Bayesian error analysis of Rutherford backscattering spectra

N.P. Barradas, C. Jeynes*, M. Jenkin, P.K. Marriott

School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey GU2 5XH, UK

Abstract

Rutherford backscattering spectrometry has been used to obtain quantitative and traceable information about homogeneous films. However, inhomogeneous films cannot be analysed with the same manual calculation methods. We have previously demonstrated that a machine algorithm using simulated annealing is available to extract depth profiles from spectra obtained from such inhomogeneous samples. In this work we show how a Bayesian error analysis can be used to construct error bounds on these depth profiles reflecting the uncertainty introduced by Poisson noise on the data. Hence we present, for the first time, RBS analysis which not only gives fully quantitative depth profiles for complex samples, but also reliable estimates of the errors due to collection statistics on these profiles. © 1999 Elsevier Science S.A. All rights reserved.

1. Introduction

Rutherford backscattering spectrometry (RBS) is a well established and powerful technique for determining the elemental composition of thin films using an energetic light ion beam (say, 1.5 MeV ^{4}He$^+$). Because the data analysis is very simple in principle, RBS can be used to obtain quantitative and traceable information about homogeneous films. It has been used in a round robin exercise organised by the National Physical Laboratory [1] to characterise the 100 nm Ta_2O_5 standard for electron spectroscopy. Another recent example of the use of RBS in standards work is the characterization of a secondary implanted standard by the Institute for Reference Materials and Measurements in Geel [2]. We demonstrated 1% accuracy recently in the determination of the In content of InGaAs films [3].

However, most interesting real films are not homogeneous. For example, for a metal film on silicon some of the silicon may move toward the surface which would mean that scattering events from it will be added to the scattering from the metal. In such a case it is usually very hard to devise a computational method transparent enough to obtain depth profiles from the spectrum whose accuracy is traceable. Moreover, most spectra obtained from real samples are sufficiently complex to preclude manual extraction of any accurate depth profiles at all in a reasonable time.

We have recently demonstrated that the general inverse RBS problem (obtaining the depth profile from the spectrum) can be efficiently solved by machine using a simulated annealing (SA) algorithm [4–10]. This work is illustrative of very hard cases that baffle analysts using traditional manual methods.

In this paper we extend this work to show how the computational method is able to generate information about the errors involved in the depth profiles obtained. These methods based on Bayesian statistics have previously been applied to RBS data, but not to solve the general inverse problem, and not in a way which is convenient for routine analysis [11–13]. We are concerned to give analysts a tool which they can use routinely to determine the reliability of the information they extract from the data.

2. Bayesian inference

Suppose that the result of an RBS analysis of a sample gives a spectrum which we denote by $X = (X_1, ..., X_k)$ where k is the number of channels in the spectrum. This spectrum can be viewed as a stochastic vector, at least in the sense that it has been contaminated with noise and/or modelling errors. However we assume an underlying deterministic model which depends on a set of quantities $\Theta = (\theta_1, ..., \theta_n)$. These parameters include some over which we have experimental control, such as the initial beam energy, incident angle, and detector solid angle. It also contain parameters about which we would like to learn, these include the number of layers and the proportions

* Corresponding author. Tel.: + 44-1483-259829; fax: + 44-1483-534139.

