

# Multicomponent Signal Unmixing from Nanoheterostructures: Overcoming the Traditional Challenges of Nanoscale X-ray Analysis via Machine Learning

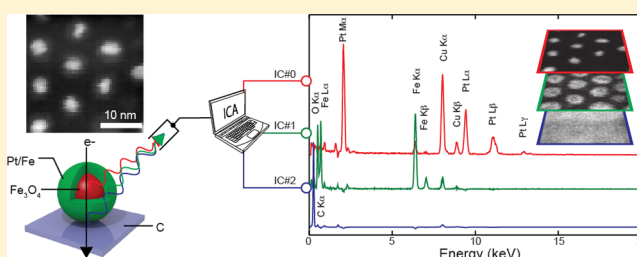
David Rossouw,<sup>\*,†</sup> Pierre Burdet,<sup>†</sup> Francisco de la Peña,<sup>†</sup> Caterina Ducati,<sup>†</sup> Benjamin R. Knappett,<sup>‡</sup> Andrew E. H. Wheatley,<sup>‡</sup> and Paul A. Midgley<sup>†</sup>

<sup>†</sup>Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge CB3 0FS, United Kingdom

<sup>‡</sup>Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

**ABSTRACT:** The chemical composition of core–shell nanoparticle clusters have been determined through principal component analysis (PCA) and independent component analysis (ICA) of an energy-dispersive X-ray (EDX) spectrum image (SI) acquired in a scanning transmission electron microscope (STEM). The method blindly decomposes the SI into three components, which are found to accurately represent the isolated and unmixed X-ray signals originating from the supporting carbon film, the shell, and the bimetallic core. The composition of the latter is verified by and is in excellent agreement with the separate quantification of bare bimetallic seed nanoparticles.

**KEYWORDS:** ICA, EDX, TEM, electron microscopy, nanoparticle



The transmission electron microscope (TEM) is a popular analytical tool for nanoscale characterization due to its high flexibility and capacity to perform spectroscopy at high spatial resolution.<sup>1</sup> As such, it plays a vital role in the continued growth of the nanosciences. In recent years, energy dispersive X-ray spectroscopy (EDX), performed in the TEM, has experienced a renaissance due to the development of new, more efficient X-ray detectors combined with an increased solid angle of detection.<sup>2,3</sup> It is now possible, and becoming commonplace, to acquire EDX spectra from thin specimens in a reasonable time frame. However, heterogeneous volumes remain particularly challenging to characterize by TEM techniques in regions where there is a spatial overlap of different phases within the beam path, such as a second phase precipitate embedded inside a matrix. The signals from different depths in the beam path are mixed when they are detected and must therefore be separated in order to find the composition of the individual phases present. In electron energy-loss spectroscopy (EELS), blind source separation (BSS) methods such as independent component analysis<sup>4</sup> (ICA) and non-negative matrix factorization<sup>5</sup> have been applied to the separation of components from a mixture.<sup>6–9</sup> In principle, the same approach could be taken for EDX spectrum image (SI) analysis, as has been demonstrated previously.<sup>10,11</sup> However, although in EELS, the energy loss near edge structure (ELNES) can be used to verify the accuracy of the decomposition,<sup>6</sup> in EDX the lower energy resolution prevents a similar corroboration method. Here, we apply ICA to EDX spectrum images of core–shell nanostructures to recover the composition of the buried cores,

and we verify the result by the separate quantification of bare bimetallic seed particles.

**Materials and Methods.** Like Co@Fe<sub>3</sub>O<sub>4</sub>, which recently underwent extensive study to elucidate both the oxidative stability of the core and the phase of the shell,<sup>12</sup> FePt@Fe<sub>3</sub>O<sub>4</sub> core–shell nanoparticles can be used as building blocks to form nanocomposites with enhanced magnetic properties with the potential for novel applications. These include magnetic data storage, catalysis, and targeted drug delivery.<sup>13–15</sup> In terms of nanomagnetism specifically, our hope is that the presence of an inert oxide shell may function to inhibit agglomeration of the cores upon annealing; this is a necessary step in creating an ordered L1<sub>0</sub> FePt bimetallic structure with a higher magnetic coercivity. The determination of the core composition is important for tailoring their synthesis in order to ultimately achieve the desired 50:50 FePt alloy composition.

