temperature. Typical specification temperatures ($\leqslant 100\,^\circ\mathrm{C}$) could therefore cause statistical fluctuations in thermal energy, which would adversely affect the desired the V_T value. Thus, even in this simplified approximation, a reduction in the channel length or an increase in the gate dielectric capacitance will result in an increased $I_{D,\mathrm{sat}}$.

In the case of increasing the gate capacitance, consider a parallel plate capacitor (ignoring quantum mechanical and depletion effects from a Si substrate and gate)[6]

$$C = \frac{\kappa \epsilon_0 A}{t}, \tag{3}$$

where κ is the dielectric constant (also referred to as the relative permittivity in this article) of the material,[7] ϵ_0 is the permittivity of free space ($= 8.85 \times 10^{-3}\,\mathrm{fF}/\mu\mathrm{m}$), A is the area of the capacitor, and t is the thickness of the dielectric. This expression for C can be rewritten in terms of t_{eq} (i.e., equivalent oxide thickness) and κ_{ox} ($= 3.9$, dielectric constant of SiO_2) of the capacitor. The term t_{eq} represents the theoretical thickness of SiO_2 that would be required to achieve the same capacitance density as the dielectric (ignoring issues such as leakage current and reliability). For example, if the capacitor dielectric is SiO_2, $t_{\mathrm{eq}} = 3.9\epsilon_0(A/C)$, and a capacitance density of $C/A = 34.5\,\mathrm{fF}/\mu\mathrm{m}^2$ corresponds to $t_{\mathrm{eq}} = 10\,\text{Å}$. Thus, the physical thickness of an alternative dielectric employed to achieve the equivalent capacitance density of $t_{\mathrm{eq}} = 10\,\text{Å}$ can be obtained from the expression

$$\frac{t_{\mathrm{eq}}}{\kappa_{\mathrm{ox}}} = \frac{t_{\mathrm{high}-\kappa}}{\kappa_{\mathrm{high}-\kappa}}$$

$$\text{or simply,}\quad t_{\mathrm{high}-\kappa} = \frac{\kappa_{\mathrm{high}-\kappa}}{\kappa_{\mathrm{ox}}}t_{\mathrm{eq}} = \frac{\kappa_{\mathrm{high}-\kappa}}{3.9}t_{\mathrm{eq}}. \tag{4}$$

A dielectric with a relative permittivity of 16 therefore affords a physical thickness of ~40 Å to obtain $t_{\mathrm{eq}} = 10\,\text{Å}$. (As noted above, actual performance of a CMOS gate stack does not scale directly with the dielectric due to possible quantum mechanical and depletion effects.)[6]

From a CMOS circuit performance point of view, a performance metric considers the dynamic response (i.e., charging and discharging) of the transistors, associated with a specific circuit element, and the supply voltage provided to the element at a representative (clock) frequency. A common element employed to examine such switching time effects is a CMOS inverter.[1] This circuit element is shown in Fig. 1 where the input signal is attached to the gates and the output signal is connected to both the *n*-type MOS (nMOS) and *p*-type MOS (pMOS) transistors associated with the CMOS

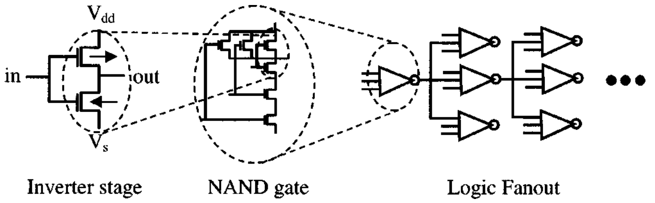

FIG. 1. Components used to test a CMOS FET technology. V_{DD} and V_S serve as the source and drain voltages, respectively, and are common to the NAND gates shown. Each NAND gate is connected to three others resulting in a fanout of 3.

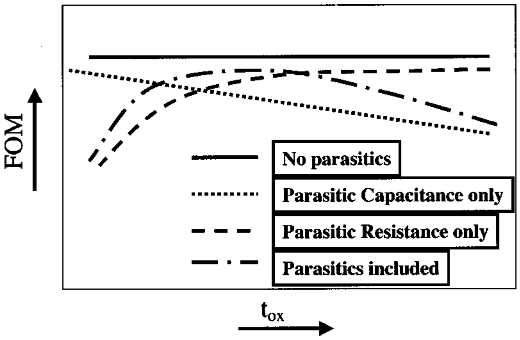

FIG. 2. FOM as a function of equivalent oxide thickness, t_{eq}. Parasitic capacitances and resistances result in transistor design tradeoffs to optimize performance.

stage. The switching time is limited by both the fall time required to discharge the load capacitance by the n-FET drive current and the rise time required to charge the load capacitance by the p-FET drive current. That is, the switching response times are given by[1]

$$\tau=\frac{C_{\mathrm{LOAD}}V_{DD}}{I_D}, \quad \text{where} \quad C_{\mathrm{LOAD}}=FC_{\mathrm{GATE}}+C_j+C_i, \tag{5}$$

and C_j and C_i are parasitic junction and local interconnection capacitances, respectively. The ''fan out'' for interconnected devices is given by the factor ''F.'' Ignoring delay in gate electrode response, as $\tau_{\mathrm{GATE}} \ll \tau_{n,p}$, the average switching time is therefore

$$\bar{\tau}=\frac{\tau_p+\tau_n}{2}=C_{\mathrm{LOAD}}V_{DD}\left\{\frac{1}{I_D^n+I_D^p}\right\}. \tag{6}$$

The load capacitance in the case of a single CMOS inverter is simply the gate capacitance if one ignores parasitic contributions such as junction and interconnect capacitance. Hence, an increase in I_D is desirable to reduce switching speeds. For more realistic estimates of microprocessor performance, the load capacitance is connected (''fanned out'') to other inverter elements in a predetermined fashion. When coupled with other NMOS/PMOS transistor pairs in the configuration shown in Fig. 1, one can create a logic ''NAND'' gate which can be used to investigate the dynamic response of the transistors and thus examine their performance under such configurations. For example, in microprocessor estimates, a fan out of $F=3$ is often employed, as shown in Fig. 1.[1]

One can then characterize the performance of a circuit (based on a particular transistor structure) through this switching time. To do this, various ''figures of merit'' (FOM) have been proposed which incorporate parasitic capacitance as well as the influence of gate sheet resistance on the switching time.[8] For example, a common FOM employed is related to Eq. (6) simply by

$$\mathrm{FOM}\cong\frac{1}{\bar{\tau}}=\frac{2}{\tau_p+\tau_n}. \tag{7}$$

In the case where parasitics are ignored, it is easily seen then that an increase in the device drive current I_D results in a decrease in the switching time and an increase in the FOM value (performance). Even in this simple model, however, the incorporation of parasitic effects, results in the ''clamping'' of FOM improvement, despite an increase in the gate dielectric capacitance. This can be seen in Fig. 2 where various FOM calculations are plotted as a function of an ''equivalent oxide thickness,'' t_{eq}, as described earlier.

Each FOM calculation shown in Fig. 2 corresponds to specific assumptions on the values of parasitic capacitance and gate sheet resistance, as indicated (gate length is kept constant in this analysis). Important aspects such as gate induced drain leakage and reliability are ignored in this simple model.[1] Nevertheless, the result of the FOM calculation shown in Fig. 1 indicates that tradeoffs on all aspects of the transistor design and scaling, including parasitics, must be carefully considered in order to increase the circuit performance.[8]

III. METAL–INSULATOR–SEMICONDUCTOR (MIS) GATE STACK STRUCTURES

Figure 3 provides the reader a schematic overview of the various regions associated with the gate stack of a CMOS FET (regions are separated simply to clarify the following discussion). The gate dielectric insulates the gate electrode (gate) from the Si substrate. Gate electrodes in modern CMOS technology are composed of polycrystalline Si (poly-Si) which can be highly doped (e.g. by ion implantation) and subsequently annealed in order to substantially increase conductivity. The selection of the dopant species and concentra-

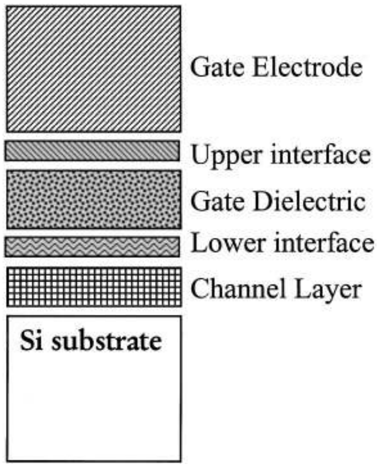

FIG. 3. Schematic of important regions of a field effect transistor gate stack.

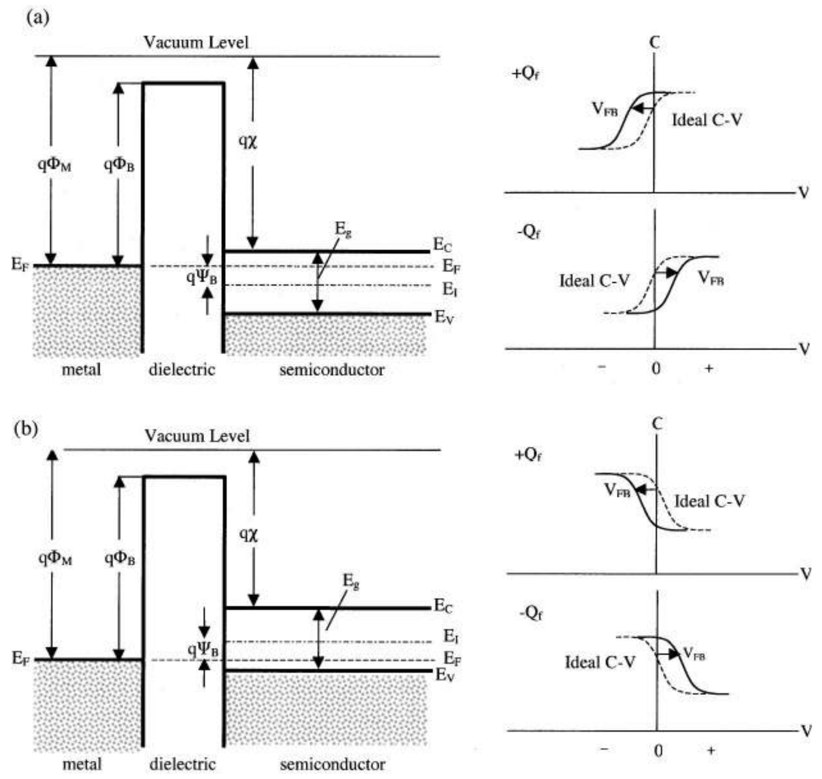

FIG. 4. Energy-band diagrams and associated high-frequency $C-V$ curves for ideal MIS diodes for (a) n-type and (b) p-type semiconductor substrates. For these ideal diodes, $V=0$ corresponds to a flatband condition. For dielectrics with positive $(+Q_f)$ or negative $(-Q_f)$ fixed charge, an applied voltage (V_{FB}) is required to obtain a flatband condition and the corresponding $C-V$ curve shifts in proportion to the fixed charge. (after Refs. 9 and 10).

tion permits the adjustment of the poly-Si Fermi level for either nMOS or pMOS FETs. Metals can also be used as the gate electrode, and, in fact, are commonly used for evaluation of capacitor structures. Work is underway to find suitable metal gates for CMOS (see Sec. VI E).

The interfaces with either the gate or the Si channel region are particularly important in regard to device performance. These regions, ~5 Å thick, serve as a transition between the atoms associated with the materials in the gate electrode, gate dielectric and Si channel. As will be discussed, these interface regions can alter the overall capacitance of the gate stack, particularly if they have a thickness which is substantial relative to the gate dielectric. Additionally, these interfacial regions can be exploited to obtain desirable properties. The upper interface, for example, can be engineered in order to block boron outdiffusion from the p^+ poly-Si gate. The lower interface, which is in direct contact with the CMOS channel region, must be engineered to permit low interface trap densities (e.g. dangling bonds) and minimize carrier scattering (maximize mobility) in order to obtain reliable, high performance.

It is instructive to consider the band diagrams for the MIS structures discussed in this review. Figure 4 shows the energy-band diagrams for ideal MIS diode structures using (a) n-type and (b) p-type semiconductor substrates.[9,10] For these ideal structures, at $V=0$ applied voltage on the metal gate, the work function difference between the metal and semiconductor, Φ_{MS}, is zero

$$\Phi_{\mathrm{MS}}=\Phi_{\mathrm{M}}-\left(\chi+\frac{E_g}{2q}-\Psi_B\right)=0; \quad n\text{-type}$$

$$\Phi_{\mathrm{MS}}=\Phi_{\mathrm{M}}-\left(\chi+\frac{E_g}{2q}+\Psi_B\right)=0; \quad p\text{-type}, \tag{8}$$

where Φ_{M} is the metal work function, χ is the semiconductor electron affinity, E_g is the semiconductor band gap, Φ_B is the potential barrier between the metal and dielectric, and ψ_B is the potential difference between the Fermi level E_F and the intrinsic Fermi level, E_I. Under these conditions, the energy bands are flat across the structure as shown in Fig. 4 and $V=V_{\mathrm{FB}}=0$, where V_{FB} is the flat band voltage (i.e., the voltage required to bring the Fermi levels into alignment). A more typical case is that the Fermi levels of the electrode and substrate are misaligned by an energy difference, and a voltage $(V_{\mathrm{FB}}\neq 0)$ must be applied to bring the Fermi levels into alignment.

Many dielectrics exhibit a fixed charge (Q_f), however, resulting in a required applied voltage $V=V_{\mathrm{FB}}\neq 0$ to achieve a flat band condition. The amount of fixed charge can be related to the measured V_{FB} value by the expression[9]

$$V_{\mathrm{FB}}=\Phi_{\mathrm{MS}}\pm Q_f/C_{\mathrm{acc}}, \tag{9}$$

where C_{acc} is the measured capacitance in accumulation. Thus, a value for fixed charge density Q_f can be determined from measured values of V_{FB}, Φ_{MS} and C_{acc}. The sign of the fixed charge is also important, as negative fixed charge correlates with the plus sign in Eq. (9), and positive fixed charge correlates with the minus sign. These expressions will be discussed further in Sec. V C 2.

The source of such fixed charge, often though *not always* positive, is thought to originate from the detailed bonding of the atoms associated with the dielectric near the dielectric/

semiconductor interface. Several proposed explanations for the cause of the observed fixed charge will be discussed in Sec. V C 2. Figure 4 shows that for positive Q_f, a negative shift in the V_{FB} from ideal conditions (where $V=0$) is required for both *n*-type and *p*-type MIS structures. Similarly, a positive V_{FB} is required for negative Q_f.

Most of the alternate dielectric candidates examined to date appear to have a substantial amount of fixed charge, which could present significant issues for CMOS applications. Given the scaling limitations on applied voltages due to power consumption, shifts in the V_{FB} value are undesirable and must be minimized. In some applications, biasing the substrate to compensate for the fixed charge has been proposed.[5] Moreover, a reproducible V_{FB} (correspondingly V_{T} for transistors) value is also required for stable, reliable transistor operation. Thus, hysteretic changes in the V_{FB} from voltage cycling of less than 20 mV are often required.

Some dielectrics which incorporate aluminum, however, thus far suggest that a negative fixed charge is present. It has been recently proposed to combine Al ions with some alternate dielectric candidates in order to compensate positive and negative fixed charges to achieve a neutral state or, at least, minimize such fixed charge effects.[11] If fixed charge is determined to be large and difficult to minimize and control in high-κ dielectrics, it will be a significant issue for obtaining the desired device performance on both *n*MOS and *p*MOS transistors. The magnitude of measured V_{FB} shifts for many alternate dielectrics will be discussed later.

IV. SCALING LIMITS FOR CURRENT GATE DIELECTRICS

The previous sections outlined the need to scale oxide thicknesses to improve performance. The next two sections describe the present understanding in the field regarding the limits of scaling current gate dielectric materials, SiO_2 and Si-oxide-nitride variations, for CMOS. Issues include band offset, interfacial structure, boron penetration and reliability. Beyond this scaling limit, another material will be required as the gate dielectric to allow further CMOS scaling.

A. Ultrathin SiO_2 properties

Experiments and modeling have been done on ultrathin SiO_2 films on Si, as a way to determine how the SiO_2 band gap or band offsets to Si change with decreasing film thickness.[12–15] In the study by Muller *et al.*,[12] electron energy loss spectroscopy (EELS) was carried out on 7–15 Å SiO_2 layers on Si. It was found that the density of states (as measured by the oxygen *K*-edge in EELS, with a probe resolution <2 Å) transition from the substrate into the SiO_2 layer indicated that the full band gap of SiO_2 is obtained after only about two monolayers of SiO_2. This indicates that within two monolayers of the Si channel interface, oxygen atoms do not have the full arrangement of oxygen neighbors and therefore cannot form the full band gap that exists within the ''bulk'' of the SiO_2 film.

An earlier *ab inito* model by Tang *et al.*[13] of extremely thin SiO_2, which was modeled as a modified beta-cristoballite phase, showed an important result, in that the band gap of SiO_2 did not begin to decrease until there were fewer than three monolayers of oxide. Moreover, estimates of the changes in the associated conduction and valance band offsets for these systems indicated that a minimum of 7 Å of SiO_2 is required to obtain bulk properties. The recent first principles study by Neaton *et al.*[14] determined that the local energy gap in SiO_2 is directly related to the number of O second nearest neighbors, for a given O atom. The *last row* of O atoms (next to the Si substrate) *by definition* cannot have the full six nearest neighbor O atoms. The second row of O atoms from the Si interface is thus the first layer of O atoms that have the required six second-nearest neighbor O atoms. The distance required to obtain the full band gap of SiO_2 at each interface is therefore given by 1.6 Å (the spacing of one Si–O bond length) +2.4 Å (the distance between neighboring O atoms is 2.7 Å, but this is variable because of Si–O bond bending. The distance is typically in the range ~2 to 2.4 Å). The thickness at each interface required for the full SiO_2 band gap is therefore ~3.5–4.0 Å. Counting both interfaces, the total thickness of 7–8 Å is required, in agreement with Tang *et al.*[13] and with the experiment.[12] These results set an *absolute* physical thickness limit of SiO_2 of 7 Å. Below this thickness, the Si-rich interfacial regions from the channel and polycrystalline Si gate interfaces used in MOSFETs overlap, causing an effective ''short'' through the dielectric, rendering it useless as an insulator.

The agreement between the experiment and simulation in these cases indicates that the inherent band gap of SiO_2 remains intact, even down to only a few monolayers of material. Other important properties of SiO_2 have been reported in the ultrathin, sub-20 Å regime, such as the conduction band offset ΔE_C to Si [using x-ray photoelectron spectroscopy (XPS)],[16] the tunneling electron effective mass m^* (from tunneling $I-V$ measurements),[17] and the photoelectron attenuation length.[18] These measurements have further demonstrated very little change in fundamental SiO_2 properties between bulk and ultrathin sub-20 Å films.

The apparent robust nature of SiO_2, coupled with industry's acquired knowledge of oxide process control, has helped the continued use of SiO_2 for the past several decades in CMOS technology. As experimental evidence of the excellent electrical properties of such ultrathin SiO_2 films, it has been demonstrated that transistors with gate oxides as thin as 13–15 Å continue to operate satisfactorily.[19–24] Although high leakage current densities of 1–10 A/cm^2 (at V_{DD}) are measured for such devices,[25] transistors intended for high-performance microprocessor applications can sustain these currents. As first reported by Timp *et al.*[20–22] scaling of CMOS structures with SiO_2 gate oxides thinner than about 10–12 Å results in no further gains in transistor drive current. This result has been subsequently and independently reported by other groups, thus 10–12 Å could serve as a practical limit for reducing the SiO_2 thickness.[23,24,26]

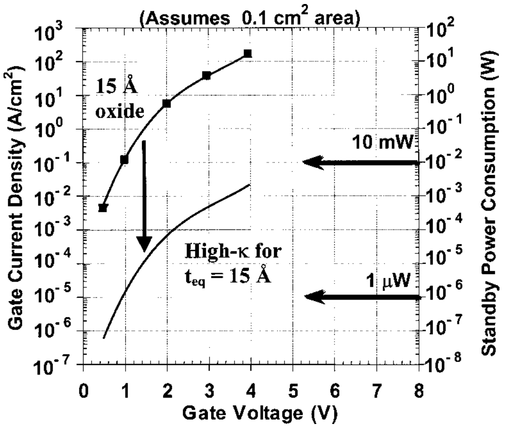

FIG. 5. Power consumption and gate leakage current density for a chip which has a 15 Å thick SiO_2 gate dielectric compared to the potential reduction in leakage current by an alternate dielectric exhibiting the same equivalent oxide thickness. Assumes at total gate area of 0.1 cm^2.

In contrast to the high performance microprocessor market, the rapidly growing market of low-power applications requires transistors with much lower ($\sim 10^{-3}$ A/cm^2) leakage currents.[5] This is illustrated in Fig. 5, where the current density and standby power consumption are plotted as a function of gate voltage. The curve for 15 Å oxide is based on measured values,[17] but the curve for a high-κ film is meant to show the potential reduction in leakage current for a high-κ dielectric with the same t_{eq} value. Depending on the specific materials and conditions, leakage current reduction may be less than shown. It is clear that a gate dielectric with a permittivity higher than that of SiO_2 is required to meet low-power application requirements.

B. Ultrathin SiO_2 reliability

An equally important issue regarding ultrathin SiO_2 gate oxides has been understanding and predicting oxide reliability. Considerable debate existed over whether SiO_2 gate oxides <22–27 Å thick would exhibit the stringent ten year reliability criteria,[27] which is required for CMOS devices. The first report of a sub-20 Å SiO_2 gate oxide to meet reliability requirements was given by Weir *et al.*,[25] where a 16 Å oxide was shown to have reliability projections at 1.6 V operation for greater than ten years. This is much thinner than projected even three to five years ago.[28,29] The high reliability of ultrathin oxides suggests that there is no *intrinsic* (i.e., not limited by intrinsic defects or thickness variations) reliability limitation to SiO_2 layers at least down to thicknesses of about 16 Å. In fact, more recent projections indicate that oxides down to 14 Å (as measured by ellipsometry) at 1.4 V operating voltage will meet ten year reliability requirements.[30] Several independent groups have also recently reached similarly encouraging reliability projections for such thin SiO_2 gate oxides.[26,31] Other *extrinsic* reliability factors, however, such as particles or contaminants, could still yield an ultimately poorer oxide reliability.

Part of the difficulty in making reliability projections arises from the difference between test conditions and operating conditions. It is clearly not feasible to test individual devices for ten years prior to product incorporation, thus testing must be ''accelerated'' at higher voltages and temperatures than are actually experienced by typical devices. Making reliability projections from accelerated to actual conditions requires proper scaling for area (from one device area to an entire chip area), voltage, temperature, and the failed fraction of devices.[27,32] Recent results by Stathis *et al.*,[31] Weir *et al.*,[25] and Nicollian *et al.*[33] show that extrapolations of reliability factors, such as the critical defect (trap) density, with voltage scaling changes dramatically at lower voltages, such as those used for testing oxides <20 Å thick. This realization of the change in voltage scaling behavior at low voltages is the largest factor contributing to the improved reliability projections described earlier. In addition, improved macroscopic oxide uniformity across the wafer and wafer-to-wafer has also been shown to give more accurate reliability projections.[25,34] This analysis should *not* be misinterpreted as meaning that the oxide reliability *itself* is improved. Rather, the reliability *projection* becomes more accurate (regardless of whether the projection is for good or poor oxide reliability) with higher macroscopic oxide uniformity.[35] Improving microscopic oxide uniformity should further produce more accurate reliability projections, according to simulation.[36] It has also been demonstrated that making reliability measurements on a large number of samples is important for obtaining better breakdown statistics and accurate projections.[34]

A fundamental mechanism for oxide breakdown in this ultrathin SiO_2 regime was first reported by DeGraeve *et al.*[37,38] as a percolation model. This model describes ultrathin oxide breakdown as the buildup of many ''defects'' within the SiO_2 layer, where after a certain amount of stress (either constant voltage or constant current through the oxide at a given temperature), a complete path of defects form across the oxide thickness.[37] This point defines breakdown or failure of the oxide. While there is general agreement on the percolation model for oxide breakdown, the defects which act as precursors to breakdown are not defined or specified. The mechanism which leads to the creation of these defects is under debate, and has been proposed as an anode hole injection model[36,39] and a hydrogen release model.[40,41]

It is also important to distinguish between previously reported leakage current projections by simulation for oxide thicknesses measured in accumulation, and the presently accepted methods.[25,31,33] Extrapolated leakage current versus gate oxide thickness data from three to five years ago was valuable at the time, when sufficient data for oxides <16 Å was not available. Caution should be used, however, when referring to such extrapolations now,[42] as data from more recent measurement methodology must be adopted for useful comparisons. Much of the understanding for ultrathin oxides have come about only in the past five years, despite decades of research on SiO_2. This suggests that understanding the reliability and failure mechanisms in high-κ dielectrics will require significant effort, especially if any material is to replace SiO_2 within five years, as most roadmaps suggest.

