in situ Deformation Studies with Scanning Nanobeam Electron Diffraction
in situ Deformation Studies with Scanning Nanobeam Electron Diffraction

by

Thomas Christopher Pekin

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Engineering – Materials Science and Engineering in the Graduate Division of the University of California, Berkeley.

Committee in charge:
Professor Andrew M. Minor, Chair
Professor Mark Asta
Professor Peter Hosemann

Summer 2018
\textit{in situ} Deformation Studies with Scanning Nanobeam Electron Diffraction

Copyright 2018
by
Thomas Christopher Pekin
in situ Deformation Studies with Scanning Nanobeam Electron Diffraction

by

Thomas Christopher Pekin
Doctor of Philosophy in Engineering – Materials Science and Engineering
University of California, Berkeley

Professor Andrew M. Minor, Chair

In this dissertation, experimental methods and data processing techniques are developed for nanobeam electron diffraction and applied to both crystalline and amorphous materials. Nanobeam electron diffraction is a technique in which a small electron probe is used to acquire diffraction patterns as the beam is rastered across the sample. We develop methods to use nanobeam electron diffraction during in situ deformation in the scanning transmission electron microscope. First, we describe the sample preparation methods used to create samples with the correct geometry for in situ experimentation. We compare and contrast in situ deformation techniques, including holders and experimental data obtained. We then develop and benchmark variations on cross-correlation algorithms for nanobeam electron diffraction strain mapping on simulated and real data. We find that Sobel filtered cross-correlation and “hybrid” correlation minimize the amount of error when diffraction patterns have uneven illumination. We also show that binning can reduce error when signal to noise is low in the diffraction patterns. We use this result to analyze in situ nanobeam electron diffraction strain mapping on a sample of 321 stainless steel being deformed in tension. We observe the motion of the first dislocation in a planar slip band, and measure the resulting lattice expansion in situ. This result is the first direct confirmation of such phenomena, and is theorized to be the reason why planar as opposed to wavy slip occurs. We then develop methodology to measure both strain and local order in amorphous materials, and demonstrate the technique on Cu$_{46}$Zr$_{46}$Al$_{8}$ bulk metallic glass. We observe strains in excess of 2% before the sample fractures along a shear band, small amounts of plasticity, as well as a decrease in local order at high strains in areas where the sample fractured. This is supported by molecular dynamics simulations, and experimentally supports the shear transformation zone model of metallic glass deformation. As a whole, this dissertation presents method developments in in situ nanobeam electron diffraction for both crystalline and amorphous samples, as well as practical results of interest to the materials science community.
Contents

Chapter 1: Introduction ................................. 1
  1.1 Electron microscopy .................................. 1
  1.2 in situ electron microscopy and the mechanical behavior of materials ....... 2
  1.3 Nanobeam electron diffraction (NBED) .................................. 2
  1.4 Dissertation overview .................................. 3

Chapter 2: Sample preparation methods for in situ microscopy ............... 5
  2.1 Introduction ............................................ 5
  2.2 Traditional methods of sample preparation ............................... 5
  2.3 Sample preparation for alternative geometries and methods for in situ deformation in the TEM .................................................. 8
  2.4 Best practices in sample preparation .................................. 13

Chapter 3: Nanobeam electron diffraction algorithm development ......... 15
  3.1 Introduction ............................................ 15
  3.2 Theory ................................................. 16
  3.3 Results and discussion .................................. 20
  3.4 Conclusion ............................................. 26

Chapter 4: in situ strain mapping of 321 stainless steel .................... 27
  4.1 Introduction ............................................ 27
  4.2 Experimental methods .................................. 28
  4.3 Results and discussion .................................. 29

Chapter 5: in situ nanobeam electron diffraction of bulk metallic glass .... 35
  5.1 Introduction ............................................ 35
  5.2 Experimental methods .................................. 36
  5.3 Results and discussion .................................. 43

Chapter 6: Summary and future work ........................................ 47
  6.1 Conclusions ............................................ 47
  6.2 Future work ............................................ 48

References .................................................. 50

Appendix A: Methods useful for crystalline diffraction patterns .......... 63
Appendix B: Methods useful for amorphous diffraction patterns .......... 71
Appendix C: Useful image processing tips and techniques .................. 79
First and foremost I would like to acknowledge my parents and brothers, whose abundant support and love has been invaluable in my personal and professional development. Mom, Dad, Matt, Dan, thank you so much for everything, without you I would never have made it this far.

I would also like to acknowledge my primary advisor, Professor Andrew Minor, and my second advisor, Dr. Colin Ophus for the help and guidance throughout graduate school. Without either of these people, this research would not exist. Specifically, I would like to thank Andy for hiring me as an undergrad, and for letting me choose a project I found interesting and enabling me to run with it once I reached graduate school. The freedom to work on what was interesting to me was invaluable. Colin, thank you for the uncountable hours teaching me all things math, image analysis, and MATLAB (this is in addition to the unfortunately countable cups of coffee).

I would like to acknowledge the many other people who have helped me in this journey: Josh Kacher, for taking me as a new graduate student and serving as a guide for my initial research, Christoph Gammer and Burak Ozdol for their support as I shifted into NBED and in situ deformation and successful collaboration on a number of papers. I will miss our long days running complex experiments, complete with post-experiment “discussions” at LaVals. I also would like to thank Claire Chisholm for showing me the ropes of the Minor lab and NCEM, and the NCEM staff for the endless support, including Jim Ciston, for helping with TEAM I, Mary Scott for the fruitful discussion and collaboration, Karen Bustillo for the endless training, instruction manuals, thankless collection and organization of necessary bits and pieces to run experiments, and help whenever assistance was needed, John Turner for FIB training and the ability to track down things around NCEM, and Peter Ercius for the depth of knowledge with regards to microscope operation as well as image analysis.

I also would like to thank the labmates I have had in the Minor and Scott groups, as their camaraderie and advice has made it possible to continue working, even when it seemed unlikely that anything was going to work. I will miss their presence both in the lab, as well as at LaVals.

I would like to thank my friends outside of graduate student life for their endless support, advice, and subsidized beer. The list is long and I am very appreciative of their role in my life. Without their support, it is questionable that I would have completed this degree.

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, of the U.S. Department of Energy under Contract No. DE-AC02-05-CH11231 within the Mechanical Behavior of Materials (KC 13) program. The electron microscopy was carried out at the Molecular Foundry, which is supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05-CH11231.

Finally, I would like to thank both the Materials Science and Engineering department as
well as UC Berkeley as a whole. I have spent 9 years here and have learned and accomplished much more than I ever though possible when I first stepped onto the campus in 2009. This has been a place of immense growth and development, both personally and professionally. I will miss this campus, its people and its atmosphere deeply. The people I have met have truly been exceptional. I have not found another place comparable, and I am happy to graduate (for the third time) a Golden Bear.
Chapter 1:
Introduction

Materials science was practiced as an art and a science, with great success, for thousands of years without characterization methods beyond visual observation and bulk mechanical testing. These practitioners brought us materials like bronze, iron, steel, engineering feats like the Eiffel tower and the Titanic. They did much with the tools they had, but often had to rely on inference and rote empirical knowledge as to why their materials behaved the way they did. The advent of advanced characterization techniques (x-ray diffraction, electron microscopy, Raman spectroscopy, synchrotron facilities) revolutionized our understanding of materials by providing concrete experimental evidence of previously invisible mechanisms. These ranged from the atomistic (point and line defects), to the micron scale (protein crystallography, grain texture). This increased understanding has rapidly changed how we approach materials science over the past 100 years, moving from trial and error, to experimentally aided computational design of materials and prediction of properties. This has enabled vast advancements in all fields of materials science, including semiconductors and structural materials, that would have previously been impossible.

1.1 Electron microscopy

Electron microscopy began in 1931 with the development of the first electron microscope by Ernst Ruska and Max Knoll [1]. This was the result of advancements in both the theoretical understanding and practical knowledge needed to use electric fields to steer a beam of free electrons in a vacuum, hitting the area of interest, and then recording the various interactions that these electrons have with the sample and vice versa. Since that day, method and instrumentation development have been active areas of graduate student (and academic as a whole) research. These advancements have yielded atomic resolution microscopy [2], differential phase contrast imaging [3], low dose imaging [4], high resolution and fast spectroscopy [5], as well as \textit{in situ} techniques [6–9], in which a full experiment is performed within the microscope. Each of these advancements has pushed different parts of materials science and biology forward, armed with new data and increased understanding.

The electron microscope is an important tool, as few other methods have the versatility offered by this platform to examine materials from the angstrom to the millimeter and obtain the wealth of information produced when electrons scatter off of atoms and other electrons. It is complemented by many techniques, and is itself complementary to many others. However, it provides a crucial link between the atomistic and the bulk material, one that is hard to replace by other experiments. This explains the constant development on the platform for the past 87 years.
1.2 *in situ* electron microscopy and the mechanical behavior of materials

When it comes to the mechanical behavior of materials, electron microscopy has been influential in confirming the existence and role of dislocations [10], precipitates [11], grain boundaries [12], and chemistry [13] to the resulting mechanical properties of the material. However, traditional *ex situ* microscopy is limited in that the morphology of the sample observed is always only a before or after snapshot of whatever transition might have occurred. *In situ* microscopy changes this by allowing the microscopist to observe changes to the sample in real time as the sample deforms. This can include watching dislocations move [14], precipitates dissolve [15], and other phenomena [16] only observable under dynamic conditions. The knowledge of the dynamic deformation of the material can then be used to develop better alloys with more desirable properties.

*In situ* electron microscopy is not without its drawbacks however. Often, samples are much more difficult to prepare, requiring difficult geometries or specific orientations that require much more preparation than a standard metallic specimen. Additionally, size effects change how the material deforms as the sample size is reduced to electron transparency [17]. More worryingly is the need for electron transparency. The need to be less than 200 nm thick can generate complex strain states not normally experienced in the bulk sample. Additionally, the effect of the electron beam cannot be overlooked. In some samples, the high energy electron beam can have enough energy to generate knock-on radiation damage that will amorphize and destroy the sample during the *in situ* test. This is known as beam damage [18], and is usually regarded as a negative effect of the electron microscope. Regardless, many microscopists and materials scientists believe the insights gained by observing deformation mechanisms far outweigh the downsides of *in situ* microscopy.

1.3 Nanobeam electron diffraction (NBED)

Often in electron microscopy, the traditional microscopist must decide between obtaining structural (diffraction), or morphological (imaging) based data. This is a hindrance, as physically switching the modes on a traditional TEM takes both time and expertise, both of which might be limited during an *in situ* experiment where processes can happen very rapidly. It therefore has been of interest to see if there was a way to capture both at the same time. Nanobeam electron diffraction is one way in which this can be obtained. In this technique a semi-parallel, but small in diameter, electron beam is rastered across the sample, similar to a scanning transmission electron microscope (STEM). However, instead of integrating diffracted electron intensity on an annular detector and obtaining a single gray scale value for each probe position, an entire diffraction pattern is acquired. This allows for the reconstruction of traditional STEM imaging data as well as the structural/orientation information also desired [19].

Often, these datasets are gigantic in file size, exceeding 256 GB for a single “capture” and often over 1 TB for an *in situ* experiment with many time-resolved frames. This type of microscopy requires a new approach to analysis, as visual inspection of individual images is both useless as well as impossible. Instead, we must rely on computational methods to
integrate the large amount of data collected into a coherent result. These methods are typically not well developed, or non-existent. Therefore it is important to develop efficient, accurate methods for NBED data analysis. These methods can be used to extract a variety of information from the diffraction patterns, including orientation [20], virtual bright and dark field images [19], strain in both crystalline and amorphous materials [21–24], amount of short range order [25], as well as grain classification. While current methods are functional and allow for effective experimentation and analysis, single step processes in a larger data processing workflow can take upwards of ten hours per capture, and weeks to fully iterate through the tunable parameters. There is still much room for future method improvement, both in accuracy and speed.

1.4 Dissertation overview

This dissertation will focus on advancements in the sample preparation, NBED algorithm development, and applications to in situ electron microscopy. Its focus will be method development, and how new methods can be utilized to acquire previously unobtainable data relevant to the study of materials science.

This dissertation is organized as follows: Chapter 2 covers some of the inherent difficulties to sample preparation and complex geometries required for in situ deformation. We will start with traditional sample preparation techniques, and then show how they can be modified for in situ experiments. We will compare and contrast the various sample preparation mechanisms, provide schematics of the processes, and give a list of some of the best practices for successful sample preparation.

In Chapter 3, we first go over the variety of methods previously developed to identify and track the location of diffraction disks in crystalline samples. This is very useful for NBED strain mapping. We then develop and benchmark a more efficient method with both real and simulated data, specifically for the large amounts of data NBED produces. We show that the robustness of the algorithm to variable disk illumination is very important. Finally, we show how to turn these diffraction disk spot locations into strain maps.

In Chapter 4, we use the methods developed in Chapters 2 and 3 to perform NBED strain mapping on a crystalline sample in situ. We show that we are able to measure experimentally a positive strain following the first dislocation in a planar slip band as it moved. This had previously only been possible during ex situ analysis after specimen deformation. However, this experiment shows that in situ strain mapping is possible, opening the door for more quantitative studies on dislocation interactions with precipitates, grain boundaries, dissolved solutes, other dislocations, and other types of mechanical strengthening mechanisms.

Chapter 5 will discuss both the algorithmic and practical developments to measure strain in amorphous materials, such as a bulk metallic glass (BMG). Additionally it will cover how to measure the degree of local order, as well as the experimental parameters that can maximize both measurements. Then, it will discuss performing these measurements in situ, and observing the dynamic change in both strain and local order, and how this increases our understanding of BMG deformation.

Finally, in Chapter 6, we will discuss the overall contributions to the field this researcher has provided, and holistically summarize the impact of the method development presented
here. Additionally, we will discuss future research directions opened by this work both in the *in situ* microscopy as well as algorithm development.

In the appendices, we will provide a more detailed examination of the parts of this work which confused the researcher the most, in hope that it helps future graduate students like himself. Appendix A will focus on crystalline diffraction patterns, while Appendix B will focus on amorphous diffraction patterns. Appendix C will collect useful tips and techniques on processing images in MATLAB.
Chapter 2:
Sample preparation methods for *in situ* microscopy

2.1 Introduction

One of the earliest lessons an aspiring microscopist learns during the development of their experimental skills is that the process of sample preparation is often orders of magnitude more important than the actual microscopy itself. Often, graduate students will spend months learning microscopy, but years learning sample preparation, and the beginning of real research starts with the development of a reliable method of sample preparation. This will be an overview of the methods used in this dissertation such that future work developed off of it will not require so much sample preparation development. It will focus on traditional conductive, metallic specimens. Of course, the classic textbook *Transmission Electron Microscopy: A Textbook for Materials Science* by Williams and Carter has a decent overview on the wider variety of sample preparation techniques [26].

2.2 Traditional methods of sample preparation

The goal of sample preparation for transmission electron microscopy is to produce a sample that is thin enough for electron transmission (electron transparent), yet still representative of the larger sample, without artifacts from the preparation itself. Electron transparency depends on the material; materials with higher atomic weight must be thinner than materials with lower atomic weight for equivalent transparency. Additionally, electron transparency is dependent on the accelerating voltage of the microscope, as higher energy electrons can penetrate further through a sample than lower energy electrons. What is electron transparent at 300 kV could very well be opaque at 80 kV. Finally, the thickness of the sample controls how much dynamic scattering and attenuation occurs as the electron wave propagates through the sample. If atomic resolution/kinematical diffraction is desired, often samples should be less than ~10 nm, and require much more delicate preparation. However, since this dissertation is focused on *in situ* deformation, sample thickness requirements are relaxed as we do not want to observe phenomena strictly due to size effects [27–29] and there are practical considerations on the minimum size of the sample when using *in situ* holders. Therefore, as a guide, electron transparency for light metals (Al, Mg, etc.) is somewhere around 200-250 nm at 300 kV, and 100-150 nm for heavier engineering metals (Fe, Cu, Zr, etc.).

The traditional geometry for a standard TEM sample is a 3 mm disk, which allows for the transfer of a standardized sample between various sample holders and microscopes.
Therefore, usually the first steps of sample preparation involve taking the bulk material and shaping it into a rod with a 3 mm diameter, which can then be sliced via an accurate saw or wire electrical discharge machining (wire EDM). If the bulk material is already a foil or thin sheet, disks can be cut either using a punch if thin enough, or wire EDM if it thicker. In order to minimize surface damage from this machining process, these slices are usually around 600 microns thick. These disks are then typically glued to a puck using crystal bond (if temperature insensitive) or super glue (if temperature sensitive), and then ground to less than 200 microns using wet 800 grit sand paper, flipping once halfway through to remove cutting damage on both sides of the disk.

Crystal bond is applied and removed under the effect of heat, and both crystal bond and superglue can be dissolved under acetone in an ultrasonic bath.

### 2.2.1 Electropolishing

![Diagram showing sample preparation and electropolishing process]

**Figure 2.1.** a) Sample sheet with disks punched out. b) Punched out disks. Often it is beneficial to have more disks than necessary, as many will be destroyed during the sample preparation process, or turn out to be less than ideal samples. By preparing many, a few will be good. Keep this in mind. c) Standard dimensions of disk before jet polishing. d) Schematic of jet electropolishing setup. e) Jet polished disk with electron transparent region. f) V-I curve of electropolishing. This can be estimated by controlling the voltage to the jet polisher and recording the current. Optimum electropolishing occurs in the green region.

