Four Dimensional Scanning Transmission Electron Microscopy: Applications and Data Processing with Machine Learning

By

Chenyu Zhang

At the

University of Wisconsin-Madison

2020

Date of final oral examination: Sepetmber 23rd, 2020

This dissertation is approved by the following members of the Final Oral Committee:

Paul M. Voyles, Professor, Materials Science and Engineering

Dane Morgan, Professor, Materials Science and Engineering

Chang-Beom Eom, Professor, Materials Science and Engineering

Jason K. Kawasaki, Assistant Professor, Materials Science and Engineering

Anru Zhang, Assistant Professor, Statistics

Contents

List of Figures

PSNR from noisy data calculated under different noise levels. (h) Spatial distribution of denoising PSNR for 1000 FPS data.

Figure 7.3 Denoising simulated 4D STEM data from a Si dislocation core: The virtual ADF image, a single CBED pattern from the periodic region of the Si structure at the position of the green dot in (g), and a single CBED pattern from the dislocation core region at the position marked by the red dot in (g) extracted from: (a)-(c) simulated 1000 FPS noisy data, (d)-(f) tensor SVD denoised data, and (g)-(i) ground truth. (j) Output PSNR from SVD denoised data vs. input PSNR from noisy data at different noise levels. (k) Spatial distribution of denoising PSNR for 1000 FPS data.

Figure 7.4 Denoising experimental 4D STEM data: (a) – (h) Virtual ADF images and single CBED patterns pairs. The right half of each CBED is shown as the square root of intensity to emphasize weak features. The color scales are for the linear left half of each image. (a) – (d) Data from STO [100]: (a), (b) as acquired, (c), (d) after tensor SVD denoising. (e) – (f) Data from a LiZnSb [$11\bar{2}0$] / GaSb [110] interface: (e), (f) as acquired, (g), (h) after tensor SVD denoising. (i) – (n) Scree plots used to determine the tensor rank for denoising. Components in blue are retained and components in orange are discarded in the denoised reconstruction of the data. The 0 component which represents the mean of the data is omitted from the graphs. (i) – (k) STO. scree plots. (l) – (n) LiZnSb / GaSb scree plots.

Figure 7.5 EDS elemental maps from (a) noisy data, (b) tensor SVD denoised data, (c) NLPCA denoised data, and (d) matrix PCA denoised data displayed in the order of Ca, Ti, Nd, from left to right. (a), (c) and (d) are replotted from the data in ref 166.

the lattice structure of LaPtSb and (d) a zoomed-in patch from the HAADF image with atomic sites marked on the image.

Figure 10.1 (A) An aligned WS2 spiral. (B) A deformed SiO2 particle. (C) A super-twisted WS2 spiral with a deformed SiO2 particle at its center. The first row (A1, B1, C1) is the top-down SEM images of the objects before FIB. The second row (A2, B2, C2) is the corresponding cross-section images with boundaries highlighted by dashed lines. The areas between the two yellow dashed lines are WS2. Fig. B2 and C2 clearly show that the deformed SiO2 nanoparticles formed cone-shaped protrusions on Si/SiO2 substrates after the growth process.

Figure 10.2 (a) low magnification HAADF-STEM image from the whole structure, (b) – (e) PACBED pattern generated from areas with 1-, 2-, 3-, and 4- layers of WS2 structure respectively, (e) relationship between layer index and geometric twist angle and lattice twist angle measured using the shape and PACBED pattern.

Figure 10.3 (a) low magnification HAADF-STEM image from the whole structure, (b) relationship between layer index and geometric twist angle and lattice twist angle measured using the shape and PACBED pattern, (c) – (i) PACBED pattern generated from areas with 1-8 layers of WS2 structure except three layers (j) relationship between layer index and the position of each set of the diffraction pattern.

Figure 10.4 (a)-(c) Morie patterns acquired from HAADF-STEM on 1-, 2-, and 3- layers of twisted WS2, (d)-(f) simulated Morie pattern by overlapping 1, 2, and 3 layers of 1T WS2 models with the twist angle measured using PACBED patterns.

List of Tables

Acknowledgment

Writing the acknowledgment is always the nicest part, a lot of people have contributed efforts to my works presented in this thesis, and I relish this opportunity to thank them.

1. I would like to first thank my advisor, mentor, and also committee chair, prof. Paul Voyles, for his infinite patience and support through this long journey.
2. All the collaborators I have worked with and learn from, thank you for help me solving all kinds of problem, and it has always been good to know that at least another people cares about what I am working on: ReBecca Willett and Albert Oh from Department of Electrical and Computer Engineering. Yizhan Wang, Kangning Zhao, Dane Morgan, Zefeng Yu, Longfei Liu, Chengrong Cao, Dongxue Du, Patrick Strobeen, Estiaque Shourov, Jason Kawasaki, Daesu Lee, and Chang-Beom Eom from Department of Materials Science and Engineering. Rungang Han and Anru Zhang from Department of Statistics. Yuzhou Zhao and Song Jin from Department of Chemistry. Benjamin Berkels from RWTH Aachen University. Bendikt Wirth from University of Munster. Luis Rangel DaCosta from University of Michigan.
3. My PhD committee, professors Dane Morgan, Chang-Beom Eom, Jason Kawasaki, and Anru Zhang, for providing valuable discussions.
4. All the past and present Voyles group members, for all your discussion and help over the years: Andrew Yankovich, Yizhang Zhai, Ashutosh Rath, Min Yu, Xing Wang, Congli Sun, Pei Zhang, Jie Feng, Dan Zhou, Jason Maldonis, Sachin Muley, Xuying Liu, Debaditya Chatterjee, Carter Francis, Shuoyuan Huang, John Sunderland, and Jingrui Wei.

5. The UW-Madison nanoscale imaging and analysis center staff, for supporting, maintaining the instrument, help with trouble shooting, and training. Especially Alexander Kvit, without whom there won't be all the wonderful microscopy images in this thesis.
6. My friends, for all the supports and listening to my complains about all kinds of staff that happened during my PhD.
7. My family and especially my partner, Lu Guo, for all the understanding, support, and discussion.
8. Last, the United States Department of Energy for the funding for the research. (DE-FG02-08ER46547)

Abstract

This thesis reports research developing different tools to help fully utilize the power of modern scanning transmission electron microscopy (STEM) and more effectively extract information from the data collected in STEM, including high angle annular dark-field (HAADF) STEM image series, hyperspectral X-ray energy dispersive spectroscopy (EDS) data, and 4D STEM data. Machine learning and big data methods have been applied to enhance the understanding of the data collected in STEM, and the techniques have been applied to various materials systems and problems.

First, atomic-resolution data collected from the conventional HAADF detector and EDS detector in STEM is discussed. The different parameters of non-rigid registration have been studied and optimized for atomic resolution data for high precision atomic site position determination. Joint denoising and distortion correction, a new tool to directly fit the atom positions from noisy image series, has been developed to find atom positions with enhanced precision under limited electron dose. For hyperspectral data collected from the EDS detector, non-local principal component analysis has been applied to denoise the data to reveal the underlying atomic structure.

Second, machine learning tools to help to understand the data collected in the newly developed 4D STEM imaging mode are discussed. Convolutional neural networks have been used to study the convergent beam electron diffraction (CBED) pattern from 4D STEM data and predict materials properties from the CBED patterns. Tensor singular value decomposition has been used to remove the random noise from the massive 4D STEM data and find the noiseless ground truth of the noisy data.

Third, 4D STEM techniques are applied to study different aspects of the three different materials systems, including: 1. Strain states around platelet-shaped Nb-rich precipitate in ZrNb alloy after irradiation. 2. Lattice twist pattern in multilayer twisted WS_2 mesostructure grown on a non-Euclidean substrate surface. 3. Structural polarization and the internal electric field generated by the polar structure in polar metal LaPtSb and LaAuGe with hexagonal Heusler structures.

1 Introduction

Modern aberration-corrected scanning transmission electron microscopy (STEM) enables the researcher to study the local structure of materials at an atomic resolution[1] by generating a well-focused electron beam with sub-Ångstrom size[2]. The integration of X-ray energy dispersive spectroscopy (EDS) and electron energy loss spectroscopy (EELS) with STEM further enabled the study of chemical and bonding information at the atomic scale[3]. Furthermore, the recent development of ultrafast direct electron detectors enabled the acquisition of the whole convergent beam electron diffraction (CBED) pattern at each probe position and a new imaging mode called 4D STEM[4]. A wealth of information can be collected in STEM with high spatial resolution, and the methods to interpret the data and extract useful information to answer materials science questions from the data becomes critical.

Chapters 3 to 5 discuss the use of three different techniques to help better extract information from atomic-resolution data collected in STEM, including both high-angle annular dark-field (HAADF)-STEM image series and the hyperspectral data collected from the EDS detector. The experimental condition and registration parameters for non-rigid registration to achieve sub-pm precision image have been discussed in chapter 3. In chapter 4, non-local principle component analysis (NLPCA), a novel non-local denoising method using the redundancy within the data itself, has been applied to denoise the EDS spectroscopy data dominated by Poisson noise to reveal the underlying atomic structure. Chapter 5 discusses a new method to directly identify the position of atomic columns from noisy image series. By doing the all-at-once fitting on the noisy image series, the new method achieved enhanced precision level on image series collected with reduced electron dose.

With the development of 4D STEM, the data size collected in STEM increased by hundreds of times as a whole 2D CBED pattern is collected at each probe position instead of a single number or a single spectrum. Each STEM scan typically generates 4D STEM data with multiple-GB size, and the total data size collected on a certain specimen could easily reach tens to hundreds of GB. To more effectively handle these massive data collected in 4D STEM, machine learning tools have been applied to pre-process and analyze the collected data. In chapter 6, the convolutional neural network (CNN), a supervised machine learning tool, has been trained using multislice simulated 4D STEM data to predict material properties on real experimental data, and achieved nm precision in local thickness estimation at unit cell resolution. An unsupervised machine learning tool, tensor singular value decomposition (SVD), is introduced in chapter 7 and applied to denoise atomic-resolution 4D STEM data. The method shows close to the best denoising performance compare to other state-of-the-art denoising methods, while performing at lease 100-folds faster, make it suitable to denoise the massive 4D STEM data.

With a whole CBED pattern collected from each probe position, 4D STEM could access materials properties that cannot be resolved in the conventional STEM. Chapter 8 discusses the strain mapping of Nb-rich precipitate in irradiated ZrNb alloy, using the variation of distances between the Bragg diffraction spots in CBED patterns. With a similar setup, chapter 10 used the twist angles in the Bragg diffraction spots to study the twist patterns in multilayer twisted WS_2 structures. In chapter 9, the movement of the electron beam under the Columb force from the internal electric field, which is captured in the CBED patterns, is used to map the internal electric field in hexagonal Heusler alloys LaPtSb and LaAuGe, which are possible polar metal candidates with high electron conductivities.

2 Characterization and Simulation Techniques

2.1 Introduction

STEM serves as my main toolbox to study different materials structure. This chapter will introduce different aspects of STEM and how they provide information about different aspects of the structure of the material. Overall, STEM rasters high energy, well-focused electron beam across the surface of an electron transparent piece of material. The electrons interact strongly with the material, and different aspects of the material's structure will be reflected in the diffracted electron beam. Chapter 2.2 introduces the conventional STEM techniques include STEM spectrum techniques. Section 2.3 and 2.4 will focus on the CBED produced in STEM and 4D STEM imaging mode and its implementation. Section 2.5 introduces the multislice simulation, which is a powerful tool to simulate the electron-sample interaction and helps researches predict and better understand the result acquired in STEM. Section 2.6 briefly summarizes all the collaborative projects that I worked on using STEM.

2.2 Conventional STEM

In conventional STEM, a well-focused electron probe, possibly with sub-Å size, is formed and rasters on the surface of an electron transparent sample. Figure 2.1, which was adapted from ref[5], shows the schematic of a STEM. The optical system of a STEM is often viewed as a reversed TEM. Typically, a STEM consists of the electron source, pre-specimen lense, sample stage, projection lenses, and annular detector. As electron interacts strongly with atoms inside the specimen, the transmitted electron composes rich information about sample structure and properties. A modern STEM often also includes an aberration corrector, EELS detector, EDS

detector, and ronchigram camera. A few of the important components will be introduced in detail in the following paragraphs.

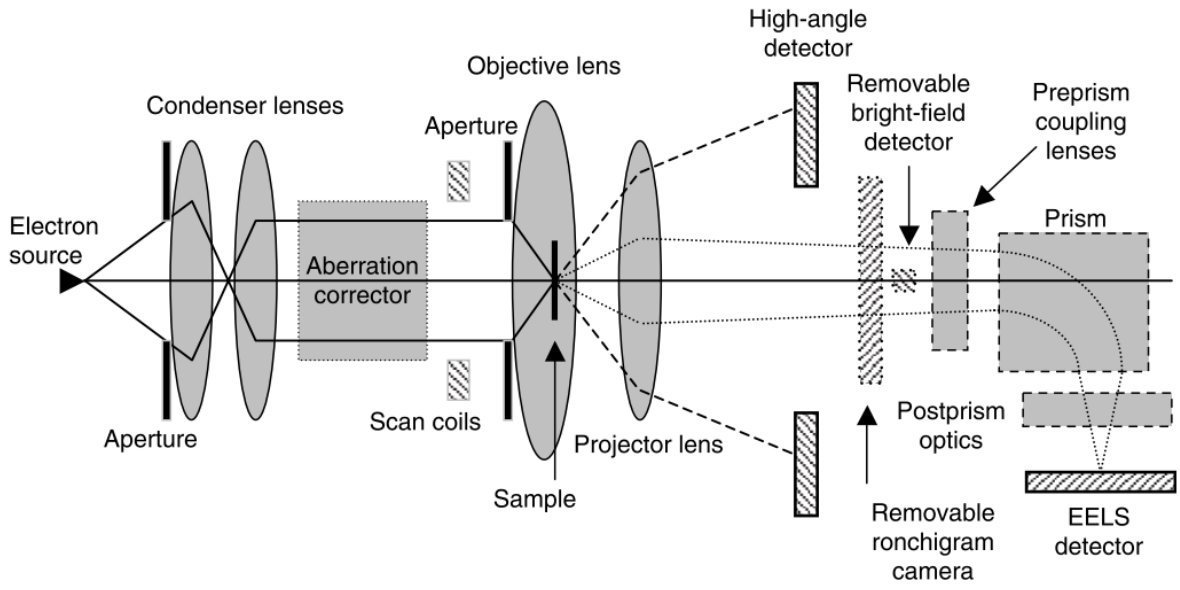

Fig 2.1 Schematic of a STEM.

The electron source produces a high energy electron with small energy dispersion. Field emission gun (FEG) is the most popular electron source for modern STEM. Recent instrument development enabled high brightness cold FEG, which generates multiple times brighter electron sources compare to traditional warm FEG. Common STEM uses an electron source that provides 200-300 kV electron, with 60-80 kV low energy electron as an alternate option. High voltage STEM that produces higher than 1 MV is available but is very rare. A small energy dispersion is favored for the electron gun both for a better imaging coherency and also a higher energy resolution in EELS. The Thermo Fisher Titan STEM installed at the University of Wisconsin-Madison has an energy dispersion of around 1 eV. A STEM with a monochromator installed can often achieve tens of meV energy resolution, with 4.2 meV energy resolution recorded[6],

revealing deeper details in EELS. But the high energy resolution often comes at the price of a smaller beam current, with an effort toward a lossless monochromator for EELS[7].

The aberration corrector, which was invented about twenty years ago[8], has become more common and hugely benefited the spatial resolution of STEM. Fig 2.2 is adapted from ref [9] shows the evolution of microscopy resolution, with aberration-corrected electron microscopy dominating the highest resolution regime. The recent improvement in imaging technique and imaging hardware has pushed the resolution to the sub-Å regime, with 0.39 Å resolution recorded with ptychography[10] and 0.41 Å resolution recorded using a novel aberration corrector[11]. The widely applied aberration corrector has made atomic resolution imaging a routine in STEM.

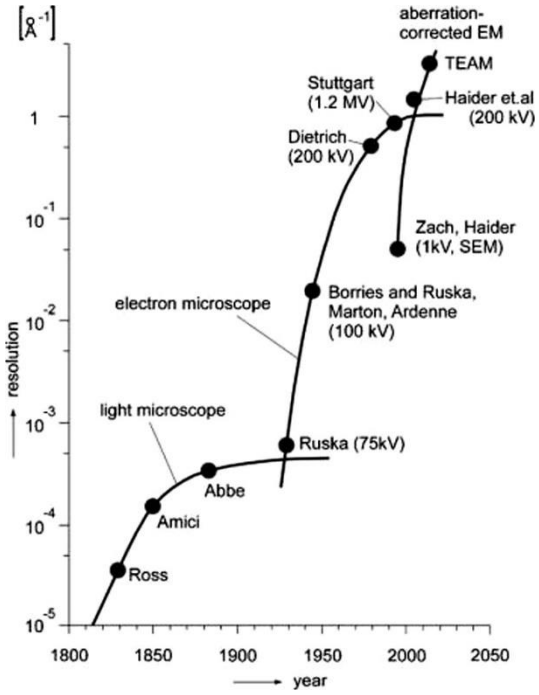

Figure 2.2 Evolution of spatial resolution for microscopy techniques

The sample stage in STEM draw increased attention recently as it can be used to achieve a wide variety of experimental conditions and achieve in-situ and in-operando experiments. A standard STEM holder can be moved along x, y, z direction within about 1 mm range, with a tilt range of about ± 40° along one or two directions depends on whether it is a single tilt or double tilt detector, and a specially designed holder can be tilted close to ±90° for tomography. Besides the standard movement and tilt, in-situ holders can apply different conditions on the specimen include heating, cooling, mechanical force, and electric field. Some other holder design also allows users to put specimen inside a liquid or gas environment.

The transmitted electrons are detected by multiple detectors, include a round-shaped HAADF detector, removable disc-shaped bright field (BF) detector, ronchigram camera, and EELS detector. HAADF detector and BF detector integrates the electron within a certain angular range and produce a single signal value for each probe position. Ronchigram camera, on the other side, collects the whole CBED pattern but is typically not synchronized to beam scan and thus produces one CBED pattern for multiple beam positions. The EELS detector, which uses an energy drift tube to distinguish electron with different energies, will produce a single spectrum at each beam position.

An important part of any STEM experiment is the sample. A basic requirement for STEM sample is the thickness has to be less than a few hundreds of nm, in order to let the high energy electron transmitted through. Sample thickness of less than 100 nm is routinely required for high-quality STEM experiments, with preferred sample thickness being less than 40 nm. A high-quality sample should also have a uniform thickness, fewer impurities, less residual strain, fewer surface contamination, and small damage introduced during sample preparation. Multiple ways have been proposed to prepare a thin and high-quality sample.