C. Boron penetration and surface preparation

In addition to leakage current increasing with scaled oxide thickness, the issue of boron penetration through the oxide is a significant concern. The large boron concentration gradient between the heavily doped poly-Si gate electrode, the undoped oxide and lightly doped Si channel causes boron to diffuse rapidly through a sub-20 Å oxide upon thermal annealing, which results in a higher concentration of boron in the channel region. A change in channel doping then causes a shift in threshold voltage, which clearly alters the intended device properties in an unacceptable way.[43] As will be discussed in the next section, incorporating nitrogen into the oxide can greatly reduce boron diffusion.

Some approaches to enhance performance have also focused on surface preparation as a way to provide a flatter, more uniform Si interface in attempts to minimize electron channel mobility degradation (due to scattering at the interface) and gate leakage.[44,45] Growing or depositing sub-15 Å oxides has also been investigated as a potential means for producing high-quality, reproducible and uniform gate oxides, either for use as the gate dielectric or as the bottom layer of a dielectric stack.[46,47] Obtaining high-quality oxides in this thickness regime is challenging, because intrinsic pinhole formation has been reported in ultrathin SiO_2 films.[48–50]

Perhaps the most significant benefit resulting from the incorporation of an alternative gate dielectric with a relative permittivity higher than SiO_2 is the dramatic reduction in the off-state tunneling (leakage) current which is observed in devices using ultrathin SiO_2 gate dielectric films. Thus, the area of alternative gate dielectrics has gained considerable attention recently because technology roadmaps predict the need for a sub-20 Å Si-oxide gate dielectric for sub-0.1 μm CMOS.[4,5]

D. SiO_xN_y and Si–N/SiO_2 dielectrics

The concerns regarding high leakage currents, boron penetration and reliability of ultrathin SiO_2 have led to materials structures such as oxynitrides and oxide/nitride stacks for near-term gate dielectric alternatives. These structures provide a slightly higher κ value than SiO_2 (pure Si_3N_4 has $\kappa \sim 7$) for reduced leakage (since the film is physically slightly thicker), reduced boron penetration and better reliability characteristics.[51–53] The addition of N to SiO_2 greatly reduces boron diffusion through the dielectric, and has been shown to result from the particular Si–O–N network bonding formed in silicon nitride and oxynitride.[54,55] Furthermore, small amounts of N ($\sim$0.1 at. %) at or near the Si channel interface have been shown to improve device performance.[56] Larger amounts of N near this interface degrade device performance, as discussed later.

The simplest approach is to use a pure nitride layer at or near the channel interface, but device performance is typically degraded in these structures. Recent work using remote plasma enhanced chemical vapor deposition (PECVD) to deposit Si-nitride directly on the Si channel[57] resulted in poor *p*MOS performance, with significant degradation of channel mobility and drive current. This degradation is attributed to several factors, including excess charge of pentavalent nitrogen atoms, a high defect density arising from bonding constraints imposed at the interface[58] (which causes increased channel carrier scattering), and from the defect levels in the Si-nitride layer which reside near the valence band of Si. In contrast, improved electrical properties have been obtained by using various oxynitrides. Yang and Lucovsky demonstrated that an oxynitride alloy with a 1:1 ratio of SiO_2:Si_xN_y can achieve $t_{eq} < 17$ Å with a leakage current of $\sim 10^{-3}$ A/cm^2 at 1.0 V bias.[56] This leakage current is $\sim 100\times$ lower than that for a pure SiO_2 layer of the same thickness, and the leakage reduction arises from both a physically thicker film and from a small amount of N at the channel interface. Song *et al.*[59] have also shown that a proper choice of thermal processing steps using a NO passivation layer, Si_xN_y deposition with SiH_4 and NH_3 and further anneals in NH_3 and N_2O can achieve oxynitride layers with $t_{eq} = 18$ Å and a leakage current density of 10^{-4} A/cm^2 at 1.0 V bias. Guo and Ma[60] have reported results with jet vapor deposited nitrides demonstrating $t_{eq} = 15$ Å with a leakage current density of $J \sim 10^{-4}$ A/cm^2 at 1.0 V.

Despite these encouraging results from a variety of deposition and growth techniques, scaling with oxynitrides/nitrides appears to be limited to $t_{eq} \sim 13$ Å.[61] Below this, the effects of gate leakage, reliability or electron channel mobility degradation will most likely prevent further improvements in device performance. According to the most recent industry roadmaps, SiO_xN_y and Si_xN_y/SiO_2 dielectrics represent current three year near-term solutions for scaling the CMOS transistor.[5]

E. Fundamental limitations

Despite the current efforts with SiO_2, oxynitrides and even high-κ gate dielectrics, several potential fundamental limitations could seriously threaten the continued scaling of all gate dielectrics, regardless of the material.[26] First, the electrical thickness of any dielectric is given by the distance between the centroids of charge in the gate and the substrate. This thickness, typically denoted by t_{eq}, therefore includes the effective thickness of the charge sheet in the gate and the inversion layer in the substrate (channel). These effects can add significantly to the expected t_{eq} derived from the physical thickness of the dielectric alone.[6] Depletion in the poly-Si gate electrode arises from the depletion of mobile charge carriers in the poly-Si near the gate dielectric interface, particularly in the gate bias polarity required to invert the channel. The result is often that $\sim$3–4 Å in the poly-Si electrode nearest to the gate dielectric interface essentially behaves like intrinsic Si, which adds $\sim$3–4 Å to the effective dielectric thickness (rather than acting as a metal with a Fermi sea of electrons right up to the dielectric interface). In the best case, the electrode depletion region can be reduced to $\sim$1–2 Å for degenerately doped poly-Si electrode right up to the interface, but this is difficult to obtain.

The nature of the inversion charge layer in the Si substrate (or channel, for transistors) contributes about 3–6 Å to the effective t_{eq} value, thus even for ideal, degenerately doped poly-Si gates, it is difficult to realize an overall $t_{eq} < 10$ Å in MOSFETs using current process technology. Metal gates offer a possible solution to the gate depletion problem, but the addition of 3–6 Å to the t_{eq} value from the Si channel will remain. We note that most of the high-κ dielectrics discussed later in this review are measured on capacitors and transistors with metal gates. Although metal gates are convenient for obtaining properties of the dielectric, current CMOS technology does not use metal gates. Therefore, for a given high-κ dielectric, any realistic device using current CMOS processing techniques should exhibit a t_{eq} value 4–8 Å larger than that reported in this review (except where poly-Si gates are noted as being used, and where no quantum mechanical correction has been performed).

Second, the ideal scaling scenario is one in which the operating voltage and transistor dimensions are reduced by the same factor, thus maintaining a constant electric field across the gate dielectric for a given technology node. In practice, however, the feature dimensions have been reduced more rapidly than the operating voltage, thereby causing a rapidly increasing electric field across the gate dielectric. The continually decreasing t_{eq} value for scaling CMOS also increases the effective electric field in the channel region. This increased electric field pulls the carriers in the channel closer against the dielectric interface, which causes increased phonon scattering of more confined carriers and thereby decreases the channel carrier mobility. At very high electric fields in the channel, such as would exist for $t_{\mathrm{eq}} < 10$ Å, interface roughness scattering further reduces carrier mobility. Thus, for $t_{\mathrm{eq}} < 10$ Å, the combination of these deleterious effects may result in not only reducing the expected performance increase for a given increase in gate capacitance, but may indeed even *decrease* device performance at fixed supply voltage. If this effect on channel mobility is in fact realized, then it may be the case that *no* dielectric will be acceptable in the required t_{eq} range (since this effect depends only on t_{eq}, not on the material).

F. Device structures

In addition to ongoing work in scaling gate dielectrics, there is also a substantial amount of research being conducted toward obtaining a device structure for CMOS, such that the demand for scaling SiO_2 or oxynitrides (and lithography) will be somewhat alleviated. In particular, efforts have been focused on structures such as vertical transistors,[62–65] and double gate planar transistors.[66–68] The premise of double gate transistors arises from the potential of achieving nearly twice the drive current over a planar, bulk CMOS device for a given channel length. Alternatively, a nearly equal drive current could be produced while using a SiO_2 layer that is twice the thickness as that required in the planar geometry (and thereby meet roadmap performance requirements with a much thicker gate oxide). This potential gain from a double gate device is explained simply by the fact that there are two parallel, aligned channels operating simultaneously, compared to one channel in a standard, planar bulk device.

One particularly promising vertical transistor uses a vertical replacement gate (VRG) structure,[62] where standard CMOS processes are used in an innovative way to remove the lithography constraint for defining the gate length. Instead, a deposited oxide layer thickness defines the gate length, which can thus be easily controlled below 50 nm. This dummy oxide layer can then be etched away, followed by growth and deposition of the gate oxide and poly-Si gate electrode, respectively. Figure 6 shows a cross-sectional transmission electron microscopy (TEM) image of a VRG device with parallel 50 nm gate lengths and 25 Å gate oxides (as measured by ellipsometry) and a single crystal, epitaxially grown Si body.[62] As an example of the potential for increased performance with thicker gate oxides, VRG transistors with 100 nm gate lengths, 25 Å gate oxides, and operating voltage 1.5 V have shown an 80% increase in drive current[63] compared to that required by the SIA Roadmap (at the same off current) using much thinner, sub-20 A oxides. High-performance planar devices reported to date have shown only about 20% increase in drive current compared to that of the roadmap, under comparable operating conditions.[23,24] For 50 nm gate length devices and 1.0 V operation, VRG transistors have been demonstrated to meet the requirements of the roadmap, while *still* using 25 Å gate oxides.[63] In contrast, *no* reported planar CMOS devices to date have been able to even approach the drive current and off-current roadmap requirements for 50 nm gate length, 1.0 V operation.

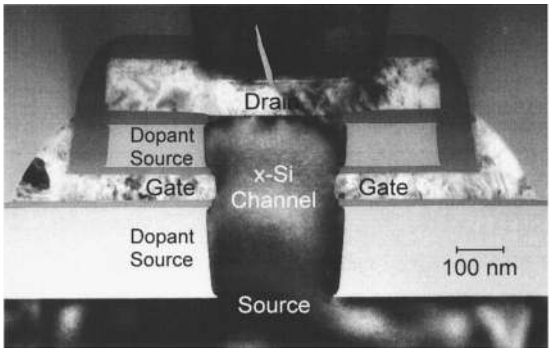

FIG. 6. Cross-sectional TEM image of a vertical replacement gate transistor, with a 50 nm gate length and 25 ÅSiO_2 gate oxide (as measured by ellipsometry) (see Ref. 63). © 2000 IEEE, reprinted with permission from IEEE.

These device approaches to stave off CMOS scaling will continue in parallel with investigations into high-κ gate dielectrics. One of these structures may ultimately prove to be a viable replacement for the planar, bulk CMOS transistor and retard the need to scale SiO_2 gate oxides. Most device structures under consideration therefore will not immediately suffer from the limitations discussed in the previous section, because a thicker gate oxide *reduces* the effects of dopant depletion in the poly-Si gate electrode, as well as reduces the electric field present in the channel region. Since these deleterious effects become amplified as the gate oxide thickness decreases, even device structures would eventually suffer from these degradation mechanisms, but in the meantime would afford more time to find other solutions to these problems. In this case, high-κ dielectrics can be combined with these device structures to further improve performance and power consumption.

V. ALTERNATIVE HIGH-κ GATE DIELECTRICS

As an alternative to oxide/nitride systems, much work has been done on high-κ metal oxides as a means to provide a substantially thicker (physical thickness) dielectric for reduced leakage and improved gate capacitance. In the search to find suitable high-κ gate dielectrics for use beyond oxynitride systems, several approaches have been used in fabricating potential materials candidates. The following sections review the current status of work in this field. We also describe the materials properties considerations that are necessary for determining the best high-κ candidate to replace SiO_2 as the gate dielectric for CMOS. It is important to note that we do not compile all of the relevant, measured electrical and physical characteristics on high-κ materials into one table. Although a table format can be helpful for organizational purposes, at the present time for high-κ gate dielectrics, it is too cumbersome (and potentially misleading) to attempt to provide all of the proper caveats (e.g., leakage current values for the same material system can vary widely depending on surface preparation, deposition method and conditions, gate electrode type, bias voltage ramp rate, stated bias for a given current ($J@V_{\text{bias}}=1$ V, or $V_{\text{FB}}+1$ V, etc.).

A. High-κ candidates from memory applications

Many of the materials initially chosen as potential alternative gate dielectric candidates were inspired by memory capacitor applications[42] and the resultant semiconductor manufacturing tool development infrastructure.

The most commonly studied high-κ gate dielectric candidates have been materials systems such as Ta_2O_5,[68–78] $SrTiO_3$,[79–84] and Al_2O_3,[85–95] which have dielectric constants ranging from 10 to 80, and have been employed mainly due to their maturity in memory capacitor applications. With the exception of Al_2O_3, however, these materials are not thermodynamically stable in direct contact with Si (this thermodynamic stability is not a requirement for memory capacitors, since the dielectric is in contact with the electrodes, which are typically nitrided poly-Si or metal). An excellent and thorough review on the Ta_2O_5 system, for both memory capacitor and transistor applications, has been given by Chaneliere *et al.*[77] The Ta_2O_5 system is known to exhibit Frenkel–Poole and Schottky transport mechanisms, depending on bias polarity, under typical voltage bias conditions. The mechanisms for relaxation or transient current of Ta_2O_5 have recently been attributed to a defect band near the conduction band in thin Ta_2O_5 films, which allows ac transient conduction leakage that follows a widely observed power law decay.[78]

Interfacial reaction, seen in Fig. 7, has been observed for the case of Ta_2O_5 on Si,[75] as is expected based on thermodynamic arguments discussed later and in agreement with previous work in dynamic random-access memory capacitor applications.[76] As evidenced by these studies, the earlier-mentioned metal oxides *require* that both the gate electrode and the channel interfaces be modified to limit the amount of reaction.

Interface engineering schemes have been developed to form oxynitrides and oxide/nitride reaction barriers between these high-κ metal oxide materials and Si in an attempt to prevent or at least minimize reaction with the underlying Si.[69–74] The passivating properties of such reaction barriers is widely reported.[94] In most cases, this amounts to further scaling the approaches used to form oxynitrides, discussed in the previous section. These barrier layers have been shown to reduce the extent of reaction between the high-κ dielectric and Si, as well as to help maintain a high channel carrier mobility.

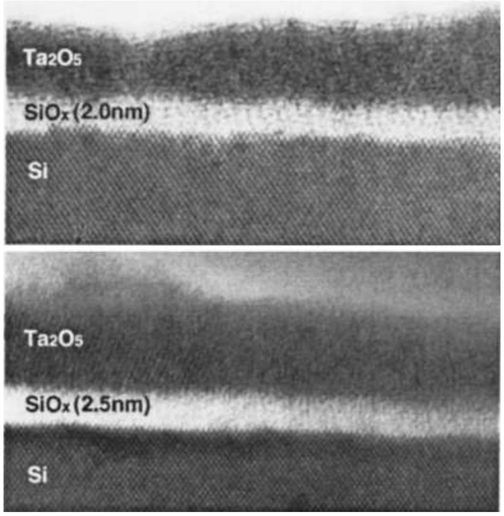

FIG. 7. Reaction at the Ta_2O_5/Si interface is observed resulting in the formation of a thin SiO_2 layer (see Ref. 75).

It is important to note, however, that using an interfacial layer of SiO_2 or another low permittivity material, will limit the highest possible gate stack capacitance, or equivalently, the lowest achievable t_{eq} value [see Eq. (4)]. In addition, the increased process complexity for the deposition and control of additional ultrathin dielectric layers, as well as scalability to later technology nodes, remains a concern.

This effect of reduced capacitance can be seen by noting that when the structure contains several dielectrics in series, the lowest capacitance layer will dominate the overall capacitance and also set a limit on the minimum achievable t_{eq} value. For example, the total capacitance of two dielectrics in series is given by

$$1/C_{\text{tot}}=1/C_1+1/C_2, \tag{10}$$

where C_1 and C_2 are the capacitances of the two layers, respectively. If one considers a dielectric stack structure such that the bottom layer (layer 1) of the stack is SiO_2, and the top layer (layer 2) is the high-κ alternative gate dielectric, Eq. (4) is simplified (assuming equal areas) to

$$t_{\text{eq}}=t_{SiO_2}+(\kappa_{\text{ox}}/\kappa_{\text{high-}\kappa})t_{\text{high-}\kappa}. \tag{11}$$

It is clear from Eq. (11) that the minimum achievable equivalent oxide thickness [defined as t_{eq} in Eq. (11)] will *never* be less than that of the lower-κ (in this case pure SiO_2) layer. Therefore, much of the expected increase in the gate capacitance associated with the high-κ dielectric is compromised. The implications of current transport through such stacked structures will be considered further later.

The largest benefit of using SiO_2 as the underlayer of a stack (at the Si channel interface) is that the unparalleled quality of the SiO_2–Si interface will help maintain a high channel carrier mobility. The prospect of using such SiO_2 interface layers was examined by Kizilyalli and Roy.[70,71] In that work, a Ta_2O_5 film was sandwiched between SiO_2 layers located at the substrate and gate (poly-Si) interfaces. The SiO_2/Si substrate interface layer was formed by a low pres-

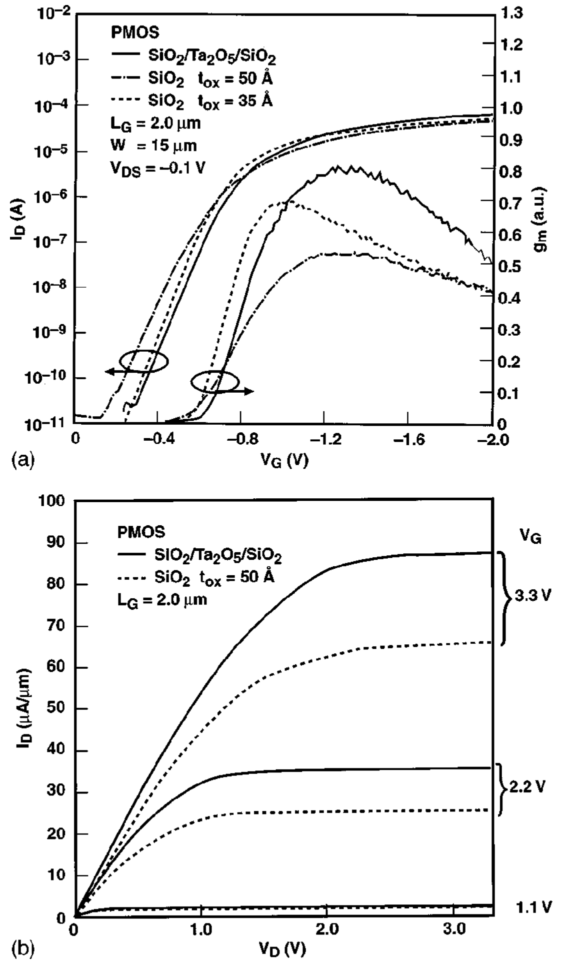

FIG. 8. (a) Drain current and transconductance for long channel PMOS transistors incorporating a $SiO_2/Ta_2O_5/SiO_2$ dielectric stack (see Ref. 70). (b) Comparison of drain current drive and saturation characteristics.

sure (850 mTorr) thermal anneal at 850 °C in O_2, resulting in a ~10-Å-thick oxide layer. A Ta_2O_5 layer was then deposited by (CVD), followed by a 10 Å capping layer of CVD SiO_2. The entire structure was then reoxidized in O_2 or N_2O at 650 °C and 850 mTorr. Given the stacked structure of the dielectric, the total capacitance of the stack will be diminished by the presence of the two lower-permittivity layers.

For the Ta_2O_5 sandwich,[70,71] using Eq. (11) for the Ta_2O_5 layer ($t \approx 50-60$ Å, $\kappa \approx 20-30$), an equivalent oxide thickness $t_{eq} = 10 + 10 + (3.9/\kappa)t = 25-30$ Å is expected. The experimentally determined value from capacitance–voltage ($C-V$) measurements was in good agreement with this estimate: 23 Å. No charge trapping was observed in the $C-V$ experiments, as evidenced by a lack of hysteresis. This study is an important attempt to demonstrate the feasibility of integrating Ta_2O_5 into a standard CMOS process, but it is clear that the presence of a SiO_2 layer (or any low-κ layer) at either interface limits the ultimate device performance.

Long channel transistor measurements utilizing this stack were also made using poly-Si gates and WSi_x contacts, as shown in Fig. 8. Figure 8(a) shows drain current and transconductance for long-channel PMOS transistors as a function of gate voltage, and Fig. 8(b) shows a comparison of drain current drive and saturation characteristics. Dopant activation anneals were accomplished with a relatively low temperature (600–800 °C) rapid thermal anneal process, presumably to avoid crystallization of the Ta_2O_5 layer and to prevent further SiO_2 formation at the interfaces within the constraints of a conventional CMOS process flow.[70,71] An interface state density (D_{it}) comparable to transistors incorporating only SiO_2 as the gate dielectric was observed based on subthreshold slope measurements. The effective dielectric thickness estimate from the transistor transconductance and drain current measurements results in a larger value, ~30 Å. The 20% discrepancy with the MIS capacitor $C-V$ measurements was attributed to partial dopant activation (from the relatively low temperature anneal process). Some degradation in mobility, relative to SiO_2, was also measured. As expected, leakage currents were well below 1 A/cm^2 as a result of the physically thicker dielectric.

It is essential at this point to distinguish between the requirements for memory[42] and transistor applications. Memory capacitors require extremely low leakage currents (typically $J < 10^{-8}$ A/cm^2) and very high capacitance density for charge storage, but the interface quality is not as critical. Memory capacitor applications require control of the interface primarily to limit interfacial reactions to keep the total capacitance high. Since the main requirement is that the capacitors store charge, however, current transport along the dielectric interface is not important. Furthermore, no electric field penetration is required below the bottom electrode, so the bottom electrode is often metal, or nitrided poly-Si (heavily doped). All of the requirements amount to the important distinction that the bottom dielectric interface quality is not as critical to capacitor performance.

In contrast, a key requirement of a FET is that the electric field penetrate into the Si channel to modulate carrier transport, and that the dielectric-channel interface be of very high quality. The channel must of course be Si, so any potential high-κ dielectric must be compatible with Si. Transistors have more lenient leakage requirements ($< 10^2$ A/cm^2 for high-performance processors, and $\sim 10^{-3}$ A/cm^2 for low-power applications), although high capacitance densities are still needed. The most critical distinction between high-κ materials requirements for capacitors versus gate dielectrics is the interface and materials compatibility: gate dielectrics *must* form an extremely high-quality interface with Si, and also be able to withstand CMOS processing conditions while in contact (or near contact) with Si.

B. Issues for interface engineering

It is well known that the industry roadmap presents a major problem for the core transistor gate dielectric, because it calls for a much thinner effective thickness for future alternative gate dielectrics: $t_{eq} \leqslant 10$ Å.[4,5] This would then require the SiO_2 layer to be ~5 Å thick. Such an extremely thin SiO_2 layer is very difficult to obtain with high quality.[22] The resulting voltage drop across the oxide could also lead to significant charge trapping in the film, especially since the interface between such stacked dielectrics may almost certainly contain a large density of traps. Furthermore, such a thin interface layer most likely will *not* prevent reaction between the substrate and any high-κ material which is not

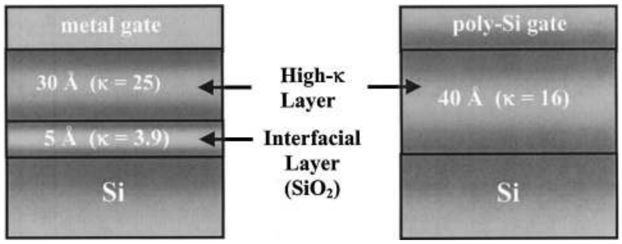

FIG. 9. Comparison of stacked and single layer gate dielectrics in a hypothetical transistor gate stack. Either structure results in the same overall gate stack capacitance or equivalent oxide thickness, $t_{\mathrm{eq}}=10$ Å.

thermodynamically stable to SiO_2 formation on Si, under standard thermal processing required for CMOS (previously discussed pinhole formation in ultrathin SiO_2 is also a concern).[48]

To illustrate this point, an example for obtaining a dielectric stack with $t_{\mathrm{eq}}=10$ Å is considered in Fig. 9. One way to achieve this would be to use 5 Å of $SiO_2(t_{\mathrm{eq}}=5$ Å) as the lower (first) layer, at the Si interface, and 30 Å of a dielectric with $\kappa=25$ ($t_{\mathrm{eq,high}\text{-}\kappa}\sim5$ Å) as the upper (second) layer. Even for low applied voltages, such a thin layer will have a large enough electric field to create a significant amount of charge trapping. In addition, an oxide layer this thin will allow a large amount of direct electron tunneling into the high-κ dielectric, likely causing further deleterious effects to the electrical performance of the stack.

