Multidimensional Data Analysis in Electron Microscopy
A NanoDTC Practical Introduction

Jędrzej Morzy
jdkm2@cam.ac.uk
Electron Microscopy Group
Department of Materials Science and Metallurgy
University of Cambridge

October 2020

This practical has been developed over a number of years within the Electron Microscopy Group. In particular the contributions of Duncan Johnstone, Caterina Ducati, Giorgio Divitini, Mathias Kobylko and Francisco de la Pena have been significant.

During this practical, you will learn some Python based methods of analysis of multidimensional data sets. In particular we will be looking at data sets from electron microscopy experiments, such as EDX/EELS (energy dispersive X-ray spectroscopy and electron energy loss spectroscopy) maps as well as scanning electron diffraction maps. Before the practical, I ask you to read this handout, in which some of the concepts of multidimensional electron microscopy and data analysis are introduced. While reading, I would like you to think about the questions included - they may prove useful during the practical. I tried to reduce the number of references to a few key ones, so I recommend reading the ones that have made it into bibliography (except the last one - book by Williams and Carter, which serves more of a further reading purpose). Near the end of this handout, you will find instructions as to how to install the software required for this practical. It will be mostly based on a Python library called Hyperspy - a tool dedicated to multidimensional data analysis (note: it is definitely not limited to electron microscopy, although focused on it). Again, I have tried limiting the amount of things you will need installed and they should be relatively easy to remove, if after the practical you decide that you never want to interact with them again. In case you have any questions or problems with installation, please send me an e-mail, I will be more than happy to help.

1 Introduction

Modern (scanning) transmission electron microscope ((S)TEM) is one of the most versatile and powerful tools for characterisation of materials and phenomena at remarkably small length scales. Its capabilities can include, but are not limited to: nano- or even atomic-scale imaging, acquiring information in reciprocal space (through electron diffraction) and with the use of appropriate detectors - chemical information (also at nano- or even atomic scale!) such as composition, valence states or bonding through electron energy-loss spectroscopy (EELS) and energy-dispersive X-ray spectrometry (EDX).[1] On top of that, it has 3D imaging[2,3] and 3D chemical mapping capabilities.[4]

Some of those signal types can be acquired simultaneously, which combined with the possibility of monitoring changes in those signals over time (with temporal resolution on the order of μs - s) creates the possibility of creating ***multidimensional data sets***. In most cases, in electron microscopy multiple dimensions mean multiple spatial (up to 3) and spectral dimensions. A diagram showing different types of multidimensional EM data sets is shown in Figure 1.