A solution of core–shell particles was drop-cast onto a 3 mm holey carbon copper grid. EDX data were acquired using an FEI Osiris TEM equipped with a high brightness Schottky X-FEG gun and a Super-X EDX system comprising four silicon drift detectors, each approximately 30 mm<sup>2</sup> in area and arranged symmetrically around an optic axis to achieve a collection solid angle of 0.9 sr<sup>3</sup>. EDX data were collected in the form of spectrum images, in which a focused electron probe was scanned in a raster across a region of interest in the scanning

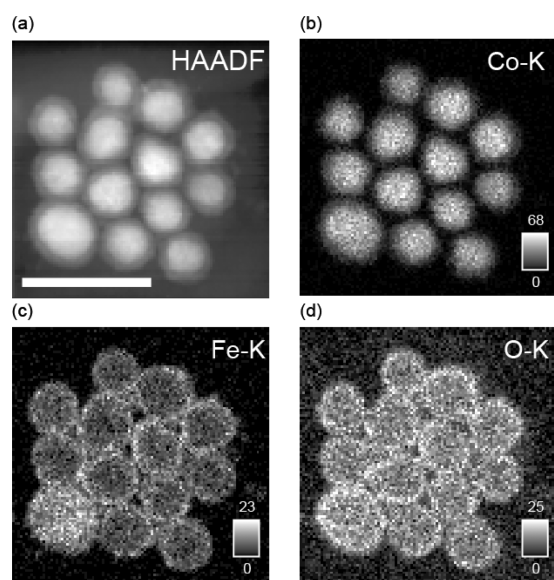
**Received:** February 3, 2015

**Revised:** March 6, 2015

**Published:** March 11, 2015

TEM (STEM). At each point in the scan, structural information was obtained from the electron scattering incident on a high angle annular dark field (HAADF) detector, and simultaneously, an EDX spectrum was obtained by collecting X-rays emitted from the local volume probed by the electron beam. The resulting EDX spectrum image was a three-dimensional data set whose ( $x, y$ ) axes correspond to the position of the probe and whose  $z$  axis corresponds to the energy of the detected X-ray. Spectrum images were acquired with a probe current of approximately 0.7 nA, an acceleration voltage of 200 kV, a spatial sampling of between 0.5 and 1 nm/pixel and 50–100 ms/pixel dwell times. TIA software was used for acquisition and HyperSpy<sup>16</sup> for data analysis. ICA was performed using the FASTICA algorithm<sup>17</sup> as implemented in Scikit learn.<sup>18</sup> X-ray intensities were obtained by fitting a model of the EDX spectra to the experimental data using weighted least-squares and atomic fractions were quantified from intensities using the Cliff–Lorimer quantification. The EDX model and the quantification were implemented in HyperSpy and will be available in future releases of the software. The  $k$ -factors used were provided by the EDX system manufacturer Bruker.

**Results.** *Co@Fe<sub>3</sub>O<sub>4</sub> Core–Shell Nanoparticles.* Figure 1a displays a HAADF STEM image obtained during the



**Figure 1.** EDX spectrum image of a Co@Fe<sub>3</sub>O<sub>4</sub> core-shell nanoparticle cluster. (a) HAADF STEM image shows that the particles have a core-shell construction. Elemental maps of (b) cobalt, (c) iron, and (d) oxygen display the location of the various elements with respect to the particle morphology (scale bar = 50 nm, greyscale = X-ray counts).

acquisition of a spectrum image enclosing a cluster of 13 Co@Fe<sub>3</sub>O<sub>4</sub> nanoparticles. Although the particle morphologies are seen to vary slightly from one particle to another, the majority of particles have a round core approximately 20 nm in diameter surrounded by a thin shell approximately 5 nm in thickness. The EDX elemental maps for cobalt, iron, and oxygen (Figure 1b–d), obtained by integration of the element's background-subtracted K-line X-ray peak, show that individual particles are comprised of a cobalt core surrounded by a shell composed of iron and oxygen, as expected. The largest particle in the lower left region of the map appears to have more iron in the core compared to the smaller particles. These conventional

EDX element maps show the location of the various elements, but the composition of the particle core, for example, cannot be determined by elemental mapping due to the presence of the shell above and below the core in projection.

