Test of the Consistency of Angle Resolved XPS Data for Depth Profile Reconstruction using the Maximum Entropy Method

A.J. Roberts,[1*] K. Macak[1] and K. Takahashi[2]
[1]*Kratos Analytical Ltd, Wharfside, Trafford Wharf Road, Manchester, M17 1GP, UK*
[2]*Kratos XPS Section, Shimadzu Corp., 380-1 Horiyamashita, Hadano, Kanagawa 259-1304, Japan*
[*]adam.roberts@kratos.co.uk

(Received: January 17, 2009; Accepted: February 24, 2009)

Angle resolved XPS is a useful method for obtaining non-destructive quantification of thin (4-6nm) layers with good absolute depth resolution. Although acquisition of ARXPS data with modern instrumentation is easy, determining the depth distribution of elements from the data is more challenging. The maximum entropy method (MEM) is a technique frequently used for solving the inversion problem in angle resolved XPS experiments. The necessary condition for the consistency of experimentally measured data with the MEM model (successful fit) is that the Laplace transform of the compositional depth profile (LTCDP) calculated from the experimentally measured normalised intensity is a monotonically decreasing function for all measured elements. We have found an efficient algorithm which can estimate the LTCDP for elements with attenuation lengths and provide an independent test of the suitability of the layered model for different sample elements. The practical application of this method is illustrated on analysis of a number of samples and conclusions on the use of these methods to generate elemental and chemical-state depth profiles are discussed.

1. Introduction

The MEMSYS algorithm used by Livesey and Smith[1] can provide an estimate of the optimum value of the regularising parameter (smoothness), of the error bars in the reconstructed profile and of the noise level in the experimental data set. The noise scaling factor represents the cumulative measure of the uncertainties in the user supplied factors such as elemental sensitivity factors, inelastic mean free paths and the overall consistency of the sample structure with the layered model. This assumes that the emission depth distribution function follows simple exponential attenuation law and that the material is constant throughout the sample (without voids). The confidence limits in the reconstructed depth profiles for all elements are proportional to the single noise scaling factor. Because the noise scaling in the current version of the MEM algorithm does not discriminate between the elements we have developed an alternative method to identify the possible major discrepancies between the measured data and the layered model.

1.1 Detected photoelectron intensity

The detected photoelectron intensity for element ***a*** at ***k***th angle is defined by:

$$I_a(\theta_k) = T \cdot A(\phi_k,\theta_k) \cdot \sigma_a \cdot L(\gamma) \int_0^{\infty} c_a(z) e^{\frac{-z}{\lambda_a \cos(\theta_k)}} dz$$

$$\approx \mathcal{L}[c_a(z); p_a(\theta_k)]$$

$$\approx \sum_{i=0}^{N_{layers}-1} n_{a,i} \exp(-p_a(\theta_k) \cdot i \cdot t)$$

Where
$A(\varphi,\theta)$ is the area of analysis on the specimen surface
T – instrument transmission / detector function
$L(\gamma)$ – asymmetry factor
σ – photoionisation cross section
λ – attenuation length
$c(z)$ – number of atoms per unit volume
$p(\theta)=(\lambda \cos(\theta))^{-1}$ – Laplace parameter
$\mathcal{L}[c(z);p(\theta)]$ – Laplace transform of $c(z)$

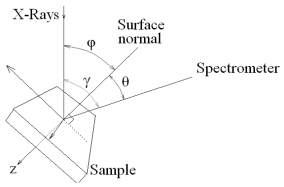
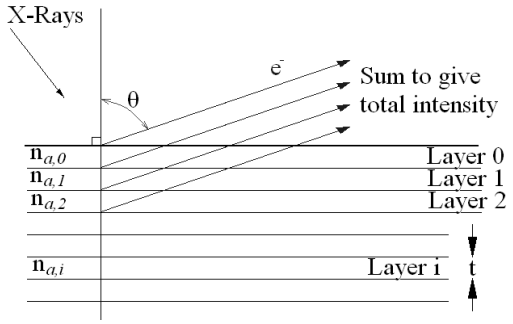


Fig. 1 Geometry of the XPS experiment.

In order to eliminate the effects of varying X-ray flux density and projection of detector area on the tilted sample $A(\varphi,\theta)$ the intensity ratios are used to calculate the apparent composition [2]
A discreet layered sample model for calculation of XPS intensities $X(\theta)$ can be described by the figure below:



$$X_a(\theta_k) = \frac{I_a(\theta_k)/R_a(\theta_k)}{\sum_i^{N_{elements}} I_i(\theta_k)/R_i(\theta_k)}$$

Fig. 2 Discreet layered sample model for calculation of XPS intensities.

where R(θ) is a sensitivity factor.

