Ellipsometry of Sol-Gel Films

54

Eric Yeatman

Contents

Abstract

Ellipsometry is an optical method for determining the properties of thin films, using manipulation and measurement of the polarization state of reflected light. This chapter introduces the method, summarizes the optical principles and analysis involved, and describes the basic experimental arrangements. Applications and limitations of the technique are presented. A specific application, the use of ellipsometry to determine the pore size distribution in thin porous films, is then described in some detail.

Introduction

In the broadest sense, "ellipsometry" is concerned with measurement and analysis of the state of elliptical polarization of light (Rothen 1974). However, it is generally used to mean a method, based on analysis of elliptical polarization, to determine the properties of thin films (particularly the thickness) on dielectric or metal surfaces. The basic theory derives from the work of Lord Rayleigh and Paul Drude in the late

E. Yeatman (✉)
Faculty of Engineering, Department of Electrical and Electronic Engineering, Imperial College London, London, UK
e-mail: e.yeatman@imperial.ac.uk

nineteenth century. Lord Rayleigh had inferred the presence of viscous films of minute thickness on water surfaces and devised an experiment, whereby the effect of these films on the polarization state of a reflected beam could be measured with great precision. Drude, meanwhile, was interested in the reflection of light from thin films on solids and derived, based on Maxwell's equations, the fundamental formulae on which ellipsometric instruments are based.

A plane wave incident on a surface at an oblique angle can in general be decomposed into two polarization components, which are typically referred to as the p and s waves, whose electric field vectors lie in and normal to, respectively, the plane of incidence. Upon reflection, these two waves may each undergo changes of amplitude and phase. In ellipsometric measurement, it is the change in the relative amplitude and phase of the two components that is measured. That is to say, if we designate phases and amplitudes by β and A, respectively, and denote p and s waves by the corresponding subscripts, then the two quantities to be determined are (Archer 1962):

$$\Delta = \left(\beta_p - \beta_s\right)_{\text{reflected}} - \left(\beta_p - \beta_s\right)_{\text{incident}} \tag{1}$$

$$\Psi = \tan^{-1}\left[\frac{\left(A_p/A_s\right)_{\text{reflected}}}{\left(A_p/A_s\right)_{\text{incident}}}\right] \tag{2}$$

Drude's formulation gives the relation between Δ and Ψ and the film properties according to:

$$\tan\left(\Psi\right)e^{i\Delta} = \left(\frac{r_{1p} + r_{2p}e^{-2i\delta}}{1 + r_{1p}r_{2p}e^{-2i\delta}}\right)\left(\frac{1 + r_{1s}r_{2s}e^{-2i\delta}}{r_{1s} + r_{2s}e^{-2i\delta}}\right) \tag{3}$$

Here r_{1p} and r_{1s} are the Fresnel reflection coefficients of the film surface, and r_{2p} and r_{2s} the Fresnel coefficients at the film–substrate interface. These in turn can be calculated from the material properties and wave directions, so that

$$r_{1p} = \frac{n_a \cos\varphi_f - n_f \cos\varphi_a}{n_a \cos\varphi_f + n_f \cos\varphi_a} \tag{4}$$

$$r_{1s} = \frac{n_a \cos\varphi_a - n_f \cos\varphi_f}{n_a \cos\varphi_a + n_f \cos\varphi_f} \tag{5}$$

with n_a and n_f as the ambient and film refractive indices and φ_a and φ_f as the angle of incidence and angle of refraction in the film, respectively. The two angles are related by Snell's law, and the coefficients r_{2p} and r_{2s} are found by substituting the subscripts f and s (film and substrate) for a and f in Eqs. 4 and 5. Note that both the indices and (therefore) the angles may be complex if the materials are lossy, as will often be the case.

In Eq. 3, δ is the phase change caused by the propagation through the film (in both directions) according to

$$\delta = k_o d \sqrt{n_f^2 - n_a^2 \sin^2 \varphi_a} \tag{6}$$

with d the film thickness and $k_o = 2\pi/\lambda$ the free-space wave number.