Once the samples have been thinned to ~200 nm thick, the careful work to thinning to electron transparency begins. Twin jet electropolishing has been one technique to do this
Table 2.1. Jet polishing parameters for materials used in this dissertation

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Electrolyte Solution</th>
<th>Temp. (°C)</th>
<th>Voltage (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Mg [30]</td>
<td>25% nitric, 75% methanol</td>
<td>-20</td>
<td>8-12</td>
</tr>
<tr>
<td>Al-Mg [30]</td>
<td>35% nitric, 65% methanol</td>
<td>-20 – -10</td>
<td>10</td>
</tr>
<tr>
<td>Al-Mg</td>
<td>10% perchloric, 90% methanol</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>321 Stainless Steel</td>
<td>6% perchloric, 39% butanol, 55% methanol</td>
<td>-15</td>
<td>30</td>
</tr>
<tr>
<td>CuZrAl BMG^2</td>
<td>33% nitric (70%), 67% methanol</td>
<td>-25</td>
<td>18</td>
</tr>
<tr>
<td>CuZr BMG [32]^3</td>
<td>9% nitric, 30% methanol, 61% butoxyethanol</td>
<td>-50</td>
<td>—</td>
</tr>
</tbody>
</table>

since the beginnings of electron microscopy [30, 31]. This process can be seen in Figure 2.1. In this process, the thinned disk is put into an acid-alcohol electrolyte solution, usually at reduced temperatures (-50 – 0°C), a bias is applied between the sample and the solution, and two jets of solution are pumped at the center of the disk (Fig. 2.1c). This causes a localized electropolishing effect where the jets are aimed, and causes a preferential thinning. This thinning proceeds until a photodetector detects light from the other side of the disk, provided by a bright source, indicating the presence of a hole.

While this process sounds simple in practice, there are several parameters which need to be optimized for successful jet polishing. The main parameters include choice of electrolyte, temperature of electrolyte, jet pump speed, and electropolishing voltage. Furthermore, these parameters are interdependent, and it often takes several samples to optimize. However, it is often helpful to turn to the literature to see what parameters previous researchers have used as a starting point, especially with regard to electrolyte.

Determining the voltage for electropolishing can be done by acquiring the curve shown in Figure 2.1f. By sweeping the voltage while electropolishing a sample a plateau can be observed. This is the voltage to use during polishing.

Temperature and pump speed can be used to control the speed of the polishing, with colder temperatures and slower pump speeds slowing the reaction, and warmer temperatures and faster pump speeds speeding it up. However, faster pump speeds could cause the thinnest areas of the sample to be destroyed before the pump can be cut off by the photodetector due to larger currents in the electrolyte bath.

Table 2.1 lists the parameters generally used for the samples in this dissertation. However, these should be taken as a starting point, as several more electrolyte solutions were tested, but did not differ markedly from the ones listed.

^1 This was not a reliable polishing solution.
^2 This was a very effective polishing solution and sample preparation parameters.
^3 This was not used, but could be a viable alternative for CuZr BMG. The voltage was not provided.
After the jet polishing is completed, samples should be dipped vertically into pure ethanol and let to dry to remove all polishing solution and water. Samples can then be stored until they oxidize, which is material dependent.

Sample quality can be initially determined optically. A good sample has a smooth bowl-like reflective surface, and a single very small symmetric hole in the center of the disk. Poor samples will have a combination of pitting, an oxide layer, multiple holes, off center holes, and a non-reflective surface.

2.2.2 Focused ion beam milling

Focused ion beam (FIB) milling is a second method traditionally used to make cross-sectional samples of specimens, especially when the amount of source material is limited. In this method, a dual-beam scanning electron microscope (SEM) and FIB are used in tandem, with a micromanipulator, to extract very small, thin sections of material, which are then placed on a special TEM grid containing stubs designed for mounting [33–35]. A schematic for this process is shown in Fig. 2.2.

There are several pros and cons with regards to FIB produced samples. The benefits include the ability to remove site-specific samples from regions of interest visible in the SEM (crack tips, precipitate-matrix boundaries, specific areas from semiconductor devices), small amounts of bulk sample consumption (for when source material is limited), and ease of removing cross sectional specimens which are difficult to do accurately with electropolishing. However, the use of FIB introduces additional constraints with respect to sample preparation. The FIB sample preparation process is often much slower on a per sample basis when compared to electropolishing, on the order of two to three hours per cross section, which can be very tedious. Additionally, care must be taken during the preparation that the sample is simply not destroyed during imaging if using the ion beam. Sample drift during preparation can render the process unusable. Finally, perhaps most importantly, the FIB causes sample damage. Traditional FIBs utilize gallium atoms as the main milling species. This Ga is implanted into the sample, and during the milling process, produces an amorphous layer proportional to the accelerating voltage of the Ga atoms. This can destroy features, embrittle aluminum samples, and cause severe contamination. To minimize this, low voltage milling is performed, but this limits the resolution, and can lead to overmilling if not careful. Additionally, to minimize the effects of Ga in susceptible materials, alternative ion beam technologies exist [36].

2.3 Sample preparation for alternative geometries and methods for *in situ* deformation in the TEM

So far, we have covered the primary methods of sample preparation for TEM studies of metallic specimens. However, these are not well suited for *in situ* deformation in the microscope, as a 3 mm disk is hard to accurately deform, and liftouts on a traditional liftout stub as shown in Fig. 2.2f would be difficult to deform in a controlled manner. Specialized TEM sample holders have been developed to solve these problems though. A few examples can be seen in Fig. 2.3.
Figure 2.2. a) Source material polished relatively flat. b) Sample before liftout occurs. Before this step, trenches were milled on both sides of the liftout specimen multiple microns deep, and with enough clearance such that the bottom of the liftout can be cut free from the bulk. During this trenching and thinning, the sample is tilted $\pm 1 - 2^\circ$ from normal to keep the flat faces roughly parallel during milling. Notice the one edge already cut free with room for the micromanipulator to come in and attach and the Pt layer deposited on top of the sample to protect the underlying sample from unnecessary beam damage. At this step the bottom of the sample is already cut off as well. c) The sample lifted out from the bulk. It is welded to the micromanipulator using deposited Pt. d) The sample before placement on liftout grid. The liftout grids have several posts which the liftouts can be attached to using platinum. Prior to the sample being brought near the post, the top view shows a notch milled out using the FIB. The sample will be placed in this notch. The dashed line shows where the sample will be cut from the micromanipulator after it is welded in the notch. e) The sample in the notch ready for additional thinning or TEM observation. Notice the Pt patch applied in the top view. This is to prevent the sample from falling out of the milled notch in the post. f) In the box on the top is a schematic of a standard liftout grid with four posts. One of these posts can be seen in d) and e). Below the box are standard dimensions of a liftout prior to additional thinning.
Figure 2.3. Some of the available geometries for the Gatan 654 straining holder, and the Hysitron/Bruker PI 95 nanoindenter. a) Dimensions for the Gatan 546 holder. This geometry uses a full bar of the sample material that has been thinned in the middle, either by electropolishing or ion milling. The black arrows show the direction of deformation. b) Sample geometry for gluing a 3 mm disk on the tensile bar (dotted circle). c) Sample geometry for laying a thin film over a slot in the tensile bar (dotted rectangle). d) Sample geometry for nanowires or FIB liftouts. The small gap is bridged by a small silicon bar (black), which is then milled in the center to a desired gap. This gap of only a few microns can be spanned by the nanowire or liftout. e) Sample geometry for PI 95 pillar compression deformation. There are a variety of pillar thicknesses and indenter tip geometries for various tests. The inset image in the box shows the overall geometry, with a half disk thinned in the center where the pillars are, and the black indenter tip approaching the pillars. f) The PI 95 push to pull (PTP) device. This is a silicon device that turns the nanoindentation compression into a tensile force. The red dotted box is shown magnified on the right, with two types of samples. Both the nanowire and the FIB milled dogbone specimen will undergo tensile deformation. They are affixed to the Si PTP with deposited Pt.
The two main *in situ* deformation holders used in this dissertation were the Gatan 654 straining holder, and the Hysitron Bruker PI 95 Picoindenter. These holders have very different capabilities, and sample geometry requirements.

The Gatan 654 holder is a screw driven tensile holder that accepts a large bar with dimensions 11.5 mm by 2.5 mm by 100 microns or less. This bar is screwed into the tensile holder by two screws, which then are driven apart, putting the bar under a tensile load. This holder is very easy to use, with the only controls being to move the screws apart or together, and has a display that roughly measures displacement in microns. Additionally, the sample preparation is very simple. The large bars can be cut using wire EDM from the source material and electropolished in the center, or 3 mm disks can be glued over a hole, or smaller specimens can be placed directly over an appropriately sized gap on a tensile bar mount. These are all shown in Fig. 2.3a-d. While the holder simplicity causes sample preparation to be relatively easy, it does have some drawbacks. The lack of control with regards to the amount of tensile deformation means that the amount of deformation is not really known. Additionally, the geometry makes it impossible to know what the strain fields are. Finally, with the easier to prepare sample geometries, while they do provide large electron transparent regions to examine, one does not know where the sample will first locally deform. This researcher’s experience predicts that the sample will usually deform first out of the field of view, sometimes catastrophically, before anything changes in the region being viewed. Therefore there is a trade off between having large areas of specimen to observe interesting static microstructural features and knowing exactly where the dynamic processes are going to concentrate.

The PI 95 can be viewed as the opposite with respect to capabilities and features of the Gatan 654 deformation holder. First, the PI 95 sample geometry is very different. There are two main types of sample geometries, compression, usually in the form of nanopillars, and tension, usually using push to pull (PTP) devices. These geometries are shown in Fig. 2.3e and f. As the size of the samples are much smaller, fabrication is much more difficult, and almost always requires the FIB. Sample fabrication methods for compression tests can be seen in Fig. 2.4, while sample preparation for PTPs involve either floating thin films on top of the device, and then milling away the excess, or using the FIB to produce planar liftouts or manipulate nanowires as shown in Fig. 2.3f. Additionally, the holder is much more fragile and experimentation is much more difficult, with several alignment and calibration steps not present with the Gatan 654. While the JEOL and the FEI holder have different internal mechanics, both can easily be broken if handled roughly. The benefits of the PI 95 though, are the high control of both load and displacement during mechanical deformation. The PI 95 can acquire load-displacement curves during deformation with extremely high resolution (<0.2 μN in force, <1 nm in displacement), which allows for both quantitative analysis, as well as knowledge of the strain state of the sample under deformation. Additionally, since the sample geometry is highly confined, the location of localized deformation is known. Therefore, there is no chance for the deformation to occur outside of the field of view, and geometries can be utilized to probe certain material characteristics (single crystal orientation effects, grain boundaries, etc.).
Figure 2.4. a) Top down method of making pillars. In this method, annular patterns are milled normal to the surface of a polished flat surface or edge of a thin specimen (shown). Subsequent annular patterns shrink in both inner and outer diameter until the desired inner diameter is reached, after which the surrounding area can be cleaned off for better visibility in the TEM. Often, the pillars made using this method have a larger diameter at the bottom when compared to the top, as the FIB process does not mill straight down. b) The author’s preferred method of pillar preparation. A thinned half disk is mounted 45° to the FIB. Straight rectangles are milled vertically, creating slots in the edge of the sample. Then, the sample stage is rotated 180°, such that the 45 degree angle with respect to the FIB causes a 90° angle with the previous cuts. The straight rectangles are milled again, leaving square pillars with constant dimensions down the whole pillar. This process can be repeated to create smaller and smaller pillars. As a side effect of the milling, often the pillars will not need secondary cleaning around them. This process is often faster and less susceptible to drift than a), as well as produces straighter pillars.
2.4 Best practices in sample preparation

To conclude, here is a short list of best practices in sample preparation.

2.4.1 Electropolishing

- If starting from the bulk, and using wire EDM to cut disks from a bar, disks should be cut to roughly 600 microns thick.

- Consider annealing the samples before electropolishing to remove excess dislocation density, especially if using an unannealed foil.

- Use 400-800 grit sandpaper to thin disks to less than 200, and preferably 150 microns. The thinner the better as it minimizes the time, and therefore the amount of things that can go wrong during electropolishing.

- Conventional wisdom says the damage layer for sanding is three times the grit size. 400-800 grit sandpaper has grit sizes between 10 and 20 microns [26].

- Take care to keep the sanded disks flat, wedge shaped disks will not electropolish correctly.

- Electropolishing has a bit of an “art” to it. Certain polishing solutions get better with age, others worsen. This takes experimentation to determine. However, if a good recipe is found, it can be very repeatable.

- When trying a new solution, try a range of temperatures. Start cold and let the solution warm up while you make many samples. Check the voltage occasionally to make sure it does not need adjustment.

- Sample yield is typically around 10 to 50%. Prepare enough blank samples accordingly.

- Half grids are often useful for in situ deformation. If care is taken, these samples can be cut using sharp scissors or a razor blade after electropolishing without destruction of the thin area.

- Well prepared electropolished specimens offer the largest possible electron transparent region, when compared to other methods.

2.4.2 Focused ion beam milling

- When producing a sample of complex geometry, plan all stage movements prior to sample insertion, to make sure the geometry is possible.

- Image the sample as little as possible using the FIB, it will continually damage your sample.

- When making liftouts or utilizing Pt deposition, make the sample larger, so that Pt contamination can be as far from the thin center of the sample as possible.
• After depositing Pt, do not turn on the electron beam for a few minutes until the pressure has noticeably dropped. This pressure drop is related to the Pt gas precursor, and deposition can occur via the electron beam if the chamber is not fully evacuated.

• Keep the FIB damage mechanisms, damage layer depth, and preparation artifacts in mind during sample fabrication [36].

• Starting with an electropolished disk for liftouts or pillar fabrication is much faster than trying to remove all the surrounding material using the FIB.

• The pillar fabrication method in Fig. 2.4b produces much higher quality square pillars, and is simpler than the method in Fig. 2.4a.
Chapter 3:
Nanobeam electron diffraction algorithm development

3.1 Introduction

Strain and its spatial distribution is important for a greater understanding of many of the relevant engineering materials currently in use. In most modern silicon devices, strain is an important parameter used to modify the properties of the device itself [37]. Likewise in metallic specimens, understanding strain and its evolution under deformation will help further the understanding and predictive capabilities of the field. While many techniques of measuring strain exist [23, 38–44], scanning convergent nanobeam electron diffraction (NBED) is attractive for a number of reasons. First, NBED strain mapping offers the potential of very high accuracy in strain measurement. Independent diffraction patterns are recorded with high reciprocal space resolution for each probe position, which limits the spatial resolution to the probe size. On a modern scanning transmission electron microscope (STEM), this probe size can easily be below one nanometer while still maintaining a small enough convergence angle to display full diffraction patterns. This can be compared with geometric phase analysis (GPA) strain mapping, in which real space images are acquired with very high spatial resolution, but do not directly sample reciprocal space. Additionally, GPA strain maps are necessarily limited to a small field of view (FOV) as atomic columns must be resolved to make accurate measurements. With NBED this is not a problem and FOV is usually limited by the storage space of the data acquisition system or the sample stability. With the introduction of high speed direct electron detectors, a large number of diffraction patterns can be obtained from a single sample, covering a very large field of view without concerns for sample drift or other instabilities [21, 38, 45–47].

Cooper et al., have noted that NBED strain measurements can lose accuracy due to non-uniform disk intensity [46]. This non-uniformity is due to experimental limitations such as sample bending, dynamical effects, or imperfect alignment, resulting in more complicated data sets. The effects of dynamical contrast can be dealt with both during and after the experiment. During the experiment, the electron beam can be precessed around the central

The results presented in this chapter and have been published as a regular article with title “Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping” in Ultramicroscopy 176, 170-176 (2017) by Thomas C. Pekin, Christoph Gammer, Jim Ciston, Andrew M. Minor and Colin Ophus [22]. The material is presented here with the permission of co-authors and publishers.
axis to obtain a radially averaged diffraction pattern. This has been shown to reduce dynamical contrast and return diffraction patterns that can be analyzed as if they were kinematical [48, 49]. However, in return for more easily processed diffraction patterns, acquisition speed is slowed to the speed at which the beam can be precessed (~0.1 s per pattern versus 0.0025 s per pattern with a Gatan K2-IS camera), and the spot size increases as microscope aberrations have a larger effect on off axis beams. Alternatively, after the data has been acquired, choices in disk position measurement, lattice fit to disk positions, and image downsampling provide a range of options to optimize the data obtained from malformed disks and recover strain information. Here we present an overview of the strain mapping technique itself and experimental options, tested on both simulated data as well as relevant experimental results.

3.2 Theory

3.2.1 Measuring Lattice Vectors from Nanobeam Electron Diffraction Patterns

Strain measurement via diffracted peak location is a well understood result of Bragg’s law, and has been successfully performed using several different experimental techniques [21, 45, 50]. For NBED, the first and most important step is proper data acquisition. While the sample does not need to be on a perfect zone axis, it must be close enough to have ideally several orders of diffraction disks illuminated. For every pixel in the reconstructed strain image, an entire diffraction pattern must be recorded, shown schematically in Figure 3.1a. From each of these patterns, the disk positions are extracted and stored as a \((x, y)\) location in reciprocal space, usually to subpixel precision. After all the disk positions are recorded, they are used to find the local lattice vectors at every probe position. This is done by solving the system of linear equations

\[
BL = P, \tag{3.1}
\]

for \(L\), where \(L\) is the matrix made up of two lattice vectors defined by pixel lengths from the (000) spot, \(P\) is a matrix of every disk position in pixels, and \(B\) is a matrix of every disk position in normalized lattice vectors. Equivalent rows of \(P\) and \(B\) should correspond to the same diffraction disk for each diffraction disk registered. Often, this calculation is overdetermined, as there ideally will be many more disk positions than lattice vectors. If the solution is overdetermined, the fit accuracy can be improved by using weighted least squares, where the weights are equal to the correlation peak value. This calculation is carried out for every diffraction pattern in the dataset. In addition, a reference lattice \(L_0\) is computed using the disk positions from either a single real space pixel, or the mean of the disk positions for a subset of pixels (the reference region of the dataset).

Once the lattice vectors have been calculated for every diffraction pattern, it is simple to calculate matrix strain using \(L_0 \text{ } T = L\), where \(L_0\) is the reference lattice, \(T\) is the transformation matrix, and \(L\) is the current lattice for the pixel in question. If infinitesimal strain theory is assumed, the resulting strain matrix (and infinitesimal rotation) is simply

\[
\begin{bmatrix}
\varepsilon_{xx} & \frac{1}{2} (\varepsilon_{xy} - \theta)
\frac{1}{2} (\varepsilon_{yx} + \theta) & \varepsilon_{yy}
\end{bmatrix}
= T \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}, \tag{3.2}
\]
Figure 3.1. Strain measurements from nanobeam electron diffraction (NBED). a) Experimental geometry showing a single NBED measurement. The correlation of the measured diffraction pattern with a center disk template produces an image with sharp peaks at each disk position. Different correlation methods are applied to synthetic disks with b) no internal structure, c) disks with signal on opposite edges and d) disks with signal along a single disk edge. Each method shows an example correlation image, as well as the horizontal and vertical error (divided by disk radius) as a function of counts in the ideal disk without internal structure. Red ellipses show the best fit standard deviations for an elliptic Gaussian function on all peaks.
where $\varepsilon_{xx}$ is the strain in the $x$ direction, $\varepsilon_{yy}$ is the strain in the $y$ direct, $\varepsilon_{xy}$ and $\varepsilon_{yx}$ are the shear, and $\theta$ is the lattice rotation. Finally, we note that the above expression describes lattice strain in reciprocal space. Because NBED is a reciprocal lattice measurement, the signs of $\varepsilon_{xx}$, $\varepsilon_{yy}$ and $\varepsilon_{xy}$ (but not the sign of $\theta$) must be flipped in order to produce the correct strain measurement of the real space sample deformation.