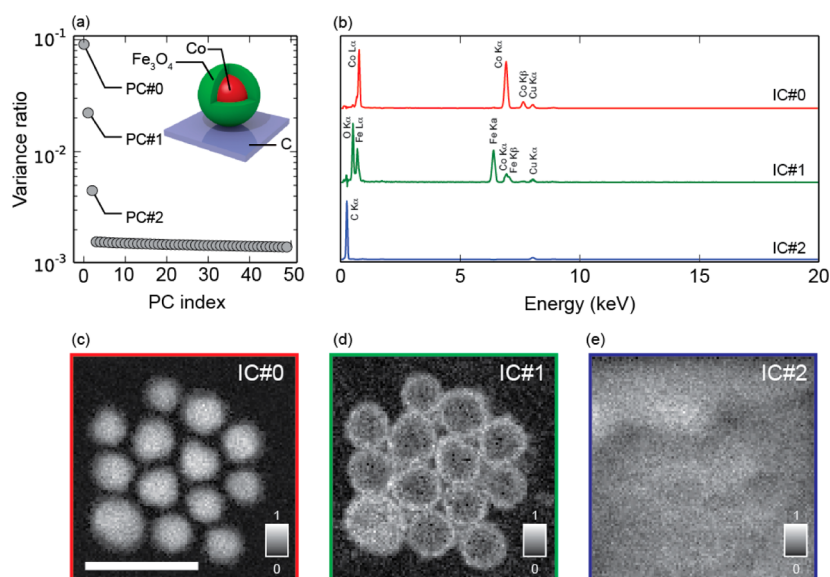
The EDX spectrum image data of the Co@Fe<sub>3</sub>O<sub>4</sub> nanoparticle cluster shown in Figure 1a was subsequently processed using BSS methods in HyperSpy. Figure 2 displays a summary of the BSS results. First, the spectral dimension in the data set was binned by four from 5 eV/channel to 20 eV/channel in order to increase the number of counts per channel. Next, a linear variance-stabilizing transformation for Poisson statistics<sup>19</sup> was applied to the data. We note that the binning step is necessary in order to optimize the accuracy of the variance stabilization channel.<sup>20</sup> Then, we performed PCA for dimensionality reduction purposes. The first three principal components, PC#0, PC#1, and PC#2, exhibited significantly greater variance than the remaining components (Figure 2a), which suggests that there are only three phases present in the sample. That being the case, those three PCA components should be a linear combination of the spectra and distribution maps of those phases, but the mixing matrix is unknown. Next, we compute numerically the first derivative of the PCA spectral components in order to diminish the correlation caused by the EDX background, and we use FastICA<sup>17</sup> to estimate the mixing matrix and compute the independent components (ICs) IC#0, IC#1, and IC#2 (Figure 2b) and their distribution maps (Figure 2c–e) from the PCA results. Component independence is a much more stringent property than uncorrelatedness imposed by PCA. Further details on the ICA method can be found in the literature.<sup>21</sup> If we disregard the small copper peaks contained in all the independent components, likely originating from the copper support mesh, we see that IC#0 contains cobalt X-ray peaks, IC#1 iron and oxygen peaks, and IC#2 a carbon peak. The three ICs appear to belong to the three phases present in the originally scanned area: the core, shell, and supporting film. This hypothesis is explored further in the next section. At this point, however, it is important to note that, unlike in the conventional EDX mapping shown in Figure 1, no elements were selected prior to performing ICA, and thus, the analysis is free of external bias, except for the choice of the number of components.

*FePt@Fe<sub>3</sub>O<sub>4</sub> Core–Shell Nanoparticles.* We now move on to the analysis of a second cluster of particles comprised of a bimetallic iron/platinum core surrounded by an iron oxide shell. A crystalline core surrounded by a polycrystalline oxide shell is observed in the representative high-resolution STEM HAADF image shown in Figure 3.

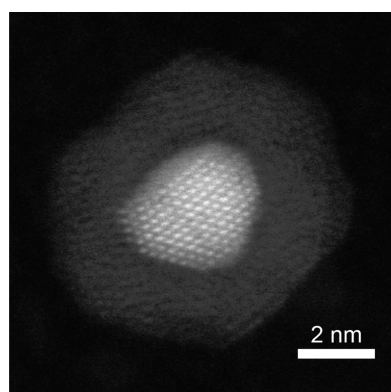
Across the sample the particle morphologies were found to have a mean core diameter of approximately 3.3 nm and mean shell thickness of approximately 1.7 nm (Figure 4a). Also visible were pure iron oxide particles (in the lower right-hand corner of Figure 4a).