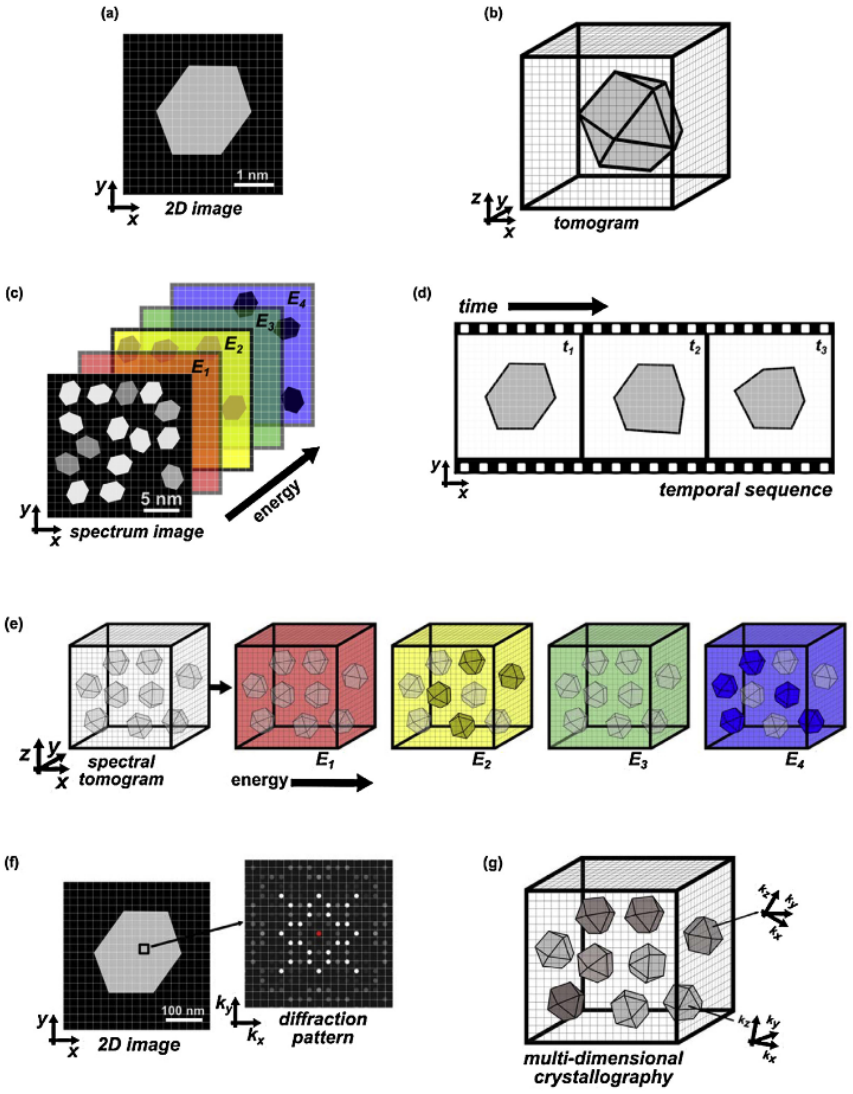

Figure 1: Forms of multidimensional electron microscopy: (a) 2D imaging 3D structures in projection; (b) 3D imaging typically achieved via tomography; (c) spectrum-imaging in which a spectrum (energy dimension) is acquired at every position in an image; (d) temporal image sequences; (e) spectral tomography in which a spectrum is associated with every voxel in 3D; (f) crystallographic mapping in which a diffraction pattern is recorded at each real space pixel; (g) crystallographic tomography in which three dimensional reciprocal space information is associated with every voxel. Reproduced from Thomas et al.[5]

With such complex data sets, appropriate software tools had to be developed to transform unintuitive, raw multidimensional data into figures, graphs and plots that will be understandable to a wider audience. This practical aims to show some of the ways of accessing, analysing and visualising data from multidimensional electron microscopy.

Within this practical we will use open-source Python-based software to investigate and analyse:

- Morphology
- Chemical composition (EDX/EELS map)
- Nanoscale crystallography (scanning electron diffraction map)

It is worth noting that general analysis methodology of those multidimensional data sets is not limited to electron microscopy. It can be used for other types of experiments yielding multidimensional data such as Raman spectral maps, cathodoluminescence maps, AFM indentation maps or any spectral time series. So, even if you are not planning to use electron microscopy in your research career, you may find it useful to adapt some of the concepts introduced during this practical to your needs, which should be relatively straight-forward thanks to the versatility of Python and its libraries.

2 Morphology

TEM and STEM are incredibly versatile and robust approaches to imaging the morphology of materials at the nanoscale. Images produced are relatively intuitive to understand, can be acquired with high speeds for high throughput or *in situ* experiments. Of course, (S)TEM is not the ultimate imaging approach that you might use in every case - there are some specialised optical, X-ray and scanning probe microscopy techniques that allow you to reach nanometre scale resolutions (in some cases even better than for (S)TEM), but every one of them, including (S)TEM has its own advantages and limitations.

Questions to think about: What are the limitations of TEM and STEM? What types of materials are the most challenging to image with it?

A typical problem of TEM imaging is that it presents us with 2D images of 3D specimens viewed in transmission mode. This may be unintuitive to imagine as our brains are used to 3D analysis of objects based on reflected and not transmitted light images. This problem can be illustrated through the use of the photo presented in Fig. 2. What we see is a projection of two rhinos that appear to be merged into one two-headed beast. With real objects like animals, we have a clear expectation (prior knowledge) of what we should be seeing, so we understand that this is just an optical illusion, but with TEM images things may not be that simple and often transmission images might be misleading.

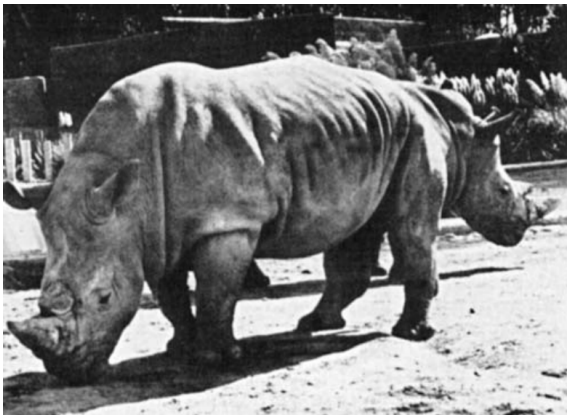

Figure 2: Photograph of two rhinos. Reproduced from Williams and Carter[6]

This limitation (also called projection-limitation) leads to the fact that all information acquired in the TEM is averaged through the thickness of the sample. In other words, TEM does not have depth sensitivity. On the other hand, this enables the use of a technique called **electron tomography** in which the sample is tilted and a series of 2D projection images is taken at each tilt angle and then reconstructed into a 3D model. It should also be noted that this technique can be applied to sequences of spectral maps as well.[4] In Figure 3 you can see a schematic representation of how electron tomography and reconstruction works.

