EXAFS and principal component analysis: a new shell game

Stephen R. Wasserman,[a*] Patrick G. Allen,[b] David K. Shuh[c], Jerome J. Bucher[c], and Norman M. Edelstein[c]

[a]*Advanced Photon Source, Argonne National Laboratory, 9700 S. Cass Ave., Argonne, IL 60439, USA,* [b]*G. T. Seaborg Institute for Transactinium Science, Lawrence Livermore National Laboratory, MS-231, Livermore, CA 94551, USA,* [c]*Chemical Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road Mailstop 70A-1150, Berkeley, CA 94720, USA. Email: srw@anl.gov*

The use of principal component (factor) analysis for the interpretation of EXAFS spectra is described. The components derived from EXAFS spectra share mathematical properties with the original spectra. As a result, the abstract components can be analyzed using standard EXAFS methodology to yield bond distances and other coordination parameters. The number of components that must be analyzed is usually less than the number of original spectra. The method is demonstrated using a series of spectra from aqueous solutions of uranyl ions.

Keywords: principal component analysis; factor analysis; EXAFS; uranyl

1. Introduction

Recently principal component analysis (PCA) has been used to examine several series of related x-ray absorption spectra, both XANES and EXAFS, from samples that exhibit changes in structure or composition (Coulston *et al.*, 1997; Fernandez-García *et al.*, 1995; Wasserman *et al.*, 1996, 1997, and 1998). PCA provides a systematic method for analyzing such series and, when used in conjunction with the conventional methods of XAS analysis, can simplify the process of extracting information from the various spectra.

PCA provides an alternative approach to determining the models or standards that are used in the interpretation of XAS data. In the traditional analysis of XANES or EXAFS, the model edges used, or the scattering phases and amplitudes necessary for the extraction of the coordinative structure, are taken from known compounds or computed using *ab initio* multiple-scattering calculations. In PCA, however, the standards used to reproduce the observed spectra are derived from the original spectra themselves. The resultant mathematical constructs are referred to as components or factors. The methodology of PCA imposes several requirements on these artificial standards. No two standards reproduce the same features in the spectra. There is a weighting factor associated with each standard which quantifies that standard's importance in reproducing the entire spectral series. Finally, the derived set of standards must be unique. Each of these requirements has a direct correspondence with basic concepts from linear algebra, including those of a vector space, orthogonality, eigenvalues, and normalized eigenvectors.

The number of components that are constructed equals the number of original spectra. This set is then reduced in size to the minimum necessary to reproduce the spectra to within the experimental error. The order in which the standards are included in the reproduction step is determined by the associated weighting factor. The number of components that regenerates the data equals the number of distinct species in the series. This process results in an alternative representation of the data, while preserving the underlying information. The mathematical details of the method are explained in detail elsewhere (Malinowski *et al.*, 1980; Wasserman, 1997).

In this paper we discuss the specific application of PCA to EXAFS spectra. To illustrate the unique aspects of PCA in analyzing EXAFS, we use a series of spectra from uranyl (UO_2^{2+}) ions in solution as a function of increasing chloride concentration (Allen *et al.*, 1997). The chloride ions replace the water molecules at the equatorial position of the uranyl. The axial oxygen atoms, which are bound to the uranium through double bonds, remain intact in each sample. We also demonstrate how PCA illuminates some of the inherent limitations in standard EXAFS analysis.

2. Experimental

The experimental details for the acquisition of the 10 absorption edge spectra and the extraction of the EXAFS have been described previously (Allen *et al.*, 1997). The chloride concentration varied from 0 to 14 M, using 0, 1, 2, 4, 6, 8 and 10 M solutions of HCl and 12 and 14 M solutions of LiCl. In addition, a sample that had only chloride in the equatorial position was prepared using a Dowex anion exchange resin. Transmission data were collected for the Dowex sample and each of the HCl solutions. For the two LiCl solutions, fluorescence detection was used. Scattering phases and amplitudes were determined from single scattering calculations within FEFF7. In the analysis, components 1 and 2 were fit simultaneously. The bond distance, disorder, and energy parameters for the equatorial oxygen and chloride ligands were constrained to be the same for these two components.