From the selected element maps (Figure 4c–e), it is clear that, with the exception of the two particles in the lower right-hand corner, the particles are comprised of a platinum rich core surrounded by an iron oxide shell. However, it is not clear from the maps alone whether iron is present in the core. By conventional elemental mapping, one cannot tell whether the particles contain a pure platinum core surrounded by an iron oxide shell or whether iron is alloyed with platinum to form a bimetallic core.

We now address these questions by performing BSS on the same EDX spectrum image using the same procedure detailed



**Figure 2.** Result of BSS by PCA and ICA of an EDX SI of a Co@Fe<sub>3</sub>O<sub>4</sub> core-shell nanoparticle cluster. (a) Scree plot of the first 50 principal components. (b) Corresponding independent component spectra contain the expected X-ray lines for the elements present. Independent component maps (c–e) show that (c) IC#0 is concentrated in the nanoparticle cores, (d) IC#1 in the shells, and (e) IC#2 everywhere on the carbon supporting film (scale = 50 nm, greyscale = normalized weighting).



**Figure 3.** High-resolution STEM HAADF image of a FePt@Fe<sub>3</sub>O<sub>4</sub> bimetallic core-shell nanoparticle.

before. Once again, by inspecting the scree plot (Figure 4b), we find that the sample consists of three phases. The spatial distribution of IC#0 is concentrated in the particle cores, IC#1 in a shell around the cores, and IC#2 is approximately uniformly distributed over the map (Figure 4f–h). If we again disregard the spurious copper peak, IC#0 contains iron and platinum X-ray peaks, IC#1 iron and oxygen peaks, and IC#2 a single carbon peak (Figure 4i). From this analysis it appears that the ICA components represent the different phases present in the EDX spectrum image; IC#0, IC#1, and IC#2 genuinely represent the bimetallic FePt cores, iron oxide shells and the carbon support, respectively.

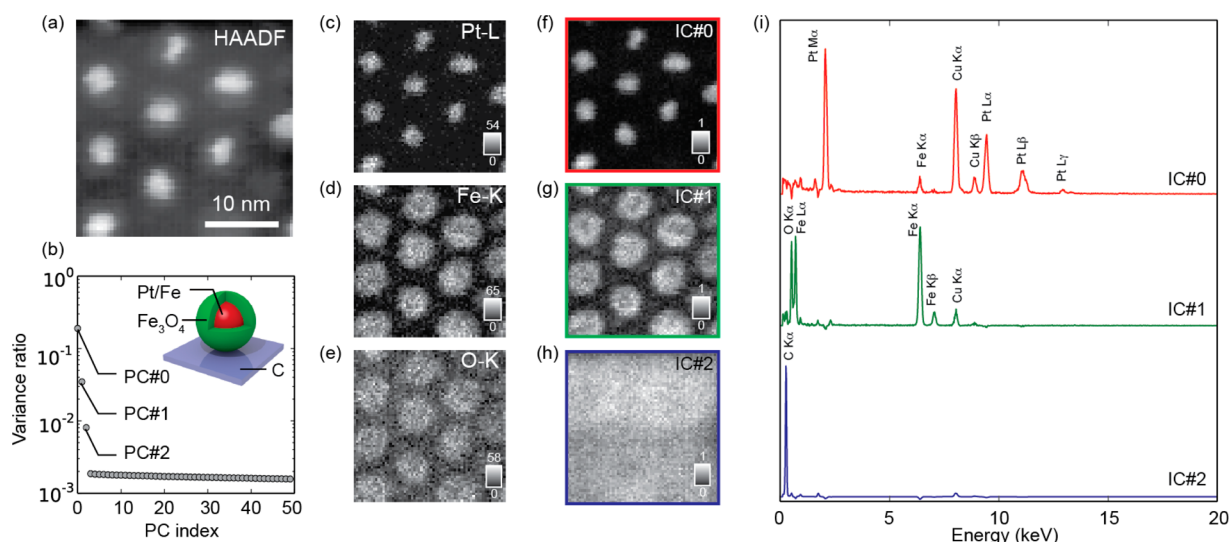
In order to verify the accuracy of the ICA results, we evaluated whether IC#0 represented the true composition of the core by analyzing bare FePt seed particle clusters obtained from the same chemical synthesis but extracted prior to the shell addition step. Being made from the same synthesis, the composition of the bimetallic seed particles are expected to match to the composition of the bimetallic cores in the core-shell particles analyzed in Figure 4.

