Microscopy AND
Microanalysis

Original Article

Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part II: Post-Acquisition Data Processing, Visualization, and Structural Characterization

Gary W. Paterson[1]* (iD), Robert W. H. Webster[1] (iD), Andrew Ross[1],† (iD), Kirsty A. Paton[1] (iD), Thomas A. Macgregor[1] (iD), Damien McGrouther[1] (iD), Ian MacLaren[1] (iD) and Magnus Nord[1,2]*‡ (iD)

[1]SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK and [2]EMAT, Department of Physics, University of Antwerp, Antwerp 2000, Belgium

Abstract

Fast pixelated detectors incorporating direct electron detection (DED) technology are increasingly being regarded as universal detectors for scanning transmission electron microscopy (STEM), capable of imaging under multiple modes of operation. However, several issues remain around the post-acquisition processing and visualization of the often very large multidimensional STEM datasets produced by them. We discuss these issues and present open source software libraries to enable efficient processing and visualization of such datasets. Throughout, we provide examples of the analysis methodologies presented, utilizing data from a 256×256 pixel Medipix3 hybrid DED detector, with a particular focus on the STEM characterization of the structural properties of materials. These include the techniques of virtual detector imaging; higher-order Laue zone analysis; nanobeam electron diffraction; and scanning precession electron diffraction. In the latter, we demonstrate a nanoscale lattice parameter mapping with a fractional precision $\leq 6 \times 10^{-4}$ (0.06%).

Key words: 4D-STEM, fast pixelated detector, higher-order Laue zone, scanning precession electron diffraction, virtual detectors

(Received 6 April 2020; revised 9 June 2020; accepted 2 August 2020)

Introduction

One of the greatest revolutions in scanning transmission electron microscopy (STEM) in recent years is the development and use of fast pixelated detectors incorporating direct electron detection (DED) technology, and these are rapidly becoming a key component of the imaging system for a modern STEM. Ophus (2019) has provided an excellent review of the area, and Tate et al. (2016), Yang et al. (2015), and Krajnak et al. (2016) have described some suitable detectors for different applications in pixelated STEM imaging. Part I (Nord et al., 2020) of this work briefly discussed the general benefits of fast pixelated detectors for use in STEM and presented software solutions for their hardware control, and for the acquisition, real-time processing and visualization, and storage of data from them. An increasing number of Python packages are being developed by the electron microscopy community with capability to process four-dimensional (4D) STEM (4D-STEM) data, including HyperSpy (de la Peña et al., 2018), LiberTEM (Clausen et al., 2019), py4DSTEM (Savitzky et al., 2019, 2020), pycroscopy (Somnath et al., 2019), and pyXem (Johnstone et al., 2019).

In this present article, Part II, we discuss post-acquisition data exploration and analysis of a variety of STEM datasets acquired with a fast pixelated detector, using the fpd (fpd devs, 2015; fpd demos devs, 2018) and pixStem[1] (pixStem devs, 2015) Python libraries, with a focus on their use in mapping the structural properties of materials. These include the techniques of virtual detector imaging (Rauch & Veron, 2005; Gammer et al., 2015), higher-order Laue zone STEM (HOLZ-STEM; Huang et al., 2010; Nord et al., 2019), nanobeam electron diffraction (NBED; Béché et al., 2009), and scanning precession electron diffraction (SPED; Vincent & Midgley, 1994). Throughout, we provide examples of the application of these techniques to data recorded with a Medipix3RX hybrid counting DED detector (henceforth referred to as Medipix3) (Ballabriga et al., 2013)

‡Current address: Department of Physics, Norwegian University of Science and Technology, Trondheim 7491, Norway.

*Authors for correspondence: Gary W. Paterson, E-mail: dr.gary.paterson@gmail.com; Magnus Nord, E-mail: Magnus.Nord@ntnu.no

†Current address: Institute of Physics, Johannes Gutenberg Universität Mainz, Staudingerweg 7, 55128 Mainz, Germany.

[1]Toward the very end of preparing this paper, it was decided to merge pixStem with pyXem (Johnstone et al., 2019). All of the features detailed in Part I and II of this series of papers that are related to pixStem are in the processes of being added to pyXem and will continue to be available. The features of the fpd library remain unaffected.

CrossMark

mounted in STEMs with 200 kV electron sources, though the techniques and software packages discussed here extend to data from other detectors. The names of the specific functions, classes, methods, modules, and packages used in the examples given are specified in the main text or the figure captions in typewriter font.

Both of the software libraries presented are made available under the free and open source GPLv3 license, allowing transparency of the implemented algorithms, and the ability for anyone to use and to further improve upon them. The libraries themselves draw on the rich ecosystem of mature Python libraries (Oliphant, 2007), including ones for optimized numerical (Oliphant, 2006) and scientific (Jones et al., 2001) computing, image processing (van der Walt et al., 2014; Gouillart et al., 2016), and data visualization (Hunter, 2007).

The pixStem library is built upon HyperSpy and extends its capabilities to common pixelated STEM tasks, with a focus on processing HOLZ data. A key difference between the pixStem and fpd packages is that the latter library is a lower level one, similar in style to NumPy but with added classes, GUIs and plotting capabilities built-in where needed. A higher-level object-oriented interface will be added in a future version. A detailed description of every feature of the pixStem and fpd libraries can be found in the documentation and example Jupyter (Kluyver et al., 2016) notebooks available online (fpd devs, 2015; pixStem devs, 2015; fpd demos devs, 2018) and will not be replicated here. Instead, in the following sections, we describe some of the general techniques employed by the libraries for efficient data processing, for data visualization, and the main features of the libraries for the structural characterization of materials.

The paper is organized as follows. Section "General Techniques" outlines general techniques for handling postprocessing of the large 4D-STEM datasets typically produced by fast pixelated detectors. In the section "Data visualization", ways of visualizing the datasets are discussed. Section "Virtual detector images" covers the various methods of generating virtual images from the 4D datasets, using annular dark-field (ADF) contrast from a noncrystalline soft material as an example. In the section "Higher-order Laue zone analysis", HOLZ-STEM data analysis is explained, which allows for the retrieval of information about the periodicity of a crystalline sample in the direction parallel to the electron beam. Lastly, general lattice image processing is discussed in the section "Lattice analysis", with a focus on diffraction imaging. Using test data obtained in the SPED mode of acquisition (Vincent & Midgley, 1994) of a custom system (MacLaren et al., 2020), we demonstrate that a lattice parameter fractional precision of 6×10^{-4} is possible using a probe with a spatial resolution of 1.1 nm. This precision value is approximately 2–3× larger than the best ones reported in the literature using standard probes in a SPED mode (Rouvière et al., 2013) and patterned probes in standard STEM mode (Guzzinati et al., 2019; Zeltmann et al., 2020).

A future Part III of this work will cover post-acquisition processing and visualization of data from fast pixelated detectors for differential phase contrast imaging.

General Techniques

The primary challenge to the post-acquisition processing of data from fast pixelated detectors is the size of the data, which can easily be much greater than the available system memory on current generation computers. For example, a spatial scan with 1k × 1k points recording data from a 512 × 512 pixel detector would occupy ~977 GB when stored in 32-bit integer format. In addition to this, in some types of processing, at least the same amount of memory is required to process the data, putting the potential memory requirements well into the terabyte range. For STEM imaging, smaller scan sizes of 256 × 256 or 512 × 512 points with a 256 × 256 pixel detector at lower bit depth are often adequate to observe the specific feature of interest. In these cases, and for 16-bit data, the memory requirements are more modest, being 8.6 and 34.4 GB (in SI units), but the problem of available system memory and efficient computation currently remains.

One solution to these issues is out-of-core processing, where the data are stored on disk and only parts of it are loaded into memory when needed. As an example, to generate a bright-field image from a STEM dataset, each reciprocal space image may be loaded one at a time from a hard drive into memory, a sum of each image performed, with the resulting single values stored in memory, and the diffraction pattern memory reused to store the next image. This is an extreme example and, in reality, multiple images may be loaded into memory at once and processed across multiple CPU cores in parallel to achieve significant performance improvements.

Out-of-core processing is termed "lazy-loading" in HyperSpy and was recently added through the use of Dask (Dask Development Team, 2016), a library which abstracts away the complexities of out-of-core processing. The pixStem library relies on HyperSpy for out-of-core processing, whereas the fpd library was implemented before out-of-core processing was available in HyperSpy and so implements its own methodology, relying on HyperSpy for Gatan Digital Micrograph file access.

The two libraries can both also process in-memory data. When processing out-of-core data on disk, there are several ways of making the data exploration and analysis more efficient. Chunking the dataset in a compressed HDF5 (The HDF Group, 1997–2018) file, as discussed in Part I (Nord et al., 2020), in the scan and detector dimensions by partitioning the diffraction pattern data into several pieces, can greatly speed up data processing by both reducing the data reading and decompression time, and by reducing the in-memory array sizes. For example, if one wants to make a bright-field image by using a small virtual aperture, most of the diffraction pattern can be ignored. Without chunking the detector dimensions, each full image and, thus, the entire dataset must be read. However, when a 256 × 256 pixel diffraction pattern is partitioned into 16 × 16 chunks and the bright-field disc is located within one such chunk, only 0.39% of the dataset has to be read and processed. The use of HDF5 files for data storage brings with it many additional benefits, and these are discussed more fully in Part I (Nord et al., 2020).

In both packages, reading of chunked data from disk may be aligned to the chunk size when possible to optimize read times. The size of the data read into memory can be set to be independent of the chunk size or be multiples thereof, allowing data to be processed even on systems with limited memory and processing power. Additionally, many of the algorithms in the fpd and pixStem library can be set to operate using one CPU core or, by default, to employ all available cores to process the in-memory data chunk in parallel, enabling significant speed increases over single core processing.

Binning the data is another implemented method that can increase data processing speed. Detector resolution is an important factor when considering the influence of binning on data analysis. Most fast DEDs have relatively few pixels compared to

conventional CCD-based detectors and the reduction in pixel counts on rebinning can limit the ability to extract the maximum signal from the data. In many cases, however, modest rebinning can be applied without significant changes in the signal-to-noise ratio (SNR), or large rebinning can be applied to vastly increase the processing speed. If the data were recorded using a detector that has no readout noise and that counts electrons only in the pixel in which they entered the detector's sensor, then the noise would be Poissonian and, for some analyses, such as center of mass, rebinning has almost no affect on the SNR of the analysis result.

The Medipix3 detector (Ballabriga et al., 2013) has a noise-free readout when the counting threshold is set above the level of the detector's intrinsic noise and, at electron beam energies up to 80 keV, the detector is capable of a near perfect response, counting electrons mainly in the pixel of entry (Mir et al., 2017). However, 200 keV electrons can travel long distances in the sensor and are counted by multiple pixels (McMullan et al., 2007), giving rise to a point spread function of a few pixels. This correlation modifies the noise profile of the detector from pure Poisson statistics. However, rebinning the data, even by a small extent, nullifies the correlated component of the noise, resulting in a Poissonian noise profile at the expense of lower detector resolution. While rebinning tends to have little effect on the SNR of the result of many analysis techniques, it can improve it in some cases, and the speed improvements it brings are particularly useful when performing live data processing, as discussed in Part I (Nord et al., 2020), or for rapid post-acquisition assessment of the data while on the microscope.

The binning of data can also be used to reduce the dataset size to the point where it can fit in memory. For example, binning the detector dimensions of a $256 \times 256 \times 256 \times 256$ 12-bit dataset by 2 along each axis reduces the size by a factor of 4 from 8.6 GB down to a more manageable 2.1 GB.

Data Visualization

Once a dataset has been acquired, it is important to be able to visualize the raw data and to be able to correlate it with the results produced from processing the dataset. The DataBrowser class of the fpd.fpd_file module allows basic GUI-based data inspection, using by default the pre-rendered sum image for navigation if one of our HDF5 files (Nord et al., 2020) is supplied, and loading images on-the-fly as needed. The class can also display data from other sources, such as in-memory NumPy arrays and memory mapped files, and so it is also useful when inspecting data immediately after acquisition and before conversion of the data to the HDF5 format.

To demonstrate the DataBrowser class, we use it to display a recently published nanobeam diffraction dataset (Temple et al., 2018b) from a cross section of a patterned, epitaxially grown FeRh/NiAl/MgO sample (Temple et al., 2018a). FeRh, in its B2 structure, exhibits a tunable first-order transition at a temperature of 370 K, which is accompanied by a ferromagnetic to antiferromagnetic ordering transition (Lewis et al., 2016). These properties make FeRh of interest for a number of applications, including data storage and sensors. In the following, we present an analysis of the crystal structure, a key property in determining the magnetic ordering, using the methodology detailed later, in the section "Lattice analysis".

The top row of Figure 1 shows the screenshots of the two GUI windows of the DataBrowser class. The navigation image in Figure 1a is a user-supplied one showing the diffraction pattern lattice parameter, k, parallel to the growth direction, as indicated by the black arrows in Figures 1a and 1b (the other annotations in Fig. 1a have been added to the GUI image for this presentation of the data). The diffraction pattern at the location in the scan marked by the bottom white square in Figure 1a is shown in Figure 1b. The displayed diffraction pattern can be selected by clicking anywhere in the navigation image or by dragging the white marker with the mouse, enabling the dataset to be examined.

The extent of the FeRh and NiAl layers and the MgO substrate are readily identifiable from regions of uniform color in the navigation image of Figure 1a. The black region in this image is from nanocrystalline platinum deposited prior to forming the cross section and has been masked in the image. Beneath the Pt layer, the top and side sections of the FeRh layer have been modified by a combination of the 1 kV Ar^+ etching used in the sample patterning (Temple et al., 2018a) and damage done during focused ion beam cross-section preparation, creating disordered regions and grains of different FeRh structures.

While the changes in the mapped lattice parameter are indicative of different grains, the other lattice parameters must be considered when defining a grain. In Figures 1c–1j, we show diffraction patterns outlined in the same color as the regions from which they are taken in Figure 1a, with the exact single scan pixel location of each pattern marked by the equivalently labeled white or black boxes within each region. The diffraction patterns from the NiAl layer (Fig. 1j) and the epitaxial section of the FeRh layer (Fig. 1i) are both consistent with a B2 structure when viewed along the [011] direction, as reported previously (Temple et al., 2018a), while the MgO pattern (Fig. 1b) is consistent with the NaCl structure when viewed along the [100] zone-axis. The regions labeled in Figures 1d–1h each have a diffraction pattern consistent with a chemically disordered bcc crystal structure (Temple et al., 2018a), with the rotation of the crystal orientation, lattice parameter, and strain varying between regions (cf. Figs. 1e–1g). Over the imaged area, approximately half of the FeRh material is in a disordered phase.

In addition to the structural analysis of the type discussed above, the ability to inspect the 4D dataset in this manner is particularly useful in magnetic imaging, where the source, either magnetic or crystalline, of apparent beam-shifts can be interrogated by navigating the 4D dataset using a color vector magnitude image produced from the analysis [examples of vector magnitude plots of this type may be seen in Paterson et al. (2019)]. This type of beam-shift processing will be discussed further in Part III.

More configurable plotting of data from our HDF5 files (Nord et al., 2020), such as plotting along different axes, may be achieved by opening the EMD formatted (EMD authors, 2019) datasets embedded in the HDF5 file in HyperSpy (de la Peña et al., 2018). This can be performed with the fpd_to_hyperspy function of the fpd.fpd_file module or by loading them directly through the pixStem or HyperSpy libraries. Furthermore, the datasets may also be loaded into any custom analysis or visualization code written in Python using the fpd_to_tuple function which relies on the h5py library (Collette, 2013), or indeed any of the many other languages supporting the HDF format (The HDF Group, 1997–2018).