Since we have two measurable variables Δ and Ψ, we can derive two unknown quantities of the system, for example, d and n_f (if we assume n_f is real). While this gives an immediate benefit over more straightforward interferometric measurements, it is mainly the attainable precision that has made ellipsometry an important technique. Interest in very thin films was greatly stimulated by the development of the Langmuir–Blodgett method of transferring monolayer films onto solid surfaces in the 1930s. These were initially measured by purely interferometric techniques, giving a thickness precision of at best a few Å (Rothen 1974). The use of ellipsometry improved this by an order of magnitude. This was first achieved by Rothen (1945), who devised a technique to determine the polarization parameters precisely without a photomultiplier tube; while the availability of these have made his particular technique redundant, the name he coined for the instrument, "ellipsometer", has come into general use.

Ellipsometric Measurements

The basis of the ellipsometer measurement is to determine the state of polarization of the reflected beam, i.e., its ellipticity and the orientation of the ellipse. If a $\lambda/4$ plate is placed in the beam path such that its principal directions are aligned with the major axis of the ellipse, the light emerging is linearly polarized and can be extinguished by an analyzer at the appropriate angle. The angles of the $\lambda/4$ plate and analyzer thus indicate Δ and Ψ.

In practice this simple nulling approach lacks precision, and a number of variants of the instrument architecture and operation have been derived, such as that shown in Fig. 1a. Here the quarter-wave plate (or compensator) is placed in the incident beam and is kept at a fixed position, with the analyzer and an initial input beam polarizer being rotated. Measurement of detected intensity vs. these two angles allows the nulling positions to be interpolated with high precision, and the instrument is highly suited to automation. Since only two parameters can be determined by such a measurement, as stated above, additional parameters must be known in advance. Normally these will be the substrate optical constants and the film absorption coefficient (usually this will be approximated as zero for glass films). Silicon is a useful substrate material because its optical constants are well characterized, wafers are available at low cost with excellent surface flatness and smoothness, and it is highly absorbing in the visible. The latter property is useful because it avoids the risk of additional reflections from the bottom of the substrate, as well as ensuring a strong reflection at the film–substrate interface (important for signal-to-noise ratio in the measurement). Thin films on silicon, particularly thermal oxide films, are also of major technological importance in their own right.

Fig. 1 Geometry of standard null ellipsometer (**a**) and rotating analyzer ellipsometer (**b**). L, P, C, A, D, and S represent the laser, polarizer, compensator, analyzer, detector, and sample, respectively (After Jenkins 1999)
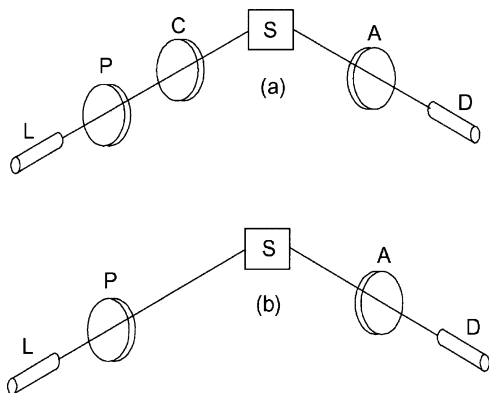
Additional information can be acquired if the angle of incidence (Jenkins 1999) or wavelength is varied, and this information can allow a larger number of film parameters to be calculated. Angle variation tends to add complexity to the operation of the instrument because of the need to know the angle very precisely at every measurement point, so that use of a few fixed angles rather than continuous variation is often preferred. In addition, the incident angle in an ellipsometer is usually set to be near Brewster's angle, where the reflection coefficients differ the most between the two polarization states. Varying over a large range reduces this advantage. At Brewster's angle, given by $\theta_{B} = \tan^{-1}(n_2/n_1)$, the reflection coefficient for a p-polarized wave incident at the boundary between media 1 and 2 is zero. For measurement of silica films on silicon, and taking 1.46 and 3.5 as the refractive indices of SiO_2 and Si, respectively, the Brewster angle at the air–silica boundary is 55.6°, while θ_{B} at the SiO_2–Si boundary is 67°. Most ellipsometry is carried out in the angular range 55–80°.