3. Results and Discussion

Figure 1 shows the first five of the 10 components calculated from the original EXAFS, as well as their weighting factors. The first three, the primary components, are sufficient to reproduce the spectral series to within experimental error (Figure 2). This result indicates that there are three distinct species within the entire series. Further analysis, which is beyond the scope of this paper, demonstrates that one of these species is the product of the use of the anion-exchange resin.

In the case of XANES, the derived components, with the exception of the first or most significant one, usually bear little re-
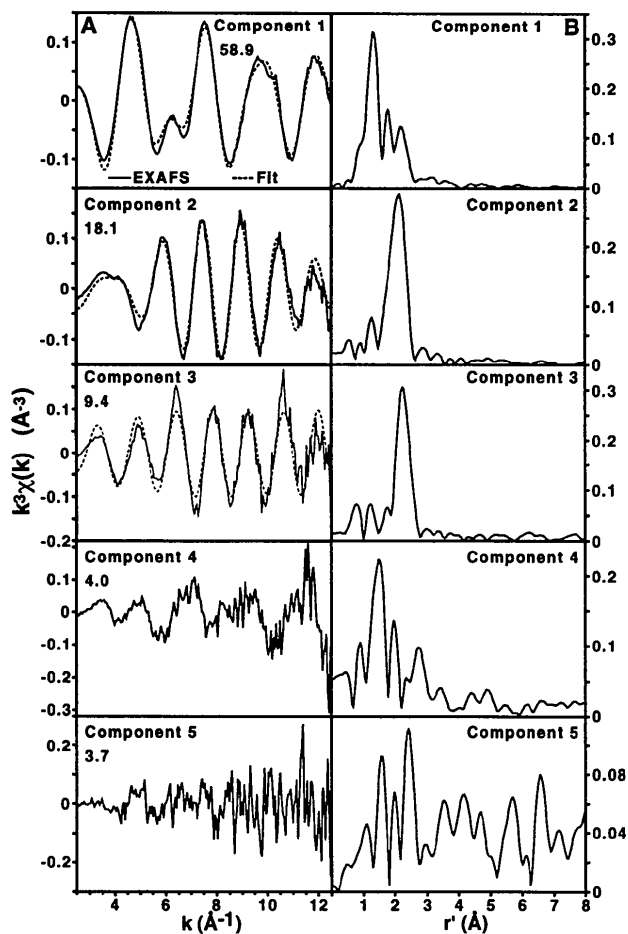
Figure 1

(A) The first five components derived from ten uranium L_3 edge k^3-weighted EXAFS spectra of uranyl ions in solutions with and without chloride anions. The relative weighting factors are given for each component. For the first three components, fits using scattering phases and amplitudes from FEFF7 are shown. The fits were performed using three shells (O_{ax}, O_{eq}, and Cl) for component 1, two (O_{eq}, and Cl) for component 2, and 1 (Cl) for component 3 in the range from 2.5 to 12.5 $Å^{-1}$. (B) Fourier transforms of the components. Transforms were taken in the range $\Delta k = 2$ to 13 $Å^{-1}$.

semblance to real absorption edge spectra (Wasserman *et. al*, 1996). However, the three primary components from the uranyl system each exhibit an oscillatory behavior similar to that found in EXAFS data. This property arises from the fact that each component is orthogonal to the others, just as the Fourier terms in each EXAFS spectrum are orthogonal. As a result, the components can be transformed to radial structure functions and fit to extract the coordination parameters. The Fourier transforms of the components are also shown in Figure 1.

