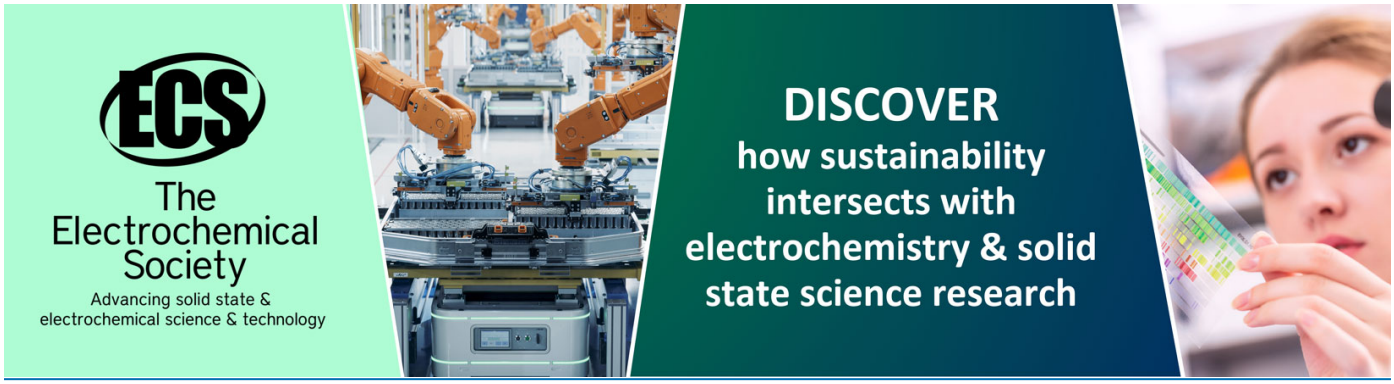
Predetermining acceptable noise limits of EXAFS spectra in the limit of stochastic noise



Journal of Physics: Conference Series **190** (2009) 012029 doi:10.1088/1742-6596/190/1/012029

Predetermining acceptable noise limits of EXAFS spectra in the limit of stochastic noise

Yung-Jin Hu[1] and Corwin H. Booth[2]

[1]Nuclear Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

[2]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

E-mail: YHu@lbl.gov

Abstract.

The effect of stochastic noise on Extended X-ray Absorption Fine Structure (EXAFS) data measurement, analysis, and fitting is discussed. Stochastic noise reduces the ability to uniquely fit a calculated model to measured EXAFS data. Such noise can be reduced by common methods that increase the signal-to-noise ratio; however, these methods are not always practical. Therefore, predetermined, quantitative knowledge of the level of acceptable stochastic noise when fitting for a particular model system is essential in maximizing the chances of a successful EXAFS experiment and minimizing wasted beamtime. This paper outlines a method to estimate, through simulation, the acceptable level of stochastic noise in EXAFS spectra that still allows a successful test of a proposed model compound.

1. Introduction

EXAFS measurements are used to probe a variety of experimental systems, but excel at elucidating local structure in samples which have slight disorder or no long-range crystalline order. Of special interest to the authors is the use of EXAFS in understanding the molecular-level binding structure and characteristics of actinides on the surface of environmental minerals and model mineral analogs [1]. In environmental systems the element of interest can be on the order of 10^{-7}% by weight of the total sample. Obviously such samples would be impossible to measure using EXAFS techniques. It is therefore essential to increase the concentration of the element of interest while still preserving a sample's ability to represent environmental conditions. Under such low concentration limits it is expected that the collected data is count-rate, or stochastically limited. This condition occurs as we approach the signal-to-noise (S/N) limit of the technique where the random noise of the measurement process dominates over possible systematic errors [2, 3]. When stochastic error is expected to dominate systematic error, it is possible to predict, with the use of simulations, the ability of model fits to tolerate a certain level of stochastic noise. Elsewhere in these proceedings, we discuss how to tell when systematic errors dominate in measured EXAFS spectrum [4].

Here, we outline a technique for determining the number of EXAFS scans necessary to test the relevance of a given structural model. Appropriate stochastic noise levels are determined for each point in r-space by collecting data on a real system. These noise levels are then applied to EXAFS simulations using a test model. In this way, all significant systematic error sources

Table 1: Proposed EXAFS model for plutonium sorbed at the mineral-water interface of pyrolusite in solution.

path	N	R(Å)	σ^2(Å^2)
$Pu-O_1$	9	2.3800	0.0049
$Pu-O_2$	2	3.6912	0.0081
$Pu-Mn$	4	3.8086	0.0081
$Pu-O_3$	2	3.9943	0.0081

are eliminated in the simulated data. The structural model is then fit to the simulated data, decreasing the noise and increasing the k-range of the fit until the veracity of the model passes an F-test [5, 6].

