Lawrence Berkeley National Laboratory

Title
Confirmation of standard error analysis techniques applied to EXAFS using simulations

Permalink
https://escholarship.org/uc/item/13z2d2x7

Author
Booth, Corwin H

Publication Date
2009-11-05

Peer reviewed

Confirmation of standard error analysis techniques applied to EXAFS using simulations

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

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

E-mail: `chbooth@lbl.gov`

Abstract. Proper application of error analysis techniques remains uncommon in most EXAFS analyses. Consequently, many researchers in the community remain distrustful of parameter-error estimates. Here, we demonstrate the accuracy of conventional methods through r-space fits to simulated data. Error estimates are determined as a function of r by averaging many scan simulations in r-space. The statistical-χ^2 value can then be calculated. Since $\langle \chi^2 \rangle$ corresponds to the degrees of freedom in a fit, we check Stern's rule for the number of independent data points in an EXAFS spectra. Finally, we apply these simple methods to real data from a Cu foil, highlighting the overwhelming role of systematic errors in theoretical backscattering functions and pointing to the ultimate power of the EXAFS technique if such errors could be removed.

1. Introduction

Systematic uncertainties, such as those in calculated backscattering amplitudes, crystal glitches, etc. (e.g. Ref. [1]), not only limit the ultimate accuracy of the EXAFS technique, but also affect the covariance matrix representation of real parameter errors in typical fitting routines [2]. Despite major advances in EXAFS analysis and in understanding all potential uncertainties [1, 3–8], these methods are not routinely applied by all EXAFS users. Consequently, reported parameter errors are not reliable in many EXAFS studies in the literature. This situation has made many EXAFS practitioners leery of conventional error analysis applied to EXAFS data. However, conventional error analysis, if properly applied, can teach us more about our data, and even about the power and limitations of the EXAFS technique. Here, we describe the proper application of conventional error analysis to r-space fitting to EXAFS data. Using simulations, we demonstrate the veracity of this analysis by, for instance, showing that the number of independent data points from Stern's rule [9] is balanced by the degrees of freedom obtained from a χ^2 statistical analysis. By applying such analysis to real data, we determine the quantitative effect of systematic errors. In short, this study is intended to remind the EXAFS community about the role of fundamental noise distributions in interpreting our final results.

2. Conventional χ^2-based error methods in EXAFS

Most applications of error analysis revolve around determining the statistical-χ^2 parameter [10],

$$\chi^2 = \sum_{i=1}^{N_{\text{ind}}} \frac{(y_i - y_{\text{f},i})^2}{s_i^2}, \tag{1}$$

where y_i are the measured data, $y_{\text{f},i}$ are the values from a fit using a model with p_1, p_2,...,$p_{N_{\text{fit}}}$ fitting parameters, s_i are the standard deviations of the data, and N_{ind} is the number of independent data points. It is this last quantity that generates confusion regarding χ^2 methods in EXAFS, and will be further discussed below. In any case, if one applies equation 1 for known $y_{\text{f},i}$, each term in the sum should average to unity assuming a normal error distribution, and the sum should total to N_{ind}, on average. For every fit parameter used to determine $y_{\text{f},i}$, the sum will be reduced by one; that is, the degrees of freedom in a fit are $\nu = N_{\text{ind}} - N_{\text{fit}}$.

As χ^2 is calculated from data with a statistical distribution, χ^2 itself has a distribution with $\langle \chi^2 \rangle = \nu$. If one has $y_{\text{f},i}$, s_i, and N_{ind}, one can determine χ^2, and hence the errors on the final fitting parameters p_i. These errors are typically calculated as in the Levenberg-Marquardt method [2] from the covariance matrix, based on the curvature of χ^2 at its minimum value with respect to p_i. This procedure assumes not only that χ^2 is quadratic from $p_i = \langle p_i \rangle$ to $\langle p_i \rangle \pm \sigma_i$, where σ_i is the standard deviation error on $\langle p_i \rangle$, but also does not account for correlations with other parameters, except by examining the off-diagonal elements of the full covariance matrix. We have therefore implemented an alternate, more general method [2, 11] where one performs several fits, holding a given p_i at some Δ_i away from its best-fit value, $p_i = \langle p_i \rangle \pm \Delta_i$. Since all the other parameters are allowed to vary in such a fit, ν is increased by one. When one finds the maximal deviation from $\langle p_i \rangle$ where $\chi^2 = \langle \chi^2 \rangle + 1$, $p_i = \langle p_i \rangle \pm \sigma_i$, where σ_i is one standard deviation. Note that this method can easily be extended to yield asymmetric error bars.

