Fourier-ratio deconvolution and its Bayesian equivalent

R.F. Egerton a,b,*, F. Wang b, M. Malac a,b, M.S. Moreno c, F. Hofer d

a Physics Department, University of Alberta, Edmonton, Canada
b National Institute of Nanotechnology, 11421 Saskatchewan Drive, Edmonton, Canada
c Centro Atomico Bariloche, 8400 San Carlos de Bariloche, Argentina
d Graz University of Technology, Steyrergasse 17, A-8010 Graz, Austria

Abstract

We discuss how an inner-shell electron energy-loss spectrum can be processed using Bayesian (maximum-entropy or maximum-likelihood) deconvolution to simultaneously remove plural scattering and improve the energy resolution. As in Fourier-ratio deconvolution, a low-loss spectrum (recorded from the same area of specimen) is used as a kernel or resolution function. This procedure avoids the need to record the zero-loss peak in the absence of a specimen and uncertainties related to the width of the zero-loss peak. Unlike the case of Fourier-ratio deconvolution, we find that core-loss data do not require pre-edge background subtraction and extrapolation towards zero intensity; simply matching the intensity at both ends of the region is usually sufficient to avoid oscillatory artifacts. Using the low-loss spectrum as both data and kernel yields a zero-loss peak whose width provides an indication of the energy resolution as a function of the number of iterations. Finally, we argue that Fourier-ratio deconvolution or its Bayesian equivalent is the correct way to remove the substrate or matrix contribution to an energy-loss spectrum recorded from a particle on a substrate or embedded in a matrix.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: EELS; TEM; Deconvolution; Bayesian sharpening; Plural scattering

1. Introduction

Fourier-ratio deconvolution is a technique used in electron energy-loss spectroscopy (EELS) to remove the effect of plural scattering from the spectrum. It is usually applied to background-subtracted core-loss data, making use of a low-loss spectrum (recorded from the same region of the specimen) as if it were an energy-broadening function describing the energy resolution of the spectrometer system. In other words, the measured spectrum is taken to be a convolution of low-loss spectrum with an ideal core-loss spectrum, recorded using a spectrometer with perfect energy resolution and from a vanishingly thin specimen in which plural scattering would be negligible.

Transforming to Fourier space, the convolution becomes a product and the ideal spectrum is obtained by dividing Fourier coefficients, followed by an inverse transform. This procedure is effective in removing plural scattering but usually cannot eliminate the energy broadening of the spectrometer system because high-frequency spectral noise becomes amplified during the Fourier division. Therefore, a deconvolution function is introduced and the improvement in energy resolution is limited to typically a factor of two (Egerton and Crozier, 1988). In this respect, the procedure resembles Fourier-log deconvolution (Johnson and Spence, 1974), which removes plural scattering from an entire energy-loss spectrum but is often difficult to implement when core losses are involved because the full spectrum then has a very large dynamic range.

By adding constraints, such as the requirement for positive electron intensity, Bayesian methods (maximum-entropy or maximum-likelihood deconvolution) promise greater resolution enhancement. They have been used successfully in astronomy and more recently in EELS (Overwijk and Reefman, 2000; Kimoto et al., 2003; Gloter et al., 2003; Ishizuka et al., 2003; Lazar et al., 2006; Egerton et al., 2006). However, these procedures are iterative and as the number of iterations increases, so do the noise and the visibility of spectral artifacts (Egerton et al., 2006; Lazar et al., 2006; Fister et al., 2007).

The zero-loss peak can be used as an energy-resolution function (Bayesian kernel) but must be recorded separately (with specimen displaced out of the electron beam) or else extracted from the low-loss spectrum, and both of these procedures involve approximations, as we discuss below. In the
case of core-loss data, however, combined information about the instrumental resolution and the extent of plasmon (plural) scattering is available in the form of the low-loss spectrum. In fact, a recorded core-loss spectrum can be regarded as a convolution of the low-loss spectrum with an ideal core-loss spectrum, in which plural (plasmon + core-loss) scattering and instrumental broadening are absent (Egerton and Whelan, 1974). If so, it should be permissible to use the low-loss spectrum as a kernel in Bayesian deconvolution and simultaneously remove the effects of plural scattering and instrumental broadening, but with less noise penalty than with the Fourier-ratio procedure.