It is important to note, however, that if a *single* layer dielectric can be used, then $t_{\mathrm{eq}}=10$ Å can be achieved with 40 Å (physical thickness) of a material which only has a moderate permittivity of $\kappa=16$ (see Fig. 9). This physical thickness is greater than the total physical thickness of the stack in the earlier example (=35 Å), even though the permittivity of the single layer gate dielectric ($\kappa=16$) is *much lower* than that of the alternate dielectric in the stack ($\kappa=25$). In addition, any potential charge trapping at a dielectric-dielectric interface would be avoided. These considerations for the choice of the best high-κ materials will be covered in more detail later.

The approach of using an epitaxial high-κ gate dielectric, such as $SrTiO_3$, requires submonolayer control of the channel interface for dielectric deposition.[79–84] As with the use of all perovskites for dynamic random access memory (DRAM) applications (e.g., $Ba_xSr_{1-x}TiO_3$,$Pb_xZr_{1-x}TiO_3$, etc.), the dielectric must be crystalline (usually polycrystalline) to obtain the enormous permittivities typically observed ($\kappa>300$). The work done thus far on gate dielectrics therefore has required molecular beam epitaxy (MBE), to obtain interface control and layer-by-layer deposition. Since it is difficult to attain a crystalline oxide on Si, interface engineering has been employed to provide submonolayer deposition of several initial ''template'' Sr–Si–O layers.[79,82,84] This interface helps reduce reaction due to the thermodynamic instability of $SrTiO_3$ on Si, and also helps accommodate the difference in lattice constants between Si and $SrTiO_3$. Transistors showing encouraging results have been fabricated using $SrTiO_3$ gate dielectrics, with metal gates (to prevent reaction between $SrTiO_3$ and poly-Si gates) and modified device processing.[80,82]

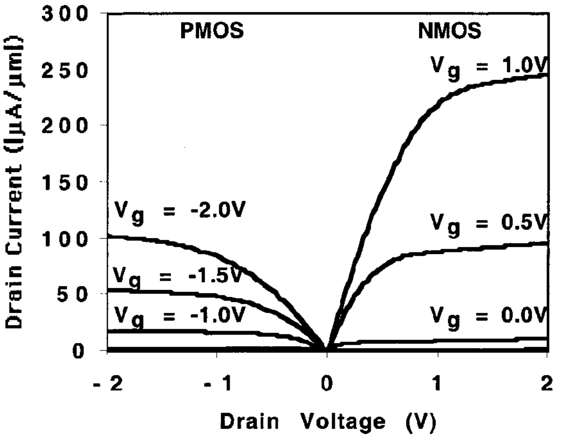

FIG. 10. Transistor results for effective gate length $L_{\mathrm{eff}}=1.3$ μm which incorporate a high-κ $SrTiO_3$ gate dielectric (see Refs. 82–84).

In the study by Eisenbeiser *et al.*, TaN gate electrodes were used, and devices exhibited $t_{\mathrm{eq}}\approx8$ Å (with quantum mechanical correction to the $C-V$ data) with a leakage current density $J\sim10^{-2}$ A/cm^2 at 1 V bias, for a 110 Å physically thick $SrTiO_3$ film.[82] Transistors with a 1.2 μm effective channel length showed electron and hole mobilities of 221 and 62 cm^2/V s, respectively. Device characteristics are shown in Fig. 10, with a subthreshold slope of 103 mV/decade for *n*-channel devices, 95 mV/decade for *p*-channel devices. The extracted interface state density showed a low $D_{\mathrm{it}}\sim6\times10^{10}$/cm^2, but the fixed, low work function ($\Phi_B=4.2$ eV) of the TaN gates produced undesirable threshold voltages for *p*-channel devices.[82] Figure 11 shows that the dielectric deposition process can cause interfacial reaction to occur with Si, resulting in an amorphous SiO_x-containing layer.[84] This layer most likely provides a better interface to the Si channel than the $SrTiO_3$ would, thereby exhibiting encouraging device properties. This MBE approach and the implications of the UHV conditions required for MBE will be discussed further in Sec. VI F.

C. Recent high-κ results

Considering the potential problems and limitations in using a ~5 Å SiO_2 layer in a dielectric stack, it is highly desirable to employ an advanced gate dielectric which is stable on Si, *and* exhibits an interface quality to Si which is comparable to that of SiO_2. This would avoid the need for an interfacial layer and at the same time, the high permittivity of the material could be fully realized. Table I is a compilation of several potential high-κ dielectric candidates, with the columns indicating the most relevant properties, which will be discussed in detail throughout the remainder of this review. Although a substantial amount of work has been reported on Ta_2O_5 as a gate dielectric,[77] and it clearly has many attributes for memory capacitor applications, the inherent thermal instability when in direct contact with Si is a severe limitation as a gate dielectric. We now examine the available results from the literature on such metal oxide and pseudobinary systems, and have categorized them for chemical similarities by group in the periodic table.

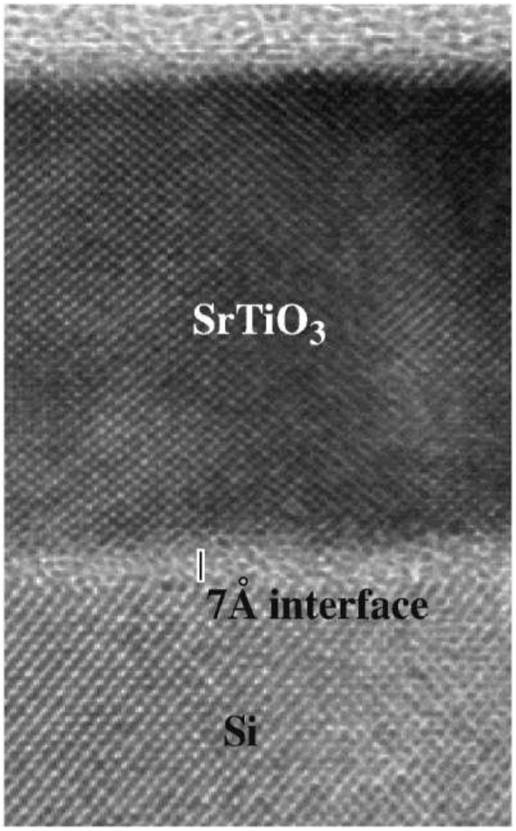

FIG. 11. Reaction at the interface of the $SrTiO_3$ gate dielectric and the Si interface (see Ref. 82).

1. *Group IIIA and IIIB metal oxides*

Among the group III candidate dielectrics, alumina (Al_2O_3) is a very stable and robust material, and has been extensively studied for many applications. Regarding its usefulness as an alternate gate dielectric, Al_2O_3 has many favorable properties, as shown in Table I, including a high band gap, thermodynamic stability on Si up to high temperatures, and is amorphous under the conditions of interest. The drawback is that Al_2O_3 only has $\kappa \sim 8-10$, and would therefore make it a relatively short-term solution for industry's needs (1–2 generations). If no longer-term solution is available by the time that a replacement is required, however, such a short-term solution may indeed be suitable. The otherwise desirable attributes of Al_2O_3 have resulted in several recent studies of both its physical and electrical properties as a gate dielectric.

TABLE I. Comparison of relevant properties for high-κ candidates.

Material	Dielectric constant (κ)	Band gap E_G (eV)	ΔE_C (eV) to Si	Crystal structure(s)
SiO_2	3.9	8.9	3.2	Amorphous
Si_3N_4	7	5.1	2	Amorphous
Al_2O_3	9	8.7	2.8[a]	Amorphous
Y_2O_3	15	5.6	2.3[a]	Cubic
La_2O_3	30	4.3	2.3[a]	Hexagonal, cubic
Ta_2O_5	26	4.5	1–1.5	Orthorhombic
TiO_2	80	3.5	1.2	Tetrag.[c] (rutile, anatase)
HfO_2	25	5.7	1.5[a]	Mono.[b], tetrag.[c], cubic
ZrO_2	25	7.8	1.4[a]	Mono.[b], tetrag.[c], cubic

[a] Calculated by Robertson (See Ref. 153).

[b] Mono.=monoclinic.

[c] Tetrag.=tetragonal.

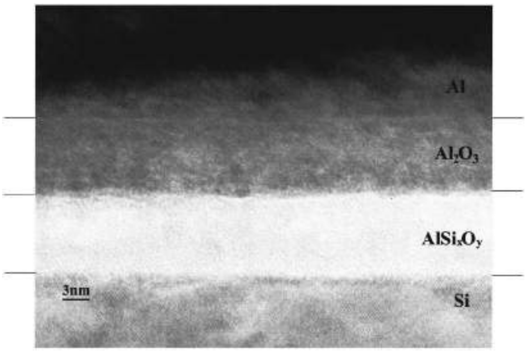

FIG. 12. HRTEM image of CVD Al_2O_3 on Si with an interfacial Al-silicate reaction layer evident (see Ref. 85).

As mentioned before, Al_2O_3 is thermodynamically stable on Si against SiO_2 formation, but any deposition technique of interest typically operates under nonequilibrium conditions. This means that reactions can occur, and phases can form which are not predicted by equilibrium phase diagrams, although further thermal processing (depending on the temperatures and times involved) will tend to drive the system toward the equilibrium state. Klein *et al.*[85] studied thin Al_2O_3 films deposited by CVD (using triethyldialuminum for the Al precursor) at temperatures below 400 °C, and found evidence of an aluminum silicate phase formed at the interface with Si, as seen in Fig. 12. Nuclear resonance profiling (NRP) (which measures the Al concentration) of the films showed a two-layered structure, with the expected Al concentration for Al_2O_3 in the top layer, and a marked decrease in Al content in the interface layer. XPS indicated the presence of Al–O–Si bonds in an interface layer for a 35 Å deposited Al_2O_3 film. A significant amount of carbon was also observed in the films (~20 at. %), but it is nevertheless apparent that kinetics can play an important role in determining the interface structure, which is the most critical region for thin gate dielectrics.

Atomic layer CVD (ALCVD) Al_2O_3 has been studied by Gusev *et al.*[87] both physically and electrically, in particular to better understand the interface formed between Si and Al_2O_3 deposited by this technique. Using NRP, medium energy ion scattering (MEIS) and high-resolution TEM, it was determined that ALCVD-deposited Al_2O_3 (using trimethylaluminum and water as the Al and O precursors, respectively) could be deposited on H-terminated Si *without* forming an interfacial SiO_2 layer, as shown in Fig. 13.[87] This is an important result, because even though Al_2O_3 is thermodynamically stable on Si (as mentioned previously), all of the deposition and growth techniques discussed here occur under nonequilibrium conditions. It is therefore usually found that an interfacial SiO_2-containing layer forms during deposition, between the high-κ material (in this case Al_2O_3) and the Si substrate. The combination of the three physical analyses used in the study by Gusev *et al.*[87] show that it is possible to control the interface reactions, at least for this case of precursors and deposition conditions.

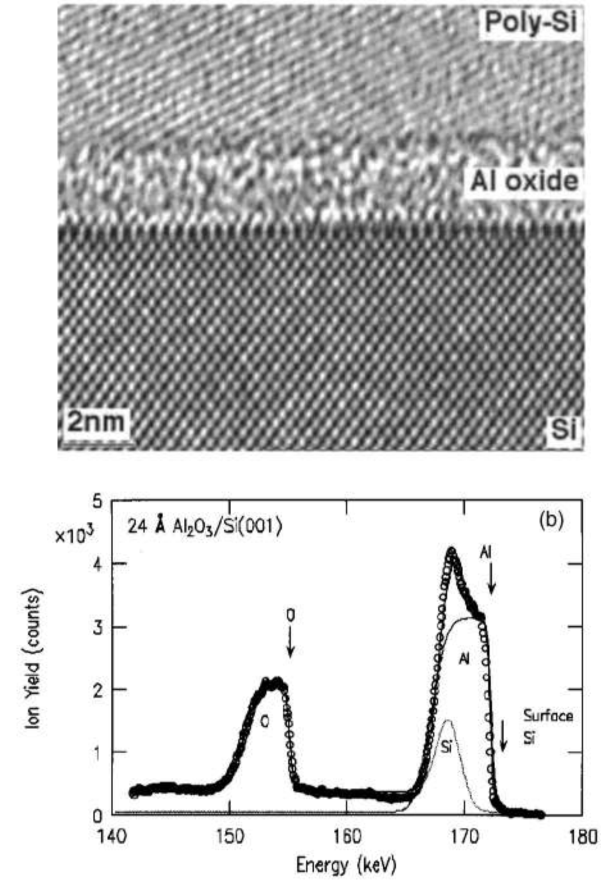

FIG. 13. (a) TEM image of Al_2O_3 film deposited by ALCVD methods on a hydrogen passivated surface from a HF-last process. (b) Corresponding MEIS profile of the film. No interfacial layer is detected (see Ref. 87). © 2000 IEEE, Fig. 13(a) reprinted with permission from IEEE.

Transistor results for 48 Å of Al_2O_3 (thermally evaporated Al followed by thermal oxidation) were reported by Chin *et al.*,[88] which exhibited $t_{eq} = 21$ Å with a leakage current of $\sim 10^{-8}$ A/cm^2 at 1 V gate bias, compared to $\sim 10^{-1}$ A/cm^2 for 21 Å of pure SiO_2. The Al_2O_3 films exhibited low stress-induced leakage current (SILC) effects, but did show a high interface state density, $D_{it} > 10^{11}\,\mathrm{eV^{-1}\,cm^{-2}}$. A follow-up study by Chin *et al.*[89] achieved a thinner physical thickness of 21 Å Al_2O_3 to produce $T_{eq} = 9.6$ Å, with 22 mV of hysteresis and $D_{it} \geqslant 3 \times 10^{10}\,\mathrm{cm^{-2}}$, and a flatband shift $\Delta V_{FB} \sim +600$ mV, suggesting negative fixed charge in the film. Very good transistor properties were reported for these films, again with low SILC effects.

Buchanan *et al.*[91] reported *n*MOSFET results on ALCVD Al_2O_3 for 0.08 μm gate length transistors with poly-Si gates, using standard processing conditions, including a rapid-thermal dopant activation anneal at $T > 1000\,°C$. A leakage current of $J \sim 10^{-1}$ A/cm^2 (at $V_{bias} = V_{FB} + 1$ V) was measured for $t_{eq} = 13$ Å, showing a reduction in leakage current of $\sim 100\times$ compared to SiO_2 of the same t_{eq} value. A trend in fixed charge correlated with Al_2O_3 thickness was demonstrated, showing that fixed charge increases with decreasing film thickness. Extrapolation of that data indicated that the fixed charge is concentrated near the top, poly-Si interface. Furthermore, an interfacial layer was determined to contribute ~ 8 Å to the overall the overall t_{eq} value. The composition of the interface reaction layer is currently not known, but apparently has a κ value larger than that of pure SiO_2. At an effective field of 1 MV/cm, the channel carrier mobility value for Al_2O_3 was measured to be smaller by a factor of $\sim 2\times$ (100 cm^2/V s compared to 220 cm^2/V s for the universal mobility curve) than that expected by the universal mobility curve.[91] Encouraging drive currents and reliability characteristics were demonstrated for these devices, but the significant mobility degradation clearly indicates some deleterious effects of the ALCVD Al_2O_3 which warrant further investigation.

The study by Park *et al.*[90] demonstrated that boron diffuses through ALCVD Al_2O_3 during dopant activation anneals, and indeed this may be a serious issue for any alternative dielectric. It was reported that dopant activation anneals of 800–900 °C performed on boron implanted poly-Si gates on top of ~ 60 Å Al_2O_3 caused significant diffusion of boron through the Al_2O_3 film and into the *n*-Si substrate, as evidenced by a flatband shift of ~ 1.5 V. Secondary ion mass spectroscopy profiles also indicated a significant amount of boron in the substrate after anneal. Furthermore, the addition of an oxynitride layer, grown by an N_2O anneal before Al_2O_3 deposition, greatly reduced the flatband shift to 90 mV. In a different study by Lee *et al.*,[92] phosphorous diffusion from the n^+ poly-Si electrode into ALCVD Al_2O_3 was observed under reasonable annealing conditions of 850 °C for 30 min. $C-V$ analysis showed a flatband shift $\Delta V_{FB} = 670-740$ mV (depending on the particular dopant incorporation process), which corresponds to $> 10^{12}\,\mathrm{cm^{-2}}$ of negative fixed charge in the film. These results indicate that in this case, phosphorous not only diffused through the Al_2O_3 layer, but also introduced fixed charge into dielectric. The authors propose that phosphorous modifies the Al_2O_3 network, causing negatively charged Al–O dangling bonds. It will continue to be extremely important to identify and understand dopant diffusion in any potential alternative gate dielectric.

Most of the high-κ films thus far exhibit a flatband voltage different from that expected for the given choice of electrode and substrate type. As can be seen in Fig. 4, the flatband voltage is ideally determined *only* by the electrode work function and the electron affinity of the substrate. We therefore very roughly estimate flatband shifts from the data reported in the literature. (Note that this shift is different from the hysteresis flatband shift, which is required to be at least an order of magnitude smaller, and arises from sweeping the $C-V$ curve in opposite polarity directions.) In most studies, ΔV_{FB} values are not reported by the authors. We estimate the value of ΔV_{FB} from published $C-V$ curves, however, using work function values reported by Michaelson[95] for a given electrode, along with electron affinity values of 4.18 and 5.3 eV for *n*-Si and *p*-Si substrates, respectively. The ΔV_{FB} values we estimate are not intended to be extremely accurate, but rather to show an approximate value, within a few hundred millivolts.

In the studies mentioned above for Al_2O_3 on Si, the measured flatband voltage shift ΔV_{FB} is about +300 to +800 mV, compared to that expected by the electrode and substrate types used. This shift is typically interpreted as fixed charge within the film, although it can also arise from oxide damage associated with gate electrode deposition or other forms of processing treatments. As mentioned previously, a positive ΔV_{FB} value corresponds to a negative fixed charge. Considering that large ΔV_{FB} values have been measured by several independent groups, using different processing conditions and electrodes, this is currently being interpreted as fixed charge within the films.

Several groups have studied the group IIIB metal oxides Y_2O_3,[96–100] La_2O_3[89,99] and Pr_2O_3[101] for the purposes of high-κ gate dielectrics. Manchanda and Gurvitch[96,97] thermally oxidized sputtered yttrium films to form Y_2O_3, for structures with and without an intentional SiO_2 layer between the Y_2O_3 and Si substrate. It was found that structures with ~260 Å Y_2O_3 showed very low leakage of $<10^{-10}\,A/cm^2$ at 5 V bias and breakdown fields of E_{BD} ~4 MV/cm. Capacitors accumulated well with little hysteresis and dispersion, but showed an interface charge density of $\sim 6\times 10^{11}\,cm^{-2}$ and showed an interface trap density of $\sim 10^{11}\,cm^{-2}\,eV^{-1}$. The dielectric constant of the Y_2O_3 grown on SiO_2 was found to be $\kappa \sim 17-20$, but for Y_2O_3 grown directly on Si, it was found that $\kappa \sim 12$. This lower measured permittivity value likely resulted from growth of interfacial SiO_2 during the thermal oxidation step.

Kwo *et al.* investigated Y_2O_3 and Gd_2O_3 films deposited by molecular beam epitaxy (see Secs. VI D and VI F), in both crystalline and amorphous phases (as measured by x-ray diffraction).[98] Amorphous films showed lower leakage current densities than crystalline layers, and capacitors showed $t_{eq}=10-15$ Å with leakage current densities of $J=10^{-6}-10^{-3}\,A/cm^2$, depending on morphology, deposition and postannealing conditions. Breakdown fields were measured as $E_{BD}\sim 3$ MV/cm, with interface state density $D_{it}<10^{12}\,cm^{-2}\,eV^{-1}$. The $C-V$ curves for these films also exhibited frequency dependence between 100 Hz and 1 MHz.

Guha *et al.* recently investigated relatively thick (t_{eq} ~45 Å) Y_2O_3 films using an atomic O and Y metal beam epitaxy approach.[99] It was found that quasistatic and high frequency (100 kHz) $C-V$ measurements using Al electrodes overlap well upon postmetallization annealing suggesting that interface state densities less than $10^{11}\,cm^{-2}$ were obtained. Little or no flatband voltage shift was reported. Leakage currents at 1 V in thinner films (t_{eq}~21 Å) were reported to be $\sim 10^{-8}\,A/cm^2$. High resolution TEM and MEIS studies, however, indicate the formation of a columnar polycrystalline structure with ~10% Si in the film and a 15-Å-thick interfacial SiO_x layer which was attributed to oxygen diffusion (possibly along grain boundaries) and/or catalytic reaction.

Chambers and Parsons[100] recently investigated the interface associated with Y_2O_3, Si and oxidized/nitrides Si surfaces. They found that silicides can readily form depending on the kinetics of the deposition process, and that subsequent oxidation of the silicide yields an interfacial silicate layer. They also point out that such reaction kinetics may well be expected for other metal–oxide systems used for alternate gate dielectrics. The flatband voltages for the Y_2O_3 films reported to this point are shifted by $\Delta V_{FB}=300-600$ mV from that expected for an ideal capacitor using the respective electrodes and substrate types in these studies. As with Al_2O_3, these are very large voltage shifts measured in a broad array of sample conditions.

Chin *et al.*[89] reported results from La_2O_3 films which were formed by evaporation of La onto Si, followed by low-temperature thermal oxidation and *ex situ* Al gate deposition. The actual composition and structure in the films may require further examination, as pure La is well known to be very volatile and reactive in air, and La_2O_3 will absorb water vapor from air. These characteristics signify that any *ex situ* exposure of these films to air will certainly result in an uncontrolled reaction. Nevertheless, remarkable device results were shown for these films, as a physical thickness of 33 Å La_2O_3 produced $t_{eq}=4.8$ Å, $J\sim 10^{-1}\,A/cm^2$ at 1.0 V gate bias, and $D_{it}\sim 3\times 10^{10}\,eV^{-1}\,cm^{-2}$ (almost no quantum mechanical effect on the t_{eq} value because of the Al gate). Long-channel transistors with the same films for the gate dielectric exhibited very good turn-on characteristics, a subthreshhold slope of 75 mV/decade, and a high field effective mobility (at 1 MV/cm) of $\mu_{eff}>300\,cm^2/V\,s$. These films also showed a flat band voltage shift of ~+700 mV.

Guha *et al.* also recently investigated thin (t_{eq} ~10–14 Å) La_2O_3 films[99] where leakage currents of $10^{-4}-10^{-7}\,A/cm^2$ at 1 V are reported. An amorphous film structure with ~10% Si and a thin interfacial SiO_x layer was reported. A large flat band voltage shift of −1.4 V was also observed.

There are clearly many attributes to the group III metal oxides, as many promising and encouraging properties have already been demonstrated using these materials systems. As will be discussed in Sec. VI, the beneficial properties of M_2O_3 metal oxides arise in part because the mole fraction of cations is higher (40 mol.%) compared to MO_2 (Group IV) metal oxides, with a cation fraction of 33.3 mol %.

The magnitude of the flatband voltage shifts measured in all of these high-κ systems, suggests a substantial fixed charge density of $Q>10^{12}\,cm^{-2}$ in the films, and must be further investigated. Interface stability under standard processing conditions remains a concern, as the earlier cases illustrate that uncontrolled reaction tends to occur during the nonequilibrium deposition process through kinetics (with the apparent exception of the ALCVD-deposited Al_2O_3), despite some thermodynamic predictions. Control of the Si channel interface is a key issue for high-κ gate dielectrics, and will be revisited later.

2. Group IVB metal oxides

A substantial amount of investigation has gone into the group IVB metal oxides, specifically TiO_2,[102–109] ZrO_2,[110–121] and HfO_2,[110,120,122–126] as these systems have shown much promise in overall materials properties as candidates to replace SiO_2.

