Critical metrology for ultrathin high k dielectrics.

W. Vandervorst\textsuperscript{a,b}, B. Brijs, H. Bender, T. Conard, J. Petry, O. Richard, X. Blasco\textsuperscript{c}, M. Nafria\textsuperscript{c}

\textsuperscript{a}IMEC, Kapeldreef 75, B-3001 Leuven, Belgium
\textsuperscript{b}also : INSYS, KU Leuven, Belgium
\textsuperscript{c}Departament d’Enginyeria Electrònica, Universitat Autònoma de Barcelona, Edifici Q. 08193. Bellaterra. Spain

Abstract. Targeting very thin equivalent oxides (<1 nm) requires the deposition of (very) thin dielectrics onto silicon surfaces with minimal interfacial oxide. Typically, high-k dielectric layers are deposited using ALD or MOCVD with, at present, a prime emphasis on Hf-based high-k dielectrics, either as pure HfO\textsubscript{2}, as silicate, or mixed with Al\textsubscript{2}O\textsubscript{3}. Depending on the deposition conditions, serious deficiencies in terms of film closure and material density occur for the ultra thin (<3 nm) films which are required for very small EOT’s. As such, critical metrology needs arise enabling one to study details of the film growth and its evolution upon thermal anneal. As discussed in this paper, many tools are required such as Rutherford Backscattering Spectrometry and high-resolution elastic recoil detection (ERDA), Low Energy Ion Scattering, Time-of-flight SIMS, (spectroscopic) ellipsometry and X-ray photoelectron spectroscopy. When trying to correlate these different methods one must be aware of potential discrepancies due to non-homogeneous growth and reduced material density. Local electrical measurements based on tunneling atomic force microscopy reveal very fine scale inhomogeneities, which can be correlated to local structural defects.

INTRODUCTION

High-k insulating metal oxides are currently under consideration as novel gate dielectric materials for silicon MOS devices because they provide the required specific capacitance at a considerably larger physical thickness than SiO\textsubscript{2}, thus allowing the reduction of the gate leakage current by suppression of direct tunneling. Precise control of the layer growth, the evolution of the interface between the silicon and the high-k dielectric and the thermal stability of the gate stack are key elements towards meeting the target of the equivalent oxide thickness being less than 0.5 nm for very advanced applications. Even when using high k materials such as Al\textsubscript{2}O\textsubscript{3}, ZrO\textsubscript{2}, HfO\textsubscript{2}, and their mixtures with Si or Al which have a value of k ranging from 9-25, obtaining an EOT of 0.5 nm still requires very thin films i.e. in the order of 2-3 nm ideally with no, or minimal, interfacial oxide.

The depositions of high k films, as discussed in this work, are either based on Atomic Layer Deposition (ALD) or metal-organic chemical vapor deposition (MOCVD). In its most simplistic way, ALD can be described as a series of sequential saturating gas-solid reactions by alternating H\textsubscript{2}O and precursor pulses (in this work (Al(CH\textsubscript{3})\textsubscript{3}, ZrCl\textsubscript{4} or HfCl\textsubscript{4}) which should result in excellent uniformity of the deposited layers and accurate film thickness control. In MOCVD HfO\textsubscript{2} and Hf-silicates are deposited respectively using tetrakis (diethylamino) hafnium (TDEAH) and TDMAS tetrakis (dimethylamino) silicon) as precursor. In all cases it is clear that achieving a high quality thin film requires a complete control over the growth process and the means to actually monitor the film growth and its structural and electrical properties. A complete review of all possible characterization methods of value in this context is beyond the scope of this paper. Instead we will focus on the issues of determining film thickness and growth mode, film composition and density, and local electrical properties.

STRUCTURAL FILM METROLOGY

Film thickness and density

One of the obvious approaches to study the basic properties of these deposition processes is to look at growth curves i.e. the evolution of film thickness as a function of growth/deposition time. The most unambiguous way is to rely on techniques like Rutherford Backscattering (RBS) or X-ray Fluorescence (XRF) which measure directly the number of atoms deposited per cm\textsuperscript{2}. Typical examples of growth curves (for ALD deposition of Al\textsubscript{2}O\textsubscript{3}, HfO\textsubscript{2} or ZrO\textsubscript{2} on an HF-last surface) are shown in Fig. 1. They clearly indicate that film growth can occur in a very