Two ways have been used most frequently in the studies in this thesis are mechanical polishing[12] and focused ion beam (FIB) lift-out method. The mechanical polishing method has the advantage of creating a large uniform thin area with small damage, with the possibility of later modification and surface refinement using an ion-mill or nano-mill. The disadvantage of this method is that the sample has to be water-resistance or organic solution-resistance, and the preparation process consumes a relatively large piece of the sample. Also, it will be hard if you want to thin a very specific part of the whole sample. FIB lift-out method, on the other hand, compensates the disadvantages of mechanical polishing as it operates under vacuum with no usage of water or organic solvent. As FIB builds on top of an SEM, it allows the user to visualize the sample that you are working on and thus can create a sample from a very specific region and only consume a small portion of the sample (about 10 μm^3). FIB process is also possible to be programmed and runs automatically. The disadvantage of FIB is that it often introduces beam damage due to the usage of the high energy beam to thin the sample, and the final sample is typically thicker (50-80 nm) with amorphous layers on each side.

The FIB used for the studies in this thesis is a Zeiss Auriga dual-beam FIB-SEM system operates with a Ga ion beam. 30 kV ion beam is used at the beginning stage to create a trench and lift out a small lamella of the sample with about 15 × 2 × 10 μm size. The lamella is then thinned under 30 kV ion beam with a decreasing beam current from 300 pA to 20 pA. The sample is then final thinned using 5 kV ion beam and 100 pA beam current to get a uniform surface. The sample is then finished in a Fishione 1040 nanomill with 900 eV Ar ion on both sides to further remove the amorphous layer and get a smooth surface. The final thickness of the sample is expected to be less than 40 nm, with a thickness of less than 20 nm routinely achieved.

The mechanical polishing tool used for the studies in this thesis is an Allied High Tech Multiprep polishing wheel. Wedge-shaped samples are polished using diamond paper with decreasing particle size from 30 μm to 0.1 μm. The samples are then further thinned using a Fishione 1050 ion-mill starting from 3 kV Ar ion beam with voltage gradually decreased and finished at 100 eV. The samples are thinned in Fishione 1040 nanomill with 900 eV Ar ion on both sides to further remove amorphous layers if necessary. The thickness of amorphous layers on each side is expected to be decreased on each step with a lower beam voltage with a final thickness of less than 5 nm.

Many other ways have been proposed to prepare TEM samples. Another way that has been used in the studies in this thesis is the dispersion of nanoparticles. The nanoparticles are dispersed in absolute ethanol and then dropped on to the top of carbon film supported by a Cu grid. The nanoparticles will be dispersed on top of carbon film in random orientation after ethanol evaporated. Other ways that have proposed by other people but are not utilized in this thesis include electropolishing[13], small-angle cleavage[14], dimple grinding[15], single-crystal crushing[16], and ultramicrotome[17]. Each method has its own field application and advantage, and users should choose the correct method base on the experiment design and specific material.

Sample contamination induced by electron beam irradiation has long been a serious problem for high-resolution imaging in TEM and STEM. Carbon contamination from hydrocarbon on the surface of the sample or in the environment deposits to the surface hitted by electron[18],[19], and the accumulated carbon contamination generates strong contrast and a loss of resolution. In EELS, carbon contamination also generates a strong carbon-K edge in the spectrum and makes it impossible to analyze the peaks that are close to the carbon edge. The problem is more serious

for STEM[20] as it has a much higher electron dose than the conventional TEM. And it could be more serious on the samples with a high-energy surface such as 2D materials.

The most common effort to solve this problem is to remove the source of hydrocarbon. Samples are typically treated in oxidizing plasma to remove the hydrocarbon on the sample surface before being transferred into the TEM column. At UW-Madison, a GV10x Gentle Asher cleaner was used to clean the sample for 10 minutes under 20 W each time before the sample is transferred into the TEM column. The gentle plasma generated by the Asher cleaner does not damage fragile samples such as 2D material or the lacey carbon film. For the Titan STEM at UW-Madison, an on-column Gentle Asher cleaner is installed to clean the whole column together with the installed sample. This setup was also reported elsewhere[21] for a better chemical imaging of carbon in polymer samples. All the tools for sample preparation, include tweezers, TEM grids, sample storage boxes, and glassware, are all cleaned and baked if possible to prevent possible introduction of hydrocarbon onto the sample.

A less common way to solve this problem is to shower a larger area of sample under a high current electron beam. By doing this, carbon contamination will be uniformly deposited on a large area of the sample, typically μm size, to prevent it from builds up locally when doing high-resolution imaging. This method was used occasionally by showing the beam under 160 MX magnification with about 100 pA beam current for 30 minutes before acquiring high-resolution images.

2.3 Four-dimensional STEM (4D STEM)

Starting in recent years, with the development of fast direct electron detectors, a new imaging mode called 4D STEM[4] has become possible and gained popularity. 4D STEM refers to the

experiments where a whole diffraction pattern is collected for each probe position. Unlike conventional STEM, where a single value is generated from an integrating detector at each probe position, a whole 2D CBED pattern is collected, and the collected data is indexed in (x, y, k_x, k_y), thus four-dimensional. Fig 2.3 shows the schematic of data acquisition in 4D STEM. As a whole diffraction pattern is collected from each probe position, 4D STEM data contains much more information than conventional STEM and has been used in many different applications. Section 2.3.1 will introduce the detectors for 4D STEM, three different cases of application on crystalline materials will be introduced in section 2.3.2 to 2.3.4 categorized by the relationship between convergence angle and Bragg diffraction angle, then the applications to amorphous material will be introduced in section 2.3.5.

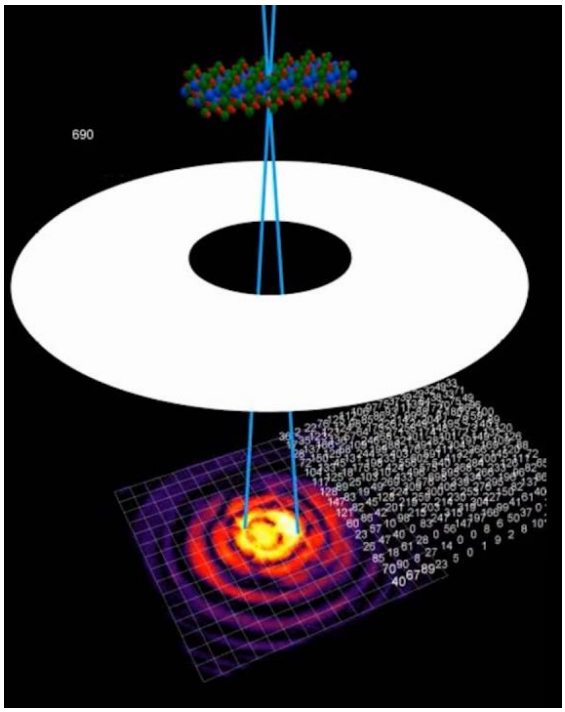

Fig 2.3 schematic of diffraction pattern acquisition on a pixelated detector in 4D STEM.

2.3.1 Detectors

The advent of electron detectors is the key to enable 4D STEM imaging mode. Traditionally, diffraction patterns are either visualized on a phosphorous screen or being recorded on a charge-coupled device (CCD)[22] with digital readout. CCD camera's acquisition rate (typically tens of Hz) is much lower than the STEM scan rate, which is typically at tens to hundreds kHz. To synchronize STEM scan to a CCD camera, one is limited to a coarse scan grid with low resolution. Otherwise, sample drift would dominate the whole acquisition and make the acquired image uninterpretable.

	Detector Pixels	FPS at full pixels	Maximum FPS	Detector px at max FPS	Single e^- SNR@200kV	Dynamic range (e^-)	Physical pixel size (μm)	Detector type
FEI EMPAD[23]	128x128	1100	1100	128x128	142	$1.85x10^6$	150x150x500	PAD
Jeol 4Dcanvas[24]	264x264	1000	10000	264x24	380	5000	48x48	CCD
Quadro by Dectris[25]	514x514	4500	18000				75x75	
Gatan K2 Summit[26],[27],[28]	3838x3710 (sub-px readout possible)	400 (40 FPS transfer speed)	1600	1920x460	39	136 (300kV)	5x5	CMOS
Gatan K3 IS[29]	5760x4092	1500						CMOS
DE-16[30]	4096x4096	92	4237	2048x128	67.5	20	6.5x6.5	CMOS
LBL 4DSTEM	512x512 + integration detector		100000					CMOS
UW-DE Ultrafast	1024x1024	10000-	120000	128x128				CMOS

		30000						
Merlin Medipix 3[31],[32]	256x256	1200	24414	256x256, 1bit			55x55x300	CMOS

Table 2.1 List of different electron detectors and their parameters, Quadro, LBL 4DSTEM, and UW-DE Ultrafast detectors are not commercialized yet.

There are two primary design routes for electron detectors. One is adopted from the detectors for cryo-TEM, which uses monolithic active pixel sensors (APS) from complementary metal-oxide-semiconductor (CMOS) chips. This type of detector is very sensitive to electron hitting and typically consists of many small pixels. Gatan K2, K3, DE-16, and the two ultrafast cameras under development at LBL and UW in table 2.1 used this design. The drawback of this design is that the sensor could be easily saturated at high beam current. This is not a problem for conventional TEM as the beam current is spread over a large area, but imposes a major problem for STEM as the electron dose is much higher in STEM and the majority of the dose is inside the undiffracted beam. A solution to this problem is to use a higher frame rate to reduce the electron dose within the exposure time of a single frame. At extremely high frame rate, counting mode could be achieved, which is widely used in conventional TEM that generates binary frames, and the number at each pixel simply indicates whether an electron hits this pixel or not. This mode could reduce the data size of collected data by replacing 8-bit or 16-bit data with 1-bit binary data, and simplifies the noise pattern in collected data to pure Poisson noise. Up to now, no detector is fast enough to achieve ubiquitous counting mode for all STEM experiments, but efforts are made to develop ultrafast detectors, and there have been some successful researches on ptychography in counting mode[33] using Merlin Medipix-3 detector with 1-bit output.

The other design uses thick pixels and has a high dynamic range optimized for STEM and is called a hybrid pixel array detector (PAD). EMPAD and Medipix in table 2.1 are two sensors with this design. The benefit of a PAD detector is that it has an almost infinite high dynamic range, and signal at a very high collection angle can be collected simultaneously with undiffracted beam without saturating the pixels exposed to undiffracted beam. This design is especially beneficial for ptychography as a larger collection range can be used to enhance the spatial resolution in the reconstructed images[10]. One of the disadvantages of this design is the large pixel size, which could cause the loss of some fine features in the CBED pattern. However, so far, this is not the limiting factor for the 4D STEM experiment as the number of pixels on any pixel array detector is more than enough to resolve the features on CBED patterns. Other disadvantages for this detector design is unclear for now, as this design is rather new with very few different PAD detectors and related studies. There could be some limitation on frame rate as the electron signals need to travel through the thick active layer to reach the circuit bump bonded at the bottom.

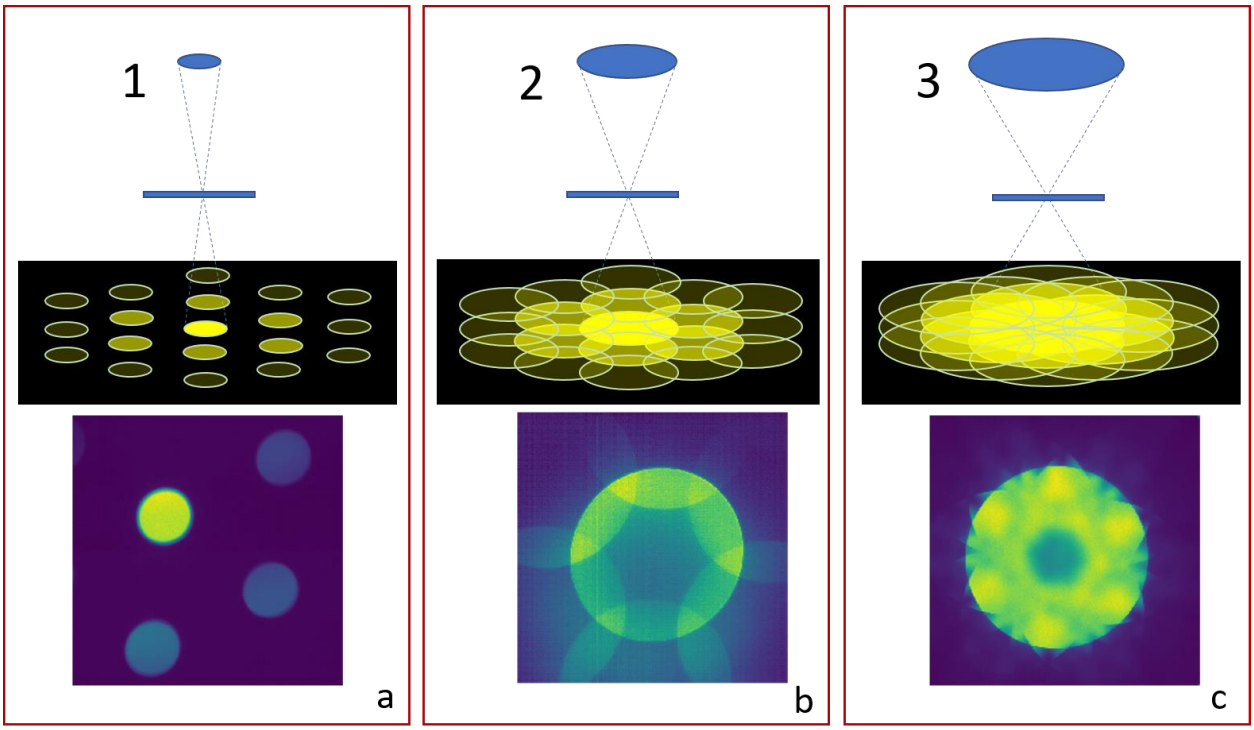

Fig 2.4 Schematic and CBED pattern acquired on multilayer WS_2 sample for a) Bragg angle > convergence angle, b) Bragg angle ≈ convergence angle, and c) Bragg angle < convergence angle

2.3.2 4D STEM Mode 1: Bragg Angle > Convergence Angle

In this 4D STEM mode, a small objective aperture needs to be used to get a small convergence angle, typically less than one mrad to a few mrad, as a typical Bragg angle for crystalline material is about 10 mrad. When the Bragg angle in the crystalline material is larger than the convergence angle of 4D STEM, the CBED patterns contain a set of well-separated diffraction disks, and each one of them can be indexed. In this case, any structural change that would change the Bragg angle could be detected by 4D STEM.

The most popular application for this mode is to detect the strain inside samples, which is the change in the lattice parameter and directly changes the magnitude of Bragg angles. The benefit of detecting strain using diffraction patterns instead of high-resolution real space image is that a large area (μm size) can be detected for each time with the same strain precision[34] as high-resolution images[35]. This tool has been applied to multiple materials systems, including irradiated ZrNb alloy[36], GaAs/GaAsP multilayer heterostructure[34], and two-dimensional heterostructures[37]. Multiple refinements have been reported include better detection of diffraction disks position[38], cancel out the effect of sample tilt[39], application of patterned probe[40], and the usage of exit-wave power-cepstrum[41].

This 4D STEM mode was also used to study the twist in multilayer Van der Waals structure such as GaN[42]. The same setup has also been used to reveal the structural polarization in oxides. The results agreed well with high-resolution STEM study and were used to study room-temperature polar skyrmions[43] and negative capacitance[44].

Although with the advantage of being able to detect and index each diffraction spot, the most significant disadvantage of this 4D STEM mode is the resolution as the spatial resolution is typically limited to nm to a few Å due to the usage of a small convergence angle to get separated diffraction disks. Also, it worth mention that the measurement precision under this mode depends on how accurately one can locate the diffraction disks, which could be hard if the disk has a fuzzy edge, irregular shape, or unclear edge due to low electron dose or thick sample.

The application of 4D STEM on amorphous materials also often adapt this image mode, as a high spatial resolution is not necessary for the samples without repeating structure. There have been few studies focus on apply 4D STEM to study amorphous material. A representative study

focuses on determining the structural heterogeneity in metallic glasses using this new technique[45].

2.3.3 4D STEM Mode 2: Bragg Angle ≈ Convergence Angle

With a moderate size objective aperture and a convergence angle of about 10 mrad, this 4D STEM imaging mode with simple overlapping between diffraction disks could be achieved. Under this 4D STEM mode, diffraction disks overlap with each other and result in interference in the overlapping region, but the overlap patterns are simple, so one can still identify and index each diffraction disk.

This imaging mode is especially helpful for a phase retrieval technique called ptychography[46],[47]. Ptychography seeks to recover the lost phase information in the transmitted electron probe at each probe position using the interference between the undiffracted beam and diffracted beam. The theory was proposed about thirty years ago[46] with simple Wigner deconvolution algorithm[48] solving ptychography problem and some early year experimental implementation[49],[50] showing resolution from phase image beyond the microscope's information limit. Ptychography started to draw some considerable attention with the recent development of the fast direct electron detector, which enabled the simultaneous acquirement of the whole diffraction pattern, which makes ptychography data much more accessible.

Using the fast direct electron detector, researchers used psychography to achieve enhanced contrast on light elements[51], higher spatial resolution on low-voltage STEM[52], and even brake the world record for the highest spatial resolution achieved in a TEM[10]. Studies have also been carried out to optimize the ptychography technique, includes the new iterative

ptychography solving engine using multiple diffraction patterns[53], graphic processing unit (GPU) based ptychographic solver with much higher processing speed[54], and systematic study of optimized imaging condition using pixelated detectors[55],[56].

There are also significant limitations to ptychography. First, ptychography solves the phase information in the transmitted electron probe, and the phase information may not be directly related to the atomic structure when the sample is thick with complex dynamic diffraction. And it would be especially confusing when there is a phase wrap (phase change by 2π) and make ptychography result uninterpretable. Also, in many cases, ptychography problems are solved under the weak phase object assumption, which assumes the sample only changes the phase of the electron wave function and does not change the magnitude. This also only holds true for very thin samples. As a result, ptychography is very useful for thin samples, such as 2D materials, or materials with light elements, such as polymer and biomaterials. Recently, there has been some effort to extend psychography to thicker samples and solve the whole 3D electrostatic potential using a more complex algorithm[57], but to do this would require a huge dataset acquired from different defocus values and consumes lots of computational power to solve the problem.

2.3.4 4D STEM Mode 3: Bragg Angle < Convergence Angle

With a large objective aperture and a large convergence angle (often > 20 mrad), this 4D STEM mode with high spatial resolution can be achieved. This setup is the same as conventional high-resolution STEM and high-resolution STEM spectroscopy. The benefit of this image mode is an extremely high spatial resolution, often sub-Å, that can be used to resolve individual atom in the sample. As a high convergence angle is used, there are lots of overlap between different diffraction disks. When the sample is thick with complex dynamical diffraction, it becomes hard to identify and index each diffraction disk, and the CBED patterns often consist of a bright

undiffracted disk and lots of arcs around it. As a result, it could be hard to study the material structure using the information from a certain diffraction disk.

