1. INTRODUCTION

The promise of nanotechnology [1] can only be fully realised if characterisation techniques are available to study structures and devices at the nanometre scale. In particular, many of the proposed nano-devices are truly three-dimensional in their design and high spatial resolution microscopy is needed to assess a device in all three dimensions. Transmission electron microscopy (TEM), be it in the form of high resolution electron microscopy (HREM) [2] or scanning transmission electron microscopy (STEM) [3] can provide images with extremely high spatial resolution (sub Å) in two dimensions. However, all TEM images are formed by the propagation of the electron beam through the specimen and as such they are two-dimensional projections of a three-dimensional object. Often such projections (particularly of cross-sectional specimens) have been sufficient to determine the structure of simple devices [4]. Where the third dimension has been recognised as important, for example in the study of dislocation networks, then ‘stereo pairs’ can be used in which a pair of images is recorded at ~10° tilt to mimic the eyes’ angular separation. By viewing the two images simultaneously an illusion of three dimensions can be achieved but in reality provides very little 3D information [5].

In structural biology, there has been a need to image highly complex 3D structures at the nanoscale for many years. 3D TEM techniques have been developed to study protein structures [6], viruses [7], ribosomes [8] and larger cellular structures, such as the mitochondria [9]. Three approaches have been employed. Firstly, if the protein
structure can be crystallised then standard electron crystallography techniques (involving HREM and electron diffraction) can be used to solve the crystal structure and from an electron density map retrieve the asymmetric unit that describes the unique protein structure [10]. If the crystal is sufficiently large, X-ray methods are used routinely, if not the electron microscope must be employed. Secondly, if the structure of interest is repeated many times on a specimen grid (as is the case for many biological structures, e.g. viruses), then a single TEM image will contain a large number of sub-images each of which, in general, a projection of the structure, for example, at a different orientation. By determining the exact orientation of each sub-image, the three dimensional structure can be reconstructed [11]. The third approach is employed for unique cellular structures and involves recording a series of images (projections) of the same object at successive tilts and then reconstructing the object from the series [12]. For the purpose of this chapter and simply to aid clarity, I will describe the three methods as (i) electron crystallography, (ii) single particle analysis and (iii) electron tomography, respectively. This division is somewhat artificial and readers should bear in mind that each method is in essence an electron tomographic technique, all are based on the ‘Radon transform’ [13] and the ‘projection requirement’ [14] (described in detail later) and all use comparable reconstruction algorithms.

Similar tomographic techniques exist in the physical sciences. In materials science and engineering, X-ray tomography has been used successfully to reconstruct relatively large three-dimensional structures, such as metallic foams [15] or to probe the stress in engineering structures [16]. However, in general the wavelength of X-rays, coupled with the relatively poor quality of X-ray optics, is such that a resolution of \( \sim 2 \) microns is normally the best achievable. As a caveat to this, however, are recent claims that a resolution of a few 10’s nm is now achievable using soft X-rays from synchrotron sources and Fresnel lenses (zone plates) [17]. Such an improvement in resolution is very exciting and promises great things. At the opposite end of the resolution scale, the atom probe field ion microscope (APFIM), designed around a time-of-flight spectrometer and a position sensitive detector, is able to reconstruct three-dimensional maps of atom positions and determine each atomic species [18]. Atom probe tomography is the only tomographic technique that allows single atom counting of a three-dimensional structure. Such remarkable sensitivity is however also a limitation in that it is very time consuming to examine large objects using this technique—for example a 100 nm cube of crystalline silicon contains \( 5 \times 10^7 \) atoms! More problematic is the requirement for the sample to be conducting and withstand high field stresses exerted at the tip of the needle-shaped sample needed for the APFIM technique [19].

In nanotechnology, the structures and devices designed to take advantage of the mechanical, physical or chemical changes that occur at these length scales will have features at around the nanometre level but whose overall size may be tens or hundreds of nanometres. Such nanoscale design is of course already underway to a large extent in the microelectronics industry where the three-dimensionality of, for example, the metallisation or the dopant profiling becomes increasingly critical to the performance of the device [20]. In the magnetic recording industry, the magnetic ‘bits’ are becoming ever smaller and a need is developing to examine the composition and magnetic
microstructure in three dimensions [21]. The latter requires accurate measurement of the magnetic induction and this can be achieved with a 3D form of electron holography [22]. Indeed, three-dimensional analysis will also become increasingly important not only for functional materials but also for nanoscale structural and engineering materials, such as ultra-fine cerments [23]. In the catalysis industry, heterogeneous catalysts [24] are now designed with nanometre-sized active particles distributed in three dimensions on or within a porous support structure—the tomography of such catalysts will be discussed later in the chapter.

Thus a microscope technique is needed that will allow relatively large structures and devices to be studied (say up to 500 nm in diameter) but with a 3D resolution of \(\sim 1\) nm to allow the intricate detail of the internal nanostructure to be unravelled. Such requirements are remarkably similar to those demanded by structural biologists studying cellular structures and so it is natural to turn to electron tomography, used so successfully in the life sciences, as a means by which the 3D structure of nanoscale devices can be elucidated.

2. TOMOGRAPHY

2.1. A History of Tomography

Before describing the technique in detail, it is worth spending some time reviewing the birth and subsequent development of tomography, and of electron tomography in particular. The need to obtain ‘structures’ using data of lower dimensionality is present in many different fields of physical and life sciences. It was in the field of Astronomy that in 1956 Bracewell [25] proposed a method of reconstructing a 2D map of solar emission from a series of 1D ‘fan beam’ profiles measured by a radio telescope. This pioneering work covered the mathematical formulation of projection and reconstruction but despite its clear potential, this work had little impact beyond its immediate field. However, in 1963, interest in tomography was rekindled by its possible use in medical sciences [26]. This led to the development of the X-ray computerised tomography (CT) scanner [27], known more commonly as the CAT-scan (computer assisted tomography or computerised axial tomography). This remarkably successful technique is undoubtedly the most well known application of 3D tomographic reconstruction and its pioneers, Cormack and Hounsfield, were awarded jointly the Nobel Prize for Medicine in 1979. The success of the CAT-scan was mirrored by the development of similar techniques such as positron emission tomography (PET) [28], ultrasound CT [29] and zeugmatography (reconstruction from NMR imaging) [30]. Outside the medical field, tomography was applied in many other disciplines to allow, for example, 3D stress analysis [16], geophysical mapping [31] and non-destructive testing [32, 33].