Virtual Detector Images

Traditional STEM detectors are routinely used to generate image contrast by collecting the signal from electrons scattered through
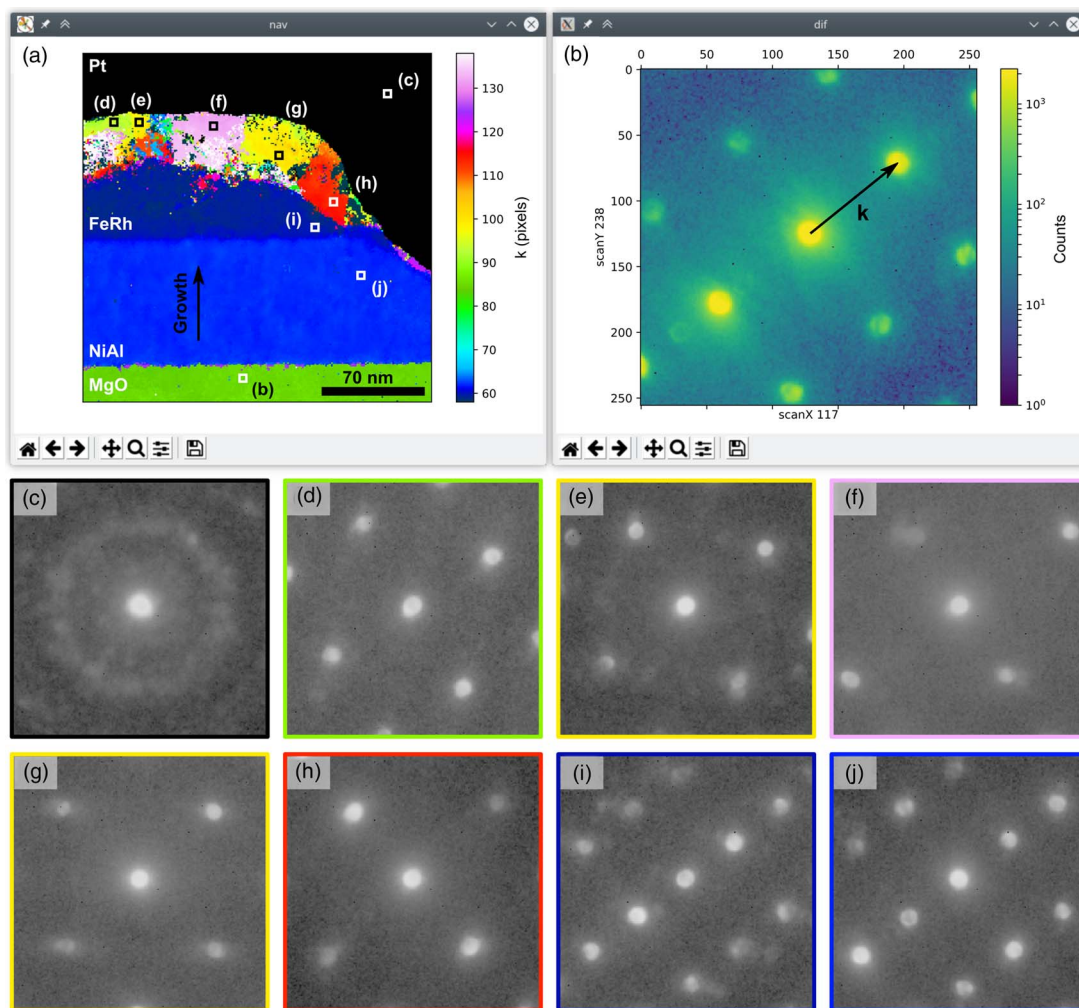
Fig. 1. Example of the DataBrowser class showing (**a**) a user-supplied navigation image and (**b**) the diffraction pattern from the scan point marked by the white square at the bottom of (**a**). The dataset is from an FeRh sample (Temple et al., 2018*a*, 2018*b*) and the navigation image is the diffraction pattern lattice parameter in pixels along the material layer growth direction, obtained by analyzing the data using the methodology detailed in the section "Lattice analysis". This direction is marked by the vertical arrow in (**a**). The equivalent axis is marked by a similar arrow in (**b**) and is rotated with respect to the scan axes. The black region in (**a**) is from electron and ion beam deposited platinum and has been masked. All other annotations in (**a**) have been added to the GUI screenshots to mark the locations of the diffraction patterns shown in panels (**c–j**). Details of the acquisition from the referenced publications are a scan size of 256 × 256 points, a probe step size of 0.93 nm, and a semi-convergence angle <1 mrad.

different angles by the sample under study, from bright-field (BF; LeBeau et al., 2009; MacLaren et al., 2013), through annular bright-field (ABF; Hammel & Rose, 1995; Findlay et al., 2010) to ADF, and high-angle annular dark-field (HAADF; Pennycook & Jesson, 1991; Hartel et al., 1996). However, these detectors offer a limited range of often mutually exclusive fixed collection angles at any given camera length. By using a pixelated detector, the range of scattering angles is resolved and this provides great flexibility in the both the number and range of collection angles from which images may be constructed by applying "virtual" apertures to the dataset in software (Rauch & Veron, 2005; Rauch & Véron, 2014). In addition to allowing greater insight into the sample properties, pixelated detectors have the potential benefit of being more dose efficient than segmented diode or photomultiplier tube detectors (Shibata et al., 2010) by avoiding repeated scans when multiple collection angles are required, and also do not require the careful calibration (Jones et al., 2018) that traditional STEM detectors do in order to minimize errors in quantitative imaging. Furthermore, when there are

no de-scan coils in a system, or where they are not perfectly set up, movement of the probe position on the detector will occur; an effect that is especially apparent in large area scans. This can be corrected for in software when using a pixelated detector, whereas with fixed geometry detectors, a range of collection angles depending on the probe position would be sampled, potentially giving rise to nonintuitive or nonquantitatively-interpretable contrast in the resultant image.

In many cases, this effect is not a significant one, but when it is, the data may be corrected during HDF5 file creation (Nord et al., 2020) or the data may be corrected afterwards. Multiple methods may be used to determine the direct beam position. Thresholded center of mass and edge-filtered phase correlation techniques, such as those implemented in the fpd.fpd_processing module, can be used in many cases and will be discussed in detail in Part III with regards to differential phase contrast imaging. Canny edge detection combined with Hough transforms may also be used, but this tends to be much slower than other methods due to the increased computational requirements of the technique.

This approach, implemented in the find_circ_centre function of the same module, is demonstrated in the section "Direct beam characterization" along with an edge-fitting technique.

Accurate knowledge of the center position is clearly important when using it as basis for making radial profiles in order to prevent mixing of scattering angles and, thus, maintain angular resolution. All of the aforementioned methods for center finding work well for diffraction patterns acquired with a low convergence angle beam in cases where there is no overlap between the diffraction discs. When there is a large degree of overlap between the discs, such as in data acquired with a high convergence angle beam, the task becomes more difficult (examples of these sorts of images are shown in Figures 4a–4c, discussed later). If the sample is on-axis, the aforementioned methods should work well. However, a small misalignment of the zone-axis is a common issue in crystalline TEM samples, due to the sample thinning process imparting a slight bend in the samples. This can cause the intensity in the center part of the diffraction pattern to shift, resulting in a misidentification of the center position of the direct beam. For some cases, edge detection followed by Hough transforms can work surprisingly well, as can thresholded center of mass, if the threshold is chosen carefully. Another way of mitigating this is scanning a smaller area; this limits the misalignment (and also de-scan), but is not always desirable when large fields of view are required. A potential solution to this issue in particularly difficult samples is to acquire a reference dataset in vacuum with exactly the same experimental settings, except for the sample stage position. Here, only the direct beam is visible, so calculating the center position is trivial, but care must be taken to ensure that the sample does not deflect the beam due to electromagnetic fields, electrostatic charging, or the sample being slightly tapered. Another possibility is having reference areas within the dataset, for which the center position can be interpolated for the whole field of view. Aspects of these issues, in particular, de-scan correction, will be covered more thoroughly in Part III.

Some idea about the characteristic scattering angles of a sample may be obtained from direct inspection of a single or averaged diffraction pattern, or by calculating the radial profiles of their intensities. Depending on the sample properties, this process may be improved by examining other measures of contrast, such as the variance or maximum values across the scan axes. For the biological sample used here, there is little difference in the apparent contrast from these measures and, as we will show, it is not always simple to determine which scattering angles yield the best contrast in virtual images. As an example, in Figure 2, we plot the radial distribution of scattering intensity from a microtome section of a mouse liver. The inset shows the diffraction pattern summed over all scan positions (this is the same as what is sometimes referred to as a position averaged convergent beam electron diffraction (PACBED) pattern, without dividing by the number of scan points) that was pre-calculated during conversion of the 4D data to our HDF5 format (Nord et al., 2020). We will stay in pixels for the simplicity of the discussion, but the images may be calibrated in scattering angle or, equivalently, reciprocal space units (such as k or Q). Marked by arrows in the diffraction pattern are two rings of different widths; these are more easily visible at around pixels 63 and 102 in the mean intensity profile (also arrowed). The drop in intensity beyond in the corners of the image marks the aberration-limited field of view of the probe-corrected JEOL ARM200cF STEM (McVitie et al., 2015) in the objective-off mode that was used to record the data.

The scattering angles marked by the peaks in Figure 2 may be used to define one or more virtual apertures. These are simply arrays with values between 0 and 1, corresponding to regions outside and inside the collection angles of interest, respectively, and are applied to the whole dataset by multiplication and then summing the total counts for each mask. Methods for doing this are provided in the synthetic_aperture and synthetic_images functions from the fpd.fpd_processing module, and by the virtual_annular_dark_field method of pixStem's PixelatedSTEM class. The resulting virtual detector images allow the contributions to different scattering angles to be spatially resolved. This approach also works well when using arbitrary, nonannular masks, such as when selecting one or more diffraction spots, where the contrast produced may be used to determine the location and extent of different phases or lattice defects (Gammer et al., 2015).

With out-of-core processing from compressed HDF5 files, these calculations may take several seconds, depending on the scan size, the computer hardware, and the data chunking. This approach is thus well suited to generating images from known collection angles. While inspection of the diffraction pattern is suitable for identifying the relevant angles in simple cases, it is far more useful to gain live feedback on the changes in spatial contrast from alteration of the virtual detector collection angles. To enable this, the fpd.fpd_processing module also provides the VirtualAnnularImages class. This class maps the radial_profile function, discussed above, across all scan points using the map_image_function function of the same module, and then generates weighted cumulative sums of the radial profiles that may be looked-up at a later point to form the virtual detector image from any given detector geometry. This intermediate dataset is typically ~40–50 times smaller than the 4D one and is currently stored within the class since it can easily fit in memory. For the data in Figure 2, the size is reduced from 8.6 GB to 195 MB. The significantly smaller size of 3D datasets also makes them much more amenable to multivariate analyses such as non-negative matrix factorization (NMF), independent component analysis (ICA), and principal component analysis (PCA).
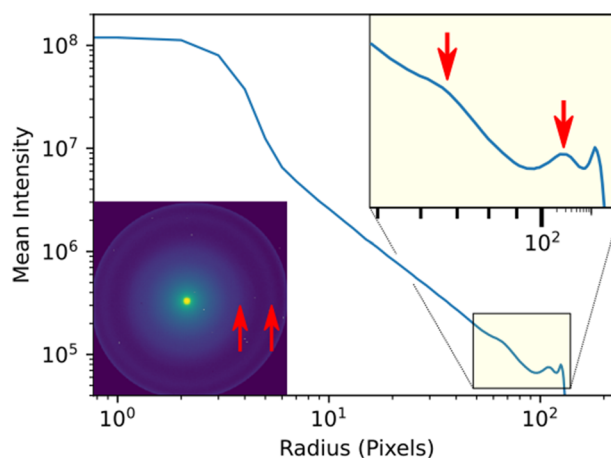

Fig. 2. Radial distribution of scattering intensity from the summed diffraction pattern, shown in the inset, of a 256 × 256 probe position scan of a mouse liver microtome section, calculated using the fpd.fpd_processing.radial_profile function. The pixel spacing was 1.2 nm, the exposure was 4 ms, the semi-convergence angle was 0.436 mrad, the camera length was 180 cm, and the condenser aperture was 10 μm. The arrows mark the location of peaks of intensity.

Once instantiated, virtual images with any possible continuous set of scattering angles may be generated near-instantaneously using the annular_slice method of the class. This calculation time is on the order of milliseconds and, unlike the more flexible synthetic_images approach, the image calculation time does not change with aperture size. A current limitation of this efficiency saving is that center position is taken to be constant across the scan; however, it is not a fundamental one and it could be removed in a future version. As was noted above, in many cases, this limitation is not an issue and, when it is, techniques are available to align the data.

Live feedback on the influence of the virtual detector angles on image contrast may be obtained using the plot method of the VirtualAnnularImages class. Figure 3 shows the two plots created by the plot method (Figs. 3a, 3b) applied to the same data as was used in Figure 2, and images created from the dataset (Figs. 3c–3i). The two vertical features are endoplasmic reticula, each studded with ribosomes. Figure 3a shows a navigation window containing the virtual detector controls, while Figure 3b shows the live image produced by the virtual detector. The navigation window contains a user-supplied navigation image (in this case, a summed diffraction pattern), a measure of the real-space image