### 3.2.2 Methods for Determining Diffraction Disk Positions

It is clear from the above discussion that the resolution of any resulting strain measurement depends on the accuracy and precision of the diffracted disk position measurement. If only whole pixel disk shifts were to be used, an artificial lower bound would be placed on the accuracy of the local strain measurements, effectively binning the measured strain values from a continuous distribution to a discrete one. Therefore, in order to obtain the most accurate strain measurements, subpixel shifts in disk position must be measured. While many subpixel registration algorithms exist [51–54], several considerations must be taken into account when choosing which to use. Due to the incredibly large size of the datasets obtained, both in file size and number of disk positions to be measured, only the fastest algorithms are considered. For example, while other authors have used edge detection via Canny or Prewitt filters followed by circle fitting to find the center of the disk [23], that approach is very computationally intensive, requiring several iterative steps per disk in order to resolve a measurement with subpixel accuracy. Additionally, that method relies on accurate filter thresholds which can vary with the type of diffracted disk structure observed, decreasing general robustness and ease of use.

Radial gradient maximization is another method used [23, 55, 56], in which concentric circles or ellipses are placed around an estimated disk center and their rotational averages are calculated. Since most disks have sharp edges, when the difference between the rotational average of concentric shapes are a maximum, the concentric shapes are properly placed around the center of the disk. Through an iterative process, the estimated center is shifted until this maximum is found. However, this method performs best when there are whole disks without missing portions or strong contrast changes along the disks’ edges, which is not always experimentally possible. Finally, this is an iterative method, and while it is faster than edge detection [56], it still requires several computational steps per disk.

Cross correlation with respect to image registration is a fast technique in which a template and base image are aligned via a convolution-like operation. While intuitively this might seem computationally intensive, the Cross Correlation Theorem reduces the problem to a matrix element multiplication the size of the image, two discrete Fourier Transforms (DFTs) and a single inverse (fast) Fourier Transform (IFFT). For images $g_1$ and $g_2$ with DFTs of $G_1$ and $G_2$,

$$\mathcal{F}\{g_1 \otimes g_2\} = \frac{G_1 \circ G_2^*}{|G_1 \circ G_2^*|^n}$$

where $\otimes$ represents the cross correlation operation, $\mathcal{F}$ denotes the discrete Fourier transform, $^*$ indicates the complex conjugate, $\circ$ is the element-wise product of the two matrices, and $1 - n$ is the power to which the Fourier coefficients are weighted.

Standard cross correlation is performed when $n$ is set equal to zero. In this case, the
intensity distribution in the real space image is fully taken into account. In the Fourier transforms, this manifests itself as a large weight put into low spatial frequency signals, resulting in a center-of-mass like registration between the image and the template.

Phases correlation is the natural extension of Eq. 3.3, where $G_1 \circ G_2$ is normalized element-wise to unit magnitude, thereby only retaining the phase information. This occurs when $n$ is set equal to 1. Phases correlation results are insensitive to image intensity, but also more sensitive to noise, sharp edges, and other high frequency components that previously had low image intensity. Finally, both cross and phase correlation can be seen as the limits of a more general correlation, in which image intensity is weighted by some arbitrary fraction $n$ from 0 to 1. In this study, we have tested correlation measurements using $n = 0.5$, which we hereafter refer to as the "hybrid" correlation method.

After taking the IFFT of $\mathcal{F}\{g_1 \otimes g_2\}$, the resulting matrix is the computed correlation, where the pixel of highest intensity corresponds to the shift required to align one image to the other. Two primary methods exist to increase the accuracy of the alignment to subpixel values. The first is curve-fitting the correlation surface to find the function’s maximum [57, 58]. This method can require iteratively curve-fitting a function (usually parabolic or Gaussian) to the correlation in the spatial domain and can be made arbitrarily accurate, but is computationally intensive.

The second, faster, algorithm involves Fourier upsampling in the frequency domain and is well described in [52, 59]. Using the original correlation peak as its center, the method directly computes subpixel IFFT values a fixed range around the brightest pixel with an arbitrary upsampling factor $\kappa$. This computation is performed as the matrix multiplication of three matrices and is very computationally efficient. The drawback of this method when compared to curve fitting is the lack of peak shape information and granularity of peak location measurement dependent on $\kappa$. However, to improve the final peak location, a last step is performed in which a simple parabola is fit to the nearest neighbors of the upsampled peak location, giving the measurement a final slight improvement. Due to the lack of a two step iterative curve-fitting step, we observed Fourier upsampling to be approximately three times faster when compared to fitting a Gaussian curve to the peak location for subpixel registration.

A variety of image filters can also be applied prior to correlation in order to emphasize features of interest that will aid in accurate registration. Often, this is the disk edge. A number of iterative edge detection filters exist, including the previously mentioned Prewitt and Canny filters, but they require user defined thresholds and several attempts to correctly find an edge. A simpler approach is the Sobel filter, which is a simple $3 \times 3$ kernel convolved with the image which approximates the gradient of intensity in a certain direction. Since the Sobel filter is a simple matrix, it can be applied across datasets without concern for effects caused by changing user defined thresholds, and is much faster and impartial than an iterative method. Sobel filtering has been shown to improve cross correlation accuracy of the forward scattered disk, useful in differential phase contrast imaging, but has not been robustly tested with diffracted disks containing dynamical contrast [60].

Simulated examples of correlation measurements using a circular template with perfect noisy circular disk signals (with and without regions of zero intensity) are given in Fig. 3.1b-d. In all cases, as the image intensity weighting factor $n$ in Eq. 3.3 is increased from 0 to 1,
the correlation peak sharpens. In a noisy measurement, this represents a trade-off between measurement precision and robustness of the disk position measurement. Additionally, we have simulated the relative disk position error (measured displacement from true position divided by disk radius $R$) as a function of number of counts (electrons) calculated using Poisson statistics in the total circular disk area. For perfect disks, or disks with intensity modulations that have inversion symmetry, cross correlation performs the best, and approaches the ideal measurement root-mean-square (RMS) error along one direction $\sqrt{\langle \Delta x^2 \rangle}$ of

$$\frac{\sqrt{\langle \Delta x^2 \rangle}}{R} \approx \frac{1.2}{\sqrt{\lambda R}} \quad (3.4)$$

where $\lambda$ is the total number of counts. For the case in Fig. 3.1d without inversion symmetry, cross correlation fails badly. In Figs. 3.1c and d, the horizontal error is approximately double the vertical error for most cases. This is because the disk edges are well-defined at the top or top and bottom of the circular boundary. Small horizontal shifts in the position measurements are not well-constrained due to the lack of horizontal edges.

Another trend apparent in Figs. 3.1b-d is that all of the correlation methods follow a power law that scales with the usual value of $1/\sqrt{\lambda}$, provided the number of counts are high enough relative to the disk size. Below this threshold, large position errors are always present. To prevent these errors, the diffraction space measurements should be sufficiently binned to smaller disk radii, as shown in Fig. 3.2.

Figs. 3.2a-d show that for low signal-to-noise ratios, the power law error dependence can be recovered in all cases by down-sampling the measurements, i.e. binning the resulting images. However binning also increases measurement error, and so one must be careful to limit the amount of down-sampling. Phase correlation and the Sobel-filtered cross correlation measurements are particularly sensitive to over-sampling-induced position errors. The reason why is demonstrated in Fig. 3.2e for the Sobel case. When the number of counts is low enough to produce separated point measurements, the edge filtering no longer produces an accurate measurement. A similar effect happens in phase correlation where erroneous high spatial frequencies increase the measurement error when separated points are apparent. We therefore recommend that all measurements are down-sampled until all disks resemble the center column of Fig. 3.2e or better.

## 3.3 Results and discussion

### 3.3.1 Position Measurements of Simulated Disks With Fine Structure

Often, the diffraction patterns recorded do not exhibit perfect kinematical scattering from an ideal sample, which would result in round, evenly illuminated diffraction disks. This is due to sample thickness, bending and other kinematical and dynamical diffraction effects, all of which increase the inner disk structure. This poses a challenge for the correlation methods previously discussed due to their reliance on specific similar features between the template and target disk. To test the effects of fine structure in the disks, one ideal disk and four imperfect disks were simulated. These are shown inset in Figure 3.3. For each
Figure 3.2. Effect of binning on disk position measurement error, using a) phase correlation, b) hybrid correlation, c) cross correlation and d) Sobel-filtered cross correlation, for different numbers of counts. e) One example simulation with Sobel filtering for various numbers of counts.
measurement, the disk was shifted a random amount \((\Delta x, \Delta y)\) and the effect of number of
electron counts per disk being tested was applied using Poisson statistics. This simulated
diffraction disk was then correlated with an ideal disk using one of the four methods of
correlation discussed here, and difference between actual and measured shift was recorded.
This process was simulated 1000 times for each signal to noise ratio level.

The results of these simulations are shown in Figure 3.3. For a perfect disk, it is clear
that cross correlation is the best and most accurate method across all SNR levels. However,
when the disks are unevenly illuminated, cross correlation has a strong systematic error in
position due to its inherent center-of-mass weighting. While cross correlation gets worse
when inner structure is present, the other methods do retain their accuracy for disks with
a high number of counts. This is a result of the correlation methods themselves and the
type of image feature each method emphasizes. Since cross correlation does not normalize
the Fourier transform, it is overly reliant on intensity variations in the image, leading to the
systematic errors seen. The other methods rely much more on high frequency components
of the image, namely edges, and the sharp edges present around the outer edge of the disk
provide an accurate feature for registration. The results of this simulation make it clear that
when inner structure appears in the diffracted disks, hybrid or Sobel filtered cross correlation
should be used over pure phase or cross correlation.

3.3.2 Experimental Strain Measurements on an Unstrained Silicon Lattice

Pure Si TEM support grids were purchased with 35 nm thin windows. A 50 by 50 pixel
scanning nanodiffraction dataset was obtained using a convergence angle of 1.01 milliradians
with a diffraction pattern image size of 512 by 512 pixels. The images were acquired with a
binning of 4, and not further binned. We expect this sample to have a uniform zero strain
across the field of view, which spans \(250 \times 250\) nm. In this dataset, the central beam was
obscured by a physical beam stop, preventing its use as a template. This often occurs in
overly thick samples as the beam current must be raised to a level beyond the camera’s
Figure 3.4. A comparison between disk registration methods on unstrained Si. a) A virtual dark field image showing thickness or bending contrast, but otherwise contrast-free. Streaks are indicative of poor camera readout during the scanning process. b) The mean diffraction pattern showing the x and y lattice vectors chosen. c-f) Line profiles of the regions marked with black rectangles. g-j) Strain maps of the area in a), highlighting differences in measured strain as a result of the correlation method used. The strain scale is from -0.6 to 0.6% strain. Inset is the root mean square value of the frame from zero strain.
damage threshold to fully illuminate the diffracted disks. To form a template for the various correlation methods, a best-fit ellipse to the disks was used (defined as the disk with the highest normalized correlation peak). Once the peak template was created, we computed strain maps using each of the four correlation methods, compiled in Figure 3.4, using the mean diffraction pattern as the reference.

The results confirm the prediction from the simulations. For the standard cross correlation, large changes in strain state exist in the maps from left to right and top to bottom in the reconstructed images, and the maps are fairly free of noise. The line profile in Fig. 3.4c shows both low noise and an unrealistic strain gradient. Cross correlation therefore has a high precision (small changes in adjacent diffraction patterns), but has large systematic errors in accuracy due to uneven illumination in the diffracted disks. This signal could be due to changes in sample thickness, small local sample tilts due to bending, or beam tilt.

In comparison, the phase correlation strain maps are very noisy, with large discontinuous jumps in strain state from pixel to pixel. The inset numbers in Fig. 3.4j correspond to the root-mean-square of the image strain, and a comparison between the other methods show that phase correlation is significantly noisier. This corresponds to the relative low accuracy of phase correlation and its susceptibility to background noise, which agrees with the simulations.

Finally, both hybrid and Sobel filtered cross correlation offer a decent compromise between the two extremes, with the Sobel filtered results having a minimum of noise and a very uniform measure of the strain in all directions. This mirrors the simulated results, in which these were the two best methods when the diffracted disks began to show signs of inner structure.

3.3.3 Experimental Measurements of an Austenitic Stainless Steel Sample

An austenitic 304 stainless steel sample was prepared via jet electropolishing using a perchloric acid solution. A 250 by 250 pixel scanning nanodiffraction dataset was obtained using a convergence angle of 2 milliradians, with a diffraction pattern image size of 1920 by 1792 pixels and a pixel step size of 2.5 nm. The final dataset size was approximately 450 gigabytes. The data was binned by a factor of 8 to reduce dataset size to a more manageable 7 GB. The four disk registration methods were then applied to the dataset and the resulting strain maps computed via the process described previously. The results are shown in Fig. 3.5 with a scale showing strain from -2 to 2%. A low angle annular dark field (LAADF) image was taken at 720 mm with a collection semiangle of 4.5-22 mrad. This emphasized the presence of a number of isolated dislocations and a (111) twin boundary, which was also seen in the diffraction pattern.

Again, there are a number of features that experimentally match the simulated data. First, in the pure phase correlation, the effects of noise and other outliers in the diffraction pattern images cause significant noise in the disk registration. This suggests that phase correlation lacks the robustness needed to ignore subtle variations in background noise.

It is harder to recognize the failure of pure cross correlation at a glance, but upon closer examination, large changes in strain are measured across flat regions of the sample (where the LAADF and ADF images have little to no change in intensity). The LAADF image does
Figure 3.5. A comparison of the four different disk registration methods and the resulting strain measurements around a twin boundary in an austenitic stainless steel. a) A LAADF image emphasizing dislocation contrast, taken prior to the NBED scan. b) A map of the summed correlation values for all diffracted peaks. The green box is the chosen zero strain reference region used in computing the strain maps. c-d) Average diffraction patterns for the right and left side of the twin boundary, showing the symmetry of the system. Additionally, x and y directions are shown in red and blue, correctly oriented with the strain maps. e-h) Strain maps computed using different correlation methods of the region of interest.
not support the dislocation density to realistically achieve a 4-6% change in strain across the FOV. Instead, this erroneous measurement can be attributed to the effect of changing inner disk illumination conditions as shown in the simulated data.

Finally, the two intermediate disk registration methods both fared much better on this dataset. While the hybrid method was noisier than the Sobel-filtered cross correlation image, the strain measurements across the image lack the large changes in strain typical to cross correlation, and measurement noise is minimized. This is apparent from the relatively smooth changes in strain and realistic values seen in their strain maps. For both the hybrid and Sobel-filtered maps, the strain fields around the individual dislocations can clearly be seen, without abnormal variations in uniformly-illuminated regions. Therefore, we found both the hybrid and Sobel-filtered cross correlation methods to provide the best strain field reconstruction of the nanobeam diffraction dataset shown here.

### 3.4 Conclusion

Four separate correlation methods were tested against simulated disks, an unstrained Si sample, and an engineering alloy. We show that when the diffracted disks have inner structure, methods such as pure phase or cross correlation often include artifacts that can be eliminated by Sobel filtered cross correlation or hybrid correlation disk registration methods. Additionally, we show that for low signal to noise disks, binning can help the accuracy of the disk position measurement when using Sobel filtered, phase or hybrid correlation. This is increasingly relevant due to advancements in detector acquisition speed, which can lower the dose per diffraction pattern for a given total beam current. Finally, we compared the four selected methods on two different experimental datasets, with similar results. Both cross and phase correlation had irrecoverable flaws (large systematic errors and a high degree of noise respectively), while hybrid and Sobel filtered cross correlation yielded optimal disk registration results, and therefore had the most accurate strain measurements.
Chapter 4:

\textit{in situ} strain mapping of 321 stainless steel

4.1 Introduction

The deformation of crystalline materials is highly dependent on defects and their response to applied stress [62, 63]. Understanding the motion and dynamic response of these defects, as well as their interactions with each other, is key to future alloy development. A new tool for quantifiably probing these interactions over a large area is scanning nanobeam electron diffraction (NBED). NBED [19, 21, 45, 55, 64] is a scanning transmission electron microscopy (STEM) technique that allows for the quantification of strain fields around features of interest at the level of single nanometers. This method, which utilizes a semi-converged electron beam, is faster than methods utilizing electron beam precession [23, 46, 55], and can acquire a larger field of view when compared to methods requiring atomic resolution [45, 46, 65]. Additionally, while NBED requires a sample close to a zone axis, it is relatively robust to imperfect sample tilt, which is all but guaranteed during an \textit{in situ} deformation experiment [45]. This combination of speed, robustness to sample orientation [22], and increased field of view has allowed for the development of \textit{in situ} strain mapping, in which the sample is mechanically deformed while successive strain maps are acquired [66]. While traditional TEM based \textit{in situ} mechanical testing has been a useful technique in understanding the mechanisms of deformation [67, 68], by combining nanoscale deformation with simultaneous strain mapping, we now acquire quantitative data on the local strain field evolution and its relationship with dislocation motion.

It is generally known that the dislocation slip character can be classified in one of two ways [69]. Typically, dislocations either move in three dimensions via wavy slip, or in two-dimensional oriented ensembles known as planar slip [70]. There is a wide body of both experimental and theoretical literature that has shown that the dislocation slip character has a large effect on the resulting mechanical properties of the material, including fracture behavior, fatigue resistance, and stress corrosion cracking [71–78]. In close packed materials (both face-centered cubic (FCC), and hexagonal (HCP)), the presence of short range order (SRO) has been found to be the main factor leading to planar slip [70, 79, 80].

