



Chapter 5.9

Keywords: accuracy; information content.

Unknown systematic errors and their impact on the information content of the data

Corwin H. Booth*

Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. *Correspondence e-mail: corwin.booth@gmail.com

Understanding the ultimate accuracy of EXAFS results includes performing statistical analysis of the estimated parameter errors and minimizing/accounting for the effect of known systematic errors, such as pin holes, self absorption *etc.* Such known sources of error are considered elsewhere in this volume. Even in the event that all such sources of error are understood and minimized, high-quality EXAFS data are often still not fitted statistically well. In this short chapter, the effect of such unknown errors is considered with ultimate limitations on the accuracy of EXAFS data analysis.

1. Introduction

The techniques outlined in Booth (2024) consider the effect of normal, random noise on the estimated parameter uncertainties. Unfortunately, there are many sources of error that are not normal or random, or may have timescales that are too long to effectively average. For instance, sample inhomogeneities, pin holes, detector nonlinearity and many other non-random sources of error may not be properly understood within a given experiment. Even if one could remove all sources of errors that can be controlled, ultimately, fitting standards are limited in their accuracy. Even experimental lineshape standards are limited by how well the concept of chemical transferability works. All of these known sources of systematic errors are discussed elsewhere in this volume. Here, the ultimate absolute accuracy of EXAFS measurements is considered, both by direct comparison to standard materials and by the effect on real data where all of the currently known systematic error sources have been minimized or eliminated.

2. Absolute error estimates

As alluded to above, absolute errors in EXAFS are ultimately limited by the reference material used to determine the lineshapes in a fitting routine. In cases where the material itself can be used to provide the lineshape, femtometre-scale displacements have been observed in differential EXAFS measurements (Pettifer *et al.*, 2005). For actual bond-length determinations, absolute bond-length measurements can be accurate to within 0.001 Å in well defined cases such as simple molecular systems in gas-phase samples (Filipponi & D'Angelo, 1998). For more general absolute error estimates, limitations have been established through comparisons between crystallographic results and detailed EXAFS fit results (Li *et al.*, 1995). These comparisons were made using both experimentally derived standard lineshapes and those calculated by *FEFF5* (Mustre de Leon *et al.*, 1991). Considering first the comparison with theoretical lineshapes, nearest-neighbour pair distance differences between the EXAFS fits and the crystallographic data were generally better than

Related chapters

Volume I: 3.14, 3.43, 4.6, 4.7, 5.1, 5.2, 5.3, 5.6, 5.7, 5.8, 5.12, 5.13

0.01 Å and often better than 0.005 Å. Experimental lineshapes derived from one standard compound and applied to another do even better, with differences of only 0.001 Å between the measured and known pair distances. Values of σ^2 were also considered for the simplest cases where a correlated Debye model should apply, namely single-element metals, and the measurements generally agree with the prediction of the model within about 20%. Estimates of the number of neighbours are considered to be within 10% (Li *et al.*, 1995; Ravel & Kelly, 2007).

These comparisons were made under ideal conditions, only considering nearest neighbours with little overlap from further scattering shells. Making comparisons between theory and experiment for further shells becomes much more dependent on the fitting model, as overlap from other scattering shells, including those due to multiple scattering, affects the further shells to varying degrees in different materials. The copper-foil fits in Fig. 1 in Booth (2024) are not typical, in that the difference between the pair distances from the crystallographic and EXAFS results are all within 0.01 Å for the three fitted scattering shells out to $R = 4.4$ Å. Generally, a rule of thumb of ± 0.02 Å is employed for shells as distant as 4.4 Å.

3. Effects of unknown systematic errors on conventional error analysis

These ultimate limitations are evident in, for instance, the fit example in equation (1) of Booth (2024) to data from a copper foil. While the fit appears to be practically perfect, in fact the estimated error bars are small enough that virtually none of the data points agree within four standard deviations of the fit model, yielding a fit-quality estimate of $\chi^2/\nu \simeq 13\,000$, a quantity which should trend towards unity in a fit with many degrees of freedom such as this one (Booth & Hu, 2009).

In the copper-foil example (Booth & Hu, 2009), assuming only the estimated random uncertainty on all three single-scattering paths out to the Cu–Cu path at about 4.4 Å, the estimated uncertainty on R for each of these paths is 0.0001 Å, and likewise the estimated uncertainty on the σ^2 values of these same paths is 0.00001 Å², both of which are more than an order of magnitude smaller than the estimated absolute uncertainty on these quantities, which may be as small as

0.005 Å and 0.0005 Å², respectively (Li *et al.*, 1995). In a case such as this one, it may not make sense to report, or otherwise focus on, the random errors, which may then be misconstrued.

A common method for considering total errors, especially in such cases, although it is much more widely applied, is to base the data error estimate not on an actual estimate from multiple scans or some other method, but on the required error to force $\chi^2/\nu = 1$. That is, one can define a single $e = e_i$ for all i in equation (1) of Booth (2024) or alternatively scale the actual estimate errors $e_i = De_i$, such that $\chi^2/\nu = 1$. It is very important to realize that this method is not statistically sound and will only partially account for differences, whatever the source, between the model lineshapes and the data. Insofar as these differences are not normally distributed, such an estimate will fail, and so should be used with caution. For instance, if the difference is a phase shift $\varphi_d = kc_d$, where c_d is a constant, then the fitted bond length will be shifted in the final fit result and there will be no contribution to χ^2 and no contribution to the estimated uncertainty.

As an example of this scaling method, consider again the fit in equation (1) of Booth (2024) using a scaled error of $e = 2.1$ such that $\chi^2/\nu = 1$. With such a fit, the estimated uncertainties on the bond length increase from ± 0.0015 Å for the nearest Cu–Cu pairs to about ± 0.02 Å for the Cu–Cu pairs near 4.4 Å. Likewise, the uncertainties on σ^2 increase from 2.4% to 4.3% for these same scattering paths.

Funding information

This work was supported by the US Department of Energy (DOE), Office of Science (OS), Office of Basic Energy Sciences (OBES) under Contract No. DE-AC02-05CH1123.

References

- Booth, C. H. (2024). *Int. Tables Crystallogr. I*, ch. 5.8, 672–675.
- Booth, C. H. & Hu, Y.-J. (2009). *J. Phys. Conf. Ser.* **190**, 012028.
- Filipponi, A. & D'Angelo, P. (1998). *J. Chem. Phys.* **109**, 5356–5362.
- Li, G. G., Bridges, F. & Booth, C. H. (1995). *Phys. Rev. B*, **52**, 6332–6348.
- Mustre de Leon, J., Rehr, J. J., Zabinsky, S. I. & Albers, R. C. (1991). *Phys. Rev. B*, **44**, 4146–4156.
- Pettifer, R. F., Mathon, O., Pascarelli, S., Cooke, M. D. & Gibbs, M. R. J. (2005). *Nature*, **435**, 78–81.
- Ravel, B. & Kelly, S. D. (2007). *AIP Conf. Proc.* **882**, 150–152.