These procedures rely on an accurate determination of N_{ind} to properly evaluate equation 1. In a purely k-space situation, N_{ind} is merely the number of data points collected in a spectrum, assuming good energy resolution [12]. However, since data are typically fit over some range Δ_r in an r-space fit, $N_{\text{ind}} = 2\left(\Delta_r/\delta_r + 1\right)$, where δ_r is the the step-size in r-space, and the factor of 2 comes about because each data point in r-space has a real and an imaginary component. Since a Fast Fourier Transform is typically used, data are collected with a constant step size in k. Therefore, the number of data points in k-space is $N_k = k_{\text{max}}/\delta_k + 1$, where k_{max} is the maximum k value measured. Note that in any Fourier transform (FT), the total number of independent data points cannot change; thus, $N_k = N_r = 2r_{\text{Ny}}/\delta_r$, where r_{Ny} is the maximum r value given by the Nyquist theorem, $r_{\text{Ny}} = \pi/(2\delta_k)$. Combining these equations, one can write

$$N_{\text{ind}} = \frac{2\Delta_k \Delta_r}{\pi} + 2 + \frac{2\Delta_r \delta_k}{\pi}. \tag{2}$$

The first two terms, colloquially known as "Stern's rule", give N_{ind} as derived in a famous paper [9] that was the first to point out that the constant offset, which we refer to as c, should be two for fits that don't extend to either $r = 0$ or r_{Ny}. The third term is not significant until $\Delta_r \sim 30$ Å, and is therefore ignored in the following discussion in favor of Stern's result.

Before moving on, some particular points must be made. The ability to resolve two peaks in a FT is actually given by half the step-size in r-space, owing to the complex (i.e. real and imaginary part) nature of the FT. Since this degree of resolution isn't obvious in a plot of the FT magnitude, virtually every researcher resorts to a common trick that make FTs easier to *visually* interpret: data are "padded" with $\chi(k) = 0$ to a k well beyond k_{max} in the real data. The sum in equation 1 is therefore carried out over some number of data points in, say, r-space, $N_r > N_{\text{ind}}$. However, since these points are no longer independent, the final sum must be scaled:

$$\chi^2 = \frac{N_{\text{ind}}}{N_r} \sum_{i=1}^{N_r} \frac{(y_i - y_{\text{f},i})^2}{s_i^2}. \tag{3}$$

Finally, it is important to note that when one minimizes χ^2 as the goodness-of-fit parameter, the weighting of the data by s_i eliminates any effect from k-weighting in the final fit result.

3. Example application: simulated noise and Stern's rule

Random noise can be determined in real EXAFS data by collecting many scans to determine $\langle y_i \rangle$ and s_i (here, the standard deviation of the mean) for each data point. Even some of the quasi-statistical noise in the background absorption functions can be accounted by performing the averages after background removal on individual scans. Determining s_i in r-space analytically from k-space measurements can be performed [13]. We circumnavigate this issue in the RSXAP codes [14] by transforming individual k-space scans and calculating $\langle y_i \rangle$ and s_i directly in r-space.

Here, we test the methods in section 2 by adding Gaussian noise to simulated $\chi(k)$ data generated with the FEFF code [15]. In this way, systematic errors, including in the background removal, are negligible. Noise is added to 100 scans using the Mersenne Twister algorithm [16] which are used to calculate the average r-space spectra and s_i. It is necessary to use a fairly large number of scans so that the determination of the s_i are reliable, and to ensure that no s_i estimates are erroneously too small, causing a singularity in equation 3. Such simulations have other uses. Elsewhere in these proceedings, we show how one can use simulations in conjunction with the F-test [17, 18] to judge the probable success of an experiment given the expected data quality [19]. In addition, we have determined that the method for parameter error determination described in section 2 is accurate by comparing errors determined with that method to the $\langle p_i \rangle$ distribution determined by fitting each of the 100 simulated scans individually.

Given the demonstrated efficacy of these methods, we now perform a set of simulations to generate the χ^2 distribution, and therefore determine ν and verify Stern's rule (equation 2), and especially the need for the factor of $c = 2$. By performing many fits to simulated spectra, we can vary ν (and thus $\langle \chi^2 \rangle$) by varying Δ_r. A plot of $\langle \chi^2 \rangle$ vs. Δ_r should then be a straight line, with a slope of $2\Delta_k/\pi$ and an intercept of $2 - N_{\text{fit}}$, according to Stern's rule.

