

## MULTIVARIATE STATISTICAL ANALYSIS OF SPECTRUM LINES AND IMAGES

Ian M. Anderson and Jim Bentley

Metals & Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831

Recent developments in instrumentation and computing power have greatly improved the potential for quantitative imaging and analysis. For example, products are now commercially available that allow the practical acquisition of spectrum images, where an EELS or EDS spectrum can be acquired from a sequence of positions on the specimen. However, such data files typically contain megabytes of information and may be difficult to manipulate and analyze conveniently or systematically. A number of techniques are being explored for the purpose of analyzing these large data sets. Multivariate statistical analysis (MSA) provides a method for analyzing the raw data set as a whole. The basis of the MSA method has been outlined by Trebbia and Bonnet.<sup>1</sup>

MSA has a number of strengths relative to other methods of analysis. First, it is broadly applicable to any series of spectra or images. Applications include characterization of grain boundary segregation (position-),<sup>2-4</sup> of channeling-enhanced microanalysis (orientation-),<sup>5</sup> or of beam damage (time-variation of spectra). Second, no information available in the raw data is discarded *a priori*, since all independent spectral variations that can be resolved from statistical fluctuations are identified. By discarding all spectral components but those that have been resolved from the noise level, the large raw data files can be reduced to a manageable size without loss of information that is potentially available from the data set. Third, MSA generates its own basis spectra to characterize the data, rather than comparing the raw data to a set of basis spectra that have been arbitrarily chosen. There are also limitations for MSA relative to other techniques. Practical limitations on the size of the data sets to which it can be applied result from the computational demands of the method. This analysis has been successfully applied to a series of 256 spectra, each composed of 1024 channels; using Mathematica® code, the analysis required 10 min. of CPU time on a Macintosh 9500 PowerPC, running at 200 MHz and equipped with 64 Mbytes of memory. This number of spectra should be more than sufficient for a spectrum line ( $256 \times 1$ ), but marginal for a spectrum image ( $16 \times 16$ ) where, generally, more pixels would be desirable. A second limitation of MSA is its sensitivity to spectral artifacts and noise propagation. A third limitation involves interpretation of the spectral components identified by MSA, which becomes increasingly problematic as the number of significant components increases.

The information that can be extracted by MSA from a series of spectra is illustrated by an application to a TEM spectrum-line acquired with a Gatan Imaging Filter (GIF) at the Co-L edge for a phase boundary between the periclase- (CoO) and spinel- (Co<sub>3</sub>O<sub>4</sub>) structured phases of cobalt oxide, as shown in Fig. 1a. A series of 64 spectra, each of 512 channels, has been analyzed with MSA; the details of the acquisition have been given elsewhere.<sup>3,4</sup> The exponential variation (linear on the logarithmic plot) of the information content of all but the first three components of the variance in Fig. 1b is consistent with noise due to the Poisson statistics of the acquisition. The plot shows that three components are distinct from the noise level, but that the first accounts for 97.2% of the information content of the series of spectra. The second component accounts for 0.8% of the variation, and the third for 0.2%, with a signal-to-noise ratio of less than one. The noise contribution to a component can be estimated by extrapolating the exponential variation of the higher-order components. For the first component spectrum, shown in Fig. 1c, strong positive (negative) features correlate with intense features in the upper (lower) portion of the GIF spectrum line in Fig. 1a, which was acquired from the spinel (periclase); in fact, the component spectrum is approximately the difference of the spectral features of the two phases. The corresponding amplitudes of the first component (Fig. 1d) show that, in addition to the transition of ~12 nm width at the interface, there are gradients in this component in the two adjacent phases. The amplitude gradients in the adjacent phases correspond to chemical gradients that feed the oxidation reaction at the interface. The low noise levels of the first component spectrum and its amplitudes are consistent with the information content of this component being orders of magnitude greater than its noise level. In contrast, consider the spectrum in Fig. 1e, which is a

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

linear combination of the second and third principle components identified by MSA. The higher noise level evident in the spectrum is consistent with the proximity of these components to the noise level in Fig. 1a. The amplitudes of this component, shown in Fig. 1f, exhibit a lower noise level than its spectrum and the range of positions in Fig. 1f over which the amplitudes vary appreciably are clearly similar to those in Fig. 1d. Figs. 1e and f illustrate that the amplitude variations of a component may be clear even when the corresponding spectrum is quite noisy.<sup>6,7</sup>

#### References

1. P. Trebbia and N. Bonnet, *Ultramicroscopy* 34(1990)165.
2. J.M. Titchmarsh and S. Dumbill, *J. Microscopy* 184(1996)195.
3. J. Bentley and I.M. Anderson, *Proc. Microscopy and Microanalysis* 1996, 532.
4. I.M. Anderson and J. Bentley, *Mater. Res. Soc. Symp. Proc.* 458(1997) in press.
5. I.M. Anderson and J. Bentley, *Proc. Microscopy and Microanalysis* 1996, 550.
6. P.M. Rice, K.B. Alexander, and I.M. Anderson, *these proceedings*.
7. Research at the Oak Ridge National Laboratory (ORNL) SHaRE User Facility was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp., and by an appointment (IMA) to the ORNL Postdoctoral Research Associates Program, which is administered jointly by the Oak Ridge Institute for Science and Education and ORNL.