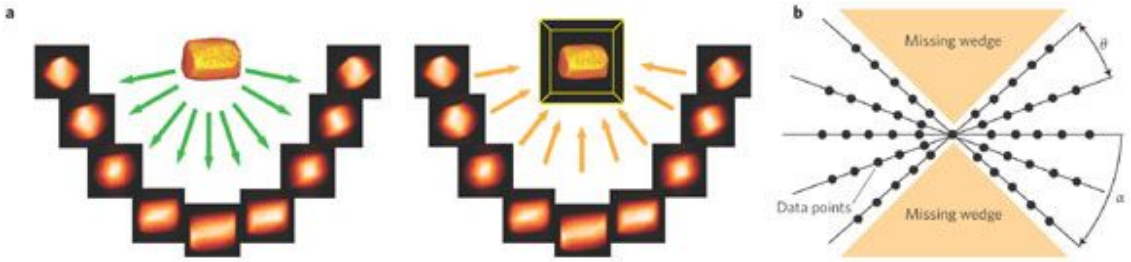

Figure 3: (a) Illustration of a series of 2D images projected from a 3D sample (left) and the back-projection of these images into 3D model (right). (b) Representation in Fourier space of the ensemble of projections and the missing wedge of information stemming from experimental constraints. Reproduced from Midgley et al. [x]

In general, TEM imaging modes can be sub-divided into bright field and dark field imaging (if you are unsure what those terms mean, I suggest going through NE.01 lecture notes), which means that TEM allows the user to access different types contrast mechanics, which in turn allows to extract different types of information from the specimen. Furthermore, using parallel beam illumination (TEM) and convergent probe scanning illumination (STEM) will yield yet again different contrast mechanisms, which in turn means, that by combining some of these techniques and understanding the image generation mechanisms, (S)TEM imaging can be much more informative than just a way to generate pretty pictures of the sample.

Question to think about: What is the different types information that can be gained from (S)TEM BF and DF imaging techniques? Is projection limitation also a problem for (S)TEM spectroscopy? If so, in what way?

3 Atomic Structure and Crystallography

In the (S)TEM it is possible to achieve atomic resolution images, videos (Under this link you can find a short note and a video on imaging crystal growth layer by layer with atomic resolution) and 3D reconstructions (A Nature video on 3D tomography reconstruction of a platinum particle). The possibility of imaging in real space at atomic resolution allows (S)TEM to probe local atomic structure and its defects (dislocations, point defects etc.). However, in many cases, diffraction approach can be more useful. Through electron diffraction it is possible to acquire information on the crystallography or, in other words, more statistically-averaged atomic structure information. This is undoubtedly a very powerful tool for crystalline materials. Thanks to the periodicity of crystals (let's forget about quasicrystals in this practical, as they are more problematic), definition of a single unit cell is enough to describe atomic structure throughout the crystal. Much insight into the structure of non-crystalline materials can also be gained materials using diffraction mode in (S)TEM, but this is beyond the scope of this practical.

Within the practical we will briefly explore the multidimensional data sets acquired using scanning electron diffraction (SED). In this technique, a small electron probe is scanned over the sample and electron diffraction patterns coming from the sample are registered at each probe position. The diffraction patterns include information about the atomic structure of an averaged crystallographic structure from a region of real space dictated by the size of the electron probe. Said size can be easily controlled and changed from sub-nanometre to micrometre level, thanks to the flexibility of electron optics in (S)TEM. By making the probe relatively small and scanning over an area, spatially resolved diffraction data can be collected.

Questions to think about: Is atomic resolution in 2D (S)TEM images really atomic resolution? What does the resolution of 0.5Å mean? What are we actually seeing at the atomic positions? What about in tomography reconstructions?

4 Spectral mapping and chemistry

Spectral maps can be obtained in STEM by scanning the electron beam across the specimen (in the same manner as in SED) and acquiring a spectrum at each pixel of the scan. Two typical spectrum types include energy dispersive X-ray spectroscopy (EDX) and electron energy-loss spectroscopy (EELS). Both yield chemical information, although in different ways. If you need a refresher on how they work, I again suggest going back to NE.01 lecture notes.

Each data set is then three-dimensional, with two real space dimensions and one spectral dimension (every pixel in a 2D map also includes a energy spectrum adding the third dimension) forming a *spectrum image*. As with conventional images, it is possible to do a tomographic tilt series and reconstruct a 3D models + 1D spectra resulting in 4D *spectrum tomogram*. There is of course, the possibility of adding further dimensions such as time or a second spectrum, but creating tomograms is quite time consuming, so the most common type is 2 spatial dimensions + 1-2 spectral dimensions.

