Limitations of x-ray reflectometry in the presence of surface contamination

D L Gil\(^1\) and D Windover

National Institute of Standards and Technology, 100 Bureau Dr., Stop 8520, Gaithersburg, MD 20899-8520, USA

E-mail: windover@nist.gov

Received 9 January 2012, in final form 13 March 2012
Published 22 May 2012
Online at stacks.iop.org/JPhysD/45/235301

Abstract

Intentionally deposited thin films exposed to atmosphere often develop unintentionally deposited few-monolayer films of surface contamination. This contamination arises from the diverse population of volatile organics and inorganics in the atmosphere. Such surface contamination can affect the uncertainties in determination of thickness, roughness and density of thin-film structures by x-ray reflectometry (XRR). Here we study the effect of a 0.5 nm carbon surface contamination layer on thickness determination for a 20 nm titanium nitride thin film on silicon. Uncertainties calculated using Markov-chain Monte Carlo Bayesian statistical methods from simulated data of clean and contaminated TiN thin films are compared at varying degrees of data quality to study (1) whether synchrotron sources cope better with contamination than laboratory sources and (2) whether cleaning off the surface of thin films prior to XRR measurement is necessary. We show that, surprisingly, contributions to uncertainty from surface contamination can dominate uncertainty estimates, leading to minimal advantages in using synchrotron-over laboratory-intensity data. Further, even prior knowledge of the exact nature of the surface contamination does not significantly reduce the contamination’s contribution to the uncertainty in the TiN layer thickness. We conclude, then, that effective and standardized cleaning protocols are necessary to achieve high levels of accuracy in XRR measurement.

(Some figures may appear in colour only in the online journal)

1. Introduction

X-ray reflectometry (XRR) is widely used for characterizing the thickness, roughness and density of nanometre-scale thin films. Because it uses wavelengths of a similar or smaller scale relative to the thicknesses of the layers being studied, the resulting data have a relatively direct connection to the structure and therefore has a rather straightforward traceability to the International System of Units (SI)\([1–3]\). This is a significant advantage over other techniques like spectroscopic ellipsometry (SE), whose results are somewhat more difficult to interpret. But XRR is still somewhat sensitive to surface contamination. A comprehensive study by Seah\( et\ al\) on ultrathin SiO\(_2\) layers on Si has shown offsets between XRR, neutron reflectometry (NR), SE and x-ray photoelectron spectroscopy (XPS)\([4]\). These offsets are believed likely due to the different effects contamination has for each technique. To use XRR for high-accuracy measurements of film thickness and roughness—as the National Institute of Standards and Technology (NIST) plans to do for thin-film standards—these effects must be quantified.

The primary difficulty in quantifying the effects of contamination is that thin surface contamination layers are hard to measure. The surface contamination with which we are concerned typically takes the form of rough, near-monolayer-thickness, carbonaceous compounds. This is challenging to measure with x-ray or neutron reflectivity-based methods: carbon’s low scattering factor for both techniques and the poor quality of the carbon layers produce low-contrast fringes.

XPS and other inelastic scattering techniques can be used to measure the quantity per surface area—and thus relative thicknesses—of contamination layers, but generally...
require calibration in order to measure absolute thicknesses [4]. This calibration is often, for denser materials, conducted by reflectometry measurements. In addition, these experiments are typically conducted in ultra-high-vacuum, which may change the properties, or even the very presence, of the adsorbed volatile contamination layers. This makes using these experiments somewhat challenging for achieving high accuracy in absolute thickness determination.

But a crucial question is whether these experiments are needed at all. How sensitive are other parameters determined by modelling of XRR data to the presence of an unknown, hard-to-measure layer of surface contamination? (E.g. how much do you have to know about contamination to achieve a given level of uncertainty in thickness for other layers within a structure?) And what sort of data quality (i.e. dynamic and q-space/angular range) is needed to achieve desired levels of uncertainties? (E.g. do you need synchrotron data?)