The simulations for this purpose used only the first three single-scattering paths from the Cu structure ($R_{Cu-Cu} = 2.55$, 3.61, and 4.42 Å). Typical constraints were employed, such as holding the threshold energy E_0 shifts equal for each path. Ultimately, $N_{\text{fit}} = 8$, with each fit
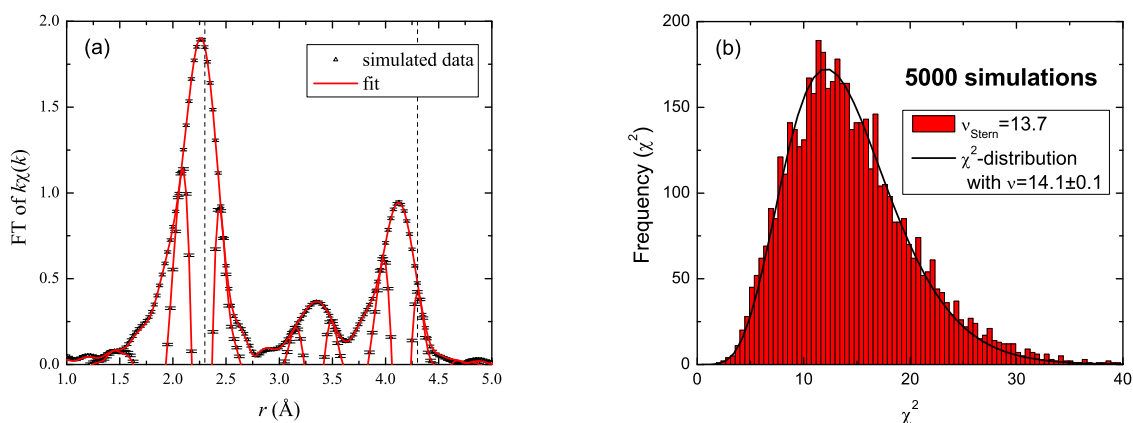
Figure 1. (a) Example of a single simulation with error bars, together with a fit between the limits shown by the vertical dashed lines. Outer envelope is the FT magnitude, and the modulating inner line is the real part of the FT, calculated between 2.5 and 18.0 Å^{-1}, with a 0.3 Å^{-1}-wide Gaussian window. (b) Histogram of χ^2 values with a fit to the χ^2 distribution.
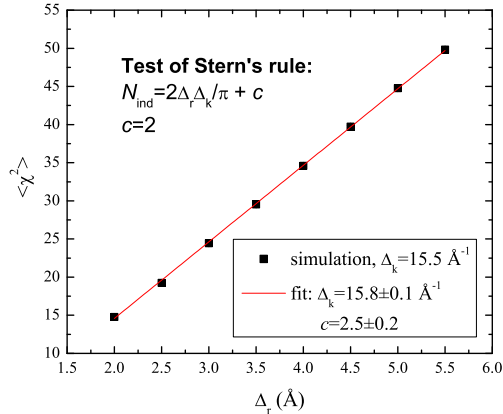
Figure 2. $\langle \chi^2 \rangle$ vs. r-range for fits to simulated Cu data. Linear fit gives ν as a function of Δ_r, with $\nu = N_{\text{ind}} - N_{\text{fit}}$ and $N_{\text{fit}} = 8$. Each data point represents 40,000 simulations.

ranging from $2.5 - 18$ Å^{-1}, using a Gaussian window of width 0.3 Å^{-1}. The r-range was varied from $\Delta_r = 4.3 - 2.3 = 2.0$ Å to $\Delta_r = 6.5 - 1.0 = 5.5$ Å in 0.5 Å increments. A total of 320,000 such simulations were used for this experiment.

Figure 1 shows an example of one simulated data set with error bars (based on 100 individual simulations, each), together with a histogram of χ^2 values from 5000 fits to such spectra. Figure 2 shows $\langle \chi^2 \rangle$ as a function of Δ_r for all these fits, together with a linear fit. This fit determines $\Delta_k = 15.8 \pm 0.1$ Å^{-1} and $c = 2.5 \pm 0.2$, in good agreement with the known value of $\Delta_k = 15.5$ Å^{-1} and the Stern value of $c = 2$. Although we do not currently have an explanation for the small discrepancy between these results and those expected from Stern's rule, we find, at worst, that Stern's rule gives a lower limit for N_{ind}, and therefore provides a good working estimate.