The most popular application of this 4D STEM mode is to visualize the electric field or magnetic field. This application takes advantage of the fact that the charged electron probe drifts under Columb force or Lorentz force, and visualize the field by detecting the shift in electron probe's center of mass. Related work could date back to about ten years ago using the segmented STEM detector. Then the results are refined with the pixelated detector to better resolve the center of mass shift[58],[23]. Nowadays, an atomic resolution electric field map can be routinely obtained on a stable enough STEM with a moderately fast pixelated detector. Furthermore, there have been works to quantitatively measure the electric field[58],[59], measure large scale polarization field after the cancellation of the strong electric field surrounding nuclei[60], and directly visualize the irregular electric field distribution around each atom site[61].

There have also been studies on the magnetic field under a similar idea using so-called Lorentz STEM. In a Lorentz STEM, the magnetic field surrounding the sample is turned off so that the magnetic structure will not be damaged by the strong magnetic field from objective lenses (around 1 T). Meanwhile, the spatial resolution is limited to tens of nm without the critical objective lenses. With the limit on spatial resolution, researchers have visualized relatively large magnetic structures such as magnetic skyrmion and magnetic domain boundaries[62],[63]. Efforts have been made to develop a magnetic field-free environment around the sample without turning off objective lenses[64], which achieves a magnetic free environment while achieving high spatial resolution. The system is proved to be capable of acquiring atomic resolution images on strong magnetic materials but has not been able to achieve atomic resolution magnetic field

imaging, as the magnetic field is much weaker compared to the electric field surrounding atom nuclei.

Besides, the ptychography reported with record-breaking resolution actually also belongs to this imaging mode as the convergence angle is 21.7 mrad, but each diffraction disk and the overlapping region can be well resolved and indexed due to the usage of the thin 2D material.

2.4 4D STEM System at UW-Madison

A 4D STEM system is developed and employed at the University of Wisconsin-Madison on top of a Direct Electron DE-16 pixelated detector and an aberration-corrected Titan STEM. The system consists of two major parts: a customized scan control system to synchronize the STEM scan and detector acquisition, and a self-developed software to control the whole 4D STEM data collection. The system is designed to have enough bandwidth to accommodate the next generation ultrafast detector, which could acquire 120,000 frames per second. The specifications of the Titan STEM at UW-Madison can be found from ref[65], and the specifications of the DE-16 detector can be found from refNN. This section will introduce the components of the 4D STEM system at UW-Madison, its different modes, and the typical operation procedure of the system.

Name	Manufacturer and model	Bandwidth	Other parameters
Arbitrary wave generator (AWG)	Keysight M3201	200 MHz	2 GB memory, 16 bit, 1GSamples memory, 0.5 V output range

Digitizer	Keysight M9217A	32 MHz	2 channel, 16 bit, 32 MSample memory, 256 V detection range
Amplifier	Keysight 3222	20 MHz	
PCIe Chassis	Keysight M9010A		
External Scan Switchbox	Thermo Fisher Scientific		Two external ports with Lemo connection

Table 2.2 Specifications of different hardware parts in the UW-Madison 4D STEM System

The components of the scan control system can be found in Table 2.2. All different parts are ready for the next generation ultrafast camera, which will be operating at a maximum of 120 kHz frequency, still much lower than the part with the lowest bandwidth (20 MHz from the amplifier) and thus bandwidth is not a concern for the ultrafast camera. A more practical concern for the 4D STEM system is the somewhat limited memory size in the AWG and the digitizer. As the 4D STEM scan is relatively slow, a large number of samples are needed to save the scan pattern and the data generated from the HAADF detector (the latter one is less of a concern as the sampling rate of the digitizer can be changed).

The AWG takes control of the whole 4D STEM acquisition with the camera runs in slave mode. The whole procedure starts when the AWG receives the acquisition command from the scan software, then it starts to generate signals to control digitizer acquisition, camera acquisition, and STEM scans. The STEM scan signal performs a conventional row-wise pattern like the internal scan system of Titan. The camera stays ready and exposed during the acquisition and does a single readout for each trigger signal it receives from the AWG. The digitizer starts a single

readout from the HAADF detector with the length covering the whole 4D STEM acquisition when it detects the rising edge of the trigger signal.

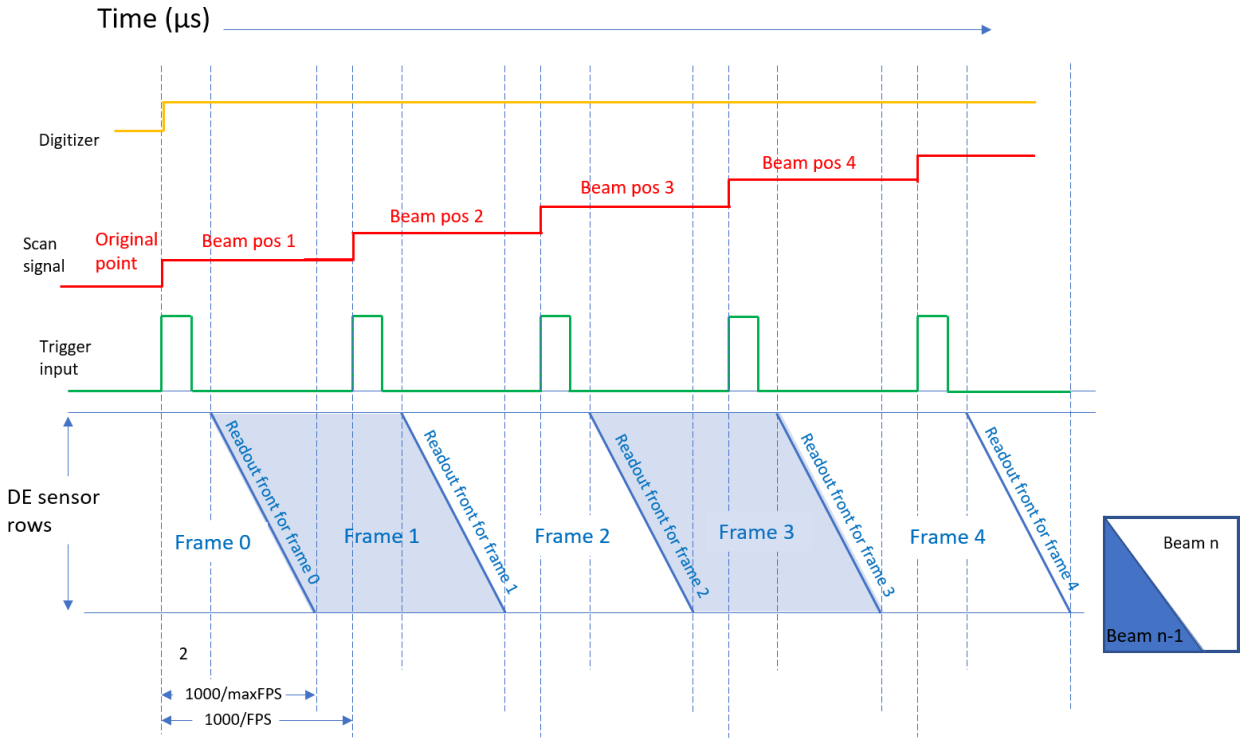

Figure 2.5 UW 4D STEM system timing diagram under the rolling shutter mode.

The DE-16 camera can run in two different shutter modes, namely rolling shutter and global shutter. Rolling shutter adapts a sequential readout structure, where different parts of the sensor expose at different points in time (see the readout front in figure 2-5). The majority of CMOS detectors utilize a rolling shutter. The rolling shutter for the DE-16 camera does the readout in a row-wise pattern. It more effectively uses the limited readout bandwidth by generating a continuous readout stream, instead of doing all at once readout after exposure in a global shutter. Rolling shutter generates artifact when it is used to take the image of a changing object, and this artifact is well studied in ref[66],[67]. In 4D STEM, the moving beam generates the changing CBED pattern. More specifically, the usage of a rolling shutter means each frame acquired

contains a mixture of information from two consecutive beam positions, and different row contains a different fraction of information from the first and second beam position. With this taken into consideration, the beam movement is not synchronized to the start of the readout of each frame, as each frame is going to contain a mixture of two beam positions anyway.

The artifact of the rolling shutter is subtle and can be ignored when the beam scan is oversampled, which means the step size between two beam positions is much smaller than the size of the object, and thus the difference between two consecutive CBED patterns are small. Also, as a result of the beam instability, sample drift, and a finite source size, the image acquired on each frame is from a small local area instead of a single infinite small point. (This actually suggest that we are always collecting PACBED patterns instead of CBED patterns in a real experiment, but we are going to keep using the term CEBD pattern to refer to the data collected from a single frame) The effect of the rolling shutter will be minimized when the scan step size is comparable or smaller than the size of this small area in each acquisition. Another case when the rolling shutter artifact does not matter is when extracting information from PACBED pattern after averaging over multiple CBED patterns, where the information from different beam positions are averaged in the end when generating the PACBED pattern and thus it does not matter if each image contains information from two beam positions. This also suggests that one way to avoid effect from the rolling shutter artifact is to use the information from PACBED patterns instead of from a single CBED pattern.

The DE-16 camera is also capable of running under the global shutter mode at the cost of acquisition speed[67], which eliminates the artifact from the rolling shutter. The timing diagram under the global shutter mode is shown in figure 2-6. Under the global shutter mode, it is important to set the timing of beam movement to be exactly the same as the start of readout,

which is 203.8 µs after the rising edge of the trigger signal after capacitor reset and charge transfer from sensor to the capacitor.

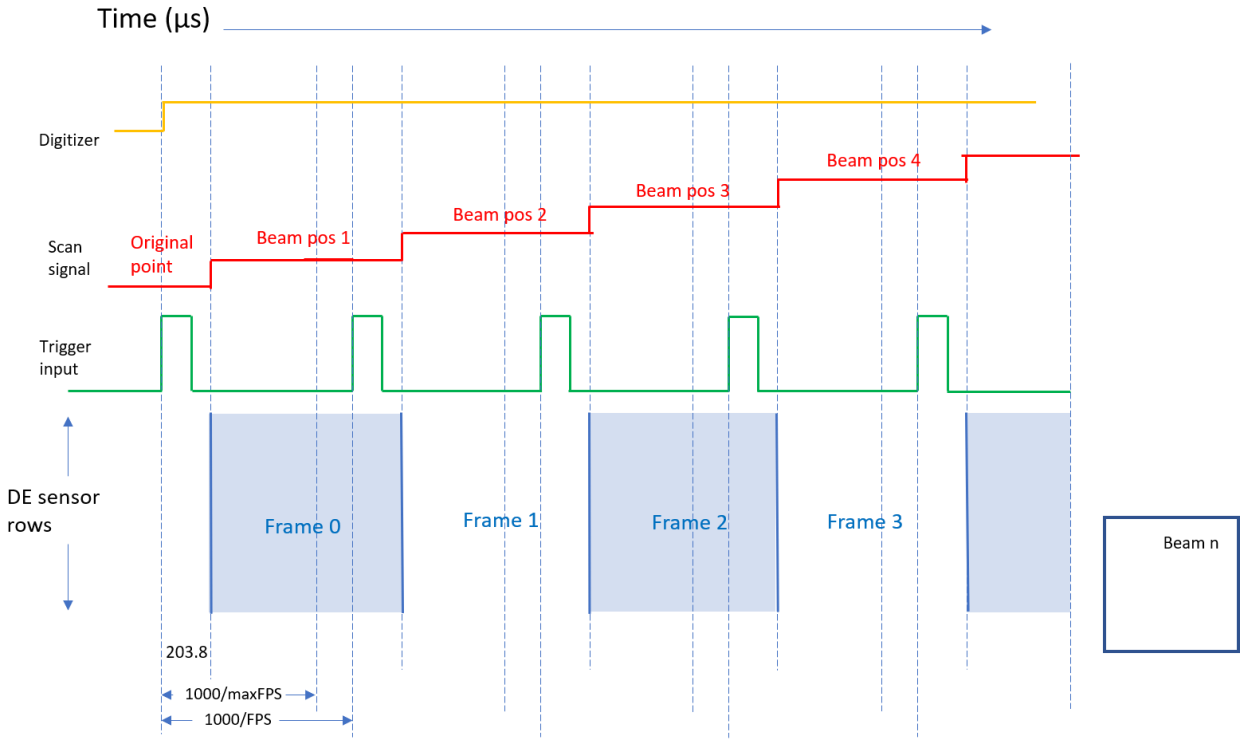

Figure 2.6 UW 4D STEM system timing diagram under the global shutter mode.

A customized scan software coded in C# is used to control the whole scan procedure, a screenshot of the software can be found in figure 2-7. The software can make three different types of scans: 2D acquisition, full-frame 4D STEM, and ROI 4D STEM. The first one runs a beam scan with the HAADF detector and generates a 2D STEM image without using the DE-16 camera. Both 4D STEM acquisition mode runs the DE-16 camera to collect a 4D STEM dataset on Streampix and simultaneously save a 2D HAADF image to the DE computer. The full-frame 4D STEM mode uses the same scan area as the internal scan system of Titan, while the ROI 4D STEM selects a sub-area on the full-frame area using the selection box shown in the software. Although the acquisition is controlled and triggered by the scanning software, the data generated

from the DE-16 camera is collected by a different software called Streampix, which simply stream data from the camera to solid-state drive (SSD) array and has no inter-talk with the scan software. Streampix is used for its ability to directly streaming data from the camera to the SSD array, and the maximum data size is determined by the full size of the SSD array. Another possible option, which was tested before, is to collect data using a software called DE-Server. DE-Server first saves everything to the computer memory, which means the maximum data size is limited by the size of the computer memory and is much smaller than the size of the SSD array, but it also brings the advantage of being able to process and manipulate the data during the acquisition process to show some real-time feedback. The DE-Server option is not adopted mostly for the limitation it imposed on the data size.

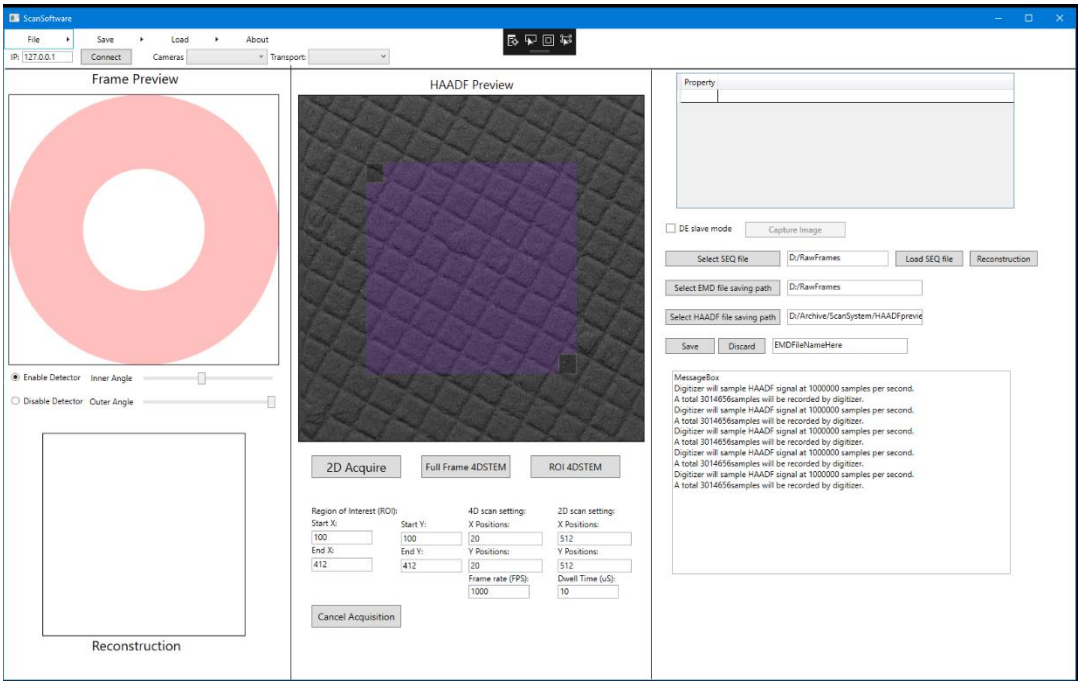

Figure 2.7 A screenshot of the scan control software used for UW 4D STEM system

Using the scan software and Streampix, a typical 4D STEM acquisition procedure includes the following steps.

1. Change the scan source from ‘Internal’ to ‘EELS’ on the external scan switch box.
2. Set the 4D STEM scan parameters on the scan software.
3. Set the save directory for Streampix, change acquisition mode to ‘Slave mode,’ then start the acquisition.
4. Unblank the beam if the beam was blanked before to protect the sample, then click ‘Full-frame 4D STEM’ or ‘ROI 4D STEM’ on scan software.

2.5 Multislice Simulation

Multislice simulation has been applied to simulate the interaction between material and electron and thus simulate the STEM image and diffraction patterns[68], which helps people design experiments, predict experiment results, and interpret the collected data[69]. In multislice simulation, a thick specimen is divided into multiple slices. For each slice, the electron wave function gets a phase shift due to the projected atomic potential of all the atoms within that slice, then propagate for the thickness of the slice. This simulation scheme is very efficient for large structures with many atoms compare to the Bloch wave method, which is typically used for small simulations. So far, there have been multiple implementations of multislice simulation, some of them are CPU parallelized, or GPU parallelized, a complete list can be found in ref[70]. In the works reported here, Prismatic[70],[71] was used to simulate both 2D STEM image and CBED patterns in 4D STEM with the acceleration of an Nvidia P100 GPU.

Although typically being qualitative and often quantitative accurate, two factors are typically not considered in multislice simulations, which could generate some discrepancy between simulation and experiment. One factor is the atomic bonding effects, as almost all the multislice simulation software consider the specimen as a collection of independent atoms with no chemical bonding

considered. To include the bonding information, one has to first draw the bonding information from sources such as density functional theory (DFT), then feed it to the multislice algorithm to replace the simple independent atom model. An implementation could be found in ref[72],[73] with discussions on the effects in the annular dark field (ADF) image. This effect could be more dramatic in some parts of the CBED patterns collected in 4D STEM experiments.

Another factor that is often not considered in multislice simulations is the inelastic scattering[74]. This procedure is typically simulated using a Monte Carlo procedure and is currently implemented μSTEM[75]. The simulation of inelastic scattering convolves three more parameters: characteristic angle, which is the characteristic of the Lorentzian distribution of the inelastic scattering; critical angle, which is the angular limit on the inelastic scattering; and mean free path of the material. The inelastic scattering is shown to be critical for a good quantitative interpretation for both conventional STEM image[76] and the annular intensity in 4D STEM[77].