The results presented in this chapter and have been published as a regular article with title “\textit{In situ} nanobeam electron diffraction strain mapping of planar slip in stainless steel” in Scripta Materialia \textbf{146}, 87-90 (2018) by Thomas C. Pekin, Christoph Gammer, Jim Ciston, Colin Ophus and Andrew M. Minor [61]. The material is presented here with the permission of co-authors and publishers.
Planar slip features can be readily observed in lightly deformed stainless steels, as shown in Figure 4.1. Figure 4.1a is a TEM bright field micrograph of a single instance of planar slip showing multiple dislocations. Figure 4.1b is an ADF STEM image showing the long range and regular orientation of these defects.

In alloys containing SRO, planar slip begins when the leading dislocation moves through the matrix, and as it moves, shifts the lattice by a Burgers vector, \( \mathbf{b} \). This is thought to lead to the local destruction of the SRO, as the locally ordered atoms are shifted out of their energetically favorable positions into a more disordered state. The result of this shift is glide plane softening, which allows further dislocations to move more easily if they follow the first, thereby forming characteristic planar slip bands [70, 80]. These highly local, heavily disordered regions then have a deleterious effect on macroscopic mechanical properties due to the resulting stress concentrations as opposed to alloys with more homogeneous wavy slip.

4.2 Experimental methods

![Figure 4.1](image)

**Figure 4.1.** a) A bright field TEM image of a single instance of planar slip in 321 stainless steel. Note the orderly array of dislocations. b) An ADF STEM image of the same alloy at a lower magnification. Multiple large arrays of planar slip can be seen oriented along specific crystalline directions. Arrows highlight the location of a few of the more prominent examples.

In this experiment, a AISI 321 FCC austenitic stainless steel was pulled in tension *in situ* in a TEM using a Gatan 654 straining holder. To accurately determine composition, samples were sent to Luvak Inc. for analysis. To determine the majority of the elements in the foil, direct current plasma emission spectroscopy was used [81]. Inert gas fusion was used determine nitrogen composition [82], and combustion infrared detection was used to determine carbon composition [83]. The composition is shown in Table 4.1. The as rolled foil was electrical discharge machined (EDM) into 11 mm by 2.5 mm blanks, annealed at 1060 C.
Table 4.1. The alloy composition as tested.

<table>
<thead>
<tr>
<th>Element</th>
<th>Fe</th>
<th>Cr</th>
<th>Ni</th>
<th>Mn</th>
<th>Si</th>
<th>Ti</th>
<th>C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wt. %</td>
<td>bal. 17.74</td>
<td>9.25</td>
<td>1.71</td>
<td>0.53</td>
<td>0.33</td>
<td>0.031</td>
<td>0.018</td>
<td></td>
</tr>
</tbody>
</table>

for 30 minutes to anneal out many of the dislocations, and then jet polished to electron transparency using a solution of 6% perchloric acid, 39% butanol and 55% methanol at −15°C and approximately 30 volts. The scanning NBED was performed on the TEAM 1 microscope which is double aberration corrected, and the diffraction patterns were acquired using the Gatan K2 IS electron detector at 400 frames per second. The microscope was operated at 300 kV, with a convergence angle of 2.62 mrad and a camera length of 230 mm, resulting in a collection angle of 32-160 mrad. Eight loading steps were applied. As the Gatan 654 holder does not measure load or accurate displacement, the tensile bar was elongated until dislocation motion commenced. Loading was then paused and a 100 by 100 pixel NBED map with a 2.2 nm step size was acquired. At 400 frames per second, each map took 25 seconds. Post-deformation ex situ dislocation characterization was performed on the same sample, but in a different region as the in situ region was destroyed during the in situ test. The dislocation characterization was performed on a JEOL 3010 TEM at 300 kV, and the dislocations were found to be [011] type perfect dislocations, as were all dislocations analyzed in similarly prepared samples. The images for \( g \cdot b \) analysis can be seen in Fig. 4.2. The strain mapping analysis was processed using MATLAB, using the hybrid method detailed in [22]. The reference lattice chosen was the mean diffraction pattern for each frame, resulting in a map showing the deviation from the mean strain at each discretely applied load. Post-processing included cropping to remove exceptionally noisy areas and smoothing. The smoothing was applied using the VBM4D MATLAB package [84], which is optimized for video data.

4.3 Results and discussion

The results can be seen in Figure 4.3. Three discrete time-steps are shown, with a, e and i showing the annular dark field (ADF) images acquired simultaneously at each step. Several features can be seen in the ADF images, including a strong bend contour, several dislocations, and a circular precipitate used as a reference for aligning the video frames. However, only one dislocation is caught moving in the field of view from left to right. This is the top dislocation in the image, and it moves a small amount from a to e, while in i it has fully propagated across the field of view. The white arrows denote the location of the leading dislocation. As the dislocation moves, it leaves behind a region of increased contrast in the ADF images. This region of higher contrast matches the band directly beneath it, which was observed previously in the same experiment to be another instance of planar slip in parallel with the moving dislocation shown here.

Strain maps are also shown in Figure 4.3. The maps are rather complex as we are measuring the transient strains around multiple dislocations and precipitates in situ. Examining the evolution of strain with respect to deformation, there are some very interesting obser-
Figure 4.2. a) An instance of planar slip. The black arrow shows the direction of slip. b) Visible dislocations. The $\mathbf{g}$ direction is given in the top right. c-d) Dislocations satisfying the invisibility criterion. These dislocations are all part of the same planar slip feature in a). e) Same area as d) showing the dislocations. The white arrows show landmark dislocations, while the black arrow shows the same instance of planar slip as in a) that is invisible in d). Through $\mathbf{g} \cdot \mathbf{b}$ analysis, the dislocations were found to have a Burgers vector direction of [011].

...vations. As the leading dislocation of planar slip moves from left to right, the strain in the 200 direction increases behind it. However, in the perpendicular 022 direction, strain remains constant, showing clear directionality when it comes to the lattice expansion. The evolution during loading of the lattice expansion $((\epsilon_{200} + \epsilon_{022})/2)$ is shown in Figure 4.3d, h, and l. Color coded line profiles (10 pixel integration width) from between the arrows in the expansion maps are shown in subpanel n. While the consistent large dips in the line profiles are due to deleterious effects of sample bending during strain mapping analysis, the profiles reveal that in the second loading step, there is a small lattice expansion as the dislocation moves from left to right, and then at the final loading step, the entire profile is shifted up roughly 0.15%. This corresponds to lattice expansion of 0.3%. This small, directional lattice expansion is not unexpected. In fact, when observing planar slip with standard TEM techniques, it is somewhat characteristic of planar slip to leave behind weak fringing. While we do not see the fringing in our ADF images during strain mapping due to the zone axis imaging condition, we can see them in Figure 4.4, which is a representative ex situ 2-beam TEM image of a similar planar slip band from an intact region after releasing the tensile...
Figure 4.3. a,e,i) ADF images obtained at increasing levels of deformation. The white arrows denote the leading dislocation location. b,f,j) Strain mapped in the 200 and c,g,k) 022 directions. The black arrows correspond to the leading dislocation location. d,h,l) Lattice expansion behind the leading dislocation of planar slip. m) Correctly oriented mean diffraction pattern showing the crystallographic directions chosen for strain axes. n) 10 pixel integrated line profiles of the area between the color coded arrows in d, h, and l. The horizontal axis corresponds to position from left to right between the arrows, and the vertical axis corresponds to lattice expansion. A lattice expansion can be seen behind the planar slip front.
**Figure 4.4.** Representative TEM image of planar slip in 321 stainless steel. The arrows denote the locations of perfect (011) dislocations, and the fringe pattern perpendicular to the arrows/dislocations is indicative of a residual displacement, or strain, in the planar slip plane.

load for deformation.

The weak fringing observed in Figure 4.4 is another manifestation of the residual displacement or strain that was observed behind the first dislocation measured in Figure 4.3. This fringing can be correlated to the destruction of SRO [85] by a dislocation passing through the matrix, and leaving behind some residual lattice displacement. In previous studies, by using quantitative image matching with simulations, the estimated displacement that results in residual fringing was found to be between $\frac{1}{150}$ and $\frac{1}{200} \langle 111 \rangle$ in a FCC Cu-Al alloy [86], $-\frac{1}{50} \langle 110 \rangle$ in a FCC Pd-Ce alloy, and between $\frac{1}{104} [11\overline{2}0]$ and $\frac{1}{145} [11\overline{2}0]$ in an HCP Ti-Al alloy [85]. Particularly for the FCC Cu-Al alloy, the displacement when converted to strain ($\text{displacement}/\sqrt{3}$) is in the range of 0.28-0.38%, which is very similar to the values measured using our NBED technique. While the fringing in the two previously studied FCC alloys is attributed to atomic size differences and unfavorable atomic positions after the destruction of SRO, in the Ti-Al system this cannot be the case due to relatively similar atomic radii, and instead is hypothesized to be caused by unfavorable nearest neighbor interactions [85, 87]. It is important to note that the sample studied here has a much more complex chemistry than the previously studied materials, and thus we must be cautious about a direct comparison between lattice expansion of the Cu-Al system and ours. The stainless steel contains multiple substitutional elements with roughly similar atomic size and weights, as well as several light interstitial elements, and thus making strong conclusions about the exact mechanisms behind this observed strain difficult. However, we believe the general mechanism should be the same as the previous studies, in which the destruction of SRO and a new, unfavorable atomic configuration causes a local lattice expansion. Hindering our ability to confirm this hypothesis is the difficulty of observing or measuring SRO in the alloy studied, in which all the substitutional species scatter equivalently in an electron microscope, and in
which the primary interstitial elements (C and N) cannot be directly imaged.

More importantly, we showcase in this paper a technique in which a lattice expansion is directly measured during *in situ* deformation via a nanobeam electron diffraction experiment. This is important for two reasons. First, it shows that it is possible to record the dynamic processes that occur during the deformation of a material and extract quantitative data from a previously qualitative technique. This has relevancy for *in situ* deformation as shown here, but it should be able to be extended to many other *in situ* techniques, including temperature control and electrical biasing. Second, particular to this experiment, a direct measurement of the expansion due to planar slip can be made without resorting to quantitative image matching at multiple zone axes. Using the *in situ* scanning NBED technique, we have for the first time directly measured this expansion in a commercially relevant stainless steel, and shown that it is due to dislocation motion *in situ*, rather than inferring it from ex situ measurements.
Chapter 5:  

*in situ* nanobeam electron diffraction of bulk metallic glass

5.1 Introduction

Bulk metallic glasses (BMGs) are an interesting class of materials noted for their wide variety of mechanical properties, associated most notably with their lack of long-range crystallographic order. BMGs include alloys that exhibit extremely high strength in excess of most engineering materials [88, 89], as well as low stiffness and high elastic strain limits [90, 91]. Because of these wide-ranging properties, BMGs are attractive alloys for future applications as they offer a potential for the development of stronger and tougher structural materials [92–95].

One of the main impedances to the adoption of high-strength BMGs is their limited ductility, which is restricted by single shear band formation and rapid propagation at low strains, which often results in catastrophic failure [96–98]. Accordingly, of key importance to further alloy development is understanding how such shear bands originate at the nanoscale because, although single shear-band formation can cause BMGs to fail at near-zero tensile ductilities, multiple shear-band formation represents the fundamental essence of plasticity in these alloys. As BMGs invariably display high strength, the creation of tensile ductility — via multiple shear banding — is thus essential to their fracture toughness, and hence damage-tolerance, in terms of their potential role as future structural materials.

Several mechanisms have been proposed for the initiation and propagation for shear bands, the predominant hypotheses being free volume softening [99–104], adiabatic heating softening [105–107], and shear transformation zones (STZs) [104, 108, 109], with the latter currently being the most widely accepted. Recently, a mechanism based on geometrically unstable motifs (GUMs) of atoms has provided a link between local atomic arrangements of bulk metallic glasses, and STZs [104, 110, 111]. In this theory, GUMs are disordered clusters of atoms that, due to their configuration and poor packing, can serve as nucleation sites for...
shear transformation zones, which propagate into shear bands. This theory has important implications as it provides a directly tunable parameter that could be used to change the mechanical behavior of BMGs, akin to reducing the grain size in polycrystalline materials.

To date, observing the mechanisms of BMG shear band formation, while possible in molecular dynamics (MD) simulations, has been experimentally challenging due to the high rate of the catastrophic shear band propagation and the current experimental limits of electron microscopy. However, observing shear band nucleation and dynamics at the scales possible in transmission electron microscopy (TEM) is crucial to linking our understanding of deformation mechanisms provided by MD simulations to the macroscale mechanical behavior. Previous TEM experiments in bulk metallic glasses have largely been limited to ex situ qualitative imaging studies [112–115], or more quantitative fluctuation electron microscopy (FEM) studies on the structure of BMGs [32, 116–118] that fall below the spatial resolution needed for shear band characterization. in situ experiments to date have been qualitative and at too low of a spatial resolution to be comparable to MD models [6]. Recent advancements in techniques and hardware have, however, allowed for the observation of strain [24] and as we will show here, the evolution of local atomic order, with nanometer resolution during in situ deformation, providing much more comparable information to the significant modeling efforts which have been performed.

5.2 Experimental methods

In this study, we designed an in situ sample to study the effect of local order and strain during tensile deformation. The BMG used in this study was a member of the model glass family [119, 120] Cu_{x}Zr_{y}Al_{100-(x+y)}, which has been extensively studied for its high glass-forming ability [121], and relative ease of computational modeling. These glasses have clusters of atoms that pack into stable icosahedral structures [122, 123], which due to their two-, three- and five-fold symmetry axes in projection, have characteristic symmetric diffraction patterns [25].

Specifically, the sample used was Cu_{46}Zr_{46}Al_{8}, which was thinned to electron transparency and then milled using a focused ion beam (FIB) into an in situ tensile bar specimen. The angular dark field images (ADF) of the resulting sample and subsequent deformation can be seen in Fig. 5.1a. Unique to this in situ experiment, at each step of deformation a nanobeam electron diffraction (NBED) dataset was acquired, in which a full diffraction pattern was acquired for each annular dark field image pixel, for a total of 167,440 diffraction patterns. The diffraction patterns were then used to measure the evolution of both local strain [24] and order [25, 124] as the sample was mechanically deformed, with a spatial resolution of 2 nm.

5.2.1 Detailed experimental methods

The initial sample was received as a cylindrical bar with an outer diameter of ~4 mm. The bar was mechanically machined down to an outer diameter of 3 mm, and electrical discharge

---

1 We would like to acknowledge Linzhi Zhao, Yanhui Liu and Weihua Wang for providing the as-processed BMG samples.
Figure 5.1. a. Annular dark field (ADF) scans taken before, during and after in situ deformation. The change in the size and shape of the hole is indicative of plastic deformation. The lack of contrast change across the specimen suggests constant sample thickness. The green dashed box shows where NBED was performed. b. Schematic showing the NBED process in the metallic glass. As the beam rasters over the area, a full map consisting of over 33,000 nanobeam electron diffraction patterns is recorded. The three patterns shown are examples containing two-fold, four-fold and zero symmetries, respectively.

machined into 600 µm thick slices. The samples were further mechanically thinned to ~150-200 µm, and then jet polished to electron transparency. The jet polishing solution used was 33% nitric acid in methanol, cooled to -25 C, and a polishing voltage of 18 V.

in situ samples were then cut and lifted out using a FEI Strata 235 dual beam focused ion beam (FIB) equipped with an Omniprobe. The samples were welded on to a Hysitron push-to-pull chip using deposited platinum. A hole was cut in the center of the tensile bar to concentrate stress and strain, allowing for higher resolution scans in the critical area. The sample after liftout can be seen in Fig. 5.2.

The in situ nanobeam electron diffraction (NBED), pre- and post-experiment imaging was performed on the TEAM 1 transmission electron microscope [125], a \( C_s \) and \( C_a \) corrected FEI microscope. The microscope was operated in scanning transmission electron microscopy (STEM) mode with an accelerating voltage of 300 kV. The microscope was operated in three condenser lens nanoprobe mode, with a spot size of 10, 10 µm aperture, and a convergence
Figure 5.2. The sample after focused ion beam milling on the Hysitron push to pull chip. a) Top down view. b) View at 52 degrees of stage tilt. The nominal thickness is approximately 100 nm. Holes were milled in the center of the tensile specimen in order to concentrate stress and strain.

angle of 0.91 mrad. The probe was measured in real space to have a full width half maximum diameter of 1.47 nm.

In situ deformation was performed in the TEM using a Hysitron PI-95 in situ deformation holder, with the sample mounted to a push to pull device. The sample was deformed under displacement control, in steps of 10 nm, up to a total displacement of 40 nm when the sample broke. NBED datasets with a size of 182 by 184 probe positions were acquired with a Gatan K2-IS camera at 400 frames per second, with a probe step size of 2.5 nm. Each scan took 70 seconds while deformation was paused. The camera acquired a 1792 by 1920 pixel full-frame diffraction pattern at each real space pixel location. Five complete datasets were acquired, corresponding to tensile deformations of 0, 10, 20, 30, and 40 nm. The sample broke at some point between 30 and 40 nm of deformation. The load-displacement curve acquired during deformation can be seen in Fig. 5.4f. The vertical jumps in load correspond to periods in which scans were acquired. The load-displacement does not show pure elasticity, instead at high deformations, plasticity is observed.