1.2 MEM Deconvolution of depth profile

The MEMSYS5 [3] algorithm fits the measured data using the Bayesian statistics. The overall probability of the reconstructed depth profile is evaluated from the following contributions:
- Gaussian likelihood of the normally distributed experimental data for the proposed reconstruction profile **n**
- Entropic prior model probability (which includes the parameter called "regularisation constant" by analogy with the statistical regularisation theory).
The optimum values α and **n** are obtained as a result of a search for maximum overall probability. Using the Bayes theorem, it is possible to formally expand the probability by considering the uncertainty in the knowledge of additional user supplied parameters such as inelastic mean free path, elemental sensitivities and indirectly also the validity of layered model assumptions. All these factors contribute to the noise scaling constant c, which is adjusted so that the resulting χ^2 is less than the number of measurements.

1.3 Confidence limits

The confidence limits estimates $\Delta n_{j,i}$ for the depth profile are the individual error bars at each reconstruction point, assuming that all other points are free to take up new values [4]. These are calculated using local curvature of the objective function Q (similar to estimation of the parameter uncertainties in the least square fitting [5]). The noise scaling factor c serves as a scaling factor for both experimental uncertainties σ and the calculated confidence limits $\Delta n_{j,i}$. If there is one outlying point in the measured data , which will increase the overall noise scaling, then the error bars ($\Delta n_{j,i}$) for all elements are affected.

1.4 Laplace Transform of the Compositional Depth Profile

The apparent composition can be expressed in terms of Laplace Transform [6] of the compositional depth profiles. For each θ_l we have the set of linear $N_{elements}$– 1 equations from which we can estimate the $\mathcal{L}[c_a(z);p_a(\theta_k)]$. The set of equations can be inverted using the following two assumptions:
1.) The sum of the concentrations of all elements throughout the profile remains constant.
2.) The $\mathcal{L}[c_y(z);p_y(\theta_l)]$ term can be linearly interpolated or extrapolated from the values of its nearest neighbours $\mathcal{L}[c_y(z);p_y(\theta_{\lambda(y,l)})]$, $\mathcal{L}[c_y(z);p_y(\theta_{\lambda(y,l)+1})]$.
The $\mathcal{L}[c_q(z);p_q(\theta_l)]$ is used to verify the consistency of the measured data with layered model, where the following property is used: The Laplace transform $\mathcal{L}[f(z);s]$ of any non-negative function f(z):f(z)≥ 0 for z≥0; is monotonically decreasing. **If there is a peak in $\mathcal{L}[c_q(z);p_q(\theta_l)]$ then there is no such depth profile which will fit the measured data using the simple exponential form of the depth distribution function.**

1.5 Algorithm Implementation

Calculation of Laplace transform of compositional depth profile was integrated into the MEM solver program allowing immediate check on the consistency of the data. The MEM solver performs the depth profile reconstruction according to the procedure described in papers by Livesey and Smith [2,4]. The program includes the following features:
- optimisation of the regularising parameter
- automatic noise scaling
- calculation of the confidence limits in reconstructed depth profiles
- user defined model for the entropic prior (including model probabilities [7])
- Laplace transform of compositional depth profiles
- measured data can be entered as intensities or atomic fractions