\textsuperscript{a}contacting author : vdvorst@imec.be
non-linear mode and apparently is very sensitive to the surface preparation. With such a non-linear growth behavior it is clear that tuning of the deposition systems and recipes becomes important and thickness control will be a key activity in a production line and during process development. For that purpose a fast non-contacting technique like ellipsometry would be a preferred choice above (expensive) large accelerators, in particular since ellipsometry is, with success, already extensively used in the normal semiconductor processing environment [1]. As the gate stack is typically a multilayer structure (some interfacial oxide is almost always present) multilayer fitting is a necessity [2]. Despite the use of spectroscopic ellipsometry and extending the wavelengths to the vacuum UV range (135-620 nm range), a large uncertainty on the interfacial oxide thickness remains and its value usually needs to be fixed in order to obtain a reliable value for the high k film thickness [3].

An inherent limitation with the application of the fitting models is that use is made of optical data extracted from thicker films, which may or may be not be applicable to very thin films. In principle an alternative approach would be to use angular resolved X-Ray Photoelectron Spectroscopy (XPS) providing information on the high k film thickness as well as on the interfacial oxide. The latter can presently be done on a full wafer scale either by repeating the measurements with different acceptance angles or by using instruments which provide directly multi-angular detection and thus all angles in one single measurement [4]. One should note however that when using angular resolved XPS (AR-XPS) and the Hf/Si intensities to extract the thickness \( t = \lambda \cos \theta \ln \left( 1 + K \frac{t_{Si}}{t_{Hf}} \right) \) (eq.1), the angular range needs to be limited (depending on film thickness) to less than 60 degrees as deviations from the expected linear relationship with \( \cos(\theta) \) are observed for larger angles (probably due to elastic scattering) (Fig. 2) [6].

When trying to establish a correlation between techniques like (spectroscopic) ellipsometry (SE), X-Ray Photoelectron Spectroscopy (XPS), or Transmission Electron Microscopy (TEM), i.e. between techniques measuring directly film thickness as a length scale (SE, TEM,...), versus techniques measuring the number of atoms/cm\(^2\) (RBS), the uncertainty on the density of the very thin films becomes an essential parameter. In this context it is also important to emphasize that the thickness as determined by XPS based on the attenuation of the Si-substrate signal, inherently contains an assumption on the density as long as standard attenuation lengths are used in the calculations [7].

In principle X-ray reflectometry (XRR) allows the simultaneous determination of thickness, density and roughness of (multi-)layers by fitting of the measured spectra with calculated ones. Unfortunately for very thin layers (< 1-2 nm) the XRR spectrum shows little structure and the correlation between the different parameters becomes large, i.e. the accuracy of the fitting might be questionable. Nevertheless, when comparing the XRR-results with the density extracted from the number of atoms as measured.
with RBS and the thickness of TEM, quite a remarkable agreement is obtained (Fig. 3). The latter holds not only for films with a density close to the normal bulk values but equally well for very thin films where the density appears to be only 25% of the expected value.

Compared to the normal bulk density values (monoclinic, tetragonal or cubic phase of HfO₂) it is clear that high k films with a very different density can be produced and their density determination will be essential to understand their electrical properties. In particular in the early phase of film growth ("incubation" period?), a very significant difference can be observed between the length scale measurements and the thickness calculated using bulk density from the number of atoms deposited (Fig. 4a). For thicker films the two curves show the same slope and differ by about one nm. Note as discussed above, that the XPS-results agree with the RBS data since they are calculated with the standard attenuation lengths which implicitly implies the assumption of the normal bulk density. However when using for each thickness, the density as derived from XRR (notwithstanding its limitations for very thin layers), the agreement is much better (Fig. 4b). The convergence in the early phase of the film deposition and the disappearance of the 1 nm offset arises from the unusually low density values which are measured in that region by XRR (< 25% of the bulk value). One must admit that such a low density is hard to interpret in terms of a different crystal structure and most probably must be viewed as indicative for a non-closed film containing holes or voids.

**Figure 3.** Density of HfO₂ films (ALD and MOCVD) as determined from RBS/TEM or XRR. The bulk values for some phases (C=cubic, M=monoclinic) of HfO₂ are shown for comparison [8].