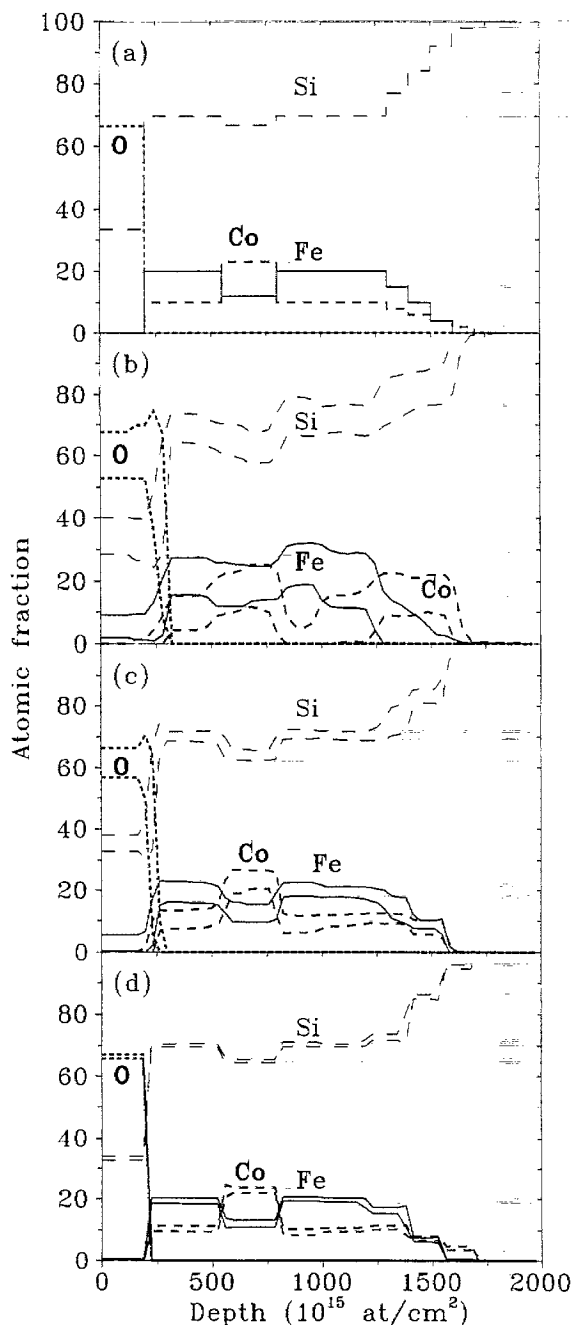*E-mail address:* c.jeynes@surrey.ac.uk (C. Jeynes)

Fig. 1. (a) Theoretical depth profile used to generate the test RBS spectra analysed, with (b) high, (c) medium, and (d) low level of noise. Confidence limits (± 1 standard deviation) of the posterior probability distribution obtained with MCMC for the different elements. Oxygen was only allowed to exist at the surface of the sample, and Co and Fe were restricted to depths below 2×10^{18} at./cm².

of elements in each layer. By combining the standard theory of RBS (see e.g. Ref. [14]) and statistical models of the error mechanisms we can connect X and Θ into a single statistical model, which we denote by $p(X|\Theta)$ where p is a probability density function for the observations X given the parameters Θ. The Bayesian inference problem is to infer information about the parameters Θ given a statistical model and the observed X. There is an exact mathematical answer to this problem given by Bayes' theorem [15] which states

$$p(\Theta|X) = p(X|\Theta)p(\Theta)/p(X) \qquad (1)$$

Using a standard theory of RBS we have a deterministic model which takes the parameter Θ to an expected spectrum. We call this the *forward model*. Let us denote the forward model by the function $\mu(\Theta) = (\mu_1, ..., \mu_k)$ where μ_i is the expected signal in channel i of the spectrum. Assuming for simplicity independent Poisson noise on the channels gives

$$p(X|\Theta) = \Pi_i \exp(-\mu_i(\Theta))\mu_i(\Theta)^{X_i}/X_i! \qquad (2)$$

The second factor we need, the so called *prior* distribution we calculate by conditioning on the number of layers in the sample to get

$$p(\Theta) = \sum p(\Theta^n|n)p(n) \qquad (3)$$

where Θ^n are the parameters when there are exactly n layers. The simplest possibility for a prior on n is to have a uniform distribution on $\{1,...,Max\}$ where Max is some pre-set maximum number of layers. We finally define the prior on Θ^n conditional on a fixed n. In general we have little or no prior knowledge of the proportion of each element in each sample hence we will assume a uniform distribution on the simplex which defines the proportions. Finally, $p(X)$ does not depend on Θ and can be treated as a constant.