The use of both multiple wavelengths and continuously variable wavelength, usually designated spectroscopic ellipsometry, has become increasingly popular. In this case it is of course important to include chromatic dispersion of the film when fitting parameters to measurements; more specifically, the film dispersion must be adequately modeled with fewer parameters than can be effectively extracted from the measurements if any additional film properties are to be determined. Naturally, variation of both incident angle and wavelength in one instrument provides an increased capacity for parameter fitting (Snyder et al. 1986).

As stated above, determination of the two ellipsometric angles Δ and Ψ allows two film parameters, usually thickness and index, to be calculated, and this calculation can of course be automated so that the film parameters are provided by the instrument directly. However, it will generally be useful to obtain the angles as well, as there are a number of important factors affecting the calculations and their interpretation. In Fig. 2 a typical plot is shown of Δ and Ψ as a function of the film index and thickness. Two particular features are clear from this plot. Firstly, the function is periodic with respect to thickness, just as in a more straightforward interferometric measurement. Thus, the thickness is not determined absolutely but
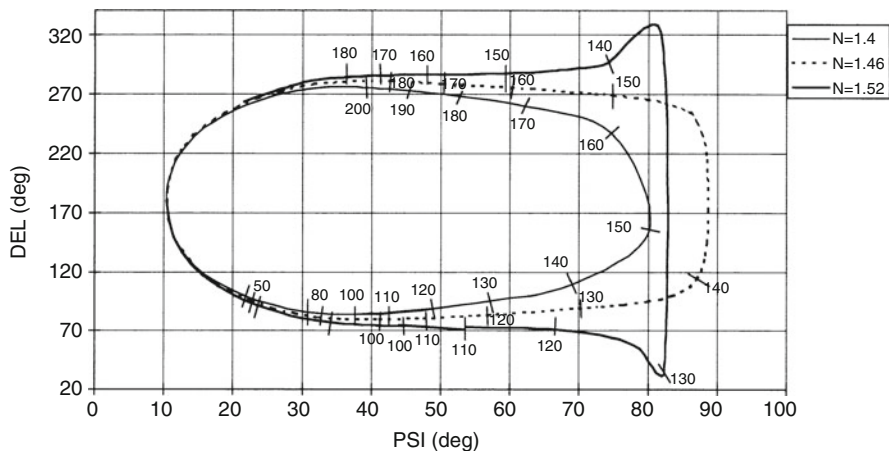
Fig. 2 Example plot of ellipsometry angles vs. film thickness for three values of film index, for transparent films on Si. Numbers on plot show film thicknesses in nm

in the form $(t_0 + mt_c)$, where t_0 is the smallest thickness giving this set of angles, t_c is the cycle period (i.e., the thickness change giving a complete rotation around the plot to return to the same point on the Δ–Ψ graph), and m is an unknown integer. Effectively, then, a single measurement can return three parameters: t_0, t_c, and refractive index. Normally m is inferred from an approximate thickness estimate, based on another measurement technique or from the deposition parameters. For glass films on silicon, visual inspection provides a reasonable estimate in the thickness range where interference colors are strong, i.e., below 1 μm.

The second consideration evident from Fig. 2 is that certain regions in Δ–Ψ space provide less reliable measurements than others, namely, where the lines of constant index converge. In such a region, the inferred film parameters are very sensitive to changes in the ellipsometric angles and therefore to errors in the reading. In the example shown in the figure, in the low Ψ-value side of the graph, there is little variation between the different index plots; for $\Psi < 20$, the index cannot be inferred with any accuracy, so that there is also uncertainty about the thickness in this range. One important source of errors affecting the reading is alignment; the calculations depend strongly on angle of incidence, and while this is usually set with high precision in the instrument, lack of flatness in the sample (e.g., stress-induced substrate bow) may affect the results significantly. This suggests that if a process is being calibrated, thicknesses should be used corresponding to "good regions" in the ellipsometric plot, while if the film properties are fixed, the wavelength and/or incident angle of the instrument should be chosen to give a reliable reading.