The TiO_2 system has been heavily studied for high-κ applications both for memory capacitors and in transistors.[102,106,107] It is attractive because it has a high permittivity of $\kappa=80-110$, depending on the crystal structure and method of deposition. This anomalously high permittivity, which arises through a strong contribution from soft

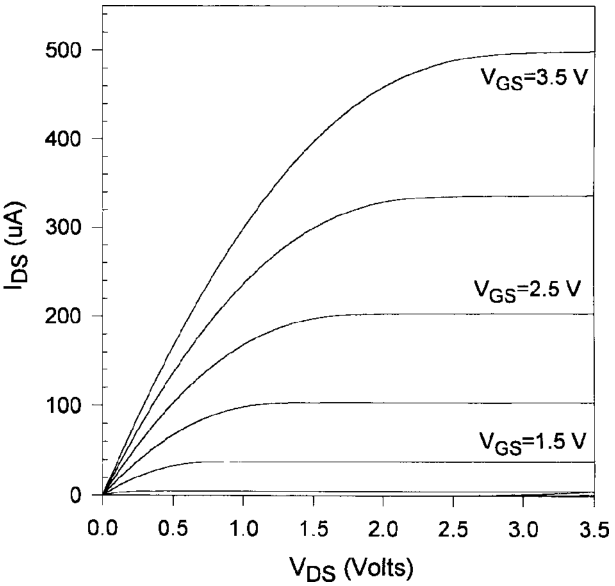

FIG. 14. $C-V$ curves for a 190 Å TiO_2 gate dielectric in a Pt/TiO_2/p-type Si MIS capacitor structure (150 μm^2 area). A gate bias-dependent accumulation capacitance was observed. Analysis which incorporates an estimate of the accumulation layer thickness indicates the presence of a SiO_2 interfacial layer, resulting in an overall $\kappa \sim 30$ for the stack (see Ref. 102). © 1997 IEEE, reprinted with permission from IEEE.

phonons involving Ti ions, is not exhibited by the other group IVB metal oxides. On the other hand, Ti has several stable oxidation states of Ti^{3+} and Ti^{4+} which lead to a well-known problem with materials containing Ti–O bonds: a reduced oxide. Such a reduced oxide has many oxygen vacancies which act as carrier traps and high leakage paths. Since TiO_2 crystallizes at temperatures much above 400 °C, it is also expected to exhibit a polycrystalline morphology, which is undesirable, as will be discussed in Sec. VI D. It is important to note that since TiO_2 is not stable on Si during deposition by CVD,[103] all of the studies on this high-κ system contain both a reaction layer at the channel interface (either intentional or unintentional) and metal electrodes/gates to prevent reaction at the gate interface.

Transistors fabricated with TiO_2 as the gate dielectric have shown results with $t_{eq} \leqslant 20$ Å, with a physical thickness of 80–120 Å, and leakage current varies widely depending on deposition method and postdeposition processing.

Full transistors using CVD TiO_2 as the gate dielectric were first reported by Campbell *et al.*,[102] using Pt electrodes to prevent reaction at the gate interface. A subthreshold slope of 83–91 mV/decade was achieved for very large transistors ($100\times 100\,\mu$m and $1\times 6\,\mu$m, respectively) and a transconductance of 322 μA/V was reported, indicating that the mobility is only ~160 cm^2/V s. The transistor performance in these devices, as illustrated in Fig. 14, was limited by a thick interfacial oxide layer which formed as a result of the CVD process and postannealing of the structure, and the relatively large interface state density ($10^{12}/cm^2$ eV). Leakage currents were also unacceptably high in the transistors, and a relatively large interface state density ($10^{12}/cm^2$ eV) was reported. Although the role of carbon in high-κ dielectrics is not yet well understood in relation to leakage current, it is possible that carbon incorporated into the film creates defect states that allow higher leakage currents. A carbon-free precursor[104] for Ti was synthesized and used to prevent carbon incorporation in the film, thereby attempting to reduce leakage currents. Furthermore, a thermal nitride was grown at the channel interface by an initial anneal in NH_3. The resulting devices showed values of $t_{eq} < 20$ Å, and leakage currents which were reduced by 100× compared to comparable SiO_2 films.[105] A hysteresis of 80 mV measured in the devices indicates the effect of fixed charge from the nitride layer at the channel interface. The flatband voltage shift was $\Delta V_{FB} \sim -200$ mV, perhaps arising from the presence of fixed charge in the film.

The jet vapor deposition (JVD) method has been used by Guo *et al.*[106] to make long-channel transistors with a TiO_2/SiN stacked gate dielectric. An interfacial nitride barrier (15 Å thick) was first deposited by JVD at 25 °C and annealed by rapid thermal annealing (RTA) in N_2 at 600 °C, in order to minimize any interfacial layer containing SiO_2 which may form during postannealing of the structure. The TiO_2 layer was subsequently deposited at a thickness of ~120 Å, followed by RTA in O_2. These structures were seen to be stable up to 900 °C, but at higher temperatures, the leakage and capacitance measurements were severely degraded. Devices using Al electrodes achieved $t_{eq} < 20$ Å, and the nMOS drive current obtained was 0.18 mA/μm for a 5 μm gate length and 10 μm gate width, at $V_D = 2.5$ V and $(V_G - V_T) = 1.5$ V. No mobility values were reported.

Ma *et al.*[107] reported that after forming TiO_2 layers by CVD, a postprocessing treatment sequence of plasma ozone (O_3) at 200 °C followed by rapid thermal anneals in N_2O or O_2 served to minimize the interfacial reaction between TiO_2 and Si. This process had the effect of reducing the leakage current by four orders of magnitude (for TiO_2 films less than 100 Å thick). The performance of MOSFETs with TiN gates was also reported, and the relatively low drive currents of ~40 μA/μm (at $V_G = 1.2$ V, $V_D = 1$ V) obtained for a 0.5 μm device were attributed to nonuniformity in the polycrystalline TiO_2 layer and corresponding roughness at the interface.

Work on alternate dielectric materials systems of group IVB metal oxides such as ZrO_2 and HfO_2 was reported in the 1970's and 80's for the purpose of optical coatings and DRAM applications.

Balog *et al.*[110] performed $C-V$ measurements on very thick (3000–4000 Å) polycrystalline HfO_2 films in Al/HfO_2/Si structures. For films grown (by CVD) at 450 °C in pure O_2, a large hysteresis was observed. It was indicated that subsequent anneals in inert ambients, such as in N_2 at 800 °C, led to a reduction in the observed hysteresis. However, HfO_2 films grown at $T \geqslant 500$ °C resulted in no observed hysteresis and $D_{it} \sim 1-6\times 10^{11}/cm^2$. Breakdown fields for these films were reported to be ~1–2 MV/cm and $\kappa = 22-25$ at 1 MHz. Balog *et al.* also found essentially identical results for ZrO_2 films in Al/ZrO_2/Si structures.[110] The permittivity for these films ranged from 17 to 18. The flatband voltage shift for both HfO_2 and ZrO_2 thick films ranged from $\Delta V_{FB} = -600$ to $+200$ mV. However, we note that the reported Si surface preparation procedure most likely resulted in the formation of a native oxide prior to HfO_2 and ZrO_2 deposition.

Shappir *et al.* examined thinner (300–600 Å) ZrO_2 films (deposited by metalorganic CVD at 450 °C) in various MIS structures (with a native oxide on the substrate) that included Al, Mo, and poly-Si gates.[111] For Al gate capacitors, annealing at $T \geqslant 800\,^\circ\mathrm{C}$ in N_2 resulted in a large reduction in leakage current. Interestingly, anneals in O_2/HCl at 800 °C resulted in the largest reduction in leakage although no reduction in the measured dielectric constant (due to potential SiO_2 formation) was observed. Breakdown fields of 4 MV/cm were reported for the subsequently annealed films. A large hysteresis was observed in all $C-V$ curves, with a shift of ~600 mV for unannealed and N_2-annealed samples, and a shift of ~200 mV for O_2/HCl-annealed samples. These films were estimated to have $D_{\mathrm{it}} \sim 1\times10^{12}/\mathrm{cm}^2$, which may have been even higher from the polycrystalline ZrO_2 were it not for the presence of the native oxide. A dielectric constant at 1 MHz ranging from 14 to 19 was observed (κ decreased with increasing anneal temperature), even with the native oxide present.[111] Leakage currents in these very thick films were reported to be as low as 8×10^{-9} A/cm^2 at 1.5 MV/cm.

Mo gate devices exhibited similar dielectric constants with a higher (~10×) loss tangent than that for the Al gate structures. Leakage currents were somewhat higher and breakdown was observed at a slightly lower value (~3.7 MV/cm). Since Mo has a larger work function ($\Phi_B = 4.5\,\mathrm{eV}$)[95] than that of Al ($\Phi_B = 4.3\,\mathrm{eV}$),[95] Mo is expected to result in a lower tunneling current for gate injection, but apparently reaction at the gate interface altered the tunneling properties. The increase in the leakage current was attributed to MoO or MoN formation during the postdeposition annealing process.

It was found that poly-Si gates produced good performance when the amorphous Si was deposited at 550 °C, followed by a $POCl_3$ doping (885 °C) and drive step (920 °C for 2.5 h), which also served to crystallize the Si electrode.[111] [It was reported that poly-Si deposition at higher temperatures (620 °C) results in the reduction of ZrO_2 to form a Zr-rich layer near the poly-Si/ZrO_2 interface, leading to very high leakage currents.] Leakage currents were comparable to that observed for the Al or Mo gate structures, but the reported breakdown field was lower than Al, ~3 MV/cm. A dielectric constant of $\kappa \sim 16$ was measured, again with the presence of a native oxide.

The large hysteresis and other shifts in the $C-V$ curves were observed for all of the electrodes and annealing conditions used in this study, which indicates that the shifts probably arise from mobile ion transport in the pure ZrO_2 films.[112] All of these studies on ZrO_2 and HfO_2 indicate crystallization of the metal oxides either during the deposition process or upon moderate postannealing conditions.

More recent work has shown encouraging results for both CVD and sputtered films. Using ALCVD, Copel *et al.*[113] demonstrated that a highly uniform layer of ZrO_2 can be deposited as thin as 20 Å on top of an SiO_2 layer (Fig. 15). In that work, the 15 Å SiO_2 layer was intentionally grown by a thermal anneal, to form a very high quality interfacial oxide.[113] As mentioned previously, this layer limits the minimum achievable t_{eq} value, but it is also required for the $ZrCl_4$ precursors used in this ALCVD process. Without a reactive SiO_2 layer on the surface, the $ZrCl_4$ precursors cannot easily displace the Si–H bonds present on the surface of a standard HF-last wet clean process. This initial oxidation therefore serves a dual purpose by providing a very high quality interface and at the same time a reactive surface on which to deposit the ultrathin layer of ZrO_2. These ZrO_2 layers were found to be thermally stable under vacuum annealing up to 900 °C, which is important, but were observed to decompose at higher-temperature vacuum anneals of 1000 °C.

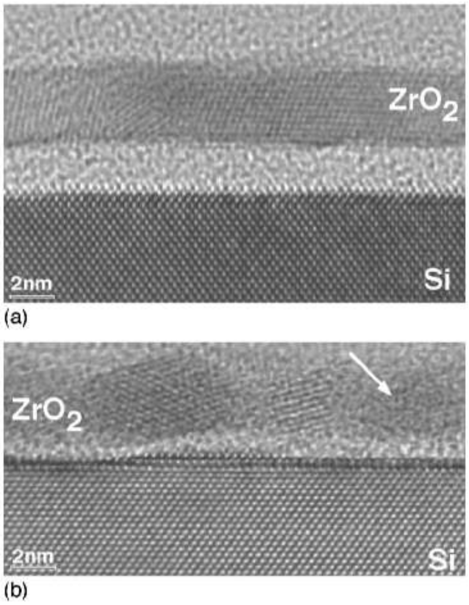

FIG. 15. HRTEM images of ZrO_2 deposited by ALCVD methods on (a) oxidized and (b) HF-last treated Si surfaces. Deposition on the H-passivated surface appears to result in island nucleation (see Ref. 113).

The uniformity of these layers was found to be remarkably good, with thickness variations less than a few percent across the wafer. It was additionally noted that ALCVD films deposited on Si surfaces prepared by a HF-last process did not produce flat and uniform films, but rather resulted in the nucleation of ZrO_2 islands interspersed with disordered material. It was suggested that this undesirable morphology might be controlled though further investigation of deposition parameters. Perkins *et al.*[127] reported the electrical characteristics of polycrystalline ZrO_2 ALCVD films deposited on chemically grown oxides. Using TiN electrodes, encouraging results of $t_{\mathrm{eq}} < 14$ Å with a leakage current density $J \sim 2\times10^{-4}$ A/cm^2 at $(V_G - V_{\mathrm{FB}}) = 1$ V were achieved. The flatband voltage shift in the ZrO_2 films was $\Delta V_{\mathrm{FB}} \sim -600$ mV, and hysteresis in the films varied with bias sweep range, as ±3 V sweeps showed hysteresis of ~80 mV and ±2 V sweeps showed ~10 mV hysteresis, respectively, for the same film.[127] The combination of the observed flatband voltage and hysteresis suggests the presence of positive charge in the films, either at the interface or in the ZrO_2 layer.

Device properties have been measured on sputtered ultrathin films of ZrO_2 and HfO_2 using Pt electrodes.[114,122] By initially sputtering only Hf metal onto the Si substrate, followed by reactive sputtering of Hf in an Ar/O_2 ambient, low t_{eq} values of 11.5 Å with a leakage current $J\sim 1\times 10^{-2}\,\mathrm{A/cm^2}$ (at a bias of $V_G-V_{\mathrm{FB}}=1\,\mathrm{V}$) were reported (no quantum corrections taken into account) with Pt electrodes.[122] Negligible hysteresis was reported for certain processing conditions, but a relatively high ΔV_{FB} value of $\sim+600$ mV was reported in one case,[122] and ~-300 mV was reported in another,[124] with slightly different processing conditions. These significant flatband shifts perhaps arise from a large amount of negative fixed charge (for positive ΔV_{FB}) and positive fixed charge (for negative ΔV_{FB}), respectively, in the films. The breakdown appears to occur in the HfO_2 layer, as somewhat low breakdown fields of $E_{\mathrm{BD}}\sim 4\,\mathrm{MV/cm}$ were measured, as was previously reported for thicker films.[110,134] An interfacial layer always forms through this sputtering deposition and postannealing process (Fig. 16), because of the well-known fast diffusion of oxygen through ZrO_2 and HfO_2.[112] In this case, oxygen which diffuses through these metal oxides reacts with Si at the interface to form an uncontrolled interfacial layer.

Under optimized processing conditions,[122] it was reported that the contribution from the interfacial layer is $t_{\mathrm{eq}}\sim 5$ Å, although a significant level of interface states is still apparent in the $C-V$ curves at low frequencies, as seen in Fig. 17. Although x-ray diffraction shows the HfO_2 films to be amorphous as-deposited, annealing above 700 °C was observed to cause crystallization in these thin films, as was previously reported for thick films.[110] Polycrystalline gate dielectric films will most likely have higher leakage current, be less uniform and less reproducible than amorphous films, but the correlation between gate dielectric morphology and device performance requires more investigation.

In the case of ZrO_2, it has been reported[114,115] that sputtering Zr in a process similar to that described for HfO_2 above leads to capacitors (using Pt electrodes) with $t_{\mathrm{eq}}=16$ Å, and leakage current density $J\sim 3\times 10^{-2}\,\mathrm{A/cm^2}$ (at a bias of $V_G-V_{\mathrm{FB}}=0.5\,\mathrm{V}$). The interface once again shows a reaction layer between the ZrO_2 and Si, and this layer was reduced to a physical thickness of about 10 Å by optimizing the deposition conditions. The interface was seen to yield a hysteresis of 50 mV, a flatband voltage shift $\Delta V_{\mathrm{FB}}\sim+200$ to $+300$ mV and $D_{\mathrm{it}}\sim 10^{11}\,\mathrm{cm^{-2}\,eV^{-1}}$, indicating the limiting quality of the reaction layer. The particularly large flatband voltages may arise in part from processing conditions during deposition and postannealing.

Houssa *et al.*[118] reported a systematic attempt at understanding the flatband voltage and fixed charge in ZrO_2 layers deposited by ALCVD on native silicon oxide. Their study showed that the net fixed charge density could be altered significantly depending only on the post-annealing conditions. Using Eq. (9) to obtain values for the amount and sign of the fixed charge in the dielectric, it was found that as-deposited ZrO_2/SiO_x stacks exhibit negative fixed charge, and that postannealing in O_2 introduced compensating positive fixed charge *and* at the same time increased the density of midgap interface states.[118] These results are shown in Fig. 18(a), where the as-deposited fixed charge density is estimated to be $-2.5\times 10^{12}\,\mathrm{cm^{-2}}$ (with $\Delta V_{\mathrm{FB}}\sim+200$ to $+600$

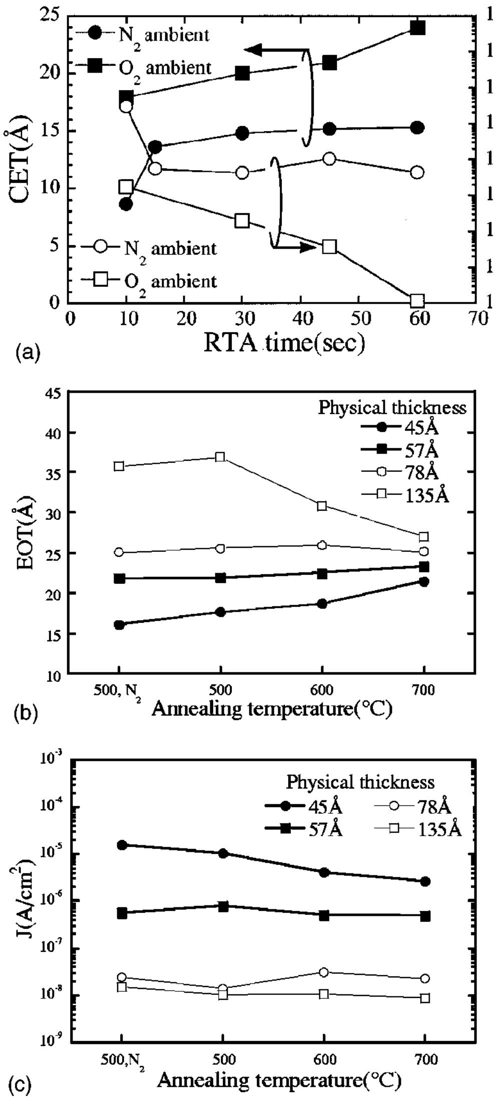

FIG. 16. Effect of postdeposition annealing on HfO_2 gate dielectrics. (a) N_2 and O_2 anneals, (b) equivalent oxide thickness, and (c) leakage current after O_2 anneals (see Refs. 122 and 124). © 2000 IEEE, reprinted with permission from IEEE.

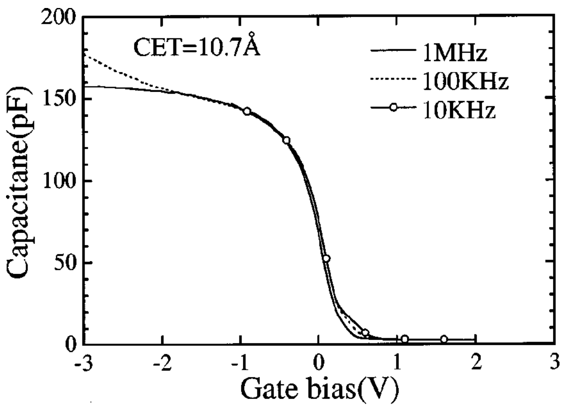

FIG. 17. $C-V$ curve of HfO_2 after a rapid thermal anneal at 600 °C. The dispersion for $V_G<-2$ V is attributed to higher leakage. Interface states are noted near $V_G\sim 0.5$ V (see Ref. 122). © 2000 IEEE, reprinted with permission from IEEE.

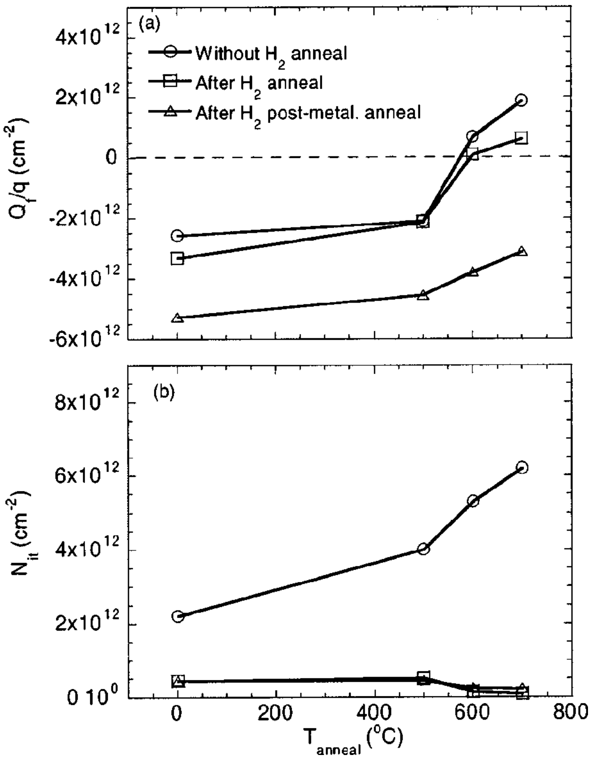

FIG. 18. Extrapolated values of (a) fixed charge density Q_f and (b) midgap interface state density N_{it} as a function of postannealing conditions in O_2 and H_2 for ALCVD ZrO_2 (see Ref. 118).

mV, depending on film thickness). Postannealing in O_2 at progressively higher temperatures is seen to decrease the net amount of negative fixed charge (presumably through compensation by the introduction of positive fixed charge), until a net *positive* fixed charge is observed for O_2 anneals above 600 °C (a net positive fixed charge density of $2\times10^{12}\,\mathrm{cm}^{-2}$ results after a 700 °C anneal in O_2).[118]

Figure 18(b) shows that for the same annealing conditions, the midgap interface state density also increases. Under subsequent annealing in H_2, the midgap interface states are minimized, while little effect is seen on the net fixed charge. If the subsequent H_2 anneals are carried out after deposition of the metal gate, however, the net fixed charge is found to become significantly *more* negative, such that even the higher temperature anneals in O_2 do not produce a net positive fixed charge. It is proposed that the resulting sign and amount of fixed charge is dependent on the annealing conditions, and that the positive fixed charge arises from over coordinated oxygen atoms, possibly induced by the presence of hydrogen,[128] as has been proposed for the case of pure SiO_2.[129]

We note that hafnium suffers the misconception that it exhibits sufficient radioactivity to cause sensitivity to single event upsets in integrated circuits. It is well known that Hf has been used in nuclear reactors as the control rod material, because it has a high thermal neutron capture cross section, and shows little decrease in capture cross-section after long periods of radiation exposure.[130] Hafnium also shows high corrosion resistance in hot water, which makes it a low-maintenance material for control rod purposes. Both zirconium and hafnium have naturally occurring, long-lived radioactive isotopes ^{96}Zr and ^{174}Hf. The following analysis suggests, however, that neither element should be a concern with respect to radioactive decay.[131]

TABLE II. Thermal neutron cross section data for device relevant materials (after Ref. 132).

Element	Isotope	Natural abundance (at. %)	Thermal neutron cross section (b)	Weighted average cross section (b)
Zr	90	51.46	0.1	
	91	11.23	1	
	92	17.11	0.2	0.213
	94	17.4	0.08	
	96	2.8	0.05	
Hf	174	0.163	400	
	176	5.21	30	
	177	18.6	370	
	178	27.1	50	97
	179	13.75	65	
	180	35.22	10	
Cu	63	69.1	4.5	
	65	30.9	2.3	3.82
Si	28	92.18	0.08	
	29	4.71	0.3	0.0913
	30	3.12	0.11	
B	10	19.7	3837	
	11	80.3	0.005	756

The naturally occurring radioactive isotope for zirconium is ^{96}Zr. The natural abundance is 2.80 at. %. The half-life $T_{1/2}$ is $>3.6\times10^{17}$ years. Assuming reasonable values for transistor dimensions, the number of Zr atoms resulting from a $ZrSiO_4$ stoichiometry associated with the gate dielectric is $\sim1.5\times10^{6}$, resulting in 4.2×10^{4} ^{96}Zr atoms/gate. Assuming 10^9 transistors in a device, this results in 4.2×10^{13} ^{96}Zr atoms per device. The number of disintegrations/year is given by $A^*\lambda$, where $A=$number of atoms and $\lambda=$probability for decay, *viz.* $\lambda=0.693/T_{1/2}=1.9\times10^{-18}\,\mathrm{yr}^{-1}$. As a result, one might expect 8×10^{-5} disintegrations per device per year, or one disintegration every 1.2×10^{4} years per device. A similar analysis for ^{174}Hf indicates that one disintegration every 1.6×10^{3} years per device if $HfSiO_4$ is used as the gate dielectric.

There is general agreement that the flux of particles that might interact with a device consists of more than 97% neutrons at sea level.[132] The relevant parameters (neutron cross sections and natural abundances) are given in Table II for hafnium and zirconium, as well as other materials that will be found in future generation devices. Hafnium has a larger thermal neutron cross section than zirconium, but to determine whether or not this cross section is significant, one must also consider the number of hafnium atoms present in the gate dielectric. To determine if hafnium is more of a liability than copper for example, one must take into account the relative number of hafnium atoms to copper atoms, or even relative to silicon in the silicon substrates. Ziegler[132] describes this as the active atoms/chip, given by

$$\begin{aligned}&\text{active atoms/chip}\\&\quad=(\text{active area})\cdot\text{electrical depth}\cdot\text{atom density.}\end{aligned}$$

TABLE III. Active cross section for different elements, using the thermal neutron cross section (after Ref. 132).