contrast as a function of the scattering angle, and controls for the starting, ending and center radii for the virtual detector. The detector position and extent are indicated by the red lines in the two plots, and the radii may be changed via the horizontal slider controls with instant updates to the virtual detector image. The real-space image contrast is measured here by the ratio of the variance in the virtual detector image, σ^2, to the image mean, N. For a flat image with Poissonian noise, σ^2 is purely from the noise and is equal to N, so the contrast measure, σ^2/N, would give a value of 1. For nonflat images with the same noise properties, contrast values greater than 1 indicate image signal. Additional noise in the source images will increase the base level and alter the linearity of the contrast parameter of the virtual detector images but it should remain monotonic with image signal. Similar normalized variance measures are routinely applied in fluctuation electron microscopy (Voyles & Muller, 2002), where the variation is calculated across the azimuth at each radii (scattering angle) in each single image (Hart et al., 2016). The contrast measure used here is different; it is for the image produced from each single pixel width detector, calculated across all radii near instantaneously using the class itself.
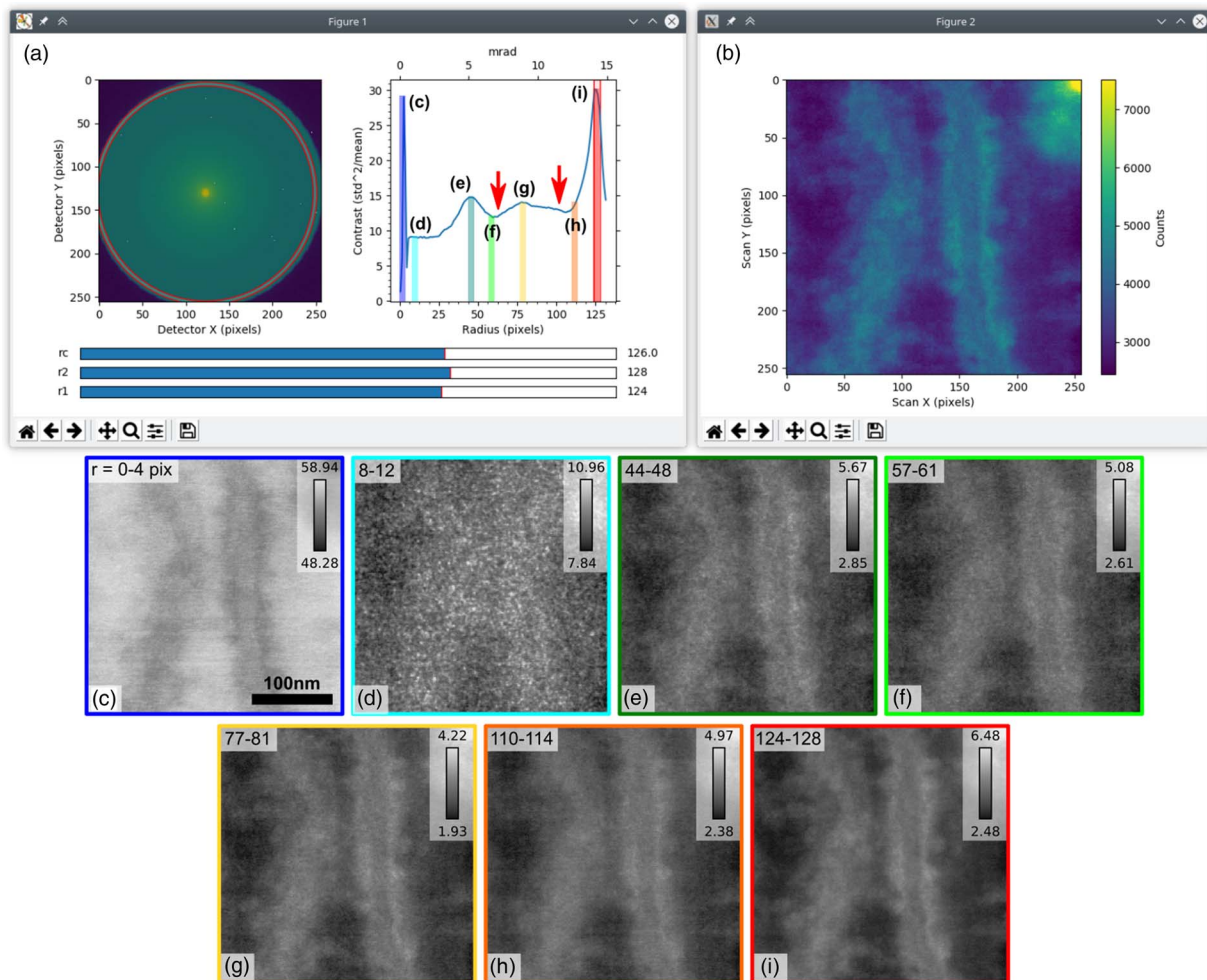

Fig. 3. (**a,b**) Example of the VirtualAnnularImages class, allowing interactive plotting of virtual annular detector images with instantaneous updates. The sample data are the same as that used for Figure 2, a scan of a mouse liver microtome section showing endoplasmic reticula studded with ribosomes. The partial height colored windows and annotations in (**a**) have been added to the regular display to mark the locations of the detector geometries used to generate the virtual aperture images in (**c–i**), from BF to DF. The arrows in (**a**) show the location of the peaks in the radial distribution presented in Figure 2. The intensity ranges in (**c–i**) are set to 0.2–99.8% of the range for each panel, and the color bars in each panel show the counts in thousands.
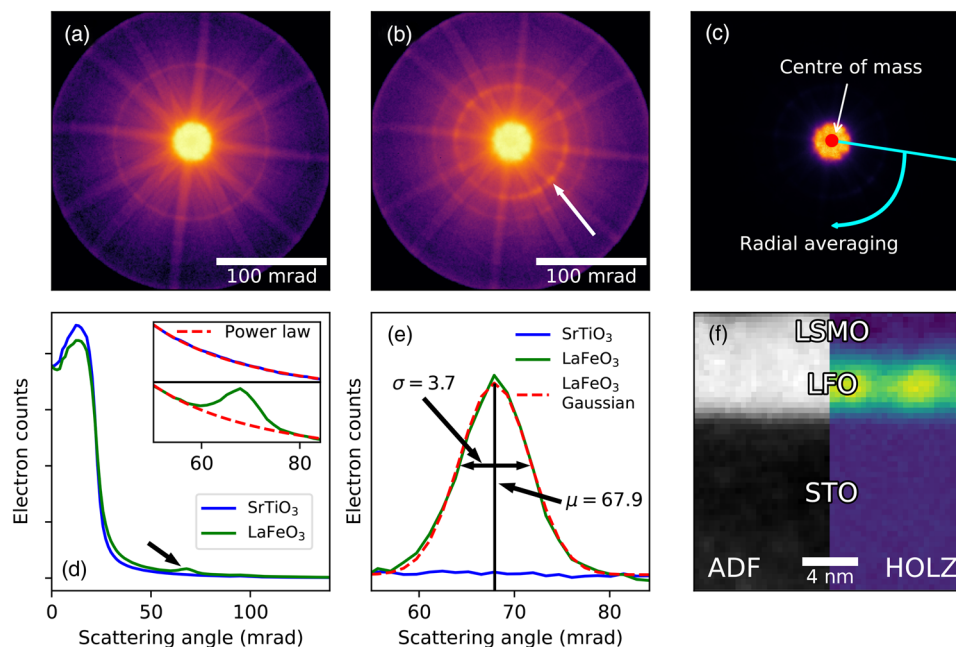
Fig. 4. Processing of HOLZ diffraction rings from a La$_{0.7}$Sr$_{0.3}$MnO$_3$ (LSMO) and LaFeO$_3$ (LFO) bilayer film grown on SrTiO$_3$ along the [111] direction. The data were acquired on a probe corrected JEOL ARM200cF, with an acceleration voltage of 200 kV, a convergence semi-angle of 20.4 mrad, a scan step size of 0.37 nm and with a total scan size of 64 × 64 points. STEM diffraction pattern of the (**a**) STO and (**b**) LFO, plotted on a logarithmic scale, with the electron beam parallel to the [110] direction. LaFeO$_3$ has an extra inner HOLZ ring due to doubling of the unit cell (arrowed). (**c**) Schematic of the processing, first using thresholded center of mass to find the center of the pattern, then radial averaging across the azimuth around this point. (**d**) Radial distribution of (**a**) and (**b**). The arrow shows the inner HOLZ peak from (**b**). The insets to (**d**) show the region around the HOLZ peak, with a power law fitted to the background. (**e**) Region around the HOLZ peak, with the power law from (**d**) subtracted, showing the background is accurately removed, and with a 1D Gaussian profile fitted to the data. Numbers showing the center position and sigma of the Gaussian. (**f**) Result of the processing for the bilayer system, showing the bright LSMO and LFO regions in the virtual ADF (106–163 mrad) image, and bright LFO region in the HOLZ contrast.

The peaks in the scattering angle dependence plot show which angles produce strong contrast, and are a useful aid in determining the range of angles on which to focus. In the case of the biological sample shown, several peaks are visible, but none of them align with those seen in the radial distribution of Figure 2 (these are also marked by red arrows in Fig. 3a), contrary to what one would expect if the local peaks in diffraction intensity produce images with more contrast. Figures 3c–3i show the virtual images produced at the geometries indicated in the annotations, from BF to ADF, at locations between and at the peaks of maximum contrast. The geometries are also indicated in Figure 3a by the partial height windows matching the order of the images and their colored outlines.

Within the DF, there are three peaks of the contrast parameter, at approximately 46, 79, and 126 pixels (Figs. 3e, 3g, 3i), and these are where the images do indeed show the highest contrast and range of spatial frequencies, allowing the ribosomes to be better resolved, with the best contrast produced from the highest of these scattering angles. To estimate the SNR of each individual image, we apply to them the single image autocorrelation method (Thong et al., 2001) implemented in the snr_single_image function of the fpd.utils module. This approach exploits the fact that noise is spatially uncorrelated, whereas the signal *is* correlated, allowing extrapolation of the autocorrelation function to zero image displacement to be used to estimate the noise-free signal power and, thus, the noise power. The optimum extrapolation method depends of the properties of the image and our implementation supports a number of methods. For the data in Figure 3, a linear extrapolation method is adequate and gives

values of 20–37 for the first six images (Figs. 3c–3h), and 71 for the highest scattering angle image (Fig. 3i), confirming that it does indeed have a much higher SNR. These SNR values are the power values with the signal magnitudes taken from the changes in the images, as defined in the original method, and thus give measures of the useful contrast.

For the biological sample used here, there is very little Bragg scattering and the ADF contrast changes in the angular range acquired of up to ~15 mrad is likely to be due to incoherent Rutherford scattering from light elements. The rings visible in the diffraction pattern and its radial distribution are equivalent to d-spacings of approximately 0.35 and 0.22 nm and may be indicative of characteristic bond lengths or periodicities in molecules but, as we have shown, do not always yield images of the greatest useful contrast.

The ability to interrogate and reconstruct STEM images showing different contrast from a single scan, post-acquisition, as outlined above, makes the technique of virtual detectors a very powerful one. In addition, unlike the similar hollow cone technique (Krakow & Howland, 1976; Tsai et al., 2016), only one scan is required to generate many types of contrast, making it dose efficient, which is especially useful in beam-sensitive materials.

While we demonstrate the utility of applying virtual detectors with data from a biological sample, the technique is useful in many other applications, where different navigation images may be more appropriate. For example, in polycrystalline samples, the navigation image may be a diffraction pattern from a specific grain, identified from navigating the dataset with the

DataBrowser class (see Fig. 1) or a similar method. It is then straightforward to set annular apertures that select specific spots, thereby identifying the extent of the grain of interest and the locations of similar grains. More advanced analysis for out-of-plane and in-plane structural analysis is discussed next, in the sections "Higher-order Laue zone analysis" and "Lattice analysis", respectively.

Higher-Order Laue Zone Analysis

The intersection of the Ewald sphere with parallel reciprocal planes of a crystal gives rise to concentric circles of reciprocal space where the Bragg condition is met, leading to constructive interference in the diffraction pattern. The central region of the diffraction pattern is formed by the crystal structure *perpendicular* to the electron beam and is called the zeroth-order Laue zone (ZOLZ). Diffraction spots in HOLZ correspond to intersection of Ewald's sphere with parallel planes of reciprocal lattice points offset *along* the electron trajectory (formally the electron wavevector in reciprocal space) (Emslie, 1934; Jones et al., 1977). Thus, for an on-axis diffraction pattern, periodicity parallel to the beam results in rings of intensity in the diffraction pattern, as shown in Figures 4a and 4b. The radius, r, of the ring in reciprocal space for the first-order Laue zone is given by:

$$r = 2 \sin^{-1}\left(\sqrt{\frac{\lambda}{2d_z}}\right), \tag{1}$$

where λ is the wavelength of the electron and d_z is the lattice period parallel to the electron beam. An equivalent relation was originally derived by Emslie (1934) in a different form that incorporated the camera length. Similar relations may be derived for higher order Laue zone rings. It has long been recognized that high-angle coherently scattered intensity concentrated in HOLZ rings may affect the contrast in HAADF images (Spence et al., 1989).

Because of the dependence of the radius on the out-of-plane structure, the presence of HOLZ rings can be used to obtain information about the structure parallel to the electron beam: the smaller the radius of the Laue zone ring, the larger is the distance between atomic planes parallel to the electron beam. This was originally proposed as a method for studying altered periodicity in dislocation core structures (Spence & Koch, 2001) and was later successfully used to determine the periodicity along the beam direction in sodium cobaltite using a simple thin annular detector setup (Huang et al., 2010).

Alternative approaches to obtaining such information from crystals are to tilt the sample to high angles, or take multiple specimens containing the same kind of feature cut at different angles and reconstruct using either diffraction tomography (Kolb et al., 2007; Mugnaioli et al., 2009) or atomic resolution discrete tomography (MacLaren & Richter, 2009; Van Aert et al., 2011; MacLaren et al., 2013). These approaches are not applicable to all samples (e.g., they could not be used on a dislocation core). They are also time-consuming and either use statistical reconstruction from multiple specimens or multiple sample areas (and so assume that the different areas imaged all contain exactly the same structure, just imaged along different directions), or, in the case of the tomographic approach, use a large radiation dose on one area. Moreover, using a thin annular detector is inflexible and needs camera length tuning to get this to work, thus using a

fast pixelated detector and detecting and processing the HOLZ data after acquisition is far more flexible for a variety of systems.

One example of a class of materials with structural distortions that cause a change in the size of the unit cell is perovskite oxides, which often exhibit a doubling of the unit cell as a result of octahedral tilting (Glazer, 1972). These distortions are important to characterize, as they can heavily influence functional properties such as ferromagnetism and ferroelectricity. This is especially important in thin film systems, as much of the interesting physics resides in the detailed crystallographic structure of these films. One such example is $La_{0.7}Sr_{0.3}MnO_3$ (LSMO) and $LaFeO_3$ (LFO) bilayer films grown on $SrTiO_3$ (STO) (Hallsteinsen et al., 2016; Nord et al., 2019). While previous work has been performed using ABF or BF imaging in STEM (Aso et al., 2013; Wang et al., 2016; Kim et al., 2017; Nord et al., 2017) to characterize the in-plane oxygen atom movements associated with the octahedral tilting, it is also possible that there are atom movements resulting in a periodicity change along the beam direction. Specifically, while researchers in perovskites and related oxides often talk about octahedral tilting, this rarely happens in isolation and is usually associated with the modulation of cation positions, a feature that is likely to result in a significant diffraction signal.

Figure 4a shows the STEM diffraction pattern of the nondistorted cubic STO substrate imaged along the [110] direction, which has one visible Laue zone ring. The diffraction pattern of the LFO film is shown in Figure 4b, which is noncubic due to structural distortions, and is similar to the LSMO pattern except for having an extra Laue zone ring at lower scattering angles (arrowed). The intensity of this additional ring is proportional to the amount of distortion and can, for example, be used to characterize the strength of the atomic movements as a function of position in thin film structures (Nord et al., 2019) as described below. Additionally, a fine structure can appear in the HOLZ rings and they can split into more than one ring, which may reveal additional subtle details about the structure (Peng & Gjønnes, 1989), and this can be recorded in scanned diffraction datasets at a suitable camera length.

The first step in the processing (Fig. 4c) is finding the center position of the diffraction pattern for each probe position. This is necessary due to the fact that the shift and the tilt of the electron beam is not always perfectly separated in the scanning coils, especially for larger scan areas, and there is no de-scan coil setup in the microscope used in this work, resulting in the diffraction pattern center moving as a function of beam position. As discussed in the section "Virtual detector images", finding the center position can be nontrival for high convergence angle diffraction patterns like we have here. However, for these samples, the zone-axis remained sufficiently aligned across the scan area for it to be possible to find accurate center positions using a thresholded and masked center of mass procedure. The mask removes the higher scattering angles, so only the region containing the BF disc is included. Then, the thresholding sets all values below the chosen threshold to zero, and all above it to one. The center of mass calculation is done on this masked and thresholded image, resulting in the center position of the diffraction pattern. For more difficult sample data, acquiring a reference dataset in vacuum and performing the same center of mass calculation on this dataset can often provide accurate center positions. Some alternatives techniques for center finding are demonstrated in the section "Direct beam characterization".