2.6 Collaborative projects

This section presents all the collaborative projects that I have worked on utilizing the power of aberration-corrected STEM. High-resolution STEM, STEM-EELS, STEM-EDS, and CBED has been used to give insights into different materials, include low-dimensional materials, battery materials, and alloy materials.

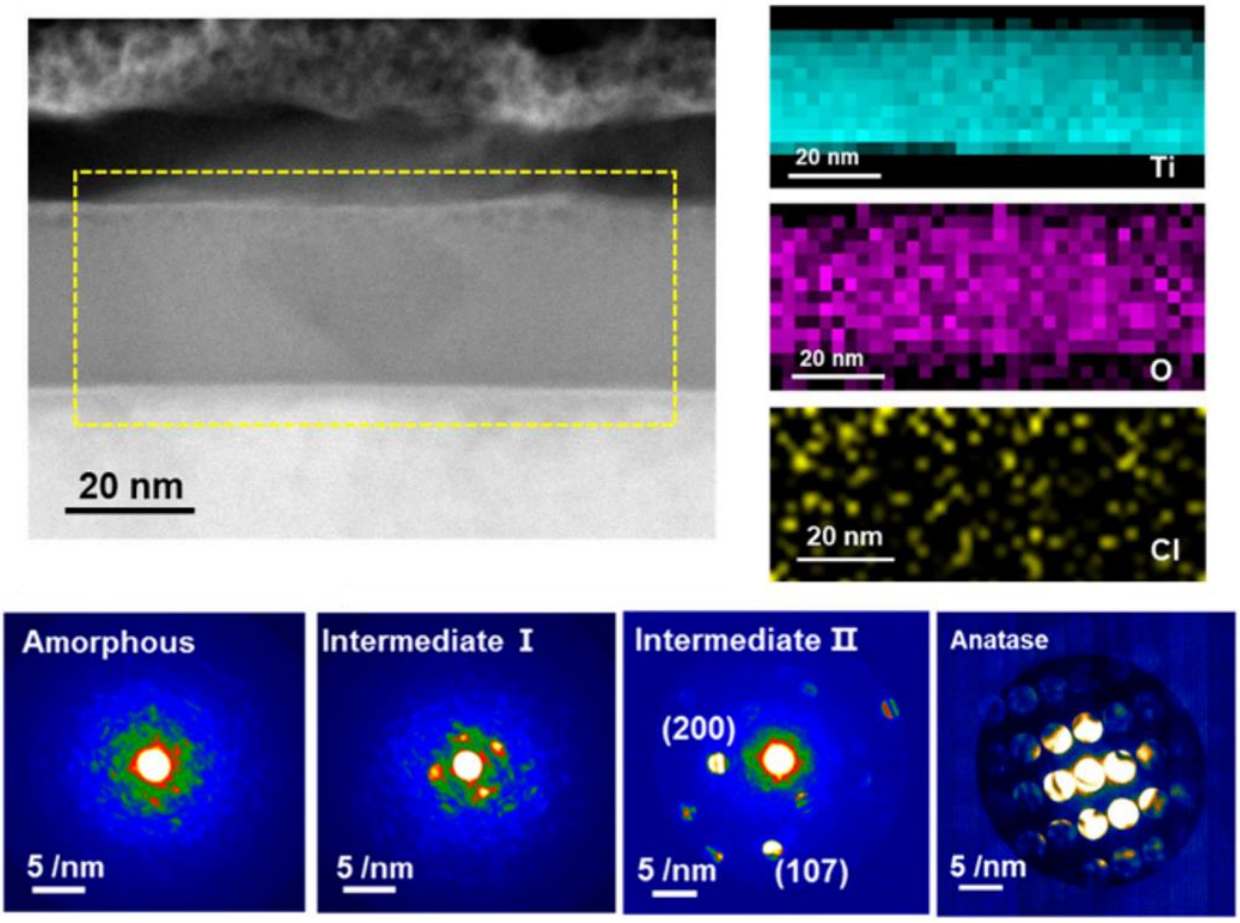

Fig 2.8 STEM-EDS elemental maps of Ti, O, and Cl on the TiO_2 thin film and CBED pattern acquired from different TiO_2 structures.

A hidden metastable intermediate in amorphous TiO_2 thin film was discovered with a critical role in electrochemical damage as they induce localized high electrical current that expedites the etching of TiO_2 thin film when it is used as Si photoelectrochemical photoanodes protection layer[78]. STEM was used to characterize the metastable intermediates in the thin film at high resolution with chemical mapping. CBED patterns acquired on the intermediate phase suggest that there exist two different intermediate phases, and both of them are different from crystalline TiO_2 with an Anatase structure and the pure amorphous structure.

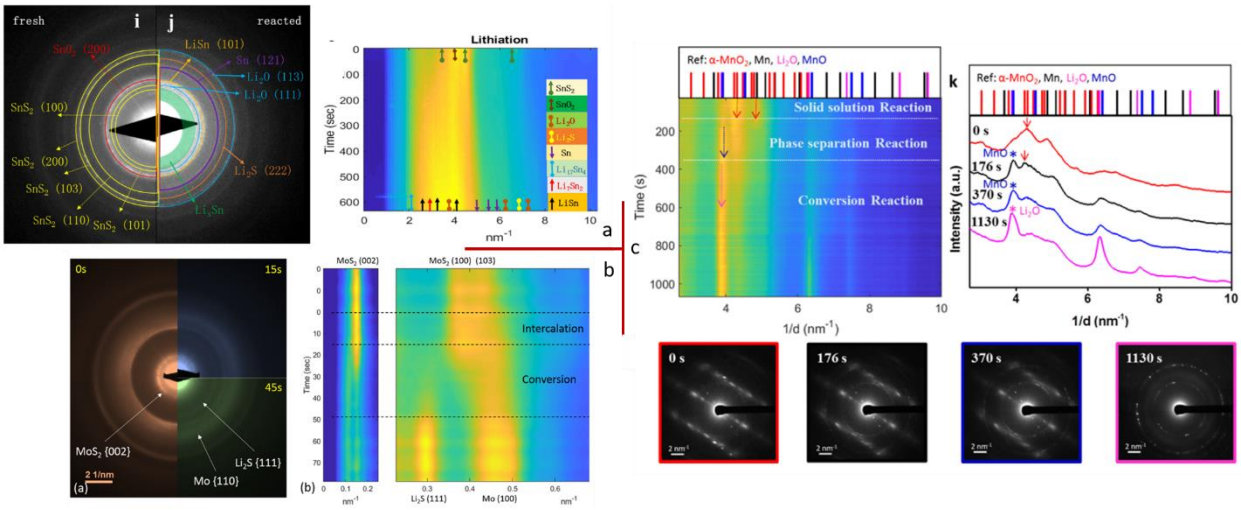

Figure 2.9 In situ TEM diffraction characterization for different transition metal dichalcogenides and oxide used as anode materials for Li-ion battery during lithiation step: (a) graphene-supported SnS2, (b) interconnected vertically stacked MoS_2, (c) surface gradient Ti-doped MnO_2 nanowires.

Nanostructured transition metal oxide and dichalcogenides, include Ti-doped MnO_2 nanowires[79], vertically stacked MoS_2[80], and graphene supported SnS_2[81], were used as anode materials for high-rate and long-life Li-ion batteries. STEM was used to study the microstructure of these materials, and in-situ TEM was used to study the phase transition of the anode materials during the lithiation step. Diffraction patterns were collected during the lithiation step, then the radial intensity profiles of diffraction patterns were derived and used to identify different phases present in different stages of the reaction. The identified phases were used to offer insight into the litiation step and explain the observed performance of the anode materials.

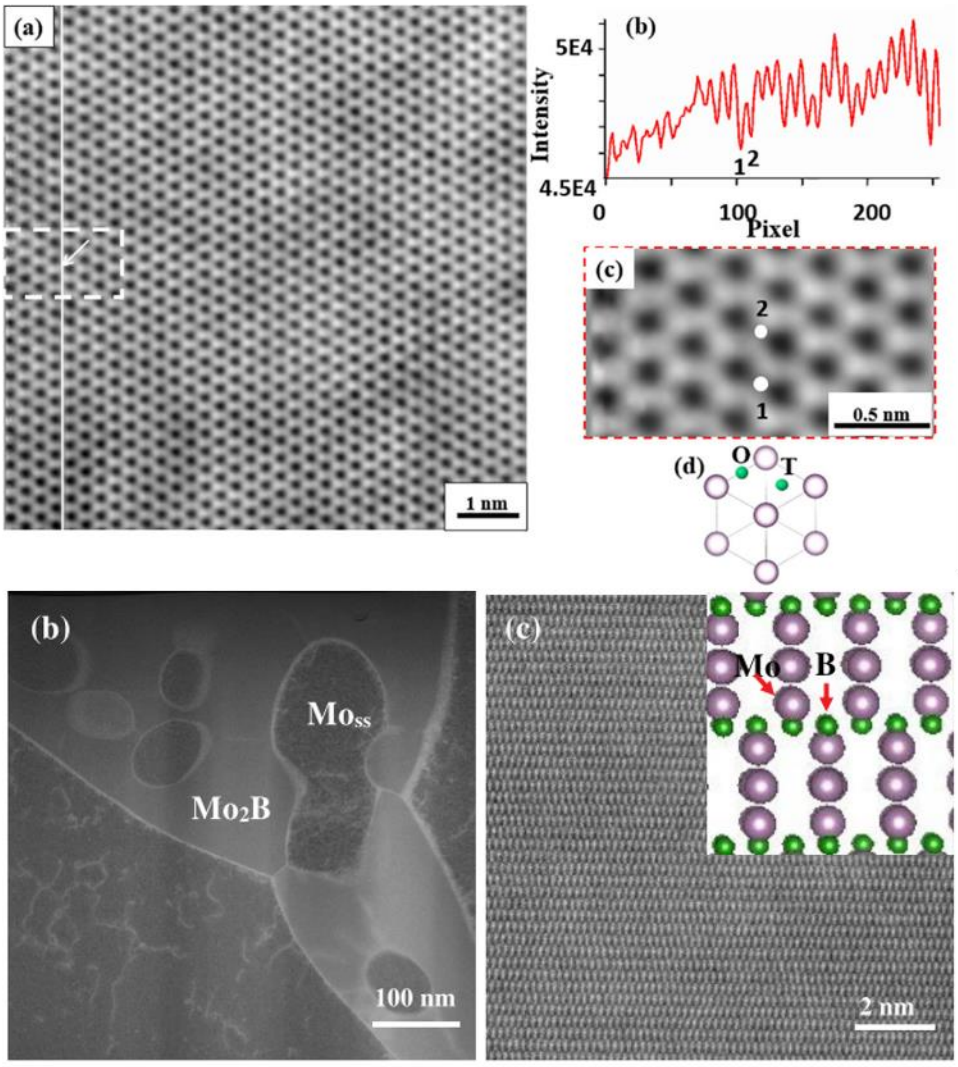

Figure 2.10 Atomic-resolution annular bright-field (ABF) image of Mo solid solution (Moss) with line profile across possible interstitial solute locations (upper) and atomic-resolution HAADF image of Mo_2B phase (lower) in Moss + Mo_5SiB_2 + Mo_2B alloys.

Aberration-corrected STEM has been applied to characterize Moss + Mo_5SiB_2 + Mo_2B alloy[82],[83]. Atomic resolution ABF image was used to analyze the Moss phase in the alloy, and results suggest that B atoms occupy the octahedral interstitial sites of the Mo lattice. Atomic resolution HAADF image was used to analyze the whole alloy which shows phase separation and the Mo_2B

Lattice structure which agrees well with the expected lattice structure.

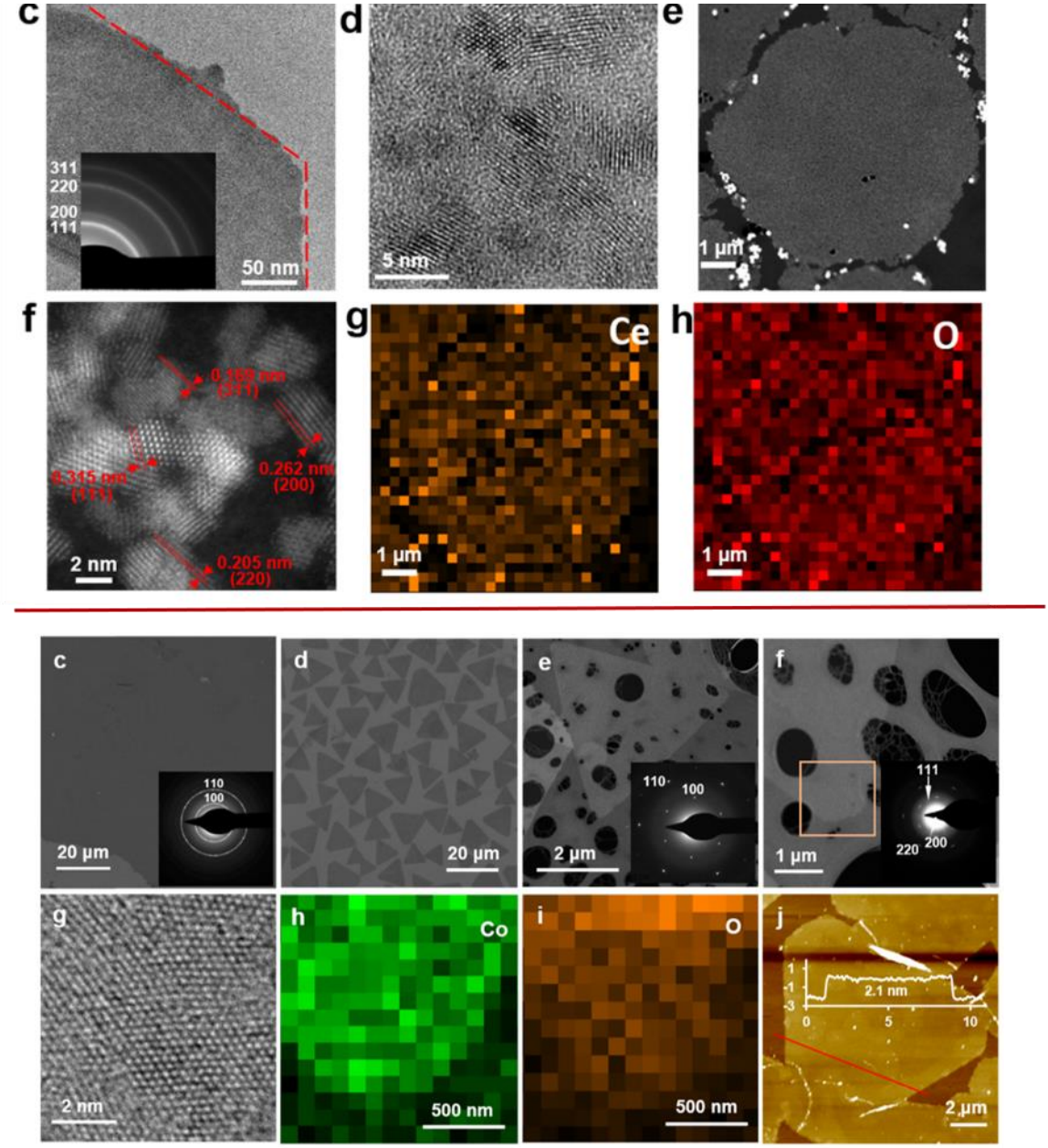

Figure 2.11 STEM characterization of CeO_2 nanosheets and bio-inspired single-crystal CoO nanosheets.

CeO_2 and CoO nanosheets have been synthesized under the help of polymers and characterized using aberration-corrected STEM. Atomic resolution HAADF images suggest that the bio-

inspired CoO nanosheets[84] have monocrystalline structure throughout the µm sized nanosheets and the CeO_2 nanosheets[85] show polycrystalline structure. STEM-EDS was used to study the composition in the nanostructures. EDS elemental maps on both CeO_2 and CoO structure shows uniform distribution of Ce, Co, and O, suggesting a homogeneous structure on the whole nanosheet.

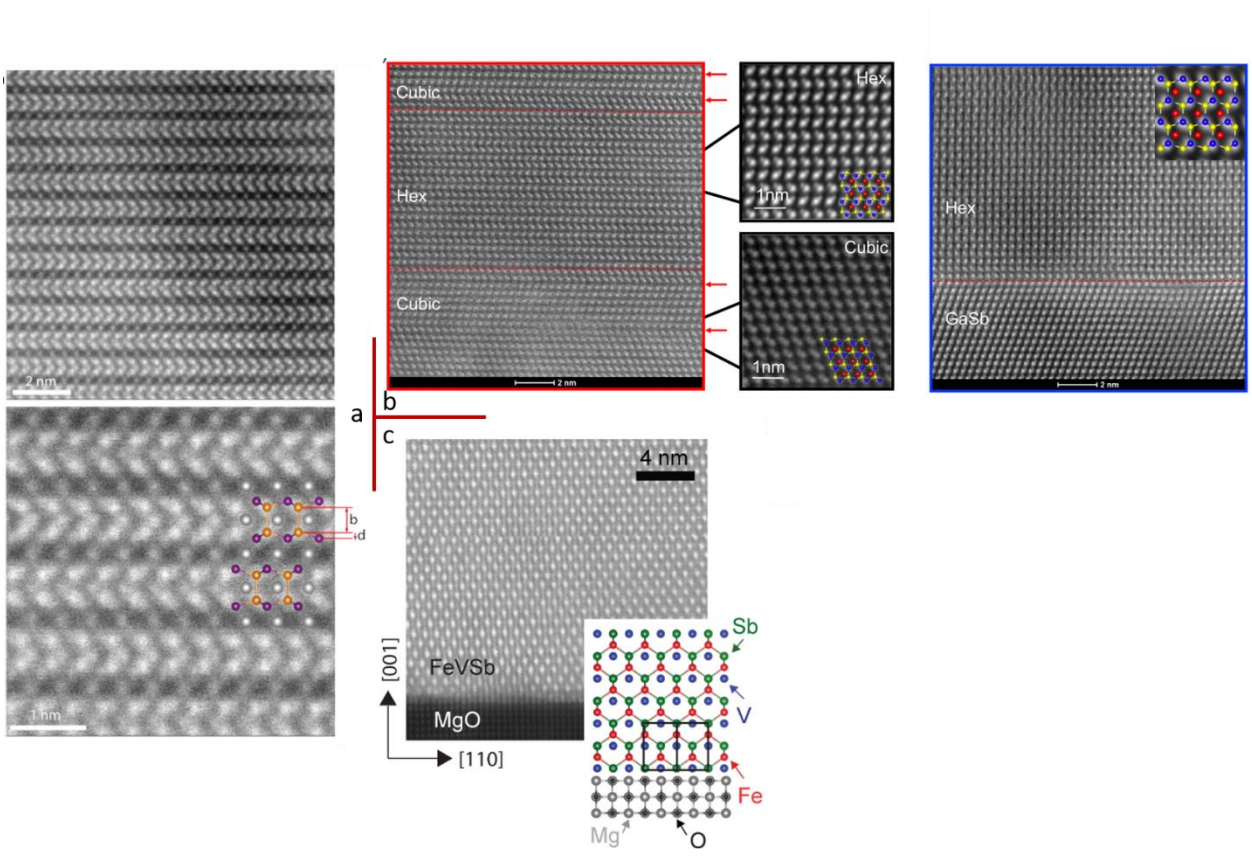

Figure 2.12 High-resolution HAADF-STEM images of (a) LaAuSb, (b) LiZnSb, and (c) FeVSb thin films

Half-Heusler compounds with composition XYZ has drawn attention for its electronic, magnetic, and thermoelectric properties that are tunable by controlling stoichiometry, defects, and structure. High-resolution HAADF-STEM has been used to characterize Heusler alloy thin films of LaAuSb[86], LiZnSb[87], and FeVSb[88] using molecular beam epitaxy. All three samples

show good epitaxial alignments between the film and the substrate, and the lattice structures agree with expected cubic or hexagonal structures.