Element	Active atoms	Neutron cross section	Active cross section
Si	2.0×10^{18}	9.0×10^{-26}	1.8×10^{7}
Zr	1.5×10^{15}	2.1×10^{-25}	3.2×10^{-10}
Hf	1.1×10^{15}	9.7×10^{-23}	1.1×10^{-7}
Cu	3.0×10^{17}	3.8×10^{-24}	1.1×10^{-6}
B	1.0×10^{15}	7.6×10^{-22}	7.6×10^{-7}

Assuming a 1 cm^2 silicon-based device with 4% of the device active and a 1 μm electrical depth, one arrives at $(0.04\,\text{cm}^2)\cdot(10^{-3}\,\text{cm})\cdot(5\times10^{22}\,\text{atoms/cm}^3)=2\times10^{18}$ Si atoms. One must also determine the active cross section (=active atoms·neutron cross section), which takes into account the differences in neutron cross sections for the different elements (shown in Table II). For Si, the weighted average neutron capture cross section is $9.13\times10^{-26}\,\text{cm}^2$ resulting in an active cross section of $2\times10^{-7}\,\text{cm}^2$. Other relevant elements are shown in Table II.

While this approximation does not provide an estimate of the absolute single event upset rate, it should provide an estimate of the relative sensitivity for different elements in the device to cosmic rays. Thermal neutron cross-sections were used as an estimate of the overall neutron cross section which is likely an overestimate of the interaction cross section.

Table III compares these sensitivities for the different elements and their associated isotopes discussed earlier. Even with the relatively high thermal neutron cross sections for Zr and Hf, these data would suggest that neither Hf nor Zr should cause as many upsets as Cu. The number of Cu atoms was determined by assuming a copper metallization process consisting of 0.3-μm-thick layers, five levels of metal, and 4% of each layer is made up of Cu. Boron was estimated assuming an implant dose of $1\times10^{15}\,\text{atoms/cm}^2$.

The results obtained on the ZrO_2 and HfO_2 metal oxide systems described earlier indeed are important demonstrations of the ability to achieve low t_{eq} values with lower leakage currents than would be achieved using comparable SiO_2 films. The interface quality of these systems remains a critical issue, however, since the materials in particular are extremely susceptible to O diffusion and reaction at the channel interface. The ultimate device performance will depend heavily on the channel carrier mobility, which in turn can be easily degraded by high interface trap levels. A large hysteresis or shift in flatband voltage for any gate dielectric will also result in unacceptable transistor threshold voltage shifts. Finally, these metal oxides have not yet demonstrated the required stability under the high temperatures required for CMOS processing, including dopant activation anneals up to 1050 °C. At this point, lower thermal budget processes have been considered, such as the replacement gate process,[69] but no proven manufacturable flow has yet emerged. Until this occurs, it should be considered a critical requirement for high-κ dielectrics to withstand such high-temperature anneals.

In the cases where a flatband voltage has been reported for the earlier oxide systems, the measured flatband voltage shifts ΔV_{FB} are substantial (compared to that expected for the given electrode work function), ranging from -600 to $+800$ mV for the high-κ films discussed thus far. As mentioned previously, in this case the term ΔV_{FB} does *not* refer to a flatband shift arising from $C-V$ voltage sweeps in opposite polarities, but rather refers to the *difference* in the measured flatband voltage (for a single sweep) from that expected for an ideal capacitor (for a given electrode and substrate doping level).

This shift is normally interpreted as fixed charge within the film, although it can also arise from oxide damage associated with gate electrode deposition or other forms of processing treatments. Considering that large ΔV_{FB} values have been measured by so many independent groups, using different processing conditions and electrodes, it is currently being attributed to fixed charge within the film. A large fixed charge (e.g., that which corresponds to $\Delta V_T > 50\,\text{mV}$) in a gate dielectric can have a serious, deleterious effect on the transistor performance, because the threshold voltage V_T becomes too large for adequate compensation by dopant implants. Smaller fixed charges (~10–20 mV) can be acceptable, and even desirable, since this can adjust the V_T such that smaller channel implant doses can be used (which results in less ion impurity scattering of carriers in the channel).

Although it is not yet known whether the observed flat band shifts arise from a fixed charge within the high-κ dielectrics or from some other phenomenon, a large amount of fixed charge in these films could have an enormous influence on the viability of their insertion into a CMOS process. If these high-κ gate dielectrics indeed prove to contain significant amount of fixed charge, then the sign of the charge will also become extremely important. The band diagram in Fig. 4 illustrates the different effects on a device threshold voltage for the cases of positive and negative flatband voltage shifts. This observed phenomenon has been suggested[11] to arise from the detailed bonding structure of the various cations near the Si interface, while another report by Houssa *et al.*,[118] mentioned previously, suggested that the sign and amount of fixed charge depends only on the annealing conditions. It should be noted that only positive ΔV_{FB} values have been measured for Al_2O_2, but both positive and negative ΔV_{FB} values have been measured for several high-κ dielectrics.

For example, TiO_2, ZrO_2, and HfO_2 show a range of $\Delta V_{\text{FB}}=-600$ to $+800$ mV, while Y_2O_3 and La_2O_3 show $\Delta V_{\text{FB}}=+300$ to $+1400$ mV (on much less data than for group IVB oxides), but Al_2O_3 shows $\Delta V_{\text{FB}}\sim+300$ to $+800$ mV. Since a negative value corresponds to positive fixed charge (assuming again that these shifts are due to fixed charge) and vice versa, Al_2O_3 seems to show opposing behavior in that it may possess negative fixed charge. These two cases clearly have different impacts on nMOS and pMOS devices. Much more work must be done to determine if indeed there is a substantial amount of fixed charge in these high-κ dielectrics on Si, and if so, how to remove it or at least make manageable adjustments for acceptable device performance.

To summarize the many results reported thus far on the unary, simple oxide systems, all high-κ metal oxides investigated to this point have demonstrated encouraging electrical properties. All of the unary oxides reported have achieved $t_{\mathrm{eq}} < 15$ Å, with $J < 10^{-2}$ A/cm^2 (at $V_G - V_{\mathrm{FB}} = 1$ V). Nearly all of the metal oxides, however, show significant frequency dependence, hysteresis and flatband voltage shifts, as well as concerns regarding oxygen diffusion and interface stability during subsequent high-temperature thermal processing.

The following section focuses on alloys of these oxides, in an attempt to combine and complement the desirable properties from several materials, and thereby overcome the deficiencies associated with the individual materials. These alloy systems are predominantly non-stoichiometric mixtures, and are therefore termed *pseudobinary alloys*.

3. Pseudobinary alloys

All of the earlier mentioned materials clearly have many advantages as high-κ gate dielectrics, but they also have undesirable properties which remain a concern regarding replacement of SiO_2. An encouraging system of materials are pseudobinaries, such as $(ZrO_2)_x(SiO_2)_{1-x}$ and $(HfO_2)_x(SiO_2)_{1-x}$, which for this purpose combine two oxides, typically (although not necessarily) in nonstoichiometric compositions. In this way, it is possible to combine the desirable properties from two different oxides while eliminating the undesirable properties of each individual material.

It is instructive to consider the chemical and electrical characteristics associated with Zr and Hf that serve to create such favorable electrical properties for the CMOS applications (and conversely why Ti is not favorable). The four-fold coordination of Si and the desire to minimize the presence of dangling bonds at the interface suggests that high-Z elements in group IVB are desirable when incorporated into an insulator oxide or silicate.

Some work in the 1970's and 80's (as mentioned previously, the work was primarily intended for use in optical lens coatings) was carried out on such pseudobinary materials. Examination of sputtered ZrO_2-SiO_2 codeposited, thick (~3000 Å) films has been reported by Russack and co-workers.[133] X-ray diffraction analysis of the films indicated that the films with $>10\%\,SiO_2$ remain amorphous even upon annealing at 500 °C in air for 60 h. Films with $<10\%\,SiO_2$ exhibited a crystalline $ZrSiO_4$ diffraction pattern under similar annealing conditions. Composition analysis of their films, as deduced by Rutherford backscattering spectroscopy, indicated that the films have a metal:oxygen ratio of 2:1, signifying that they are fully oxidized.

Roberts *et al.* investigated HfO_2:SiO_2 films also prepared by sputter techniques.[134] They found that Hf-silicate films provided superior performance to sputtered Ta_2O_5 films for thin capacitor insulator film thicknesses (100–150 Å range). A preferred stoichiometry of Hf_4SiO_{10} was identified as a result of breakdown measurements and $\kappa = 13$ was reported. It was noted that interfacial SiO_2 growth occurs upon subsequent O_2 annealing above 800 °C for Hf_xSiO_{2x+2}, where $x<2$ or $x>4$. Additionally, investigation of interface engineering using $SiO_2/Hf_4SiO_{10}/Si_3N_4$ insulator stacks provided improved performance compared to similar structures incorporating Ta_2O_5.

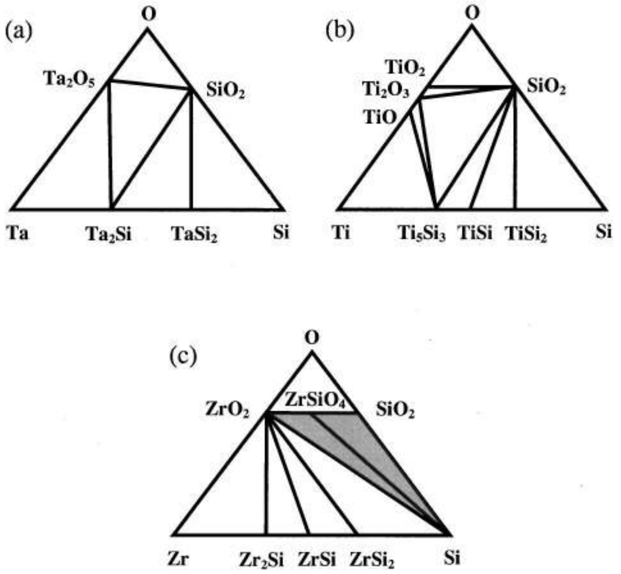

FIG. 19. Ternary phase diagrams for (a) Ta–Si–O, (b) Ti–Si–O, and (c) Zr–Si-O compounds (after Refs. 156 and 158).

Recent work on two such systems of pseudobinaries, namely silicates (M–Si–O), with M=Zr, Hf,[115,116,135–138,143(a)] La[143(b)] and Gd,[143(c)] and aluminates (M–Al–O), with M=Zr,[139,140] indicate that such materials systems already exhibit encouraging gate dielectric properties toward this end. Both materials systems have the same underlying principle of mixing a high-κ (crystalline) metal oxide with an amorphous, stable, lower-κ material (SiO_2 or Al_2O_3) to obtain a desirable morphology with suitable properties for a CMOS gate dielectric. The effect of adding SiO_2 and Al_2O_3 to metal oxides such as ZrO_2 is to produce an amorphous film that is thermodynamically stable on Si (for low leakage currents). The overall permittivity of the pseudobinary alloy is inevitably lower than that of the pure metal oxide, but this tradeoff can be very adequate, for the improved stability. In particular, ZrO_2:SiO_2 and HfO_2:SiO_2 silicates within an appropriate composition range have been demonstrated to exhibit very low leakage currents and improved κ values with only small amounts of ZrO_2 or HfO_2 in the material.[135–137]

Since Figs. 19(a) and (b) show that there are no apparent thermodynamically stable ternary compounds for the Ta–Si–O and Ti–Si–O systems, glassy silicates ($Ta_xSi_yO_z$ or $Ti_xSi_yO_z$) of these materials may be obtained. The subsequent thermal processing that these materials experience will be of key importance, though, as these systems are likely to separate into more stable M_xO_y and M_xSi_y phases. Furthermore, an extension of constraint theory for glasses, as formulated by Phillips[141] has recently been applied to glassy phases of high-κ gate dielectrics. It has been found[142] that there is excellent agreement between the expected stability and low defect densities specifically for certain composition ranges of ZrO_2:SiO_2 and HfO_2:SiO_2 silicates.

Many other silicate systems are possible, but the main concerns for all of them are whether the κ value is large enough for gate dielectric applications, and whether or not they will tend to phase separate when placed next to Si at the temperatures of interest. Ternary systems such as Al–Si–O, Y–Si–O, or Sc–Si–O are possibilities, because their metal oxide phases are stable on Si, but these silicates may suffer from relatively low κ values. Other possible silicate candidates which may work well are those which contain rare earth metals with high atomic numbers, such as La–Si–O, Ce–Si–O, and Ba–Si–O. There may be some stability concerns, however, for Ce–Si–O on Si, and BaO is notorious for its hygroscopic nature.

Glassy phases of a ternary or quaternary system containing only metals and oxygen may also be obtained, but this again raises concerns of uncontrolled reaction at the Si channel interface. For example, TiO_x films have been doped with various metals to obtain an amorphous film, as a means of reducing leakage and improving thickness uniformity.[107,109]

In the study by Ma *et al.*,[139] Al and Zr were cosputtered to form $(ZrO_2)_x(Al_2O_3)_{1-x}$ amorphous alloys with $x = 0.75$. Using TiN gates, these pseudobinary alloys achieved $t_{eq} \sim 12$ Å with $J \sim 0.1$ A/cm^2 at 1 V gate bias. The flatband shift on capacitors was $\Delta V_{FB} = +200$ to $+300$ mV. Transistors fabricated using these alloys showed good characteristics, with a subthreshold swing of 72 mV and $I_d = 1.3$ mA for a 0.6 μm$\times$1 μm *p*MOSFET, with $V_D = V_G = -1.5$ V. Hole mobilities were found to be ~30% lower than for comparable SiO_2 FETs at $E_{eff} \sim 1$ MV/cm. The work by Manchanda *et al.*[140] also used sputtered Zr–Al–Si on 5 Å SiO_2 or Si–O–N films, followed by thermal oxidation, to form $(ZrO_2)_x(Al_2O_3)_{1-x}$ pseudobinary alloys with $x \sim 0.75$ (there was also ~1.7% Si in the films). The films were found to remain amorphous up to 800 °C for 30 min. TiN gates were used to form capacitors, and $t_{eq} = 12$ Å was achieved with $J = 0.1$ A/cm^2 at 1.5 V gate bias. An interface state density of $D_{it} \sim 1-5 \times 10^{11}$ cm^2 was reported for the Zr–Al–Si–O films. No channel carrier mobilities were reported.

Other work by van Dover[109] examined the effect of a wide range of dopants in sputtered TiO_x films on film structure and leakage. The incorporation of up to 10% Nd, Tb, and Dy lanthanide dopants in Ti–O films produced amorphous structures with low leakage and high specific capacitance (no postannealing was done). Using Pt top electrodes and TiN bottom electrodes, permittivities ranging from 50 to 135 were measured for 350-Å-thick films, while maintaining an amorphous layer. Leakage currents of 5×10^{-7} A/cm^2 were obtained at an applied field of 2 MV/cm for each of the cases of Nd, Tb, and Dy doping. Although these materials systems are better-suited for memory capacitors than for transistors (due to thermal instability on Si), the results indicate the importance of considering dopants as a means of reducing leakage current without substantially degrading other electrical properties of the high-κ dielectric. The particular dopants presumably tie up dangling bond deep trap levels associated with unsatisfied cation bonds (Ti in this case), while not introducing deleterious trap levels of their own. A performance improvement is clearly attainable using high-κ dielectrics such as TiO_2. The ultimate success of these materials, however, depends on an extremely high level of quality and control over the channel interface, which has not yet been demonstrated.

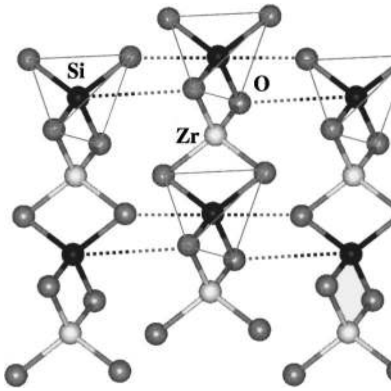

FIG. 20. Structure of crystalline $ZrSiO_4$ showing the Zr bonding to SiO_2 units. Zr–O bonding also exists in and out of the plane of the page.

Quaternary silicate films containing Zr and Hf, and possibly other elements, are certainly conceivable, but the material which satisfies the demands of technology while requiring the least change, is preferred. At this point, we will focus on the Hf–Si–O and Zr–Si–O systems because of the combination of all their desirable properties listed earlier and the promising results recently reported.[115,116,135–137,143(a),143(b),143(c)]

One potentially large advantage for silicates is that this class of materials should have a silicate-Si interface that is chemically similar to the SiO_2–Si interface, which is unparalleled in quality for transistor channel regions. This is especially important since the channel interface is playing a dominant role in determining device performance, and because almost any simple oxide (M_xO_y) deposited on Si will form a silicate interfacial layer, even if it is very thin. In this light, the tetravalent transition metal cations such as Zr and Hf offer the advantage of substituting well at Si sites, which form SiO_4 tetrahedra. For the case of forming nonstoichiometric silicates [e.g., $(ZrO_2)_x(SiO_2)_y$], where *x* and *y* are not integers), a tetravalent cation such as Zr^{4+} or Hf^{4+} ion should substitute well for Si^{4+}, to provide a favorable bonding for a silicate network with low defect densities. In the case where stoichiometric compounds are formed, then other cations may also work, such as La^{3+} in a compound $(La_2O_3)_x(SiO_2)_{1-x}$ silicate, where $x = 0.5$.[42,143(b)]

There has been much less information reported on $(HfO_2)_x(SiO_2)_y$ than on $(ZrO_2)_x(SiO_2)_y$, but the chemical similarities between Hf and Zr allow for comparison between the respective silicate systems. As described by Blumenthal,[144] and Bragg *et al.*,[145] the Bravais lattice for the stoichiometric compound $ZrSiO_4$ (zircon) is body-centered tetragonal, and belongs to point group D_{4h}. The crystal is composed of SiO_4 tetrahedra interspersed with Zr atoms, but can be considered as parallel chains of ZrO_2 and SiO_2 structural unit molecules, as shown in Fig. 20. Each Zr and Si atom shares bonds to four O atoms within each chain, and each successive pair of O atoms is oriented in a transverse configuration, forming ZrO_2 and SiO_2 units. The Si–O bond length is shorter than the Zr–O bond length within a chain,

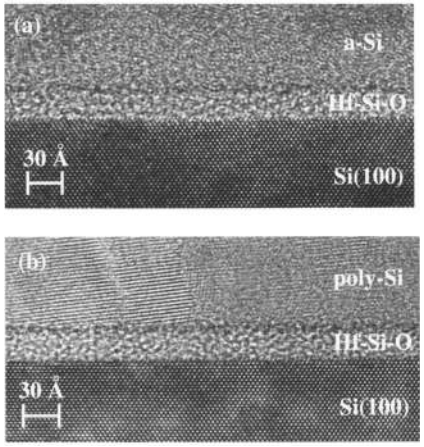

FIG. 21. HRTEM image of Hf-silicate between Si layers. (a) After deposition of amorphous Si at 500 °C and (b) after annealing at 1050 °C for 20 s in N_2 (see Ref. 137).

as is represented in the figure. Figure 20 shows that each Zr atom also shares bonds with other O atoms in neighboring chains (only two of the four bonds to neighboring chains are shown for each Zr atom, for clarity), providing a three-dimensional stability to the material.

It is important to note that each Zr and Si atom has only O atoms as nearest neighbors. Chemical analysis of homogeneous silicate films is therefore expected to show only Zr–O nearest-neighbor bonding, with a slight effect from Si as a next-nearest neighbor. Although the films presented in this study are amorphous, it is reasonable to assume that for Hf and Zr concentrations less than that of the stoichiometric MO_2-SiO_2 compound, nearly all bonds will be Zr–O (or Hf–O) and Si–O bonds.[115,116,135–137,143(c)]

Using coordination chemistry arguments between Hf and Zr, $HfSiO_4$ should have the same structure as $ZrSiO_4$. A value of $\kappa = 12.6$ for bulk $ZrSiO_4$ was reported by Blumenthal.[144] Similarly, since HfO_2 has reported values[110,146] of $\kappa = 21-25$ (Ref. 146 also reports $\kappa = 40$, but this value has not been confirmed in more recent studies), a $HfSiO_4$ compound is expected to have a range of $\kappa \sim 13-20$. The exact value of κ will certainly depend strongly on film composition, density and structure (amorphous materials typically have less lattice polarizability than their crystalline counterparts). Considering all of the desired properties, $(ZrO_2)_x(SiO_2)_{1-x}$ and $(HfO_2)_x(SiO_2)_{1-x}$ should be excellent materials candidates for advanced gate dielectrics, and indeed some very encouraging results have been previously reported.[115,116,135–137,143(c)]

In the first demonstration of a stable gate dielectric with a $t_{\mathrm{eq}} < 20$ Å deposited directly on Si, a smooth interface was demonstrated in the work by Wilk and Wallace for both Hf and Zr silicate films.[135,136] Figure 21 shows this interface for a HfO_2:SiO_2 dielectric with Si at both interfaces where it is seen that the annealing required to activate and crystallize the amorphous Si [Fig. 21(a)] does not result in a detectable interfacial layer. Figure 22 shows the results of using Al and Au electrodes in MIS a capacitor evaluation of the Hf-silicate layer. It is seen that somewhat lower accumulation capacitance and somewhat higher leakage current is observed for the Al gate electrode sample. High-resolution TEM (HRTEM) results, not shown, also indicate a contrast change at the Al/silicate interface, similar to that observed in the case of Al/Zr-silicate structures.[137] These results suggest that Al/silicate reactions likely occur with postdeposition processing, as may be expected for silicate materials systems. Leakage current densities, however, remained well below that for equivalent SiO_2 gate dielectric films, with $J < 10^{-5}$ A/cm^2 at $V_G - V_{\mathrm{FB}} = 1$ V.

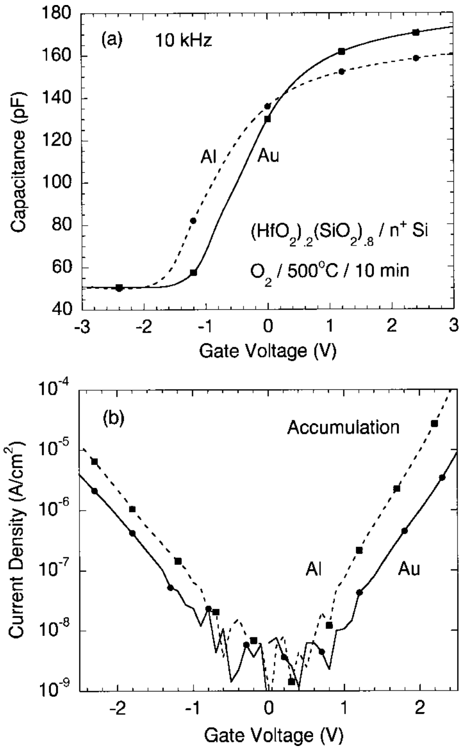

FIG. 22. Comparison of $C-V$ and $J-V$ performance for a MIS structure incorporating Hf silicate. The Al electrode results in interfacial reactions (see Ref. 137).

We further note that the potential use of silicate (and perhaps to a lesser extent aluminate) pseudobinary alloy systems as a replacement for SiO_2 gate dielectrics, in conventional CMOS processing, arguably represents the historical evolution of dielectrics in CMOS technology. Silicon dioxide has been unimaginably ideal for integrated circuits, because it has been used for decades both as the interlevel dielectric (ILD) between successive metal interconnect levels and as the gate oxide for MOSFETs. Recent technology demands, however, have required that SiO_2 in the ILD be modified to create a lower dielectric constant, so that lower RC time delays can be achieved. These changes have been made by simply incorporating H, F, and C into the SiO_2 to lower the κ value of the oxide. These modifications, while certainly not trivial, have thus far proved to be manageable in the integration process.

Similar demands for higher drive currents and lower leakage have required SiO_2 in the gate dielectric to be altered by adding N, both in the form of doping and as a silicon nitride layer on top of SiO_2. This process slightly increases the effective dielectric constant, allowing use of a physically thicker film, and thereby achieves a higher drive current with less leakage. With these trends in mind, silicate and aluminate gate dielectrics should also be viewed as only another modification of SiO_2. These pseudobinary systems allow for control in obtaining the desired materials properties by ''doping'' SiO_2 or Al_2O_3 with Zr or Hf, to further improve device performance while remaining compatible with the rest of well-established CMOS processing procedures. In fact, if these silicate films can be used, then by simply increasing the doping level of Zr or Hf from one technology node to the next for gradually improved device performance, this same materials system should fulfill the requirements on current roadmaps until the year 2010.

4. High-κ device modeling and transport

The incorporation of stacked structures has been investigated by a number of researchers in recent years. As discussed previously, the layer within the stack that has the lower dielectric constant will limit the overall capacitance of the stack. However, the minimization of interface states may require suitable interfacial layers that serve as a transition region between the Si substrate and the dielectric. One can also obviously envision a graded composition throughout the dielectric permitting control of interfacial state formation which preserves, to some extent, the high-κ properties sought for the alternative gate dielectric in the gate stack.

Vogel *et al.*[147] considered such effects in a model of potential gate stack materials. As acknowledged in the work, the model does not incorporate trap assisted tunneling mechanisms but does provide an indication of the trends associated with stack layers and scaling.