Next, integration is performed along the azimuth for each probe position to generate radial dependencies, as previously

discussed in the section "Virtual detector images", using the aforementioned center positions of the diffraction patterns. This reduces the 4D dataset to 3D: 2 probe dimensions and 1 dimension with intensity as a function of scattering angle, as shown in Figure 4d. By removing one dimension, the data size is reduced by a factor of around the number of pixels on the longest axis of the detector (256, in our case), making it much more manageable and easier to fit into the computer memory. This radial averaging is also useful for data exploration of large datasets like this, due to the much smaller data size, as demonstrated in the section "Virtual detector images". One complication is distortions caused by the projection system, leading to features which should be round becoming elliptical. When doing the radial averaging, this distortion causes a broadening of the HOLZ peak (Figs. 4d, 4e). However, for sufficiently small distortions, this does not heavily affect the total intensity in the peak. Ideally, this should be corrected for by calculating the diffraction distortions from a reference dataset. However, as the focus of this analysis was the *intensity* of the peak, no such corrections were performed here.

After the radial distribution has been calculated, the HOLZ ring is reduced to a HOLZ peak shown by the arrow in Figure 4d. To extract the intensity from only the peak, a power law is simultaneously fitted to both the regions before and after the peak. The fits are shown in the insets in Figure 4d, both for the STO which does not have the extra HOLZ peak, and for LFO which has an extra HOLZ peak at lower scattering angles. While a power law is not the optimal function to fit to this type of background over long distances, it works well for extracting the HOLZ peaks over a relatively short range of scattering angles, as shown by the level background of the corrected profiles in Figure 4e.

After background correction, the profiles may be analyzed by simply summing the intensity in the HOLZ peak to give a map of the level of structural distortion, as shown in Figure 4f. Pairing the HOLZ processing with virtual ADF generated from the same radially averaged data, one can see that the extra HOLZ ring is indeed present only in the LFO layer. More information can also be extracted by fitting a 1D Gaussian to the peak, as shown in Figure 4e, giving information about changes in the scattering angle (center position) and variation in scattering angle (standard deviation). Further information on the material aspects from a detailed analysis of this system are published elsewhere (Nord et al., 2019). This type of analysis is best suited for monocrystalline materials, especially epitaxial thin films and heterostructures, or polycrystalline materials with large grains (e.g., domain-structured functional oxides with different orientations of the principal axes of the cell), as the technique requires the crystal to be imaged along a zone axis. While it may be performed at atomic resolution, this is not necessary, and this technique has the advantage that it yields information about the modulation of atomic positions along a column without requiring atomic resolution, as demonstrated in previous analyses of this effect (Borisevich et al., 2010; Azough et al., 2016).

Lattice Analysis

Analyzing the crystal structure of materials has been a part of electron microscopy since its inception. A number of techniques in TEM can be used for atomic resolution imaging, most frequently phase contrast TEM (high-resolution TEM, HRTEM) and HAADF STEM. Nevertheless, diffraction-based techniques have been the main method for structure solution or analysis

throughout the history of electron microscopy, notwithstanding the advantages of the image-based techniques for giving local crystallographic information, especially where this varies with position such as in thin films, heterostructures, and around defects and precipitates.

Selected area electron diffraction (SAED) can localize the sample region from which data are collected to ~100s nm, much higher than the nanometer or even Ångström length scales that imaging with high-energy electrons can potentially allow. Recently, however, scanned diffraction techniques have been on the increase, especially with the rise of faster electronic detectors (initially CCD/scintillator arrangements, but more recently direct electron detectors). Moreover, the development of SPED (Vincent & Midgley, 1994) has allowed collection of spatially resolved diffraction data down to areas of 2 nm or less with very high precision of the reciprocal lattice parameters.

Convergent beam techniques provide the highest spatial resolution by imaging with a probe formed by a large diffraction-limiting aperture, and it is these techniques that have seen the greatest application in materials science in a number of areas, including grain size and orientation determination, molecular structure solutions, and material strain measurement. Strain is a particularly important property that influences the functionality of a wide range of materials, with perhaps the most notable class of materials being semiconductor devices (Cooper et al., 2016; Bashir et al., 2019).

A number of techniques for strain measurements have been developed (Béché et al., 2013), including dark-field electron holography (DFEH; Hÿtch et al., 2008; Cooper et al., 2009; Béché et al., 2011), NBED (Béché et al., 2009), SPED (also referred to as nanobeam precession electron diffraction (NPED); Rouvière et al., 2013; Midgley & Eggeman, 2015), scanning moiré fringe (SMF) analysis (Su & Zhu, 2010; Naden et al., 2018), HRTEM geometrical phase analysis (GPA; Hÿtch et al., 1998), and atomic column spacing displacement characterization (Nord et al., 2017). Of these, the best fractional strain precision reported is with SPED at 2×10^{-4} (Rouvière et al., 2013); however, very recent work with patterned probes (Guzzinati et al., 2019; Zeltmann et al., 2020) have achieved precisions approaching the same value.

Analysis of such 2D lattice images can be performed with a range of packages, including Atomap, optimized for atomic resolution STEM imaging (Nord et al., 2017); CrysTBox for HRTEM, SAED, and convergent beam electron diffraction (CBED) imaging (Klinger, 2017); library-based approaches to crystal phase and orientation identification (Rauch et al., 2010); and ones recently developed specifically for 4D-STEM (Savitzky et al., 2019; Zeltmann et al., 2020). With the rise in the use of fast pixelated detectors in TEM and STEM and the consequent ability to record large data volumes in short times, the need arises for the ability to analyze large numbers of diffraction patterns accurately and automatically. These include images with point-like lattice vertices, created by imaging under Fraunhofer conditions, and in CBED imaging, where the lattice vertices are discs. In the following sections, we report simple processing methodologies that are applicable to general lattice analysis and, in particular, to diffraction patterns produced by techniques, including CTEM, CBED, NBED, and SPED.

To demonstrate our technique, we apply it to a commercially available crystalline MgO substrate imaged in a JEOL ARM200cF TEM operated at 200 kV in SPED mode using a custom (MacLaren et al., 2020) DigiSTAR precession system from

NanoMEGAS. MgO is a widely studied material (Yang et al., 2006) that is used in magnetic tunnel junctions (Parkin et al., 2004; Zhu & Park, 2006) and is of interest as a room temperature ferromagnetic insulator when strained (Jin et al., 2015). Unstrained MgO has an NaCl-type cubic crystal structure which produces a diffraction pattern with a square projection of the reciprocal space lattice when viewed along the ⟨100⟩ zone axes. The sample was prepared by application of a standard focused ion beam milling procedure (Schaffer et al., 2012) to a multilayered structure. Electron energy-loss spectroscopy measurements in a probe-corrected STEM mode with convergence and collection semi-angles of 29 and 36 mrad, respectively, using a GIF Gatan QuantumER 965 spectrometer, showed that the ratio of thickness, t, to inelastic mean free path, λ, (t/λ) of the sample was 0.59. Using an MgO density of $3.58\ \mathrm{g/cm^2}$ (Egerton, 2011), λ is 137 nm, giving the thickness of our MgO sample as 81 nm, sufficiently thick for there to be significant dynamical effects from multiple elastic scattering.

The standard precession system was modified (MacLaren et al., 2020) to employ a Medipix3 detector instead of a fast video recording of the fluorescent screen using an external CCD camera. This modification enables improved imaging fidelity and efficiency through the higher detective quantum efficiency and the noise-free readout properties of the direct electron counting detector. Compared to static probes, precessing the electron beam reduces dynamical effects by incoherently averaging over a range of diffraction conditions to give a pseudo-kinematical diffraction pattern, resulting in more uniform diffraction pattern discs (Midgley & Eggeman, 2015).

In the experiment, test data was acquired by imaging MgO along the [100] axis under the following conditions: the beam was precessed at an angle of 1.5° over a 10 ms exposure; the microscope was operated in TEM-L mode, with a spot size of 5; the condenser aperture was $20\ \mu\mathrm{m}$; the semi-convergence angle was determined from the MgO lattice to be 1.1 mrad; the camera length was 100 cm; and the scan pixel size was 2.2 nm. No exploration of the acquisition parameter space or subsequent optimization of the parameters was performed.

Correcting for the overestimation of the electron counts recorded by the Medipix3 by a factor of 4 due to scattering of 200 kV primary electrons between pixels [see McMullan et al. (2007) for a discussion], we estimate the dose to be approximately $4.5 \times 10^5\ \mathrm{e^-/nm^2}$. While this is above the low-dose regime needed for most soft or beam-sensitive materials of 10^2–$10^5\ \mathrm{e^-/nm^2}$ (Yakovlev & Libera, 2008), there is great scope to reduce the dose through reduction in beam current or, as pointed out by MacLaren et al. (2020), by increasing the precession rate in other microscopes that support it.

To enable the SPED datasets to be used more easily, the topspin_app5_to_hdf5 function of the fpd.fpd_io module allows conversion of data originally recorded in the flat native NanoMEGAS TopSpin app5 format to the multidimensional HDF5 format outlined in Part I (Nord et al., 2020). Alternatively, the Merlin system (Plackett et al., 2013) through which the Medipix3 data are acquired can be programmed to output the data directly to a raw file, while the acquisition is being performed and controlled by the TopSpin software. These datasets can be processed by the fpd and pixStem libraries in a number of ways, including the previously covered virtual detector imaging and Laue zone analyses, and differential phase contrast (DPC) analyses for field mapping, which will be covered in Part III. Indeed, the benefits of SPED brought about by averaging over a range of diffraction conditions has great potential to also improve DPC imaging, as others have very recently also noted and investigated (Mawson et al., 2020).

Machine learning (ML) approaches to structure analysis have recently been applied to SPED data (Martineau et al., 2019). In contrast, the methodology described below is a "bottom up" one that optimizes precision without knowledge of the context of each image. The processing methodology is modular in design, allowing modification or addition of processing steps, as needed for the specific application and, indeed, could be complemented by ML approaches. Below, we outline the main steps without going into the many options provided to tailor the analysis to the sample data and, subsequently, assess the precision of the technique applied to this dataset.

The general process can be divided into the four main stages of (i) direct beam detection and characterization, (ii) feature extraction and filtering, (iii) lattice parameter estimation, and (iv) synthetic lattice inlier detection and fitting. These steps are outlined in Figures 5–8, which include slightly modified versions of the plots optionally produced by the relevant analysis functions applied to the first diffraction pattern of the MgO SPED scan. Each figure will be discussed in turn; the final results of the analysis are shown in Figure 9 and will be discussed later.

Direct Beam Characterization

The first stage in the process is to find the direct beam position and size. Figure 5a shows three images produced in this analysis. The first shows the original image, the second shows the image processed with Canny edge detection, and the third shows the results of a circular Hough transform (Gouillart et al., 2016), where the red circle represents the position and size of the detected disc. This transform generates a 2D Hough space for each radius, with values proportional to the correlation of the edge image and the nominal circle centered at each possible coordinate (the dimensions of the space). The center of the optimized circle is obtained with sub-pixel resolution by fitting 2D Gaussians to each peak in Hough space, or the images may be upscaled. If the lattice vertices are point-like, then the size of the circle represents an effective radius. If the vertices are discs, such as those produced by a convergent beam (or are of any other shape), then a template of the beam image may optionally be formed for use in edge-filtered cross-correlation in order to estimate the position of all vertices in the next steps. As several others have noted, procedures of this type can improve the accuracy of the image registration results by relying more on feature shapes than image intensities (Schaffer et al., 2004; Krajnak et al., 2016; Pekin et al., 2017). An important parameter in the analysis is the disc edge profile, and this can be estimated by converting the central beam image to polar coordinates, as shown in the inset in Figure 5b, and then fitting error functions to the edge region, as shown in the main panel. Through this process, the disc diameter and a new center position are also obtained from fitting to the angle dependence of the extracted error curve center positions. Finally, a filtered reference image may be created from the central beam image, using the extracted disc edge width, as shown in Figure 5c. This step further reduces the influence of intensity diffracted from the BF disc, which will not be uniform across all scan points and within all diffraction discs. These nonuniformities in intensity are most easily visible as texture in the rightmost image in Figure 5c, which shows the difference between the real disc and the uniform processed one.
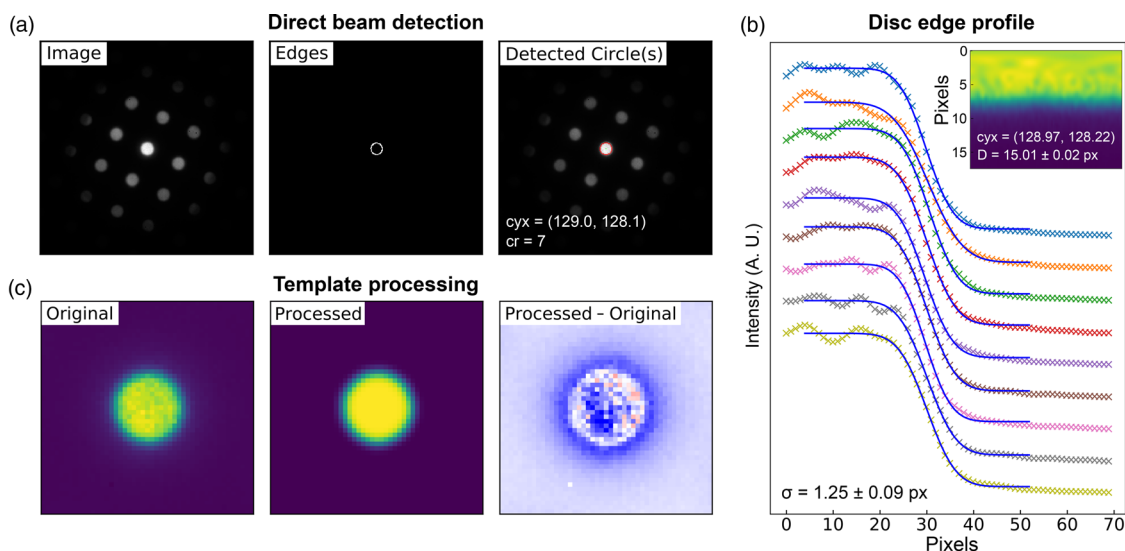
Fig. 5. Lattice analysis stage 1: (**a**) finding the direct beam, (**b**) characterization of the disc edge properties, and (**c**) forming a template using, respectively, the find_circ_centre, disc_edge_properties, and make_ref_im functions from the fpd.fpd_processing module. The lines in (**b**) are error functions fitted to the data (symbols) taken from the direct beam image converted to polar coordinates, plotted in the inset. The inset also shows the measured diameter of the disc, D, and the optimized center coordinates of the extracted edges. The error in the latter is 0.01 pixels.

This first stage need only be done once for each dataset, unless the direct beam moves a significant distance during the scan. If this is the case, the direct beam position may be estimated at each scan position using the technique described above, by center of mass as discussed in the section "Higher-order Laue zone analysis", or by any other means.