3 Optimizing Non-Rigid Registration for High-Resolution STEM Image Series

3.1 Introduction

Precise atomic position determination in STEM provides critical information for materials science research in several topics, including ferroic distortions[89], strain states[35],[90], octahedral rotation angles[91],[92], catalytic activities[93], and point defect[94],[95],[96],[97]. To factors need to be considered to get . First is the image signal-to-noise ratio (SNR), which is determined by electron dose and proportional to exposure time under fixed beam current. The second one is image distortion, which refers to the difference between the actual probe scan grid and an ideal regular grid. Image distortions mostly arise from two sources: sample drift during acquisition and scan distortions. The scan distortions are mostly dominated by the random offset deviation at the beginning of each row in the row-wise scan pattern, which results in flagging and skipping artifacts in the collected image.

To obtain both high electron dose and minimum image distortion, the most common way is to acquire a series of fast acquisition frames with small distortion, register the frames and compensate for distortions, then average the registered frames to get a single high-quality image with high precision. Multiple registration methods have been proposed to compensate for the sample drift during acquisition. Rigid registration was first applied by Kimoto et al. .[98], which moves all the pixels in a single frame by the same vector to correct the overall sample drift and achieved precision better than 10 pm in the final averaged image. Berkels et al[99] applied non-

rigid registration (NRR) to register STEM image series by assigning a separate drift vector for each pixel and better fits the pixel-wise acquisition of STEM. Sub-pm precision has been reported using NRR on high electron dose image series[100]. Both registration methods do not use any prior knowledge about the input image and can be applied to any image series. The disadvantage of these methods is that they only correct the low-frequency distortion caused by sample drift, high-frequency distortions such as the flagging and skipping during the scan are not fitted nor corrected.

Methods have been proposed to correct the flagging and skipping at the beginning of each scan row using prior knowledge about the input image series. By assuming periodic lattice structure in the image, Braidy et al.[101] detected the random offset using the phase image generated from the streaks in Fourier space with the peaks from repeating structure as a reference. By assuming a smooth peak shape in the atomic resolution image, Jones and Nellist[102],[103] developed SmartAlign, which corrected the distortions by minimizing the difference between consecutive rows and then reordering rows to get a round shaped appearance of the peaks. Both methods effectively removed the flagging and skipping artifacts, but the assumptions made on the underlying image series means the method should be applied with extra care and they may not apply to all image series.

A group of mode dedicated methods to correct the flagging and skipping artifacts use images acquired from different scan directions. As the fast scan direction has negligible distortion and high fidelity, acquiring data with at least two different fast scan directions can reconstruct a final image with small distortion across the whole frame. Sang and LeBeau[104] developed revSTEM, targeting at constant linear sample drift and is able to resolve accurate lattice spacing and angle on less stable samples. Ophus el al.[105] improved on revSTEM to

include the correction of nonlinear drift, including flagging and skipping using two images with orthogonal scan directions. Ning et al. .[106] further enhance the method for atomic resolution image by emphasizing the atomic position match between images collected with different scan directions.

In this study, we focus on the NRR algorithm used by Berkels et al. and study the conditions to achieve sub-pm precision on the HAADF-STEM image series. Both experimental conditions and registration parameters are studied using image series acquired on $SrTiO_3$ [100] single-crystal samples. This NRR algorithm capture and corrects the overall sample drift pattern during the acquisition without fitting the discontinuous distortions such as the flagging and skipping. However, these distortions could be averaged out in the final averaged image if the magnitude of these distortions has a zero mean value.

The registration quality is measured by calculating the precision of the final averaged image. In this work, the image precision is defined as the standard deviation of the distance between Sr atoms, under the assumption that the strain-free sample should have uniformly distributed Sr atoms and the standard deviation on an ideal image should be zero. The position of Sr atoms is determined by fitting each peak to an asymmetric 2D Gaussian function. Another popular method to determine the image quality on the atomic resolution image is to calculate the residual strain on a nominally strain-free specimen, which was utilized in ref[107].

3.2 Experimental Condition Optimization

To achieve sub-pm precision, it is important to keep the sample close to the zone axis and keep the mistilt less than 1 mrad. The registration result on $SrTiO_3$ [001] sample with a small mistilt angle is shown in figure 3.1 (b). The tilt is estimated to be less than 1 mrad from the [001] zone axis using the PACBED pattern shown in figure 3.1 (c), and the final precision is measured

to be 0.75 pm along the fast scan direction ([100] direction) and 0.77 pm along the slow scan direction ([010] direction). The registration result from the image series acquired with a larger mistilt angle is shown in figure 3.1 (e), where the sample is tilted by about 6 mrad from the zone axis. The precision measured on this dataset is 1.67 pm along the fast scan direction and 1.79 pm along the slow scan direction, both are worse than the precision from the sample with small tilt and over 1 pm. This loss of precision might be caused by the loss of resolution or lower SNR from the less off-zone axis tilt.

Another important factor in the experiment is the drifting rate. We suggest to stabilize the sample for long enough and achieve better than 1 nm/min drift rate. A small sample drift rate is important to guarantee a large enough field of view in the final image. One could also obtain a large field of view by using a larger scan with more scan points.

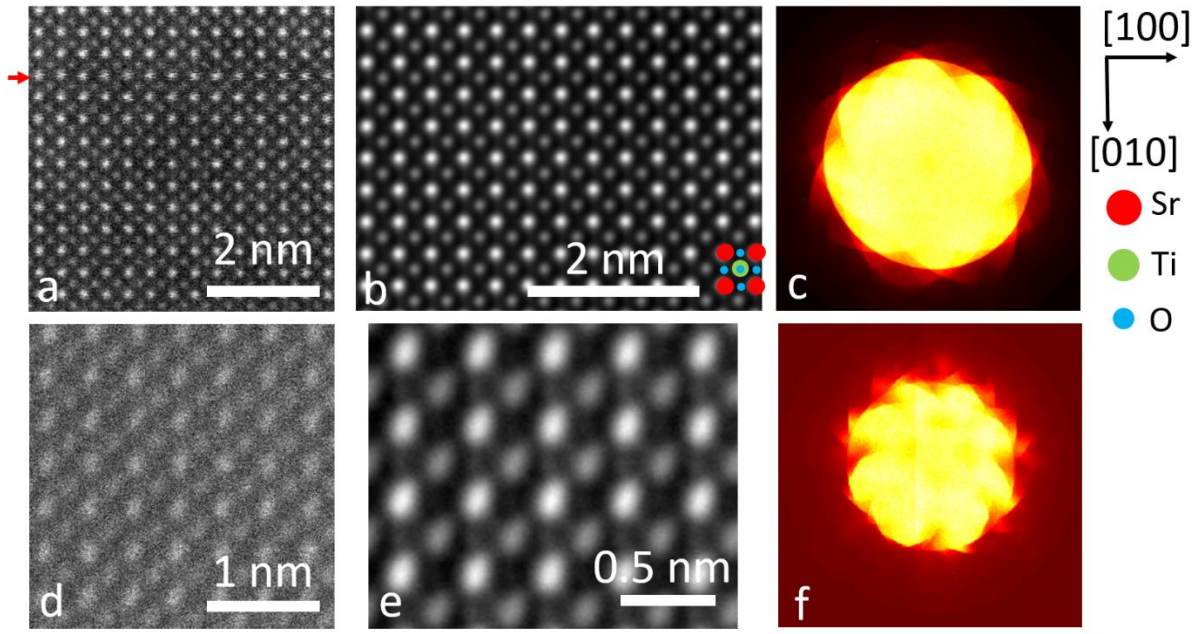

Figure 3.1 (a) single HAADF acquisition, with an example of slicing artifact marked by the red arrow, (b) averaged image after registration and cropped according to sample drift, (c) PACBED

pattern with the center bright disk corresponding to 24.5 mrad convergence angle from an on zone axis $SrTiO_3$ sample. (d-f) shows the single frame, final image, and PACBED pattern for a tilted sample

3.3 Parameter Optimization

The tunable parameters in the NRR algorithm are listed in Table 3.1 with their definition, range, and a recommended value. There is a total of five tunable parameters: the smoothness factor λ, start level and λ multiplier which controls the multilevel registration scheme and the number of registration iterations and the convergence criterion. The detailed explanation of each parameter and why they are being used can be found in ref[99]. Each parameter's effect on the precision of the final averaged image from NRR is tested by varying one parameter at a time with other parameters fixed at the recommended value.

Name	Definition (assume single raw frame has 2^n by 2^n pixels)	Range	Recommend value
Smoothness factor (λ in eq [1])	Controls the relative importance of the smoothness term in the energy, eq [1]	> 0	200
Start level	Images for the first level of registration will be downsampled to $2^{\text{start level}} \times 2^{\text{start level}}$ pixels	An integer between 1 and n for input images of $2^n \times 2^n$ pixels	n-1

λ multiplier	Multiplier to λ for registration of data on a grid downsampled by a factor of 2	> 0	5
Convergence criterion	Stopping threshold for energy minimization	> 0	10^{-6}
Total iterations	The number of total NRR iterations including the initial stage and following refining stages.	integer ≥ 1	3

Table 3.1. List of all parameters in NRR with their definitions, possible ranges, and the recommended values.

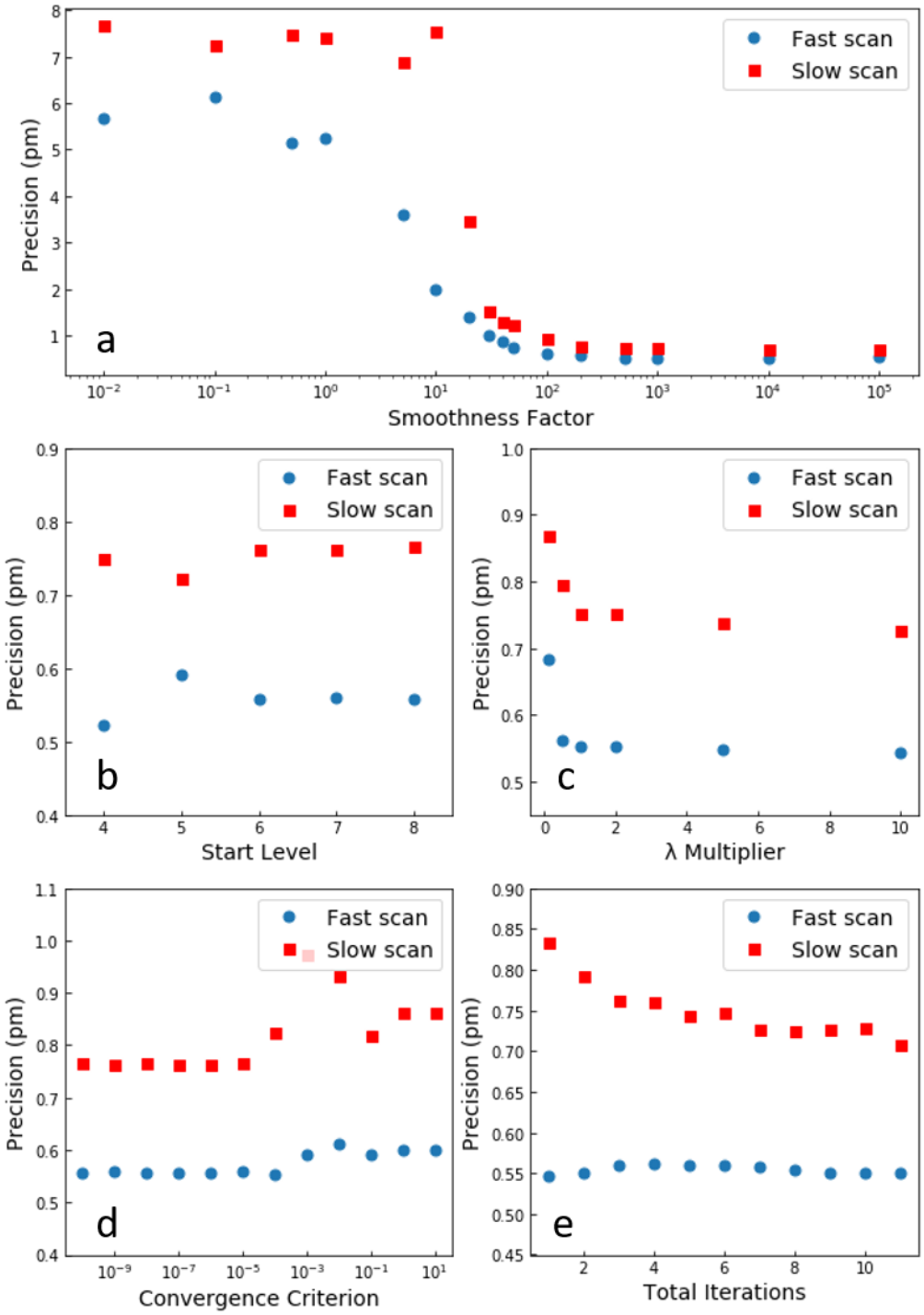

Figure 3.2 Effect of (a) λ, (b) registration start level, (c) λ multiplier, (d) convergence criterion, and (e) total iterations on the precision of the final image.

Figure 3.2 shows the relationship between image precision and the selections of each parameter. Figure 3.2 (a) shows that the smoothness factor λ is the most important factor for a high precision registration. The data used for NRR show a preference for large λ values, with a small λ value causing a loss of precision. There is no significant improvement in precision with λ larger than 200. The preference for a large smoothness factor suggests that the distortion we have on the input image series is smooth and can be well fitted to a distortion field with high smoothness.

Figure 3.2 (b) shows precision from different registration start level, which controls the downsampling in the multilevel registration process and is used to avoid unwanted local maxima. With 2^n pixels along each side in the raw input images and a registration start level of *k*, the number of pixels at the beginning of registration is

$$n_{downsampled} = \frac{n_{original}}{2^{n-k}}, \quad [1]$$

where $n_{original}$ is the side length of the raw input image in the unit of pixels. Registration results show sub-pm precision for all different registration start levels. The registration start levels smaller than 4 were not tested as the downsampled image at the beginning of registration may lose all the features that could be used in registration and cause unphysical sudden jumps between frames. This is especially likely if the image contains only periodic lattice structures. An example of this sudden jump will be shown in section 3.4.

Figure 3.2 (c) shows the effect of λ multiplier. In the multilevel registration scheme, downsampled registrations are performed with a smoothness factor of

$$(\lambda \text{ multiplier})^r \times \lambda,\ r = \log_2\left(\frac{n}{n_{downsampled}}\right), \quad [2]$$

where λ is the smoothness factor used at full pixel size, and λ' is the smoothness factor used on the downsampled image. The result shows a preference of λ multiplier larger than 1, resulting in a larger smoothness factor λ and a higher requirement for registration smoothness at a coarse pixel grid.

Figure 3.2 (d) and (e) show the effects of the convergence criterion and the number of total iterations. The convergence criterion is used as the exit condition when minimizing the energy function, and the image series will be re-registered multiple times depends on the number of total iteration parameters. Result suggests that a convergence criterion of 10^{-6} with a total of 3 iterations is enough to reach sub-pm precision in the final result.

3.4 Artifact Visualization

In this section, we directly visualize the distortion fields in the registration under different registration parameters, in order to get a better understanding of the registration behaviors and how they affect the precision in the final image. In each case, NRR seeks the deformation to register the target frame shown in figure 3.3(b) to the keyframe shown in figure 3.3(a). Figures 3.3(c), (e), and (g) show the vector plots of non-rigid pixel deformations determined under the default parameter, a (too) small start level, and a (too) small lambda value, respectively. All three plots are shown at reduced pixel density (344 pm spacing between vectors) for the whole frame. Figures 3.3(d), (f), and (h) show the deformed frames after applying the deformation.

The pair in figure 3.3 (c) and (d) shows the correctly fitted deformation field, which captures the correct overall drift pattern and the registered patch looks similar to the keyframe shown in figure 3.3 (a). Figure 3.3 (e) shows the deformation field detected when the start level is too low, which generates a half-unit cell jump and causes the Sr sites in the registered patch to overlap with Ti sites in the keyframe. Figure 3.3 (g) shows the deformation field determined

when the smoothness factor λ is too small. Without the smoothness constraint, NRR detects the unphysical deformation field with lots of abrupt changes. In this case, all the distortions in the keyframe will be duplicated in the registered frames and eventually in the final averaged result, causing a loss of precision.

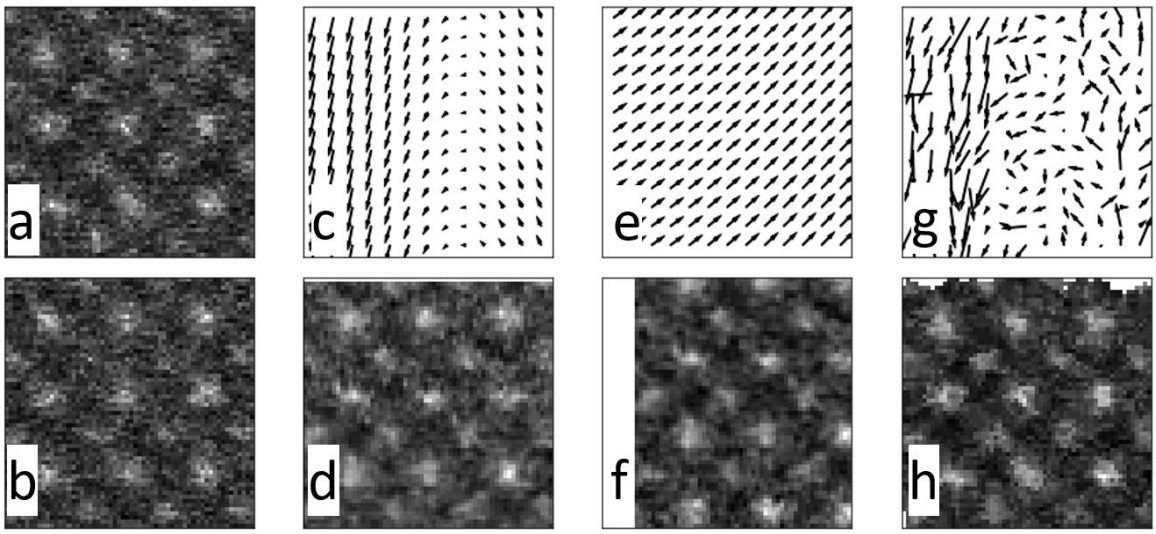

Figure 3.3 Example of the deformation field for registration between two frames under various conditions. (a) and (b) are small patches of the original frames. (a) is the keyframe, and (b) will be deformed to match it. (c) and (d), (e) and (f), and (g) and (h) are pairs of deformation vector fields (c, e, and g) and deformed images (d, f, and h) processed with various choices of registration parameters. (c) and (d) use the default parameter, (e) and (f) used a too small start level, and (g) and (h) used a too small lambda value.