A total of 12 EDX SIs were acquired in order to capture and analyze multiple FePt bimetallic seed clusters. Image segmentation was performed using thresholding<sup>22</sup> or a watershed algorithm<sup>23</sup> where appropriate. The segmentation of one of the EDX SIs is shown in Figure 5a,b.

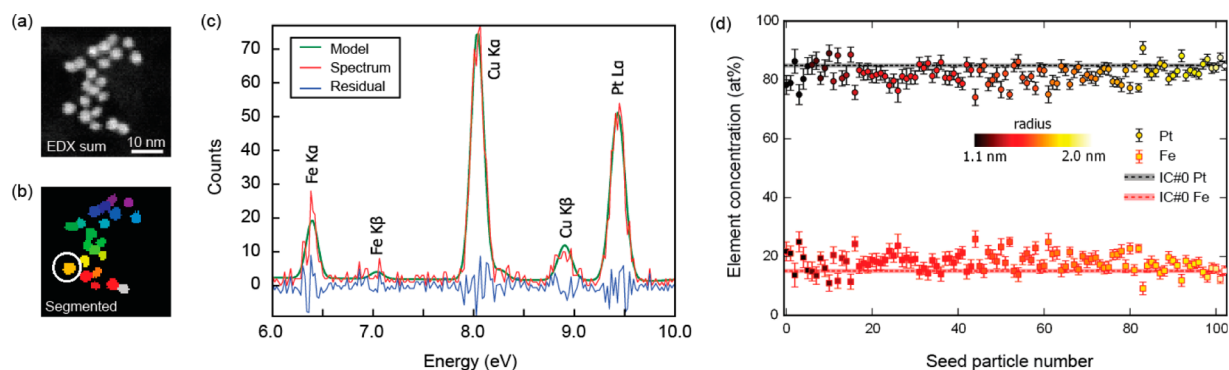
To accurately extract the intensity of the Fe K $\alpha$  and Pt L $\alpha$  peaks, a model composed of one Gaussian per X-ray line and a background based on Kramers and Small expressions as developed elsewhere<sup>24</sup> was used, as shown in green in Figure 5c. The only free parameters of the model were the area of the Gaussian and the height of the background, which was negligible. The mean reduced  $\chi^2$  over the fit of all particles was 1.01, indicating that the discrepancies between the model and the data are in accordance with the Poisson noise variance. The intensities of Fe K $\alpha$  and Pt L $\alpha$  peaks in the fitted model were quantified using the Cliff–Lorimer method. The obtained compositions are plotted for each particle in Figure 5d along with the fitting error estimated from Poisson statistics. The raw data were decomposed using PCA on all FePt seed data and the first three components were retained for noise reduction. The 103 individual FePt bimetallic seeds analyzed were found to have a mean composition of 82.0 at. % Pt and 18.0 at. % Fe with a standard deviation of 3.3 at. % Pt. By comparison, the composition of IC#0 was calculated to be 84.9 at. % Pt, which lies well within one standard deviation of the average bimetallic seed composition. The data points in Figure 5d are displayed in order of ascending particle size. As such, the calculated compositions on the right-hand side tend to have a smaller error bar on account of the signal–noise ratio being higher for larger particles. The compositions of the larger seed particles are also closer to the mean composition and to the composition of IC#0, confirming the homogeneity of the composition and the validity of the BSS analysis.

**Discussion.** A comparison of raw EDX spectra extracted from FePt bimetallic seed particles and from pure Fe<sub>3</sub>O<sub>4</sub> particles with IC#0 and IC#1, respectively, is provided in Figure 6. In both cases, the ICs were scaled by a constant to obtain a best fit to the raw spectra. Despite the strong overall





**Figure 4.** ICA of a cluster of bimetallic platinum/iron nanoparticle seeds coated by  $\text{Fe}_3\text{O}_4$  shells. (a) HAADF STEM image displays the core-shell structure of the nanoparticles. (b) Scree plot of the first 50 principal components showing the first three components lying above the noise. (c–e) Element maps of (c) platinum, (d) iron, and (e) oxygen. (f–h). The IC maps (f) IC#0, (g) IC#1, and (h) IC#2 and (i) the corresponding IC spectra.



