J. Synchrotron Rad. (1999). **6**, 264–265

Estimation of uncertainties in XAFS data

Matthew Newville,[a*] Boyan I Boyanov[b] and Dale E Sayers[b]

[a] *Consortium for Advanced Radiation Sources, University of Chicago, Chicago, IL USA, and* [b] *Department of Physics, North Carolina State University, Raleigh, NC USA. E-mail: newville@cars.uchicago.edu*

The estimate of measurement uncertainties in XAFS spectra is one of the most important and difficult aspects of error analysis for XAFS data. Since several different analysis methods are used to determine structural parameters and their uncertainties from XAFS data, and no universally agreed-upon standard exists, the ability to compare estimates of uncertainties from different methods is vital. We discuss numerical aspects of the propagation of measurement uncertainties from raw absorption coefficient to analyzed signal in $k-$ and $R-$space. We focus our attention on the conversion of measurement uncertainties between k- and R-space representations of XAFS data, and on the testing of this conversion procedure for data dominated by random fluctuations.

Keywords: XAFS Analysis; Error Analysis; Fourier Transform.

1. Using the high-R portion of $\chi(R)$ to measure the random fluctuations in XAFS data

Though XAFS data dampens with R due to disorder, mean-free-path, and the $1/R^2$ term in the XAFS equation, the random fluctuations in the data persist. Therefore, by looking at the high-R components of an XAFS spectrum, we can measure the random fluctuations in the data, and use the size of these fluctuations as an estimate of the uncertainty in χ, even for a single spectrum. This estimate of the measurement uncertainty can be used in the statistical analysis of the signal, including the sort of χ^2 tests recommended by the International XAFS Society Standards and Criteria committee (Lytle *et al.*, 1989) and for the estimation of uncertainties in parameters determined in XAFS analysis.

Figure 1 shows $\chi(R)$, the R-space transform of the XAFS, for Cu metal at 10K out to 25Å. Data for this sample was collected past 20 Å^{-1}, and the Fourier transform window used extended to 19 Å^{-1}. There is clearly no significant XAFS signal past 15 Å even for this highly ordered system, and we can be confident that $\chi(R)$ beyond this point will be dominated by the random fluctuations for nearly all systems.

Since the R components beyond 15Å contain no meaningful structural information, we use the root-mean-square (rms) average of $\chi(R)$ between 15 and 25Å as ε_R, the measured random fluctuations in $\chi(R)$. We note that this R-range is a somewhat arbitrary choice for evaluating the rms fluctuations in $\chi(R)$. It has been chosen so as to avoid any structural XAFS signal at lower R, and yet to be large enough to get a reliable rms average even for relatively clean spectra.

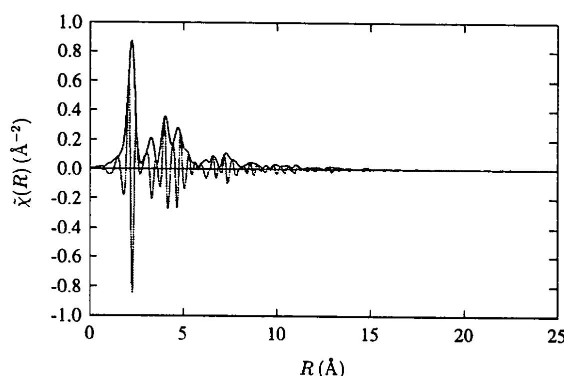

Figure 1
$\chi(R)$ for Cu metal at 10K. Even for high-quality XAFS on well-ordered systems, the signal is dominated by noise beyond 15Å.

Though not required for XAFS analysis or for the principles of the arguments discussed, we consider here the common case of using a discrete Fourier Transform (Brigham, 1974), with conjugate pairs of k and $2R$ and an evenly-spaced k-grid of $\delta k = 0.05\ \text{Å}^{-1}$, and require that $\chi(R)$ is zero below $R = 0$, which defines the relationship between the real and imaginary parts of $\chi(R)$ (Sayers & Bunker, 1988). With this form of the XAFS Fourier transform, R components up to $\pi/2\delta k \sim 31.4$ Å are sampled. Though we have not seen any evidence for aliasing R beyond 25Å, we take care to avoid the highest-R values, and recommend using the range from 15 to 25Å for measuring ε_R.

Because XAFS is routinely viewed and analyzed in both k- and R-space, it is desirable to be able to convert ε_R into ε_k, the rms fluctuations of $\chi(k)$. To do this, we note that the random fluctuations in k- and R-space are "white noise", and so extend with a constant rms amplitude over all values of k and R, respectively. With this in mind, ε_k and ε_R are simply related by Parseval's theorem (Brigham, 1974). For the symmetric XAFS Fourier transform of k-weighted spectrum using a finite k-window between $k_{\max}$ and $k_{\min}$, Parseval's theorem gives

$$\int_{k_{\min}}^{k_{\max}} |\varepsilon_k k^w|^2 dk = 2\int_0^{\pi/2\delta k} |\varepsilon_R|^2 dR. \quad (1)$$

Taking ε_R and ε_k to be constants, this gives

$$\varepsilon_k = \varepsilon_R \sqrt{\frac{\pi\,(2w+1)}{\delta k\,(k_{\max}^{2w+1} - k_{\min}^{2w+1})}}. \quad (2)$$

Of course, these ε_k and ε_R have no k- or R-dependence, which might be expected for some types systematic errors, such as large glitches and errors resulting from poor background removal. In that sense, the values determined from this simple method give a *lower bound* on the estimate of the measurement uncertainty in $\chi(k)$ and $\chi(R)$. On the other hand, the method described here is easily automated and can be applied to a single spectrum with no prior information about the data quality. In many cases these simple measurements of the rms fluctuations in $\chi(k)$ and $\chi(R)$ will give the

© 1999 International Union of Crystallography
Printed in Great Britain – all rights reserved

best available unbiased estimate of the measurement uncertainties in an XAFS spectrum, and can be used in statistical analysis of the goodness-of-fit and in the estimation of the uncertainties in fitted parameters during XAFS analysis.