2. Bayesian methods

To sharpen (i.e. improve the energy resolution of) a recorded spectrum \( J(E) \) using a Bayesian method, a resolution function (or kernel) \( R(E) \) is specified, such that:

\[
J(E) = P(E) * R(E) + N(E) \tag{1}
\]

where \( P(E) \) is a sharpened spectrum, \( * \) represents the convolution operation and \( N(E) \) represents noise in the original data. The method is iterative and produces an output \( P(E) \) that is initially a highly smoothed version of \( J(E) \) but which exhibits (during further iteration) increasingly good resolution, as seen for example by a decreasing width of the zero-loss peak (ZLP) (Egerton et al., 2006).

A sensible choice for \( R(E) \) is a zero-loss peak \( Z(E) \) recorded by the same spectrometer system, since this ZLP would be a delta function if recorded by an ideal system. \( Z(E) \) can be recorded separately, as a no-specimen ZLP, by shifting or removing the specimen. However, the ZLP of the energy-loss spectrum may be broadened relative to the no-specimen ZLP, possibly by phonon or vibrational modes of energy loss (Katterwe, 1972; Egerton, 2003) but also by instrumental factors. Energy broadening due to spectrometer aberrations or spectrometer object size is greater in the presence of a specimen, whose scattering increases the angular width of the specimen-exit beam (Egerton, 1996; p. 77).

Alternatively, the resolution function \( R(E) \) can be based on the zero-loss peak present in an energy-loss spectrum. However, use of this option requires separating the elastic and inelastic portions of the spectrum; although the ZLP can be modeled using various functions, the procedure is approximate and can lead to errors in the data, particularly for energy losses below 5 eV (Rafferty et al., 2000).

One solution to this dilemma is to use an appropriate low-loss spectrum as the kernel \( R(E) \). By analogy with Fourier-ratio deconvolution, this Bayesian procedure should remove plural scattering from the data, in addition to improving the energy resolution. An important special case is that of core-loss data, meaning that \( P(E) \) is an ionization edge, usually with its pre-edge background and then subtracted. In this study, we have used the maximum-likelihood (Richardson–Lucy) method; maximum-entropy deconvolution has previously been shown to produce similar results (Ishizuka et al., 2003). We have employed mainly an algorithm available from HREM Research Inc. but as reported earlier, other implementations give an almost identical output (Egerton et al., 2006). We begin by discussing core-loss data recorded from relatively thick specimens (and high-energy ionization edges) where plural scattering is pronounced and would need to be removed as a first step in elemental quantification.

3. Deconvolution of core-loss spectra

Fig. 1 shows the low-loss spectrum and titanium K-edge recorded from a thick sample of titanium carbide (average thickness ~200 nm), together with a sharpened single-scattering distribution (SSD) derived by the Richardson–Lucy (RL) method. The K-edge data was not background-subtracted and exhibits a difference of electron intensity at the two ends of the range, resulting in oscillatory artifacts in the SSD that are similar to those produced by Fourier-log deconvolution, although considerably weaker. In addition, the K-loss data contains a recording artifact at 5400 eV that appears amplified in the SSD. Despite these artifacts, the SSD exhibits the correct overall shape for a K-edge, as opposed to the rounded profile of the original data that arises from plural (K-loss + plasmon) scattering.

Fig. 2 shows the same K-edge data with the recording artifact removed and its pre-edge background extrapolated and subtracted. There is still an artifact at the end of the range (above 5600 eV) but this is eliminated (in the circular data points) by extrapolating the K-loss intensity to zero at an energy (around 16,500 eV) well beyond the ionization threshold.

![Fig. 1](https://example.com/fig1.png)

Fig. 1. (Above) Ti K-edge (containing a recording glitch at 5400 eV) and low-loss spectrum recorded from TiC. (Below) K-edge single-scattering profile produced by Richardson–Lucy deconvolution, showing oscillatory artifacts due to the 5400-eV glitch and due to the difference in intensity between the two ends of the K-loss range.
Even though the RL method appears relatively tolerant to the discontinuity in core-loss intensity between the two ends of the range, we tried adjusting the position and energy width of the core-loss region so that the intensity at both ends was equal. As seen in Fig. 3, which relates to a thin specimen containing iron nanoparticles embedded in silicon dioxide, this simple procedure results in a SSD that exhibits no oscillatory artifacts, despite some difference in slope at the two ends of the data. By avoiding the need for background subtraction and extrapolation, this intensity-matching procedure adds to the convenience of the spectral processing.

To check the accuracy of the RL method (using a low-loss spectrum as the kernel) we compared the output with that produced by the Fourier-ratio method on the same data. As seen in Fig. 4, the results were essentially identical, except for differences in the noise content. However, the spectral sharpening introduced by the RL method has enhanced the visibility of a pre-edge peak (at 537 eV) that originates from oxygen atoms directly bonded to iron rather than oxygen present in the surrounding SiO$_2$ (Wang et al., 2007).