The NBED datasets were acquired in a sparse electron scattering regime, which was controlled by reducing the beam current to the sample. This allowed for the counting of individual electron locations, resulting in a massive reduction in the amount of data, and a large increase of signal to noise. Once the patterns were reduced to electron locations, patterns with reduced noise were reconstructed. These patterns were shifted to remove the effect of beam sway during scanning. For the strain mapping analysis, every pattern was locally summed with its nearest neighbors, using a Gaussian weighting centered on the center pixel with a standard deviation width of two pixels. These diffraction patterns were then used for strain mapping following the methods outlined in ref. [24], in which an ellipse is fitted to every diffraction pattern, and deviations from a reference radius are converted to strains. The reference radius was chosen to minimize strain at zero loading.
The mapping of local symmetry followed the methods proposed in refs. [25, 124, 126], after correction for elliptical astigmatism produced by both microscope aberrations as well as strain. To determine the ellipticity, we have fitted the ring intensity $I(q_x, q_y)$ over the reciprocal space coordinates $(q_x, q_y)$ using the following form

$$I(q) = I_r \exp \left[ -\frac{\left(q_0 - \sqrt{Aq_x^2 + Bq_xq_y + Cq_y^2}\right)^2}{2s^2} \right] + I_0 + I_1 \sqrt{q_x^2 + q_y^2},$$  \hspace{1cm} (5.1)$$

where $I_r$, $I_0$ and $I_1$ are the intensities of the ring, constant background, and linear background respectively. $A$, $B$, and $C$ are the ellipse coefficients, and $s$ is the standard deviation of the ring, equivalent to its width. $(q_x, q_y)$ is centered on the center of the diffraction pattern. The resulting fitted ellipse can then be represented in matrix form as

$$\mathbf{M}_{\text{ellipse}} = \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix}$$  \hspace{1cm} (5.2)$$

Then, to remove elliptical distortions, each electron location $(q_x, q_y)$ can be transformed by

$$(q'_x, q'_y) = (q_x, q_y) * \mathbf{v} * \sqrt{\mathbf{d}} * \mathbf{v}^T$$  \hspace{1cm} (5.3)$$

where $\mathbf{v}$ is the eigenvectors of $\mathbf{M}_{\text{ellipse}}$, $\mathbf{d}$ is the eigenvalues corresponding to $\mathbf{v}$, and $\mathbf{v}^T$ is the transpose of $\mathbf{v}$.

The symmetry of each pattern then is measured from the binned polar transforms of the electron coordinates for each pattern to $(r, \theta)$ space, as explained in [25, 124, 126]. This method normalizes the amount of symmetry order measured by the intensity 0th order amount, which removes the effects of changing diffraction pattern intensity (sample thickness, bending, etc.).

5.2.2 Measuring strain from fitted ellipses

The diffraction pattern of an amorphous material is characterized by a diffuse ring pattern. The first ring can be used to determine the strain, as deformation induces a small deviation from an ideal circle [127, 128]. The position of the first ring as a function of scattering angle, $q(\theta)$, is compared to the radius of the ideal ring of the unstrained case, $q_0$. The resulting strain, $\epsilon$, a function of $\theta$, can be separated into the axial, tangential and shear strains $\epsilon_{xx}$, $\epsilon_{yy}$ and $\epsilon_{xy}$ [24].

$$\epsilon(\theta) = \frac{q_0 - q(\theta)}{q(\theta)} = \epsilon_{xx} \cos^2(\theta) + \epsilon_{xy} \cos(\theta) \sin(\theta) + \epsilon_{yy} \sin^2(\theta).$$  \hspace{1cm} (5.4)$$

This equation can be rewritten to

$$\frac{q_0}{q(\theta)} = 1 + \epsilon_{xx} \cos^2(\theta) + \epsilon_{xy} \cos(\theta) \sin(\theta) + \epsilon_{yy} \sin^2(\theta).$$  \hspace{1cm} (5.5)$$
Figure 5.3. Experimental geometry used in this study. A dog-bone shaped metallic glass sample is deformed in situ in the scanning electron microscope. The sample is placed at the cross-over position of the convergent electron beam. The convergent electron beam diffracts from the sample, leading to the characteristic amorphous ring pattern. A best fit ellipse is computed for each diffraction image recorded at each probe position. The elastic strain induced during loading leads to an elliptic distortion of the amorphous ring pattern.

Once Eq. 5.1 has been fitted to all probe positions, the strain can be computed from the results for $A$, $B$ and $C$. $A$ and $C$ will be very close to a value of 1 and $B$ close to 0. Therefore, we define $A = 1 + a$ and $C = 1 + c$. Converting the ellipse to polar coordinates $(q_x = q(\theta) \cos(\theta), q_y = q(\theta) \sin(\theta))$ yields

$$q(\theta)^2 ((1 + a) \cos^2(\theta) + B \cos(\theta) \sin(\theta) + (1 + c) \sin^2(\theta)) = q_0^2$$

$$\frac{q_0}{q(\theta)} = \sqrt{1 + a \cos^2(\theta) + B \cos(\theta) \sin(\theta) + c \sin^2(\theta)}$$

This expression can be approximated, by taking into account that $a$, $B$ and $c$ are small

$$\frac{q_0}{q(\theta)} \approx 1 + \frac{1}{2} (a \cos^2(\theta) + B \cos(\theta) \sin(\theta) + c \sin^2(\theta)).$$
A comparison of this equation with Eq. 5.5 allows deducing the strain

\[
\begin{align*}
\epsilon_{xx} & \approx \frac{1}{2}(A - 1), \\
\epsilon_{xy} & \approx \frac{1}{2}B, \\
\epsilon_{yy} & \approx \frac{1}{2}(C - 1).
\end{align*}
\]

(5.8)

By determining the ellipticity for each diffraction pattern in the NBED map, a full two-dimensional strain map is obtained. When the MG specimen is deformed in situ in the TEM, the local elastic strain changes resulting in a change in the elliptic distortion (see Fig. 5.3). By obtaining a time series of NBED maps, time resolved strain maps can be calculated and the change in the local strain during deformation can be mapped in situ. Note that all real experiments will contain some elliptic distortion even for non-strained samples, due to aberrations in the projector lenses of the microscope. Therefore, a reference has to be defined and the distortions measured in this region are used as reference values. We have used the measured ellipticity of the MG specimen prior to deformation as the reference strain state.

5.2.3 Computational molecular dynamics methods

Large-scale molecular dynamics (MD) simulations were performed by Jun Ding to study the Cu₄₆Zr₄₆Al₈ model metallic glasses (MGs), using the optimized embedded atom method (EAM) potential, adopted from ref. [129]. The sample contained about 5 million atoms, and the liquids of the sample were equilibrated for 1 ns at high temperature (2500K) to assure equilibrium and then quenched to room temperature at the cooling rates of 10¹² K/s employing a Nose-Hoover thermostat (the external pressure was barostated at zero) [130]. Periodic boundary conditions were applied in all three directions during the quenching [130]. The prepared Cu₄₆Zr₄₆Al₈ metallic glass sample had dimensions of 78.3 nm by 91.5 nm by 13.1 nm at 300 K. A hole with the radius of 16 nm was created in the middle of x-y plane. Then the boundary condition in x-direction was set as a free surface. The as-processed sample was then gradually heated to 680 K (below its glass transition temperature Tₙ) and annealed for 0.5 ns to reach a steady potential energy. The final step of sample preparation was to quench it from 680 K to 300 K with a cooling rate of 10¹² K/s. For the simulation of the deformation process, the Cu₄₆Zr₄₆Al₈ model metallic glass was under uniaxial tension in the y-direction with a strain rate of 10⁻⁴ /ps at 300 K under NVT ensemble. The local von Mises shear strain was analyzed using the algorithm in ref. [131], by comparing the deformed configuration with the original one. We monitor the local structural order in Cu₄₆Zr₄₆Al₈ metallic glasses by conducting Voronoi tessellation [132]. Faces of Voronoi cell with area smaller than 0.25% of the total area are discounted. Specifically, Fig. 5.6f-j are the coarse-graining plots of the fraction of full icosahedral order (with the Voronoi index (0, 0, 12, 0)) in the x-y plane, which is divided into pixels with dimensions of 2 nm by 2 nm and averaged over the whole z-direction of the sample.
Figure 5.4. a-e. Strain maps at increasing deformation steps. The top row is strain in the tensile direction, the middle row is strain perpendicular to the tensile direction, and the bottom row of strain maps is shear. By step d, plastic strain has developed in the tensile direction. f. Load-displacement plot acquired from the in situ TEM mechanical testing holder. The blue solid line is the data acquired, with the displacements corresponding to the different map acquisitions a-e marked with red squares. The orange dashed line is the least squares fit to the elastic regime, while the yellow dashed line corresponds to the plastic region. The change in slope between c and d is indicative of plastic deformation. The sample fractured before e. The drift in the plot corresponding to the hold times during the NBED acquisitions at the 10, 20 and 30 nm steps has been removed. g-k. Maps showing local order at increasing deformation steps as measured from the diffraction patterns. The top row corresponds to two-fold symmetry, while the bottom row corresponds to four-fold symmetry. l. A plot showing the relative mean amount of two-fold (blue +) and four-fold (green square) order in the corresponding areas in g at successive deformations. The plot shows roughly a 6-7% reduction order relative to the maximum.
5.3 Results and discussion

Fig. 5.4 shows the results obtained after processing the NBED data acquired during the in situ experiment. The top three rows of images (Fig. 5.4a-e) show relative strain as the experiment progresses in the orthogonal, and shear directions, while the bottom two rows of images (Fig. 5.4g-k) show the degree of two and four fold order. A clear evolution of strain is observed during deformation. The strain concentrates itself as expected around the hole in the center of the sample, as well as along the shear directions 45° to loading directions. Examining the maps more closely, during the first three steps of deformation corresponding to 0, 10, and 20 nm of displacement, very little changes with regards to strain. These steps correspond to elastic loading of the sample, which can be observed in the load-displacement curve shown in Fig. 5.4f. As the sample is deformed to 30 nm, nonlinearity of the load-displacement curve begins to occur, indicative of plastic strain. In the strain maps at this deformation (Fig. 5.4a-e), clear strain concentrations occur on both sides of the FIB milled hole. These thin areas experience strain up to 2% above the median strain of the sample, leading to local regions of plastic deformation and failure. After fracture, the strain returns to a uniform value across the field of view.

In conjunction with the evolution of local strain, we also measured the change in local order, or the atomic clustering of the metallic glass with deformation. The results can be seen in Fig. 5.4g-k. As previously shown in the literature [25], by measuring the symmetry elements found in each diffraction pattern, we can map local order. As expected, initially the sample had a uniform distribution of order across the region of interest. This is expected for a rapidly quenched bulk metallic glass. This does not change during the elastic deformation of the sample (10 and 20 nm of deformation). However, once plasticity begins, we observe the destruction of local atomic order spatially confined to the regions of high plastic strain. In the two-fold order map Fig. 5.4j, this destruction is confined to the highest strain region on the sample, on the right side of the hole. In the four-fold order map (Fig. 5.4j, bottom), this destruction is more prevalent and can be seen on both sides of the hole, but again, is confined to the high strain regions of the sample. Fig. 5.4l shows the mean amount of order in the color-coded dashed polygons shown in Fig. 5.4g for the first four steps before fracture. Within these plastically deforming regions, there is a ~7% reduction in order from the maximum for both two- and four-fold symmetries. The difference in size between the two- and four-fold symmetry areas of reduced order is most likely due to the mechanics of order destruction and diffraction, namely that the clusters of projected four-fold symmetry have more avenues to move out of Friedel diffraction than those in two-fold symmetry.

In order to statistically measure the reduction in order in the two- and four-fold order maps in Fig 5.4l, basic statistical methods were performed.

A standard two-sample $t$-test was performed in MATLAB, which tests the null hypothesis that the data in vectors $x$ and $y$ come from independent random samples from normal distributions with equal means and unequal variances, using the first and fourth step of deformation (just prior to fracture) as $x$ and $y$. In both two-fold and four-fold symmetry orders, the null hypothesis was rejected in the masked areas with a $p$-value of 0.

Finally, we have plotted histograms of the four-fold order coefficients from the masked areas at the first and fourth step of deformation in Fig. 5.5. This graphically shows the
Figure 5.5. Histogram showing the reduction in order in the four-fold order maps in Fig. 5.4, corresponding to deformations a and d. The order coefficient values follow the normal distribution, and the two-sample $t$-test rejected the null hypothesis that the data came from distributions with equal means and unequal variances, with a $p$-value of 0.

same result as the $t$-test, which is that the means are clearly different, the distributions are normal, and the variances are unequal.

To help interpret these results, large-scale molecular dynamics (MD) simulations [130] were performed by Jun Ding on a 5-million-atom sample with a similar geometry, although with reduced dimensions for practical considerations. The alloy composition was the same as tested experimentally. The parity between experimental and simulated sample geometry and composition allows for their direct correlation. The MD results can be seen in Fig. 5.6 that shows five frames of increasing deformation. The process of strain localization around the region of plastic deformation is shown in Fig. 5.6a-e; the images are colored according to the local von Mises shear strain. Under the uniaxial tension in the $y$-direction, the strain localization begins to aggregate near both sides of the hole (Fig. 5.6a); with increase in applied strain, progressively more strain localization evolves out along the two 45° directions of maximum shear stress. Finally, the continuously induced strain localizations percolate across the sample from the hole to the free surfaces, thus forming the shear bands seen in Fig. 5.6e. This deformation process revealed by the MD simulation is generally consistent with our experimental TEM characterization shown in Fig. 5.4. Similarly, Fig. 5.6f-j show the spatial distribution of local icosahedral order for the studied five frames, where a corresponding reduction in local symmetry is seen as deformation progresses. Interestingly, the regions with reduced local icosahedral order (blue regions in Fig. 5.6f-j) are almost overlapped with the strain localization in Fig. 5.6a-e. It is therefore apparent that the results of our large-scale MD simulation are fully consistent with the direct observation by nanobeam electron diffraction microscopy that correlates the strain localization and local structural transition in BMGs under deformation.

Despite experimental differences in sample size and deformation rate, previous MD experi-
Figure 5.6. Five frames of MD-simulated Cu$_{46}$Zr$_{46}$Al$_{8}$ MG at the strain of 4%, 5%, 6%, 7% and 8% respectively, under uniaxial tensile deformation. a-e a color map of local von Mises shear strain and f-j describe the spatial distribution of local icosahedral order in coarse-graining scale, for such five frames.

Experiments as well as the current simulation observe a mechanism of metallic glass deformation which relies on slightly disordered atomic clusters (GUMs) shearing and deforming under lower strains than their more ordered icosahedral cousins. In simulations, the result of the deformation was a more disordered region, and shear band initiation. Similarly, this is what we see for the first time in the NBED in situ deformation experiment. At high strains, local atomic clusters containing symmetry are destroyed before shear band propagation, and their destruction generates increasing local plasticity and shear band formation under further strain.

Our results confirm the important hypothesis of GUM formation leading to shear banding and fracture as described by prior computational models but never experimentally observed before. In order to improve the limited tensile ductility currently preventing BMGs from widespread use, resulting primarily from deformation localization in single shear bands, our experiments suggest that future alloy design should follow a pathway in which the amount of disordered atomic clusters is precisely tuned to accommodate more homogeneous ductility. We also believe that the combination of in situ microscopy and NBED as a characterization technique is well suited to image structural deformation characteristics on the nanometer scale in nominally disordered materials. The direct correlation between quantitative in situ deformation experiments and large scale MD simulations on the same length scale can serve as a crucial link between simulations and bulk mechanical properties.
Chapter 6:
Summary and future work

6.1 Conclusions

We began this dissertation with a short history of electron microscopy and its advancements, and provided the motivation behind using it as a tool for materials science. The development of characterization methods is an integral part of materials science, and the insight it provides allows for advancements in the field.

In Chapter 2, we gave a broad overview of the many sample preparation methods and in situ geometries used in this dissertation. We compared and contrasted FIB based sample preparation to electropolishing, with the conclusion that electropolishing was generally easier, but FIB fabrication allows for the use of advanced holders like the PI 95. We hope that others will find this chapter useful in their own sample preparation methods, as it is the most important part of electron microscopy.

Chapter 3 begins the original research presented in this dissertation. This chapter discusses the work published in [22], which covers algorithm development for NBED strain mapping applications. The advent of fast cameras for NBED strain mapping has led to higher data processing requirements, and faster algorithms. We showed that an algorithm with reduced complexity (correlation with a template, using either the hybrid approach, or Sobel filtering before correlation) can be robust to noise and dynamical diffraction effects. For completeness, we tested this with both simulated data representative of what is seen in experiments, as well as two different real materials, a known zero strain silicon sample and a stainless steel sample with both a grain boundary and dislocations. Additionally, we evaluated the effect of binning and signal-to-noise ratio, and found that when signal-to-noise is low, binning can improve the accuracy of the strain mapping.

The positive result in Chapter 3 enabled the results presented in Chapter 4, in which we used the strain mapping algorithm in conjunction with the Gatan 654 holder to observe planar slip in a 321 stainless steel sample. We were able to observe in situ an expansion of the lattice following the leading dislocation of planar slip. This agrees with both theory and previously observed effects ex situ, but provides valuable information on the dynamic strain effects of planar slip. Additionally, this proves it is possible to perform in situ strain mapping on complex engineering materials, opening the door for future studies of deformation and the resulting strain fields in advanced engineering materials.

In Chapter 5, we transition from analyzing NBED datasets from crystalline materials to ones from amorphous materials, characterized by a diffuse nearest neighbor diffraction ring. In this chapter we develop the methods to map both strain and local order across amorphous BMGs. We find that we are able to observe strain during in situ deformation by looking at the change in ellipticity of the first diffraction ring. Furthermore, we find that we are able to
map local order by measuring the symmetries present in the first diffraction ring, where areas that have more order have symmetries present. We then take these method advancements and apply them to a Cu$_{46}$Zr$_{46}$Al$_8$ bulk metallic glass sample mounted on a PI 95 PTP. The results show significant strain concentration around the FIB milled hole, as well as a marked reduction in local order in the high strain regions prior to shear band formation. This is indicative of STZ activation, and is an important confirmation of amorphous glass deformation theory.

### 6.2 Future work

Chapters 2, 3, 4, and 5 each have room for future development. With regards to Chapter 2 and sample preparation, there is significant room to develop more interesting and reliable geometries using FIB milling, as well as novel *in situ* tests. These might include the combination of environmental TEM with *in situ* deformation, *in situ* tests that are more comparable to traditional mechanical tests, or liquid cell deformation. Additionally, electropolishing is not, and might never be, a solved problem. While perhaps not novel enough for a PhD dissertation, there is ample room for undergraduate and masters-level research on this topic.

Algorithm development, covered in Chapter 3 has ample room for future development, and the current research has barely scratched the surface of what is possible. Automated image analysis has just begun to have an impact in materials science, but is an established field in the computer vision and machine learning communities. Materials science should take the best of these fields and apply it to the engineering challenges we find. There is no reason why the same technology Google used to find a cat in an arbitrary video in 2012 [133] cannot be used to find simple diffraction disks in 2018. Machine learning and advanced analysis techniques are the natural progression for the large NBED datasets, especially since the algorithms can be trained with highly accurate simulated data. This has started [134], but the use of advanced algorithms is in its infancy. For the impact it could have, few people are working on it. Additionally, there is room for advanced statistical techniques, such as non-negative matrix factorization (NNMF), when it comes to classification and segmentation of the NBED datasets into different experimentally important basis sets (grain classification, orientation, multiphase materials).