As seen in Fig. 23(a), tunneling currents associated with idealized gate dielectrics ($t_{eq}=2$ nm) are reduced dramatically in the voltage regime anticipated for future devices ($V_{DD}\sim 1$ V). Dielectric constants assumed for each curve are $\kappa=3.9(SiO_2)$, $7.5(Si_3N_4)$, $5.7(SiO_xN_y)$, 25 (D1), and 30 (D2). It is noted in that work, however, that it is more beneficial for a suitable alternate dielectric to exhibit a barrier slightly higher than V_{DD} in order to minimize tunneling, rather than a slightly higher dielectric constant. This is seen in the cross over point in the tunneling current for candidate dielectrics D1 and D2.

The incorporation of an interfacial SiO_2 layer (0.5 and 1.0 nm thick) as a part of a gate stack is examined in Fig. 23(b). The effect of electron injection through this interfacial layer from either the gate or the substrate is also examined by simply changing the direction of the electric field. In addition to the expected reduction in the overall tunneling current, it is seen that the tunneling current changes substantially depending on the dielectric layer first encountered by the electron.

This can be understood by consideration of the dielectric band diagrams associated with the stacked structure (Fig. 24). For a 1 nm SiO_2 interfacial layer, it can be seen in Fig. 24(a) that an electron with a suitable energy (from $V_{ins}=-2.68$ V) that first encounters the SiO_2 layer avoids the barrier associated with the high-κ layer, D1. Tunneling from the opposite direction, however, results in encounters with the barriers associated with both layers and a concomitant extensive reduction in tunneling (leakage) current.

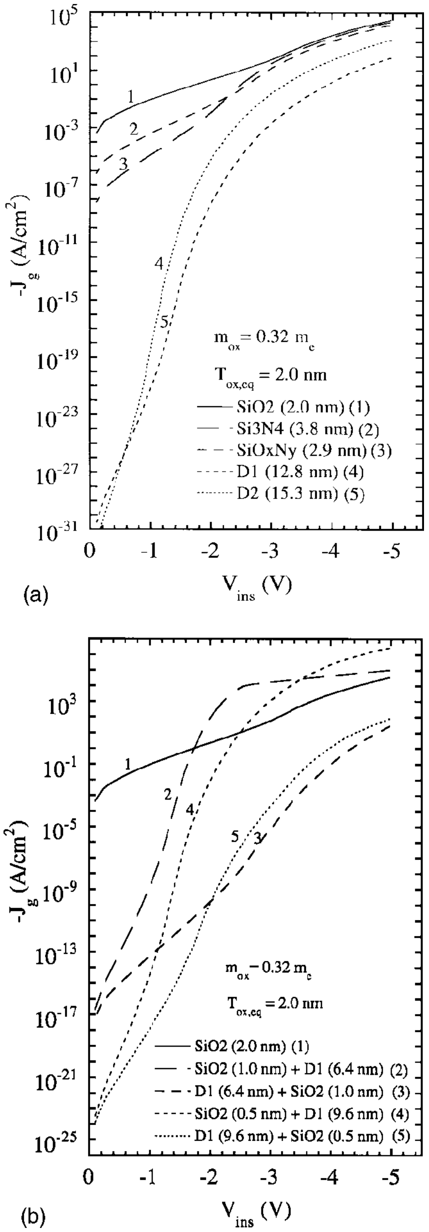

FIG. 23. (a) Tunneling current calculated for gate dielectrics in a n^+p capacitor structure exhibiting an equivalent oxide thickness of $t_{eq}=2$ nm. (b) Tunneling current calculated for stacked structures. The layer listed first is the layer initially encountered by the tunneling electron. Interfacial SiO_2 layers 0.5 and 1 nm thick are considered (see Ref. 147). © 1998 IEEE, reprinted with permission from IEEE.

For an even thinner SiO_2 layer, electrons of the same energy as that in Fig. 24(a) encounter substantial regions associated with the high-κ barrier [Fig. 24(b)]. Hence, for higher biases, the electron will eventually encounter only the barrier associated with the interfacial layer and therefore a substantial increase in tunneling current would be evident [cf. to Fig. 23(b)].

In an attempt to predict the effect of high-κ gate dielectrics on transistor performance, Frank *et al.*[148] modeled gate dielectrics with various permittivities in a planar, bulk CMOS structure. It was reported that the upper limit of permittivity would be limited to $\kappa\sim 20$ due to fringing field-induced barrier lowering at the drain region of the device.

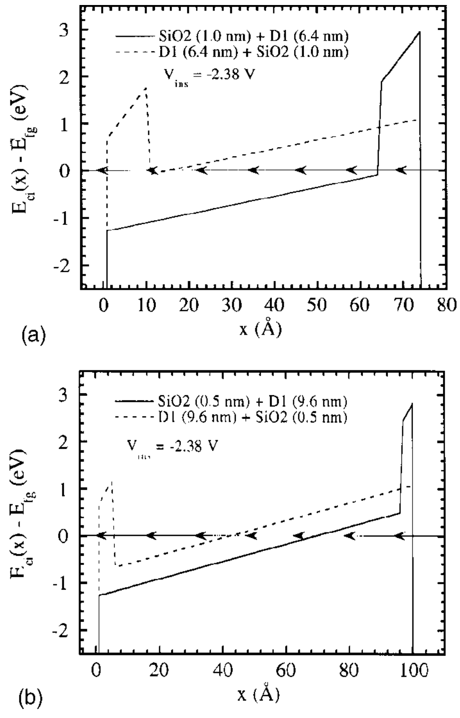

FIG. 24. Band diagrams for results in Fig. 23(b). (a) 1 nm SiO_2 interface layer, (b) 0.5 nm interface layer. The injection polarity (and energy) of the electron results in sampling different portions of the barriers associated with each dielectric (see Ref. 147). © 1998 IEEE, reprinted with permission from IEEE.

This phenomenon is a large concern, because a significant fringing field from the edge of a high-κ dielectric could lower the barrier for transport into the drain enough to seriously degrade the on/off characteristics of the device. Krishnan *et al.*[149] reported similar modeling results for high-κ gate dielectrics, but also claimed that a dielectric stack with SiO_2 at the channel interface could reduce any barrier-lowering effects from the high-κ fringing fields.[150]

Perhaps even more important is the issue of field penetration into the Si channel region.[148] The inversion charge in the channel experiences an increasing electric field with increasing gate capacitance, regardless of the gate dielectric material. At a high enough electric field penetration through the gate dielectric, channel carriers will undergo increased scattering, ultimately leading to a *decrease* in mobility and drive current. Additionally, this inversion layer will have an associated capacitance in series with the gate stack and will also eventually limit the ultimate gate stack capacitance for any high-κ dielectric.[26]

This effect was first reported by Timp *et al.*[20] using pure SiO_2, as 30 nm gate length transistors showed that for $t_{\mathrm{eq}} < 13$ Å, the drive current actually decreased. If this phenomenon causes a degradation in device performance for $t_{\mathrm{eq}} < 11-13$ Å, then tradeoffs may have to be considered for technologies using these materials.[23]

VI. MATERIALS PROPERTIES CONSIDERATIONS

All of the materials systems discussed earlier must meet a set of criteria to perform as successful gate dielectric. We now consider a summary of the appropriate materials properties for the selection of materials for gate dielectric applications.

A. Permittivity and barrier height

Selecting a gate dielectric with a higher permittivity than that of SiO_2 is clearly essential. For many simple oxides, permittivities have been measured on bulk samples and in some cases even on thin films, but for the more complex materials (more elemental constituents), the dielectric constants may not be as well known. Shannon[151] used a method involving the Classius–Mossoti equation for calculating ionic polarizabilities as a means to make predictions of permittivities for many dielectrics. Good agreement has been found for some materials, but there are also many cases of poor agreement between calculated and measured values. This discrepancy between calculation and measurement can be attributed to many factors, including film thickness, method of film deposition, and local electronic structure within the dielectrics. It is clear that much more experimental data is needed for measurements of dielectric constant for these high-κ dielectrics, particularly below the 100 Å thickness regime.

The required permittivity must be balanced, however, against the barrier height for the tunneling process. For electrons traveling from the Si substrate to the gate, this is the conduction band offset, $\Delta E_C \cong q[\chi-(\Phi_M-\Phi_B)]$; for electrons traveling from the gate to the Si substrate, this is Φ_B (see Fig. 4). This is because leakage current increases exponentially with decreasing barrier height (and thickness) for electron direct tunneling transport,[9,10] as shown in Eq. (12):

$$J_{DT}=\frac{A}{t_{\mathrm{diel}}^2}\exp\left(-2t_{\mathrm{diel}}\sqrt{\frac{2m^*q}{\hbar^2}\left\{\Phi_B-\frac{V_{\mathrm{diel}}}{2}\right\}}\right). \tag{12}$$

Here A is a constant, t_{diel} is the physical thickness of the dielectric, V_{diel} is the voltage drop across the dielectric, and m^* is the electron effective mass in the dielectric. For highly defective films which have electron trap energy levels in the insulator band gap, electron transport will instead be governed by a trap-assisted mechanism such as Frenkel–Poole emission or hopping conduction, as described by Eqs. (13) and (14), respectively,

$$J_{\mathrm{FP}}=E\exp\left(-\frac{q}{kT}\left\{\Phi_B-\sqrt{\frac{qE}{\pi\epsilon_i}}\right\}\right) \tag{13}$$

$$J_{\mathrm{hop}}=\frac{q^2l^2n^*\Gamma E}{kT}. \tag{14}$$

Here l is the interval of separation between adjacent hopping sites, n^* is the density of free electrons in the dielectric, and Γ is the mean hopping frequency. Leakage current also depends on other factors, including the structure of the layer, as discussed later in Sec. VI D. In order to obtain low leakage currents, it is desirable to find a gate dielectric that has a large ΔE_C value to Si and perhaps to other gate metals which may be used at some point. Reported values of ΔE_C for most dielectric-Si systems are scarce in the literature, but recent calculations by Robertson and Chen[152] indicate that many of the metal oxide and complex oxide materials, such as Ta_2O_5

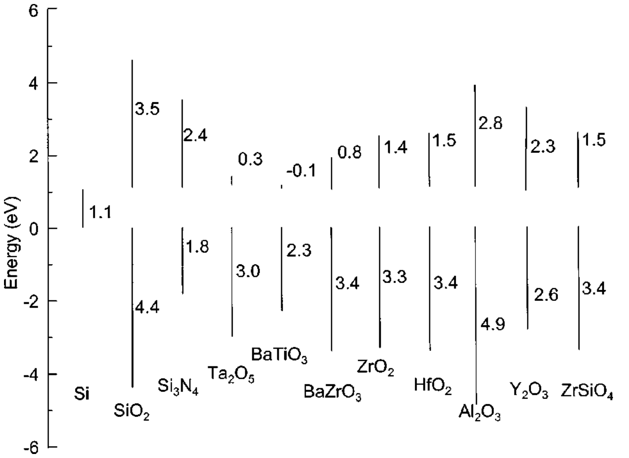

FIG. 25. Band offset calculations for a number of potential high-*k* gate dielectric materials (see Refs. 152 and 153).

and $SrTiO_3$, will have $\Delta E_C < 0.5\,\text{eV}$ on Si. On the other hand, an expanded study by Robertson *et al.*[153] includes many more high-κ dielectrics which are currently under investigation (see Fig. 25). Robertson found that $\Delta E_C \sim 2.3$–2.8 eV for Al_2O_3 and Y_2O_3, and $\Delta E_C \sim 1.5\,\text{eV}$ for ZrO_2 and $ZrSiO_4$, as illustrated in Fig. 25.[153]

These calculations are an important insight into predicting relevant barrier height (such as ΔE_C) values for many potential dielectrics based on the charge neutrality level of the material. If the experimental ΔE_C values for these oxides are even much less than 1.0 eV, it will likely preclude using these oxides in gate dielectric applications, since electron transport (either by thermal emission or tunneling) would lead to unacceptably high leakage currents. Since many potential gate dielectrics do not have reported ΔE_C values, the closest, most readily attainable indicator of band offset is the band gap (E_G) of the dielectric. A large E_G generally corresponds to a large ΔE_C (see Table I), but the band structure for some materials has a large valence band offset ΔE_V which constitutes most of the band gap of the dielectric. This uncertainty places more importance on predictive methods such as that demonstrated by Robertson.[152,153]

As noted before, it is extremely difficult to achieve the juxtaposition of these high-κ dielectrics on Si, as an SiO_2-like interface usually forms. This interface layer will of course alter the ΔE_C value of the system, and must be taken into consideration when comparing measured and calculated results.

Referring back to Table I, a list of several metal oxide (and nitride) systems, some of which have been investigated as gate dielectrics, compares values for κ and E_G, along with ΔE_C values on Si where available. For these high-κ materials, Table I indicates that E_G will be somewhat limited, since it can be seen that the dielectric constant and band gap of a given material generally exhibit an inverse relationship (although some materials have significant departures from this trend).

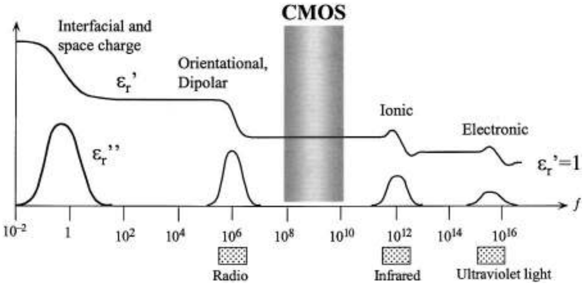

FIG. 26. The frequency dependence of the real (ε_r') and imaginary (ε_r'') parts of the dielectric permittivity. In CMOS devices, ionic and electronic contributions are present (see Ref. 154).

The detailed relationship between permittivity and band gap is by no means trivial, since many factors give rise to both properties. There are two main contributions of interest to the dielectric constant which give rise to the polarizability: electronic and ionic dipoles (the molecular dipole contribution is not relevant to this study). Figure 26 illustrates the frequency ranges where each contribution is important, and also highlights the current frequency range for CMOS operation (100 MHz–10 GHz).[154] In general, atoms with a large ionic radius (e.g., high atomic number) exhibit more electron dipole response to an external electric field, because there are more electrons to respond to the field (electron screening effects also play a role in this response). This electronic contribution tends to increase the permittivity for oxides of elements with higher atomic numbers.

The ionic contribution to the permittivity can be much larger than the electronic portion in cases such as perovskite crystals of $(Ba,Sr)TiO_3$ and $(Pb,Zr)TiO_3$ (which exhibit ferroelectric behavior below the Curie point). In these cases, Ti ions in unit cells throughout the crystal are uniformly displaced in response to an applied electric field (for the case of ferroelectric materials, the Ti ions reside in one of two stable, nonisosymmetric positions about the center of the Ti–O octahedra). This displacement of Ti ions causes an enormous polarization in the material, and thus can give rise to very large dielectric constants of 2000–3000. Since ions respond more slowly than electrons to an applied field, the ionic contribution begins to decrease at very high frequencies, in the infrared range of $\sim 10^{12}$ Hz, as shown in Fig. 26.

We note that some of the potential candidate materials may have other contributions to the permittivity, which do not exhibit the same phenomena as the perovskites. The addition of certain levels of network modifier ions (e.g., Zr or Hf) to materials such as SiO_2 can produce an increased dielectric constant even at low incorporation levels, through a discernable change in the bond order of the material. Lucovsky and Rayner[155] describe this effect for several experimental cases, where the localized bonding order, which changes with Zr (or Hf) concentration, results in substantial changes in the vibrational modes in the associated infrared spectra of the silicate. This effect can be further modeled to demonstrate that a transverse effective charge scales inversely with Zr (Hf) concentration, and the contribution of the vibrational mode to the dielectric constant is proportional to the square of this effective charge. As a result, an enhancement in the dielectric constant, over that expected by simple linear interpolation between $SiO_2(\kappa = 3.9)$ and $ZrSiO_4(\kappa = 12)$, may be expected at low concentrations for these materials (see bonding model in Fig. 20). Recent independent measurements of the enhanced dielectric constants have also been confirmed.[143(a),143(b),143(c)]

Another phenomenon which can contribute to an increased dielectric constant involving atom motion is soft phonons. This differs from a non-volatile polarization (e.g., as found in ferroelectrics), in that the atoms do not remain displaced after the electric field is removed (ferroelectricity is the limit at which certain cations remain displaced in one of two stable positions after the external field is removed). Rather, higher atomic number atoms can resonate in their bonding structures at low vibrational frequency modes, which produces a soft phonon (since it is at relatively low frequency). These phonons then make a lattice contribution to the overall polarizability, and therefore to the permittivity of the material. It appears that some of the high-κ dielectrics of interest exhibit the earlier-mentioned effects, leading to substantially increased permittivities even at compositions with low metal levels. By achieving good bonding models of these materials, some of these materials properties can begin to be better understood and predicted, as a way to optimize the choice of dielectric for this application.

It is therefore reasonable to expect any intrinsic contribution to the dielectric constant will be maintained for reasonable device operation frequencies, although other extrinsic factors such as defects may lead to an effective decrease in permittivity at much lower frequencies.

In contrast to the general trend of increasing permittivity with increasing atomic number for a given cation in a metal oxide, the band gap E_G of the metal oxides tends to *decrease* with increasing atomic number, particularly within a particular group in the periodic table. An intuitive explanation for this phenomenon is that the corresponding bonding and antibonding orbitals of the metal-oxygen atoms form a valence band and a conduction band, respectively. For the case of SiO_2, the σ bonds formed by the *sp* hybrid orbitals (which arise from Si s,p and O p orbitals) have a σ bonding orbital energy level and a higher σ^* antibonding orbital energy level. The energy separation between these levels defines a band gap, but this may or may not be the minimum band gap in the material. For even the simple case of SiO_2, where there are only *s* and *p* electron orbitals that are all filled during bonding, the oxygen electron lone pair energy level actually defines the valence band maximum (rather than the σ bonding energy level). The result is the often reported band gap of $SiO_2, E_G \sim 9$ eV. If the σ and σ^* bonds defined the valence band maximum and conduction band minimum, respectively, then the band gap of SiO_2 would be larger.

Thus for the transition metal oxides, which all have five *d* electron orbitals and other non-bonding *p* orbitals, the band gaps of these oxides can be significantly decreased by the presence of partially filled *d* orbitals, which have available states for electron occupancy. These orbital energy levels tend to lie within the gap defined by the σ and σ^* orbitals. The *d* orbital levels which lie within the gap defined by σ and σ^* levels are all therefore available for electron conduction at significantly lower energy levels than would be expected from σ and σ^* alone. This effect offers a somewhat intuitive explanation regarding the lower band gaps that are often observed for higher atomic number metal oxides, particularly for the transition metals. It is important to note that these partially filled and nonbonding levels are *not* the result of defects within the material, but rather are *intrinsic* to such atomic constituents where many orbitals are available for electron conduction. This general band gap reduction for higher-κ materials is a limitation that must be realized and expected when selecting a suitable high-κ gate dielectric.

In the cases of Ta_2O_5 and TiO_2, both materials have small E_G values and correspondingly small ΔE_C values. These small ΔE_C values directly correlate with high leakage currents for both materials, making pure Ta_2O_5 and TiO_2 unlikely choices for gate dielectrics. Table I also shows, however, that La_2O_3, HfO_2, and ZrO_2 offer relatively high values for both κ and E_G. It is important to note that all of the high-κ metal oxides listed in Table I, which includes those studied for gate dielectric applications, are crystalline at relatively low temperature (except Al_2O_3). The significance of the film structure will be further discussed in Sec. VI D.

Although many researchers originally assumed that selecting a dielectric with $\kappa > 25$ would be necessary to replace SiO_2, the more relevant consideration is whether the desired device performance (i.e., drive current) can be obtained at the prescribed operating voltage without producing unacceptable off-state (leakage) currents and reliability characteristics. It is therefore more appropriate to find a dielectric which provides even a moderate increase in κ, but which also produces a large tunneling barrier and high-quality interface to Si. With this in mind, if a single dielectric layer can be used, then even a material with $\kappa \sim 12-20$ will allow a physical dielectric thickness of 35–50 Å to obtain the t_{eq} values required for 0.1 μm CMOS and beyond.

B. Thermodynamic stability on Si

For all thin gate dielectrics, the interface with Si plays a key role, and in most cases is the dominant factor in determining the overall electrical properties. Most of the high-κ metal oxide systems investigated thus far have unstable interfaces with Si: that is, they react with Si under equilibrium conditions to form an undesirable interfacial layer. These materials therefore require an interfacial reaction barrier, as mentioned previously. Any ultrathin interfacial reaction barrier with $t_{\text{eq}} < 20$ Å will have the same quality, uniformity and reliability concerns as SiO_2 does in this thickness regime. This is especially true when the interface plays a determining role in the resulting electrical properties. It is important to understand the thermodynamics of these systems, and thereby attempt to control the interface with Si.

An important approach toward predicting and understanding the relative stability of a particular three-component system for device applications can be explained through ternary phase diagrams.[156,157] An analysis of the Gibbs free energies governing the relevant chemical reactions for the Ta–Si–O and Ti–Si–O ternary systems, as shown in Figs. 19(a) and 19(b), indicates that Ta_2O_5 and TiO_2 (or mixtures with Si), respectively, are not stable to SiO_2 formation when placed next to Si. Rather, the tie lines in (a) and (b) show that

Ta_2O_5 and TiO_2 on Si will tend to phase separate into SiO_2 and metal oxide (M_xO_y, M=metal), and possibly silicide (M_xSi_y) phases. The phase diagrams in Figs. 19(a) and 19(b) are shown for temperatures of 700–900 °C, but these relations are also valid at much lower temperatures. As mentioned in Sec. I, this instability to SiO_2 formation has been observed experimentally for both of these metal oxides,[75,103] which leads to the necessity for an additional interfacial layer.

In contrast to the Ta and Ti systems, the tie lines in the phase diagram for the Zr–Si–O system,[158] shown in Fig. 19(c), indicate that the metal oxide ZrO_2 and the compound silicate $ZrSiO_4$ will both be stable in direct contact with Si up to high temperature. The gray shaded area in Fig. 19(c) denotes a large phase field of $(ZrO_2)_x(SiO_2)_{1-x}$ compositions which are expected to be stable on Si up to high temperatures. Other $(ZrO_2)_x(SiO_2)_{1-x}$ compositions outside of the gray area are also stable on Si, but since it is desirable to prevent any Zr–Si bonding, film compositions within the gray area may be preferable. Furthermore, even within the shaded area, compositions with high O levels (closer to $ZrSiO_4$) are preferred because this will be more likely to prevent M–Si bond (and silicide phase) formation. In fact, the existence of this large phase field of stable $(ZrO_2)_x(SiO_2)_{1-x}$ compositions implies yet another potential advantage, in that the level of Zr (or Hf) incorporated into the silicate film could be gradually increased from one technology node to the next. This gradual shift in film composition could provide a continually higher κ value (up to a point), as is needed to meet device performance requirements.

This behavior is expected to be the same for the Hf–Si–O system based on coordination chemistry arguments. Although the thermodynamic information is incomplete for the Hf–Si–O system, the available data suggests that HfO_2 and $HfSiO_4$, as well as a large range of $(HfO_2)_x(SiO_2)_{1-x}$ compositions, will be stable in direct contact with Si up to high temperatures.[159–161] This fundamental difference from the Ta and Ti systems is extremely important, because it suggests that there is potential to control the dielectric-Si interface. While it is certainly the case that all deposition techniques of interest are done under nonequilibrium conditions, it is unlikely that a desirable, metastable phase, such as amorphous Ta_2O_5 (which can be obtained under nonequilibrium deposition conditions), will be maintained throughout all of the thermal cycling required for CMOS processing. Even in the case of process modifications such as a replacement gate flow,[69] which allows the high-temperature processing to be done before deposition of the final gate dielectric, the continued thermal cycling afterward is sufficient to result in poor electrical properties.[69] It is therefore important to select a materials system in which the desired final state is a stable one.

Figure 19(c) indicates that use of $(ZrO_2)_x(SiO_2)_{1-x}$ [and therefore $(HfO_2)_x(SiO_2)_{1-x}$] should allow for control of the Si interface, which may solve a key problem for the high-κ gate dielectric materials approaches. The κ values of $(HfO_2)_x(SiO_2)_{1-x}$ and $(ZrO_2)_x(SiO_2)_{1-x}$ are substantially lower than those of pure HfO_2 and ZrO_2, but as mentioned in Sec. VI A, this tradeoff for interfacial control will be acceptable as long as the resulting leakage currents are low enough.

C. Interface quality

A clear goal of any potential high-κ gate dielectric is to attain a sufficiently high-quality interface with the Si channel, as close as possible to that of SiO_2. It is difficult to imagine any material creating a better interface than that of SiO_2, since typical production SiO_2 gate dielectrics have a midgap interface state density $D_{it} \sim 2 \times 10^{10}$ states/cm^2. Most of the high-κ materials reported in this paper show $D_{it} \sim 10^{11}-10^{12}$ states/cm^2, and in addition exhibit a substantial flatband voltage shift $\Delta V_{FB} > 300$ mV (possibly from fixed charge density $\gtrsim 10^{12}/\text{cm}^2$ at the interface). It is crucial to understand the origin of the interface properties of any high-κ gate dielectric, so that an optimal high-κ–Si interface may be obtained.