**Figure 4.** Film thickness versus deposition cycle
a) TEM, SE using optical data derived for the 100 cycles sample, RBS with He beam, single angle and angular resolved XPS, and XRR fitting both the thickness and the density. RBS/XPS data are converted using bulk density values
b) after recalculation of the RBS and ARXPS data with the XRR-densities for each individual sample [8]
Growth mode

Faced with limitations of the growth techniques it is essential, at least in the early phases of materials development, to be able to assess in more detail the detailed nature of the film growth. A very effective and quantitative way to look at the film growth in the early stages is to use Low energy ion scattering (LEIS). In essence LEIS provides a way to probe the evolution of the surface coverage of the Si-surface by the growing metal oxide. Hence the (dis)appearance of the metal(Si)-atoms as a function of material deposited will be characteristic for the nature of the film growth (planar, islanding,...). When thus combining the film thickness results (RBS or XPS) with the surface coverage from LEIS (Fig. 5) for the case of Zr-oxide grown on a HF-last Si-surface, one can clearly observe the long persistence of a (small) Si-signal, indicative of Si-areas remaining uncovered with the metal oxide. The fact that some Si-atoms are not covered by ZrO\(_2\) up to a film thickness in excess of 4 nm, implies that the some holes in the film remain. Obviously this will have an impact on the electrical quality of the film. Such a delayed closure has been observed for Al\(_2\)O\(_3\), HfO\(_2\), ZrO\(_2\) when deposited by ALD and is related to the difficulties in adsorbing on the H-terminated Si-surface versus the easier adsorption on the growing metal oxide island [7]. A detailed modeling of this process has been presented elsewhere [10]. The extent of this island growth is a function of the surface preparation and becomes drastically reduced when growing on “wet” chemical oxides [11].

The existence of these holes can also be inferred from the epitaxial alignment of the poly-grains through the holes when depositing poly-silicon in-situ on top of a very thin (1.5 nm) high k film (Fig. 6).

As LEIS is rather slow and time consuming an alternative approach based on TOFSIMS has been developed for looking at film closure [12]. The technique is based on the fact that TOFSIMS is also very surface sensitive (the escape depth of a secondary ion being around 1-2 monolayers) and hence the Si-signal should decay very rapidly when becoming covered with a high k film. For an ideal layer by layer growth, the Si-intensity would decay by 0.2-0.4 nm/dec. If the film would grow in a less planar mode, a slower Si-decay will arise. Hence the steepness of the Si-intensity decay can be used as a measure for the efficiency of film closure. An example of the Si-decay for the deposition of Zr-oxide on a HF-last surface or an RTO-oxide is shown in Fig. 7. In both cases the decay of the Si-intensity is far less than the theoretical 0.2-0.4 nm/dec. In particular the very slow decay for the HF-last case indicates a very slow closure of the film, as can deduced from the LEIS-results as well.

Figure 5. Si(Zr) surface concentration as determined with LEIS for ALD-growth of ZrO\(_2\) on HF-last Si surface. Thickness is deduced from XPS measurements assuming bulk density.

Figure 6. TEM-cross section of polysilicon deposited in-situ on thin high k film. The holes in the film allow the poly to grow epitaxially aligned with the substrate.

Figure 7. Decay of Si-intensity (surface concentration) from TOFSIMS versus high k film thickness for deposition on RTO and HF-last surface. Si-surface concentration based on LEIS for the HF-last sample is shown as well.
Although in Fig. 7 almost a one-to-one correspondence between the LEIS and the TOFSIMS results can be observed, this is not universally true. When using the TOFSIMS method to study the growth behavior on different starting surfaces (Si-OH, Si-H terminated, NH$_3$ annealed surface) as well as for different metal oxides, one needs to realize that the TOFSIMS data are only semi-quantitative and possibly influenced by changes in ionization yield as a result of the increasing surface oxidation (in particular in the case of an HF-last surface) or the presence of the metal oxide. Indeed whereas one would expect a linear decrease of the Si-intensity (and a linear increase of the metal-intensity) with increasing metal-oxide coverage (as seen in LEIS), in reality these correlations are far less linear (Fig. 8). Needless to say that the non-linearity of these curves will affect the steepness of the Si-decay in addition to the effect of a decreasing metal coverage. Hence the decay curves can not be interpreted as a direct measurement of surface coverage but should solely be used in a qualitative manner. Certainly within one materials system, these non-linearities are fairly constant and a steeper decay of the Si-intensity will always be indicative for a faster closure of the film.