Vertex Identification

The second stage in the analysis is to find the position and size of all the potential lattice vertices and then, optionally, to filter these so that only those with inversion symmetry are retained. The results of one way of processing the images to extract potential lattice points are shown in Figure 6. The first step is to perform Laplacian of Gaussian (LoG) blob detection with a supplied radius range, based on the direct beam radius (or effective radius

for Fraunhofer imaging conditions). The detected features are shown in Figure 6a by the red circles. If the lattice vertices are point-like, 2D Gaussians may optionally be fitted to extract sub-pixel peak locations. For nonpoint-like or disc-shaped vertices, pixel-resolution edge-filtered cross-correlation (CC) can be used with the template image of the direct beam produced in the first stage. This result is shown in Figure 6b, where the CC peaks, marked by red crosses, indicate the location of maximum correlation, corresponding to the identified vertex locations. By the nature of lattices, cross-correlation can give rise to many false peaks, as is visible in the figure. Therefore, these are filtered by the LoG data which is in general more reliable but which has a higher noise level. Specifically, only those points located within some distance of the LoG vertex coordinates are kept. The peaks retained in this step are marked by black circles. Also at this stage, 2D Gaussian functions can be optionally fitted to the
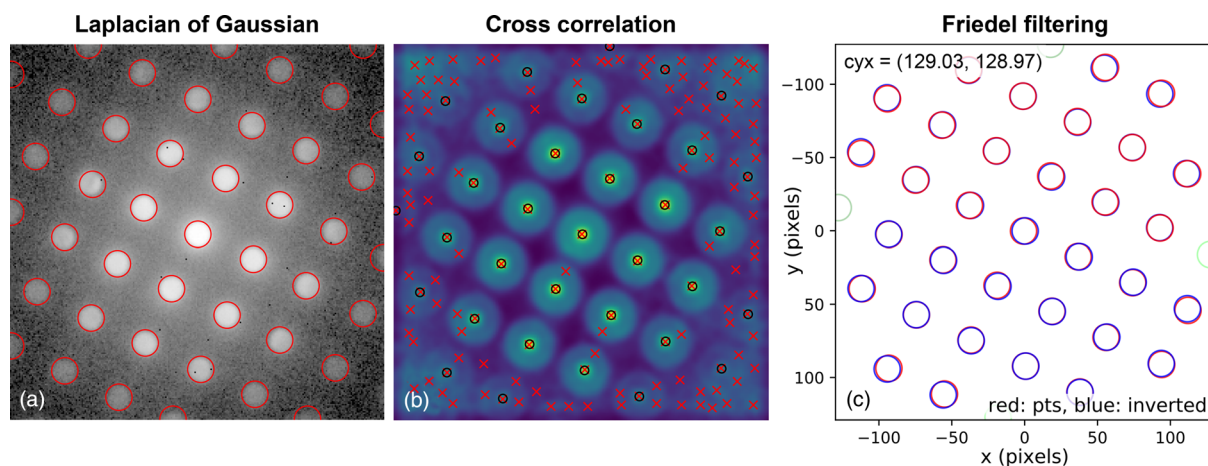

Fig. 6. Lattice analysis stage 2: extracting features and filtering them by inversion symmetry using the blob_log_detect and friedel_filter functions of the fpd.tem_tools module. The red circles in (**a**) show the blobs detected using LoG applied to a logarithmically scaled image. (**b**) The results of edge-filtered cross-correlation with a template image, with red crosses indicating peak locations and black circles those filtered for the next step by the LoG data. Green circles in (**c**) show the discs removed by "Friedel" filtering (see main text for details).

(a)

Vector combinations
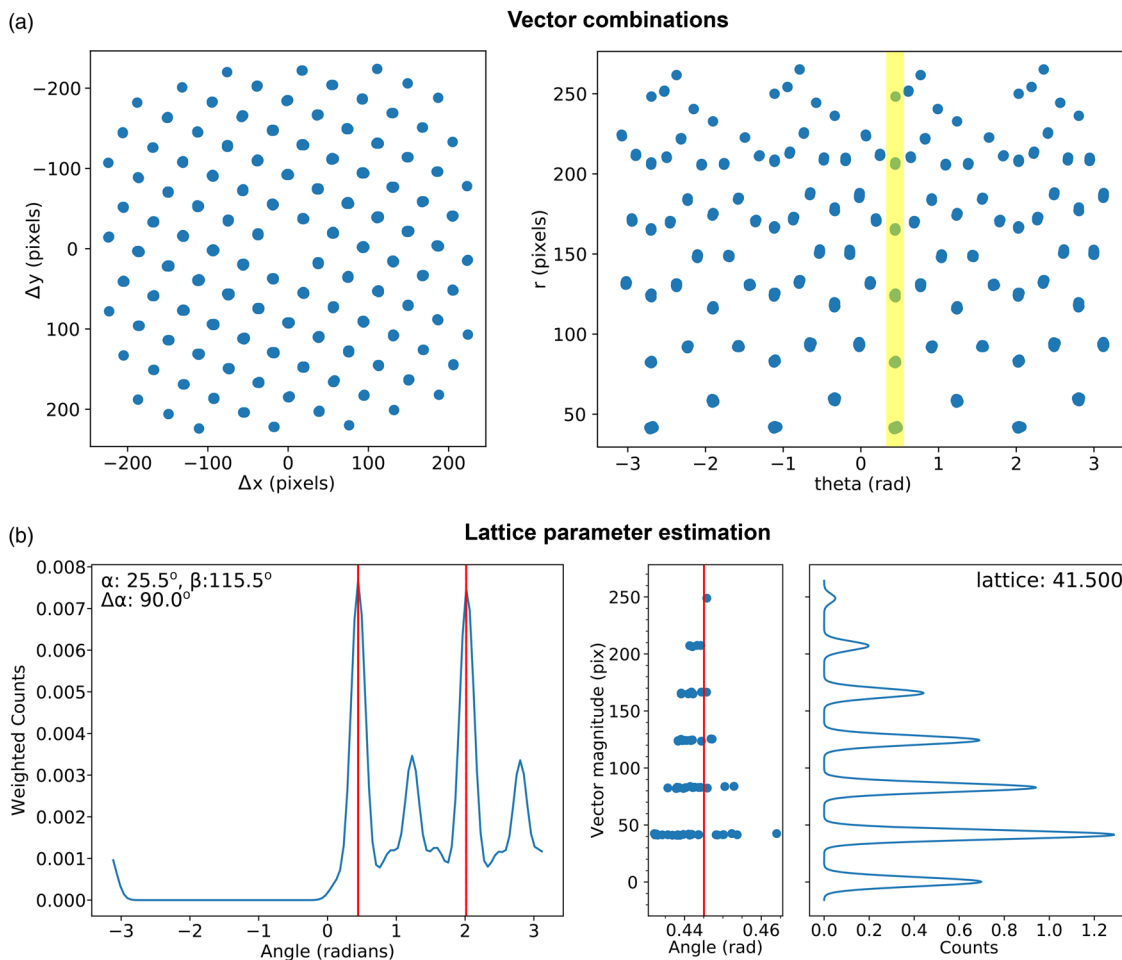

(b)

Lattice parameter estimation

Fig. 7. Lattice analysis stage 3: lattice parameter estimation. (**a**) Lattice vector combinations in Cartesian (left) and polar (right) coordinates. (**b**) Lattice angle (left) and magnitude (right) estimates (only one of the two lattice vector magnitude plots is shown).
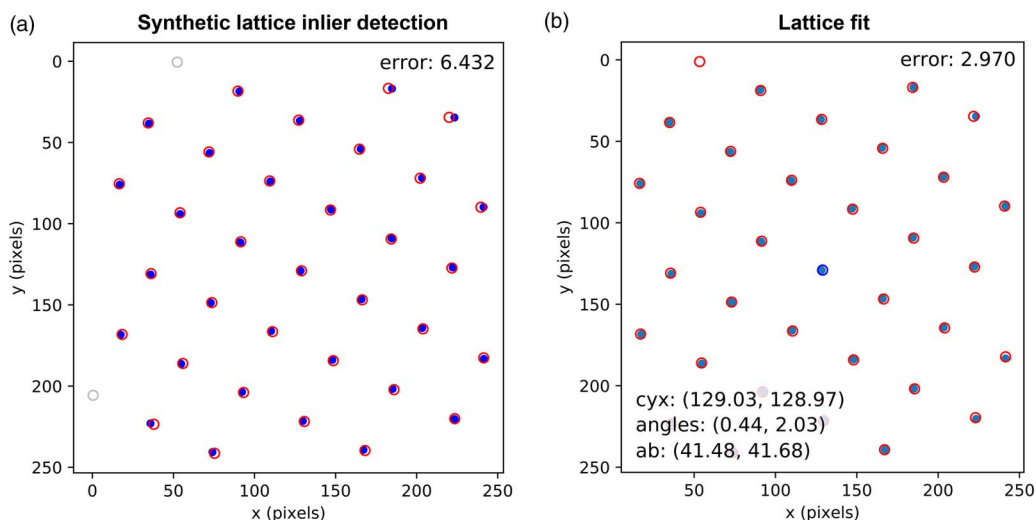

Fig. 8. Lattice analysis stage 4: lattice parameter extraction. (**a**) Analysis data (solid symbols) filtered by synthetic lattice (open symbols) inlier detection and (**b**) the lattice optimized through least squares fitting, where the total error between the two lattices (indicated in the annotations) is minimized. The blue symbol in (**b**) shows the direct beam location. Outliers would be marked by crosses; however, all data are inliers in this case.

vertices to extract the peak locations with sub-pixel accuracy, as was done here. A comparison of the different analysis levels will be given in the section "Analysis precision", with reference to Figure 10.

Next, the data are filtered according to Friedel's law (Friedel, 1913), removing points that, within some relative and absolute error, have no equivalent point at the location inverted about the center coordinate. Most of the potential vertices are retained in this example, as shown in Figure 6c, with only those at the edges being removed (shown in green). In this step, the center coordinate may be updated based on the systematic differences in positions of the pairs of diffraction spots, which allows small changes in the position of the direct beam to be accounted for. In this case, the (c_y, c_x) value was updated to (129.03, 128.97), as shown in the figure annotations.

Lattice Estimation

The previous two stages only extract the positions and sizes of potential lattice vertices and assume nothing of the relationship between these positions, the lattice properties. In the next two stages, the lattice properties are estimated and then optimized from the filtered vertex locations, as shown in Figures 7 and 8, respectively.

There are many potential ways of estimating lattice parameters and we discuss two of these. First, we use a simple statistical approach, as shown in Figure 7. To improve the statistics, all combinations of lattice vectors are computed (left of Fig. 7a) and converted into polar coordinates (right of Fig. 7a). Next, a histogram is generated from the polar data (left of Fig. 7b), with optional $1/r^2$ weighting applied to increase the significance of nearer neighbor data, where r is the Euclidean distance. An expected symmetry of the lattice may be provided at this stage to further improve the analysis. Two peaks are identified from this histogram, marked by the red lines, corresponding to two lattice vector angles. The polar data are then sliced at these angles and a histogram generated (right of Fig. 7b); the equivalent location is marked by a yellow band in the right-hand panel of Figure 7a. The histogram is processed using peak detection or Fourier analysis to extract the lattice magnitudes along these directions, yielding a complete estimate of the lattice parameters.

This simple approach to lattice estimation works best when the sample is on or near a low-order zone axis, and there is a single crystallographic phase within the beam diameter. When multiple lattices are present within a single image, alternative methods to lattice parameter estimation such as clustering and inlier detection may prove to be useful, and can easily be incorporated in the presented processing methodology without modification of the previous or next steps.

The lattice_from_inlier function of the fpd.tem_tools module provides an alternative method of lattice estimation that requires less configuration. This function generates lattices for all combinations of two of each of the first n potential lattice vertex coordinates in distance from the direct beam position, and returns the lattice parameters with the maximum number of inliers to the synthetic lattice, as determined by comparison of the Euclidean distance between matched vertices to a supplied threshold. When multiple lattices have the same number of inliers, one is selected by a user-supplied criterion which, by default, is set to the maximum of the geometrical mean of the lattice parameter magnitudes. This reduces the chances of integer divisions of the lattice parameters being found. As a consequence of this

approach, this function typically returns different equivalent lattice parameters across a uniform material. The parameters may be homogenized using the lattice_resolver function described below. Compared to the statistical approach used above with reference to Figure 7, this approach is, in general, more robust against additional spots being present in the image, less robust against there being many missing spots, and produces a slightly less precise but still perfectly usable initial estimate of the lattice.

Lattice Optimization

In the fourth and final step, using an estimate of the lattice magnitudes and angles (generated by any method), a synthetic lattice is created and data inliers to the model are selected. This synthetic lattice and experimental data are shown as solid and open symbols in Figure 8a. Next, the lattice is fitted to the inliers using least squares optimization, resulting in the final lattice shown in Figure 8b. Constraints and bounds on and between lattice parameters can be applied during the fitting process, but none were used for this test data.

The modular design of the lattice analysis makes it easy to assess each processing step, allowing optimization of the analysis parameters, and to customize the processing by inserting additional steps or removing existing ones. Once the processing conditions are established on a test image or images, the analysis may be applied to many images in parallel by passing a user created function to the map_image_function function of the fpd.fpd_processing module.

The four stages outlined above often result in a single unit cell per material. However, as mentioned earlier, it is possible for the particular unit cell returned within a material to change between a number of equivalent unit cells. To condition the data, as an additional step, the lattice parameters extracted can be resolved in specific directions using the lattice_resolver function of the fpd.tem_tools module. This generates a small synthetic lattice using the extracted lattice parameters and then identifies the new lattice vectors by finding those lattice points at angles closest to user specified values. This is especially useful when mapping lattice parameters across epitaxial interfaces between differing materials, where in- and out-of-plane strain may be of particular interest. Potential alternatives that are likely to meet with success include clustering and ML approaches applied to the extracted lattice parameters or parameters derived therefrom. Compared to applying ML approaches to images directly, applying them to the extracted data would vastly reduce the size of the data to be processed (by a factor of several thousands), and also automatically accounts for any de-scan in the measurement as the calculated basis vectors are insensitive to the pattern center position.

MgO Results

Figure 9 shows the results of analyzing the complete MgO SPED dataset, using a reference image for cross-correlation and applying sub-pixel peak fitting. A comparison of the different levels of processing applied the same data is shown in Figure 10 and will be discussed in the section "Analysis precision".

Figures 9a and 9b show the real and reciprocal sum images for the dataset. The reciprocal space coordinate system is defined in Figure 9c. The origin is located at the top left corner, the coordinates of the direct beam are (c_y, c_x), and the lattice parameters are defined by the magnitudes, a and b, and the angles, α_1 and α_2.
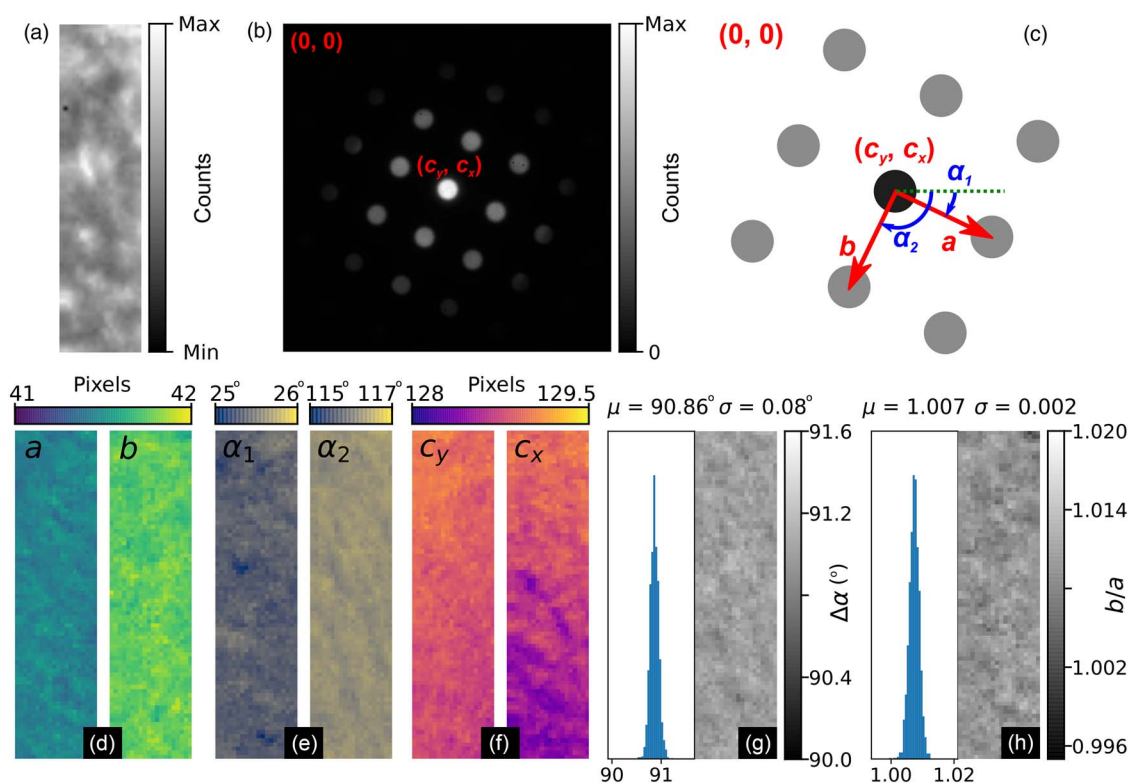
Fig. 9. MgO SPED data analysis results. (**a**) Real and (**b**) reciprocal space sum images for an 80 × 20 pixel scan (176 × 44 nm, pixel size: 2.2 nm) using a 256 × 256 pixel Medipix3 detector. The reciprocal space coordinate system is shown in the annotations in (**b**) and in the schematic in (**c**). (**d**–**f**) Spatially resolved lattice parameters and (**g**) lattice angle delta and (**h**) lattice parameter magnitude ratio histograms and maps.