Another test of accuracy was to re-convolute the RL output with the low-loss spectrum. As shown Fig. 5, we obtained a result essentially identical to the original raw data, giving us extra confidence in the procedure.

4. Deconvolution involving low-loss spectra

Although the Fourier-ratio method is most often used to remove plural scattering from core-loss data, it can be applied in other ways. For example, it has been used to remove the zero-loss tails from the low-loss region, using a ZLP recorded in vacuum as the resolution function and a tail-free Gaussian as the noise-limiting reconvolution function (Batson et al., 1992; Rafferty et al., 2000). The Richardson–Lucy method can be used as an equivalent alternative but can give rise to weak satellite wing peaks (Egerton et al., 2006) so in this instance the Fourier-ratio procedure represents a safer alternative.
If we use the low-loss spectrum with its ZLP as both the data and kernel in Bayesian deconvolution, the result should ideally be a delta function at $E = 0$. In practice, the RL output is found to contain a peak (at $E = 0$) whose width $W(N)$ is reduced as the number $N$ of iterations increases; see Fig. 6. There are also some artifact peaks that appear for large $N$.

We believe that this property can be useful where Bayesian deconvolution is being applied to data (e.g. a core-loss region) that contains no zero-loss peak and where the energy resolution is therefore uncertain. To take advantage of this option, the Bayesian deconvolution is repeated with the low-loss region used as the data as well as the kernel. $W(N)$ then provides an indication of energy resolution to be expected after a given number of iterations. Results obtained using the low-loss spectrum recorded from a 50-nm thin film of silicon nitride are shown in the inset of Fig. 6.

5. EELS of multicomponent systems

An interesting situation occurs when an energy-loss spectrum has been recorded from a particle on a substrate. Assuming for the moment that the particle is on the beam-exit side of the specimen, inelastic scattering in the substrate can be thought of as equivalent to a peculiar form of energy broadening of the electron source. Its effect can therefore be removed by Fourier-ratio deconvolution, using the spectrum recorded from the particle/substrate combination as the raw data and the spectrum from the bare substrate as an energy-resolution function. The result should be the energy-loss spectrum of the particle alone. The Bayesian equivalent would be to use the particle/substrate spectrum as the data and substrate spectrum as the kernel.

This argument assumes that the energy distributions of the electron source and of each component of the specimen convolve together to produce the measured spectrum. Because convolution is commutative, removal of the substrate contribution should work equally well if the particle is on the beam-entrance surface of the specimen. Likewise, Fourier-ratio deconvolution should be capable of removing the matrix contribution in the spectrum recorded from a particle (or overlapping particles) surrounded by a matrix of different composition.

Our discussion ignores the fact that the specimen also scatters electrons elastically and in the case of a crystalline component, this elastic scattering (diffraction) is not axially symmetric so that the convolutions involved are in principle two-dimensional. However, simple deconvolution (involving energy loss but ignoring angular distributions) has been found to be accurate enough for most spectroscopy involving an on-axis collection aperture, a property that may derive partly from the Lorentzian angular distribution of inelastic scattering (Egerton, 1996, p. 255). We also ignore the excitation of plasmon modes at the external surfaces, implying $t \gg \nu h \omega \sim 5$ nm typically ($\nu$ being the incident-electron speed and $h \omega l/2\pi \sim 20$ eV being a typical energy loss), as well as collective modes at the interface that become pronounced for layer thicknesses below about 10 nm (Bolton and Chen, 1995).

As always, experiment provides the ultimate test. Fig. 7a shows energy-loss data (solid circles) recorded from a gold particle on a carbon substrate, together with data (hollow circles) taken from a nearby area of bare substrate. The result of Fourier-ratio deconvolution (based on these two spectra) is shown as the dotted curve; the Bayesian equivalent (Richardson–Lucy deconvolution) is the solid curve. Although these two procedures give slightly different outcomes, both are similar to the spectrum of pure gold taken from the Gatan EELS atlas (solid line in Fig. 7b).

Fig. 7b also shows the result of simply subtracting the substrate spectrum from the particle/substrate spectrum. Solid circles represent the result of subtraction after scaling the substrate spectrum so that the integral $I_0$ of its zero-loss peak matched that of the particle/substrate spectrum before subtraction. Hollow circles represent a subtraction after matching the total integrals $I_0$ of the two spectra. In neither case does the result resemble the gold reference spectrum. Although subtraction might seem intuitive in terms of some additive property of electron intensity, the probabilities of an electron suffering a given energy loss in each layer are multiplicative and the energy losses themselves are additive, which leads to convolution of the spectra.