Interest in three-dimensional reconstruction using electron microscopy started with the publication of three papers in 1968. The first was by de Rosier and Klug [34] in which the structure of a biological macromolecule was determined whose helical symmetry allowed a full reconstruction to be made from a single projection (micrograph). The Fourier reconstruction methods used in this paper were akin to those developed for the determination of atomic structures by X-ray crystallography [35].
While symmetry was key to these results, it was suggested in the second of these papers, by Hoppe [36], that, given a sufficient number of projections, it should be possible to reconstruct fully asymmetrical systems, i.e with no symmetry imposed. The last of the three early papers, by Hart [37], demonstrated a method of improving the signal to noise ratio in images using an ‘average’ re-projected image calculated from a tilt series of micrographs, a technique known as a polytropic montage. Used initially as a means to combat the weak contrast in biological specimens, Hart acknowledged the 3D information generated by such an approach without extending this to the possibility of full 3D reconstruction. Shortly afterwards, a number of theoretical papers were published discussing the theoretical limits of Fourier techniques [38], approaches to real space reconstruction [39] and the use of iterative reconstruction routines [40, 41].

Until recently, the advance of electron tomography was impeded by a number of technical difficulties, in particular, the poor performance of goniometers (especially at high tilt), the length of time required to acquire a series of images and the lack of computer power for image processing and reconstruction. The improvement in electron microscope design coupled with the vast improvement in computer performance has overcome all these. However, the time taken to acquire a series is still a major problem in the life sciences as specimens damage rapidly in the electron beam [42]. To increase the longevity of samples, they are often examined at liquid helium temperatures using cryo-microscopy and at high voltages, both of which reduce the effects of inelastic scattering and subsequent damage [43]. For electron tomography, as opposed to single particle analysis, many specimens are still examined in resin or plastic sections and stained to enhance contrast [44] and such specimens are relatively robust in the beam.

2.2. The Radon Transform

Although the first practical formulation of tomography was achieved in 1956 [25], it was Radon who first outlined the mathematical principles underlying the technique in 1917 [13]. In his paper a transform, known now as the Radon transform, $R$, is defined as the mapping of a function $f(x, y)$, describing a real space object $D$, by the projection, or line integral, through $f$ along all possible lines $L$ with unit length $ds$ so that,

$$ Rf = \int_L f(x, y) \, ds $$

The geometry of the transform is illustrated in Figure 1. A discrete sampling of the Radon transform is geometrically equivalent to the sampling of an experimental object by a projection or some form of transmitted signal. As such, the structure of an object $f(x, y)$ can be reconstructed from projections $Rf$ by using the inverse Radon transform. All reconstruction algorithms are approximations of this inverse transform.

The Radon transform operation converts real space data into ‘Radon space’ $(l, \theta)$, where $l$ is the line perpendicular to the projection direction and $\theta$ is the angle of the projection. A point in real space ($x = r \cos \phi$, $y = r \sin \phi$) is a line in Radon space $(l, \theta)$ in which $l = r \cos(\theta - \phi)$. A single projection of the object, a discrete sampling
Figure 1. The Radon transform $R$ can be visualised as the integration through a body $D$ in real space $f(x, y)$ along all possible line integrals $L$, with its normal at an angle $\theta$ to the horizontal.

of the Radon transform, is a line at constant $\theta$ in Radon space. A series of projections at different angles will therefore sample Radon space and given a sufficient number of projections, an inverse Radon transform of this space should reconstruct the object. In practice the sampling of $(l, \theta)$ will be limited and any inversion will be imperfect. The goal of any reconstruction then becomes achieving the ‘best’ reconstruction of the object given the limited experimental data.

2.3. The Central Slice Theorem and Fourier Space Reconstruction

The relationship between real space and Radon space gives an understanding of the nature of a projection and its relationship with the original object. In addition, reconstruction from projections is aided by an understanding of the relationship between a projection in real space and Fourier space. The ‘central slice theorem’ or the ‘projection-slice theorem’ states that a projection of an object at a given angle in real space is a central section through the Fourier transform of that object. The relationship between the Fourier transform $F$ and the Radon transform $R$, may be summarised in operator form as:

$$F_2 f = F_1 R f = F_1 \hat{f}$$

where $\hat{f}$ is the full Radon transform of object $f$. More detail regarding the nature of the transforms and their inter-relationship can be found in [33] and for brevity will not be explored further here.

For readers familiar with electron diffraction, the central slice theorem is of course exactly that known as the ‘projection approximation’ relating the intensity of Zero Order Laue Zone (ZOLZ) reflections to the crystal potential projected parallel to the zone axis.
The shape of most objects will be described only partially by the frequencies in one section but by taking multiple images (projections) at different angles many sections will be sampled in Fourier space. This will describe the Fourier transform of an object in many directions, increasing the information available in the 3D Fourier space of the object. In principle a sufficiently large number of projections taken over all angles will yield a complete description of the object.

Tomographic reconstruction is possible from an inverse Fourier transform of the superposition of a set of Fourier transformed projections: an approach known as direct Fourier reconstruction [39] used for the first tomographic reconstruction from electron micrographs [33]. Importantly, it provides a convenient and a logical basis to describe the effects of sampling deficiencies in the original dataset. If projections are missing from within an angular range, brought about by, for example, a limit on the maximum tilt angle, then Fourier space is under-sampled in those directions and as a consequence the back transform of the object will be degraded in the direction of this missing information.