Fitting of the coordination shells in the components gives radial distances that agree well with those deduced from standard EXAFS methods (Table 1). Both approaches indicate that the distance to the axial oxygen atoms is 1.76 Å. Each method concludes that the radial distance for the equatorial oxygen from the water molecules is 2.41 Å. In the case of the equatorial chloride ions, PCA differs by only 0.01 Å from the standard method for both the resin (2.67 vs. 2.68 Å) and solution-exchanged (2.71 vs. 2.72 Å) samples. Similar agreement is found
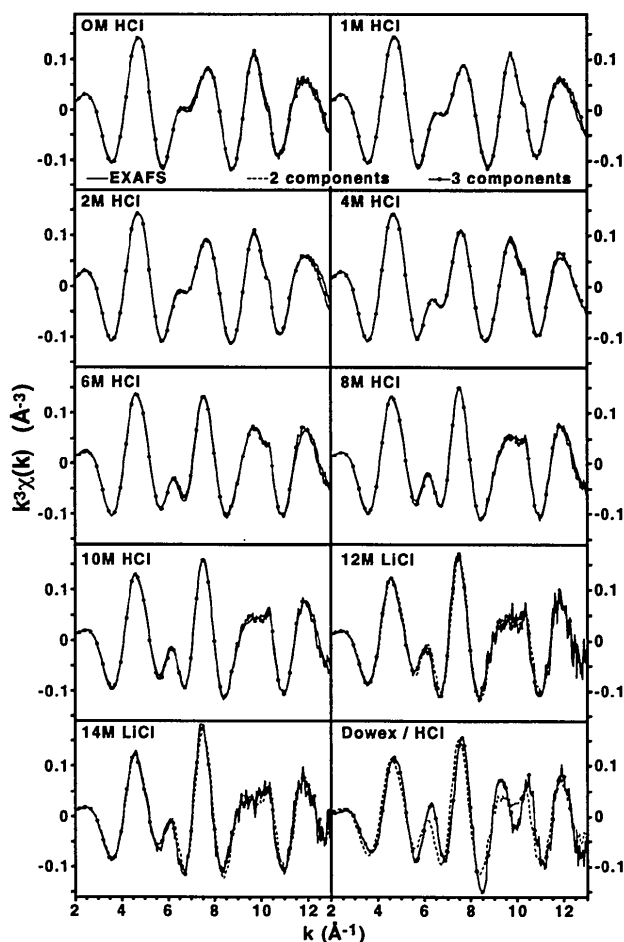
Figure 2

Reproductions of the EXAFS spectra from aqueous uranyl ions as a function of chloride concentration. Original EXAFS data (solid line), and reproductions with two (dashed line) and three components (solid line with circles.)

for the disorder and energy parameters. The coordination numbers for the equatorial ligands vary slightly between the two analytical approaches. Those from PCA range from -0.7 to 3.8 for the oxygen and from 0.2 to 5.6 for the chloride. The difference between the two methods reflects the coupling between coordination number and disorder. The best fits from PCA suggest disorder factors for the two types of equatorial ligands that are inverted from those used in the traditional analysis.

PCA is also useful for illuminating aspects of the original data that may not be evident in normal XAS analysis. One example is from the scattering due to the axial oxygen atoms. Since these atoms are always present, the accepted approach is to fit this coordination shell for the aqueous uranyl ion. The resultant parameters are then kept fixed for the analysis of other uranyl species. In our PCA analysis, we have used an equivalent procedure by normalizing the coordination numbers to the intensity of the scattering from axial ligands. However, PCA demonstrates that, in this series, the scattering due to the axial ligands is not constant. A feature that is always present will, in PCA analysis, appear only in the first component. In addition, in order to maintain constant intensity for this feature among the

Table 1

Comparison of EXAFS Fit Parameters: PCA vs. Standard

		PCA	Standard
Oxygen	r (Å)	1.76	1.76
(axial)	n	2	2
	σ^2 (Å^2)	.0015	0.0015 – 0.0022
	ΔE (eV)	-16.2	-10.2 – -12.3
Oxygen	r	2.41	2.41 – 2.44
(equatorial)	n	-0.7 – 3.8	0 – 5.3
	σ^2	0.0048	0.0070 (fixed)
	ΔE	-14.5	-10.2 – -12.3
Chloride	r	2.70	2.71 – 2.73
(equatorial)		2.68 (Dowex)	2.67 (Dowex)
	n	0.2 – 2.8	0 – 2.6
		5.6 (Dowex)	4.2 (Dowex)
	σ^2	0.0070	0.005 (fixed)
		0.0005(Dowex)	
	ΔE	-15.2	-10.2 – -12.3
		-10.4 (Dowex)	

Figure 3

Contribution of the first component to reproduction of the uranyl EXAFS spectra. The data have been normalized to a maximum value of 100.

various spectra, the amount of the first component used to reproduce each spectrum should be the same. Figure 3 shows the amount of the first component used in the reproduction of the uranyl spectra. These coefficients vary by more than 10 percent, demonstrating that the observed axial scattering changes among the samples. In standard XAS analysis, such changes are normally ascribed to variations in disorder. Indeed, standard analysis finds minimum disorder in the 6 and 8 M HCl solutions. PCA suggests that, for this system, a change in effective coordination number, rather than disorder, is more successful in describing the differences between the various EXAFS spectra. This conclusion applies only to the extracted EXAFS data. The chemistry of the uranyl system ensures that two oxygen atoms are bound to the central atom. With PCA it is possible to discern the differences in the scattering from the axial atoms before any fitting procedures are used.