If deconvolution is the correct approach, the derived spectrum of a particle should presumably contain the correct zero-loss integral $I_0$ in relation to its overall integral $I$. According to Poisson statistics, $\ln(I/I_0) = t/\lambda$, where $t$ is the thickness of a layer and $\lambda$ is a mean free path for inelastic scattering in the appropriate material. Because the scattering parameter $t/\lambda$ is additive between layers (Egerton, 1996, p. 187), its expected value for the particle is just the difference in $\ln(I/I_0)$ between the particle/substrate and substrate-only spectra. Table 1 indicates that spectra obtained by Fourier-ratio and Richardson–Lucy deconvolution give $t/\lambda$ close to the expected value, whereas those derived by scaled subtraction do not.

An essential difference between addition and convolution is that the latter transfers intensity towards higher energy loss, corresponding to an electron being inelastically scattered in both layers. This energy shift is removed by deconvolution but...
not by subtraction. The importance of this effect increases with the relative intensity of the inelastic scattering relative to the zero-loss intensity, in other words with the scattering parameter \( t/l \). Therefore, simple subtraction of a substrate or matrix component could be an acceptable approximation if \( t/l < < 1 \), where \( t \) and \( l \) are to the thickness and inelastic mean free path of the substrate.

Valence-electron (plasmon) losses in the substrate will also shift the core-loss spectrum of the particle towards higher energy loss. Therefore, the requirement \( t/l < < 1 \), where \( \lambda \) is the plasmon mean free path in the substrate, applies also to removal of a substrate core-loss component by simple subtraction. If this criterion is not met, a more correct procedure for removing the substrate contribution from core-loss data is to deconvolve the complete particle/substrate spectrum (from zero loss to the end of the core-loss region) using the complete substrate spectrum as the kernel or resolution function. Alternatively, the particle core-loss spectrum can be obtained by subtraction and then deconvolved using the substrate low-loss spectrum as kernel.

An alternative situation is where the electrons (within the analyzed area) pass through neighbouring regions of different material (the two components are then sampled “in parallel” rather than “in series”). In that case, energy-loss intensities should be additive and it seems justifiable to subtract a suitably scaled substrate spectrum from the spectrum of a particle surrounded by substrate. A more complicated situation is where an energy-loss spectrum is recorded from many particles embedded in a matrix, so that both serial and parallel paths are present and where some form of effective-medium approach may be appropriate (Howie and Walsh, 1991).

6. Discussion and conclusions

We have discussed how Bayesian deconvolution, using a low-loss spectrum as the kernel, can remove the effect of plural scattering from a core-loss spectrum and simultaneously improve the energy resolution. The improvement in resolution depends on the number of iterations and can be judged by performing a self-deconvolution of the low-loss spectra.

Typically, 15 iterations are used in Richardson–Lucy deconvolution and the instrumental energy broadening is removed only partially; a much larger number of iterations result in noise amplification, visible in the form of spectral artifacts. Further sharpening could be achieved in a second RL-deconvolution step, using a Gaussian or Lorentzian kernel and with periodic smoothing to limit noise if necessary. With data that have sufficient signal/noise ratio, it is possible to go even further and remove the effect of lifetime and final-state broadening (Filipponi, 2000), although this possibility is ultimately limited by noise amplification (Fister et al., 2007).

Fourier-ratio and Fourier-log techniques can also remove plural scattering and achieve some improvement in energy resolution, particularly if the data are coarse-sampled. However, the Bayesian procedures benefit from fine sampling and can provide more substantial sharpening, making weak spectral features visible. Even so, the output should be carefully
inspected (as a function of the number of iterations) to check for the appearance of spectral artifacts.

In the case of a composite sample, such as a particle P on a substrate S, Fourier-ratio deconvolution or its Bayesian equivalent can be used to recover the energy-loss spectrum of P, provided the spectrum of S can be recorded separately and used as an energy-resolution function or kernel. This deconvolution procedure should be used in preference to simple subtraction of intensities unless the scattering parameter $t/\lambda$ of S is small.

Acknowledgments

The authors wish to thank Kazuo Ishizuka, Alexandre Gloter, Lijun Wu and Lin de Noyer for the use of their deconvolution programs, and the Natural Sciences and Engineering Research Council and the National Research Council (NINT) for financial support. M.S.M. also thanks CONICET (Argentina) for partial support.

References