3. Calculating the posterior: MCMC

Having found the mathematical form of the distribution $p(\Theta|X)$ the following algorithm makes computation possible. It is an example of a so-called Markov chain Monte Carlo (MCMC) algorithm, in which we construct a sequence of parameter estimates $\Theta_0, \Theta_1, ..., \Theta_n$. We proceed from Θ_i to Θ_{i+1} by considering a candidate ψ generated from some random distribution $q(\Theta_i, \psi)$. Note that the candidate position depends only on the current value of the sequence. Thus we are in fact constructing a Markov chain. The candidate ψ is then accepted (i.e. Θ_{i+1} is set to ψ) with probability $\alpha(\Theta_i, \psi)$ according to the Metropolis criterion

$$\alpha(\Theta_i, \psi) = \min\{[p(\psi|X)q(\psi, \Theta_i)]/[p(\Theta_i|X)q(\Theta_i, \psi)], 1\} \qquad (4)$$

General theory [15,16] tells us that the sequence produced from such an algorithm will be a Markov chain, and that its equilibrium distribution will be the required posterior distribution $p(\Theta|X)$, which contains all the information about the sample given the data. With it, it is possible to calculate the mean solution $\langle \Theta \rangle$ as well as confidence intervals given by the standard deviation $\sigma(\Theta)$.

4. Results and discussion

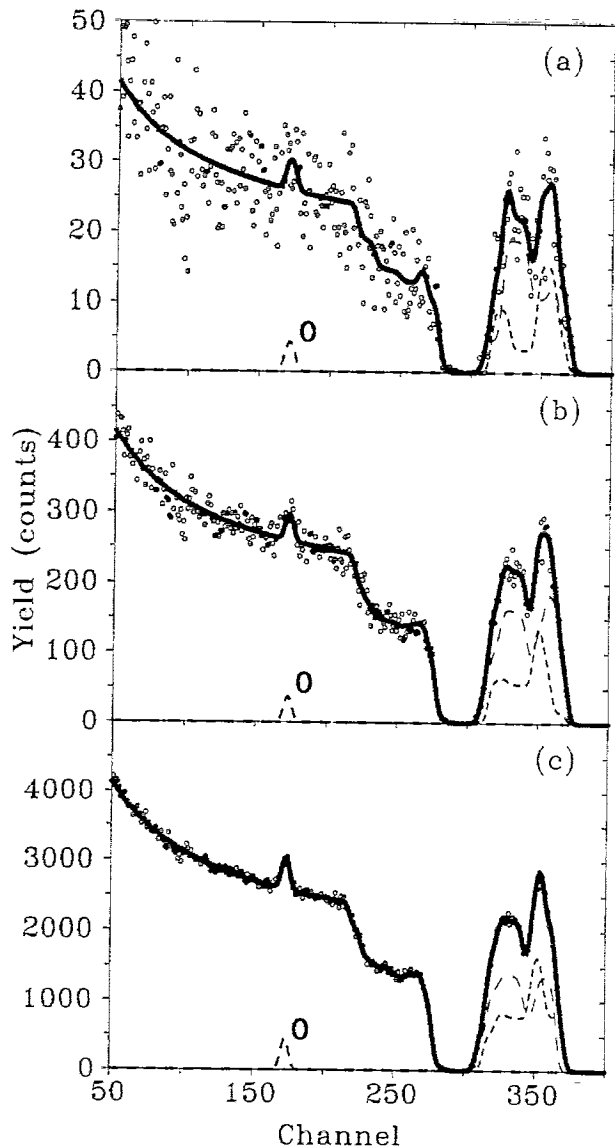We have previously studied the iron–cobalt silicide

Fig. 2. Theoretical test RBS spectra (points) and simulated annealing fit (solid line) obtained for the structure shown in Fig. 1, for (a) high, (b) medium, and (c) low level of noise, corresponding to beam fluence values of 0.1, 1 and 10 μC, respectively. The partial fitted spectra due to the Fe (long dash), Co (small dash), and surface O are also shown as dashed lines.

system with SA [4,7]. The system is interesting because the similar masses of the neighbouring Fe and Co leads to extensively superimposed signals in the RBS spectra. While SA was able to effectively deconvolute the Fe and Co depth profiles, the problem of confidence limits on the fit remained open and will be studied here using MCMC.