Although standard analytical procedures conclude that the product from the Dowex resin is different from the other samples, PCA demonstrates that the spectra from the 12 and 14 M LiCl solutions are also distinct from the other spectra in the series. The first two abstract components are necessary to describe all of the uranyl spectra, which reflects the replacement of oxygen by chloride at the equatorial position. However, the third component has a significant contribution for only the LiCl solutions and the Dowex ion-exchange product. For the former two samples, the effect of the component is negative, while for the Dowex product the contribution is positive, with a magnitude more than twice that for the LiCl solutions. Experimentally these differences are likely to come from differences in detection mode, since fluorescence detection was used only for the LiCl samples. It is also possible that the change in counter-ion from a proton affects the observed EXAFS.

4. Conclusions

PCA offers another perspective on EXAFS analysis. In this example, it is necessary to fit only three components in order to
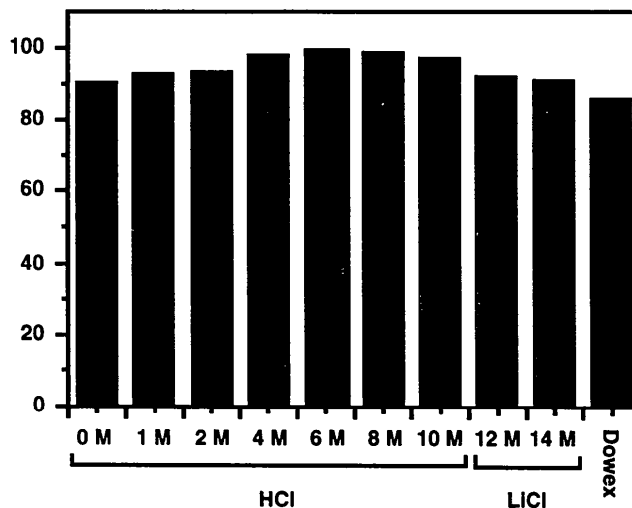
analyze the complete series of uranyl spectra. The fitting procedure is also simplified since the significant features in each component are generally intense. The PCA approach differs from traditional XAS methodology in that a single coordination shell will often be split between two or more components. It is therefore advantageous to fit simultaneously components that contain representations of common coordination shells. The fitting parameters from PCA apply to all of the spectra for which that component is significant, rather than to a single specific spectrum. A possible disadvantage of using PCA with EXAFS data is that negative coordination numbers are possible. For example, in the second component, the coordination number for the equatorial oxygen is less than zero. Usually, adding the effects of several components together will result in reasonable coordination numbers for the original spectra.

References

Allen, P. G., Bucher, J. J. , Shuh, D. K., Edelstein, N. M., & Reich, T. (1997). *Inorg. Chem.*, 36, 4676-4683.

Coulston, G. W., Bare, S. R., Kung, H. Birkeland, K., Bethke, G. K., Harlow, R., Herron, N. & Lee, P. L. (1997). *Science*, 275, 191-193.

Fernandez-García, M., Alvarez, C., M. & Haller, G. L. (1995). *J. Phys. Chem.*, 99, 12565-12569.

Malinowski, E. R. & Howery, D. G. (1980). *Factor Analysis in Chemistry*, New York, John Wiley and Sons.

Wasserman, S. R., Winans, R. E., & McBeth, R. (1996). *Energy & Fuels*, 10, 392-400.

Wasserman, S. R. (1997). *J. Phys. IV France*, 7, C2-203-205.

Wasserman, S. R., Soderholm, L. & Staub, U. (1998). *Chem. Mater.*, 10, 559-566.