3.5 NRR with Fixed Electron Dose

The total electron dose a certain sample could sustain is often limited, and the fixed total electron dose budget can be divided into more frames with a smaller frame dose, or fewer frames with higher frame dose. Essentially, we are making a choice between frames with fewer

distortions and frames with higher SNR and thus easier to register. To study the precision for different choices of frame dose, we have acquired image series with frame doses ranging from 1.2×10^3 $e^-/Å^2$ to 2.5×10^4 $e^-/Å^2$, varied by using pixel dwell times from 0.5 μs to 12 μs at a fixed beam current. Then NRR was performed on image series with a fixed total dose ranging from 2×10^5 $e^-/Å^2$ to 12×10^5 $e^-/Å^2$ with different frame doses.

Figure 3.4 shows the precision along fast scan direction and slow scan direction against the frame dose/pixel dwell time under different sets of total electron dose. For the slow scan direction, the results suggest that the highest precision can be achieved at 6 μs/px pixel dwell time. For fast scan direction, the highest precision is achieved at 0.5 μs/px dwell time, which is the fastest scan, and 8 μs/px shows a local minimum in precision. Overall, the precision gets worse when pixel dwell time increases beyond 10 μs, as more distortions exist within a single frame and cannot be corrected by NRR. At 0.5μs/px dwell time, the images capture, rather than averaging over, a broader range of distortions, at the cost of possibly larger scan distortions, which result in a good precision along the fast scan direction and slightly worse precision along the slow scan direction. 6-8 μs/px dwell time generates good precision along both scan direction, and we suspect this is the result of some distortion source being suppressed at these scan frequencies. In conclusion, we recommend 6-8 μs/px dwell time for high precision imaging using NRR, but this option could vary for different instruments and different labs.

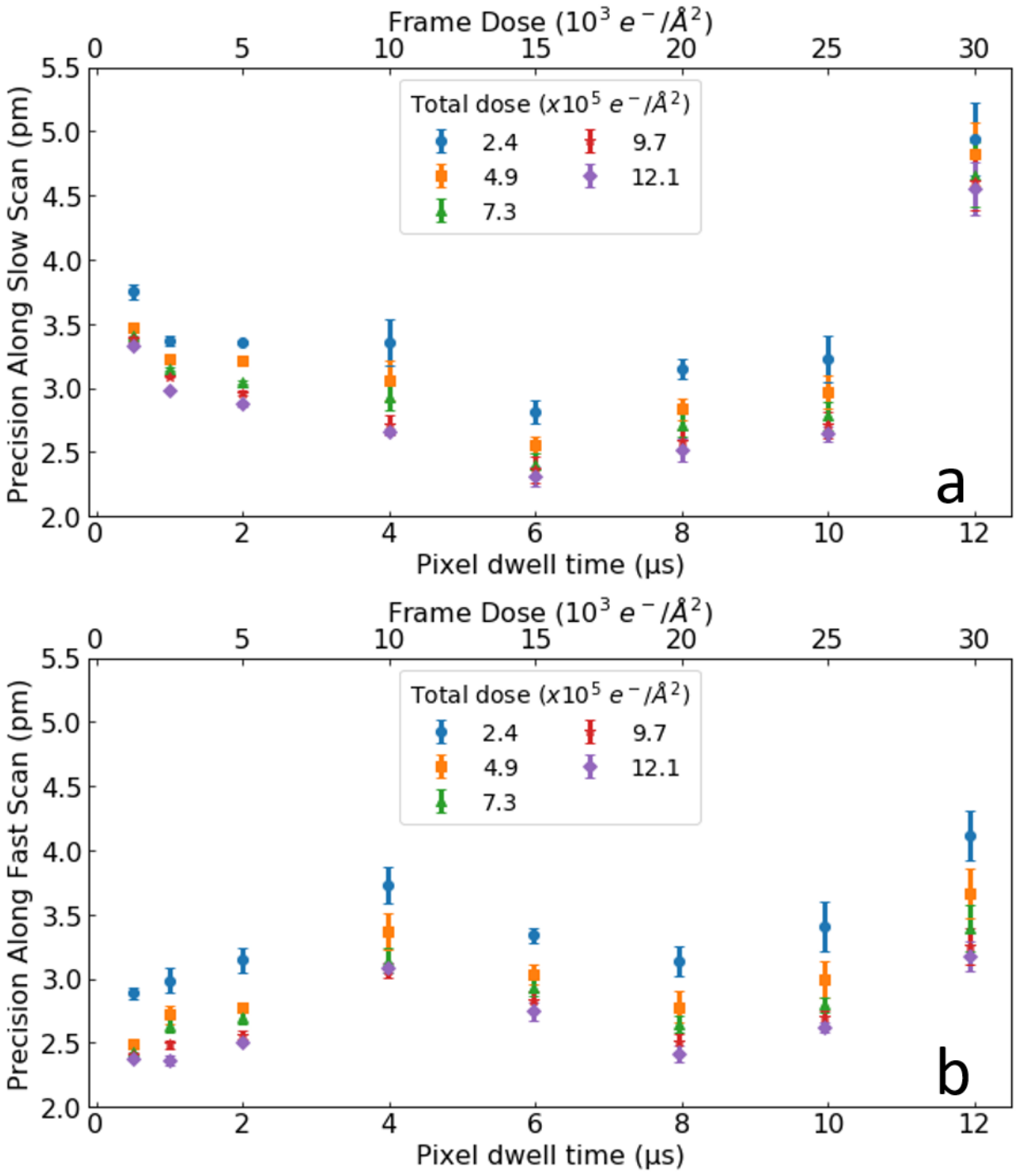

Figure 3.4 Registered averaged image precision along (a) fast scan and (b) slow scan direction under different pixel dwell time / frame dose. The total dose is kept constant by varying the number of registered images.

3.6 Conclusions

Experimental requirements and registration parameters to achieve sub-pm precision have been studied using the atomic resolution HAADF-STEM image series acquired on a $SrTiO_3$ [001] single-crystal sample. A small off-zone axis tilt of less than 1 mrad and a small sample drift of less than 1 nm/min have been found to be critical for a high precision result, while a large enough smoothness factor λ is the most important registration parameter. When applying NRR to image series with fixed total electron dose, the overall best precision is achieved at about 6 µs/px pixel dwell time. Sub-pm precision can be routinely achieved with the experimental condition and registration parameters carefully selected.

4 Non-local Principle Component Analysis to Denoise Hyperspectral EDS Data

4.1 Introduction

One capacity of modern STEM is its integration with microanalysis techniques, including electron energy loss spectroscopy (EELS) and energy-dispersive X-ray spectroscopy (EDS). With modern aberration corrector[108], advanced detector[109], and high stability sample holders and instrument, it becomes possible to study chemical information in STEM at high spatial resolution and possibly at atomic resolution. However, as a result of the limited cross-section for EDS and EELS, the collection of EDS and EELS data still requires much longer pixel dwell time (ms – seconds) to gather enough signal compare to traditional HAADF-STEM (a few to tens of µs), which makes it more important to correct sample drift and denoise to visualize the underlying information.

A few different ways have been proposed to denoise the atomic-resolution hyperspectral data collected in STEM. One simplest way is to apply a spatial filter to only keep data with certain periodicity in the denoised data. The spatial filter could effectively remove random noise and improve SNR as the random noise shows no periodicity and mostly occupies the low-frequency region in the Fourier space. However, the spatial filter is known for forcing periodicity in the denoised image and could possibly remove aperiodic features together with noise. Another widely-applied denoising method is principal component analysis (PCA)[110], which assumes the noiseless ground truth has lower dimensionality and perform the denoise by reducing the dimensionality of the noisy data. PCA has been proved to be useful in the denoising of EELS[111] and EDS[112]. The current state-of-the-art denoising method is the non-local method, which divides the data into small patches, search for similar patches, denoise the group of similar patches, then project the denoised patches back to the original data shape. Mevenkamp et al. applied non-local means[113] (NLM) and block-matching and 3D-filtering[114] (BM3D) to denoise single frame STEM image. Outside electron microscopy, block-matching and 4D-filtering[115] (BM4D) has been developed to directly denoise 3D hyperspectral data.

In this work, we adopted the NLPCA method[116] to denoise the EDS spectrum image. NLPCA builds on previous non-local methods in two ways. First, a k-means clustering[117] step is applied to the collection of all the patches to look for similar patches, instead of doing similarity search within a local search window or along a certain direction, which better help reveal the similarity and redundancy within the whole data. Second, NLPCA applied PCA to denoise groups of similar patches, instead of simply take the average or apply a spatial filter to remove the noise, which is better at removing noise with a more complex distribution. On top of that, the NLPCA we applied has Poisson statistics embedded when measuring the similarity

between image patches and perform PCA to denoise groups of image patches, which makes it a powerful method to denoise EDS data that perfectly obey Poisson statistics.

4.2 Non-local Principle Component Analysis and Parameter Selection

In this section, we will briefly introduce the NLPCA method, where further details of the method and its denoising performance on real-life photos can be found in ref[118].

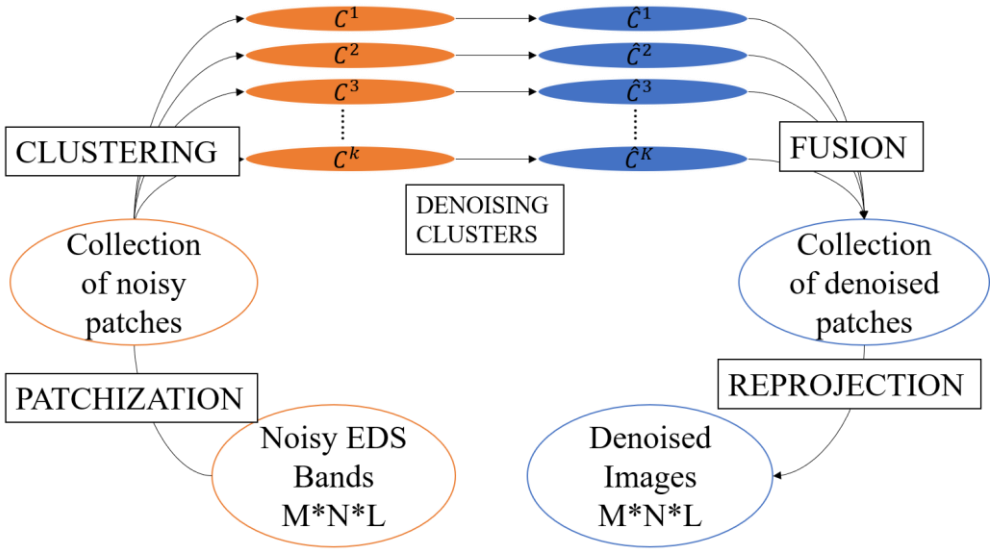

Figure 4.1 Schematic of NLPCA denoising.

The schematic of NLPCA is shown in figure 4.1. When a noisy 3D data cube is fed to NLPCA, it first divides the large datacube into lots of small overlapping cubes. The small cubes are divided into a few groups in the clustering step using k-means clustering, with the difference between two cubes defined as the Bergmann distance measured under Poisson statistic[119], which is essentially measuring the similarity between two cubes assuming both cubes are corrupted by Poisson noise. Then, within each group of similar cubes, a PCA is performed to remove the noise by finding the redundancy and reducing the dimensionality of the collection of

cubes. The denoised cubes are then reprojected back to the original size of the raw datacube to generate a denoised datacube with the same size.

Although being a powerful method and optimized for Poisson noise corrupted data, there are certain limitations on the NLPCA method. First, the denoising parameters, namely the number of clusters in the pachization step and the number of principles components to use in the PCA step, needs to be determined first, and currently, we don't have any effective data-driven method to determine the denoising parameters on specific input data. This also means that the same number of principal components will be used in the PCA step of different clusters, which is not ideal as different clusters may contain different features and may need a different number of principle components for optimal denoising performance. A second limitation is the computer resources needed for NLPCA, as it takes a large memory, often hundreds of times of the raw data size, to keep the collection of all the overlapping cubes. And the k-means clustering and PCA on the collection of overlapping cubes are computationally intensive. All these means it would be challenging, and maybe impossible, to apply NLPCA to denoise datacube with large size.

4.3 Denoising on Simulated Data and Parameter Selection

To test the reliability of NLPCA, check for reconstruction artifacts, and determine the appropriate patch size, the number of clusters, and the number of principle components, we applied NLPCA to a phantom data set with varying levels of Poisson noise. We created a noiseless phantom EDS spectrum image based on the experimental data and the ideal lattice structure of $NdTiO_3$. For every lattice site, a 2D Gaussian distribution of intensity was created in the appropriate bands, on top of an approximate Bremsstrahlung background[120]. The signal-to-background of the atomic lattice in each spectral band and the ratio of intensities from band to band was matched to the experimental data. Three different Ca distributions were considered, all

on the Nd sublattice: uniform substitution, random substitution, and nanoparticles with Ca completely replacing Nd. The O and Ti distributions are the same in all three phantoms. Poisson noise is added with a Poisson random noise generator using the intensity in each voxel as the mean of the Poisson distribution.

Denoising quality on the simulated data was calculated in terms of the peak signal-to-noise ratio using the ground truth as a reference. PSNR is defined via the mean squared error (MSE), which is the averaged squared difference between the target image and the ground truth,

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)]^2. \quad [3]$$

PSNR is then defined as

$$PSNR = 10 \cdot \log_{10}\left(\frac{{MAX_I}^2}{MSE}\right), \quad [4]$$

where MAX_I is the maximum intensity in the ground truth. The denoising parameters were varied one at a time over a wide range to denoise the noisy data and calculate the PSNR on the denoised data. Then a set of parameters that gives the best denoising quality was determined to be the optimized denoising parameters and was used in the denoising of real experimental data.

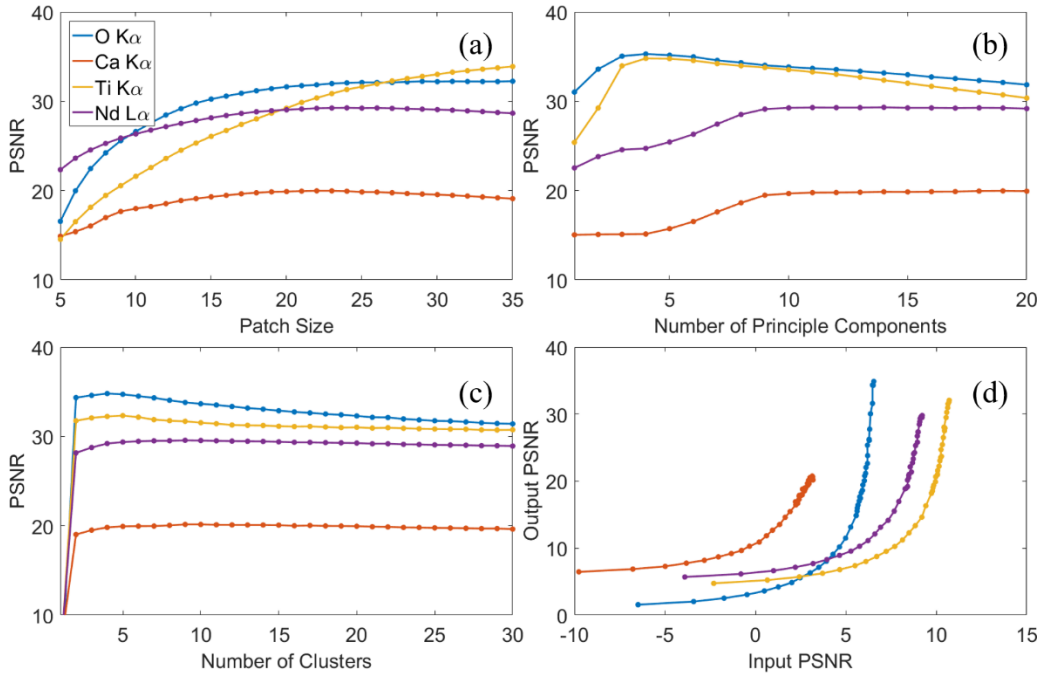

Figure 4.2: PSNR of the denoised random phantom as a function of the NLPCA parameters and the input Poisson noise level: (a)-(c) shows PSNR *vs.* patch size, number of components, and number of clusters, and (d) shows PSNR of the denoised images *vs.* PSNR of the input images.

Figure 4.2 (a)-(c) shows the relationship between denoising quality and three different denoising parameters. Figure 4.2 (a) shows that the PSNR increases with increasing patch size, as a larger patch size captures more information and periodicity inside the patch. However, this also significantly increases the computational cost. As a result, we used 17 px patch size, which is large enough to cover at least one unit cell. Figure 4.2 (b) shows that it takes only 4 components to denoise the O and Ti bands, which requires 10 components to recover the Ca and Nd band with a more complex structure. We selected 10 components for the PCA step, as there is only a small loss in PSNR in the O and Ti band between 4 and 10 components, while less than 10 components could possibly generate artifacts in the Ca and Nd bands with more complex structure. Figure 4.2 (c) shows that the denoising quality is nearly independent of the number of

clusters, as long as it is more than 2, which is reasonable as the extra clusters could be empty and won't do any harm to the denoising result. We used a total of 10 clusters for denoising, in case the data contains a more complex structure that needs to be captured using more clusters.