The experimental data is always sampled at discrete angles leaving (often, regular) gaps in Fourier space. An inverse Fourier transform requires a continuous function and so radial interpolation is required to fill the gaps in Fourier space [37]; the quality of the reconstruction is greatly affected by the type of interpolation method used [45]. Although elegant, Fourier reconstruction methods have the disadvantage of being computationally intensive and difficult to implement for electron tomography. This is not the case for single particle analysis where Fourier methods are still the norm [46]. For electron tomography of unique structures, Fourier methods have been superseded by faster and easier to implement real space backprojection methods [47].

### 2.4. Real Space Reconstruction using Backprojection

The method of backprojection is based on simple reasoning: a point in space may be described uniquely by any three ‘rays’ passing through that point—the method of triangulation. With an increase in the object’s complexity, more ‘rays’ are required to yield a unique description. Thus a projection of an object is an inverse of such a ‘ray’, and will describe some of the complexity of that object. Inverting the projection, smearing out the projection back into an object space at the angle of the original projection, generates a ‘ray’ that will describe uniquely an object in the projection direction: a method known as backprojection. Using a sufficient number of projections, from different angles, the superposition of all the backprojected ‘rays’ will return the original object: a reconstruction technique known as direct backprojection. [36, 38, 48], see Figure 2.

In principle, it is possible to reconstruct the object using backprojection in an way that is analogous to the experiment that generated the projection, i.e. by rotating the reconstruction space to the original projection angles and summing the projections along a constant reconstruction axis. However poor sampling when rotating the reconstruction will lead to artefacts. Instead the relationship between Radon space and real space, described earlier can be used to provide an algorithm that is less prone to
Figure 2. A schematic of tomographic reconstruction using the backprojection method. In (a) a series of images are recorded at successive tilts. These images are back-projected in (b) along their original tilt directions into a 3D object space. The overlap of all the back-projections will define the reconstructed object.
II. Electron Microscopy

Figure 3. An illustration of the non-uniform sampling of Fourier space, brought about by the acquisition of a tilt series. The relatively small number of data points at high frequencies results in a blurred reconstruction. The angular increment between projections is $\theta$ and the maximum tilt angle, $\alpha$.

error. Each projection is a sample of Radon space $(l, \theta)$ and a reconstruction should return an object in real space $f(r, \phi)$. The intensity of a real-space pixel $p$ from a single projection, at angle $\theta$, can be found by virtue of the relationship that such points exist within Radon space at the intersection of the line $l = r \cos(\theta - \phi)$ and $\theta$. This can be represented by [49]:

$$[Rp](r, \phi) = \int_0^\pi p(r \cos(\theta - \phi), \theta) d\theta$$

(3)

which may be evaluated using a Riemann sum [49]. For all real datasets, values for every solution of $p(r \cos(\theta - \phi), \theta)$ do not exist because of the limited sampling of Radon space and an interpolation is required to determine the unknown values. The quality of the backprojection will be dependent on what form of interpolation is applied, although typically nearest neighbour or bilinear interpolation is used [50]. The detailed algorithms behind this method can be found in the books by Deans [33] and Herman [49].

Reconstructions by direct backprojection are always blurred with an enhancement of low frequencies and fine spatial detail reconstructed poorly. This is an effect of the uneven sampling of spatial frequencies in the ensemble of original projections. Described more easily in two dimensions, as illustrated in Figure 3, each of the acquired projections is a line intersecting the centre of Fourier space. Assuming a regular sampling of Fourier space in each projection this results in a proportionately greater sampling density near the centre of Fourier space compared with the periphery. This leads to an undersampling of the high spatial frequencies of the object and a ‘blurred’ reconstruction—see the comparison later in Figure 6(a).

Since the sampling is directly related to the position in Fourier space and the number of acquired projections, it is straightforward to correct in Fourier space using a weighting filter (a radially linear function in Fourier space, zero at the centre and a maximum at the edge). In order to avoid enhancing noise at high frequencies the filter is apodised using a Gaussian function or similar so that the Fourier transform has zero value at the Nyquist frequency [51]. This weighting filter has the effect of
rebalancing the frequency distribution in Fourier space and minimising the blurring in real space; this improved reconstruction approach is known as weighted backprojection [52]. Weighted backprojection is now the most widely used reconstruction technique for electron tomography as it is simple to implement on large datasets and can be used for irregular sampling geometries [53].

Reconstructions using the backprojection method will always be ‘imperfect’ because of the limited sampling. In addition a poor reconstruction can be made worse if the number of acquired projections is small or the signal to noise ratio (SNR) is low in the original projections. However by noting that each projection is a ‘perfect’ reference the quality of the reconstruction can be improved. If the (imperfect) reconstruction is re-projected back along the original projection angles the re-projections, in general, will not be identical to the original projections (images). The difference between them will be characteristic of the deficiency of the reconstruction from the limited dataset. This difference can be backprojected into reconstruction space, generating a ‘difference’ reconstruction, which can then be used to modify the original reconstruction in order to correct the imperfections in the backprojection. This constrains the reconstruction to agree with the original projections. As the ‘difference’ is also being backprojected a single operation will not correct fully the reconstruction and the comparison operation must be repeated iteratively until a ‘best’ solution is reached [54, 55]. Such iterative methods were first developed for electron tomography in the 1970’s and they have since been recognised to be part of a family of solutions by projection onto convex sets (POCS) [56], a more generalised form of the Gerchberg-Saxton algorithm [57, 58]. There have also been attempts to use maximum entropy techniques directly. These attempt to find the simplest (least complex) reconstruction taking into account the known projections, the noise in the data, the sampling artefacts in the reconstruction and the contrast limits of the original projection [58–62].