We study these questions by a simulation-based study of the effects of a carbon contamination layer on a TiN film on Si substrate structure being considered by NIST for an XRR standard. Using a Bayesian statistical approach to XRR data analysis, we estimate uncertainties for structural parameters of the high-Z TiN layer under various contamination and data quality conditions.

2. Background

2.1. X-ray reflectometry

XRR can be used to measure the density, thickness and roughness of thin films which are laterally homogeneous at the scale of the beam [5, 6]. We limit ourselves here to the case of layered structures with fairly sharp interfaces and no density grading other than that provided by roughness. Density is measured by the critical angle for total external reflection and through the careful analysis of oscillation amplitudes; thickness is measured by the period of intensity oscillations appearing after the critical angle; and roughness is measured by the rapidity of overall intensity and oscillation intensity fall-off. The analysis here uses the Parratt recursion [7] with the perturbation for Gaussian roughness described by [8].

Analysing XRR data is an inverse problem: because only the intensity—rather than the complex amplitude—of the reflected beam can be measured, in general there is no unique structure determined by an XRR pattern. The typical approach is to fit the parameters of a multilayer structure using an optimization approach, the most common being genetic algorithms (GAs) [9–11]. Although optimization finds a best-fit solution—and thus a best estimate of the parameter values—it does not provide estimates of uncertainties on the parameters. To calculate uncertainties, a more sophisticated—and computationally expensive—approach is necessary. NIST has developed statistical Markov chain Monte Carlo (MCMC) methods in order to obtain parameter estimations and uncertainties within a Bayesian formalism [12].

2.2. Data analysis

To make inferences about physical structure from XRR data by GA or MCMC methods requires a physical model that relates structural parameters to idealized non-noisy XRR data. This is provided by the Parratt recursion with Nevot–Croce roughness described above.

A physical model is not sufficient, however, because the data collected are (at best) noisy and (at worst) corrupted by systematic instrument effects. There must be, in addition, a statistical model of the data. For the GA method, this is a $\chi^2$ cost function; for the MCMC method, a probability density function. The cost function and probability density function employed here make very similar assumptions about the statistical characteristics of the data, see [13]. Nonetheless, the GA and MCMC methods each recover different types of information from the data.

XRR data consist of pairs of angles and measured intensities. For this study, we assume data free of angular errors with counting-error-limited measured intensities. The error in measured intensities is modelled using a log-normal likelihood with standard deviation of the square root of the calculated intensities; this approximates well a Poisson (i.e. counting-statistic) likelihood [13].

We use a tiered data analysis architecture. XRR model parameters are first determined for a given structural model using a GA optimization approach [11]. In this analysis, we used a 1000 genome population evolved over 1000 generations to obtain best-fit structural parameters. This structural information was then used to initialize the starting parameters for a tuned MCMC sampler to provide us with probability distributions for each parameter within a given structural model. Each MCMC was allowed a 50 000 steps conditioning run to tune the MCMC target dimensions. The MCMCs were then run for 250 000 steps to obtain adequate statistics for inter-comparison. The tiered analysis was performed several times on each dataset to validate the refinement stability.

2.3. Simulated data

Two structures were simulated for this study: case 1—a single layer of TiN on an Si substrate (for film parameters, see table 1)—and case 2—a contaminated single layer of TiN on an Si substrate (see table 2). Simulations were performed under two different data quality conditions (see table 3) selected to compare parameter refinement results between an advanced laboratory instrument and a synchrotron measurements from, e.g., a third-generation bending magnet beamline. By way of illustration, we present XRR data simulations from our case 2 structure for laboratory (see figure 1) and synchrotron (see figure 2) data quality conditions. By considering these two cases, we can answer an oft-asked question for XRR data-collection. How many orders of magnitude data quality are needed to determine a given XRR model parameter? The Bayesian statistical approach used here can directly answer this question by determining the ‘best-case’ theoretically possible from XRR measurements for any refined parameter.