With regards to *in situ* NBED of crystalline samples, as covered in Chapter 4, this dissertation merely shows a proof of concept that it is possible to measure strain during deformation. There are many unanswered questions in materials science that this technique could help answer. For example, it would be of interest to perform a similar experiment which observes the interaction between dislocations and precipitates or grain boundaries. While this has been performed [15, 135, 136], the strain state that governs these interactions is unknown and would help greatly in the accurate modeling of these phenomena. Additionally, observing the strain fields during crystalline growth or nucleation would be a very interesting utilization of NBED strain mapping during *in situ* heating. Watching the evolution of strain with regards to precipitate nucleation and growth kinetics would be scientifically interesting.

One of the greatest hindrances with regards to NBED strain mapping during *in situ* deformation is the requirement for a zone axis diffraction pattern. This is currently experimentally difficult, as both deformation holders used in this dissertation only have one tilt
axis. While the $\alpha$ tilt range is $\pm 35^\circ$, the lack of even $\pm 10^\circ$ $\beta$ tilt means that most \textit{in situ} deformation samples will have one good zone axis for diffraction pattern imaging at most. As the sample undergoes deformation, a uniformly illuminated zone axis diffraction pattern has a propensity to tilt to poor illumination conditions in the $\beta$ direction, as this also is the direction of deformation, and is currently unrecoverable. Therefore, the development of a double-tilt \textit{in situ} deformation holder would greatly enhance the chances of experimental success. While these exist [16, 137], they have not reached widespread adoption or the sophistication of the PI 95 system. Their use in NBED strain mapping would be highly beneficial.

While the results presented in Chapter 5 were more than just a proof of concept, both the data processing techniques as well as experimental methods have wide applications for future development and discovery. With regards to the experimental methods, the optimum experimental parameters have not been discovered for both strain mapping and local order mapping. Part of this is due to the dichotomy between NBED for strain versus NBED for local order. Ideally, to measure strain the electron probe is as parallel as possible (convergence angle approximately 0.5 mrad) to concentrate the most diffracted intensity into the thinnest first diffraction ring, to most accurately measure its ellipticity. However, this leads to a larger probe not suitable for measuring local order, which is improved by probe sizes roughly equivalent to the size of local order (approximately 0.5 nm, obtainable with a 3 mrad convergence angle on TEAM 1). So far, a happy medium to obtain both has not been optimized. Additionally, no tests have been performed to determine if electron counting in a lower dose regime is more effective or prevents more sample damage than increasing the electron beam current and capturing traditional diffraction patterns.

Additionally, the use of advanced statistical techniques has not yet been used on these datasets. For example, the measurement of local order is performed using Fourier transforms of diffraction patterns in polar coordinates. This assumes and uses a basis set of sines and cosines to construct the diffraction pattern. However, other transforms utilizing more representative basis sets could give more accurate results. Additionally, only the magnitude of the Fourier components are used, which is only half of the analysis. Perhaps, by utilizing the phase of the components, rotation of the local order under perturbation (deformation or heating) could be observed!

Some further exploratory experiments have been performed on amorphous materials. Using the same material, samples were heated \textit{in situ} while NBED data was acquired. The initial results look very promising. The transition between amorphous and crystalline structure was observed, including subsequent crystalline growth. This transition occurred simultaneously with chemical segregation as well. This data still needs to be analyzed further, but points to the future of amorphous material NBED. There exists ample possibilities for crystallization nucleation and kinetics studies, in which structural information prior to and during crystal growth is also acquired. Additionally, observing the distribution of strain in a complex crystalline-amorphous alloy with NBED under deformation would be extremely interesting. As this dissertation is a proof of concept for a variety of \textit{in situ} techniques in combination with NBED, there necessarily are many avenues for future development.

Finally, overall there are several avenues for global future improvement on the techniques and experiments presented here. The first is increased multimodal imaging and optics. The
simultaneous acquisition of EDS or EELS spectra during \textit{in situ} NBED would allow for the measurement of chemical composition and diffusion, thickness, or electronic state in addition to the data acquired in this dissertation. This could be uniquely powerful for understanding the transitions that occur in materials. Additionally, the quickest improvement to all of these techniques would be the addition of both in-column energy filtering as well as controllable beam precession. Both of these hardware additions would increase the diffraction pattern quality in both amorphous and crystalline specimens, and would lead to much more experimental success during \textit{in situ} tests. Finally, current research projects at the National Center for Electron Microscopy include faster direct electron detectors. While the current camera can operate at up to 1600 frames per second (fps), this still requires acquisition times on the order of minutes for sufficiently large NBED scans. This becomes apparent during the experiment, as current experimental techniques require a pause in deformation while scans are being acquired. The new cameras, which have promised fps speeds two orders of magnitude higher than the current camera, would allow for acquisition times on the order of single seconds. This will allow acquisition of higher quality, experimentally relevant scans during continuous deformation without the need for pausing, allowing for the capture of events at timescales not currently possible (dislocation jumps, more granular crystal growth).

To conclude, this dissertation covers method development for NBED \textit{in situ} experimentation, and some interesting materials science results that the advancements in the technique have resulted in. This researcher has thoroughly enjoyed this work and believes that there are many future developments for NBED that will only increase its usefulness and impact.
References


Appendix A:
Methods useful for crystalline diffraction patterns

These appendices are used to highlight some specific points during the author’s research that caused particular trouble, or could use additional clarification in terms of examples, code, etc. All code given here is written in MATLAB R2016 and is not guaranteed to be free of bugs or errors.

A.1 Correlation methods

While Chapter 3 does a decent job of discussing the 2D correlation methods, simpler 1D versions are shown here. Figure A.1 was produced using Eq. 3.3, with weights for n of 0 for cross correlation, 1 for phase correlation, or 0.5 for hybrid, or “square root” weighted correlation. A code snippet that created the figure can be seen in Code Block A.1.

![Figure A.1. Graphical representation of a square wave correlated with itself (autocorrelation). All three correlations in the noise free case accurately recover the center of the square wave, but the shape of each correlation is different. The cross correlation is wide, and gradually converges to the maximum correlation. The phase correlation is a delta spike, while the hybrid “square root” correlation is between the two.](image)

```matlab
x = zeros(1,100);
x(25:75) = 1;
correl = conj(fft(x)).*fft(x);
cc = fftshift(ifft(correl)); \% cross correlation
pc = fftshift(ifft(correl./(abs(correl)))); \% phase correlation
hc = fftshift(ifft(correl./(abs(correl).^0.5)))); \% hybrid correlation
```

**Code Block A.1.** Naive implementation of 1D correlation, used to create Fig. A.1.

The transition to 2D correlation is straightforward, and is done by changing the 1D discrete Fourier transforms (DFTs), to 2D DFTs. In MATLAB this is as simple as changing `fft` and `ifft` to `fft2` and `ifft2`, respectively.
Since this correlation is performed in Fourier space, it is a circularly symmetric operation. Therefore, the hard edges on an image as it circularly wraps from right to left and top to bottom are unwanted artifacts that can cause errors in the correlation. To prevent this, a window function should be applied to both the template and object image, which smoothly takes the image to zero outside of some defined interval [138].

A.2 Finding disk positions using correlation

A.2.1 Creating a template

Creating an accurate template is crucial to the effective determination of disk position. As the diffracted disks originate from the unblocked beam, the best template is a vacuum image of the central beam in focus. This is simple to do monochromated microscopes, as the beam intensity can be lowered using the monochromator focus until it will not damage the camera.

If this is not possible, the next best technique involves aligning and summing the center spot from all of the diffraction patterns on a blank background. By also specifying minimum and maximum values for intensity the effects of dynamical diffraction can be minimized. Fig. A.2 shows the profile of both good and bad templates.

![Figure A.2. Illustrated profiles of two templates. On the left is the good template, with a clear maximum intensity and sharp edges. On the right, the template profile is both uneven as well as it contains soft edges, which will reduce correlation accuracy.](image)

A.2.2 Finding disk positions – to subpixel accuracy

Once the template is created, it can be correlated using the methods described in 3. However, a naive correlation in which the maximum of the correlogram is taken as the shift will only yield single pixel accuracy. Here we will present some methods used to improve the accuracy of the measurement.

Sub-pixel correlations: upsampling

The simplest way of improving the correlation is to upsample the correlogram. One simple way to do this is to pad the correlogram in Fourier space with zeros before taking the inverse
Fourier Transform. To do an upsampling by two, an image that is \([n, m]\) in size will need to be inserted into an array of zeros of size \([2n, 2m]\), which is four times larger than the original image. While this may work for upsampling values of two to four, this \(n^2\) dependence on upsampling factor quickly will outpace the memory constraints of the system.

To get around these constraints, much more efficient algorithms have been developed [52, 59]. These algorithms work by first making a correglogram at an upsampling factor of two, and then further upsampling is done by directly computing the DFT using matrix multiplication only in a small area around the peak, to an arbitrary upsampling factor. The code used by the author was a modified version of \texttt{dftregistration} which can be found at [139]. The modifications, which resulted in a function called \texttt{multiCorr} offered the ability to use different correlation methods at the initial step before upsampling.

### Fitting a parabola to correglogram

One can also fit a parabola to the resulting correglogram to maximize the accuracy when finding the peak location. If the three by three region roughly centered on the peak is known, the analytic solution is given by Code Block A.2.

```matlab
1 % icc is a matrix of size [3 3], centered as well as possible on the peak of a curve. dx and dy are the shifts for x and y, respectively, from the center.
2 dx = (icc(3,2)-icc(1,2))/(4*icc(2,2)-2*icc(3,2)-2*icc(1,2));
3 dy = (icc(2,3)-icc(2,1))/(4*icc(2,2)-2*icc(2,3)-2*icc(2,1));
```

**Code Block A.2.** Analytical fit of a parabola to a \(3 \times 3\) curve peak. Original derivation courtesy of Colin Ophus.

With these tools, then all that is left is to combine them in a \texttt{for} loop and find the location of each diffraction spot in every pattern. It is helpful to crop out each diffraction spot individually and correlate them one at a time. Additionally, safeguards such as pixels cannot move more than 1-5 pixels between frame are usually helpful in minimizing error in strain measurement. During this process, it is also helpful to save the correlation score as well as peak location, for later weighting.

### A.3 Fitting a lattice to disk positions

Now that there is a list of peak locations in pixels and correlation score for every peak in all the images, the next step is to fit a reciprocal space lattice to these peaks. It is assumed that the peak locations are stored as a list per image, in a matrix organized by \([\text{xLocation, yLocation, peakIntensity}]\). Once a set of basis vectors that represents each diffraction pattern is known, a lattice can be computed that maps the basis vectors to the disk positions in pixels. Here, we utilize standard least squares fitting methods to the system of equations represented by the equation \(Ax = B\), where \(A\) is our matrix of basis vectors, \(x\) is our lattice to be fit to our disk positions, and \(B\) is our matrix of disk positions. As will be shown below, \(x\) is a \(3 \times 2\) matrix, where the first row is the center location of the matrix, and the next two rows are the two lattice vectors, all in \((x, y)\) pixel locations. Since we typically have many more disk positions than unknowns in our lattice matrix, the system of equations is
very overdetermined. To obtain a more accurate fit, we can use the correlation values of each peak with our template as weights. This causes measurements of clear, bright disks to be weighted more heavily than disks that do not match the template as well, and is a good technique for minimizing the error of the lattice fit.

### A.3.1 Using peak shape as additional weighting for lattice determination

During the research behind Chapter 3, it was theorized that perhaps lattice fit accuracy could be improved by utilizing the shape of the correlation peak. Specifically, it was thought that a 2D Gaussian curve could be fit to the correlation peak, and each basis vector could be broken into orthogonal components aligned with the major and minor axes of this Gaussian curve. The components, during the fit of the basis vector to the disk locations, could be weighted by the lengths of their respective Gaussian axes. For example, in Fig. 3.1c, the Gaussian fit (indicated by the red oval) is very large in the direction of the fit with high uncertainty, where the disk does not have a clear edge. In this weighting scheme, the long axis direction would be weighted with less importance than the short axis direction. However, in practice the accuracy gains were minimal, as the system was often so overconstrained that these enhancements yielded very little return.

### A.4 Calculating strain from a reference lattice

Once there is a lattice containing the center spot, and two vectors in pixel locations, strain can be calculated. This is simple to do under the infinitesimal strain assumptions. To calculate this, again linear algebra is used to solve the equation $xA = B$, where $A$ is the reference lattice, $B$ is the lattice at each diffraction pattern, and the relationship between $x$ and strain can be seen by Eq. 3.2.

Put together, these steps described in sections A.3 and A.4 can be seen in Code Block A.3.

```matlab
function [sNBED] = fit_lat(sNBED, lat)
% inputs
% sNBED - struct from that contains xyI
% lat - user defined lattice, used to compute reference lattice
% outputs
% sNBED - struct containing both best fit lattices and strains

CBEDmean = sNBED.CBEDmean; % mean image from stack
diskShiftMax = 4-1+5; % maximum shift allowed for a peak
subLatticeFrac = [1 1]*2; % used if peaks are at fractional lattice
    % indicies, [0.5 1], for example
or = lat(1,:); % first row is center of pattern
u = lat(2,:); % second row is u vector
v = lat(3,:); % third row is v vector
sNBED.or = or;

sNBED.u = u;
```
sNBED.v = v;

% Get mean peak position values
Np = size(sNBED.peaksRefine);  % peaksRefine is the list of fitted peaks

% flat_region = ...; % this should be a mask of zeros and ones of where 0
% strain should be computed from

% set a reference region for 0 strain
if exist('flat_region','var')
    subReference = false(Np(3),1);
    subReference(flat_region(:)) = 1;
    sNBED.ref = flat_region;
else
    subReference = true(Np(3),1);  % otherwise set mean diffraction pattern
    as 0 strain
end

xyI = zeros(Np(1),3);
r2 = diskShiftMax^2;
% delete peaks that have shifted too much - usually due to fitting error
% sNBED.peaks is position of peaks on mean image
for a0 = 1: Np(1)
    sub = (sNBED.peaksRefine(a0,1,:) - sNBED.peaks(a0,1)).^2 
    + (sNBED.peaksRefine(a0,2,:) - sNBED.peaks(a0,2)).^2 
    < r2;
    xyI(a0,1:3) = max(mean(sNBED.peaksRefine(a0,1:3,sub(:)),3),0);
end

% Solve for the lattice vectors of each peak
for a0 = 1:5
    a = round(subLatticeFrac(1)*((xyI(:,2)-or(2))*v(1) ... 
    - (xyI(:,1)-or(1))*v(2))/(v(1)*u(2)-v(2)*u(1))) ... 
    /subLatticeFrac(1);
    b = round(subLatticeFrac(2)*((xyI(:,2)-or(2))*u(1) ... 
    - (xyI(:,1)-or(1))*u(2))/(v(2)*u(1)-v(1)*u(2))) ... 
    /subLatticeFrac(2);
    weights = xyI(:,3);  % the weights are the correlation intensities
    basis = [ones(Np(1),1) a b];  % this is now a matrix of lattice vectors
    in 1./subLatticeFrac increments
    lat = lscov(basis(:,:),xyI(:,1:2),weights(:,:));  % lat is still in
    pixels
    or = lat(1,:);
    u = lat(2,:);
    v = lat(3,:);
end

% Best fit positions, for plotting and reference
xyFit = basis*lat;

% solve reference lattice in the region chosen to be 0 strain
xObs = sNBED.peaksRefine(:,1,subReference);
xObs = median(xObs,3);
yObs = sNBED.peaksRefine(:,2,subReference);

yObs = median(yObs,3);
IObs = sNBED.peaksRefine(:,3,subReference);
IObs = median(IObs,3);

r2 = diskShiftMax^2;
sub = (xObs - sNBED.peaks(:,1)).^2 ... + (yObs - sNBED.peaks(:,2)).^2 ... < r2;

% lat = basis \ [xobs yObs IObs]
lat = lscov(basis(sub,:), [xobs(sub) yobs(sub)], IObs(sub));

sNBED.latRef = lat;

% Calculate lattice and strain for all images
sNBED.latRefine = zeros(3,2,Np(3));
sNBED.strains = zeros(Np(3),4 + 2);

% strains = [exx eyy exy theta dark_signal bright_signal]
warning off
for a0 = 1:Np(3) % loop through all the images
xp = sNBED.peaksRefine(:,1,a0);
yp = sNBED.peaksRefine(:,2,a0);
Ip = sNBED.peaksRefine(:,3,a0);

sub = (xp - sNBED.peaks(:,1)).^2 ... + (yp - sNBED.peaks(:,2)).^2 < r2; % exclude peaks that have shifted too far

% virtual bright and dark field
Ibf = Ip(1); % value of center spot (first in list)
Idf = sum(Ip([false; sub(2:end)])); % value of rest of spots

latFit = lscov(basis(sub,:),[xp(sub) yp(sub)],Ip(sub)); % fit lattice to one diffraction pattern
sNBED.latRefine(:,1,a0) = latFit;

m = (lat(2:3,:) \\ latFit(2:3,:)))'; % compute transformation matrix

sNBED.strains(a0,1:4) = ... [1-m(1,1) 1-m(2,2) ... -(m(2,1)+m(1,2))/2 ... (m(2,1)-m(1,2))/2]; % compute strains per pattern

% save virtual df and bf pixel values
sNBED.strains(a0,5:6) = [Idf Ibf];
end

warning on

end

Code Block A.3. The code for the final step of strain mapping, from fitting a lattice to each set of diffraction spot locations, to computing strain from a reference lattice.
A.4.1 Rotating strains

There is one part that is easy to overlook, but of crucial importance. Up until this point, all lattice vectors have been described by their position in the image in pixels. This means that when the strain is computed, as it is describing a transformation in the \(x, y\) pixel coordinate system, the strains are in the same coordinate system. There are a couple of methods to align the infinitesimal strain tensor to a different direction, usually one that is more crystallographically relevant. The first is to rotate the lattice fits to the correct direction using the standard rotation matrix before computing strain. This is straightforward. The second is to convert the strains after the fact. Since the strain tensor given by

\[
\mathbf{M} = \begin{bmatrix}
\epsilon_{xx} & \epsilon_{xy} \\
\epsilon_{xy} & \epsilon_{yy}
\end{bmatrix}
\] (A.1)

this can easily be converted by applying a tensor rotation, which is just \(\mathbf{M}^{\text{rot}} = \mathbf{RMR}'\), where \(\mathbf{R}\) is the standard rotation matrix, and \(\prime\) indicates the transpose. The resulting strains are then given by \(\epsilon_{uu} = \mathbf{M}_{11}^{\text{rot}}\), \(\epsilon_{vv} = \mathbf{M}_{22}^{\text{rot}}\), and \(\epsilon_{uv} = (\mathbf{M}_{12}^{\text{rot}} + \mathbf{M}_{21}^{\text{rot}})/2\).
Appendix B:

Methods useful for amorphous diffraction patterns

B.1 Counting electrons in sparse data

When using the Gatan K2-IS camera, which is a direct electron detector, in certain low intensity regimes the location of individual electrons on the sensor can be determined. This is very helpful as it removes all the effects of electrical noise and bloom from the diffraction patterns. This is often the case in weakly scattering amorphous materials. In practice, this is as simple as thresholding a background-subtracted image above a certain value, and then removing nearest neighbors after thresholding.