2. Examples of LTCDP in ARXPS Data Analysis

2.1 Very Good Fit for Artificial Data

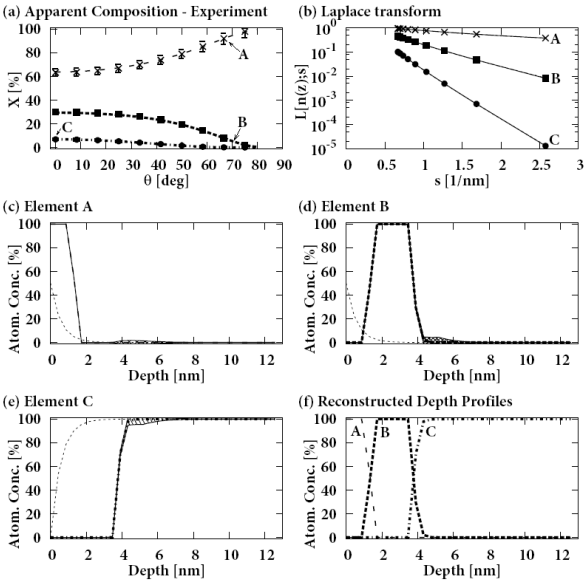


Fig. 3 Artificial ARXPS dataset from paper by Smith and Livesey [2] which is fitted accurately by MEM deconvolution algorithm.(a) Simulated apparent composition profiles which correspond to the measured ARXPS data. Artificial noise of level of 5% (relative) was superimposed on the apparent composition data. Symbols denote input to the MEM deconvolution algorithm: x – element A, ■ - element B, ●- element C. Lines are the apparent compositions calculated from the reconstructed depth profile. (b) Comparison of the Laplace transform plots obtained from reconstructed depth profiles (lines) with the estimates calculated from measured apparent composition (symbols x-element A, ■ - element B, ●- element C). (c) ,(d) ,(e) Reconstructed depth profiles for elements A,B and C with the corresponding error bars shown as shaded regions. The profiles are very close to the original hypothetical layered stack structure and the error bars are small. Entropic prior model for each element is shown as thin dotted line. (f) Comparison of reconstructed depth profiles for elements A,B and C.

Random Guassian noise of standard deviation 0.05 times the apparent composition was then added to the simulated apparent concentrations to form the simulated noisy data set. For each data point the 5% confidence limits are shown on the same graph. The maximum entropy reconstruction is shown as the solid line. The reconstruction of the depth profile is very good and the estimated error bars in the reconstruction profile are small. The LTCDP points obtained from the apparent composition data are fitted well with the LTCDP calculated from the reconstructed depth profile.

2.2 Example of Very Poor MEM fit detected by LTCDP

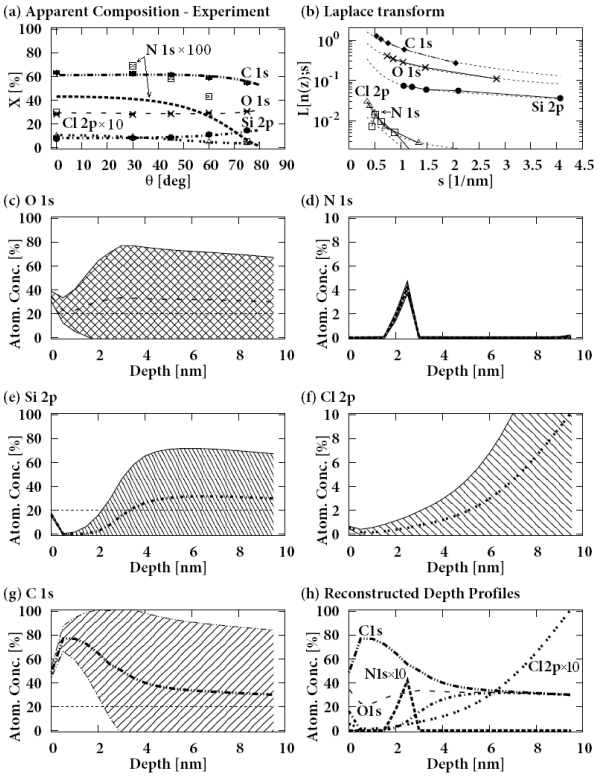


Fig. 4 PDMS overlayer test sample is an example of ARXPS data which are not possible to fit using the simple layered model with constant inelastic mean free path through the depth profile. (a) Measured ARXPS data (symbols: ♦C 1s, x-O 1s,●-Si 2p,Δ-Cl 2p,□ - N 1s) expressed as apparent composition vs. MEM fit (lines). The N 1s data point at 0 deg is difficult to fit and reduces the precision of the resulting fit by a factor of ~23. (b). Laplace transform plots – symbols are calculated directly from the measured apparent composition. The data for Laplace transform plot for N 1s (□) exhibit an unphysical shape at 0-30 deg region (0.5 nm^{-1}). Laplace transform must be always decreasing function of Laplace parameters. The peak is an indication of the problem with the fit. Lines are Laplace transforms calculated from the reconstructed depth profiles. (c),(d),(e),(f) and (g) Reconstructed depth profiles for O, N, Si,Cl and C. There is almost no useful information in the profiles due to large uncertainties. Entropic prior model for each element is shown as thin dotted line.(h) Comparison of the reconstructed depth profiles without error bars.

ARXPS from a surface modified polymeric sample was acquired. The sample is a polymer containing C, O, Cl and a little N as the substrate. Above this there is a PDMS type over-layer (containing O, Si and C only). At the very surface a

thin hydrocarbon contamination layer containing only C and O is expected. This contamination layer is probably about 1 nm thick.