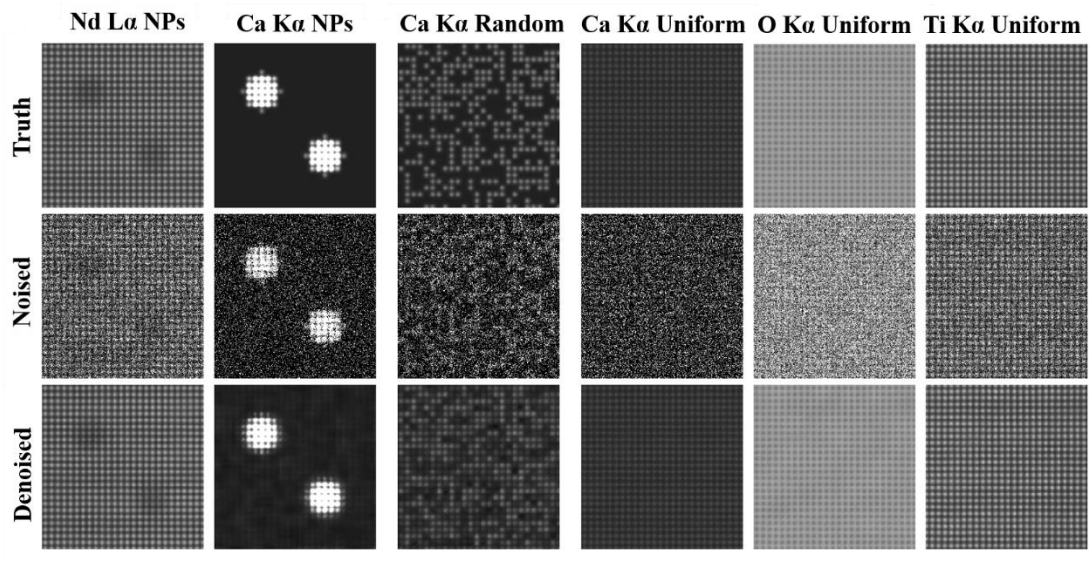

Figure 4.3: Simulated ground truth (top row), noised observation (middle row) and denoised results (bottom panel) for Ca NPs, randomly distributed Ca and uniformly distributed Ca. In every case the Ca occupies the Nd sublattice. The O and Ti distributions are the same for all three phantoms, so they are shown only once. NLPCA performs very well for the highly redundant uniform Ca distribution. Performance for the less redundant distributions is also good.

The denoising results on simulated EDS data is shown in figure 4.3 for three different Ca distributions. The NLPCA perfectly recovered the ground truth for uniformly distributed Ca. For the randomly distributed Ca and clustered Ca, NLPCA was able to recover most of the structure, with some shadow on the denoised Ca map where should be no Ca, as a result of cross-talk between different bands. The randomly distributed Ca atoms show the worst denoising performance, as there isn't much redundancy in the data that can be used for denoising. Overall, NLPCA is shown

to be capable of improving SNR on the data and recover the underlying ground truth for all three different data types.

4.4 Denoising Results and Comparison to Other Methods

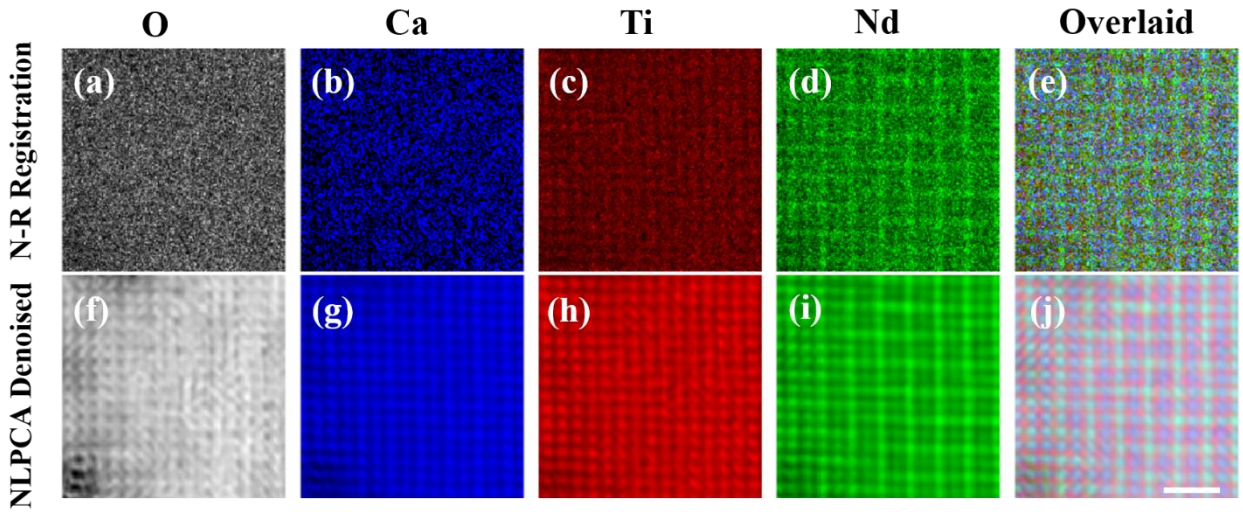

Figure 4.4: Comparison between EDS elemental maps recorded along [001] Ca-stabilized $Nd_{2/3}TiO_3$ after non-rigid registration (a-e) and NLPCA denoised elemental maps (f-j). Elemental maps from left to right display the integrated intensity from (a) O Kα, (b) Ca Kα, (c) Ti Kα+Ti Kβ, (d) Nd Lα+Nd Lβ+Nd Lβ4 peaks and (e) an overlaid composite map of Ca, Ti, and Nd. Scale bar is 2nm and applies to all panels.

With the method tested on simulated data and optimal denoising parameters determined, we applied NLPCA to experimentally collected atomic resolution EDS data from Ca-stabilized $Nd_{2/3}TiO_3$ [001] sample. The comparison of O, Ca, Ti, and Nd elemental maps from noisy data and denoised data are shown in figure 4.4. A significant SNR improvement can be seen in all four element maps, with clear atomic features in Ca, Ti, and Nd maps. The O map shows some features, but we suspect they are mostly denoising artifacts. The denoised data also clearly shows

that Ti occupies a different site than Nd and Ca, which agrees with the fact that Ca was doping the Nd site in the perovskite structure.

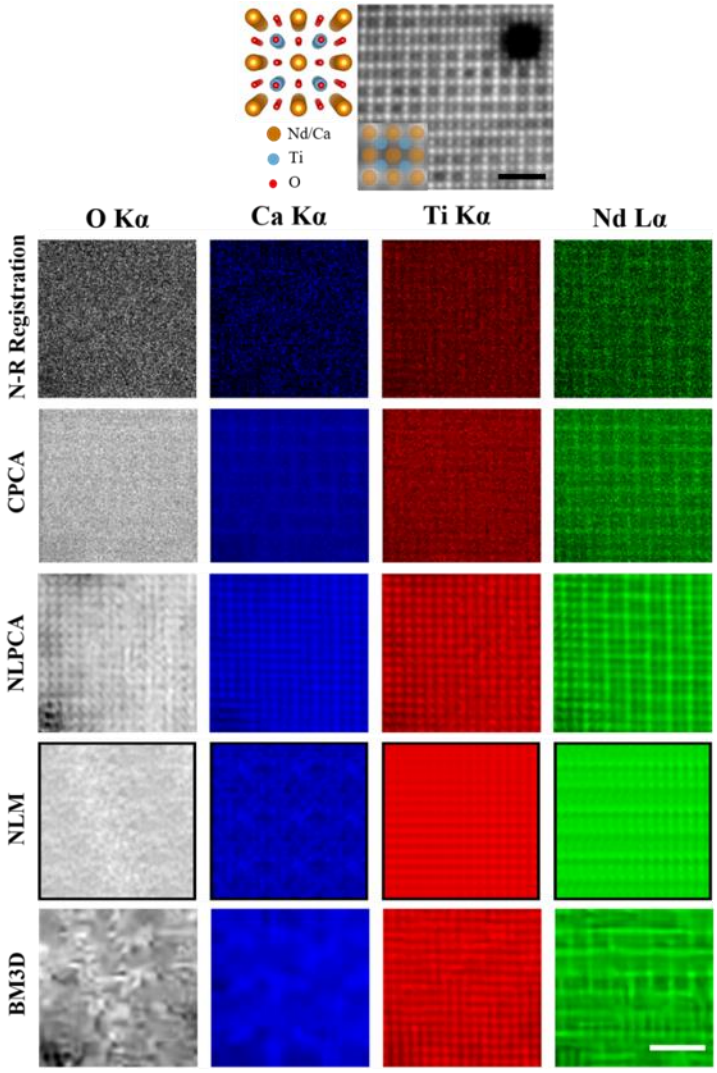

Figure 4.5: Comparison of CPCA, NLPCA, NLM, and BM3D applied to experimental data recorded along [001] Ca-stabilized $Nd_{2/3}TiO_3$. O Kα, Ca Kα, Ti Kα, Nd Lα maps denoised by CPCA, NLPCA, NLM, and BM3D are shown in the top row, second row, third row, and bottom row, respectively. The scale bar is 2 nm and applies to all panels.

We have also compared the denoising result from NLPCA to other denoising methods, including conventional weighted PCA[121],[122] (CPCA), NLM[113],[114], and BM3D[123]. The

denoising results from CPCA, as shown in the second row, shows almost no improvement in SNR. The denoising result from NLM is shown in the fourth row, which shows improved SNR and some atomic features in the Ti map, but the denoised result in Nd and Ca maps are unphysical and most likely to be denoising artifacts. The denoising results from BM3D, as shown in the last row, shows good denoising results in the Ti map and some atomic features in the Nd map, but is not able to recover the atomic features in Ca map. Also, the denoised Nd map is not clean with obvious artifacts. Overall, NLPCA has the best denoising performance on this noisy EDS dataset, and is the only one that is able to correctly recover the atomic structure in Ca, Ti, and Nd bands.

4.5 Conclusions

NLPCA has been applied to denoise atomic resolution spectrum image series collected in EDS to reduce the effects of Poisson noise and improve SNR for better visualization of the underlying microstructure. The denoised elemental maps are consistent with the structure of the material and have a significantly clearer atomic structure. Based on denoised simulated data, NLPCA with appropriate parameters can reconstruct the underlying atomic structure for several different microstructures. The comparison to other denoising methods also shows that NLPCA performs the best in terms of denoising performance and introduces the least artifact. The usage of NLPCA may enable EDS spectrum imaging experiments at a reduced dose and thus lower beam damage, extending the reach of the technique to lower concentrations and more beam sensitive materials.

5 Joint Denoising and Distortion Correction

5.1 Introduction

As discussed in section 3.1, being able to accurately determine the position of atoms in atomic resolution STEM could be beneficial on multiple research topics, and lots of efforts have been made to correct the distortions in STEM for better atom position detection. These distortion correction tools work well on STEM data with a high electron dose with a reproducible sub-pm precision result on the processed image.

Low dose STEM data is typically dominated by Poisson noise, and denoise methods are often applied to find the noiseless ground truth from noisy images to improve SNR and reveal the atomic structure. One widely adopted method is to apply filter masks in Fourier space[124],[125], with the assumption that high-resolution STEM data is sparse in Fourier space and occupies characteristic frequency. However, this could be problematic as it forces periodicity in the final processed image and would ignore non-periodic structures. More advent method applies non-local means[113] and block-matching and 3D-filtering (BM3D)[114] to denoise STEM image using redundancies within the image itself with Poisson noise statistic embedded in the method.

In this work, we present the application of joint denoising and distortion correction (JDDC), which is a Bayesian model-based method to do an “all at once” fitting on low dose atomic resolution HAADF STEM image series to directly resolve atom positions. A model is built to describe the noisy image series with atom position, distortions, and Poisson noise, then Bayesian approach is applied to look for the most probable set of atom positions and distortions that generate the noisy observation with Poisson noise introduced. Unlike the peak detection in

previous works, where a high-quality image with reduced distortion is generated in the intermediate step, JDDC directly fit the atom positions from the noisy image series. JDDC is especially helpful for noisy image series, where the pixel-wise matching between frames in the distortion correction becomes hard and unreliable.

5.2 Joint Denoising and Distortion Correction Method

The idea of JDDC is to build a model that contains all different factors in STEM image acquisition, include Poisson noise, sample drift, and flyback offset error, for a whole image series with multiple frames. A detailed description and derivation of the energy function can be found in ref[126]. The model used in this work is improved upon the previous model by using a collection of asymmetric Gaussian functions instead of symmetric Gaussian functions, which better describes the peak shape expected in STEM images, especially with imperfect optical systems. Corresponding to this change, we adapted an extra regularizer term on the peak shapes to prevent the model from using extended peaks to fit slicing artifacts in the image and results in much more stable fitting under low electron dose.

Assume we have an image series with N_3 frames, and the area of a region of interest (ROI) within each frame has pixel size N_1-by-N_2. Considered the drift between multiple frames, we only work on an ROI area that is presented in the whole image series, and the size of the ROI area is typically smaller than the size of a single raw frame. The fitted parameters can be divided into two parts: a collection of Gaussian functions $\mathbf{P} = \{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_{N3}\}$ which describes total L atom sites in the image series, and a collection of deformations $\mathbf{W}$ for each pixel (i,j) in each frame k. For the Gaussian function part, we assume that atomic columns in a HAADF STEM image can be described by a 2D asymmetric Gaussian function with constant background, such

that each atom position can be described by seven parameters: $\boldsymbol{P}_l = (x_{0,l}, y_{o,l}, A_l, c, x_w, y_w, I_0)$. Thus, the intensity on the coordinate (x, y) contributed by atom site *l* is

$$I_l(\mathrm{x},\mathrm{y} \mid \boldsymbol{P}_l) = I_0 + A_l exp[(\frac{1}{2(1-c^2)})(\left(\frac{x-x_{0,l}}{x_w}\right)^2 + \left(\frac{y-y_{0,l}}{y_w}\right)^2 + (\frac{2(x-x_{0l})(y-y_{0l})}{x_w y_w}))]. \quad [5]$$

To simplify the model and reduce the number of parameters being fitted, we set I_0, c, x_w, y_w to be the same for all the atom sites, while each atom site has its own intensity and position A_l, x_{0l}, y_{ol}.

The deformations describe the difference between an ideal regular scan grid and the actual scan grid. A single deformation vector $\boldsymbol{w}_{\mathrm{ijk}} = (\mathrm{w_{ijk,1}}, \mathrm{w_{ijk,2}})$ is assigned for each pixel (i, j) in each frame k, where the two components correspond to the deformation along the horizontal direction and vertical direction. The deformation field includes both random sample drift and the flyback offset error at the beginning of each scan row.

The parameters are fitted by minimizing the energy function, which is a sum of N_3 energy functions for N_3 frames

$$E = \sum_{k=1}^{N_3} E^k(\boldsymbol{P}, \boldsymbol{W^k}) + R_3(\boldsymbol{P})$$

$$= \sum_{k=1}^{N_3} \{d[I(x, y + \boldsymbol{W^k}|\boldsymbol{P}), G^k] + R_1(\boldsymbol{W^k}) + R_2(\boldsymbol{W^k})\} + R_3(\boldsymbol{P}) \quad [6]$$

In the function, $\boldsymbol{P}$ is the collection of Gaussian function parameters for each peak and is shared through the whole image series. $\boldsymbol{W^k}$ is the collection of all the deformation vectors in frame k. In the expanded formula, the first term on the right side measures the similarity between observation G^k and the ground truth with deformation. The R_1 term serves to fit the Brownian

motion for sample drifts and flyback offset errors. The R_2 term is a regularizer on the deformation, and R_3 term is a regularizer on peak positions and shapes. The R_2 and R_3 terms do not fit any specific distortion pattern in the image series but serves to prevent the model from unphysical fitting results.

The R_1 term on the right side that fits the sample drifts and flyback offset errors expands into

$$R_1(\boldsymbol{W}^k) = \frac{1}{2D}\left[\frac{|w^k{}_{11}|^2}{\Delta t} + \sum_{j=2}^{N_2}\left(\sum_{i=2}^{N_1}\frac{\left|w^k{}_{ij} - w^k{}_{(i-1)j}\right|^2}{\Delta t} + \frac{\left|w^k{}_{1j} - w^k{}_{N_1(j-1)}\right|^2}{\Delta T}\right)\right]. \quad [7]$$

The second term on the right side penalizes the deformation difference between two adjacent pixels within a row divided by the time interval Δt between the acquisition of two pixels. The third term on the right side penalizes the deformation difference between the start pixel of a row and the end pixel of the previous row. A large time interval $\Delta T = 1000\Delta t$ is used, which allows more flexibility, and the more significant flyback error can be fitted. D at the beginning is a balancing factor that controls the relative importance of the Brownian motion term.

The R_2 term on the right side of [6] is a regularizer penalizing the squared norm of the deformation with separate weights assigned for horizontal and vertical direction and expands into

$$R_2(\boldsymbol{W^k}) = \frac{1}{2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\nu_{hor}\left|w^k{}_{ij,1}\right|^2 + \nu_{vert}\left|w^k{}_{ij,2}\right|^2\right), \quad [8]$$

where $w^k{}_{ij,1}$ and $w^k{}_{ij,2}$ denote the horizontal and vertical component of the deformation on pixel (i, j) in frame k. This term is only related to the absolute deformation magnitude on each pixel, it does not penalize abrupt change in deformation between adjacent pixels like [7] does.

The usage of this term could prevent the model from using the Brownian motion to fit noise pixels.

The R_3 term is a regularizer on atom positions outside the region of interest and atom shapes, it expands into

$$R_3(\boldsymbol{P}) = \frac{\nu_{out}}{2} \sum_{p \in P_{out}} \left[\left(x_i - {x_i}^{ini}\right)^2 + \left(y_i - {y_i}^{ini}\right)^2 \right] + \nu_{shape} \left(\left| \frac{\sigma_y}{\sigma_x} \right| - 1 \right)^3. \quad [9]$$

The first term on the right side prevents the atoms outside the region of interest them from moving far away from initial positions $({x_i}^{ini}, {y_i}^{ini})$. These atom sites need to be considered as they have residual intensity inside the region of interest, but as they are not considered when fitting the sample drift, we simply require them to stay close to their initial position. The second term on the right side penalizes shape change from symmetric Gaussian to prevent JDDC from fitting peaks into highly extended Gaussian peaks to compensate slicing artifacts. The two terms are weighted by constant factors ν_{out} and ν_{shape}.

A full list of the tunable fitting parameters and the value we adapted is listed in Table 5.1. ν_{hor}, ν_{vert}, ν_{shape}, and D are the most important parameters for the final fitting result, and they are optimized by a global search over multiple parameters simultaneously. Low dose image series acquired from $SrTiO_3$ [100] sample with 2 frames and a total 8.56×10^2 $e^-/Å^2$ electron dose were used to search for optimal fitting parameters by looking for a set of parameters that generate the best fitting precision. The fitting precision is defined as the standard deviation between Sr atom sites in the fitted results along the fast scan and slow scan directions[100],[127]. We did not attempt to refine the fitting parameters for image series with a higher electron dose,

as the energy landscape gets smoother with higher electron dose, and JDDC is typically robust over a wider range of parameter selections.