4. Real data and systematic error effects

The above results rely on using data simulated with the actual backscattering functions used in the fits, thereby ascribing systematic error sources to the fitting codes (i.e. round-off error). Real data have many other non-random sources of error, including problems with background removal, crystal glitches, harmonic content, etc. Careful treatments of these effects are possible [20, 21]. Nevertheless, even relatively clean data will benefit from a conventional error analysis: the calculation of a real χ^2 from a fit will indicate whether the data are limited by systematic or random errors when compared to ν. Moreover, it has been shown [1] that the main effect of systematic errors is to produce an offset, χ_0^2, in an experimentally determined χ^2; that is, while varying a parameter p_i in a fit to real data, $\chi_{exp}^2(p_i) \approx \chi_{rand}^2(p_i) + \chi_0^2$. Therefore, the procedure outlined above will still generate the random component of the error on a given parameter p_i.

An example of a case where random errors play a dominant role is given in reference [19]. Here, we give an example of data that is limited by systematic errors. For this purpose, we have chosen data on a Cu foil collected at 50 K, since the FEFF code should produce its best results for this model system. A total of 8 scans were collected to calculate $\langle y_i \rangle$ and s_i in r-space (figure 3). The s_i are corrected for the limited number of scans using Student's t factor [23]. The data were fit in a similar manner to those in section 2, except that two multiple scattering paths were explicitly included, and a "kitchen sink" path was also used including all other paths with four scattering legs or less out to 6 Å. The fit results are given in table 1. The fitted χ^2 is very large, and attempts to ascertain parameter errors as in section 2 were not practical; therefore, σ_p were
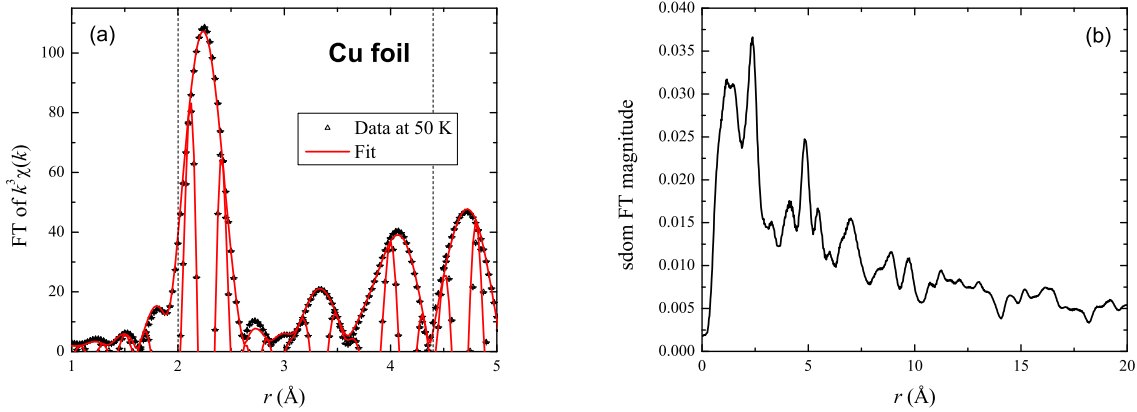
Figure 3. (a) Fit results for real Cu foil data. Data were transformed between 2.5 and 15.8 Å^{-1} with a 0.3 Å^{-1}-wide Gaussian window. Data are from the average of 8 scans. Error bars (difficult to discern in this plot, see (b)) were determined from the standard deviation of the mean (sdom) of these scans. (b) Error bar magnitude from data in (a).

estimated by performing separate fits to each of the 8 scans.

There are a number of important features of these fits with regard to the current subject. First, the fits are quantitatively excellent on the scale of EXAFS fits in general. Moreover, the absolute differences between the bond lengths implied by diffraction and those measured by EXAFS are quite small, consistent with other model compounds [1]. However, the absolute differences are very much larger than expected from the estimate of the random errors on the bond lengths. These errors are, in fact, the *maximum* estimated errors, as both FEFF and RSXAP become unreliable at higher precision due to round-off errors. In addition, the fit in figure 3 does not pass through the estimated error bars, which are quite small and difficult to see in the figure. The biggest problem in this fit is the χ^2/ν estimate, which is nominally unity in a statistically-limited fit. This fit, therefore, is strongly limited by systematic errors. In fact, these data are far better than need be, given the systematic error level: the random errors could

Table 1. Fit results from Cu foil data at 50 K, between 2.0 and 4.4 Å (figure 3). Stern's rule gives $\nu = 8.5$. R_{diff} are from diffraction measurements at 298 K, corrected to 50 K for thermal expansion [22]. Multiple scattering is included, but only single scattering paths are reported.