2. Simulation methods and results

The technique for predetermining acceptable noise limits is presented in the form of a case study where we wish to test for the presence of a near-neighbor Pu-Mn pair due to Pu sorbed onto the surface of natural minerals. Appropriate noise levels for the simulations are based on those obtained from real data acquired from samples consisting of plutonium, initially introduced as a PuO_2^{2+} solution, sorbed onto the surface of the manganese oxide mineral pyrolusite (β-MnO_2). The total atomic percent of plutonium in the sample is 0.048%. Plutonium L_{III}-edge EXAFS spectra were collected at the Stanford Synchrotron Radiation Lightsource (SSRL) on beam line 11-2 using a Si(220) $\phi - 0^\circ$ double-crystal monochromator. Data were collected in fluorescence mode using a 32-element Ge-detector. Total fluorescence counts were sufficiently low that no dead-time correction was required. Acquisition time for each spectra was approximately 15 minutes. EXAFS oscillations, $\chi(k)$, were isolated using standard procedures [7]. A total of 25 repeated experimental scans were used to estimate the stochastic noise associated with the measurement process at the beamline. The standard deviation of $\chi(k)$ was calculated at each measured k-value and fit with a constant value in $k\chi(k)$ weighting. The application of Student's t factor accounts for a limited number of experimental scans in estimating these standard deviations [8].

The simulations begin with a theoretical calculation [9] of a EXAFS spectrum from a structural model (table 1). Random (Gaussian) noise with a standard deviation characteristic of the noise previously extracted from the experimental data is applied point-by-point in k-space using pseudo-random numbers generated by the Mersenne Twister algorithm [10]. The stochastic noise in each simulated spectrum then corresponds to that expected for a single experimental scan within the original experimental conditions.

Two different chemically-feasible model systems are then fit to the simulated EXAFS data using RSFIT (part of the RSXAP package [11]): one model with both manganese and oxygen scattering paths, and one with only oxygen scattering paths. Fits were conducted varying the maximum k of the data in the fit ($k_{\max} = 10 \rightarrow 18$ Å^{-1}) and the number of simulated EXAFS scans (20-2500 files). No fits were extended below a $k_{\max}$ of 10 Å^{-1} because the number of degrees of freedom would be less than the number of parameters being fit [12]. For each model system an R-factor is calculated by RSFIT. Finally, a confidence level, α, that the fit including the manganese scattering path is statistically better than the fit without it, is obtained using a modified F-Test [5, 6, 13]. The α values are calculated 1000 times for each combination of simulated files and $k_{\max}$. The average is then plotted as a function of the two variables, forming a confidence-level surface.

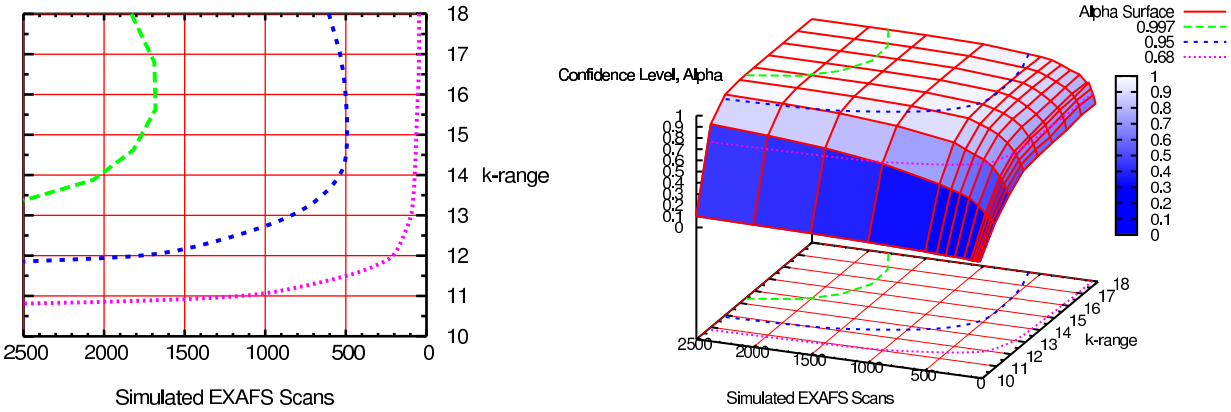


Figure 1: Contour plot (left) and surface plot (right) of the α surface as a function of simulated EXAFS scans and fit k-range in Å^{-1}. The plot represents the accumulation of F-Tests on $\sim$ 130000 individual simulations.