3. TOMOGRAPHY IN THE ELECTRON MICROSCOPE

3.1. Acquisition

The rest of the chapter will concentrate in particular on electron tomography, as defined previously, and in this section the problems that arise when acquiring a tomographic series of images in the TEM are discussed. For unique (non-repeating) structures, a series of images (projections) must be acquired at angular increments by tilting the specimen using the microscope goniometer. ‘Single-axis tilting’ is the technique normally chosen for electron tomography. ‘Single-axis tilting’ is the technique normally chosen for electron tomography. The specimen is tilted about the eucentric axis of the specimen holder rod, from one extreme of the tilt range to the other. By recording images at each tilt, Fourier space is sampled in planes whose normals are perpendicular to the tilt axis.

With the capability in modern instruments of controlling the goniometer using a computer, it is now possible to fully automate the acquisition process [63]. The small movements in the position of the sample as it is tilted through the series can be minimised if the goniometer is pre-calibrated, that is the mechanical movements as a function of tilt are known and subsequently corrected. The reproducibility of the specimen position in modern goniometers is such that calibration need be done only
infrequently [64]. To automate the acquisition the image must be re-focussed at each
tilt, achieved through the analysis of the image as a function of defocus. Such auto-focus
schemes are now well-established and relatively straightforward to implement. Further,
in STEM mode it is also possible to implement a ‘dynamic focus’ correction, in which
the probe focus is altered to account for the specimen geometry—particularly useful
when the specimen is at high tilt and one part of the specimen is at a considerably
different height to another [65].

3.2. Alignment

In the majority of electron tomography experiments in the life sciences, the alignment
of BF images within a tilt series is made difficult by the lack of distinct contrast. Two
practical methods can be used to help in this alignment: by tracking the movement
of high contrast fiducial markers (typically gold particles, a few nm in diameter) [66]
or by recording a STEM HAADF tilt series (which will be discussed in more detail
later) [67]. If the first method is used, it means that selection of an area for tomo-
graphic reconstruction is limited only to those areas that have sufficient markers for
alignment. This may be is acceptable for a specimen showing many structures dis-
persed on a carbon film but for specimens that have a small number of areas, perhaps
only one site-specific area, suitable for analysis, such as will often be the case in nan-
technology applications, this may make alignment by fiducial markers difficult, if not
impossible. An alternative is to rely on a cross-correlation alignment that, because of
the change in the projection of the object through the tilt series, must be carried out
image-by-image in a sequential fashion [68, 69]. It is important to correct for the
tilt geometry [70], whereby each image is stretched in a direction perpendicular to
the tilt-axis by \(1/\cos \psi\), where \(\psi\) is the angle of the projection relative to a refer-
ence zero tilt image, to improve the spatial relationship between successive projections.
This action converts orthogonal projections, with the specimen rotated with a fixed
source and detector, into inclined projections, which would exist if only the source was
rotated.

The direction of the tilt axis for the object must also be identified accurately before
any reconstruction is performed. For a single-axis tilt series, all objects through the
series should follow a path that is perpendicular to that tilt axis. If an accurate spatial
alignment has been achieved then a summation over all, or some, of the tilt series should
highlight the movement of any objects through the series. The path of movement
should be perpendicular to the tilt axis, as illustrated in Figure 4. Once the tilt axis has
been determined, the entire dataset is rotated to place this axis parallel to a single image
axis. The image stretch described before also has the effect of placing the tilt axis at the
centre of the zero tilt image. Any misalignment of the tilt axis will ‘spread’ the signal
from a reconstructed object and produce characteristic arcs of intensity, illustrated in
Figure 5. The direction of the arc will depend on the direction of the misalignment
away from the correct axis and the degree of ‘spread’ is dependent on the magnitude
of that misalignment [71]. These distinctive distortions can provide a very sensitive
method of refining the tilt axis.
19. Tomography using the Transmission Electron Microscope

Figure 4. Tilt axis direction determination by series summation (a) A single STEM HAADF image, extracted from a tilt series, of a catalyst composed of palladium nanoparticles on a carbon matrix. (b) The summation of the entire (aligned) tilt series showing a distinct streaking in one direction at an angle $\phi$ to the horizontal. (c) The power spectrum allows an accurate assessment of the tilt axis.

Figure 5. A demonstration of the effects of misalignment of the tilt axis on a reconstruction of a ‘head’ test object. The original object, from which the projections were generated, is shown top left. The number indicates the pixel misalignment, perpendicular to the tilt axis. The ‘head’ is 64 pixels wide.

3.3. Anisotropic Resolution

The sampling of the object controls the resolution of the tomographic reconstruction. For the single-axis tilt geometry, the resolution parallel to the tilt axis, say the $x$-axis, $d_x$, is equal to the original resolution of the projections, assuming a perfect tilt series
alignment, see later. The resolution in the other perpendicular directions is controlled by the number of projections acquired, $N$, and the diameter, $D$, of the volume to be reconstructed. This is seen most easily in Fourier space [38] and is:

$$d_y = d_z = \frac{\pi D}{N}$$  \hspace{1cm} (4)

However, this expression assumes that the $N$ projections cover the whole angular range (i.e. $\pm 90^\circ$). In practice the limited space between the objective lens pole pieces and the finite thickness of the specimen holder limits the tilt range, giving rise to the ‘missing wedge’ of information, see Figure 2. This missing information leads to the resolution in the direction parallel to the optic axis, $d_z$, being degraded further by an ‘elongation factor’ $e_{yz}$ so that

$$d_z = d_y e_{yz}$$  \hspace{1cm} (5)

which is related to the maximum tilt angle, $\alpha$ by [72]:

$$e_{yz} = \sqrt{\frac{\alpha + \sin \alpha \cos \alpha}{\alpha - \sin \alpha \cos \alpha}}$$  \hspace{1cm} (6)

Thus in order to provide the maximum 3D information, as many projections as possible should be acquired over as wide a tilt range as possible. Figure 6 illustrates this pictorially. As an example, the polepiece gap of the FEI Super TWIN objective lens is 5.2 mm. A standard FEI single tilt holder allows a maximum tilt angle of $42^\circ$, leading to an elongation factor, $e_{yz}$, of 2.29 and significant blurring of the reconstruction in the $z$-direction (parallel to the optic axis). To improve this, slimmer, narrower holders were constructed, firstly in-house [73] and more recently by commercial manufacturers [74]. These holders can now tilt to $\pm 70^\circ$ (an elongation factor of only 1.3) or even higher without undue shadowing or problems with the polepiece gap.