Another consideration affecting interpretation of the results relates to the assumptions in the model by which the ellipsometric plot is calculated. In particular, in Fig. 2 the film is assumed to be homogeneous and isotropic, while the nature of thin film deposition means that films may deviate substantially from either of these qualities. With sol-gel coatings, densification may occur from the film surface

downwards, if dominated by solvent evaporation, such that a more dense "skin" is formed before full densification is carried out, and this may leave a residual depth dependence of density. Other mechanisms may also produce a variation of structural properties with depth. If such a structure is suspected, ellipsometry does offer the possibility (as discussed above) to determine a higher number of film parameters. This can be done by modeling the film as a laminate of two or more homogeneous layers, and using multiple measurements, for example, by varying incident angle or wavelength, to acquire additional data points.

These data can then be used to calculate the multiple parameter values by a best-fit approach. For example, Gartner et al. (2003) used a Bruggeman effective index approximation to fit several parameters to spectroscopic ellipsometer measurements of multilayer Fe_2O_3 sol-gel films. By this method they were able to extract an additional characteristic, by dispensing with the assumption of isotropy found in the simplified model of Eqs. 3, 4, and 5, namely, the film birefringence. Many films will have some anisotropy, most commonly with uniaxial symmetry according with the film's geometry. This may be caused by the manner in which the film grows on the substrate, leading to microstructural anisotropy, or by an asymmetric stress in the film, for example, resulting from thermal mismatch with the substrate. In the case of Gartner et al. (2003), the polycrystalline films exhibited a preferential grain orientation leading to an index difference (between in plane and normal to the film) of up to 0.07.

In fact, the relations between the film properties and the measured ellipsometric angles for anisotropic films have been long derived, by extension of the standard theory; for the uniaxial case, with the axis normal to the film surface, the resulting equations are not greatly more complicated than the standard ones (Den Englesen 1971), and the corresponding curves can be straightforwardly plotted. One consequence on the form of the ellipsometric plots is that the $\Delta-\Psi$ curves are not closed but become spirals, and if the films are absorbing and birefringent, consecutive loops of these curves may alter considerably in shape and position. Higher degrees of anisotropy have also been analyzed; ellipsometry can, for instance, be used to determine optical constants of biaxial materials (Schubert and Dollase 2003).

An important cause of uniaxial properties in sol-gel films is in-plane tensile stress, caused by the high degree of shrinkage generally undergone during drying and sintering. This stress is a common cause of mechanical failure of the films and is the reason why the achievable thickness of single sol-gel coatings, at least for inorganic films, is rather limited. Stresses can be characterized by measuring substrate curvature, and values above 100 Mpa are not exceptional (Sengupta et al. 1998). Such levels would cause only slight birefringence in fused silica, where the opto-elastic coefficients are on the order of 2×10^{-12}/Pa. However, in the case of sol-gel films, the stresses are likely to influence the films microstructure during the densification processes and may have much higher effects. This subject does not appear to have been much investigated to date.

Finally, we can mention another variation of the conventional ellipsometer, namely, the focusing ellipsometer (Neuschaefer-Rube et al. 2003). Here, as the name suggests, focusing of the incident beam allows an increased spatial resolution

(in the in-plane directions) of the measurements. However, the focusing inherently increases the angular distribution of the incident light, which must be taken into account in the interpretation of the measured values.

Pore Analysis

Porosity is an important characteristic of sol-gel films, pores on a nanometer scale typically being produced by the removal of organic ligands during heat treatment. These pores influence the film properties directly but also play a role in the stability of the films, as they can provide pathways for atmospheric water and other reactive contaminants into the internal structure. On the other hand, the pores also present a useful site for the insertion of dopants. Therefore, knowledge of the nature of the porosity is often an important aspect of film characterization. However, most methods for assessing porosity are only effective for bulk materials; for example, the common BET method based on N_2 absorption relies on a measurable mass of N_2 being absorbed. The ability of ellipsometry to give a precise measurement of a film's optical properties despite the negligible mass suggests its use in porosity evaluation.

Such an application was introduced when adsorption isotherms of sol-gel films were recorded by ellipsometry (Martin and Green 1990). The refractive index of silica–titania films was measured as a function of relative humidity and the resulting curves used to infer pore size distribution as in the BET technique. For example, a rapid rise of index with humidity indicates small pore radii, as water will condense more readily on small radius (concave) surfaces. However, while useful qualitative information was achieved, the validity of the calculations giving the quantitative results was in some doubt, because of the exceptionally small pore radii (below 1 nm). Nevertheless, this technique continues to find uses, for example, to characterize protective vacuum deposited titania coatings (Alvarez-Herrero et al. 1999).