Question to think about: What are the main limitations of these techniques in terms of chemical analysis? Are the limitations correlated with limitations of (S)TEM imaging in general?

A typical workflow for spectrum images involves three main steps. First, the spectra need to be pre-processed through background removal, denoising, cosmic-ray removal

etc. Secondly, resulting processed data has to be analysed. For example, it could involve peak identification, fitting and quantification through eg. Cliff-Lorimer method for EDX analysis. The last, crucial step is visualising the analysed data. As the data is multidimensional, this is quite often the most challenging part.

Question to think about: What could be some examples of visualising certain aspects of a 4D data set consisting of 2 spatial dimensions, a spectral dimension and temporal one?

Quite often, it is valuable to separate the spectrum image into a set of representative spectra coming from the sample due to different chemical compositions in various regions. Then, maps (loadings) of where these representative spectra (components) are present in the data set can be created. This is generally known as the *blind source separation* problem. Machine learning methods can and have been adopted to help with this kind of analysis in electron microscopy. In HyperSpy (Python library developed largely in Electron Microscopy Group, Cambridge) there are many different algorithms implemented, including principal component analysis, independent component analysis or non-negative matrix factorisation. During the practical we will use some of them to de-noise the data sets and decompose them to separate contributions from different material phases. We will use those algorithms without delving too far into the maths behind them, but purely from an experimental point of view. If you are interested in looking into this methodology further, I would suggest looking at HyperSpy documentation, which will contain brief explanations of different algorithms as well as references for further reading.

5 Required software

1. ImageJ
   is a very useful programme for image processing and analysis. It has some powerful functionalities including thresholding, cropping, measuring distances, areas, counting particles and a wealth of plugins for various data analysis tasks. Alternatively, you can use Fiji, which is ImageJ but with more plugins installed from the start.

2. HyperSpy
   HyperSpy is a Python library for multidimensional data analysis with very good set of functions specific for electron microscopy. I suggest installing it through Anaconda by entering the following command into Anaconda prompt:

   ```
   conda install hyperspy -c conda-forge
   ```

3. pyXem (optional)
   This is a library for multidimensional analysis of electron diffraction data sets. You **do not** need to install this for the practical, but we will explore some of the capabilities briefly during the practical (time permitting).

   If you have any issues with installation of the software, let me know, as these things can be a little tricky if the creators update their codes without making sure that the installation is compatible with other libraries.

References

[1] G. Divitini, S. Cacovich, F. Matteocci, L. Cinà, A. Di Carlo, and C. Ducati, "In situ observation of heat-induced degradation of perovskite solar cells," *Nature Energy*, vol. 1, no. 2, p. 15 012, 2016, ISSN: 20587546. DOI: 10.1038/NENERGY.2015.12. [Online]. Available: http://www.nature.com/articles/nenergy201512.

[2] P. A. Midgley and R. E. Dunin-Borkowski, "Electron tomography and holography in materials science," *Nature Materials*, vol. 8, no. 4, pp. 271–280, 2009, ISSN: 14764660. DOI: 10.1038/nmat2406. eprint: arXiv:1606.02938.

[3] Z. Saghi and P. A. Midgley, "Electron Tomography in the (S)TEM: From Nanoscale Morphological Analysis to 3D Atomic Imaging," *Annual Review of Materials Research*, vol. 42, no. 1, pp. 59–79, 2012, ISSN: 1531-7331. DOI: 10.1146/annurev-matsci-070511-155019. [Online]. Available: http://www.annualreviews.org/doi/10.1146/annurev-matsci-070511-155019.

[4] G. Haberfehlner, A. Orthacker, M. Albu, J. Li, and G. Kothleitner, "Nanoscale voxel spectroscopy by simultaneous EELS and EDS tomography," *Nanoscale*, vol. 6, no. 23, pp. 14 563–14 569, 2014, ISSN: 20403372. DOI: 10.1039/c4nr04553j.

[5] J. M. Thomas, R. K. Leary, A. S. Eggeman, and P. A. Midgley, "The rapidly changing face of electron microscopy," *Chemical Physics Letters*, vol. 631-632, pp. 103–113, 2015, ISSN: 00092614. DOI: 10.1016/j.cplett.2015.04.048. [Online]. Available: http://dx.doi.org/10.1016/j.cplett.2015.04.048.

[6] D. Williams and B. Carter, *Transmission Electron Microscopy*. Springer, 2009, ISBN: ISBN 978-0-387-76500-6. DOI: 10.1007/978-3-319-26651-0.