path	N	R_{diff}	$R(\text{Å})$	$\sigma^2(\text{Å}^2)$
Cu-Cu	12	2.5456	2.5376(1)	0.00272(1)
Cu-Cu	6	3.6001	3.5906(1)	0.00406(1)
Cu-Cu	24	4.4092	4.4051(1)	0.00482(1)
ΔE_0			3.78(1)	
S_0^2			0.8210(3)	
$R(\%)$			4.62	
χ^2/ν			$\sim$13,000	

be $\sqrt{\chi^2/\nu} \sim 100$ times larger before they are of similar magnitude as the systematic errors.

Normally, one might ascribe the enhanced $\sqrt{\chi^2/\nu}$ to deficiencies in the fitting model; however, this fitting model is well founded. It has been previously shown by comparisons between fits to model-system data using experimentally and theoretically determined backscattering functions that the main source of systematic error is in the backscattering calculation [1, 5, 24]. Therefore, an important result given the extremely small errors in reproducibility, is that the true potential of the EXAFS technique is still far from being reached, and vast improvements can still be made to backscattering codes that could revolutionize the field.

5. Conclusions

After reviewing proper evaluations of the statistical-χ^2 and fit degrees of freedom ν, simulations of simplified Cu EXAFS were used to generate statistically-limited fits. By performing many such fits, the accuracy of the χ^2 distribution was verified, as was Stern's rule for determining the number of independent data points N_{ind} in a spectrum, including the extra factor of $+2$. Using the same techniques on real data from a Cu foil produced very high quality fits that were severely limited by systematic differences between the data and the theoretically calculated backscattering functions. It is suggested that routine application of these techniques will allow experimenters to properly quantify errors when limited by random noise, and to also allow for determining the magnitude of systematic errors in real data.

References

 [1] Li G G, Bridges F and Booth C H 1995 *Phys. Rev.* B **52** 6332
 [2] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in Fortran 77: The Art of Scientific Computing* (New York: Cambridge University Press) chap 15 2nd ed
 [3] Gurman S J and McGreevy R L 1990 *J. Phys.: Condens. Matter* **2** 9463
 [4] Filipponi A 1995 *J. Phys.: Condens. Matter* **7** 9343
 [5] Kvitky Z, Bridges F and van Dorssen G 2001 *Phys. Rev.* B **64** 214108
 [6] Krappe H J and Rossner H H 2002 *Phys. Rev.* B **66** 184303
 [7] Curis E and Bénazeth S 2005 *J. Synchrotron Rad.* **12** 361
 [8] Rossberg A and Scheinost A C 2005 *Anal. Bioanal. Chem.* **383** 56
 [9] Stern E A 1993 *Phys. Rev.* B **48** 9825
[10] Bevington P R and Robinson D K 1992 *Data Reduction and Error Analysis for the Physical Sciences* (Boston: WBC/McGraw-Hill) chap 11 2nd ed
[11] Arndt R A and MacGregor M H 1966 *Methods in Computational Physics* vol 6 (New York: Academic Press Inc.) p 253
[12] Morrison T I, Shenoy G K and Niarchos D 1982 *J. Appl. Cryst.* **15** 388
[13] Curis E and Bénazeth S 2000 *J. Synchrotron Rad.* **7** 262
[14] http://lise.lbl.gov/RSXAP/
[15] Ankudinov A L and Rehr J J 1997 *Phys. Rev.* B **56** R1712
[16] Matsumoto M and Nishimura T 1998 *ACM Trans. Model. Comput. Simul.* **8** 3
[17] Hamilton W C 1965 *Acta Cryst.* **18** 502
[18] Downward L, Booth C H, Lukens W W and Bridges F 2007 *AIP Conf. Proc.* **882** 129
[19] Hu, Y-J and Booth C H Elsewhere in this volume.
[20] Li G G, Bridges F and Wang X 1994 *Nucl. Instr. and Meth.* **A340** 420
[21] Curis E, Osán J, Flakenberg G, Bénazeth S and Török S 2005 *Spectrochimica Acta* Part B **60** 841
[22] Lide D R (ed) 1990 *CRC Handbook of Chemistry and Physics* (Boca Raton: CRC Press) pp 4–160 and 12–107 71st ed
[23] Student 1908 *Biometrika* **6** 1
[24] Newville M, Kas J J and Rehr J J Elsewhere in this volume.