As the evolution from islands towards a closed film represents the transition from a heterogeneous surface towards a homogeneous surface, the growth rate per cycle is continuously changing as well. Assigning the origin of the non-linearities in the growth curves to the same mechanism (delayed closure, heterogeneous surface) implies that with optimized surface preparation, whereby films will close more rapidly, also the growth curves will be more linear. The latter is evident when comparing the trend in the RBS growth curves with the TOFSIMS decays for Hf-oxide films grown on various surfaces (Fig. 9) [10].

Composition profiles

Faced with the formation of nano-crystallites in the high k films, even at moderate anneal temperatures, more complex systems based on mixed oxides (aluminates, silicates) have been proposed, as they appear to have a higher thermal stability. In ALD these mixed oxides are grown with alternating cycles of the individual oxides.
components. As the growth rate per cycle is less than one monolayer, one expects to obtain a uniform composition throughout the film. However it has been observed when growing one kind of high k material on top of another one, that the initial growth rate is strongly influenced, hence some interaction of the individual metals can not be excluded. An indication of such an effect is found when growing AlHf-oxides with different thickness and determining their composition with RBS or XPS. As these films are very thin (1-2 nm) neither of these techniques has sufficient depth resolution to determine the real depth profile and only an average composition can be determined. The results shown in Fig. 10 indicate that for very thin films the Hf composition exceeds the intended composition significantly suggesting a strong initial preference for Hf deposition versus Al. The latter appears to be a general phenomenon and dependent on the Hf/Al-ratio (higher initial deposition of Hf with increased Hf/Al ratio)

![Figure 10](image-url)  

**Figure 10.** Average Hf/Al composition ratio versus film thickness for three “bulk” Hf/Al-cycle ratios (1:2, 1:1, 2:1). As the growth rate per cycle is different for Hf and Al, the cycle ratios do not reflect the composition.

The results in Fig. 10 indicate a transient in the deposition efficiency of Hf versus Al with a higher Hf- or a reduced Al-deposition in the early stages of the film growth. Inevitably this will then lead to a non-homogenous film composition. In order to be able to probe such a compositional variation, one needs a technique with sub-nm depth resolution. One possible approach is to make a TOFSIMS depth profile with a very low energy (< 500 eV) sputter beam. The results in Fig. 11 indicate a strong non-uniformity with a Hf enrichment at the SiO₂ interface and a sloped profile for the Hf and the Al. Although the TOFSIMS profiles show the expected behavior i.e. interfacial enrichment of Hf, some inconsistencies are present as well. For instance, whereas the Al signal scales with the deposition ratio, there is very little influence on the Hf-signal despite the change in Hf-concentration. Obviously the non-linear response of the SIMS signals is linked to (unknown) changes in ionization probability with Hf/Al-concentration. These are probably smaller for Al than for Hf as the ionization potential (one of the determining factors for ionization) of Al (5.99 eV) is smaller than the one of Hf (7 eV) making the ionization less sensitive to the surface chemistry.

In order to resolve this problem, use should be made of a more quantitative technique with very high depth resolution. An excellent tool for this application is high-resolution elastic recoil detection (ERDA) [13]. The latter is the complement of RBS in the sense that now a high energy (40-170 MeV) heavy element (I-Au) is used to elastically scatter light elements out of the target, which are then analyzed according to their mass and energy. Using very advanced magnetic sector analysers, sub-nm depth resolution can be obtained (at least in the near-surface regions) while maintaining the quantification properties of nuclear methods. The results shown in Fig. 12 relate to a 7 nm HfAl-film deposited on a nominally 1 nm interfacial oxide. The gradient of the oxygen profile at the surface is a good measure for the depth resolution of the technique. In this experiment it is less than 0.5 nm/dec. The presence of the interfacial oxide can quite easily be detected when comparing the depth of the oxygen profile versus the depth of the Al and the Hf-profiles. The difference (~1 nm) determines the thickness of the interfacial oxide after high k film growth. Consistent with the results in Fig. 10 and 11, the
ERDA-data show a sloped profile for the Hf as well as for the Al, highlighting the non-linear growth rates as a function of film thickness. Quite surprising, but very clear, is the fact that the first monolayer is completely free of Hf suggesting that during the growth of the mixed oxide a phase separation occurs leading to some Al-floating on the growing surface. The different slopes (positive versus negative) of the two elements also point into the direction of a surface migration of Al and Hf-indiffusion causing an (extra) enhanced Hf-concentration at the interface. It is also interesting to look at some other contaminants such as H and C. The C-profile solely peaks at the surface and is thus representative of the contamination adsorbed on the sample surface during exposure to the ambient. The H-profile has a double peak structure: one large peak concentrated at the surface which is also indicative of the surface contamination and a smaller one (= peak concentration (~1 at %) located at the metal-oxide- silicon oxide interface. Clearly the latter is associated with remnants from the water pulses of the ALD-process, probably reflecting a poorer removal of OH-groups from the silicon oxide surface. The other contaminants H and C show a different behavior: H has a double peak structure with one peak at the surface and the other one close to the sample surface. C on the other hand shows only one peak at the surface. The presence of H and C on the sample surface is indicative of the chemical reaction between the sample surface and the ambient during exposure.
migration occurs during the anneal whereby some Si diffuses towards the surface. Needless to say that in that case the apparent interfacial oxide will be overestimated as the photoelectrons from the surface oxide are far less attenuated than the ones from the deeper lying interface. Note that concurrent with the Si-diffusion migration of Al and Hf can be observed.