A variation of the water adsorption isotherm method was then developed (Yeatman et al. 1994; Dawnay et al. 1995) called *molecular probe ellipsometry (MPE)*. In this technique, refractive index readings are made in dry nitrogen and then in nitrogen saturated with an adsorbate solvent vapor. The porous structure is modeled as an effective medium according to the Lorentz–Lorenz relation, giving a relation between the measured film index n_f and the indices of the material in the pores and of the solid skeleton, n_p and n_s, respectively:

$$\frac{n_f^2 - 1}{n_f^2 + 2} = \left(1 - \upsilon_p\right)\frac{n_s^2 - 1}{n_s^2 + 2} + \upsilon_p\frac{n_p^2 - 1}{n_p^2 + 2} \tag{7}$$

where υ_p is the fractional porosity. If the skeleton index is known, or can be assumed with confidence, then the porosity can be obtained directly from the dry measurement, as in McDonagh et al. (2002). However, if the pores are assumed filled with the adsorbate in the saturated case, then from the two measurements two parameters can be determined, namely, n_s and υ_p.

An apparatus suitable for both the MPE and the isotherm techniques is shown in Fig. 3. A source of dry N_2 is fed into a measurement chamber by two paths, one of which passes through a bubbler of the adsorbate liquid. It is assumed that the gas leaving the bubbler is fully saturated. For MPE, the appropriate valve is opened for the wet or dry reading, the other being fully closed. For taking isotherms, both lines are partially opened, with flow meters (not shown) monitoring their relative rates; the relative humidity is taken as the percentage of wet gas entering the chamber. The measurement chamber is simply a closed aluminum box with windows for the ellipsometer beam to pass in and out and an additional window in the top for alignment and monitoring purposes. Care must be taken to ensure that the beam windows do not alter the properties of the beam, particularly its polarization state; they must be flat, clean, perpendicular to the beam, and free of any stresses that might induce birefringence. A gas output line can pass through a cold trap or other mechanism to remove the adsorbate vapors if necessary. Several additional considerations should be observed in the construction and operation of this apparatus. Firstly, all parts of the system should be at the same temperature, most importantly to avoid condensation on the sample. For this reason, the gas flow across the sample should not be so high as to cause convective cooling. Uniformity in the chamber environment is also important; this can be aided by having perforated hoses for inlet and outlet running inside the chamber. Readings in wet and dry states should be taken after sufficient stabilization of the sample in the measurement environment – tens of minutes seemed to be sufficient for this purpose, while excessive times (hours), particularly in the "wet" state, can cause irreversible changes to the sample. The gas lines on the inlet side must be of a material which does not absorb any of the adsorbate species.