Alternatively, a ‘conical tilting’ approach can be used, made possible by either a second perpendicular tilt axis (double-tilt electron tomography) or a tilt-rotate holder, in which the cone angle is fixed and projections are acquired throughout a full precession of the specimen [75]. With this acquisition geometry the missing volume is a cone or pyramid, rather than a wedge, and the total volume of unsampled space is lower given the same maximum tilt angle. In that regard the conical or double-tilt approach offers a higher quality reconstruction but is technically far more demanding. Nevertheless, double-axis tomography is gradually becoming more popular as better quality reconstructions are demanded.

Whilst the Crowther criterion (equation 4) is a useful guide for the expected resolution in a reconstruction it ceases to become valid for constrained reconstruction techniques [76, 77], such as the iterative POCS-based methods and maximum-entropy methods (COMET). The reconstruction resolution of such methods is dependent on
Figure 6. (a) The two columns show the result of adding successively more projections to a tilt series for reconstruction using direct backprojection in the left hand column and weighted backprojection in the right hand column. The numbers refer to the number of projections over ±90°. (b) A montage of simulations showing the original object in the left-hand column, the direct backprojection reconstruction in the middle and the weighted backprojection reconstruction in the right-hand column. The reduction in the elongation by increasing the tilt range from ±10° to ±60°, and the improvement in resolution through weighted backprojection, is quite apparent. The tilt axis is perpendicular to the page.

the noise characteristics of the original data, the shape of the object to be reconstructed and the nature of the constraints applied. The reconstruction resolution has been the subject of much debate in the literature but a recent paper [78] has elaborated on a new way to define resolution for 3D reconstructions using a spectral signal to noise ratio method.
3.4. The Projection Requirement

Any signal used for a tomographic reconstruction must meet several assumptions of which, as stated by Peter Hawkes, ‘the most crucial is the belief that what is detected is some kind of projection through the structure. This ‘Projection’ need not be a sum or integral through the structure of some physical property of the latter; in principle a monotonically varying function would be acceptable’ [79]. This is known as the projection requirement. Until very recently, all published electron tomography results were derived from a tilt-series of bright field (BF) TEM images, the contrast in which arises due to a combination of low angle elastic and inelastic scattering. BF tomography is based on the assumption that for sufficiently thin, weakly scattering, large unit cell crystalline or amorphous objects contrast relating to the thickness and atomic number (‘mass–thickness contrast’) of the specimen dominates [79]. In structural biology, BF TEM images satisfy this criterion to a very good approximation, be they unstained cryo-specimens, embedded in ice, or stained plastic or resin sections. Of course, in principle the contrast transfer function should be taken into account but for most electron tomography experiments to date the resolution achievable has not required this correction.

However, in general, for most (crystalline) specimens in the physical sciences and certainly for most specimens of nanotechnological importance, BF contrast will depend strongly upon the diffraction condition of the crystal and this will not have a monotonic relationship with the amount of material though which the beam passes. BF images of such systems cannot be used for tomography because they are not strictly projections [79]. Further, even if the specimen is amorphous or weakly diffracting, the 3D nature of the specimen coupled with the short depth of focus in the TEM ensures that Fresnel contrast will be very apparent (especially if using a FEG-based instrument) and this again cannot satisfy the projection requirement. This is true, in principle, even for specimens in the life sciences.

It is only in the last few years [80], that electron tomography has begun to be applied to nanoscale systems in the physical sciences. Although BF imaging may not be suitable in general, there are many other signals that do satisfy the projection requirement. To overcome the problem of Fresnel contrast and diffraction effects, the signals acquired must be predominantly incoherent in nature. Both STEM HAADF (Z-contrast) imaging [81] and energy-filtered TEM (EFTEM) [82] can be seen as a good basis for electron tomography in the physical sciences. Both imaging techniques are, or can be made to be, incoherent and both are chemically sensitive enabling the 3D structure and composition to be mapped simultaneously at high spatial resolution. Very recently, STEM tomography has also been recognised within the structural biology community as an ideal means of imaging 1nm gold clusters within sections of biological material [66].

With these new imaging techniques available to the electron tomographer, it is now possible to produce high spatial resolution reconstructions of nanoscale objects relevant to nanotechnology research. A variety of examples using both STEM and EFTEM tomography are shown in the next two sections and reveal how, with care,
the 3D reconstruction of many structures and devices can be achieved with nanometre resolution.

4. STEM HAADF (Z-CONTRAST) TOMOGRAPHY

The low angle scattering of the electron beam is predominantly coherent in nature and as such conventional BF and DF images are prone to contrast reversals with changes in specimen thickness, orientation or defocus. On the other hand high angle scattering is predominantly incoherent, and STEM images formed using a high-angle annular dark field (HAADF) detector do not show the contrast changes associated with coherent scattering [83]. Within a classical description, such high angle scattering is associated with the interaction of the electron beam close to the nucleus of the atom and thus the cross-section for HAADF scattering approaches that for unscreened Rutherford scattering so that it is strongly dependent on the atomic number Z; in fact in the unscreened limit it is proportional to Z$^2$. In practice this limit is never reached and the exact dependence, particularly for crystalline specimens, is a function of many other factors, which need to be determined before any possible quantification can take place [84]. The choice of the inner angle for the HAADF detector, $\theta_{HAADF}$, is important and must be large enough to ensure coherent effects are minimal. A guide can be obtained from $\theta_{HAADF} \geq \lambda/d_{thermal}$ [85] where $\lambda$ is the electron wavelength and $d_{thermal}$ is the amplitude of atomic thermal vibration. For Si at 200 kV, $\theta_{HAADF} > 40$ mrad. For more information about STEM imaging and its uses in 2D nanotechnology, see the chapter by Cowley [86].