The appropriate threshold for clustering is usually determined by fitting a Gaussian function to a histogram of intensity values for a subset of the images in the stack. It is typical that the threshold is then $9\sigma$ to $11\sigma$ from the mean in the positive direction.

Once the images have had their background subtracted and have been thresholded, there could be a situation in which two pixels next to each other have both been above the threshold. This can occur if an electron scatters sideways into a neighboring pixel. To remove this effect, around every thresholded electron, the nearest neighbors are checked and if they are lower in intensity, are deleted. This process is shown in Fig. B.1, and the implementation is shown in Code Block B.1.

![Figure B.1](image)

**Figure B.1.** Clustering process example, going from Poissonian noise to a thresholded image, to an image of only electron locations. Notice how each bright pixel in $I_{electrons2}$ is not touching any other bright pixel. These are the electron locations.
function [] = clusterExample()
% this function will make an image with random Poissonian noise, threshold
% it above a certain value, and then perform nearest neighbor removal.

threshValue = 12;
im = poissrnd(threshValue.*rand(100)); % create random array
N = size(im);

imThresh = im >= threshValue; % perform thresholding - just as an example

% perform NN thresholding in two different ways
% first way - utilize circshift - this returns an image. To make a list of
% electron positions use the find command
imElectrons1 = ...% second way - utilize 1D indexing
[y, x] = meshgrid(1:N(2),1:N(1));
x = x(:);
y = y(:);
indsNN(:,1) = sub2ind(N, mod(x-2,N(1))+1, mod(y-2,N(2))+1);
indsNN(:,2) = sub2ind(N, mod(x-1,N(1))+1, mod(y-2,N(2))+1);
indsNN(:,3) = sub2ind(N, mod(x-0,N(1))+1, mod(y-2,N(2))+1);
indsNN(:,4) = sub2ind(N, mod(x-2,N(1))+1, mod(y-1,N(2))+1);
indsNN(:,5) = sub2ind(N, mod(x-0,N(1))+1, mod(y-1,N(2))+1);
indsNN(:,6) = sub2ind(N, mod(x-2,N(1))+1, mod(y-0,N(2))+1);
indsNN(:,7) = sub2ind(N, mod(x-1,N(1))+1, mod(y-0,N(2))+1);
indsNN(:,8) = sub2ind(N, mod(x-0,N(1))+1, mod(y-0,N(2))+1);
% this builds an array of NN indices
indsCand = find(im > threshValue);
indsElectrons = indsCand(im(indsCand) > max(im(indsNN(indsCand,:)),[],2));
% this is now a list of electrons - more compact representation of the
data
% build back the image to compare to imElectrons1
imElectrons2 = zeros(N);
imElectrons2(indsElectrons) = 1;
isequal(imElectrons1, imElectrons2) % this will return true

% plotting - these commands left here for clarity
f = figure(1);
f.Position = [112 1234 1405 420];
clf
ax1 = axes('Position', [0 0 .3 1]);
imagesc(im);
axis off image
caxis([0 151])
ax2 = axes('Position', [0.33 0 .3 1]);
imagesc(imThresh);
axis off image;
ax3 = axes('Position', [0.66 0 .3 1]);
imagesc(imElectrons2);
axis off image;
end

Code Block B.1. Code used to create Fig. B.1. Note that the clustering is done two different ways that yield the same result (line 48). The first way, resulting in Ielectrons1 is conceptually simpler.

B.2 Polar transforms of diffraction patterns

To proceed further with amorphous materials characterization, specifically measuring the amount of local order, we must correct for two-fold astigmatism in the diffraction patterns, caused by both incorrect alignment of the microscope, as well as naturally due to strain, and then convert the images to polar \((r, \theta)\) representations. This will then allow for the calculation of local order amount.

The process to remove astigmatism is described in Chap. 5, Eqns. 5.1-5.3 for electron positions stored as \((x, y)\) positions from the center of the pattern. Then the conversion to polar coordinates is straightforward, with \((r, \theta) = \left(\sqrt{x^2 + y^2}, \text{atan2}(y/x)\right)\), and then the transformed coordinates can be appropriately binned into images using \texttt{accumarray}. The stigmation correction can be seen in Fig. B.2.

Figure B.2. a) Stigmated frame made using electron positions. b) Ring fit (brighter) and data (dimmer) used to fix astigmatism. c) Corrected image.

The process is more involved for whole images, as aliasing effects appear when the naive polar transformation is performed on an entire image. To avoid these effects, bilinear re-
sampling is used during the transformation. The process of polar transformation with and without stigmation correction can be seen in Fig. B.3, where it is clear that without stigmation correction, severe artifacts would appear when looking for any type of rotational symmetry in the diffraction pattern. The code for this is shown in Code Block B.2.

![Figure B.3. a) Stigmated ring. b) Polar transform of stigmated ring. c) Downsampled polar transform of stigmated ring. Both b) and c) have clear two-fold symmetry strictly due to stigmation. d) Corrected polar transform of stigmated ring, e) downsampled.](image)

```matlab
function [ ] = imEllipseTransform2()
% April 2018 - Colin Ophus & Tom Pekin
% This function works for polar coordinate transforming whole images. Uses
% bilinear interpolation to remove aliasing.
% Testing function for elliptic polar coords resampling

% Polar coordinates variables
rMax = 200; % max radius
dr = 1; % radial step size - probably best at 1
dt = .5*pi/180; % theta step size - probably best such that ~360 fit at
% the radius of interest. so 360/(2*pi*rFit)

drBin = 30;
dtBin = 5 * pi/180;

% Test image variables:
imageSize = [1 1]*400;
xyCenter = imageSize/2 + 1;
a = 1; % do not change - makeTest does not include it!
b = 0.8;
c = .9;
r0 = 100;
sigma = 20;

% Generate test image with elliptic distortion
[Itest] = makeTest(imageSize,xyCenter, b, c, r0, sigma);

% Polar coordinates
```
\[ r = 0 : dr : (r_{\text{Max}} - dr); \]
\[ t = 0 : dt : (2\pi - dt); \]
\[ [ta, ra] = \text{meshgrid}(t, r); \]
\[ raBins = \text{floor}(ra ./ drBin) + 1; \]
\[ taBins = \text{floor}(ta ./ dtBin) + 1; \]

\% Resample into polar coordinates WITHOUT correcting elliptic distortion
\[ xa1 = ra .* \cos(ta) + xyCenter(1); \]
\[ ya1 = ra .* \sin(ta) + xyCenter(2); \]
\[ \text{imageResample} = \text{bilinearResample}(Itest, xa1, ya1); \]
\[ \text{imageResampleDown} = \text{accumarray}([raBins(:) taBins(:)], \text{imageResample}(::), [], @\text{mean}); \]

\% resample into polar coordinates WITH correcting elliptic distortion
\[ \text{ellipseMatrix} = [a b/2; b/2 c]; \]
\[ [v, d] = \text{eigs}(\text{inv}(\text{ellipseMatrix})); \% notice the inverse \text{ellipseMatrix} is used here \]
\[ x0 = ra .* \cos(ta); \]
\[ y0 = ra .* \sin(ta); \]
\[ \text{newxy} = [x0(:) y0(:)] * v; \]
\[ \text{newxy} = \text{newxy} * \sqrt{d}; \]
\[ \text{newxy} = \text{newxy} * v'; \]
\[ xa = \text{reshape}(\text{newxy}(:,1), \text{size}(x0)) + xyCenter(1); \]
\[ ya = \text{reshape}(\text{newxy}(:,2), \text{size}(y0)) + xyCenter(2); \]
\[ \text{imageResampleCorrect} = \text{bilinearResample}(Itest, xa, ya); \]
\[ \text{imageResampleCorrectDown} = \ldots \text{accumarray}([raBins(:) taBins(:)], \text{imageResampleCorrect}(::), [], @\text{mean}); \]

\textit{figure}(11) \hspace{1cm} \text{clf} \hspace{1cm} \text{imagesc}(Itest); \hspace{1cm} \text{hold on} \hspace{1cm} \text{axis} \hspace{0.1cm} \text{equal} \hspace{0.1cm} \text{off} \hspace{1cm} \text{set}(\text{gca},'position',[0 0 1 1]) \hspace{1cm} \text{figure}(12) \hspace{1cm} \text{clf} \hspace{1cm} \text{imagesc}([\text{imageResample} \ldots \text{zeros}(%0 1 1)]) \hspace{0.5cm} \text{axis} \hspace{0.1cm} \text{off} \hspace{1cm} \text{set}(\text{gca},'position',[0 0 1 1]) \hspace{1cm} \text{figure}(13) \hspace{1cm} \text{clf} \hspace{1cm} \text{imagesc}([\text{imageResampleCorrectDown} \ldots \text{zeros}(%0 1 1)]) \hspace{0.5cm} \text{axis} \hspace{0.1cm} \text{off} \hspace{1cm} \text{set}(\text{gca},'position',[0 0 1 1])
axis off
set(gca,'position',[0 0 1 1])
end

function [imageResample] = bilinearResample(im,xa,ya)
% Coordinates
N = size(im);
xa = min(max(xa,1),N(1)-1);
ya = min(max(ya,1),N(2)-1);
xf = floor(xa);
yf = floor(ya);
dx = xa - xf;
dy = ya - yf;

% bilinear resampling
imageResample = reshape( ... 
    im(sub2ind(N,xf(:),yf(:))).* (1-dx(:)).*(1-dy(:)) ... 
+ im(sub2ind(N,xf(:)+1,yf(:))) .* ( dx(:)) .* (1-dy(:)) ... 
+ im(sub2ind(N,xf(:), yf(:)+1)) .* (1-dx(:)) .* ( dy(:)) ... 
+ im(sub2ind(N,xf(:)+1,yf(:)+1)) .* ( dx(:)) .* ( dy(:)) ... ,
size(xa));
end

function [Itest] = makeTest(imageSize, xyCenter, b, c, r0, sigma)
x = 1:imageSize(1);
y = 1:imageSize(2);
x = x - xyCenter(1);
y = y - xyCenter(2);
[ya,xa] = meshgrid(y,x);
Itest = exp( -(r0 - sqrt(xa.^2 + b*xa.*ya + c*ya.^2)).^2 / (2*sigma^2));
end

Code Block B.2. Code used to create Fig. B.3. This code corrects for stigmation before downsampling, and uses bilinear resampling to remove aliasing artifacts.

B.3 Calculating local order

Once the diffraction patterns have been transformed to polar coordinates and sufficiently downsampled, the amount of order is calculated using the methods described in ref. [25]. Example code is given in Code Block B.3, and the steps are plotted in Fig. B.4.

function [] = localOrdering()
% this function will demonstrate local order measurement
% create diffraction pattern
dp = zeros(6,21);
N = size(dp);
dp(3:4,:) = [linspace(0,2*pi,N(2)); linspace(0,2*pi,N(2))];
dp(3,:) = 3.*sin(4.*dp(3,:)) + 3; % large 4 fold symmetry in 3rd row
dp(4,:) = 2.*sin(2.*dp(4,:)+pi/4) + 2; % smaller 2 fold symmetry in 4th row
dp = dp + rand(size(dp)); % adding some small noise
dpCorr = ifft(abs(fft(dp, [], 2)).^2, [], 2); % create correlogram
dpCorrFFT = abs(fft(dpCorr, [], 2)); % take FFT to get symmetry elements
dpCorrFFTNorm = dpCorrFFT./repmat(dpCorrFFT(:,1), [1 N(2)]); % normalize by 0th order intensity

% get measure of symmetries
order4 = dpCorrFFTNorm(3, 5);
order2 = dpCorrFFTNorm(4, 3); % use symmetry order + 1 for element

order4 > order2 % will usually evaluate to true - noise might cause it occasionally to evaluate to false

end

**Code Block B.3.** Code used to create Fig. B.4, and used to measure order in one polar-transformed diffraction pattern.

---

**Figure B.4.** Steps taken to measure amount of local order. Note that `dpCorrFFT` has been `fftshift`-ed so that the 0th order channel has been centered for easier inspection. The red arrows are the frequencies corresponding to two and four-fold symmetries.
Appendix C: Useful image processing tips and techniques

C.1 Standard image coordinate systems and indexing

In MATLAB and most other programming languages, images are represented as a matrix. Therefore, accessing subregions of an image should not be thought of as Image(x, y), but rather Image[row, column]. This is illustrated in Fig. C.1.

Figure C.1. Illustrations of the indexing schemes used by MATLAB in this thesis. The Image on the left shows both 1D and 2D indexing schemes. Functionally, Image(5) == Image(1,2). Stack3D and Stack4D show how 3D and 4D stacks are indexed. Note that in Stack4D, to recover the same Image as in Image and Stack3D, the appropriate command should actually be Image = squeeze(Stack4D(1,1,:,:)), but (4,2,:,:,:) was used to show “deeper” indices. In Stack4D, each of the first two indices represents an entire image, usually a diffraction pattern. squeeze is used to remove singleton dimensions.

C.2 Helpful packages

- The BM3D family of algorithms are very useful for denoising with many variations for different use cases [84, 140]. For videos, V-BM4D is very powerful — http://www.cs.tut.fi/~foi/GCF-BM3D/.
- export_fig is a very useful toolbox “for exporting figures from MATLAB to standard image and document formats nicely” [141]. The author has found it to be more flexible
and easier to use than the standard print or or other built-in dialogues — https://github.com/altmany/export_fig.

• convolve2 is a faster implementation of conv2 in MATLAB, and therefore is useful for speeding up computation in which this operation is repeatedly performed [142] — https://www.mathworks.com/matlabcentral/fileexchange/22619-fast-2-d-convolution.

C.3 Helpful scripts

Many of these scripts can be found at the author’s Bitbucket repository, found at https://bitbucket.org/tcpekin/scripts/src (the most helpful script/tool is Git — use it!). Please request access, read-only, and clone it if these scripts seem handy. Pull requests are welcomed.

C.3.1 Importing microscope image formats

Often, TEM manufacturers export the images acquired to proprietary formats that contain important metadata. While the microscope and camera manufacturers can export the images to TIFF files, often it is more convenient to use the native format for import into MATLAB or Python. Most of the MATLAB import scripts are hosted on the openNCEM Bitbucket repository (https://bitbucket.org/ercius/openncem/src/master/), which is maintained by Peter Ercius. There also exists a GitHub openNCEM repository, which is more Python focused (https://github.com/ercius/openNCEM).

Gatan file formats

To read .dm3 files, use dm3Reader.m (https://bitbucket.org/ercius/openncem/src/master/dm3Reader.m).

To read .dm4 files, use dm4Reader.m. This should be able to read 4D NBED datasets acquired on the Gatan K2-IS camera (https://bitbucket.org/ercius/openncem/src/master/dm4Reader.m).

FEI TIA file formats

TIA outputs both .emi and .ser files, but only .ser files can be successfully read by MATLAB and Python currently. The importer for this filetype can be found at https://bitbucket.org/ercius/openncem/src/master/serReader.m.

