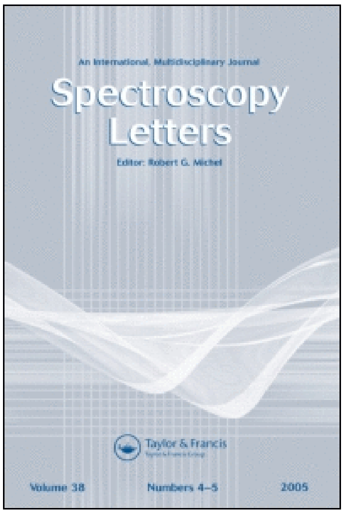

Spectroscopy Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

Bayesian Analysis of Exafs Data

Taimei Yu[a]; Scott L. Whittenburg[a]
[a] Department of Chemistry, University of New Orleans, New Orleans, LA

To cite this Article Yu, Taimei and Whittenburg, Scott L.(1996) 'Bayesian Analysis of Exafs Data', Spectroscopy Letters, 29: 1, 109 — 115

To link to this Article: DOI: 10.1080/00387019608001585

URL: http://dx.doi.org/10.1080/00387019608001585

SPECTROSCOPY LETTERS, 29(1), 109-115 (1996)

BAYESIAN ANALYSIS OF EXAFS DATA

Key words: EXAFS, Bayesian

Taimei Yu and Scott L. Whittenburg

Department of Chemistry, University of New Orleans, New Orleans, LA 70148

Abstract

In this paper we demonstrate the use of Bayesian analysis methods for the analysis of EXAFS data. We will discuss the physical parameters that may be estimated by the method and demonstrate the applicability of the method to Molybdenum coordination compounds.

INTRODUCTION

EXAFS is produced by the interference between an outgoing photoelectron wave and the back scattered wave. The theoretical description of the EXAFS process and the method of data analysis has been thoroughly developed [1]. Here we present only a sufficient amount of information to relate to the Bayesian approach.

The interference between the outgoing and back scattered waves for a single photoelectron event produces a sinusoidally varying signal in momentum-space whose real part is given by

$$\chi(k) = \sum_i \chi_i(k) = \sum_i \frac{m}{2\pi h^2 k^2} t_i(2k) \frac{\sin[2kr_i + \delta_i(k)]}{(kr_i)^2} \qquad (1)$$

where m is the mass of the electron, k is the momentum or wavenumber of the photoelectron, $2r_i$ is the distance from the origin and back from the scatterer, δ_i is the k-dependent phase shift and $t_i(k)$ is a k-dependent scattering amplitude. In eq.1 we have summed over all scatterers in the sample to yield the total EXAFS signal

THEORY

Bayes' theorem is

$$P(A|B,I)=\frac{P(A|I)\times P(B|A,I)}{P(B|I)} \tag{2}$$

where the proposition the A is true given B and any prior information, I, is related to the proposition A is true given only the prior information, the proposition B is true given A and the prior information divided by the probability B is true given the prior information. To relate this more directly to spectroscopy, consider a general k-domain function, f(k), define as

$$f(k)=\sum_{j=1}^{m} N_j G_j(k,\{R\}) \tag{3}$$

where G_j is some function of momentum which depends parametrically on a set of radial distances, the distances between the scattering centers, and N_j is an amplitude for each model function. As we will show below, these functions are chosen to model eq. 1. Bayes theorem can then be used to determine the probability that the model functions, which contain adjustable parameters, correctly describe the experimental data

$$P(N,\{R\}|D,I)=\frac{P(N,\{R\}|I)\times P(D|N,\{R\},I)}{P(D|I)} \tag{4}$$

One advantage of the Bayesian formulation is that linear parameters in the probability may be removed by marginalization, that is, the parameter-dependence may be integrated out of the probability. In the model functions used in EXAFS analysis the amplitude may be marginalized so that only the separation of each shell remain. Because of marginalization non-linear least-squares adjustment of the parameters to optimize the probability now becomes feasible.

It has previously been shown that marginalization of the amplitude leads to the following expression for the probability [2]

$$P(R|D) \equiv \left\{1 - \frac{m\overline{h}^2}{N\overline{d}^2}\right\}^{\frac{m-N}{m}} \qquad (5)$$

where m is the number of model functions, N is the number of data points, R is the shell separation in the EXAFS spectrum, $\overline{h}^2$ is the sufficient statistic, defined by

$$\overline{h}^2 = \frac{1}{m}\sum_{j=1}^{m} h_{ij} \qquad (6)$$

where h_{ij} is the projection of the orthogonalized model function onto the data and $\overline{d}^2$ is given by

$$\overline{d}^2 = \frac{1}{N}\sum_{i=1}^{N} d_i^2 \qquad (7)$$

where the orthogonalized model functions are obtained by forming the matrix, g_{ij}, of all possible products $G_i * G_j$ of the model functions given in eq. 3 and obtaining the eigenvectors of this matrix. The Bayesian method is a non-linear least-squares optimization of the probability, given by eq. 5, where the radial separation of each model function, i.e. peak in the spectrum, is adjusted. The marginalized parameters, in this case the number of atoms in each shell, may be obtained from the final finals of the estimated parameters.