**LOCAL ELECTRICAL PROPERTIES**

Most of the knowledge about the electrical behavior of high-k materials has been gained from measurements performed on macroscopic MOS capacitors or transistors using standard characterization methods at the macroscopic device level. These tests, however, provide only spatially averaged information on the electrical properties of the material and do not address the microscopic failure mechanisms. One very interesting approach to address this issue is to use Conductive Atomic Force Microscopy (C-AFM), as already demonstrated for SiO$_2$ [14]. Conductive Atomic Force Microscopy (C-AFM) is a measurement technique based on an AFM equipped with a conductive tip, whereby simultaneously topographical and electrical (tunneling current) information is collected with a typical lateral resolution of about 10nm. Moreover, local current-voltage (I-V) characteristics can also be measured in a range from 100 fA to 100 pA. As the current tunneling from the tip through the oxide is strongly dependent on the film properties, leakage paths and trapped charges, C-AFM is an excellent tool to study on a local scale the electrical quality of the film and the breakdown mechanisms.

![Figure 14: Interfacial oxide versus film thickness of mixed oxide before and after anneal.](image)

**Figure 14.** Interfacial oxide versus film thickness of mixed oxide before and after anneal.

![Figure 15. TOFSIMS depth profile of HfAl-oxide before and after anneal. Arrows indicate the element migration with increasing anneal temperature.](image)

**Figure 15.** TOFSIMS depth profile of HfAl-oxide before and after anneal. Arrows indicate the element migration with increasing anneal temperature.

![Figure 16. Current maps (-6 V bias) of HfAl- gate stack for different annealing temperature : a) not annealed sample, b) $T_A=800\,^\circ\text{C}$, c) $T_A=900\,^\circ\text{C}$, d) $T_A=1000\,^\circ\text{C}$. The area of the scanned region is 2x2 $\mu$m$^2$. The current ranges from 0pA (black) to 20pA (white). e) 1x1 $\mu$m$^2$ zoom of a), showing weak (more conductive) spots, when the current scale is decreased from 0pA (black) to 1pA (white), f) a 0.3x0.3 $\mu$m$^2$ current image of the sample exposed to $T_A=900\,^\circ\text{C}$. The current ranges from 0pA (black) to 10pA (white) [15].](image)
In the example shown in Fig. 16, HfAl-oxide films have been measured as a function of their post-deposition anneal temperature [15]. Basically the current maps show a low homogeneous leakage current distribution for the as-deposited samples with a gradually increasing leakage current for higher anneal temperatures. At the same time one observes the appearance of local regions of even higher conductivity. Obviously the latter represent “weak” spots (not yet breakdown spots!) in the film which may evolve into complete film failure. The size of these defects is close to the tip resolution and on the order of 5-20 nm. Their density increases with anneal temperature and can become as large $10^{10}$ defects/cm$^2$. Although their nature is not complete understood yet, it is tempting to correlate these spots with the formation of nanocrystallites in the high k film as observed in TEM-images (Fig. 17) and increased conduction through these grains or their grain boundaries.

**REFERENCES**


2. W.W.Chism, A.C.Diebold and J.M.Price, these proceedings

3. P.Boher, P.Evrard, C.Defrenoux, J.C.Fouere, E.Bellandi and H.Bender, these proceedings

4. R.Champanaria, P.Mack, R.White and J.Wolstenholme, these proceedings