2.4. MCMC analysis

The MCMC analysis method has initial optimal parameters input using the results of a GA. Details of MCMC methods are beyond the scope of this paper. The most important
Table 1. Case 1: clean TiN/Si structure. Composition, thickness ($t$), roughness ($\sigma$) and density ($\rho$).

|       | $t$ (nm) | $\sigma$ (nm) | $\rho$ (g cm$^{-2}$) |
|-------|---------|---------------|----------------------|
| TiN   | 20.0    | 0.5           | 4.90                 |
| Si    | —       | 0.4           | 2.49                 |

Table 2. Case 2: carbon-contaminated TiN/Si structure.

|       | $t$ (nm) | $\sigma$ (nm) | $\rho$ (g cm$^{-2}$) |
|-------|---------|---------------|----------------------|
| C     | 0.5     | 0.1           | 2.25                 |
| TiN   | 20.0    | 0.5           | 4.90                 |
| Si    | —       | 0.4           | 2.49                 |

Table 3. Parameters of XRR data quality cases used in simulations.

| Case         | $2\theta$ range ($\degree$) | Step size ($\degree$) | Maximum intensity (counts) | Background (counts) |
|--------------|-------------------------------|-----------------------|---------------------------|---------------------|
| Laboratory   | 4.0                           | 0.005                 | $10^6$                     | 1                   |
| Synchrotron  | 7.0                           | 0.005                 | $10^9$                     | 1                   |

Figure 1. Simulated (Cu radiation) XRR data for case 2 structural modal and laboratory quality data (see tables 2 and 3). XRR simulated data (plus signs) has been fit using a GA refinement (solid line) and yielded nearly identical parameters to those used in the simulation (table 2). Note refinement quality (magnified regions).

Figure 2. Simulated (Cu radiation) XRR data for case 2 structural modal and synchrotron quality data (see tables 2 and 3). XRR simulated data (plus signs) have been fit using a GA refinement (solid line) and yielded nearly identical parameters to those used in the simulation (table 2). Note refinement quality (magnified regions).

Table 4. Model 1: allowed MCMC (uniform prior) ranges for TiN/Si structure with no surface contamination layer.

|       | $t$ (nm) | $\sigma$ (nm) | $\rho$ (g cm$^{-2}$) |
|-------|---------|---------------|----------------------|
| TiN   | 15.0 to 25.0 | 0.01 to 2.5 | 4.0 to 6.0           |
| Si    | —       | 0.01 to 2.5  | 2.0 to 3.0           |

Table 5. Model 2: allowed MCMC (uniform prior) ranges for TiN/Si multilayer with surface contamination layer.

|       | $t$ (nm) | $\sigma$ (nm) | $\rho$ (g cm$^{-2}$) |
|-------|---------|---------------|----------------------|
| C     | 0.0 to 2.0 | 0.01 to 2.5 | 2.0 to 3.0           |
| TiN   | 15.0 to 25.0 | 0.01 to 2.5 | 4.0 to 6.0           |
| Si    | —       | 0.01 to 2.5  | 2.0 to 3.0           |

Table 6. Model 2a: allowed MCMC range for TiN/Si multilayer with a known surface contamination layer.

|       | $t$ (nm) | $\sigma$ (nm) | $\rho$ (g cm$^{-2}$) |
|-------|---------|---------------|----------------------|
| C     | 0.5     | 0.1           | 2.25                 |
| TiN   | 15.0 to 25.0 | 0.01 to 2.5 | 4.0 to 6.0           |
| Si    | —       | 0.01 to 2.5  | 2.0 to 3.0           |

3. Results

Statistically determined uncertainties for laboratory and synchrotron levels of data quality have been calculated using the MCMC method for a simulated clean TiN sample (case 1) with corresponding modelling ranges (model 1) and a simulated carbon-contaminated TiN sample (case 2) with its corresponding modelling ranges (model 2). Absolute and
relative uncertainties for each structural parameter and each data quality are presented. We also present a modified analysis for case 2, in which we provide the exact parameters for the contamination layer as prior information for the MCMC method (model 2a) and discuss the resulting uncertainties.