For MPE, the porosity calculation of Eq. 7 is first done with water as the adsorbate and then repeated with adsorbate species of increasing molecular size. A set of molecules was chosen for this purpose with minimal polarity and roughly spherical shape, the largest being a crown ether of just over 0.9 nm radius. These are shown in Table 1. Each calculation gives a measure of the porosity accessible to molecules of that adsorbate's size and therefore should include pores of that size or greater. A plot of the measured porosity vs. size, as in Fig. 4, is therefore effectively an integral, from right to left (decreasing pore size), of pore volume vs. diameter. For example, a
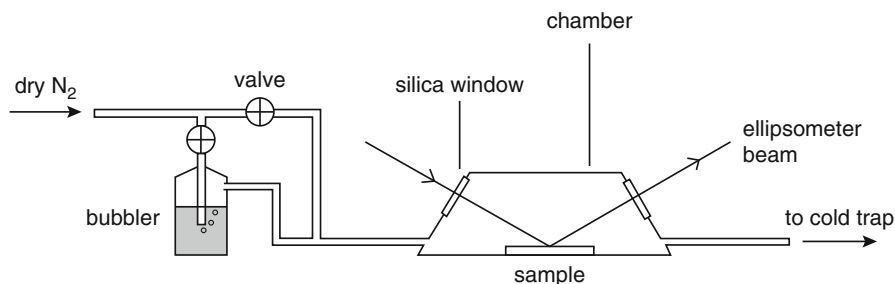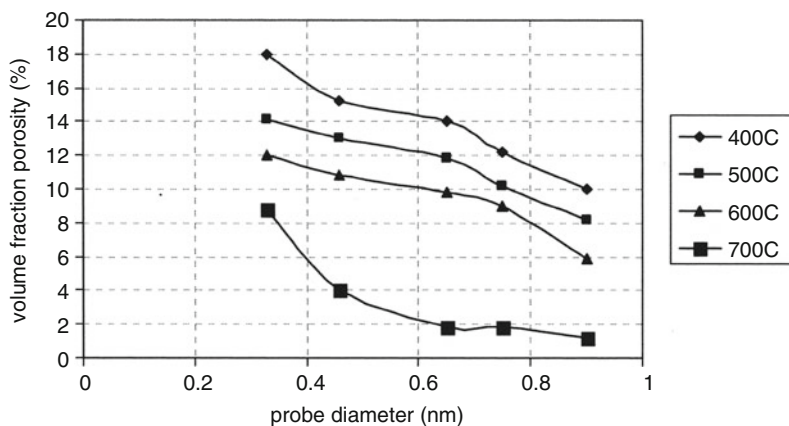

Fig. 3 Apparatus for molecular probe ellipsometry (After Yeatman et al. 1994)

Table 1 Adsorbate species and dimensions for molecular probe ellipsometry (From Yeatman et al. 1994)

Adsorbate	Critical dimension (Å)
Water	3.28
Methanol	4.73
Carbon tetrachloride	6.49
Cyclohexane	7.36
1,5,9-Cyclododecatriene ($C_{12}H_{18}$)	9.04

Fig. 4 Fractional porosity vs. molecular probe size for silica sol-gel films annealed at various temperatures as indicated (After Yeatman et al. 1994)

constant slope in such a plot indicates a uniform pore size distribution. Clearly a measured value of zero indicates that there is no accessible porosity at that diameter; there may, however, be internal pores of this size, but necking in the pore structure blocks access to them.

Figure 4 shows an example of the results obtained by molecular probe ellipsometry. The films were HCl catalyzed SiO_2 films spin coated on silicon substrates and baked for 30 min in air at different temperatures. For this material, when annealed at temperatures up to 600 °C, the accessible volume fraction drops slowly with adsorbate size, but for annealing at 700 °C, this drop is much more rapid, indicating that at this temperature the size range of the pores becomes much more restricted. This technique was used to study in detail the effect of the molar ratio of water to metal precursor in sol-gel silica films (Fardad et al. 1995) and to characterize sol-gel protective coatings on stainless steel (Fuentes-Gallengo et al. 1997). In the former case, pure silica films were prepared from HCl-catalyzed sols based on TEOS (tetra-ethyl-orthosilicate) in ethanol solution, with different molar ratios of TEOS to water. Various film parameters were then studied with respect to annealing temperature in the range 100–1000 °C, namely, film thickness and index, porosity and pore size distribution, scattered light intensity, FTIR spectra, and film stress. Water vapor isotherms were also taken as described in Martin and Green (1990). A number of conclusions could be drawn from this study:

- Pore size results from MPE and water vapor isotherms correlate reasonably well, both indicating that pores below 20 Å diameter are responsible for the bulk of the porosity.
- Increasing water ratio R decreases both porosity and pore size.
- The annealing temperature required to eliminate the larger pores within this range depends strongly on R, but annealing at above 700 °C is needed in all cases to eliminate porosity accessible to water.

Such information about pore morphology has also been very useful in developing nanoporous sol-gel films for doping with semiconductor nanocrystals (Dawnay et al. 1997). In this case, crystallites of CdS and other II–VI semiconductors are desired to provide a strong optical nonlinearity, for applications in optical switching and signal processing. In this size regime, the bandgap of the crystals varies strongly with size, and tailoring of this bandgap to the operating wavelength is needed to obtain maximum nonlinearity with minimum linear and nonlinear absorption losses. For this reason, the fabrication method should provide both defined pore sizes and a narrow pore size distribution. MPE was used to show that this was achieved, and the pores, having been stabilized at temperatures high enough to eliminate most organic content, were found to provide a strong template for the growth of the nanocrystals.