The six lattice parameters extracted from the data are shown in Figures 9d–9f, with the angles between lattice vectors, $\Delta\alpha$, and the ratio of the lattice magnitudes, b/a, shown in Figures 9g and 9h, respectively.

The mean angle between lattice vectors in the SPED data is 90.86 ± 0.08° (0.09%), and the mean ratio of the lattice vector magnitudes is 1.007 ± 0.002 (0.16%), making for a slightly skewed lattice, and is evidence of strain in our multilayer sample. By comparison, the equivalent lattice vector magnitude ratio for the nanodiffraction MgO data in Figure 1 from a different sample and analyzed by the same method was 1.020 ± 0.006 (0.6%). Although both datasets are influenced by nonuniformities in the MgO, the difference in the variance of the data will reflect to some extent the improvements obtainable by precessing the electron beam.

Care in interpreting the results must be taken in order to rule out the influence of image distortion (Mahr et al., 2019). This may be done in a number of ways, including: recording two or more scans at different sample rotations; recording a reference lattice against which strain parameters (Rouvière & Sarigiannidou, 2005) can be calculated; or by correcting the lattice vertices by the application of an affine or other transformation to the vertex coordinates from a suitable calibration.

The center coordinates approximately describe planes due to de-scan (as discussed before). However, in samples supporting magnetic or electric fields or exhibiting mean inner potential changes, the beam shifts extracted in these parameters can give useful DPC contrast. Aspects of DPC analysis will be discussed in Part III of this work.

Analysis Precision

Figure 10 shows how the cumulative addition of edge-filtered cross-correlation (Fig. 10b) and sub-pixel peak finding (Fig. 10c) each improve upon the LoG analysis (Fig. 10a). The third row contains the same data as shown in Figures 9d–9f, but the color map scale here is maximized for each panel, with the limits of the ranges shown on each color bar.

The same main features are present in the LoG (first row) data as are in the third row data, but the noise level is much higher and there is a degree of digitization, obscuring the more subtle features. Edge-filtered cross-correlation gives a large improvement, with the spread in parameter values generally reduced and the results are less digitized. This is because the processing technique reduces the influence of intensity variations within the discs and because it is inherently more sensitive to the location of the discs. This will be especially true when the beam is not precessed, such as in CBED or NBED patterns where nonuniformities in the disc intensities can be significantly larger. The improvement over the LoG plus CC results (Fig. 10b) obtained by the addition of peak fitting, shown in Figure 10c, are mainly in the deep sub-pixel range, with a reduction in the random noise level.

To estimate the SNR in these analyses, we applied the same single image autocorrelation power SNR estimation (Thong et al., 2001) methodology used in the section "Virtual detector images" to the data shown in Figure 10. Here, a second-order polynomial extrapolation of the autocorrelation function was used after de-trending the images with a plane fit, and the signal level was taken from the known zero rather than using the
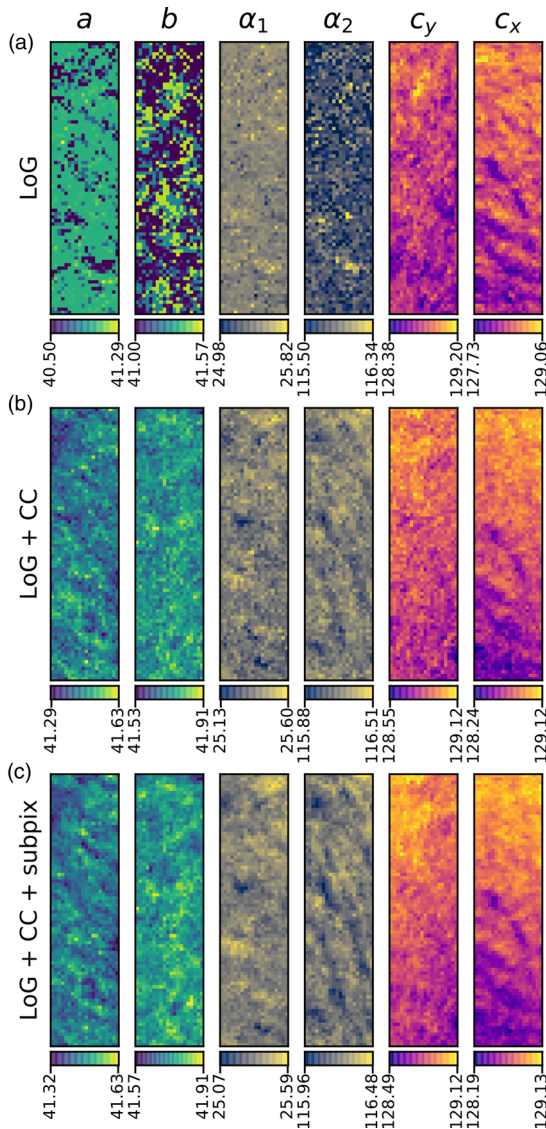
Fig. 10. Comparison of analysis of the MgO SPED data used in Figure 9 with three different approaches (in rows) of generating the diffraction disc centers: (**a**) Laplacian of Gaussian (LoG), (**b**) LoG plus cross-correlation (CC), and (**c**) LoG plus CC plus sub-pixel peak fitting. The units of the data are the same as in Figure 9, pixels and degrees. The color bar ranges are matched to the data range of each panel.

variation in the images. The resulting power SNR values for the *a* parameters are 64 and 70 dB for the LoG plus CC and LoG plus CC with sub-pixel peak fitting analyses, respectively (the results from the other parameters are similar). The pixelation in the LoG data means we cannot easily extract an accurate SNR estimate for this dataset using this method. However, an estimate of the SNR may be made using the LoG plus CC with sub-pixel peak fitting analysis data as a reference, and this yields a value of 47 dB for the LoG analysis.

The smooth, continuously varying features of our test MgO data means we cannot interpret the level of precision of our approach from the standard deviation of the parameters. However, because we have extracted an estimate of the SNR ratio, we may use this parameter to estimate the fractional precision as $1/\sqrt{\text{SNR}}$. The results give a fractional precision of 3.1 × 10^{-4} (0.03%) in the LoG plus CC with sub-pixel peak fitting

analysis which, with an *a* parameter of 41.5 pixels, corresponds to a fractional precision of 0.01 pixels of the detector. The fractional precision of the LoG plus CC analysis is about half as good at 6.6×10^{-4}. The theoretical full-width half-maximum (FWHM) spatial resolution of the Airy probe in our measurements, estimated using the airy_fwhm function of the fpd.tem_tools module, is 1.1 nm, which is consistent with direct imaging of the probe. However, at the precession angle used in the experiment (1.5°) and for our sample thickness (81 nm), the center of the probe will describe circles of a diameter of 2.1 nm at the surfaces of the sample and, thus, the actual spatial resolution of the measurement will be poorer than that defined by the static probe. In fact, ignoring the increased weighting of the sampling of the specimen in its vertical center, the FWHM defined resolution of the probe will be given by the diameter of the circle described by the beam. Importantly, if our 2.2 nm scan pixel data were oversampled, our precision estimates may be overly optimistic. As a simple check of this, we repeated the analysis using only every other pixel, giving an interpixel spacing of 4.4 nm, and found the SNR values of the LoG plus CC and LoG pluts CC with sub-pixel peak fitting analyses decreased to 62 and 64 dB, corresponding to precision values of 8.2×10^{-4} (0.08%) and 6.0×10^{-4} (0.06%), respectively.

Using the alternative approach of wavelet-based Gaussian noise estimation (Donoho & Johnstone, 1994; Gouillart et al., 2016) yielded very similar fractional precision results to those from the previous analysis with a larger inter-pixel spacing: 7.9 × 10^{-4} for the LoG plus CC analysis, and 5.6×10^{-4} for the LoG plus CC with sub-pixel peak fitting approach, giving credence to the results. Rebinning the data by a factor of 2 along each axis before performing the SNR calculations gives precision values around 4.5×10^{-4} for both methods. The improvement over using every other pixel is partly due to the increased dose at the same spatial resolution.

Further work is required to determine the optimum acquisition parameters for the accuracy of the analysis. For example, the noise-free readout of the Medipix3 and the relatively low number pixel count of 256×256 for a single die (versus several 1,000 s for a typical CCD detector), may mean that better results would be obtained by increasing the camera length so that only the lowest order spots are imaged. Alternatively, different DEDs with higher pixel counts may be used, or more pixels may be added to a Medipix3 detector by tiling the detector chips. The Medipix3 family of detectors are 3-side buttable (Ballabriga et al., 2013), and thus may be tiled in $2 \times N$ arrays (Bücker et al., 2020), with larger pixels at the joints which must be accounted for. Furthermore, employment of the through-silicon *via* feature of the Medipix3 family can allow tiling on all four sides of the sensor in even larger arrays (Tick & Campbell, 2011; Ponchut et al., 2015) with minimal dead areas for the readout circuitry.

At the small level of material distortion found here, the fractional precision of the lattice parameters is the same as that of the associated strain parameters. Very recent reports of strain measurements using patterned probes in the literature (Guzzinati et al., 2019; Zeltmann et al., 2020) approach the best values reported from standard Airy probes in SPED acquisitions (Rouvière et al., 2013), but with potential benefits of improved dose efficiency. The fractional precision of 6×10^{-4} obtained here with a DED is approximately 3× higher than the value from the latter (of $\leq 2 \times 10^{-4}$) with a similar spatial resolution, but using exposures 100× smaller and a detector with 64× fewer

pixels (256 × 256 DED versus a 2k × 2k CCD). Our results correspond to an approximately 3× higher sub-pixel precision. We cannot compare the beam dose between the two experiments, but we note that the aperture used in our experiment was 2–4× the radius of that used in the referenced work which, taking account of the different exposure times, would give a 25–12.5× reduction in dose in our experiment with a DED *if* the emission currents were the same. With the noise-free readout of the Medipix3 and the ability to operate in continuous mode, where every individual electron is counted with no gaps, there is great scope for further reducing the sample dose in SPED for use with beam-sensitive materials (MacLaren et al., 2020).

Method Applicability and Efficiency

We demonstrate the flexibility of our simple lattice analysis method by applying it to the four synthetic lattice images shown in Figure 11. These images were generated by sub-pixel Fourier shifting 2D Gaussian spots of peak intensity 1,024 and include Poissonian noise. The lattices have (a) square, (b) rectangular, (c) hexagonal, and (d) oblique unit cells. These cover four of the five shapes available in 2D space; the remaining one being the rhombic lattice which only differs from the oblique one by having equal lattice vector magnitudes. The extracted parameters are shown in the insets of Figure 11, with the nominal values in parentheses. The agreement between the numbers in this somewhat idealized case is less than around 0.006 pixels and 0.006°, corresponding to an average error of 0.003%, and gives some idea of the accuracy of the algorithm at this electron dose. Ultimately, the final noise and precision achievable will depend on the sample; the imaging conditions, including the dose and beam shape; the characteristics of the detector and its number of pixels; and the parameters used in the analysis.

Despite the widely varying lattice size, aspect ratios, and symmetries, the *same* analysis parameters were used to extract the lattice properties across all four datasets in Figure 11 at the same time, demonstrating the ability of the approach to characterize multiple materials in a dataset *without* prior knowledge or constraints on what they are. An example of the application of this feature is given in the section "Data visualization", where multiple phases were identified from experimental nanobeam diffraction data without modification of the analysis parameters. However, many analysis parameters may be specified to improve the parameter extraction in real data, including the ability to apply arbitrary constraints on the lattice parameters, potentially allowing more accurate data extraction for known materials by removal of unneeded degrees of freedom.

In data with only Poissonian noise, the relative error of the extracted parameters, $\sigma_v/\bar{v}$, will generally follow:

$$\frac{\sigma_v}{\bar{v}} = \frac{m}{\sqrt{N}}, \qquad (2)$$

where N is the total counts and m is a constant that varies with the analysis method used and the properties of the source data, such as the intensity distribution within each lattice spot or disc. To estimate m in our approach, we calculated the ratio of the standard deviation to the mean of the smaller lattice parameter, a of the synthetic lattice in Figure 11d across approximately 1,000 images with different Poisson noise, as a function of the total dose, N. This data is plotted as red circles in Figure 12 for the LoG plus sub-pixel peak fitting analysis of each image, along with a fit to the data of the above equation (black line). The value of m is 0.018 across the dose range investigated of 273 to 140k counts (see inset for examples of a typical lattice vertex as a function of dose). By comparison, the same accuracy data for an analysis by fitting a 2D Gaussian function to each of just two of the lattice vertices is shown as blue crosses and has a much higher m value of 0.178. The equivalent m value obtained by reducing N to the counts in 2 of the 21 vertices is 0.039, around twice that found by using all vertices, showing the benefit of using all available counts.

