

## Sparsity, Parsimony and Data Reduction – Applications across Multi-Dimensional Electron Microscopy

Paul Midgley<sup>1</sup>

<sup>1</sup> Department of Materials Science and Metallurgy, University of Cambridge, Cambridge, UK.

The remarkable advances in electron microscopy hardware in recent years include the development of fast and efficient spectrometers and fast, almost noise-free, cameras. The vast amount of data that can now be acquired brings with it the need for new software and new approaches that enable fast and efficient data processing. One area that has benefitted enormously from both the developments in hardware and software is ‘multi-dimensional’ electron microscopy, or MDEM, a collection of techniques in which 3D, 4D and 6D data sets may be acquired [1]. MDEM techniques allow not just the structure, but also the chemistry and crystallography to be investigated at the atomic, or nanoscale, with temporal resolution and in all three dimensions. These techniques, by their very nature, produce very large data sets that require new methods to enable the key salient information to be extracted [2]. These will be reviewed in this paper.

Figure 1 shows examples of where electron energy loss spectroscopy has been used in conjunction with electron tomography to elucidate the localised surface plasmon resonances (LSPRs) of silver nanostructures. Figure 1a shows a single EELS spectrum acquired from a silver nanocube as part of a tilt series of spectrum-images. The fit to the spectrum was undertaken using machine learning and in particular a factor analysis called non-negative matrix factorization (NMF). The series of spectra recorded (ca. 100k) can be decomposed into a linear sum of eight component spectra, five of which are associated with LSPRs, shown in color. The loading at each pixel associated with the five components can be visualized as an image and used as input for a tomographic reconstruction assuming each image is a projection of some 3D physical quantity, in this case, to a good approximation, the squared modulus of a 3D surface potential associated with each LSPR. Each tilt series was composed of only five images and as such high reconstruction fidelity was maintained only through the application of symmetry and sparsity constraints and the use of a compressed sensing approach to the reconstruction [3]. In many cases, however, a simple projection approximation is not possible and then a model-fit approach is required. In Figure 1b that approach led to the reconstruction of the surface charge densities of the LSPRs of a silver bipyramid. This was achieved through a least squares fit comparing experimental NMF component spectra with computed (quasi-static) spectra and with an extra regularization constraint to promote parsimonious solutions [4].

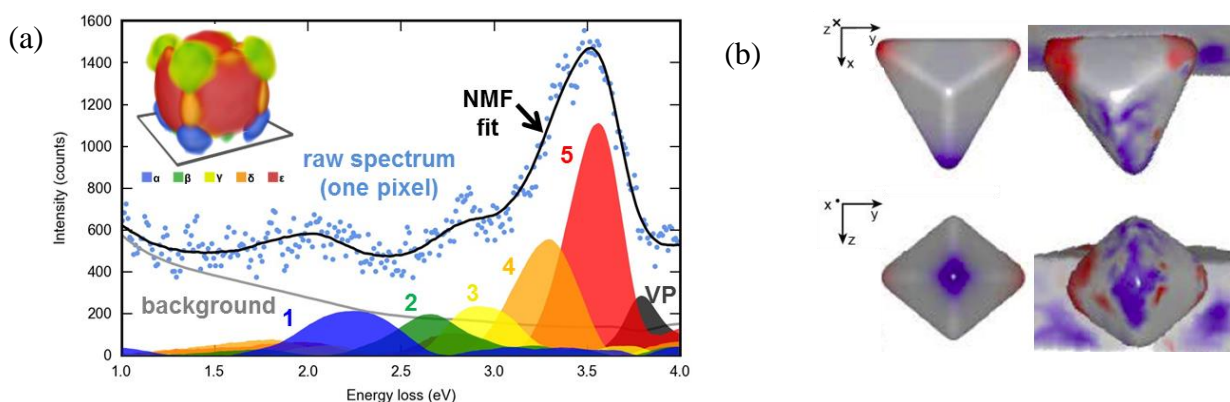
Whilst MDEM enables materials to be interrogated in fine detail the penalty is almost always the need for a high total electron dose. Many materials, however, are beam sensitive and so there is a requirement to acquire *just enough* data to answer the problem. In structural biology, similar ideas are implemented through dose fractionation in which the total dose is divided amongst the individual TEM images. In materials science, where STEM is often the technique of choice, we need to develop similar ideas but through scanning the beam we have alternative dose fractionation methodologies from which to choose [5]. Figure 2 shows longitudinal and axial slices through a 3D reconstruction of a Co phthalocyanine / ZnO nanotube. The full reconstruction (100%, 2°) is compared with one that has been undertaken with images with only 30% of the pixel number of the full reconstruction and where two out of the three images in the tilt series is discarded (30%, 6°). Overall this is an order of magnitude reduction in data

but through the use of compressed sensing reconstruction and in-painting algorithms, applying constraints in image and gradient space, the loss in resolution was only just over 25%.

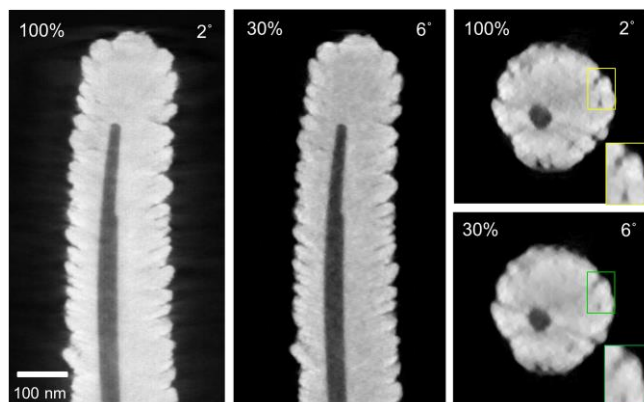
The paper will conclude by considering possible future developments, efficient ways to record data and the desire for on-the-fly analysis. Such analysis would not only provide near-instant feedback but would also guide the microscope in an automated and unsupervised fashion to probe more profitable areas and/or acquire optimized signals [6].

#### References:

- [1] P.A. Midgley and J.M. Thomas *Angew. Chem. Int. Ed.* **53** (2014), p. 8614.
- [2] T. Furnival *et al*, *Ultramicroscopy* 10.1016/j.ultramic.2016.05.005i
- [3] O. Nicoletti *et al*, *Nature* **502** (2013), p. 80.
- [4] S.M. Collins *et al*, *ACS Photonics* **2** (2015), p. 1628.
- [5] Z. Saghi *et al*, *Adv. Struct. Chem. Imag.* **1** (2015), p. 7.
- [6] The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement 291522–3DIMAGE, as well as from the European Union Seventh Framework Programme under Grant Agreement 312483-ESTEEM2 (Integrated Infrastructure Initiative -I3).



**Figure 1.** (a) Low loss EELS spectrum, NMF fit and 3D reconstruction of LSPRs excited in a silver nanocube [3]. (b) Mutually perpendicular views of a 3D tomogram of the reconstructed surface charge density of one dipole mode of a silver bipyramid on  $\text{MoO}_3$  (right) and a simulation of an isolated bipyramid for comparison (left) [4].



**Figure 2.** Longitudinal and axial slices of a 3D compressed sensing reconstruction of a CoPc/ZnO nanowire. The full reconstruction (100%,  $2^\circ$ ) is compared with one that has been undertaken with images with only 30% of the data of the full reconstruction and where two out of the three images in the tilt series is discarded [5].