References

Alvarez-Herrero A, Fort AJ, Guerrero H, Bernabeu E. Ellipsometric characterization and influence of relative humidity on TiO_2 layers optical properties. Thin Solid Films. 1999;349:212–9.

Archer RJ. Determination of the properties of films on silicon by the method of ellipsometry. J Opt Soc Am. 1962;52:970–7.

Dawnay EJC, Fardad MA, Green M, Horowitz F, Yeatman EM, Almeida RM, Vasconcelos HC, Guglielmi M, Martucci A. Control and characterisation of microstructure in sol–gel films for optical device applications. In: Vincenzini P, editor. Advanced materials in optics, electro-optics & communication technology. Forli, Italy: Techna Srl; 1995. p. 55–62.

Dawnay EJC, Fardad MA, Green M, Yeatman EM. Growth and characterisation of semiconductor nanoparticles in porous sol–gel films. J Mater Res Soc. 1997;12:3115–26.

Den Englesen D. Ellipsometry of anisotropic films. J Opt Soc Am. 1971;61:1460–6.

Fardad MA, Yeatman EM, Dawnay EJC, Green M, Horowitz F. Effects of H_2O on structure of acid-catalysed SiO_2 sol–gel films. J Non-Cryst Solids. 1995;183:260–7.

Fuentes-Gallengo JJ, Blanco E, Esquivias L. Ellipsometric characterization of an AISI 304 stainless steel protective coating. Thin Solid Films. 1997;301:12–6.

Gartner M, Crisan M, Jitianu A, Scurtu R, Gavrila R, Oprea I, Zaharescu M. Spectroellipsometric characterization of multilayer sol–gel Fe_2O_3 films. J Sol–Gel Sci Technol. 2003;26:745–8.

Jenkins TE. Multiple-angle-of-incidence ellipsometry. J Phys D Appl Phys. 1999;32:R45–56.

Martin AJ, Green M. Sol–gel nano-porous silica–titania thin films with liquid fill for optical interferometric sensors. Proc Soc Photo-Instrum Eng. 1990;1328:352–62.

McDonagh C, Bowe P, Mongey K, MacCraith BD. Characterisation of porosity and sensor response times of sol–gel-derived thin films for oxygen sensor applications. J Non-Cryst Solids. 2002;306:138–48.

Neuschaefer-Rube U, Hozapfel W, Wirth F. Surface measurement applying focusing reflection ellipsometry: configurations and error treatment. Measurement. 2003;33:163–71.

Rothen A. The ellipsometer, an apparatus to measure thicknesses of thin surface films. Rev Sci Instrum. 1945;16:26–30.

Rothen A. Measurements of the thickness of thin films by optical means. In: Passaglia E, Stromberg RR, Kruger J, editors. Rayleigh and Drude to Langmuir, and the development of the present ellipsometer, vol. 256, National Bureau of Standards. Gaithersburg, MD, USA: Miscellaneous Publication; 1974. p. 7–21

Schubert M, Dollase W. Generalized ellipsometry for biaxial absorbing materials: determination of crystal orientation and optical constant of Sb_2S_3. Opt Lett. 2003;27:2073–5.

Sengupta SS, Park SM, Payne DA, Allen LH. Origins and evolution of stress development in sol–gel derived thin layers and multideposited coatings of lead titanate. J Appl Phys. 1998;83:2291–6.

Snyder PG, Rost MC, George H, Woollam JA, Alterovitz SA. Variable angle of incidence spectroscopic ellipsometry: application to G_aA_s–Al_xGa_{1-x} As multiple heterostructures. J Appl Phys. 1986;60:3293–302.

Yeatman EM, Green M, Dawnay EJC, Fardad MA, Horowitz F. Characterisation of microporous sol–gel films for optical device applications. J Sol–Gel Sci Technol. 1994;2:711–5.