3.1. Clean sample—case 1

The power of the Bayesian analysis via MCMC is through its generation of posterior probability distributions for each parameter within a physical model, clearly showing the uncertainty ranges (for example, see figure 3). The expanded uncertainties can be directly calculated by finding the parameter bounds for the probability distribution plot area representing the 95% highest probability. For case 1, these expanded uncertainty ranges are tabulated in table 7.

For a clean, single-layer structure, there is a clear (factor of two) advantage to synchrotron measurements with regards to determination of accurate thickness and roughness information. This statistical determination method for uncertainty estimation is absent from optimization refinement methods such as GAs. Studying two-dimensional (two simultaneous parameters) posterior probability distributions allows us to qualitatively and quantitatively explore parameter correlations. The clear improvements in TiN thickness and roughness seen in table 7 are due to the orthogonal (no correlation) nature of thickness and roughness (see figure 4). As a general rule, when no correlation exists between parameters within a refinement, then better data quality will directly correspond to reduced uncertainties for the parameters in question. However, if correlations do exist between two or more parameters, as for example between film roughness and film density (see figure 5) in case 1, then correlations will introduce intrinsic parameter uncertainties which cannot be further reduced with higher data quality.

As seen by studying the ratios of uncertainty estimates (last column in table 7), when determining the density of either the TiN or the Si substrate, there is no clear advantage between the laboratory and the synchrotron levels of data quality. The relative quality of density determination is nearly identical in both cases (uncertainty ratios equal to 1). This constant nature of density uncertainty over both levels of data quality is likely caused by two factors: first, that both datasets have the same spacing between collection points, so that the critical angle is not much more precisely determined by synchrotron data than the laboratory data. Second, density correlates with other modelling parameters, for example, interface roughness (see figure 5).

3.2. Carbon-contaminated film—case 2

For case 2, we present the expanded uncertainty ranges in table 8 and see several surprising results. When one introduces a carbon contamination layer, the advantages in reduced uncertainties from using the higher data quality of a synchrotron vanishes for all but roughness determination. For ratios of unity or near unity, (e.g. 0.93, 0.75) there is no clear advantage to synchrotron data. [Apparent disadvantages to synchrotron data (i.e., ratios less than one) are artefacts to the coarseness of our sampling analysis.] This is partially a consequence of high inverse correlation between contamination layer thickness and TiN thickness (see figure 6). Correlations also exist between the contamination density and surface roughness (see figure 7), further expanding uncertainties throughout the model parameters.

When one examines only the highest quality data (synchrotron), the effect of clean versus contaminated surfaces
Figure 5. Two-dimensional histogram showing correlation between TiN density ($\rho_1$) and TiN interface roughness ($\sigma_1$) for case 1. Intensity scale of histogram shows the relative frequency with which the MCMC explores a given parameter space.

Table 8. MCMC-determined expanded uncertainties (95% probability intervals) for the model parameters of case 2, model 2, and the ratio of these uncertainties, $[U_{\text{lab}}/U_{\text{sync}}]$.

| Parameter | U (synchrotron) | U (laboratory) | $[U_{\text{lab}}/U_{\text{sync}}]$ |
|-----------|----------------|----------------|-------------------------------------|
| $t_C$     | 0.64 nm        | 0.66 nm        | 1.0                                 |
| $t_{\text{TiN}}$ | 0.67 nm        | 0.62 nm        | 0.93                                |
| $t_{\text{C} + \text{TiN}}$ | 0.32 nm        | 0.24 nm        | 0.75                                |
| $\sigma_C$ | 0.033 nm       | 0.16 nm        | 4.8                                 |
| $\sigma_{\text{TiN}}$ | 0.50 nm        | 0.48 nm        | 0.96                                |
| $\sigma_S$ | 0.027 nm       | 0.20 nm        | 7.4                                 |
| $\rho_C$  | 0.61 g cm$^{-2}$ | 0.92 g cm$^{-2}$ | 1.5                             |
| $\rho_{\text{TiN}}$ | 1.49 g cm$^{-2}$ | 1.28 g cm$^{-2}$ | 0.86                             |
| $\rho_S$  | 0.94 g cm$^{-2}$ | 0.94 g cm$^{-2}$ | 1                                 |

can be directly compared. In table 9, we see that only the TiN thickness and roughness show pronounced reductions in uncertainty ranges from a contaminated versus clean structure.