Medium-resolution ($\sim 1$ nm) STEM images, formed with a HAADF detector, are very sensitive to changes in specimen composition with the intensity varying (for the most part) monotonically with composition and specimen thickness, thus satisfying the projection requirement. Although atomic resolution HAADF images depend on the excitation of Bloch states and channelling [87], in principle even 3D atomic resolution is possible given a sufficiently thin specimen and a STEM with a high resolution (perhaps aberration-corrected) probe-forming lens. Such channelling effects are also present in medium resolution STEM images and when a crystalline specimen is at or near a major zone axis there is an increase in the STEM HAADF signal that depends on the localisation of the beam onto atomic strings. The string strength [88] dictates the level of intensity enhancement. However, in general, strong channelling will occur very infrequently during a tilt series and will have little effect on the overall intensity distribution in the reconstruction.

As with all techniques used for imaging 3D objects, attention must be paid to the depth of focus. This can be maximised by using a small condenser aperture (objective aperture on a dedicated STEM) to minimise the convergence angle but ultimately the diffraction limit will dominate and a residual blurring is inevitable. This will place a lower limit on the possible resolution achievable in all STEM tomography. In practice images are re-focused after every tilt, either manually or using computer control, to ensure optimum focus over the majority of the image. Recent results have shown how for a medium resolution STEM probe, say $\sim 1$ nm in diameter, the probe diameter,
II. Electron Microscopy

Figure 7. A typical STEM HAADF image of a heterogeneous catalyst composed of Pd₆Ru₆ nanoparticles (approximately 1 nm diameter) and an MCM-41 mesoporous silica support with mesopores of approximately 3 nm diameter.

or at least its central maximum, does not broaden significantly over a remarkably large range of defocus, particularly in an underfocus condition [89]. For STEM imaging, the tails of an underfocussed probe are not too important and in essence simply add an unwanted (and easily removed) background to the image. (This is of course not true for microanalysis where the tails of the probe can account for over half the emitted X-ray signal, for example.) The STEM image contrast can be described by a convolution of the central maximum of the probe with the object function and thus may account for the surprisingly good resolution of thick specimens in STEM mode compared to the resolution seen in an image acquired using conventional BF TEM [90].

The first example of STEM tomography is used to illustrate the resolution achievable with this technique and the ability to analyse the 3D data set in a statistical and quantitative fashion. In particular, it illustrates how the 3D distribution of nanometre-sized particles can be determined in a porous support. The specimen is a heterogeneous catalyst composed of bimetallic particles (each with a diameter of about 1 nm) within a mesoporous silica support (MCM-41) whose mesopores are hexagonal in cross-section with a diameter of about 3 nm [91]. Knowledge of the three-dimensional distribution of the metal nanoparticles, and their location at or close to the walls of the internal pores, is key to understanding the factors that govern the activity and selectivity of the nanocatalysts and their change during reaction as a possible result of sintering and coalescence [92]. Figure 7 shows a typical STEM HAADF image of one of these
catalysts, recorded with a detector inner radius of 40 mrad. With this set up, the image formed will be almost totally incoherent in nature. The nanoparticles, in this case Pt$_6$Ru$_6$, stand out very well against the light SiO$_2$ background and some appear to lie within the mesopores. However, to ensure that this is the case it is necessary to determine a 3D reconstruction of this and similar catalysts.

As an example, consider the series of 71 STEM HAADF images of the Pt$_{10}$Ru$_2$-MCM 41 catalyst [93] taken at 200 kV at tilts ranging from $+70^\circ$ to $-70^\circ$ in 2$^\circ$ steps. This catalyst has proven to be remarkably successful in hydrogenating trans-muconic acid to adipic acid, the former derived from glucose, the latter used to make nylon—a case of sugar being turned into plastic! The image series was spatially aligned using the modified cross-correlation algorithm [70]. Both a weighted back projection and iterative reconstruction were used with the iterative technique improving the fidelity of the reconstruction. All routines for the alignment and reconstruction were written in the IDL programming language [94].

Figure 8(a) shows a perspective view of the reconstruction and 8(b) and (c) show two images, displayed as multi-level voxel projections of the boxed area of the reconstructed catalyst, viewed in perpendicular directions, parallel and perpendicular to the MCM-41 pore structure. What is clear from the reconstruction is that the nanoparticles are very well resolved, in all directions, within the silica framework structure. Further the resolution is not degraded significantly in either direction. It also appears that the excellent activity of this catalyst is in part due to the relatively high filling quotient; there are a large number of particles in the pores with few if any in this view aggregated outside the pores. It is possible by sampling the 3D structure to calculate the number of particles in the volume, the internal surface area of the silica and thus the weight of active particle per unit area of support, about 20 $\mu$g·m$^{-2}$, about 3% of the initial loading. It is also possible to measure the occurrence of particles in each pore and whether the distribution is random or not. Analysis of the reconstruction reveals the mesoporous structure of the silica has been reconstructed faithfully with little sign of beam damage despite the long acquisition time (~3 hours) needed for the series. Such silica frameworks damage rapidly when examined at low voltage and/or in fixed beam (TEM) mode [95].