This approach of fitting to the complete lattice makes use of every vertex and all available signal, optimizing the accuracy for a given dose, as well as extracting more information in the form of the other lattice parameters. The imperfections in real data
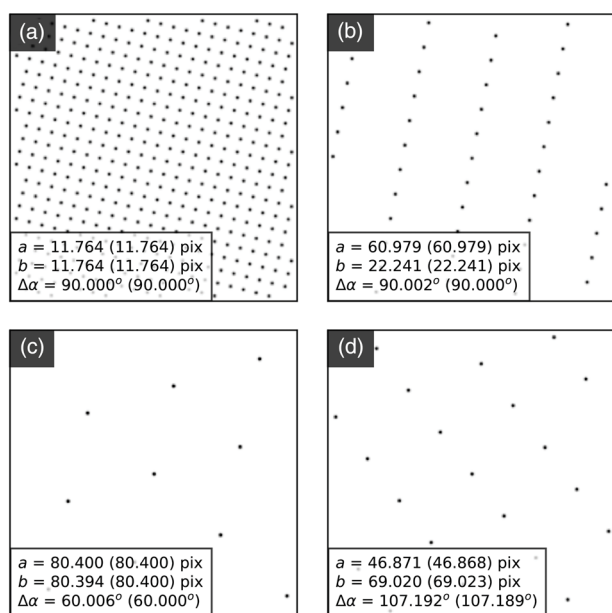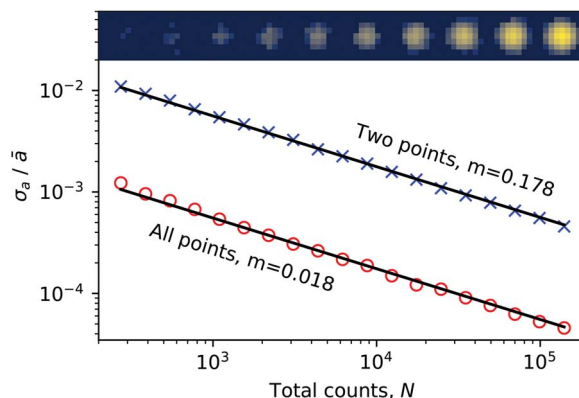

Fig. 11. Four different synthetic lattices and, in the inset text, the results of their analysis using identical processing parameters. The nominal values are given in parentheses. All spots were of equal counts before application of Poissonian noise.

a = 11.764 (11.764) pix
b = 11.764 (11.764) pix
Δα = 90.000° (90.000°)

a = 60.979 (60.979) pix
b = 22.241 (22.241) pix
Δα = 90.002° (90.000°)

a = 80.400 (80.400) pix
b = 80.394 (80.400) pix
Δα = 60.006° (60.000°)

a = 46.871 (46.868) pix
b = 69.020 (69.023) pix
Δα = 107.192° (107.189°)

Fig. 12. Relative error of the smaller lattice parameter magnitude, *a*, of the lattice in Figure 11d as a function of dose. The data were analyzed using LoG feature detection plus sub-pixel peak fitting, using the full lattice (all points) and the central spot and one other point (two points). The inset shows example images of every other spot as a function of dose, on a logarithmic intensity scale.

will mean that the precision of our approach applied to such data will most likely be significantly lower than the characteristic values reported in Figure 12, and will be dependent on the properties of the sample and the detector, but the benefits of our approach will remain; specifically, that by using all vertices, the signal is used optimally and the effect of systematic variations in vertex properties, such as background variations, will be reduced.

Regardless of the source of the data, once the shape of a lattice has been characterized, it is then possible to extract additional information. For diffraction data, this includes strain mapping, the number and spatial extent of different phases or grains (see Fig. 1), and the spot intensities for structure factor characterization (Midgley & Eggeman, 2015) and structure determination (Mugnaioli et al., 2009; Clabbers et al., 2017; Bücker et al., 2020), and for the analysis of noncentrosymmetric crystals. Intensity mapping may be done by generating a synthetic aperture based on the extracted lattice parameters and then applying it to the data, as discussed in the section "Virtual detector images", or, for point-like spots, fitting a 2D Gaussian to the peak regions using the functions provided in the fpd.utils module, which are also used in the sub-pixel peak finding described above, yielding details of the peak properties.

Summary

In this work, we discussed the key issues around post-acquisition analysis of data from fast pixelated detectors and presented software libraries to allow efficient data processing and visualization. We provided examples of their use across a number of applications in the area of structural characterization, including the techniques of virtual detector imaging for BF and DF imaging, HOLZ analysis for extraction of structural information along the path of the beam, and nanobeam and scanning precession electron diffraction for lattice parameter determination and strain analysis.

While the data analysis algorithms and libraries presented are applicable to data from any detector, the examples provided show that highly dose efficient active pixel direct electron detectors such as the Medipix3 perform excellently as universal STEM detectors, despite their relatively low pixel counts. Indeed, we have demonstrated a nanoscale lattice parameter mapping in an SPED mode with a fractional precision $\leq 6 \times 10^{-4}$, approaching the best values reported in the literature. Furthermore, in addition to being of use as a regular STEM detector, with every electron being recorded noise-free, there is excellent prospects for the application of the detector to the characterization of beam-sensitive materials.

The software packages presented are hosted in public repositories (fpd devs, 2015; pixStem devs, 2015; fpd demos devs, 2018), are under active development and contain many more features than are covered in this part or in Part 1 (Nord et al., 2020) of this work. These packages are provided under an open source licence, allowing full transparency of the algorithms implemented and for them to be continually improved upon by the community.

Part III of this work will cover post-acquisition processing and visualization of data from fast pixelated detectors for differential phase contrast imaging.

Data and Software Availability. The complete source data and scripts to analyze the data and produce the results presented here are publicly available at http://dx.doi.org/10.5281/zenodo.3903517 (Paterson et al., 2020). The software presented in this work is available in public Git repositories (fpd devs, 2015; pixStem devs, 2015; fpd demos devs, 2018).

References

Aso R, Kan D, Shimakawa Y & Kurata H (2013). Atomic level observation of octahedral distortions at the perovskite oxide heterointerface. *Sci Rep* **3**, 2214. doi:10.1038/srep02214.

Azough F, Cernik RJ, Schaffer B, Kepaptsoglou D, Ramasse QM, Bigatti M, Ali A, MacLaren I, Barthel J, Molinari M, Baran JD, Parker SC & Freer R (2016). Tungsten bronze barium neodymium titanate ($Ba_{6-3n}Nd_{8+2n}Ti_{18-O_{54}}$): An intrinsic nanostructured material and its defect distribution. *Inorg Chem* **55**, 3338–3350. doi:10.1021/acs.inorgchem.5b02594.

Ballabriga R, Alozy J, Blaj G, Campbell M, Fiederle M, Frojdh E, Heijne EHM, Llopart X, Pichotka M, Procz S, Tlustos L & Wong W (2013). The Medipix3RX: A high resolution, zero dead-time pixel detector readout chip allowing spectroscopic imaging. *J Instrum* **8**, C02016–C02016. doi:10.1088/1748-0221/8/02/c02016.

Bashir A, Millar RW, Gallacher K, Paul DJ, Darbal AD, Stroud R, Ballabio A, Frigerio J, Isella G & MacLaren I (2019). Strain analysis of a Ge micro disk using precession electron diffraction. *J Appl Phys* **126**, 235701. doi:10.1063/1.5113761.

Béché A, Rouvière J, Barnes J & Cooper D (2011). Dark field electron holography for strain measurement. *Ultramicroscopy* **111**, 227–238. doi:10.1016/j.ultramic.2010.11.030.

Béché A, Rouvière JL, Barnes JP & Cooper D (2013). Strain measurement at the nanoscale: Comparison between convergent beam electron diffraction, nano-beam electron diffraction, high resolution imaging and dark field electron holography. *Ultramicroscopy* **131**, 10–23. doi:10.1016/j.ultramic.2013.03.014.

Béché A, Rouvière JL, Clément L & Hartmann JM (2009). Improved precision in strain measurement using nanobeam electron diffraction. *Appl Phys Lett* **95**, 123114. doi:10.1063/1.3224886.

Borisevich AY, Ovchinnikov OS, Chang HJ, Oxley MP, Yu P, Seidel J, Eliseev EA, Morozovska AN, Ramesh R, Pennycook SJ & Kalinin SV (2010). Mapping octahedral tilts and polarization across a domain wall in $BiFeO_3$ from Z-contrast scanning transmission electron microscopy image atomic column shape analysis. *ACS Nano* **4**, 6071–6079. doi:10.1021/nn1011539.

Bücker R, Hogan-Lamarre P, Mehrabi P, Schulz EC, Bultema LA, Gevorkov Y, Brehm W, Yefanov O, Oberthür D, Kassier GH & Dwayne Miller RJ (2020). Serial protein crystallography in an electron microscope. *Nat Commun* **11**, 996. doi:10.1038/s41467-020-14793-0.

Clabbers MTB, van Genderen E, Wan W, Wiegers EL, Gruene T & Abrahams JP (2017). Protein structure determination by electron diffraction using a single three-dimensional nanocrystal. *Acta Crystallogr D* **73**, 738–748. doi:10.1107/S2059798317010348.

Clausen A, Weber, D, @probonopd, Caron J, Nord M, Müller-Caspary K, Ophus C, Dunin-Borkowski R, Ruzaeva K, Chandra R, Shin J & van Schyndel J (2019). Libertem/libertem: 0.2.2. Available at 10.5281/zenodo.3489385.

Collette A (2013). *Python and HDF5*. Sebastopol USA: O'reilly Media, Inc.

Cooper D, Barnes J, Hartmann J, Béché A & Rouvière J (2009). Dark field electron holography for quantitative strain measurements with nanometer-scale spatial resolution. *Appl Phys Lett* **95**, 053501. doi:10.1063/1.3196549.

Cooper D, Denneulin T, Bernier N, Béché A & Rouvière J-L (2016). Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope. *Micron* **80**, 145–165. doi:10.1016/j.micron.2015.09.001.

Dask Development Team (2016). Dask: Library for dynamic task scheduling. Available at http://dask.pydata.org.

de la Peña F, Ostasevicius T, Fauske VT, Burdet P, Prestat E, Jokubauskas P, Nord M, Sarahan M, MacArthur KE, Johnstone DN, Taillon J, Caron J, Migunov V, Furnival T, Eljarrat A, Mazzucco S, Aarholt T, Walls M, Slater T, Winkler F, Martineau B, Donval G, McLeod R, Hoglund ER, Alxneit I, Hjorth I, Henninen T, Zagonel LF, Garmannslund A, & 5ht2 (2018) hyperspy/hyperspy: HyperSpy 1.3.1. Available at 10.5281/zenodo.1221347.

Donoho DL & Johnstone IM (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81**, 425–455. doi:10.1093/biomet/81.3.425.

Egerton RF (2011). *Electron Energy-Loss Spectroscopy in the Electron Microscope*, 3rd ed. New York: Springer.

EMD authors (2019). Electron Microscopy Datasets: An HDF5-based interchange file format for electron microscopy data. Available at https://emdatasets.com/format (retrieved June 3, 2018).

Emslie AG (1934). Scattering of electrons by stibnite and galena. *Phys Rev* **45**, 43–46. doi:10.1103/PhysRev.45.43.

Findlay SD, Shibata N, Sawada H, Okunishi E, Kondo Y & Ikuhara Y (2010). Dynamics of annular bright field imaging in scanning transmission electron microscopy. *Ultramicroscopy* **110**, 903–923. doi:10.1016/j.ultramic.2010.04.004.

fpd demos devs (2018). Notebook examples for the fpd package. Available at https://gitlab.com/fpdpy/fpd-demos (retrieved June 3, 2018).

fpd devs (2015). FPD: Fast pixelated detector data storage, analysis and visualisation. Available at https://gitlab.com/fpdpy/fpd (retrieved February 6, 2018).

Friedel G (1913). Sur les symétries cristallines que peut révéler la diffraction des rayons röntgen. *C R Acad Sci* **157**, 1533–1536.

Gammer C, Ozdol VB, Liebscher CH & Minor AM (2015). Diffraction contrast imaging using virtual apertures. *Ultramicroscopy* **155**, 1–10. doi:10.1016/j.ultramic.2015.03.015.

Glazer AM (1972). The classification of tilted octahedra in perovskites. *Acta Crystallogr B* **28**, 3384–3392. doi:10.1107/S0567740872007976.

Gouillart E, Nunez-Iglesias J & van der Walt S (2016). Analyzing microtomography data with Python and the scikit-image library. *Adv Struct Chem Imaging* **2**, 18. doi:10.1186/s40679-016-0031-0.

Guzzinati G, Ghielens W, Mahr C, Béché A, Rosenauer A, Calders T & Verbeeck J (2019). Electron Bessel beam diffraction for precise and accurate nanoscale strain mapping. *Appl. Phys. Lett.* **114**, 243501 doi: https://doi.org/10.1063/1.5096245.

Hallsteinsen I, Moreau M, Grutter A, Nord M, Vullum P-E, Gilbert DA, Bolstad T, Grepstad JK, Holmestad R, Selbach SM, N'Diaye AT, Kirby BJ, Arenholz E & Tybell T (2016). Concurrent magnetic and structural reconstructions at the interface of (111)-oriented $La_{0.7}Sr_{0.3}MnO_3/LaFeO_3$. *Phys Rev B* **94**, 201115. doi:10.1103/PhysRevB.94.201115.

Hammel M & Rose H (1995). Optimum rotationally symmetric detector configurations for phase-contrast imaging in scanning transmission electron microscopy. *Ultramicroscopy* **58**, 403–415. doi:10.1016/0304-3991(95)00007-N.

Hart MJ, Bassiri R, Borisenko KB, Véron M, Rauch EF, Martin IW, Rowan S, Fejer MM & MacLaren I (2016). Medium range structural order in amorphous tantala spatially resolved with changes to atomic structure by thermal annealing. *J Non-Cryst Solids* **438**, 10–17. doi:10.1016/j.jnoncrysol.2016.02.005.

Hartel P, Rose H & Dinges C (1996). Conditions and reasons for incoherent imaging in STEM. *Ultramicroscopy* **63**, 93–114. doi:10.1016/0304-3991(96)00020-4.

Huang F-T, Gloter A, Chu M-W, Chou FC, Shu GJ, Liu L-K, Chen CH & Colliex C (2010). Scanning transmission electron microscopy using selective high-order Laue zones: Three-dimensional atomic ordering in sodium cobaltate. *Phys Rev Lett* **105**, 125502. doi:10.1103/PhysRevLett.105.125502.

Hunter JD (2007). Matplotlib: A 2D graphics environment. *Comput Sci Eng* **9**, 90–95. doi:10.1109/MCSE.2007.55.

Hÿtch M, Houdellier F, Hüe F & Snoeck E (2008). Nanoscale holographic interferometry for strain measurements in electronic devices. *Nature* **453**, 1086–1089. doi:10.1038/nature07049.

Hÿtch MJ, Snoeck E & Kilaas R (1998). Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* **74**, 131–146. doi:10.1016/S0304-3991(98)00035-7.

Jin Z, Nori S, Lee Y-F, Kumar D, Wu F, Prater JT, Kim KW & Narayan J (2015). Strain induced room temperature ferromagnetism in epitaxial magnesium oxide thin films. *J Appl Phys* **118**, 165309. doi:10.1063/1.4934498.

Johnstone DN, Crout P, Høgås S, Martineau B, Smeets S, Laulainen J, Collins S, Morzy J, Prestat E, Ånes H, phillipcrout, Doherty T, Ostasevicius T & Bergh T (2019). pyxem/pyxem: pyxem 0.9.2. Available at https://doi.org/10.5281/zenodo.3407316.

Jones E, Oliphant T, Peterson P (2001). SciPy: Open source scientific tools for Python. Available at http://www.scipy.org (retrieved October 30, 2018).

Jones L, Varambhia A, Sawada H & Nellist PD (2018). An optical configuration for fastidious STEM detector calibration and the effect of the objective-lens pre-field. *J Microsc* **270**, 176–187. doi:10.1111/jmi.12672.

Jones PM, Rackham GM & Steeds JW (1977). Higher-order Laue zone effects in electron-diffraction and their use in lattice-parameter determination. *Proc R Soc Lond A* **354**, 197. doi:10.1098/rspa.1977.0064.

Kim Y-M, Pennycook SJ & Borisevich AY (2017). Quantitative comparison of bright field and annular bright field imaging modes for characterization of oxygen octahedral tilts. *Ultramicroscopy* **181**, 1–7. doi:10.1016/j.ultramic.2017.04.020.

Klinger M (2017). More features, more tools, more *CrysTBox*. *J Appl Crystallogr* **50**, 1226–1234. doi:10.1107/S1600576717006793.

Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, Ivanov P, Avila D, Abdalla S & Willing C (2016). Jupyter notebooks: A publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, Loizides F & Schmidt B (Eds.), pp. 87–90. IOS Press.

Kolb U, Gorelik T, Kübel C, Otten M & Hubert D (2007). Towards automated diffraction tomography: Part I–Data acquisition. *Ultramicroscopy* **107**, 507–513. doi:10.1016/j.ultramic.2006.10.007.