An astute observer may wonder why the uncertainties for Si and TiN densities are not improved through the removal of the carbon contamination layer. This is because XRR, in some cases, is simply not sensitive to a given model parameter. This sensitivity issue can be distinguished from a correlation phenomena, again using the MCMC posterior probability densities or two-dimensional histograms, and looking for parameters which produce uniform posteriors out of the analysis. In figure 8, we see that the Si density probability density is nearly uniform over the allowed range of the parameter. This lack of a pronounced peak demonstrates very little sensitivity to Si density in our data.

3.3. Contamination of known thickness, roughness, and density—case 2a

In table 10, we introduce the results from our known parameter carbon contamination case. We compare the uncertainties between clean, unknown contamination, and exactly known contamination cases. The most interesting feature is the TiN thickness. Even for the case where the contamination thickness, roughness and density are known a priori, the model still has five times higher TiN thickness uncertainty over the clean surface case. There is a reduction of a factor of 2 over the unknown contamination case; this reduction indicates that not knowing the exact properties of the contamination layer does have an effect on uncertainties. (Or, identically, that refining an unknown contamination layer increases the uncertainties.) But this effect is much smaller than the effect of the mere presence of the contamination layer. This is because the presence of the contamination layer causes a decrease in the contrast of the TiN layer thickness fringes, decreasing the ability of higher quality data to provide more information. This manifests itself in correlations in the model which can substantially increase the overall uncertainties for the TiN layer thickness.

The very presence of contamination on the structure—rather than the need to fit the contamination—has the largest effect on the uncertainty.

4. Conclusions

There are some caveats to this conclusion: Only simulated data have been considered. All systematic instrumental errors
showing lack of sensitivity for this parameter within the XRR model
synchrotron data is made only for the case of a Cu Kα
have been neglected. No instrument response functions have
been modelled. The comparison between laboratory and
laboratory source radiation and an synchrotron beamline set to
beam for specific materials and structures has been neglected.
(For an example of XRR fit improvement through judicious
source energy tuning, as not being likely to
improve data quality, the
MCMC XRR analysis technique provides a powerful tool for
studying the theoretical limitations of XRR measurements of
a structure before taking measurements.

In this case of a carbon contamination layer, we see that the
theoretical uncertainty estimates for parameters are dominated by
correlations between the surface contamination thickness and the TiN thickness increasing the uncertainty estimates for
our thin film of interest whenever the carbon contamination
layer is present. Even when armed with prior knowledge of all parameters for the contamination layer, we see a five-
fold increase in the TiN thickness uncertainty caused by the
introduction of the contamination layer. Higher data quality will provide significant reductions in parameter uncertainties for simple XRR models, such as the clean TiN thin film. However, in the presence of contamination, we see minimal gain through enhanced data quality for the determination of thin film thickness.

This MCMC simulated data study has shown that removing the contamination is essential to significantly
reducing the uncertainties in the high-Z layer thickness measurement.

Acknowledgments

We would like to thank Victor Varntanian of the International SEMATECH Manufacturing Initiative (ISMI) (Albany, NY) for providing pre-standard test structures and XRR measurements for XRR SRM development at NIST. We would also like to thank P Y Hung of Sematech (Albany, NY) for extensive and very helpful discussions on thin-film characterization and for providing numerous interesting sets of XRR data for the development of analysis techniques.