Recent work by Lucovsky *et al.*[58] previously has shown that bonding constraints must also be considered at the Si-dielectric interface. It is shown empirically[58] that if the average number of bonds per atom $N_{av} > 3$, the interface defect density increases proportionally, with a corresponding degradation in device performance. Metal oxides which contain elements with a high coordination (such as Ta and Ti) will have a high N_{av}, and form an overconstrained interface with Si. Degradation in leakage current and electron channel mobility is therefore expected. Similarly, cations with low coordination (e.g., Ba, Ca) compared to that of Si lead to underconstrained systems in the corresponding metal oxides. These systems (metal oxides, ternary alloys, etc.) which are either over- or underconstrained with respect to SiO_2, lead to formation of a high density of electrical defects near the Si-dielectric interface, resulting in poor electrical properties.

These bonding arguments can be extended to silicide formation in the gate dielectric, or even to any M–Si bonding (not necessarily a full silicide phase). Any silicide bonding which forms near the channel interface (as may result when a metal oxide is placed in direct contact with, or near, Si), will tend to produce unfavorable bonding conditions, leading to poor leakage current and electron channel mobilities. In order to maintain a high-quality interface and channel mobility, it is expected to be important to have no metal oxide or silicide phases present at or near the channel interface.

It is interesting to consider the binary oxides which are more thermally stable than SiO_2, some of which are shown in Table I, including Al_2O_3, Y_2O_3, La_2O_3, ZrO_2, and HfO_2 (also covered in a thorough study by Hubbard and Schlom).[157] The requirement on bonding constraints mentioned earlier, however, may significantly reduce the list of materials. Although Al_2O_3 is a special case in the way that Al cations bond within the network (alternating Al–O tetrahedra and octahedra), Y_2O_3 and La_2O_3 are underconstrained and may therefore form very high defect densities at the channel interface. More highly coordinated metals such as Nb, V, and Mn are not only overconstrained, but each also has several different stable oxidation states, and will therefore lead to oxygen vacancies and electron trap sites. In addition to binary systems, pseudobinary systems are excellent candidates for gate dielectrics. Adding a third component to an alloy or network may favorably affect the material in terms of thermal stability, bonding constraints and in morphology.

Several simple oxides such as ZrO_2 and HfO_2 have been previously reported as having high oxygen diffusivities.[112] This is a serious concern regarding control of the interface once it is initially formed. Any annealing treatments which have an excess of oxygen present (either from the ambient or from a sidewall oxide, for example), will lead to rapid oxygen diffusion through the oxides, resulting in SiO_2 or SiO_2-containing interface layers. Although we have already shown that SiO_2 is an ideal interface with Si, an uncontrolled amount of SiO_2 formation at the interface will severely compromise the capacitance gain from any high-κ layers in the gate stack. Caution must therefore be used in assessing the interface stability of high-κ dielectrics, as resistance to oxygen diffusion in annealing ambients should be characterized. Another annealing ambient of concern is forming gas (typically 90% N_2:10%H_2), which is a standard final anneal in the CMOS process and is believed to passivate interfacial traps (dangling bonds) with hydrogen. Since many high-κ dielectrics will be reduced in the presence of H_2 (the Ti-containing perovskites are all severely reduced by even low temperature anneals in forming gas), high-κ gate dielectrics also need to be characterized with respect to the effect of anneals in reducing ambients.

The ideal gate dielectric stack may well turn out to have an interface comprised of several monolayers of Si–O (and possibly N) containing material, which could be a pseudobinary layer, at the channel interface. This layer could serve to preserve the critical, high-quality nature of the SiO_2 interface while providing a higher-κ value for that thin layer. The same pseudobinary material could also extend beyond the interface, or a different high-κ material could be used on top of the interfacial layer.

D. Film morphology

Most of the advanced gate dielectrics studied to date are either polycrystalline or single crystal films, but it is desirable to select a material which remains in a glassy phase (amorphous) throughout the necessary processing treatments. As shown in Table I, nearly all metal oxides of interest, with the exception of Al_2O_3, will form a polycrystalline film either during deposition or upon modest thermal treatments: HfO_2 and ZrO_2 are no exceptions. It is important to note, however, that the phases listed in Table I are bulk properties, and there will certainly be some suppression of crystallization for very thin films such as gate dielectrics, at temperatures where crystallization would otherwise be expected to occur. The extent of crystallization suppression for a given oxide will depend on composition and thermal processing.

Polycrystalline gate dielectrics may be problematic because grain boundaries serve as high-leakage paths, and this may lead to the need for an amorphous interfacial layer to reduce leakage current. In addition, grain size and orientation changes throughout a polycrystalline film can cause significant variations in κ, leading to irreproducible properties. The previously mentioned studies by Houssa *et al.*[118] and Perkins *et al.*,[127] however, appear to be counter-examples, as both reported very encouraging electrical properties for ALCVD ZrO_2. It should be noted, however, that both studies also used a gate dielectric stack, with the ZrO_2 film on top of an amorphous SiO_2 layer. It is unclear at this point to what extent the amorphous SiO_2 layer affords the encouraging electrical properties, but this issue will become important, as the SiO_2 layer presents a limit to the minimum achievable t_{eq} value for these structures.

Work by van Dover,[109] as previously discussed, used lanthanide dopants in TiO_x films to create and maintain an amorphous film for capacitor applications, even though TiO_2 is typically crystalline even at low temperatures (see Table I). Very encouraging results were obtained, as both high permittivities and low leakage currents were achievable with Nd, Tb, and Dy dopants. Although these particular films are not stable on Si, similar approaches may be useful for gate dielectrics.

Single crystal oxides grown by MBE methods[82,98] can in principle avoid grain boundaries while providing a good interface, but these materials also require submonolayer deposition control, which may only be obtainable by MBE approaches (see Sec. VI F for further discussion on the MBE deposition method). Kwo *et al.*[98] formed capacitors with single-domain, crystalline Gd_2O_3 films on Si by MBE, which had no apparent interfacial layer according to infrared absorption spectroscopy (not shown in the article). The leakage current for these films was 10^{-3} A/cm^2 at 1 V bias, and $C-V$ analysis showed $\kappa \sim 14$, while some frequency dependence was observed and the permittivity decreased with decreasing film thickness.[98] For perovskite materials such as $SrTiO_3$, where the structure consists of alternating SrO and TiO_2 planes, each single atomic-height step edge (which always exist on the surface of Si wafers) may possibly serve as a nucleation site for an antiphase boundary and possibly a grain boundary. As will be discussed in Sec. VI E, any of these dielectrics which require interfacial layers on the Si channel (to avoid reaction between the high-κ material and Si) will also require metal gate, or perhaps a buffer layer at the poly-Si gate interface. In contrast, amorphous films will exhibit isotropic electrical properties, will not suffer from grain boundaries, and can easily be deposited by manufacturable techniques.

Given the concerns regarding polycrystalline and single crystal films, it appears that an amorphous film structure is the ideal one for the gate dielectric (although some initial electrical results for a polycrystalline ZrO_2 layer on an amorphous SiO_2-containing layer look encouraging). This is yet another clear virtue of SiO_2. The Zr–Si–O system in Fig. 17(c), and analogously the Hf–Si–O system, indicates that the only stable ternary crystalline compound is $ZrSiO_4$ (or $HfSiO_4$). While it is possible that other ternary crystalline compounds do exist, it is unlikely. Bulk thermodynamics therefore suggests that a phase field [gray shaded area in Fig. 17(c)] exists for relatively low levels of Zr or Hf in which a $(ZrO_2)_x(SiO_2)_{1-x}$ or $(HfO_2)_x(SiO_2)_{1-x}$ composition can be obtained which will remain amorphous and stable on Si up to high temperatures, without phase separating into crystalline $MSiO_4$ or MO_2 and SiO_2.

E. Gate compatibility

A significant issue for integrating any advanced gate dielectric into standard CMOS is that the dielectric should be compatible with Si-based gates, rather than *require* a metal gate. Si-based gates are desirable because dopant implant conditions can be tuned to create the desired threshold voltage V_T for both nMOS and pMOS, and the process integration schemes are well established in industry. Nearly all of the potential advanced gate dielectrics investigated to this point, however, require metal gates. This is expected because the same instability with Si, mentioned in Sec. VI B, will exist at both the channel and the poly-Si gate interfaces.

Specifically, metal gates such as TiN and Pt have been used with most of the high-κ gate dielectrics mentioned above to prevent reaction at the gate interface. Attempts have been made to use doped poly-Si gates with Ta_2O_5, by depositing a CVD SiO_2 reaction barrier, which is of lower electrical quality than thermal oxide, on top of the Ta_2O_5 layer.[71] The presence of SiO_2 at both the channel and gate interface predictably limited the device performance, and the lowest obtainable oxide equivalent was $t_{\mathrm{eq}} = 23\text{Å}$. Even initial attempts to use poly-Si gates with ZrO_2[116,120] have been unsuccessful (recent work, however, has shown improved poly-Si stability on HfO_2),[125,126] as reaction layers have been observed at the interface. On the other hand, another benefit of a pseudobinary system such as a silicate[116,137] in contact with a poly-Si electrode is more inherent stability, since a significant amount of Si is already contained within the dielectric. Al_2O_3 has been shown to be stable with respect to reaction with the poly-Si gates throughout typical CMOS processing, as expected.[90–92] As mentioned previously in Sec. V C 1, however, both boron and phosphorous dopant diffusion have been observed with Al_2O_3 gate dielectrics, which cause significant, undesired shifts of V_{FB} and V_T values.[90,92] These results place more emphasis on the need to better understand dopant diffusion through all potential high-κ candidates.

Metal gates are very desirable for eliminating dopant depletion effects and sheet resistance constraints. In addition, use of metal gates in a replacement gate process[69,139] can lower the required thermal budget by eliminating the need for a dopant activation anneal in the poly-Si electrode. There are two basic approaches toward achieving successful insertion of metal electrodes: a single midgap metal or two separate metals. The energy diagrams associated with these two approaches are shown in Fig. 27.

The first approach is to use a metal (such as TiN) that has a work function that places its Fermi level at the midgap of the Si substrate, as shown in Fig. 27(a). These are generally referred to as ''midgap metals.'' The main advantage of employing a midgap metal arises from a symmetrical V_T value for both nMOS and pMOS, because by definition the same energy difference exists between the metal Fermi level and the conduction and valence bands of Si. This affords a simpler CMOS processing scheme, since only one mask and one metal would be required for the gate electrode (no ion implantation step would be required).

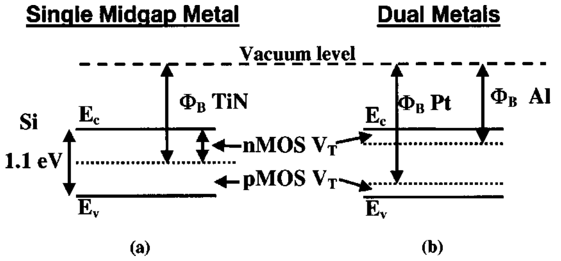

FIG. 27. Energy diagrams of threshold voltages for nMOS and pMOS devices using (a) midgap metal gates and (b) dual metal gates.

For the case of sub-0.13 μm bulk CMOS devices, however, a major drawback is that the band gap of Si is fixed at 1.1 eV, thus the threshold voltage for any midgap metal on Si will be ~0.5 V for both nMOS and pMOS. Since voltage supplies are expected to be $\leqslant 1.0$ V for sub-0.13 μm CMOS technology, $V_T \sim 0.5\,\mathrm{V}$ is much too large, as it would be difficult to turn on the device. Typical threshold voltages for these devices are expected to be 0.2–0.3 V. Compensation implants can be made in the channel to lower the V_T, but other concerns then arise regarding increased impurity ion scattering, which would degrade the channel carrier mobility. Furthermore, midgap work function metal gate systems have been predicted not to provide a performance improvement worthy of the added process complexity to replace Si-based gates.[162]

The second main approach toward metal electrodes involves two separate metals, one for pMOS and one for nMOS devices. As shown in Fig. 27(b), two metals could be chosen by their work functions, Φ_M, such that their Fermi levels line up favorably with the conduction and valence bands of Si, respectively. In the ideal case depicted in Fig. 27(b), the Φ_M value of Al could achieve $V_T \sim 0.2\,\mathrm{V}$ for nMOS, while the higher Φ_M value of Pt could achieve $V_T \sim 0.2\,\mathrm{V}$ for pMOS. In practice, Al is not a feasible electrode metal because it will reduce nearly any oxide gate dielectric to form an Al_2O_3-containing interface layer. Other metals with relatively low work functions, such as Ta and TaN, however, are feasible gate metals for nMOS. Similarly for pMOS, Pt is not a practical choice for the gate metal, since it is not easily processed, does not adhere well to most dielectrics, and is expensive. Other elemental metals with high Φ_M values such as Au are also not practical, for the same reasons as for Pt.

As an alternative to elemental metals, conducting metal oxides such as IrO_2 and RuO_2, which have been studied for years in DRAM applications,[1] can provide high Φ_M values in addition to affording the use of standard etching and processing techniques. Alloys of these and similar conducting oxides can also potentially be fabricated to achieve a desired work function. Regarding potential gate electrodes for pMOS devices, Zhong *et al.*[163,164] made initial measurements of the important properties of RuO_2, including thermal stability up to 800 °C, a low resistivity of 65 $\mu\Omega$ cm, and a measured work function of $\Phi_M = 5.1\,\mathrm{eV}$. All of the measurements were made on capacitors with both SiO_2 and Zr-silicate dielectrics. More work must be done to better understand alternative metal electrodes, both for midgap metal and dual metal approaches, as a means to alleviate potentially limiting properties of doped poly-Si.

It should also be noted that for the scaling trend to sub-0.1 μm CMOS, other efforts are focused on using poly-$Si_{1-x}Ge_x$ gates for achieving higher boron activation levels[165,166] and therefore better performance in *p*MOS devices, and potentially better performance in *n*MOS devices as well.[166] It is therefore still desirable to employ a gate dielectric which will be compatible in direct contact with Si-based gates. The pseudobinary alloys mentioned in Sec. V C 3 are predicted to be stable next to Si-based gates *as well as* metal gates. Since doped poly-Si is the incumbent gate electrode material, however, dopant diffusion studies must be carried out to determine how dopants in poly-Si diffuse through any potential high-κ gate material.

For CMOS scaling in the longer term, however, current roadmap predictions indicate that poly-Si gate technology will likely be phased out beyond the 70 nm node, after which a metal gate substitute appears to be required.[5] It is therefore also desirable to focus efforts on dielectric materials systems which are compatible with potential metal gate materials. A key issue for gate electrode materials research will be the control of the gate electrode work function (Fermi level) after CMOS processing.

F. Process compatibility

A crucial factor in determining the final film quality and properties is the method by which the dielectrics are deposited in a fabrication process. The deposition process for the dielectric must be compatible with current or expected CMOS processing, cost, and throughput. Since all of the feasible deposition techniques available occur under nonequilibrium conditions, it is certainly possible to obtain properties different from those expected under equilibrium conditions. It is therefore important to consider the various methods for depositing the gate dielectrics, and the following techniques will be discussed here: physical vapor deposition (PVD) (e.g., sputtering and evaporation), CVD, ALCVD, and MBE.

PVD methods have provided a convenient means to evaluate materials systems for alternate dielectric applications. The damage inherent in a sputter PVD process, however, results in surface damage and thereby creates unwanted interfacial states. Additionally, device morphology inherent to the scaling process generally rules out such line-of-sight PVD deposition approaches. For this reason, CVD methods have proven to be quite successful in providing uniform coverage over complicated device topologies.

As pointed out, the reaction kinetics associated with film CVD deposition require careful attention in order to control interfacial layer formation. The precursor employed in the deposition process must also be tailored to avoid unwanted impurities in the film as well as permit useful final compositions in the dielectric film. Indeed, a graded composition for dielectric films may be a key requirement in order to control interface state formation to a level comparable to SiO_2. The recent application of ALCVD methods for depositing Al_2O_3,[87] ZrO_2,[113,118,120,127] and HfO_2[120] appears to provide much promise, where self-limiting chemistries are employed to control film formation in a layer-by-layer fashion. As discussed before, attention to the surface preparation and the resultant chemistry must be carefully considered.

In addition to the examples discussed above which employ PVD and CVD methods, very high-κ dielectrics, such as $SrTiO_3$ have been deposited directly on Si using MBE methods (Fig. 11). An equivalent oxide thickness of less than 10 Å was reported (physical thickness 110 Å) with dramatic improvements in transistor performance (Fig. 10). More recently, crystalline ZrO_2 (stabilized by Y_2O_3) has also been examined by MBE methods.[167] However, a manufacturable scaled CMOS process incorporating MBE methods, with the inherent poor throughput relative to present Si-based fabrication operations, remains a clear challenge. Further advances in ALCVD approaches, however, may make such polycrystalline dielectrics a reality in the manufacturing environment.

G. Reliability

As previously discussed in Sec. IV B, the electrical reliability of a new gate dielectric must also be considered critical for application in CMOS technology. The determination of whether or not a high-κ dielectric satisfies the strict reliability criteria requires a well-characterized materials system—a prospect not yet available for the alternate dielectric materials considered here. The nuances of (1) the dependence of voltage acceleration extrapolation on dielectric thickness and (2) the improvement of reliability projection arising from improved oxide thickness uniformity, both discussed in Sec. IV B, have only recently become understood, despite decades of research on SiO_2. This further emphasizes the importance and urgency to investigate the reliability characteristics of alternative dielectrics, as these materials are sure to exhibit subtleties in reliability that differ from those of SiO_2.

This being stated, some preliminary projections for reliability, as determined by stess-induced leakage current (SILC), time-dependent dielectric breakdown, and mean time to failure measurements, appear to be encouraging for Al_2O_3[89–91] and HfO_2 films.[122,124] Similar preliminary measurements of ZrO_2 films appear to be somewhat mixed, as both very promising[114] and less encouraging[139] results have been reported. Despite excellent ten-year reliability projections at high applied voltages of 2.0–2.5 V reported for these materials,[89–91,114,122,124] it is essential to carry out proper area scaling conversion, to account for the difference between individual capacitors in these cases, and a full chip area in a realistic situation.[32] Proper area scaling can significantly degrade the ten-year reliability projection. Results of reliability investigations for pseudobinary alloys are not yet published. It is clear, however, that the determination of the preferred dielectric constituent composition has yet to be completed thus making even initial reliability extrapolations problematic. Moreover, recent lessons from the scaling changes associated with ultrathin SiO_2 may come into play with these new materials. Clearly, more work in production-worthy development fabrication facilities will be required to critically address these issues.

VII. CONCLUSIONS

A review of the state-of-the-art of the ongoing research for an alternative gate dielectric to SiO_2 for Si-based CMOS was presented. Key materials considerations discussed include: (a) permittivity, band gap, and barrier height, (b) thermodynamic stability on Si, (c) interface quality, (d) film morphology, (e) gate compatibility, (f) process compatibility, and (g) reliability. A material which satisfies all of these considerations has yet to be determined, but several promising candidates have been identified. The pseudobinary materials systems, however, currently offer the most promise toward the ultimate goal of integrating a gate dielectric into future CMOS technology nodes. Whether pseudobinary materials are best-suited as an interfacial layer to Si with a higher-κ layer on top, or are able to comprise the entire gate dielectric stack remains to be determined. Extensive research and development efforts are underway to narrow the field of candidates further.

Any potential alternative dielectric faces several fundamental concerns, in addition to those outlined earlier, regarding CMOS scaling which must be better understood and overcome before successful insertion of a new material will occur. These fundamental limitations include fixed charge, dopant depletion in the poly-Si gate electrode, and an increasing electric field in the channel region, which decreases device performance. Furthermore, dopant diffusion characteristics, failure mechanisms, and reliability of any potential high-κ dielectric need to be understood. The semiconductor industry has enjoyed the excellent reliability characteristics of SiO_2 for many years, but any new material will certainly exhibit different behavior, which may or may not have deleterious effects on device performance. The stringent requirements for ten year reliability of CMOS devices will be a demanding challenge on any high-κ materials candidates.

An even greater challenge is the adoption of a new candidate in the time frame required by the industry roadmap (~4–5 years) in order to maintain cost/performance trends. The industry has enjoyed the fruits of over 30 years of research and development on the SiO_2/Si materials system—a fact not always recognized in technology development planning. A new generation of scientists and engineers will be challenged by not only integrating these new materials in a timely manner, but also by avoiding the mistakes of the past. Opportunities to revolutionize an industry like these are indeed rare!

ACKNOWLEDGMENTS

The authors acknowledge useful discussions with Scott Summerfelt of Texas Instruments in the early work on Hf and Zr silicates. The authors also gratefully acknowledge the discussions and data provided by Professor Bruce Gnade of University of North Texas, and Professor Gerry Lucovsky and Professor Greg Parsons of North Carolina State University prior to publication. The authors thank Don Monroe and Jack Hergenrother of Agere Systems (formerly of Bell Laboratories), David Muller of Bell Laboratories, Lucent Technologies, and Luigi Colombo of Texas Instruments, for many useful discussions. The authors are grateful to Charles Wilk for several insightful suggestions regarding the manuscript. One of the authors (R.M.W.) gratefully acknowledges the support of the Texas Advanced Technology Program, the Semiconductor Research Corporation, and the Defense Advanced Research Projects Agency in the preparation of this review.

[1] T. Hori, *Gate Dielectrics and MOS ULSIs* (Springer, New York, 1997).
[2] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, J. IEEE **SC–9**, 256 (1974).
[3] G. Baccarani, M. R. Wordeman, and R. H. Dennard, IEEE Trans. Electron Devices **31**, 452 (1984).
[4] P. A. Packan, Science **285**, 2079 (1999).
[5] See *The International Technology Roadmap for Semiconductors*, Semiconductor Industry Association; see also http: //public.itrs.net/ for the most recent updates (1999).
[6] R. Rios and N. D. Arora, Tech. Dig. Int. Electron Devices Meet. **1994**, 613 (1994).
[7] The relative permittivity of a material is often given by ϵ or ϵ_r, such as with the expression $C=\epsilon\epsilon_0 A/t$. The relation between κ and ϵ varies depending on the choice of units (e.g., when $\epsilon_0=1$), but since it is always the case that $\kappa\propto\epsilon$, in this review we use the definition $\kappa=\epsilon$.
[8] A. Chatterjee, M. Rodder, and I-C. Chen, IEEE Trans. Electron Devices **45**, 1246 (1998).
[9] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981).
[10] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982).
[11] G. Lucovsky (private communication).
[12] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, Nature (London) **399**, 758 (1999); D. A. Muller, ''Characterization and Metrology for ULSI Technology,'' Intl. Conf. 2000, edited by D. G. Seiler, A. C. Diebold, T. J. Shaffner, R. McDonald, W. M. Bullis, P. J. Smith, and E. M. Secula, p. 500.
[13] S. Tang, R. M. Wallace, A. Seabaugh, and D. King-Smith, Appl. Surf. Sci. **135**, 137 (1998).
[14] J. B. Neaton, D. A. Muller, and N. W. Ashcroft, Phys. Rev. Lett. **85**, 1298 (2000).
[15] A. A. Demkov and O. F. Sankey, Phys. Rev. Lett. **83**, 2038 (1999).
[16] J. L. Alay and M. Hirose, J. Appl. Phys. **81**, 1606 (1997).
[17] B. Brar, G. D. Wilk, and A. C. Seabaugh, Appl. Phys. Lett. **69**, 2728 (1996).
[18] Z. H. Lu, J. P. McCaffrey, B. Brar, G. D. Wilk, R. M. Wallace, L. C. Feldman, and S. P. Tay, Appl. Phys. Lett. **71**, 2764 (1997).
[19] H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S.-I. Nakamura, M. Saito, and H. Iwai, IEEE Trans. Electron Devices **43**, 1233 (1996).
[20] G. Timp, A. Agarwal, F. H. Baumann, T. Boone, M. Buonanno, R. Cirelli, V. Donnelly, M. Foad, D. Grant, M. Green *et al.*, Tech. Dig. Int. Electron Devices Meet. **1997**, p. 930.
[21] G. Timp, K. K. Bourdelle, J. E. Bower, F. H. Baumann, T. Boone, R. Cirelli, K. Evans-Lutterodt, J. Garno, A. Ghetti, H. Gossmann *et al.*, Tech. Dig. Int. Electron Devices Meet. **1998**, p. 615.
[22] G. Timp, J. Bude, K. K. Bourdelle, J. Garno, A. Ghetti, H. Gossmann, M. Green, G. Forsyth, Y. Kim, R. Kleimann *et al.*, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 55.
[23] B. Yu, H. Wang, C. Riccobene, Q. Xiang, and M.-R. Lin, VLSI Tech. Dig. **2000**, p. 90.
[24] T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr, Tech. Dig. VLSI Symp. **2000**, p. 174; R. R. Chau, J. Kavalieros, B. Roberds, R. Schenker, D. Lionberger, D. Barlage, B. Doyle, R. Arghavani, A. Murthy, and G. Dewey, Tech. Int. Electron Devices Meet. **2000**, p. 45.
[25] B. E. Weir, P. J. Silverman, M. A. Alam, F. Baumann, D. Monroe, A. Ghetti, J. D. Bude, G. L. Timp, A. Hamad, T. M. Oberdick *et al.*, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 437.
[26] H. Iwai, H. S. Momose, and S. Ohmi, Proc.-Electrochem. Soc. **2000-2**, p. 3.
[27] J. H. Stathis and D. J. DiMaria, Tech. Dig. Int. Electron Devices Meet. **1998**, p. 167; Appl. Phys. Lett. **71**, 3230 (1997).
[28] C. Hu, Tech. Dig. Int. Electron Devices Meet. **1996**, p. 319.
[29] Y. Taur, D. A. Buchanan, W. Chen, D. J. Frank, K. E. Ismail, S.-H. Lo,

G. A. Sai-Halasz, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind, and H.-S. P. Wong, Proc. IEEE **85**, 486 (1997).