Eq. 1 is difficult to apply to the analysis of EXAFS data due to two unknown quantities, the amplitude factor, t(2k) and the phase shift, δ(k). We follow Cramer, et al [2] and write our model function as

$$\chi(k) = \frac{N}{R}A(k)\sin[2kR + \delta(k)] \qquad (8)$$

where N, R, A(k) and δ(k) apply to each shell in the data. N is the number of scatterers in each shell, i.e. the coordination number, A(k) is the amplitude factor and δ(k) is the phase shift. A(k) contains collectively the contributions from the

Debye-Waller factor and multiple scattering effects in addition to t(2k). The function chosen to model A(k) is

$$A(k) = c_o \exp(-c_1 k^2) / k^{c2} \tag{9}$$

where c_0, c_1 and c_2 are adjustable parameters. The phase shift is assumed to follow the form

$$\delta(k) = a_o + a_1 k + a_2 k^2 \tag{10}$$

Therefore, there are six adjustable parameters, a_0, a_1, a_2, c_0, c_1 and c_2 and two parameters of physical interest, N and R, for each shell represented in the experimental data. The problem of fitting the data with this large number of adjustable parameters in further complicated by the fact that several of the parameters are highly correlated. Thus, the uniqueness of the resulting set of parameters is dubious.

We follow Cramer, et. al in suggesting that the six adjustable parameters are transferable between compounds consisting of the same metal centers as scatterers. Thus, the six adjustable parameters may only need to be determined for a single "standard" compound, preferably a simple one-shell system, and may then be used as constants in the remaining compounds in this family.

RESULTS AND DISCUSSION

Along these lines we have developed a simplex optimization program in this lab [3] to fit the above set of equations for the one-shell compound, MoS. The simplex method was chosen was chosen because, although slow, it is robust and does not require analytic derivatives of the model function. In this method we fixed the shell distance, R, at the X-ray [4] value of 2.367 A^o and fixed the coordination number, N, at the expected value of 6.0. We then minimized the sum-square of residuals for the optimal values of the six adjustable parameters. The results are given below.

We also allowed the coordination number to be an adjustable parameter. It has often been found from conventional FFT analysis that the coordination number is slightly less than the expected value. Usually this is ascribed to the strength of the

Table 1. Results of simplex fit of EXAFS equations to MoS data.

ao	a1	a2	co	c1	c2	N
-0.357	-1.563	0.035	0.293	0.018	0.812	6.0
-0.357	-1.563	0.035	0.317	0.018	0.812	5.5(4)

binding although it more likely represents coupling to other parameters or effects of prior massaging of the experimental data. The value of N is close to the experimental value, however, not it has resulted in a change in the c_0 parameter. A plot of the fit spectrum to the experimental data is not shown as obviously the fit is good as the number of adjustable parameters is quite large.

As a test of the Bayesian method we have analyzed the one-shell data of MoS by fixing the values of a_0, a_1, a_2, c_0, c_1, c_2 to the values in the first row of Table 1 and using a program developed in our laboratory [3] to maximize the sufficient statistic, and therefore the probability, using the Bayesian method described above treating the coordination number as a nuisance parameter, i.e. it is marginalized out of the probability. Thus, for the Bayesian analysis there only one parameter per shell, the radial separation, R. At the end of the Bayesian analysis we can "back-calculate" the coordination number from the projection of the orthogonalized model functions on to the data. The resulting values of the physical parameters are R=2.371 and N=5.339 which agree well with the literature values [2]. A plot of the resulting fit is given in Fig. 1.

A more stringent test of the Bayesian method is the two-shell compound, $MoO(S_2CNEt_2)_2$ hereafter referred to as MoO. In applying the Bayesian technique we now have two adjustable parameters, R1 and R2, the two shell distances. As described above, we use the values of the six adjustable parameters as fixed to those of MoS. The resulting values are R=1.923 and 2.486 A^o and the coordination numbers, obtained from "back-calculation" of the marginalized amplitudes, are found to be N=1.44 and 3.30. The curve-fit values in reference 2 are 1.660 and 2.427 A^o for the distances and 1.3 and 3.7 for N. The Fourier-fit conventional method suggests values of 1.73 and 2.42 A^o for the separations and 1.4 and 3.7 for N. Thus, the values obtained from the Bayesian analysis are within the experimental error of the literature values. The Bayesian fit of the MoO data is given in Fig. 2.