References

[1] Bowen D K, Matney K M and Wormington M 1998 X-ray metrology for ULSI structures AIP Conf. Proc. 449 928–32
[2] Bowen D K and Deslattes R D 2001 X-ray metrology by diffraction and reflectivity AIP Conf. Proc. 550 570–9
[3] Hasche K, Thomsen-Schmidt P, Krumrey M, Ade G, Ulm G, Stuempfel J, Schaeldich S, Frank W, Procop M and Beck U 2003 Metrological characterization of nanometer film thickness standards for XRR and ellipsometry applications Proc. SPIE—Int. Soc. Opt. Eng. 5190 165–72
[4] Seab M P and Spencer S J 2003 Ultrathin SiO₂ on Si: I. Quantifying and removing carbonaceous contamination J. Vac. Sci. Technol. A 21 345–52
[5] Chason E and Mayer T M 1997 Thin film and surface characterization by specular x-ray reflectivity Crit. Rev. Solid State Mater. Sci. 22 1–67
[6] Pietsch U, Vaclav H and Baumbach T 2004 High-Resolution X-Ray Scattering: from Thin Films to Lateral Nanostructures (New York: Springer)
[7] Parratt L G 1954 Surface studies of solids by total reflection of x-rays Phys. Rev. 95 359
[8] Nevot L and Croce P 1980 Characterization of surfaces by grazing x-ray reflection—application to study of polishing of some silicate-glasses Rev. Phys. Appl. 15 761–79
[9] Dane A D, Veldhuis A, de Boer D K G, Leenaers A J G and Buydens L M C 1998 Application of genetic algorithms for characterization of thin layered materials by glancing incidence x-ray reflectometry Physica B 253 254–68
[10] Ulyanenkov A and Sobolewski S 2005 Extended genetic algorithm: application to x-ray analysis J. Phys. D: Appl. Phys. 38 A235–8
[11] Wormington M, Panacccione C, Matney K M and Bowen D K 1999 Characterization of structures from x-ray scattering data using genetic algorithms Phil. Trans. R. Soc. Lond. A 357 2827–48
[12] Windover D, Gil D L, Cline J P, Henins A, Armstrong N, Hung P Y, Song S C, Jammy R and Diebold A 2007 NIST

Table 9. Comparison of MCMC-determined expanded uncertainties (95% probability intervals) between clean versus contaminated cases for synchrotron quality data.

| Parameter | U(clean) (synchrotron) | U(with carbon) (synchrotron) | [U(carbon) / U(clean)] |
|-----------|------------------------|----------------------------|-----------------------|
| $t_{\text{TiN}}$ | 0.048 nm | 0.67 nm | 14 |
| $\sigma_{\text{surface}}$ | 0.033 nm | 0.033 nm | 1 |
| $\sigma_{\text{TiN}}$ | 0.050 nm | 0.5 nm | 15 |
| $\rho_{\text{TiN}}$ | 1.49 g cm⁻² | 1.49 g cm⁻² | 1 |
| $\rho_{\text{Si}}$ | 0.94 g cm⁻² | 0.94 g cm⁻² | 1 |

Table 10. Comparison of MCMC-determined expanded uncertainties (95% probability intervals) between clean, unknown, and known contamination layer cases for synchrotron quality data.

| Parameter | U(clean) | U(carbon) | U(known carbon) |
|-----------|---------|----------|-----------------|
| $t_{\text{TiN}}$ | 0.048 nm | 0.67 nm | 0.31 nm |
| $\sigma_{\text{surface}}$ | 0.033 nm | 0.033 nm | Fixed |
| $\sigma_{\text{TiN}}$ | Same as $\sigma_{\text{surface}}$ | 0.5 nm | 0.36 nm |
| $\rho_{\text{TiN}}$ | 0.050 nm | 0.027 nm | 0.027 nm |
| $\rho_{\text{Si}}$ | 1.49 g cm⁻² | 1.49 g cm⁻² | 1.49 g cm⁻² |

Figure 8. Posterior probability density for Si substrate density ($\rho_{\text{Si}}$) showing lack of sensitivity for this parameter within the XRR model for case 2.
method for determining model-independent structural information by x-ray reflectometry. *AIP Conf. Proc.* 931 287–91

[13] Sivia D S 1996 *Data Analysis A Bayesian Tutorial* (Oxford: Oxford University Press)

[14] Krumrey M, Gleber G, Scholze F and Wernecke J 2011 Synchrotron radiation-based x-ray reflection and scattering techniques for dimensional nanometrology *Meas. Sci. Technol.* 22 094032