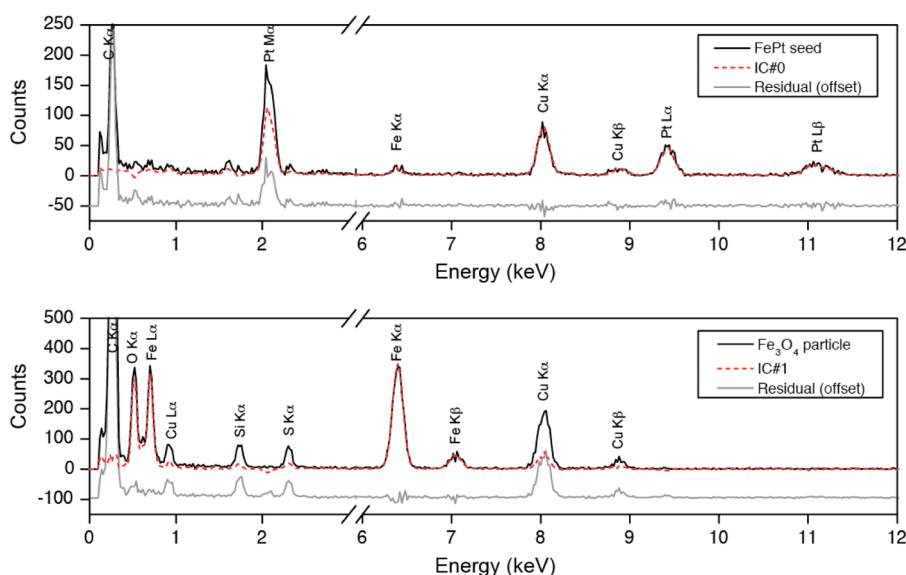
**Figure 5.** Summary of the composition of 103 bare FePt bimetallic nanoparticles extracted from a synthesis prior to the shell addition step. (a) Selected cluster of FePt seed particles. (b) Segmentation of the EDX spectrum image prior to quantification of each particle. (c) Fitting of the EDX spectrum from a single seed (circled in (b)) to a model spectrum to determine the Fe  $K\alpha$  and Pt  $L\alpha$  peak intensities. (d) Particle seed compositions obtained by quantifying the fitted intensities from 103 different particles (error bars = 1 standard deviation).

agreement in each case, some differences are seen for individual X-ray peaks. The carbon peak difference in both cases is caused by the separation of carbon into a different component (IC#2). The difference in the shell Cu peaks are not due to a compositional difference as the Cu signal is spurious in origin. The iron oxide particle spectrum also contains small silicon and sulfur peaks which likely originate from residue on the carbon film. The difference in the Pt  $M\alpha$  core peak may be due to the attenuation of Pt  $M\alpha$  X-rays in the shell of the core-shell nanoparticles and in nearby particles along the trajectory to the detector. The strong overall similarity between the raw and IC spectra provide direct evidence showing that the spectral components extracted by ICA from the core-shell spectrum image data are strongly representative of the buried core, surrounding shell, and carbon support compositions.

When analyzing beam-sensitive materials, the main limitation to the accuracy of the BSS analysis method that we propose is the intensity of the EDX signal achievable without inducing significant sample damage. In our case, despite the use of a high efficiency EDX system, the number of collected X-rays is low. In order to avoid the artifacts that arise when using the variance-stabilizing transformation out of its domain of

application, we have binned the data<sup>20</sup> by four in the spectral dimension from 5 eV/channel to 20 eV/channel. Given that the resolution of our EDX detector is approximately 130 eV at Mn  $K\alpha$  and that there were no overlapping X-ray lines, in our case, the increase in the number of counts per channel (and hence in the accuracy of the analysis) comes without significant resolution loss and, therefore, should not have any adverse effect in the analysis.

**Conclusions.** A blind source separation method based on PCA and ICA has been applied to the analysis of EDX spectrum images of core-shell nanoparticle clusters. The analysis has accurately determined the number of phases in the analyzed volume (core, shell, and supporting film) as well as their spectra and distribution maps. We have confirmed the accuracy of the analysis by comparing the calculated spectra from the platinum-iron core and the iron oxide shell to those obtained from these structures in isolation, and the excellent agreement suggests that BSS, therefore, can be used to accurately analyze EDX data. The use of ICA on EDX spectrum image data promises to be a powerful technique for extracting buried compositions at the nanoscale in a variety of



**Figure 6.** Comparison between raw EDX spectra extracted from a FePt bimetallic seed particle (top) and from an iron oxide particle (bottom) with IC#0 and IC#1, respectively. The FePt seed and  $\text{Fe}_3\text{O}_4$  X-ray signals are summed over several particles.

materials, and further testing on the method's applicability to different systems is now being initiated.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: rossoud@mcmaster.ca.