A further example of how STEM HAADF tomography can be used is in the determination of the external ‘shape’ of a nanoscale object. To demonstrate this we focus on the magnetite (Fe$_3$O$_4$) nanocrystals found in the ‘backbone’ of magneto-tactic bacteria. Such organisms use this ‘backbone’ of magnetite crystals, which are ferromagnetically aligned, to sense the earth’s geomagnetic field and thus aid navigation and feeding [96]. Recently, they have become of great interest as similar magnetite chains have been observed on the surface of martian meteorites [97–100]. To determine whether the similarity is more than superficial, enormous efforts are being made world-wide in order to characterize these crystals in particular the crystal habit and any variation in the composition within the crystal. As such, 3D analysis is vital.

Figure 9 shows a phase image reconstructed from an electron hologram of such a bacterium that illustrates quite convincingly the ferromagnetic alignment. The bottom left inset is a BF image of a typical bacterium that highlights the backbone of crystals.
Figure 8. (a) A perspective view (voxel projection) of a reconstruction of a heterogeneous catalyst composed of Pt_{10}Ru_{2} active nanoparticles supported within an MCM–41 framework. The reconstruction was undertaken on a series of STEM HAADF images acquired every 2° between +/-70°. (c) and (d) two perpendicular voxel projections of the reconstruction volume boxed in (b). In (c) the hexagonal order of the mesoporous silica is evident (inset shows power spectrum) and in (d) it is possible to see how the pores are filled with the nanoparticles.

surrounded by the cellular ‘envelope’. Figure 10(a) shows a tomographic reconstruction (surface render) from a series of STEM HAADF images recorded between +/-76° with a 2° interval. Both the organic envelope and the backbone have been shown in the upper figure, the backbone alone in the lower figure. Note the helical arrangement of the crystals, known to exist in these systems. One of the nanocrystals has been boxed in the lower figure and two slices perpendicular to the main axis of the crystal are shown in Fig 10(b), one from near the end of the crystal, the other from near the centre.
Figure 9. A reconstructed phase image from an electron hologram of a magneto-tactic bacterium showing the magnetic field lines of the ferromagnetic chain of magnetite crystals. The inset shows a BF image of the bacterium revealing the backbone of crystals within the organic membrane (figures courtesy of R. E. Dunin-Borkowski).

The cubic nature of the magnetite phase allows the facets revealed by tomography to be indexed unambiguously as shown. Further the near perfect hexagonal symmetry of the central slice is revealed with great clarity by the reconstruction, showing the 6 symmetrically equivalent \{110\} facets.

Although common, this 3D morphology (or habit) is not unique to these systems and occasionally ‘trigonal’ prismatic crystals are seen, an example of which is shown reconstructed in Figure 11(a). In this crystal two \{111\} facets are dominant. This crystal is seen, arrowed, in Figure 11(e), part of a series of images, seen in Figures 11(b)–(e) used to illustrate the effects of channeling that exist in STEM HAADF imaging of crystals. The figure is a montage of four STEM HAADF images recorded at different tilts. As the bacterium is tilted each nanocrystal will, in general, be at a different orientation to the incoming electron beam. If the crystal planes are at, or close to, a strong diffraction condition (for example close to a low order zone axis) then the strength of the scattering to high angles as recorded in the HAADF image, will increase because the incoming beam will be localised on atom strings and propagate through the crystal as Bloch states. This can be seen in the montage of figures as a sudden increase in the HAADF signal for certain crystals at certain orientations, for
II. Electron Microscopy

Figure 10. (a) A tomographic reconstruction of a magneto-tactic bacterium showing very clearly the backbone of magnetite crystals surrounded by the bacterium’s organic ‘envelope’. (b) Two slices from the crystal boxed in (a), the left hand slice is taken from the end of the crystal, the right hand crystal from the middle. Note the excellent fidelity of the reconstruction and the perfection of the crystal facetting.

example, crystal X in (b), crystal Y in (c), crystal Z in (d) and a small crystal behind that labelled Y in (e). As this dynamical enhancement occurs only once or twice in a series of perhaps 140 images it makes little difference to the final reconstruction, particularly if the exterior shape is all that is required. The prismatic crystal, shown in Figure 11(a), is seen very clearly in (e). Animations of these reconstructions can be viewed on our web site [101].

For all STEM tomography reconstructions, it is difficult to determine an overall 3-D resolution but as a rule of thumb, for large objects, \( D > 100 \) nm, we have found by experience that the 3-D resolution is approximately \( D/100 \) [73].
5. **EFTEM TOMOGRAPHY**

With the advent of both post-column and in-column energy filters [102], energy filtered transmission electron microscopy (EFTEM) has become a routine analytical tool that allows rapid quantitative mapping of elemental species over wide fields of view with a spatial resolution of $\sim 1\,\text{nm}$. [103–105]. If an energy slit (window) is used that allows only the zero-loss part of the spectrum to be transmitted then images can be formed using only (predominantly) elastically-scattered electrons (typically $+/-5\,\text{eV}$). By removing electrons that have undergone inelastic scattering of greater than about 5 eV, the contrast of BF images is improved enormously, particularly for thick specimens as used often in structural biology [106], for 2D or 3D imaging. Chemical analysis by core loss imaging (using energy losses characteristic of a particular atomic species) is rarely used in biology because of the high electron doses necessarily involved [107]. However most physical sciences specimens are several orders of magnitude more beam stable than their biosciences counterparts and therefore by using a tilt series of core-loss images it should be possible to reconstruct a three-dimensional elemental distribution map [82, 108].