Krajnak M, McGrouther D, Maneuski D, O'Shea V & McVitie S (2016). Pixelated detectors and improved efficiency for magnetic imaging in STEM differential phase contrast. *Ultramicroscopy* **165**, 42–50. doi:10.1016/j.ultramic.2016.03.006.

Krakow W & Howland LA (1976). A method for producing hollow cone illumination electronically in the conventional transmission microscope. *Ultramicroscopy* **2**, 53–67. doi:10.1016/S0304-3991(76)90416-2.

LeBeau JM, D'Alfonso AJ, Findlay SD, Stemmer S & Allen LJ (2009). Quantitative comparisons of contrast in experimental and simulated bright-field scanning transmission electron microscopy images. *Phys Rev B* **80**, 174106. doi:10.1103/PhysRevB.80.174106.

Lewis LH, Marrows CH & Langridge S (2016). Coupled magnetic, structural, and electronic phase transitions in FeRh. *J Phys D Appl Phys* **49**, 323002. doi:10.1088/0022-3727/49/32/323002.

MacLaren I, Frutos-Myro E, McGrouther D, McFadzean S, Weiss JK, Cosart D, Portillo J, Robins A, Nicolopoulos S, del Busto EN & Skogeby R (2020). A comparison of a direct electron detector and a high speed video camera for scanning precession electron diffraction phase and orientation mapping. *Microsc Microanal.* doi:10.1017/S1431927620024411.

MacLaren I & Richter G (2009). Structure and possible origins of stacking faults in gamma-yttrium disilicate. *Philos Mag* **89**, 169–181. doi:10.1080/14786430802562132.

MacLaren I, Wang L, Morris O, Craven AJ, Stamps RL, Schaffer B, Ramasse QM, Miao S, Kalantari K, Sterianou I & Reaney IM (2013). Local stabilisation of polar order at charged antiphase boundaries in antiferroelectric $(Bi_{0.85}Nd_{0.15})(Ti_{0.1}Fe_{0.9})O_3$. *APL Mater* **1**, 021102. doi:10.1063/1.4818002.

Mahr C, Müller-Caspary K, Ritz R, Simson M, Grieb T, Schowalter M, Krause FF, Lackmann A, Soltau H, Wittstock A & Rosenauer A (2019). Influence of distortions of recorded diffraction patterns on strain analysis by nano-beam electron diffraction. *Ultramicroscopy* **196**, 74–82. doi:10.1016/j.ultramic.2018.09.010.

Martineau BH, Johnstone DN, van Helvoort ATJ, Midgley PA & Eggeman AS (2019). Unsupervised machine learning applied to scanning precession electron diffraction data. *Adv Struct Chem Imaging* **5**, 3. doi:10.1186/s40679-019-0063-3.

Mawson T, Nakamura A, Petersen TC, Shibata N, Sasakif H, Paganin DM, Morgan MJ & Findlay S (2020). Suppressing dynamical diffraction artefacts in differential phase contrast scanning transmission electron microscopy of long-range electromagnetic fields via precession. arXiv:2002.01595.

McMullan G, Cattermole D, Chen S, Henderson R, Llopart X, Summerfield C, Tlustos L & Faruqi A (2007). Electron imaging with Medipix2 hybrid pixel detector. *Ultramicroscopy* **107**, 401–413. doi:10.1016/j.ultramic.2006.10.005.

McVitie S, McGrouther D, McFadzean S, MacLaren DA, O'Shea KJ & Benitez MJ (2015). Aberration corrected Lorentz scanning transmission electron microscopy. *Ultramicroscopy* **152**, 57–62. doi:10.1016/j.ultramic.2015.01.003.

Midgley PA & Eggeman AS (2015). Precession electron diffraction: A topical review. *IUCrJ* **2**, 126–136. doi:10.1107/S2052252514022283.

Mir JA, Clough R, MacInnes R, Gough C, Plackett R, Shipsey I, Sawada H, MacLaren I, Ballabriga R, Maneuski D, O'Shea V, McGrouther D & Kirkland AI (2017). Characterisation of the Medipix3 detector for 60 and 80 keV electrons. *Ultramicroscopy* **182**, 44–53. doi:10.1016/j.ultramic.2017.06.010.

Mugnaioli E, Gorelik T & Kolb U (2009). "Ab initio" structure solution from electron diffraction data obtained by a combination of automated diffraction tomography and precession technique. *Ultramicroscopy* **109**, 758–765. doi:10.1016/j.ultramic.2009.01.011.

Naden AB, O'Shea KJ & MacLaren DA (2018). Evaluation of crystallographic strain, rotation and defects in functional oxides by the moiré effect in scanning transmission electron microscopy. *Nanotechnology* **29**, 165704. doi:10.1088/1361-6528/aaae50.

Nord M, Ross A, McGrouther D, Barthel J, Moreau M, Hallsteinsen I, Tybell T & MacLaren I (2019). Three-dimensional subnanoscale imaging of unit cell doubling due to octahedral tilting and cation modulation in strained perovskite thin films. *Phys Rev Mater* **3**, 063605. doi:10.1103/PhysRevMaterials.3.063605.

Nord M, Vullum PE, MacLaren I, Tybell T & Holmestad R (2017). Atomap: A new software tool for the automated analysis of atomic resolution images using two-dimensional Gaussian fitting. *Adv Struct Chem Imaging* **3**, 9. doi:10.1186/s40679-017-0042-5.

Nord M, Webster RWH, Paton KA, McVitie S, McGrouther D, MacLaren I & Paterson GW (2020). Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part I: Data Acquisition, Live Processing, and Storage. *Microsc. Microanal.* **26**(4), 653–666. doi: 10.1017/S1431927620001713.

Oliphant TE (2006). *A guide to NumPy*. USA: Trelgol Publishing.

Oliphant TE (2007). Python for scientific computing. *Comput Sci Eng* **9**, 10–20. doi:10.1109/MCSE.2007.58.

Ophus C (2019). Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microsc Microanal* **25**, 563–582. doi:10.1017/S1431927619000497.

Parkin SSP, Kaiser C, Panchula A, Rice PM, Hughes B, Samant M & Yang S-H (2004). Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nat Mater* **3**, 862–867. doi:10.1038/nmat1256.

Paterson GW, Macauley GM, Li Y, Macêdo R, Ferguson C, Morley SA, Rosamond MC, Linfield EH, Marrows CH, Stamps RL & McVitie S (2019). Heisenberg pseudo-exchange and emergent anisotropies in field-driven pinwheel artificial spin ice. *Phys Rev B* **100**, 174410. doi:10.1103/PhysRevB.100.174410.

Paterson GW, Webster RWH, Ross A, Paton KA, Macgregor TA, McGrouther D, MacLaren I & Nord M (2020) Dataset. 10.5281/zenodo.3903517

Pekin TC, Gammer C, Ciston J, Minor AM & Ophus C (2017). Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping. *Ultramicroscopy* **176**, 170–176. doi:10.1016/j.ultramic.2016.12.021.

Peng L-M & Gjønnes JK (1989). Bloch-wave channeling and HOLZ effects in high-energy electron diffraction. *Acta Crystallogr A* **45**, 699–703. doi:10.1107/S0108767389005982.

Pennycook S & Jesson D (1991). High-resolution Z-contrast imaging of crystals. *Ultramicroscopy* **37**, 14–38. doi:10.1016/0304-3991(91)90004-P.

pixStem devs (2015). pixStem: Analysis of pixelated STEM data. Available at https://gitlab.com/pixstem/pixstem (retrieved October 3, 2018).

Plackett R, Horswell I, Gimenez EN, Marchal J, Omar D & Tartoni N (2013). Merlin: A fast versatile readout system for Medipix3. *J Instrum* **8**, C01038. doi:10.1088/1748-0221/8/01/C01038.

Ponchut C, Collet E, Hervé C, Caer TL, Cerrai J, Siron L, Dabin Y & Ribois JF (2015). SMARTPIX, a photon-counting pixel detector for synchrotron applications based on Medipix3RX readout chip and active edge pixel sensors. *J Instrum* **10**, C01019–C01019. doi:10.1088/1748-0221/10/01/c01019.

Rauch EF, Portillo J, Nicolopoulos S, Bultreys D, Rouvimov S & Moeck P (2010). Automated nanocrystal orientation and phase mapping in the transmission electron microscope on the basis of precession electron diffraction. *Z Kristallogr* **225**, 103–109. doi:10.1524/zkri.2010.1205.

Rauch EF & Veron M (2005). Coupled microstructural observations and local texture measurements with an automated crystallographic orientation mapping tool attached to a TEM. *Materialwiss Werkst* **36**, 552–556. doi:10.1002/mawe.200500923.

Rauch EF & Véron M (2014). Virtual dark-field images reconstructed from electron diffraction patterns. *Eur Phys J Appl Phys* **66**, 10701. doi:10.1051/epjap/2014130556.

Rouvière JL, Béché A, Martin Y, Denneulin T & Cooper D (2013). Improved strain precision with high spatial resolution using nanobeam precession electron diffraction. *Appl Phys Lett* **103**, 241913. doi:10.1063/1.4829154.

Rouvière JL & Sarigiannidou E (2005). Theoretical discussions on the geometrical phase analysis. *Ultramicroscopy* **106**, 1–17. doi:10.1016/j.ultramic.2005.06.001.

Savitzky BH, Hughes LA, Zeltmann SE, Brown HG, Zhao S, Pelz PM, Barnard ES, Donohue J, DaCosta LR, Pekin TC, Kennedy E, Janish MT, Schneider MM, Herring P, Gopal C, Anapolsky A, Ercius P, Scott M, Ciston J, Minor AM & Ophus C (2020). py4DSTEM: A software package for multimodal analysis of four-dimensional scanning transmission electron microscopy datasets. arXiv:2003.09523.

Savitzky BH, Zeltmann S, Barnard E, lerandc, Brown HG, Henderson M & Ginsburg D (2019). py4dstem/py4DSTEM: DOI release. https://doi.org/10.5281/zenodo.3333960.

Schaffer B, Grogger W & Kothleitner G (2004). Automated spatial drift correction for EFTEM image series. *Ultramicroscopy* **102**, 27–36. doi:10.1016/j.ultramic.2004.08.003.

Schaffer M, Schaffer B & Ramasse Q (2012). Sample preparation for atomic-resolution STEM at low voltages by FIB. *Ultramicroscopy* **114**, 62–71. doi:10.1016/j.ultramic.2012.01.005.

Shibata N, Kohno Y, Findlay SD, Sawada H, Kondo Y & Ikuhara Y (2010). New area detector for atomic-resolution scanning transmission electron microscopy. *Microscopy* **59**, 473–479. doi:10.1093/jmicro/dfq014.

Somnath S, Smith CR, Laanait N, Vasudevan RK, Ievlev A, Belianinov A, Lupini AR, Shankar M, Kalinin SV & Jesse S (2019). USID and pycroscopy: Open frameworks for storing and analyzing spectroscopic and imaging data. arXiv:1903.09515.

Spence J & Koch C (2001). On the measurement of dislocation core periods by nanodiffraction. *Philos Mag B* **81**, 1701–1711. doi:10.1080/13642810108223113.

Spence J, Zuo J & Lynch J (1989). On the HOLZ contribution to STEM lattice images formed using high-angle dark-field detectors. *Ultramicroscopy* **31**, 233–239. doi:10.1016/0304-3991(89)90218-0.

Su D & Zhu Y (2010). Scanning moiré fringe imaging by scanning transmission electron microscopy. *Ultramicroscopy* **110**, 229–233. doi:10.1016/j.ultramic.2009.11.015.

Tate MW, Purohit P, Chamberlain D, Nguyen KX, Hovden R, Chang CS, Deb P, Turgut E, Heron JT, Schlom DG, Ralph DC, Fuchs GD, Shanks KS, Philipp HT, Muller DA & Gruner SM (2016). High dynamic range pixel array detector for scanning transmission electron microscopy. *Microsc Microanal* **22**, 237–249. doi:10.1017/S1431927615015664.

Temple RC, Almeida TP, Massey JR, Fallon K, Lamb R, Morley SA, Maccherozzi F, Dhesi SS, McGrouther D, McVitie S, Moore TA & Marrows CH (2018a). Antiferromagnetic-ferromagnetic phase domain development in nanopatterned FeRh islands. *Phys Rev Mater* **2**, 104406. doi:10.1103/PhysRevMaterials.2.104406.

Temple RC, Almeida TP, Massey JR, Fallon K, Lamb R, Morley SA, Maccherozzi F, Dhesi SS, McGrouther D, McVitie S, Moore TA & Marrows CH (2018b). 4D STEM data set supporting "Antiferromagnetic-ferromagnetic phase domain development in nanopatterned FeRh islands". University of Glasgow, Enlighten: Research Data.

The HDF Group (1997–2018). Hierarchical Data Format, version 5. Available at http://www.hdfgroup.org/HDF5.

Thong JTL, Sim KS & Phang JCH (2001). Single-image signal-to-noise ratio estimation. *Scanning* **23**, 328–336. doi:10.1002/sca.4950230506.

Tick T & Campbell M (2011). TSV processing of Medipix3 wafers by CEA-LETI: A progress report. *J Instrum* **6**, C11018–C11018. doi:10.1088/1748-0221/6/11/c11018.

Tsai C-Y, Chang Y-C, Lobato I, Van Dyck D & Chen F-R (2016). Hollow cone electron imaging for single particle 3D reconstruction of proteins. *Sci Rep* **6**, 27701. doi:10.1038/srep27701.

Van Aert S, Batenburg KJ, Rossell MD, Erni R & Van Tendeloo G (2011). Three-dimensional atomic imaging of crystalline nanoparticles. *Nature* **470**, 374–377. doi:10.1038/nature09741.

van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T & the scikit-image contributors (2014). scikit-image: Image processing in Python. *PeerJ* **2**, e453. doi:10.7717/peerj.453.

Vincent R & Midgley P (1994). Double conical beam-rocking system for measurement of integrated electron diffraction intensities. *Ultramicroscopy* **53**, 271–282. doi:10.1016/0304-3991(94)90039-6.

Voyles PM & Muller DA (2002). Fluctuation microscopy in the STEM. *Ultramicroscopy* **93**, 147–159. doi:10.1016/S0304-3991(02)00155-9.

Wang Y, Salzberger U, Sigle W, Suyolcu YE & van Aken PA (2016). Oxygen octahedra picker: A software tool to extract quantitative information from STEM images. *Ultramicroscopy* **168**, 46–52. doi:10.1016/j.ultramic.2016.06.001.

Yakovlev S & Libera M (2008). Dose-limited spectroscopic imaging of soft materials by low-loss EELS in the scanning transmission electron microscope. *Micron* **39**, 734–740. doi:10.1016/j.micron.2007.10.019.

Yang H, Jones L, Ryll H, Simson M, Soltau H, Kondo Y, Sagawa R, Banba H, MacLaren I & Nellist PD (2015). 4D STEM: High efficiency phase contrast imaging using a fast pixelated detector. *J Phys Conf Ser* **644**, 012032. doi:10.1088/1742-6596/644/1/012032.

Yang Q, Sha J, Wang L, Wang J & Yang D (2006). MgO nanostructures synthesized by thermal evaporation. *Mater Sci Eng C* **26**, 1097–1101. doi:10.1016/j.msec.2005.09.082.

Zeltmann SE, Müller A, Bustillo KC, Savitzky B, Hughes L, Minor AM & Ophus C (2020). Patterned probes for high precision 4D-STEM bragg measurements. *Ultramicroscopy* **209**, 112890. doi:10.1016/j.ultramic.2019.112890.

Zhu J-G & Park C (2006). Magnetic tunnel junctions. *Mater Today* **9**, 36–45. doi:/10.1016/S1369-7021(06)71693-5.