Figure 1 (right) shows a plot of the average confidence level, α, as a function of both the measured k-range and number of simulated EXAFS scans. Lines of constant confidence are drawn at the 68%, 95%, and 99.7% level by fitting a B-spline through the surface.

3. Discussion

Manual fitting of various parts of the α surface was conducted before automatic calculations were started to ensure that fits were likely to converge. Even in fits that had not completely converged the α metric had typically converged to better than 0.1%. Large numbers of calculations were conducted to ensure that our mean estimate for α was precise. In any case, since individual monitoring of each of the $\sim$ 130000 fits would have been exceedingly difficult the large number of simulations serves to reduce the weight of the occasionally un-converged fit.

According to figure 1 under our initial experimental conditions, assuming only stochastic noise, we would have had to collect over 500 experimental scans (about 125 hours) out to a $k_{\max}$ of greater than 14.5 Å^{-1} before we would expect to be 95% confident that our fitting model containing the manganese scattering path represented our data statistically better than the fitting model with only oxygen scattering paths. Various other effects which could have limited our $k_{\max}$ would have resulted in the need to take even more experimental scans. However, within the limit of stochastic error, the effect of increasing experimental scans is simply to increase the S/N ratio. Other options to increase the S/N ratio would be to increase the concentration of the element of interest. For example, a two-order of magnitude increase in the element of interest will increase S/N by a factor of 10, assuming counting statistics. In other words, a 95% confidence level could be reached after 50 experimental scans, or about 13 hours.

The contour plot in figure 1 also shows an extrema in the contour lines at a point between $k_{\max} = 14.5$ Å^{-1} and 16.5 Å^{-1} depending on the confidence level. As the fit is increased to higher $k_{\max}$ values, α degrades slightly, showing that more simulated EXAFS scans are needed to keep the same confidence level at higher $k_{\max}$. This degradation is likely due to the lower S/N ratio at large $k_{\max}$ and is the reason why most EXAFS practitioners collect for longer times in at high k values. However, as the contour plot shows for this particular model, a better use of beamtime would be to collect data using a shorter k-range than attempt to measure data out to a large $k_{\max}$.

4. Conclusion

This paper outlines a method of testing model systems in EXAFS fitting before measurements are conducted to determine the quality of measured data required for fitting of a particular model system with statistical confidence.

It is important to reiterate that the calculated α surface in figure 1 is only applicable to the particular model presented in this paper. Furthermore, this procedure only takes into account stochastic noise; consequentially any confidence levels calculated should be viewed as upper limits to the confidence levels in systems which also contain a significant amount of systematic noise.

Acknowledgments

The authors thank Prof. Heino Nitsche for many helpful discussions and his support throughout the thesis work of Y. Hu. The research was performed under appointment of the Office of Civilian Radioactive Waste Management Graduate Fellowship Program administered by Oak Ridge Institute for Science and Education under a contract between the U.S. Department of Energy (DOE) and Oak Ridge Associated Universities, and is also supported by the DOE under Contract No. DE-AC02-05CH11231. X-ray absorption data were collected at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the DOE/OBES.

References

[1] Nitsche H 1995 *J. Alloys Compd.* **223** 274–279
[2] Curis E, Osn J, Falkenberg G, Bnazeth S and Trk S 2005 *Spectrochim. Acta,* Part B **60** 841 – 849
[3] Li G, Bridges F and Wang X 1994 *Nucl. Instrum. Meth. A* **340** 420 – 426
[4] Booth C H and Hu, Y-J 2009 "Conventional parameter-error estimates in EXAFS analysis: Lessons from Nyquest and Fourier" Elsewhere in this volume
[5] Hamilton W C 1965 *Acta Cryst.* **18** 502
[6] Downward L, Booth C H, Lukens W W and Bridges F 2006 *AIP Conference Proceedings* vol 882
[7] Li G G, Bridges F and Booth C H 1995 *Phys. Rev. B* **52** 6332
[8] Student 1908 *Biometrika* **6** 1–25 ISSN 00063444 URL `http://www.jstor.org/stable/2331554`
[9] Zabinsky S I, Rehr J J, Ankudinov A, Albers R C and Eller M J 1995 *Phys. Rev. B* **52** 2995–3009
[10] Matsumoto M and Nishimura T 1998 *ACM Trans. Model. Comput. Simul.* **8** 3–30 ISSN 1049-3301
[11] Booth C H 2006 http://lise.lbl.gov/RSXAP/
[12] Stern E A 1993 *Phys. Rev. B* **48** 9825–9827
[13] Bacchi A, Lamzin V S and Wilson K S 1996 *Acta Cryst. D* **52** 641