The intensity observed in an image formed using an energy loss window is a complex combination of inelastic scattering (through changes in composition and electronic structure) and elastic effects (via crystal thickness and orientation). The true compositional information encoded in an energy loss image may be isolated by generating either a background subtracted elemental map (from three or more images) or a jump-ratio map (from two images). Both maps will show intensity that is related to the amount of an atomic species at a given pixel. However, elemental maps often show residual diffraction contrast that will, in general, mean they do not conform to the projection

---

**Figure 11.** (a) A reconstruction of a ‘prismatic’ crystal. (b)–(e) HAADF images taken from a tilt series used to generate (a). Note the changes in the HAADF intensity in some crystals, labelled X, Y and Z that arise when the beam direction is close to a major zone axis. The triangular crystal arrowed in (e) is shown as a 3D reconstruction in (a).
requirement, in the same way as it would for a conventional BF image. Diffraction contrast can be removed partially by dividing the map by a zero-loss image, but this can also introduce artefacts associated with changes in the diffraction contrast itself as a function of energy loss: the diffraction contrast in a zero loss image is considerably sharper than that of a core-loss image. However, jump-ratio images are a convenient and simple way of removing residual diffraction contrast. They can show higher sensitivity than an elemental map but of course the intensity values of a jump-ratio map cannot be related in an absolute (quantitative) way to the composition. The jump ratio signal changes monotonically with thickness up to approximately the overall inelastic mean free path, $\lambda$ [107]. Beyond this value the jump-ratio actually falls as the specimen thickness increases as the true elemental signal increases more slowly than the underlying background. Thus the jump-ratio signal can only be used for tomography so long as the specimen thickness (in projection) is less than one inelastic mean free path, typically 100 nm at 200 kV. This places some constraints on the sample preparation or the types of specimens examined in this fashion.

To illustrate the advantages of EFTEM tomography, consider a typical problem seen in the metallurgical field namely the precipitation of chromium carbides at a grain boundary in stainless steel. Although larger in scale than cases considered previously, it illustrates here the advantage of EFTEM over STEM tomography in that the atomic number contrast between the precipitates and the matrix is small. It should be possible to analyse the shape of the carbides from a tomographic reconstruction of a tilt series of chromium jump-ratio images. A series of EFTEM images were acquired using a Philips CM300, with a Gatan Imaging Filter (GIF) fitted with a $2k \times 2k$ CCD camera. The dataset was acquired at 24 tilt increments over a tilt range of $\pm 58^\circ$, an increment of just under $5^\circ$. At each tilt three energy loss images, each with a 10 eV window, were acquired over the chromium L$_{23}$ edge (onset at 575 eV), two pre-edge at 545 eV and 565 eV and one post-edge at 580 eV. Jump-ratio and elemental maps were determined at each tilt but the latter showed considerable diffraction contrast and therefore only the jump-ratio signal was used for reconstruction.

Each group of energy loss images were corrected for any shift relative to the first pre-edge image of each group using a cross-correlation routine. Spatial and rotational alignment through the tilt series was corrected by sequential cross-correlation and series averaging. Towards the extreme ends of the tilt series there was a loss of contrast in the jump-ratio images, which was perhaps because the amount of material through which the beam passes may have increased beyond the upper thickness limit of one mean free path. The tomographic reconstruction was carried out using weighted backprojection and is shown as three perpendicular voxel projections in Figure 12. These projections clearly show that chromium carbides have complex 3D shapes and orientations, the nature of which could only be surmised from a single EFTEM elemental distribution image. For example, it becomes clear from (b) that the upper boundary between the precipitates and the matrix is (at least partially) coherent, the lower boundary predominantly incoherent in nature. An animation of this reconstruction is shown on our web site [101].

We return to the magnetotactic bacteria crystals for a second example to show the polytropic montage described earlier. In this case, reconstruction of the magnetite
19. Tomography using the Transmission Electron Microscope

Figure 12. (a) BF zero loss image of a grain boundary in stainless steel which shows carbide precipitation at the boundary. (b)-(d) Voxel projections of a tomographic reconstruction using Cr jump-ratio images of the grain boundary carbide structure. The carbides are viewed in three perpendicular directions, emphasizing the morphology of the precipitates. The reconstruction has been smoothed with a 2×2×2 pixel Gaussian filter prior to visualization in order to reduce the effects of noise. In addition, the voxel projections have been contrast selected to show only the chromium carbides.

The ‘backbone’ of crystals was carried out using an iterative algorithm for both the oxygen and iron datasets to help improve the reconstruction which would otherwise suffer from the low signal-to-noise ratio (SNR) in some of the original projections, especially in the oxygen tilt series. For more details of this reconstruction see [108]. Re-projections of both the iron and oxygen reconstructions in the zero degree direction show much higher SNR than the original projections; this increase in SNR is the polytropic montage effect. This is most clearly seen in the case of the oxygen data, shown in Figure 13 in which the projected reconstruction is compared with the original oxygen jump-ratio image at zero degree tilt. The improvement is remarkable.

6. CONCLUSIONS

It is evident that electron tomography offers a means to determine the three-dimensional structure and composition of many different materials at the nanometre level. In general, tomography using BF TEM for materials science applications will not yield true reconstructions because of the coherent nature of the scattering process seen in such images. BF images contain contrast that does not satisfy the projection requirement for tomography. Incoherent signals, such as those used to form STEM HAADF images or core-loss EFTEM images, do satisfy the projection requirement, at least within certain limits. Further, by using these imaging techniques, it is possible to simultaneously record three-dimensional compositional information, either indirectly through the atomic number dependence of HAADF imaging or directly, by choosing a window that corresponds to a energy loss (electronic transition) within a
particular atomic species. This one-to-one correspondence of structure and composition in three dimensions should give the physical scientist a very powerful method to analyse nanoscale structures and devices in the future.

ACKNOWLEDGEMENTS

The author would like to acknowledge the invaluable contribution of Dr Matthew Weyland to this work. He would also like to thank Prof. Sir John Meurig Thomas, Prof Peter Buseck and Drs Chris Boothroyd, Rafal Dunin-Borkowski and Ron Broom for their help and interest. He thanks the EPSRC, FEI, Isaac Newton Trust and the Royal Commission for the Exhibition of 1851 for their financial support over a number of years.

REFERENCES

19. Tomography using the Transmission Electron Microscope


II. Electron Microscopy