### Present Address

<sup>§</sup>McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4L7, Canada.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

D.R. acknowledges support from the Royal Society's Newton International Fellowship scheme. B.R.K. thanks the U.K. EPSRC for financial support (EP/J500380/1). F.d.I.P. and C.D. acknowledge funding from the ERC under grant no. 259619 PHOTO EM. P.A.M. and P.B. acknowledges financial support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 291522-3DIMAGE. P.A.M. also acknowledges financial support from the European Union's Seventh Framework Programme of the European Commission: ESTEEM2, contract number 312483.

## REFERENCES

- (1) Thomas, J. M.; Midgley, P. A. *Chem. Commun. (Cambridge, U.K.)* **2004**, 1253–1267.
- (2) Newbury, D. E. *Scanning* **2005**, *27*, 227–239.
- (3) Harrach, H. S. Von; Dona, P.; Freitag, B.; Soltan, H.; Niculae, a; Rohde, M. *J. Phys. Conf. Ser.* **2010**, *241*, 012015.
- (4) Jutten, C.; Herault, J. *Signal Process.* **1991**, *24*, 1–10.
- (5) Lee, D. D.; Seung, H. S. *Nature* **1999**, *401*, 788–791.
- (6) De la Peña, F.; Berger, M.-H.; Hocheplied, J.-F.; Dynys, F.; Stephan, O.; Walls, M. *Ultramicroscopy* **2011**, *111*, 169–176.
- (7) Duchamp, M.; Lachmann, M.; Boothroyd, C. B.; Kovács, a; Haug, F.-J.; Ballif, C.; Dunin-Borkowski, R. E. *Appl. Phys. Lett.* **2013**, *102*, 133902.
- (8) Nicoletti, O.; de la Pena, F.; Leary, R. K.; Holland, D. J.; Ducati, C.; Midgley, P. A. *Nature* **2013**, *502*, 80–84.
- (9) Dobigeon, N.; Brun, N. *Ultramicroscopy* **2012**, *120*, 25–34.
- (10) Kotula, P.; Keenan, M. *Microsc. Microanal.* **2006**, 538–544.
- (11) Lucas, G.; Burdet, P.; Cantoni, M.; Hebert, C. *Micron* **2013**, *52–53*, 49–56.
- (12) Knappett, B. R.; Abdulkina, P.; Ringe, E.; Jefferson, D. a; Lozano-Perez, S.; Rojas, T. C.; Fernández, A.; Wheatley, A. E. H. *Nanoscale* **2013**, *5*, 5765–5772.
- (13) Sun, S.; Murray, C. B.; Weller, D.; Folks, L.; Moser, A. *Science* **2000**, *287*, 1989–1993.
- (14) Kim, J.; Lee, Y.; Sun, S. *J. Am. Chem. Soc.* **2010**, *132*, 4996–4997.
- (15) Fuchigami, T.; Kawamura, R.; Kitamoto, Y.; Nakagawa, M.; Namiki, Y. *Langmuir* **2011**, *27*, 2923–2928.
- (16) HyperSpy Home Page. [www.hyperspy.org](http://www.hyperspy.org) (accessed Feb 2015).
- (17) Bingham, E.; Hyvärinen, A. *Int. J. Neural Syst.* **2000**, *10*, 1–8.
- (18) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (19) Keenan, M. R.; Kotula, P. G. *Surf. Interface Anal.* **2004**, *36*, 203–212.
- (20) Kotula, P.; Keenan, M. *Microsc. Microanal.* **2003**, *9*, 2002–2003.
- (21) Hyvärinen, A.; Oja, E. *Neural Networks* **2000**, *13*, 411–430.
- (22) Jones, T. R.; Kang, I. H.; Wheeler, D. B.; Lindquist, R. a; Papallo, A.; Sabatini, D. M.; Golland, P.; Carpenter, A. E. *BMC Bioinformatics* **2008**, *9*, 482.
- (23) Soille, P.; Ansoult, M. *Signal Process.* **1990**, *20*, 171–182.
- (24) Williams, D. B.; Carter, C. B. *Transmission electron microscopy: a textbook for materials science*, 2nd ed.; Springer: New York, 2009.