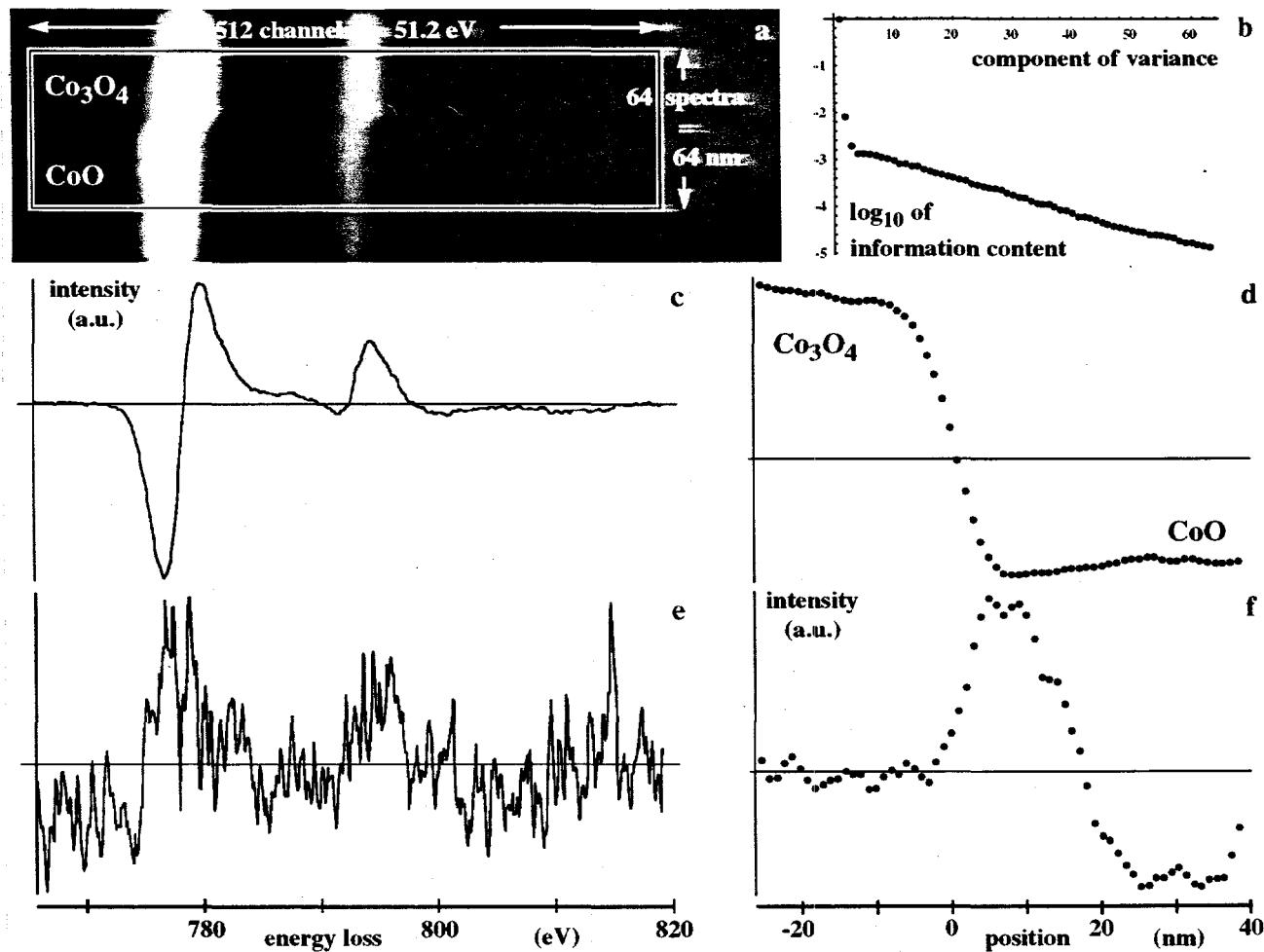


FIG 1 - An application of MSA: (a) logarithmic plot of the information content of the MSA-identified principle components of the series of spectra; (b) the spectrum line of the Co-L edge acquired with the GIF; (c) the first component (spectrum) of the variation; (d) the amplitude of the first component in each spectrum of the raw data; (e) a second component and (f) its amplitudes.

## **DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**DISCLAIMER**

**Portions of this document may be illegible  
in electronic image products. Images are  
produced from the best available original  
document.**