We generated theoretical RBS spectra, from the depth profile shown in Fig. 1a, to which we added different levels of random statistical noise on each spectrum, obtaining the results shown in Fig. 2. The spectra correspond to beam fluence values of 0.1, 1, and 10 μC for a solid angle of 2.6 msr and the experimental conditions given in Ref. [7]. The solid lines in Fig. 2 are the fits obtained with SA, during which some constraints were imposed on the parameter

space: oxygen was only allowed to exist at the surface of the sample, as it comes from a surface oxide, and Co and Fe were restricted to depths below 2×10^{18} at./cm^2.

Fig. 1b–d shows the confidence limits (± 1 standard deviation) of the posterior probability distribution for the different elements as obtained with MCMC for the three levels of noise (Markov chain length 3000) where the restrictions to the parameter space described above were enforced. In all cases the profiles match the original profile, within error. As expected, the confidence intervals become narrower for higher statistics. The standard deviation for Fe (at 0.4×10^{18} at./cm^2 depth) is 5.9, 3.0 and 0.9 at.% for the low, medium and high statistics data, respectively. It should be noted, however, that the confidence intervals shown are due only to the statistical error in X, as other sources of uncertainty were not introduced in Eq. (2). For instance, we have previously shown that a 1% uncertainty in the value of the beam fluence can lead to an error in the Fe and Co of up to 20% [7], but less than 3% in the Si. Other uncertainties in the forward model $\mu(\Theta)$, due, e.g. to the limited accuracy of the experimentally determined stopping powers, would also lead to increased error.

Fig. 3 shows the confidence limits (± 1 standard deviation) of the posterior probability distribution for the different elements, as obtained for the low statistics spectrum, after lifting all restrictions to the parameter space. Below 2×10^{18} at./cm^2 the results for Fe and Co are equivalent to those obtained enforcing the parameter space restrictions. This means that for Fe and Co the problem is well determined, i.e. unambiguous, up to that depth. That is due to the fact that the concentration of Fe and Co in that depth range is determined from the data between channels 280 and 380, in which no counts due to Si or O can exist due to the backscattering kinematics. However, the counts in channel 270 can be due to Si at the surface, or metal at about 3×10^{18} at./cm^2. Thus the amount of Si is ambiguous; taking into account the higher cross sections for heavier elements, about 5 at.% metal at 3×10^{18} at./cm^2 leads to about 20 at.% Si at 0.5×10^{18} at./cm^2 (instead of the original 70 at.%). These results mean that the data are consistent with up to 10 at.% metal below 3×10^{18} at./cm^2, which leads to a large ambiguity on the concentration of Si and O due to the lower sensitivity of RBS for light elements.

5. Conclusions

We demonstrate that statistical bounds on the variability of the solution obtained by simulated annealing due to counting noise can be obtained with a sound and robust mathematical procedure using Markov chain Monte Carlo integration. We show incidentally that for the case studied the solution is quite well defined even for very short collection times.