2. Test with data dominated by random fluctuations

To investigate whether Eq. 2 gives a reasonable estimate of the fluctuations in the data, we use XAFS data that is dominated by random fluctuations so that independent estimates of the noise level can be made. The data used for such tests was Ti K-edge spectra of Ti implanted into a 200Å thick layer of SiC. 22 fluorescence spectra collected using a 13 element solid-state Ge detector over a 15 hour period at NSLS beamline X11-A were measured, each having a fairly high noise level. Statistical tests were done to check the normality of the fluctuations, which showed the fluctuations in this data to be normally distributed.

The uncertainties in the $\chi(k)$ data were estimated using 4 different techniques. First, the method described above was used, in which ε_k was determined from the high-R components of $\chi(R)$ using Eq. 2. Second, Poisson counting statistics were done using the counts in the solid-state detector elements. Third, the rms fluctuations in the pre-edge absorption data were used. Fourth and finally, the rms fluctuations in the post-edge absorption data were used to estimate ε_k. For the last two methods described, the data was fitted with a smooth polynomial and the rms deviations from that function were used. All four methods were applied to averages of different numbers of scans (from 1 to 22). This not only gave more comparisons of the different estimates of the uncertainties, but also allowed the expected $1/\sqrt{N}$ fall-off in noise to be checked, and for any base-line noise level to be identified. The four methods for estimating ε_k are presented in Fig 2, which compares the results of these methods as a function of number of individual scans used in the average.

The results show that the expected $1/\sqrt{N}$ trend is followed fairly well for all measures of noise. There does appear to be a tendency towards saturating at a finite lower limit of noise (probably at $\varepsilon_k \approx 0.003$ for this data), which may indicate systematic errors in the measurement dominating the measurement error when the fluctuations from counting statistics become sufficiently small. The evidence for that is not conclusive, but is suggestive that even fairly noisy XAFS data may have an appreciable noise level that does not come from random fluctuations. It is also seen that the four different methods agree rather well in their estimate of the measurement uncertainty at all noise levels. We should point out that the estimate from the Fourier Transform method tends to give a value that is slightly smaller than the estimate from Poisson statistics, which might be due to smoothing of the XAFS data during interpolation from E to evenly spaced values of k. Despite this effect (which still keeps the different measurements within a factor of 2 of each other), these results demonstrate that the estimate of ε_k from the Fourier transform method described here gives a reasonable estimate of the measurement uncertainty for data dominated by random fluctuations.

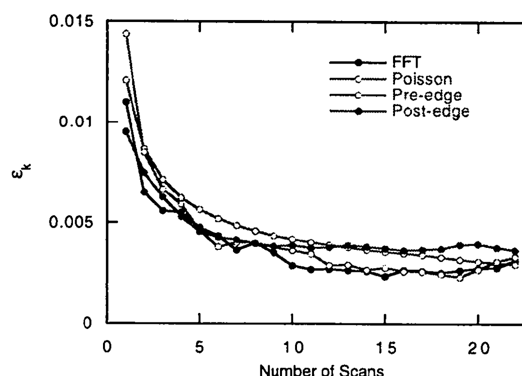

Figure 2

Comparison of 4 different estimates of the measurement uncertainty for Ti K-edge XAFS data dominated by counting statistics as a function of the number of individual scans averaged. The four estimates are from: the high-R method described here (FFT); Poisson counting statistics using the counts in the solid-state detector (Poisson); the rms deviations of the pre-edge absorption data (Pre-edge); and the rms deviation of the post-edge absorption data (Post-edge).

3. Conclusion

Using the high-R portion of a single XAFS spectrum gives a simple and reliable estimate of the measurement uncertainty that is needed for the complete statistical analysis of that spectrum, such as setting the scale for χ^2 tests and estimates of the uncertainties in parameters determined in XAFS analysis. We have shown this measure to be consistent with other estimates of the measurement uncertainty for data dominated by random errors. The method assumes that the measurement uncertainties do not vary with k or R and so will not include some kinds of systematic errors, but is simple enough to use in a routine way.

We thank Prof. Edward Stern and Dr. Peteris Livins for helpful discussions.

References

Brigham, E. O. (1974). *The Fast Fourier Transform.* Englewood Cliffs, NJ: Prentice-Hall.

Lytle, F. W., Sayers, D. E. & Stern, E. A. (1989). *Physica B*, **158**, 701–722.

Sayers, D. E. & Bunker, B. (1988). In *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES*, edited by D. C. Koningsberger & R. Prins, vol. 92 of *Chemical Analysis*, p. 211. New York: John Wiley & Sons.

(Received 10 August 1998; accepted 21 December 1998)