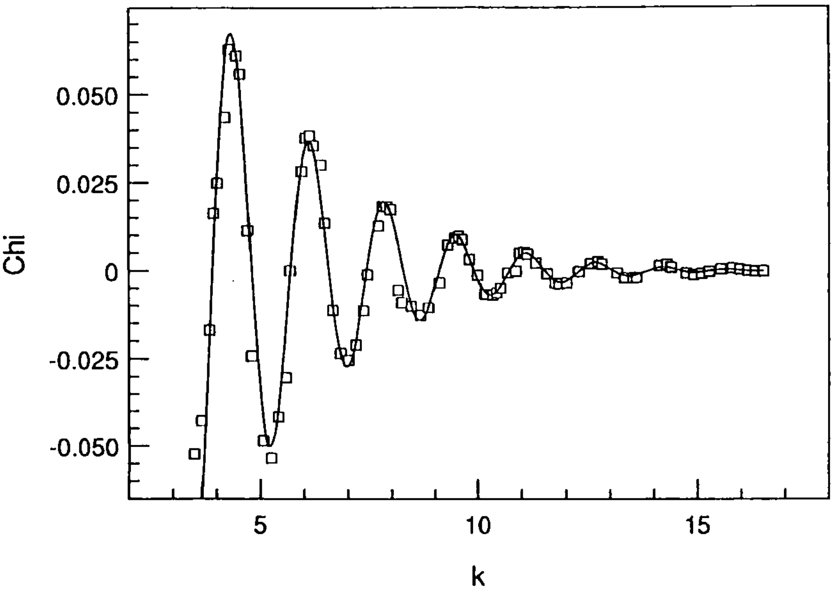

Fig 1. Bayesian fit to the single-shell EXAFS data of MoS.

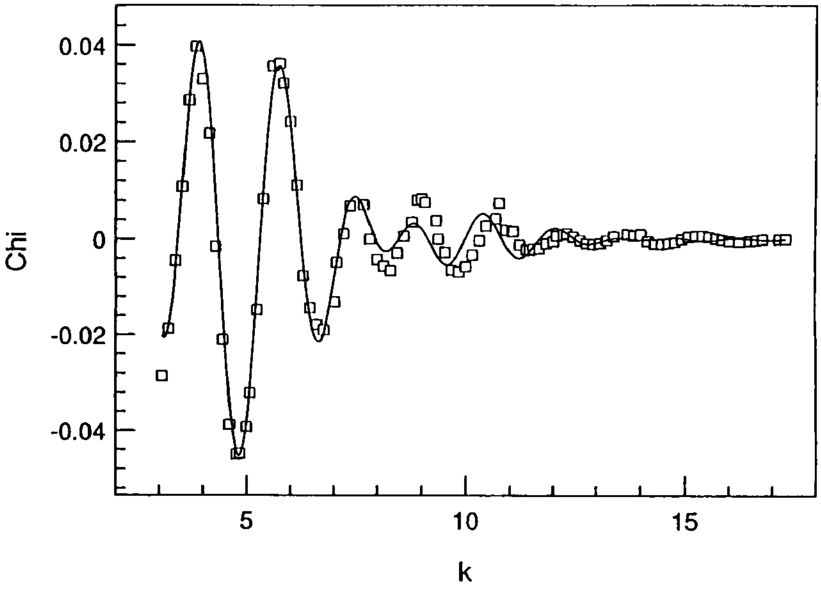

Fig 2. Bayesian fit of model functions to experimental data of MoO two-shell compound.

In conclusion, the Bayesian method has been shown to provide reliable estimates of shell-separations and coordination numbers from EXAFS data using the method of transferable amplitude and phase functions. This application, however, has only been the initial study of its application. EXAFS data undergoes a lot massaging and manipulation prior to get the data in a form that conventional FFT analysis can handle [5]. Each of these steps reduces the reliability of the final estimates of the physical parameters. A full Bayesian analysis of EXAFS data can incorporate in to the model functions all of this prior manipulation and can therefore treat the experimental data as it comes off the instrument. Such an analysis is in progress.

ACKNOWLEDGMENT

This work is supported by NSF and the Louisiana Educational Quality Support Fund through grant EHR 9108765.

REFERENCES

[1] Bretthorst, G. L; *Bayesian Spectrum Analysis and Parameter Estimation*, Vol. 48; Springer-Verlag, New York, 1988.

[2] Cramer, S. P.; Hodgson, K. O.; Stiefel, E. I.; Newton, W. E. *J. Am. Chem. Soc* **1978**, 100, 2748-2760.

[3] T. Yu, *Bayesian Probability Analysis of EXAFS* M.S. Thesis, University of New Orleans, 1994.

[4] Richard, L.; Estienne, J.; Karagiannidis, P.; Toledano, P.; Fischer, J.; Mitschler, A.; Weiss, R. *J. Coord. Chem.*, **1974**, 3, 277-294.

[5] Koningsberger, D. C. Prins, R.; *X-Ray Absorption, Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*, Vol. 92, Wiley and Sons, New York, 1988.

Received: June 28, 1995
Accepted: August 1, 1995