We also demonstrate that the intrinsic ambiguity of RBS is not well characterised by the methods we present, since
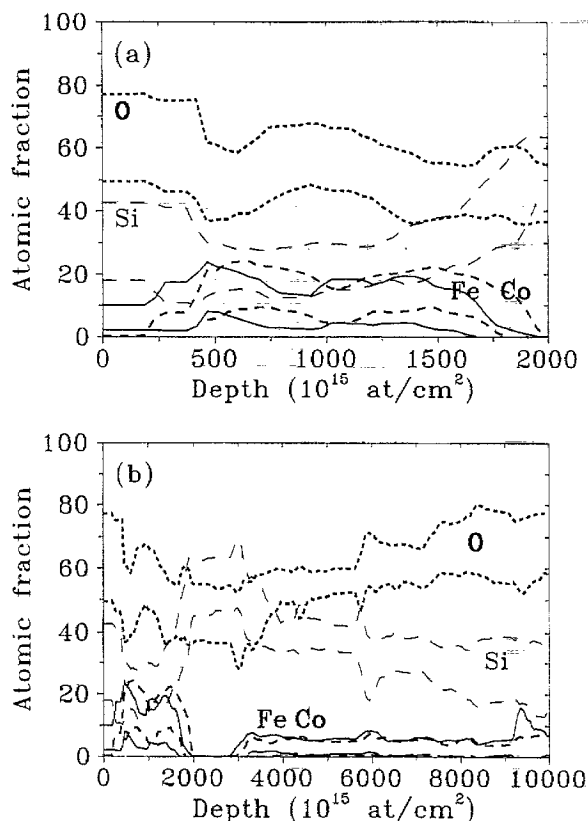
Fig. 3. Confidence limits (±1 standard deviation) of the posterior probability distribution obtained with MCMC for the different elements, for the high noise test spectrum. No restrictions were imposed on any element.

the variance of Θ reflects only the density of states in the solution space. Many similar solutions may exist which correctly yield the observed spectrum, but there may be a few quite different solutions which are equally good, although outside the calculated standard deviation. We believe that this problem is usually easily overcome by simply restricting the solution space searched: this would

be equivalent to introducing more specific priors, but a proper treatment of this is outside the scope of this paper.

As in our previous treatment [4], theoretical spectra were used for simplicity, since artefacts of errors in the forward model (such as stopping power uncertainty and multiple and plural scattering) do not have to be considered.

References

[1] N.P. Seah, D. David, J.A. Davies, C. Jeynes, C. Ortega, C. Sofield, G. Weber, Nucl. Instrum. Methods B 30 (1988) 140.

[2] U. Wätjen, H. Bax, P. Rietveld, Surf. Interface Anal. 19 (1992) 253.

[3] C. Jeynes, Z.H. Jafri, R.P. Webb, M.J. Ashwin, A.C. Kimber, Surf. Interface Anal. 25 (1997) 254.

[4] N.P. Barradas, C. Jeynes, R. Webb, Appl. Phys. Lett. 71 (1997) 291.

[5] N.P. Barradas, P.K. Marriott, C. Jeynes, R.P. Webb, Nucl. Instrum. Methods B 136–138 (1998) 1157.

[6] N.P. Barradas, C. Jeynes, R.P. Webb, U. Kreissig, R. Grötzschel, Nucl. Instrum. Methods B (1999) accepted.

[7] N.P. Barradas, C. Jeynes, M.A. Harry, Nucl. Instrum. Methods B 136–138 (1998) 1163.

[8] N.P. Barradas, C. Jeynes, S.M. Jackson, Nucl. Instrum. Methods B 136–138 (1998) 1168.

[9] N.P. Barradas, C. Jeynes, K.P. Homewood, B.J. Sealy, M. Milosavljevic, Nucl. Instrum. Methods B 139 (1998) 235.

[10] The Data Furnace: Manual for Version v1.0 (October 1997) and update to Version v2.1 (March 1998), University of Surrey, http://www.ee.surrey.ac.uk/Research/SCRIBA/ndf/.

[11] R. Fischer, M. Mayer, W. von der Linden, V. Dose, Phys. Rev. E 55 (1997) 1.

[12] R. Fischer, M. Mayer, W. von der Linden, V. Dose, Nucl. Instrum. Methods B 136–138 (1998) 1140.

[13] V.M. Prozesky, J. Padayachee, R. Fischer, W. von der Linden, V. Dose, R.A. Weller, Nucl. Instrum. Methods B 136–138 (1998) 1146.

[14] W.-K. Chu, J.W. Mayer, M.-A. Nicolet, Backscattering Spectrometry, Academic Press, New York, 1978.

[15] P.M. Lee, Bayesian Statistics: An Introduction, 2nd, Arnold, London, 1997.

[16] C.R. Smith, W.T. Grandy (Eds.), Maximum Entropy and Bayesian Methods in Inverse Problems Reidel, Dordrecht, 1985.