Data for O 1s, Si 2p, Cl 2p are fitted well, but the N 1s fit is very far from the experimental points. The LTCDP plot reveals that there is a peak in N 1s LTCDP plot which violates the condition that all LTCDP plots should be decreasing functions of the Laplace parameter. The LTCDP plots for the remaining elements fulfil this criteria. So the N 1s measurement at 0 degrees is a major contributor to the misfit function and the noise scaling factor in the solution is ~23. The error bars in the reconstructed depth profile are thus 23 times bigger than expected from the estimate of the experimental uncertainties.

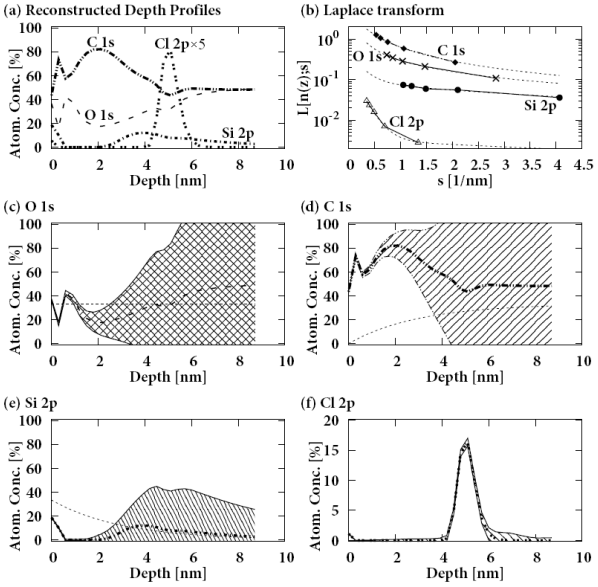


Fig. 5 An improved fit of the data for sample on Figure 4.– low concentration element N 1s was removed. (a) Comparison of reconstructed depth profiles for O,C, Si and Cl. (b) Laplace transform plots – symbols ((symbols: ♦C 1s, x-O 1s,●-Si 2p,Δ-Cl 2p) are calculated directly from the measured apparent composition. All Laplace transform plots are now monotonically decreasing with Laplace parameters. Lines are Laplace transforms calculated from the reconstructed depth profiles. (c),(d),(e) and (f) Reconstructed depth profiles for O,C,Si and Cl shown with error bars which are ~7 times smaller in the surface region than for the original dataset with N1s (Figure 4). Entropic prior model for each element is shown as thin dotted line.

2.3 Improved fit for sample explored in 2.2

The same data as in the previous example with N 1s signal removed. The N 1s element has a low concentration, so it should not affect the depth profiles of the remaining elements. The noise scaling factor is reduced 7 times, and as a result the depth profiles have much smaller error bars.

3. Conclusions

The automatic noise scaling of the MEMSYS algorithm accumulates the errors from all measurements and does not discriminate between individual elements. The reduced accuracy of the fit for one element (even low concentration dopants) thus affects the accuracy of the reconstructed depth profiles of all elements.

The exponential form of the depth distribution allows to extract the plot of the Laplace transform of the compositional depth profile for individual elements which can be used to identify the outlying points with major contributions to the χ^2 sum.

If the Laplace transform plot is not a monotonically decreasing function, then the simple model with exponential depth distribution function is not suitable for the measured data, and there will be large deviation in the fit.

4. References

[1] G.C. Smith and A.K. Livesey, Surf. Interface Anal. **19,** 175-180 (1992).
[2] P. J. Cumpson, J. Electron Spectrosc. Relat. Phenom. **73**, 25-52 (1995).
[3] S.F. Gull and J. Skilling, "Quantified Maximum Entropy MemSys5 User's Manual", Version 1.2, September 8, Maximum Entropy Data Consultants Ltd. (1999).
[4] A. K. Livesey and G.C. Smith, J. Electron Spectrosc. Relat. Phenom. **67,** 439-961 (1994).
[5] W.H. Press, S.A. Teukolsky, W. T. Vetterling and B.P. Flannery, "Numerical Recipes in C", Chapter 15, "Modelling of Data", ISBN 0-521-43108-5, Cambridge University Press (1992).
[6] K. Macak and A.J. Roberts, J. Vac. Sci. Technol. A, submitted.
[7] J. Skilling, "Classic Maximum Entropy" in "Maximum Entropy and Bayesian methods" ed. J. Skilling, pp. 45-52, Kluver Academic Publishers (1989).