[30] M. Alam, B. Weir, P. Silverman, J. Bude, A. Ghetti, Y. Ma, M. M. Brown, D. Hwang, and A. Hamad, Proc.-Electrochem. Soc. **2000-2**, p. 365.

[31] J. H. Stahis, A. Vayshenker, P. R. Varekamp, E. Y. Yu, C. Montrose, J. McKenna, D. J. DiMaria, L.-K. Han, E. Cartier, R. A. Wachnik, and B. P. Linder, Tech. Dig. VLSI Symp. **2000**, p. 94.

[32] D. A. Buchanan, IBM J. Res. Dev. **43**, 245 (1999); J. H. Stathis and D. J. DiMaria, Microelectron. Eng. **48**, 395 (1999).

[33] P. E. Nicollian, W. R. Hunter, and J. Hu, Tech. Dig. IRPS Symp. 2000, p. 7.

[34] E. Wu, E. Nowack, L. Han, D. Dufresne, and W. Abadeer, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 441.

[35] B. E. Weir, M. A. Alam, J. D. Bude, P. J. Silverman, A. Ghetti, F. Baumann, P. Diodato, D. Monroe, T. Sorsch, g. L. Timp, Y. Ma, M. M. Brown *et al.*, Semicond. Sci. Technol. **15**, 455 (2000).

[36] M. Alam, J. Bude, and A. Ghetti, Tech. Dig. IRPS Symp., 2000, p. 21.

[37] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, Tech. Dig. Int. Electron Devices Meet. **1995**, p. 863.

[38] R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, Ph. Roussel, and H. E. Maes, IEEE Trans. Electron Devices **45**, 904 (1998).

[39] M. A. Alam, B. E. Weir, J. D. Bude, P. J. Silverman, and D. Monroe, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 449.

[40] D. J. DiMaria and J. H. Stathis, Appl. Phys. Lett. **74**, 1752 (1999); Proc.-Electrochem. Soc. **2000-2**, p. 33.

[41] J. H. Stathis, J. Appl. Phys. **86**, 5757 (1999).

[42] A. I. Kingon, J. P. Maria, and S. K. Streiffer, Nature (London) **406**, 1032 (2000).

[43] M. Cao, P. V. Voorde, M. Cox, and W. Greene, IEEE Electron Device Lett. **19**, 291 (1998).

[44] J. Sapjeta, T. Boone, J. M. Rosamilia, P. J. Silverman, T. W. Sorsch, G. Timp, and B. E. Weir, Mater. Res. Soc. Symp. Proc. **477**, 203 (1997).

[45] G. D. Wilk, Y. Wei, H. Edwards, and R. M. Wallace, Appl. Phys. Lett. **70**, 2288 (1997).

[46] G. Lucovsky, A. Banerjee, B. Hinds, B. Clafflin, K. Koh, and H. Yang, J. Vac. Sci. Technol. B **15**, 1074 (1997).

[47] G. D. Wilk and B. Brar, IEEE Electron Device Lett. **20**, 132 (1999).

[48] Y. Wei, R. M. Wallace, and A. C. Seabaugh, Appl. Phys. Lett. **69**, 1270 (1996).

[49] Y. Wei, R. M. Wallace, and A. C. Seabaugh, J. Appl. Phys. **81**, 6415 (1997).

[50] S. Tang, Y. Wei, and R. M. Wallace, Surf. Sci. Lett. **387**, L1057 (1997).

[51] S. V. Hattangady, R. Kraft, D. T. Grider, M. A. Douglas, G. A. Brown, P. A. Tiner, J. W. Kuehne, P. E. Nicollian, and M. F. Pas, Tech. Dig. Int. Electron Devices Meet. **1996**, p. 495.

[52] Y. Wu and G. Lucovsky, IEEE Electron Device Lett. **19**, 367 (1998).

[53] X. W. Wang, Y. Shi, and T. P. Ma, Tech. Dig. VLSI Symp. 1995, p. 109.

[54] K. A. Ellis and R. A. Buhrman, Appl. Phys. Lett. **74**, 967 (1999).

[55] K. A. Ellis and R. A. Buhrman, J. Electrochem. Soc. **145**, 2068 (1998).

[56] H. Yang and G. Lucovsky, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 245.

[57] V. Misra, H. Lazar, M. Kulkarni, Z. Wang, G. Lucovsky, and J. R. Hauser, Mater. Res. Soc. Symp. Proc. **567**, 89 (1999).

[58] G. Lucovsky, Y. Wu, H. Niimi, V. Misra, and J. C. Phillips, Appl. Phys. Lett. **74**, 2005 (1999).

[59] 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. **1998**, p. 373.

[60] X. Guo and T. P. Ma, IEEE Electron Device Lett. **19**, 207 (1998).

[61] S. Song, W. S. Kim, J. S. Lee, T. H. Choe, J. K. Choi, M. S. Kang, U. I. Chung, N. I. Lee, K. Fujihara, H. K. Kang *et al.*, Tech. Dig. VLSI Symp. **2000**, p. 190.

[62] J. M. Hergenrother, D. Monroe, F. P. Klemens, A. Kornblit, G. R. Weber, W. M. Mansfield, M. R. Baker, F. H. Baumann, K. J. Bolan, J. E. Bower *et al.*, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 75.

[63] S.-H. Oh, J. M. Hergenrother, T. Nigam, D. Monroe, F. P. Klemens, A. Kornblit, W. M. Mansfield, M. R. Baker, D. L. Barr, F. H. Baumann *et al.*, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 65.

[64] L. Risch, W. H. Krautschneider, F. Hofmann, H. Schäfer, T. Aeugle, and W. Rösner, IEEE Trans. Electron Devices **43**, 1495 (1996).

[65] C. P. Auth and J. D. Plummer, IEEE Device Res. Conf. Tech. Dig. **1996**, p. 108.

[66] H.-S. P. Wong, K. K. Chan, and Y. Taur, Tech. Dig. Int. Electron Devices Meet. **1997**, p. 427.

[67] D. Hisamoto, T. Kaga, and E. Takeda, IEEE Trans. Electron Devices **38**, 1419 (1991).

[68] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y. K. Choi, K. Asano *et al.*, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 67.

[69] A. Chatterjee, R. A. Chapman, K. Joyner, M. Otobe, S. Hattangady, M. Bevan, G. A. Brown, H. Yang, Q. He, D. Rogers *et al.*, Tech. Dig. Int. Electron Devices Meet. **1998**, p. 777.

[70] P. K. Roy and I. C. Kizilyalli, Appl. Phys. Lett. **72**, 2835 (1998).

[71] I. C. Kizilyalli, R. Y. S. Huang, and P. K. Roy, IEEE Electron Device Lett. **19**, 423 (1998).

[72] Q. Lu, D. Park, A. Kalnitsky, C. Chang, C. C. Cheng, S. P. Tay, Y.-C. King, T.-J. King, C. Hu, IEEE Electron Device Lett. **19**, 341 (1998).

[73] D. Park, Y.-C. King, Q. Lu, T.-J. King, C. Hu, A. Kalnitsky, S.-P. Tay, and C.-C. Cheng, IEEE Electron Device Lett. **19**, 441 (1998).

[74] H. F. Luan, B. Z. Wu, L. G. Kang, B. Y. Kim, R. Vrtis, D. Roberts, and D. L. Kwong, Tech. Dig. Int. Electron Devices Meet. **1998**, p. 609.

[75] G. B. Alers, D. J. Werder, Y. Chabal, H. C. Lu, E. P. Gusev, E. Garfunkel, T. Gustafsson, and R. S. Urdahl, Appl. Phys. Lett. **73**, 1517 (1998).

[76] Y. Nishioka, N. Homma, H. Shinriki, K. Mukai, K. Yamaguchi, A. Yuchida, K. Higeta, and K. Ogiue, IEEE Trans. Electron Devices **ED-34**, 1957 (1987); Y. Nishioka, H. Shinriki, and K. Mukai, J. Appl. Phys. **61**, 2335 (1987).

[77] C. Chaneliere, J. L. Autran, R. A. B. Devine, and B. Balland, Mater. Sci. Eng., R. **22**, 269 (1998).

[78] R. M. Fleming, D. V. Lang, C. D. W. Jones, M. L. Steigerwald, D. W. Murphy, G. B. Alers, Y. H. Wong, R. B. van Dover, J. R. Kwo, and A. M. Sergent, J. Appl. Phys. **88**, 850 (2000).

[79] R. A. McKee, F. J. Walker, and M. F. Chisholm, Phys. Rev. Lett. **81**, 3014 (1998).

[80] R. A. McKee, F. J. Walker, and M. F. Chisholm, Mater. Res. Soc. Symp. Proc. **567**, 415 (1999).

[81] Z. Yu, R. Droopad, J. Ramdani, J. A. Curless, C. D. Overgaard, J. M. Finder, K. W. Eisenbeiser, J. Wang, J. A. Hallmark, and W. J. Ooms, Mater. Res. Soc. Symp. Proc. **567**, 427 (1999).

[82] K. Eisenbeiser, J. M. Finder, Z. Yu, J. Ramdani, J. A. Curless, J. A. Hallmark, R. Droopad, W. J. Ooms, L. Salem, S. Bradshaw, and C. D. Overgaard, Appl. Phys. Lett. **76**, 1324 (2000).

[83] Z. Yu, J. Ramdani, J. A. Curless, J. M. Finder, C. D. Overgaard, R. Droopad, K. W. Eisenbeiser, J. A. Hallmark, W. J. Ooms, J. R. Conner, and V. S. Kaushik, J. Vac. Sci. Technol. B **18**, 1653 (2000).

[84] Z. Yu, J. Ramdani, J. A. Curless, C. D. Overgaard, J. M. Finder, R. Droopad, K. W. Eisenbeiser, J. A. Hallmark, W. J. Ooms, and V. S. Kaushik, J. Vac. Sci. Technol. B **18**, 2139 (2000).

[85] T. M. Klein, D. Niu, W. S. Epling, W. Li, D. M. Maher, C. C. Hobbs, R. I. Hedge, I. J. R. Baumvol, and G. N. Parsons, Appl. Phys. Lett. **75**, 4001 (1999).

[86] L. Manchanda, W. H. Lee, J. E. Bower, F. H. Baumann, W. L. Brown, C. J. Case, R. C. Keller, Y. O. Kim, E. J. Laskowski, M. D. Morris *et al.*, Tech. Dig. Int. Electron Devices Meet. **1998**, p. 605.

[87] E. P. Gusev, M. Copel, E. Cartier, I. J. R. Baumvol, C. Krug, and M. A. Gribelyuk, Appl. Phys. Lett. **76**, 176 (2000).

[88] A. Chin, C. C. Liao, C. H. Liu, W. J. Chen, and C. Tsai, Tech. Dig. VLSI Symp. 1999, p. 135.

[89] A. Chin, Y. H. Wu, S. B. Chen, C. C. Liao, and W. J. Chen, Tech. Dig. VLSI Symp. 2000, p. 16.

[90] D.-G. Park, H.-J. Cho, C. Lim, I.-S. Yeo, J.-S. Roh, C.-T. Kim, and J.-M. Hwang, Tech. Dig. VLSI Symp. 2000, p. 46.

[91] D. A. Buchanan, E. P. Gusev, E. Cartier, H. Okorn-Schmidt, K. Rim, M. A. Gribelyuk, A. Mocuta, A. Ajmera, M. Copel, S. Guha *et al.*, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 223.

[92] J. H. Lee, K. Koh, N. I. Lee, M. H. Cho, Y. K. Kim, J. S. Jeon, K. H. Cho, H. S. Shin, M. H. Kim, K. Fujihara *et al.*, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 645.

[93] J. Kolodzey, E. A. Chowdhury, G. Qui, J. Olowolafe, C. P. Swann, K. M. Unruh, J. Suehle, R. G. Wilson, and J. M. Zavada, Appl. Phys. Lett. **71**, 3802 (1997).

[94] See, for example, R. M. Wallace and Y. Wei, J. Vac. Sci. Technol. B **17**, 970 (1999), and references therein.

[95] H. B. Michaelson, J. Appl. Phys. **48**, 4729 (1977).

[96] L. Manchanda and M. Gurvitch, IEEE Electron Device Lett. **9**, 180 (1988).

[97] M. Gurvitch, L. Manchanda, and J. M. Gibson, Appl. Phys. Lett. **51**, 919 (1987).

[98] J. Kwo, M. Hong, A. R. Kortan, K. T. Queeney, Y. J. Chabal, J. P. Mannaerts, T. Boone, J. J. Krajewski, A. M. Sergent, and J. M. Rosamilia, Appl. Phys. Lett. **77**, 130 (2000).

[99] S. Guha, E. Cartier, M. A. Gribelyuk, N. A. Borjarczuk, and M. A. Copel, Appl. Phys. Lett. **77**, 2710 (2000).

[100] J. J. Chambers and G. N. Parsons, Appl. Phys. Lett. **77**, 2385 (2000).

[101] H. J. Osten, J. P. Liu, P. Gaworzewski, E. Bugiel, and P. Zaumseil, Tech. Dig. Int. Electron Devices Meet. **2000**, 653 (2000).

[102] S. A. Campbell, D. C. Gilmer, X. Wang, M. T. Hsich, H. S. Kim, W. L. Gladfelter, and J. H. Yan, IEEE Trans. Electron Devices **44**, 104 (1997).

[103] C. J. Taylor, D. C. Gilmer, D. Colombo, G. D. Wilk, S. A. Campbell, J. Roberts, and W. L. Gladfelter, J. Am. Chem. Soc. **121**, 5220 (1999).

[104] D. C. Gilmer, D. G. Colombo, C. J. Taylor, J. Roberts, G. Haustad, S. A. Campbell, H.-S. Kim, G. D. Wilk, M. A. Gribelyuk, and W. L. Gladfelter, Chem. Vap. Deposition **4**, 9 (1998).

[105] B. He, T. Ma, S. A. Campbell, and W. L. Gladfelter, Tech. Dig. Int. Electron Devices Meet. **1998**, 1038 (1998).

[106] X. Guo, X. Wang, Z. Luo, T. P. Ma, and T. Tamagawa, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 137.

[107] Y. Ma, Y. Ono, and S. T. Hsu, Mater. Res. Soc. Symp. Proc. **567**, 355 (1999).

[108] C. Hobbs, R. Hedge, B. Maiti, H. Tseng, D. Gilmer, P. Tobin, O. Adetutu, F. Huang, D. Weddington, R. Nagabushnam *et al.*, Tech. Dig. VLSI Symp. **1999**, p. 133.

[109] R. B. van Dover, Appl. Phys. Lett. **74**, 3041 (1999).

[110] M. Balog, M. Schieber, S. Patai, and M. Michman, J. Cryst. Growth **17**, 298 (1972); M. Balog, M. Schieber, M. Michman, and S. Patai, Thin Solid Films **41**, 247 (1977); M. Balog, M. Schieber, M. Michman, and S. Patai, *ibid.* **47**, 109 (1977); M. Balog, M. Schieber, M. Michman, and S. Patai, J. Elec. Chem. Soc. **126**, 1203 (1979).

[111] J. Shappir, A. Anis, and I. Pinsky, IEEE Trans. Electron Devices **ED-33**, 442 (1986).

[112] A. Kumar, D. Rajdev, and D. L. Douglass, J. Am. Chem. Soc. **55**, 439 (1972).

[113] M. Copel, M. A. Gribelyuk, and E. Gusev, Appl. Phys. Lett. **76**, 436 (2000).

[114] W.-J. Qi, R. Nieh, B. H. Lee, L. Kang, Y. Jeon, K. Onishi, T. Ngai, S. Banerjee, and J. C. Lee, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 145.

[115] W.-J. Qi, R. Nieh, B. H. Lee, K. Onishi, L. Kang, Y. Jeon, J. C. Lee, V. Kaushik. B.-Y. Nguyen, L. Prabhu *et al.*, Tech. Dig. VLSI Symp. **2000**, p. 40.

[116] C. H. Lee, H. F. Luan, W. P. Bai, S. J. Lee, T. S. Jeon, Y. Senzaki, D. Roberts, and D. L. Kwong, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 27.

[117] T. Ngai, W.-J. Qi, R. Sharma, J. Fretwell, X. Chen, J. C. Lee, and S. Banerjee, Appl. Phys. Lett. **76**, 502 (2000).

[118] M. Houssa, V. V. Afanas'ev, A. Stesmans, and M. M. Heyns, Appl. Phys. Lett. **77**, 1885 (2000).

[119] M. Houssa, M. Tuominen, M. Naili, V. Afanas'ev, A. Stesmans, S. Haukka, and M. M. Heyns, J. Appl. Phys. **87**, 8615 (2000).

[120] R. C. Smith, N. Hoilien, C. J. Taylor, T. Z. Ma, S. A. Campbell, J. T. Roberts, M. Copel, D. A. Buchanan, M. Gribelyuk, and W. L. Gladfelter, J. Electrochem. Soc. **147**, 3472 (2000).

[121] H. Zhang, R. Solanki, B. Roberds, G. Bai, and I. Banerjee, J. Appl. Phys. **87**, 1921 (2000).

[122] B. H. Lee, L. Kang, W. J. Qi, R. Nieh, Y. Jeon, K. Onishi, and J. C. Lee, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 133; **2000**, p. 39.

[123] B. H. Lee, L. Kang, R. Nieh, W.-J. Qi, and J. C. Lee, Appl. Phys. Lett. **76**, 1926 (2000).

[124] L. Kang, B. H. Lee, W.-J. Qi, Y. Jeon, R. Nieh, S. Gopalan, K. Onishi, and J. C. Lee, IEEE Electron Device Lett. **21**, 181 (2000).

[125] L. Kang, K. Onishi, Y. Jeon, B. H. Lee, C. Kang, W.-J. Qi, R. Nieh, S. Gopalan, R. Choi, and J. C. Lee, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 35.

[126] S. J. Lee, H. F. Luan, W. P. Bai, C. H. Lee, T. S. Jeon, Y. Senzaki, D. Roberts, and D. L. Kwong, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 31.

[127] C. M. Perkins, B. B. Triplett, P. C. McIntyre, K. C. Saraswat, S. Haukka, and M. Tuominen (unpublished).

[128] A. H. Edwards, Phys. Rev. B **44**, 1832 (1991).

[129] V. V. Afanas'ev and A. Stesmans, Phys. Rev. Lett. **80**, 5176 (1998).

[130] From W. T. Adams in *Zirconium and Hafnium*, A Chapter from Mineral Facts and Problems (U.S. Dept. of Interior, Bureau of Mines Preprint from Bulletin 675, U.S. Govt. Printing Office, Washington, D.C., 1985), pp. 946–947.

[131] B. E. Gnade and R. M. Wallace (private communication).

[132] J. F. Ziegler, IBM J. Res. Dev. **40**, 19 (1996).

[133] M. A. Russack, C. V. Jahnes, and E. P. Katz, J. Vac. Sci. Technol. A **7**, 1248 (1989).

[134] S. Roberts, J. G. Ryan, and D. W. Martin, in *Emerging Semiconductor Technology*, ASTM STP 960, edited by D. C. Gupta and P. H. Langer (ASTM, Philadelphia, 1986), p. 137.

[135] G. D. Wilk and R. M. Wallace, Appl. Phys. Lett. **74**, 2854 (1999).

[136] G. D. Wilk and R. M. Wallace, Appl. Phys. Lett. **76**, 112 (2000).

[137] G. D. Wilk, R. M. Wallace, and J. M. Anthony, J. Appl. Phys. **87**, 484 (2000).

[138] T. Yamaguchi, H. Satake, N. Fukushima, and A. Toriumi, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 19.

[139] Y. Ma, Y. Ono, L. Stecker, D. R. Evans, and S. T. Hsu, Tech. Dig. Int. Electron Devices Meet. **1999**, p. 149.

[140] L. Manchanda, M. L. Green, R. B. van Dover, M. D. Morris, A. Kerber, Y. Hu, J. P. Han, P. J. Silverman, T. W. Sorsch, G. Weber *et al.*, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 23.

[141] J. C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979); **47**, 203 (1983).

[142] J. C. Phillips, J. Vac. Sci. Technol. B **18**, 1749 (2000).

[143] (a) W.-J. Qi, R. Nieh, E. Dharmarajan, B. H. Lee, Y. Jeon, L. Kang, K. Onishi, and J. C. Lee, Appl. Phys. Lett. **77**, 1704 (2000); (b) M. Copel, E. Cortier, and F. M. Ross, Appl. Phys. Lett. **78**, 1607 (2001); (c) J. A. Gupta, D. Londheer, J. P. McCaffrey, and G. I. Sproule, Appl. Phys. Lett. **78**, 1718 (2001).

[144] W. B. Blumenthal, *The Chemical Behavior of Zirconium* (Van Nostrand, Princeton, 1958), pp. 201–219.

[145] L. Bragg, G. F. Claringbull, and W. H. Taylor, *Crystal Structures of Minerals* (Cornell University Press, Ithaca, 1965), p. 185.

[146] P. J. Harrop and D. S. Campbell, Thin Solid Films **2**, 273 (1968).

[147] E. M. Vogel, K. Z. Ahmed, B. Hornung, W. Kirklen Henson, P. K. McLarty, G. Lucovsky, J. R. Hauser, and J. J. Wortman, IEEE Trans. Electron Devices **45**, 1350 (1998).

[148] D. Frank, Y. Taur, and H.-S. P. Wong, IEEE Electron Device Lett. **19**, 385 (1998).

[149] S. Krishnan, G. C.-F. Yeap, B. Yu, Q. Xiang, and M.-R. Lin, Proc. SPIE **3506**, 65 (1998).

[150] B. Cheng *et al.*, IEEE Trans. Electron Devices **46**, 1537 (1999).

[151] R. D. Shannon, J. Appl. Phys. **73**, 348 (1993).

[152] J. Robertson and C. W. Chen, Appl. Phys. Lett. **74**, 1168 (1999).

[153] J. Robertson, J. Vac. Sci. Technol. B **18**, 1785 (2000).

[154] S. O. Kasap, *Principles of Electrical Engineering Materials and Devices*, 2nd ed. (McGraw-Hill, New York, 2002).

[155] G. Lucovsky and B. Rayner, Appl. Phys. Lett. **77**, 2912 (2000).

[156] R. Beyers, J. Appl. Phys. **56**, 147 (1984).

[157] K. J. Hubbard and D. G. Schlom, J. Mater. Res. **11**, 2757 (1996).

[158] S. Q. Wang and J. W. Mayer, J. Appl. Phys. **64**, 4711 (1988).

[159] I. Barin and O. Knacke, *Thermochemical Properties of Inorganic Substances* (Springer, Berlin, 1973).

[160] L. B. Pankratz, *Thermodynamic Properties of Elements and Oxides* (U.S. Dept. of Interior, Bureau of Mines Bulletin 672, U.S. Govt. Printing Office, Washington, D.C., 1982).

[161] S. P. Murarka, *Silicides for VLSI Applications* (Academic, New York, 1983).

[162] S. Murtaza, J. Hu, S. Unnikrishnan, M. Rodder, and I-C. Chen, Proc. SPIE **3506**, 49 (1998).

[163] H. Zhong, G. Heuss, and V. Misra, IEEE Electron Device Lett. **21**, 593 (2000).

[164] H. Zhong, G. Heuss, V. Misra, H. Luan, C. H. Lee, and D. L. Kwong, Appl. Phys. Lett. **78**, 1134 (2001).

[165] T.-J. King, J. P. McVittie, K. C. Saraswat, and J. R. Pfiester, IEEE Trans. Electron Devices **41**, 228 (1994).

[166] K. Uejima, T. Yamamoto, and T. Mogami, Tech. Dig. Int. Electron Devices Meet. **2000**, p. 445.

[167] S. J. Wang, C. K. Ong, S. Y. Xu, P. Chen, W. C. Tjiu, J. W. Chai, A. C. H. Huan, W. J. Yoo, J. S. Lim, W. Feng, and W. K. Choi, Appl. Phys. Lett. **78**, 1604 (2001).