Converting to 4D .dm4 files

Gatan’s Digital Micrograph GMS 3 (the latest version) does have some convenient tools for 4D stack manipulation, such as relatively simple strain mapping, virtual darkfield and bright field image creation, as well as the standard image analysis tools contained. Digital Micrograph also has a C++-based scripting language [143]. The first script provided here in
Code Block C.1 can convert 3D stacks acquired by the FEI TitanX and the Libra at NCEM to 4D stacks to use the latest tools provided by Gatan. Additionally, the second script (Code Block C.2 can do this with a nested folder structure, which is included for archival reasons. This stack format is not longer in vogue. These scripts should be opened in Digital Micrograph and executed in the program itself. For more reading on scripting in Digital Micrograph, please refer to [144, 145].

```plaintext
1  // begin user defined numbers
2  // offset - i added this. Image number at which real DPs start
3  number offset = 0 // 0 for 3D stacks from TitanX
4  // binning factor
5  number binningfactor = 1 // binning factor to bin diffraction patterns by,
   // should be multiple of two and DPs should be sized such that both
   // dimensions are evenly divisible by binningfactor
6  if( binningfactor < 1 )
7      {   
8         Throw("binning factor is less than 1, change please")
9      }
10  // end of user defined numbers
11
12  // debugging section
13  /* string filename = "Z:\bustillo\20170206 For Tom 4DSTEM dataset from
   TitanX small\Winey 05 12 best_01 10x10 30nm steps spot 11 0p51mrad -4
   p7um defocus 33ms CL=380.dm3"
14  
15  image front = OpenImage(filename)
16  ShowImage( front )
17  */
18
19  // get basic image on already opened 3D stack
20  image front := getfrontimage()
21
22  taggroup imtags = front.imagegettaggroup()
23
24  number width, height
25  string imWidthTag = "Series:nimagesx"
26  string imHeightTag = "Series:nimagesx"
27  number dimx, dimy, dimz
28  get3Dsize(front, dimx, dimy, dimz)
29
30  imtags.taggroupgettagasnumber(imWidthTag, width)
31  imtags.taggroupgettagasnumber(imHeightTag, height)
32
33  number cameraWidth = dimx
34  number cameraHeight = dimy
35
36  // Change these numbers to the numbers you acquired and uncomment if error
37  // shows up with dimension = 0. This usually happens on the Libra. width
38  // and height // are scan dimension, cameraWidth and cameraHeight are diffraction
39  // pattern dimensions.
40  //width = 30
41  //height = 30
```
// cameraWidth = 256
// cameraHeight = 256
// Result(width + height + cameraWidth + cameraHeight)

if( binningfactor != 1 )
{
    cameraWidth = cameraWidth / binningfactor
    cameraHeight = cameraHeight / binningfactor
}

image img4D := NewImage( "4D_image", 10, width, height, cameraWidth, cameraHeight )

ShowImage( img4D )
IMDSetImageFormatDiffractionImage( img4D )

number rowIndex = 0
number columnIndex = 0
number needMoreData = 1

// Function that performs batch action
void PerformBatchAction( image img0 )
{
    if( rowIndex == height )
    {
        Result( "\t\t\tNOT NEEDED\n" )
        needMoreData = 0
        return;
    }

    img4D.sliceN( 4, 2, columnIndex, rowIndex, 0, 0, 2, cameraWidth, 1, 3, cameraHeight, 1 ) = img0
    columnIndex += 1
    if( columnIndex >= width )
    {
        columnIndex = 0
        rowIndex += 1
    }
}

void StackProcess( )
{
    number dimx, dimy, dimz
    get3Dsize( front, dimx, dimy, dimz)
    for ( number i = offset ; i < dimz && needMoreData; i += 1 )
    {
        image img0 := slice2( front, 0, 0, i, 0, dimx, 1, 1, dimy, 1 )
        if( binningfactor != 1 )
        {
            // Script taken from DM documentation
Code Block C.1. Create4DStack_3DStack.s This function will open a 3D stack and then convert it to a 4D stack. It provides options for binning as well as applying an offset.

```
// binning factor
number binningfactor = 2

// size of scan in real space
number width = 101 // add 1 for K2 datasets
number height = 100

// offset - i added this. Image number at which real DPs start
```
number offset = 6584

// size of acquired diffraction image
number cameraWidthTotal = 1920 // I think this line just was left in here by accident
number cameraWidth = 1792 // current unbinned values
number cameraHeight = 1920 // current unbinned values

// John Balk Data
// cameraWidth = 1920
// cameraHeight = 1920

if( binningfactor != 1 )
{
   cameraWidth = cameraWidth / binningfactor
   cameraHeight = cameraHeight / binningfactor
}

image img4D := NewImage( "4D_stack", 10, width, height, cameraWidth, cameraHeight )
ShowImage( img4D )
IMDSetImageFormatDiffractionImage( img4D )

number rowIndex = 0
number columnIndex = 0
number needMoreData = 1

void PerformBatchAction( image img0 )
{
   if( rowIndex == height )
   {
      Result( "\t\t\tNOT NEEDED\n" )
      needMoreData = 0
      return;
   }
   number sx1, sy1, sx2, sy2
   GetSize( img0, sx1, sy1 )
   GetSize( img4D.sliceN( 4, 2, columnIndex, rowIndex, 0, 0, 2, cameraWidth, 1, 3, cameraHeight, 1 ), sx2, sy2 )
   result(""+sx1+" "+sy1+" "+sx2+" "+sy2+" ")
   img4D.sliceN( 4, 2, columnIndex, rowIndex, 0, 0, 2, cameraWidth, 1, 3, cameraHeight, 1 ) = img0
   columnIndex += 1
   if( columnIndex >= width )
   {
      columnIndex = 0
      rowIndex += 1
   }
}

// Function to create a list of file entries with full path
void CreateFileList( TagGroup fileList, string folder )
{ // for each directory, make a new folder name and call recursively
  TagGroup directoryTG = GetFilesInDirectory( folder, 2 ); // 1 = Get files, 2 = Get folders, 3 = Get both

  for (number i = 0; i < directoryTG.TagGroupCountTags(); i++ )
  {
    TagGroup entryTG
    if ( directoryTG.TagGroupGetIndexedTagAsTagGroup( i , entryTG ) )
    {
      string directoryName
      if ( entryTG.TagGroupGetTagAsString( "Name" , directoryName ) )
      {
        CreateFileList( fileList, folder + "\" + directoryName )
      }
    }
  }

  // for each file, add to list
  TagGroup filesTG = GetFilesInDirectory( folder, 1 ); // 1 = Get files, 2 = Get folders, 3 = Get both

  for (number j = 0; j < filesTG.TagGroupCountTags(); j++ )
  {
    TagGroup entryTG2
    if ( filesTG.TagGroupGetIndexedTagAsTagGroup( j , entryTG2 ) )
    {
      string fileName
      if ( entryTG2.TagGroupGetTagAsString( "Name" , fileName ) )
      {
        fileList.TagGroupInsertTagAsString( fileList.TagGroupCountTags() , folder + "\" + fileName )
      }
    }
  }

  // Open and process all files in a given fileList
  void BatchProcessList( TagGroup fileList , string name )
  {
    number nEntries = fileList.TagGroupCountTags()
    if ( nEntries > 0 )
    {
      result( "Processing file list <" + name + "> with " + nEntries + " files.\n" )
    }
    else
    {
      result( "File list <" + name + "> does not contain any files.\n" )
    }

    for ( number i = offset ; i < nEntries && needMoreData; i += 1 )
    // CHANGE I = OFFSET??
{
  string str0, str1, str2, str3
  fileList.TagGroupGetIndexedTagAsString( i, str0 )

  result(  "\t open: " + str0 +  "\n"  )

  image img0 := OpenImage( str0 )

  //Debugging
  number sx, sy

  // Actual batch-action
  result(  "process...  \n"  )
  number j
  j = i - offset
  result(  j +  "\n"  )

  if( binningfactor != 1 )
  {
    // Script taken from DM documentation
    // Create exact copy (including tags, calibration...)
    image out
    out := ImageClone( img0 )
    // Resize image to binned size.
    // - sets pixel values to zero
    // - adjusts calibrations to keep FOV
    number sx, sy

    GetSize( img0, sx, sy )
    result("Test" +sx +sy +"\n")

    ImageResize( out, 2, sx/binningfactor, sy/binningfactor )
    // Sum over all possible sub-sections

    for ( number j = 0; j < binningfactor; j++ )
    {
      for ( number i = 0; i < binningfactor; i++ )
      {
        out += Slice2( img0, i, j, 0, 0, sx/binningfactor, binningfactor
        , 1, sy/binningfactor, binningfactor )
      }
    }
    DeleteImage(img0)
    img0 = out
    DeleteImage(out)
  }

  PerformBatchAction( img0 )

  DeleteImage(img0)
}
void BatchProcessFilesInFolder()
{
    string folder, outputFolder
    if (!GetDirectoryDialog( "Select folder" , "" , folder ) )
        return
    TagGroup fileList = NewTagList()
    CreateFileList( fileList, folder )
    BatchProcessList( fileList, "DM4 list" )
}

Code Block C.2. Create4DStack_NestedFolder.s This function will open a nested folder structure of .dm3 files and then convert it to a 4D stack. It provides options for binning as well as applying an offset. This file structure with the nested folders was common with the early K2-IS NBED acquisitions.

C.3.2 Image manipulation

When Gatan K2-IS data is acquired, it is often convenient to remove the background from each diffraction pattern. The background consists of lines due to electrical noise. Luckily, the HAADF detector in the NBED patterns provides a convenient reference. This process is shown in Fig. C.2, and the code for this operation is shown in Code Block C.3.

Figure C.2. Process of background subtraction. a) The mean diffraction, with red box showing where the reference background is measured. b) The reference background, which is subtracted from a). c) The resulting pattern. This is often done for every pattern in the stack of images.

```
im = im_orig - repmat( mean( im_orig(:,206:end),2) , [1, size(im_orig,2)] );
% im is the background subtracted image, im_orig is the original image, and the reference background is created by using repmat to expand the
```
mean in the HAADF aperture to the size of the original image. the 206 index is where column where the aperture starts.

**Code Block C.3.** Standard background subtraction code using the aperture edge as background reference.

```matlab
function [qx, qy] = makeFourierCoords(N, pixelSize)
% This function generates Fourier coordinates.
% Updated 2017 to fix odd sizes, and allow 2D output.
if mod(N(1),2) == 0
    qx = circshift(((N(1)/2):(N(1)/2-1))/N(1)*pixelSize),[0 -N(1)/2]);
else
    qx = circshift(((1-N(1))/2):((N(1)-1)/2))/(N(1)*pixelSize),[0 (1-N(1))/2]);
end
if nargout == 2
    if length(N) == 1
        [qy, qx] = meshgrid(qx);
    else
        % Add second dimension
        if mod(N(2),2) == 0
            qy = circshift(((N(2)/2):(N(2)/2-1))/N(2)*pixelSize),[0 -N(2)/2]);
        else
            qy = circshift(((1-N(2))/2):((N(2)-1)/2))/(N(2)*pixelSize),[0 (1-N(2))/2]);
        end
        [qy, qx] = meshgrid(qy, qx);
    end
end
end
```

**Code Block C.4.** makeFourierCoords is one of the most widely used functions in the image processing done here. It is useful for when you want to operate in Fourier space.

The Fourier Shift Theorem is very helpful when shifting around images a non-integer amount. The theorem simply states that circularly shifting a signal by an amount \( m \) can be performed by multiplying its Fourier transform by \( \exp(-2\pi iq_x m) \), where \( q_x \) is \( k/N \), where \( k \) is the frequency spectrum, and \( N \) is the size of the image. This can be recognized as a linear phase plane wave. The inverse DFT will recover the shifted image. This is implemented in Code Block C.5.

```matlab
function [imageOut] = imageFourierShift(imageIn, xShift, yShift)
% this function takes in an image, and uses a fourier shift to move it % xShift and yShift pixels. xShift moves down, yShift moves right.
imageSize = size(imageIn);
qx = makeFourierCoords(imageSize(1), 1)';
if imageSize(2) == imageSize(1)
    qy = qx;
else
    qy = makeFourierCoords(imageSize(2), 1);
end
imageOut = real(ifft2(fft2(imageIn).*exp(-2i*pi*qx*xShift) ...
Code Block C.5. This function utilizes makeFourierCoords and the Fourier Shift Theorem to smoothly shift an image by an arbitrary amount. This shift does not have to be done in integer pixel amounts.

Image binning is often very helpful to reduce file size, increase signal to noise, and speed up processing. A simple function to do this is shown in Code Block C.6.


C.4 Image viewing

The next function in Code Block C.7 allows you to quickly use the arrow keys view a 3D or 4D stack of images in MATLAB without unnecessary plotting.
im = squeeze(im_stack(numx, numy, :, :));
a = gca;
title(a, [numx, numy]);
end

h = imagesc(im); axis image; colorbar
if nargin == 2
caxis(range)
end

function figScroll(src, callbackdata)
if length(size(im_stack)) == 3
    if strcmp(callbackdata.Key, 'rightarrow') == 1
        num = num + 1;
        if num == size(im_stack, 3) + 1
            num = 1;
        end
        im = im_stack(:, :, num);
        re_eval(im, num)
    elseif strcmp(callbackdata.Key, 'leftarrow') == 1
        num = num - 1;
        if num == 0
            num = size(im_stack, 3);
        end
        im = im_stack(:, :, num);
        re_eval(im, num)
    elseif strcmp(callbackdata.Key, 'uparrow') == 1
        numx = numx - 1;
        if numx == 0
            numx = size(im_stack, 1);
        end
        im = squeeze(im_stack(numx, numy, :, :));
        re_eval(im, [numx, numy])
    elseif strcmp(callbackdata.Key, 'downarrow') == 1
        numy = numy + 1;
        if numy == size(im_stack, 2) + 1
            numy = 1;
        end
        im = squeeze(im_stack(numx, numy, :, :));
        re_eval(im, [numx, numy])
    elseif strcmp(callbackdata.Key, 'leftarrow') == 1
        num = num - 1;
        if num == 0
            num = size(im_stack, 3);
        end
        im = im_stack(:, :, num);
        re_eval(im, num)
    elseif strcmp(callbackdata.Key, 'uparrow') == 1
        numx = numx - 1;
        if numx == 0
            numx = size(im_stack, 1);
        end
        im = squeeze(im_stack(numx, numy, :, :));
        re_eval(im, [numx, numy])
    endif
endif
```matlab
elseif strcmp(callbackdata.Key, 'downarrow') == 1
    numx = numx + 1;
    if numx == size(im_stack, 1) + 1
        numx = 1;
    end
    im = squeeze(im_stack(numx, numy, :, :));
    re_eval(im, [numx, numy])
end
end

function re_eval(im, num)
    h. CData = (im); % this will not work in versions prior to r2014b.
    if you use that version, comment out this line, and uncomment the next line.
    % set(h, 'CData', im);
    title(a, num)
    drawnow
end
end
```

**Code Block C.7.** A function to view 3D and 4D stacks easily, in the figure window with the arrow keys.

The function virtualDF.m takes in a 4D image stack and outputs a struct that contains a virtual darkfield image.

```matlab
function [out] = virtualDF(cube4D, center, radius)
    % virtualDF takes in a 4D_cube and creates a (for now) circular aperture
    % around a center point with a user defined radius
    % cube4D is a dataset in dimensions [x y qx qy], where real-space
    % coordinates occur first
    % center is a [x y] location in image coordinates, and can be a row
    % delimited list [x1 y1; x2 y2; ...]
    % radius is a number

    N = size(cube4D);
    Nc = size(center);

    [xgrid, ygrid] = meshgrid(1:N(4), 1:N(3));
    mask = zeros(N(3:4));

    for k = 1:Nc(1)
        r_mask = sqrt((xgrid-center(k,2)).^2 + (ygrid-center(k,1)).^2);
        r_mask = r_mask <= radius;
        mask = mask | r_mask;
    end

    img = zeros(N(1:2));
    CBEDmean = squeeze(mean(mean(cube4D, 1), 2));
```
% figure(9)
% imagesc([CBEDmean./max(CBEDmean(:))) r_mask]); axis image; colormap gray
figure(10)
ap_image = log(CBEDmean .* (mask+0.1));
imagesc(ap_image); axis image; colormap gray

proceed_flag = input('Proceed with calculation? [y/n] \n', 's');
while strcmpi(proceed_flag, 'n') || ~strcmpi(proceed_flag, 'y')
    if strcmpi(proceed_flag, 'n')
        return
    else
        proceed_flag = input('Incorrect input, proceed with calculation? [y/n] \n', 's');
    end
end
disp('Proceeding with calculation');
tic
for i = 1:N(1)
    for j = 1:N(2)
        img(i,j) = sum(cube4D(i,j,mask));
    end
end
out.img = img;
out.center = center;
out.radius = radius;
out.ap_image = mask;

% figname = '[';
% for i = 1:Nc(1)
%    figname = [figname ' ' num2str(center(i,:)) ' '];
% end
% figname = [figname ']';

f = figure;%(’Name’, figname);
imagesc(img); axis image off; colormap(gray)
h = gca;
h.Position = [.25 0 .75 1];
j = axes(’units’, ’normalized’, ’position’, [-.02 .7 .3 .3]);
imagesc(ap_image); axis image off; colormap(gray)
hold on
for i = 1:Nc(1)
    scatter(center(i,2), center(i,1), ’r+’, ’LineWidth’, 2)
end
hold off
toc
function [imRGB] = plotMaskedImage(im, mask, caxis_scale, flag_plot)
    % this function will take in an image, and a mask, and output a plot in
    % which the masked off areas are black.

    if size(im) ~= size(mask)
        error('Image size ([%d %d]) does not match mask size ([%d %d]).', ...
            size(im, 1), size(im, 2), size(mask, 1), size(mask, 2));
    end

    if nargin <= 3
        flag_plot = 1;
    end

    if nargin == 2 | isempty(caxis_scale)
        caxis_scale = [min(im(mask(:))) max(im(mask(:)))];
    end

    cmap = parula(256);
    % cmap = hot(256);
    sigma = 0; % for smoothing with gaussian kernel - maybe this shouldn’t
    % exist in this function and smoothed image should just be passed in

    % smooth mask
    mask = sin(mask*pi/2).^2;

    if sigma > 0
        k = fspecial('gaussian', 2*ceil(2*sigma)+1, sigma);
        % mask = conv2(mask, k, 'same') ./ conv2(ones(size(mask)), k, 'same');
        kNorm = 1./ conv2(mask, k, 'same');
        im = conv2(mask.*im, k, 'same') .* kNorm;
    end

    im = (im - caxis_scale(1)) ./ (caxis_scale(2) - caxis_scale(1));
    im = min(max(im, 0), 1);  
    imRGB = ind2rgb(round(255.*im)+1, cmap);
    temp = imRGB;

    if flag_plot == 1
        % figure(999);
        % clf;
        % imagesc(mask);

    % intermediate plots
    imRGB = rgb2hsv(imRGB);
    imRGB(:,:,3) = imRGB(:,:,3).*mask;
    imRGB = hsv2rgb(imRGB);

end

Code Block C.8. A function to create virtual darkfield images.

The function plotMaskedImage.m in Code Block C.9 allows you to plot matrices with areas specified by mask as black. This is not easily possible using built in functions such as imagesc. This is useful for when there is vacuum in the image, which should be represented as black. This function uses the HSV (hue, saturation, value) representation of the image, and sets the V to 0 or black where the mask specifies.
Code Block C.9. A function to create a masked image.

This last function plotClusterImage.m takes in a list of \((x, y)\) coordinates that are the locations of electrons in a clustered diffraction pattern, and outputs an image for inspection, or further processing.

```matlab
function [Ik] = plotClusterImage(sCluster, Ind)
    N = circshift(sCluster.stackSize,0);
    x = sCluster.electrons{Ind}(:,:1);
    y = sCluster.electrons{Ind}(:,:2);
    I = zeros(N(1:2));
    I(sub2ind(N(1:2),x,y)) = 1;
    I = imresize(I,1/imageScale,'bilinear');
    k = fspecial('gaussian',2*ceil(4*sigma)+1,sigma);
    Ik = conv2(I,k,'same');
end
```

Code Block C.10. This function takes a list of electrons in a struct and outputs an image.