Notation	Description	Value
ν_{hor}	Weight factor penalizing the magnitude of horizontal deformation.	1
ν_{vert}	Weight factor penalizing the magnitude of vertical deformation.	100
ν_{shape}	Weight factor penalizing the asymmetrical shape.	0.3
ν_{out}	Weight factor penalizing peak movement outside the region of interest.	0.05
D	Weight factor for Brownian motion term.	0.002
Δ_T/Δ_t	Time length ratio between Brownian motions for flyback offset errors and sample drifts.	1000

Table 5.1 List of all fitting parameters in JDDC and adapted values

When an image series is fed to the JDDC, the algorithm first performs a rigid registration to align multiple frames, then average to get a single frame with improved SNR. Atom site positions are initialized by a local-maxima searching on the smoothed image. A region of interest that presents in all frames inside the image series is then selected. Starting from the initialized atom position and an empty distortion field where distortions are zero everywhere, a trust region Newton method is used to minimize the energy function [6].

5.3 High Precision Atom Position Fitting at Low Dose

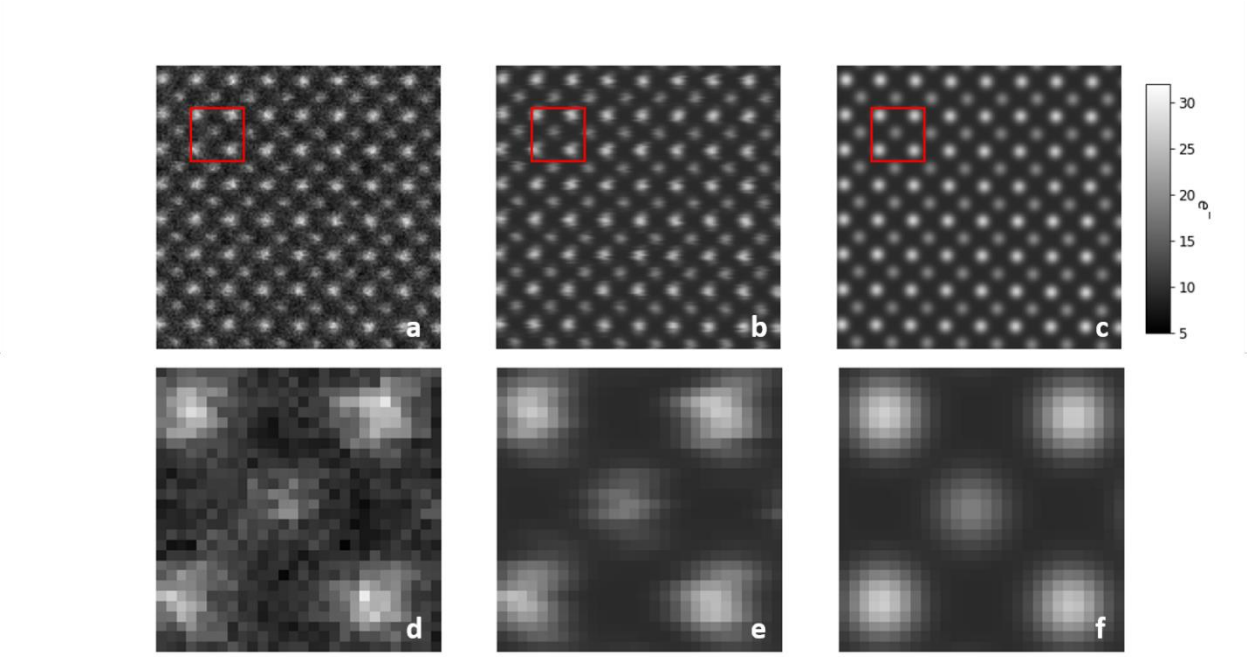

Figure 5.1 Single frame of (a) raw image series, (b) JDDC fitted peaks with distortions, and (c) JDDC fitted peaks. (d) – (f) are the zoomed-in patch from the single frame shown in (a) – (c).

The peak detection of JDDC has been tested on the HAADF-STEM image series collected from $SrTiO_3$ [001] single-crystal sample. The image series contains 10 frames of 256-by-256 px images, with 21.5 pm pixel size and 10 µs pixel dwell time. A single frame from the raw image series is shown in Figure 5.1 (a) with a zoomed-in patch from the region in the red box shown in Figure 5.1 (d). The atomic columns are clearly distorted in the raw data, with clear flagging artifacts from the scan distortions. The collection of 2D Gaussian functions with distortion fitted by JDDC is shown in Figure 5.1 (b) with a zoomed-in patch from the same area shown in Figure 5.1 (e). The fitted atomic sites with distortion added from JDDC closely resembles the distorted atomic sites in the raw data. The sample drift term from JDDC fitting shows zero mean drift during the acquisition of each pixel with a standard deviation of 1.5 pm. The scan offset term

from JDDC fitting shows zero mean offset for each row along both x and y direction, with 10.8 pm and 15.1 pm standard deviation along the x and y direction, respectively. The fitting result suggests that the probe offset at the beginning of each row is more serious along the slow scan direction, in other words, the skipping artifact should be more obvious than the flagging artifact.

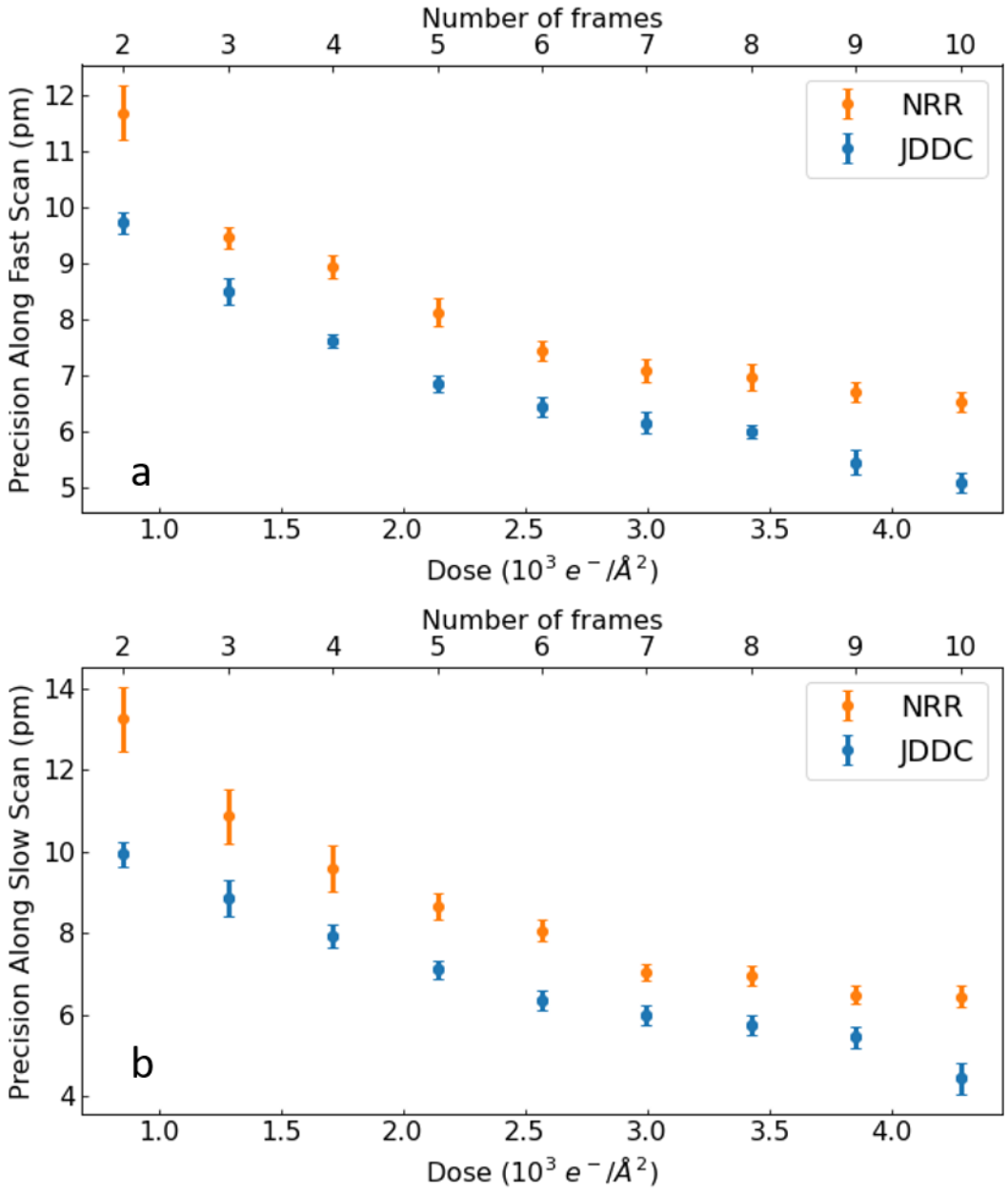

Figure 5.2 The precision of JDDC fitting results on low dose image series compared to NRR from (a) fast scan direction and (b) slow scan direction. The error bars come from the standard deviation of the mean from ten repetitions using different image frames.

To test the performance of JDDC at a low electron dose, 0.5 µs pixel dwell time was used to collect the HAADF image series, which is the shortest possible pixel dwell time without causing a significant, systematic scan distortion on our microscope[128]. Fitted atom positions from JDDC were directly used to calculate the precision with peaks from the edge 10 px excluded. The same ROI used in JDDC is used in the final averaged image from NRR to calculate the precisions from fitted peak positions. Details of determining peak positions from NRR results can be found in ref[100]. 0.5 µs pixel dwell time was used to collect the low dose image series to get the lowest possible electron dose, and the total dose was controlled by using a different number of frames under a fixed frame dose.

Figure 5.2 shows the precision of fitted atomic sites positions from JDDC at low electron dose benchmarked against the NRR method along the fast scan direction and the slow scan direction. JDDC shows improved precision from NRR over the whole range of tested electron dose less than 42.8×10^3 $e^-/Å^2$ and achieved better than 10 pm precision along both scan directions with 856 $e^-/Å^2$ electron dose. Such a low dose would cause minimum beam damage to most of the inorganic crystal materials but is still outside the useful range for biological materials or organic crystal.

5.4 Conclusions

A new model-based method called JDDC, which performs all-at-once fitting to detect atom sites and distortions from the HAADF-STEM image series, was tested on experimental image series collected from $SrTiO_3$ [100] single-crystal sample. JDDC fitting is shown to be able to

capture the correct drift pattern from noisy image series, and the fitted images with distortions match well with experimental images by eye. JDDC also shows improved precision compare to NRR at low dose image series with a total dose of less than 4.5×10^3 $e^-/Å^2$, and achieved better than 10 pm precision with 856 $e^-/Å^2$ electron dose. When applying JDDC to fit image series with fixed total electron dose, the best precision was achieved at a pixel dwell time of 2 μs/px, which largely eliminates the effect from detector response time while the image still captures most distortions. Overall, JDDC is shown to be an effective tool to directly identify atomic sites at relatively high precision on a fast scan, low dose image series.

6 4D STEM Data Analysis with Convolutional Neural Networks

6.1 The Motivation for Using Convolutional Neural Networks

The full 2D CBED patterns in an atomic-resolution 4D STEM data set contain a wealth of information about the specimen. The information is difficult to access directly, partially due to the data complexity and volume, and partially due to uncertainty in the precise, sub-Ångstrom positioning of the probe on the sample. Position-averaged CBED (PACBED)[129] has been the most common method to reduce position uncertainty by giving up the spatial resolution. PACBED has been used to determine local STEM sample thickness[129], ferroelectric polarity[130], octahedral tilt in oxide superlattices[131], and sample composition with a well-defined alloy system[132].

The typical way of PACBED pattern analysis compares the experimentally collected PACBED patterns to a simulated library covering all possibilities. One way to make the comparison is by using the judgment of human eyes, which is slow, limited to resources, and

prone to bias. An alternative way is to perform least square fitting (LSF) [133],[134] between experiment data and all the possible ones in the simulation library, which is more objective but requires a full match of all the experimental parameters. Parameters that need to be matched include the size of the zero beams, the position of the zero beams, and the ellipticity of the beam, which requires a very large simulated library or a library that is repeatedly recalculated to fit the experimental parameter change. Neither of them is well-suited to the massive data that is collected in 4D STEM.

As an alternative to the LSF, we propose to do this matching using CNNs, which is a supervised deep-learning tool developed for computer vision and image recognization[135]-[136], mostly for image classification. CNN has the advantage of extracting characteristic features from input images and using the features for image recognization. It does not require a full match between the input image and the image it was trained on, which solves the problem of full matching mentioned above. CNNs have been applied to solve a variety of materials science problems. As just a few examples: Ziadinov *et al.* .[137] trained a CNN to locate atom sites and identify defects from noisy atomic-resolution STEM images. Xie *et al.* .[138] built CNN on top of element and bonding information in perovskite structure and trained it to predicted materials properties. Li *et al.* .[139] trained CNN to identify dislocation loops in low-resolution diffraction-contrast TEM images at an accuracy comparable to a human being.

Pioneered by Xu and LeBeau, CNNs have been applied to analyze PACBED patterns[140]. They studied PACBED patterns acquired from 60 nm side length region of a $SrTiO_3$ [001] single crystal sample, and trained a series of different CNNs to analyze different aspects of the sample, including pattern size, center, rotation, local thickness, and off-zone axis tilt. In this work, we build on their work in three different ways. First, we target the PACBED

patterns averaged from 4D STEM dataset that covers just one atomic column, instead of multiple unit cells, resulting in a much higher spatial resolution. The schematic of this sub-unit cell PACBED generation and thickness measurement is illustrated in Figure 6.1. Second, we use a single CNN, instead of a combination of several CNNs, to predict the property that we are interested in while ignoring unrelated properties. Third, regression-CNN (R-CNN) is tested besides the conventional classification CNN, which generates a continuous number as output instead of separate classes. R-CNN has been shown to be better suited for many materials science problem and has higher prediction accuracy when predicting the local sample thickness from PACBED patterns.

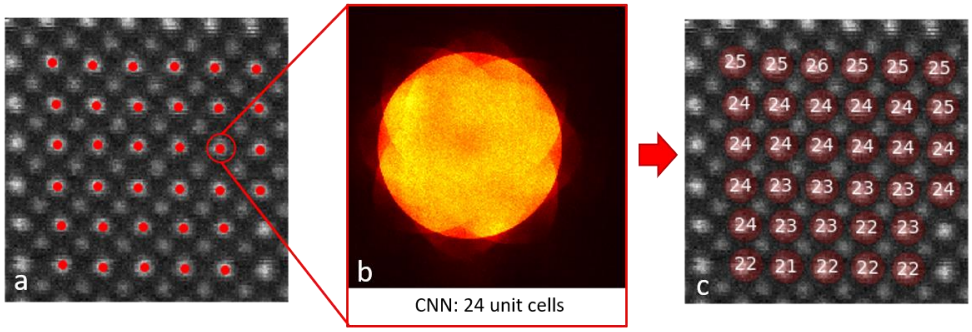

Figure 6.1 Schematic of sub-UC resolution thickness prediction, a) simultaneously acquired HAADF image with peaks marked b) example of Sr-PACBED pattern and CNN estimated thickness c) CNN derives thickness map overlapped on HAADF image

6.2 CNN Structure and Training

The structure of CNN can be divided into two parts, a convolutional base and a top fully connected layer for decision making. The convolutional base uses multiple layers of convolutional kernels to extract features in the input image, and result in a map of high-level

features. Then the top fully connected layers use the detected high-level features to make a final decision about the input image, either the class the image belongs to or a single continuous number that describes the image.

Two top fully connected layers structure has been studied, lead to classification CNN (C-CNN) and R-CNN. The direct output from the fully connected layers a C-CNN is a single probability for each class the network is trained on, then the class with the highest probability will be used as the final output and to classify the input image. For R-CNN, the fully connected layers directly generate a continuous number as the output to describe the input image.

For the materials science-related properties, some of them can only be predicted using a C-CNN structure, as they cannot be expressed by a continuous number, such as lattice symmetry and different types of atomic sites. On the other hand, many properties we are interested in can be described using a continuous number and can be predicted using an R-CNN, such as sample thickness, off-zone axis tilt, doping fraction, lattice constant/strain. All these properties can also be predicted using C-CNN by dividing the continuous number into many classes.

For the properties that can be expressed using a continuous number, C-CNN is not the best fit as it assumes the classes cannot be arranged in any meaningful order. Practically speaking, when training a C-CNN, the same penalty will be applied as long as the predicted class is not the right class, no matter how far the wrong prediction is from the correct value. For R-CNN, the penalty applied depends on the difference between the prediction and the desired value, which better fits the prediction of a continuous number. However, so far, there are much fewer studies on R-CNN, partially because real-life problems are mostly classification problems, and also R-CNN is much harder to train as it requires the model to learn not only the difference between different images but also the overall changing trend across different images. In this work, we

trained both C-CNN and R-CNN to predict local sample thickness from a single atomic site using the sub-unit cell PACBED pattern generated from atomic-resolution 4D STEM data. We trained C-CNN to estimate sample thickness between 0-100 nm with 2 nm estimation accuracy and trained R-CNN to estimate sample thickness between 0-35 nm. A smaller estimation range is used for R-CNN as it is harder to train, and it is more challenging to train it to learn a wide variety of PACBED appearance.

A good training data is the key to train an effective CNN model, and a good training dataset should satisfy two aspects: correctly labeled, and covers all the possibilities within each class. As multislice simulated data is used as the training data, and the simulation parameters are used as the labels, the data is guaranteed to be correctly labeled. The challenges become how to generate simulated data that is realistic and how to cover all the possibilities in the simulation.

Parameter name	Parameter value
Model size	49 Å × 49 Å × 1000 Å
Output thickness increment	5 unit cells
Probe step size	0.175 Å
Potential space sampling	0.06 Å
Frozen phonon configurations	20
Root mean square thermal displacement for Sr, Ti, O	Sr: 0.0887Å;Ti: 0.0746 Å O: 0.0947 Å

Table 6.1 Multislice simulation parameters used for CNN training data.

The multislice simulation parameters are listed in Table 6.1. To generate realistic data, we have included the effect of thermal vibration, effective source size, and Poisson noise to the

simulated data. Effective source size has been previously studied for a quantitative match of HAADF-STEM images[69] with different methods proposed to measure the source size in experiments[141],[142]. This factor has also been implemented in different STEM simulation packages[143],[70]. We found that the effective source size is crucial for the appearance of the sub-unit cell PACBED pattern and has to be considered for a visually good match. Figure 6.2 shows a panel of simulated PACBED patterns under different combinations of effective source size and integration radii. The appearance of the PACBED patterns changes significantly with different source size when the integration radius is smaller than the size of a single unit cell. And by matching the simulated PACBED patterns to the experiment PACBED pattern, we have determined that a 110 pm effective source size should be applied to all the simulated 4D STEM data. Besides the effective source size, random, Poisson-distributed noise was added to each CBED pattern after rescaling every single CBED pattern to a maximum value of 5, which is close to the actual intensity in the number of electrons we collected in the experiment.

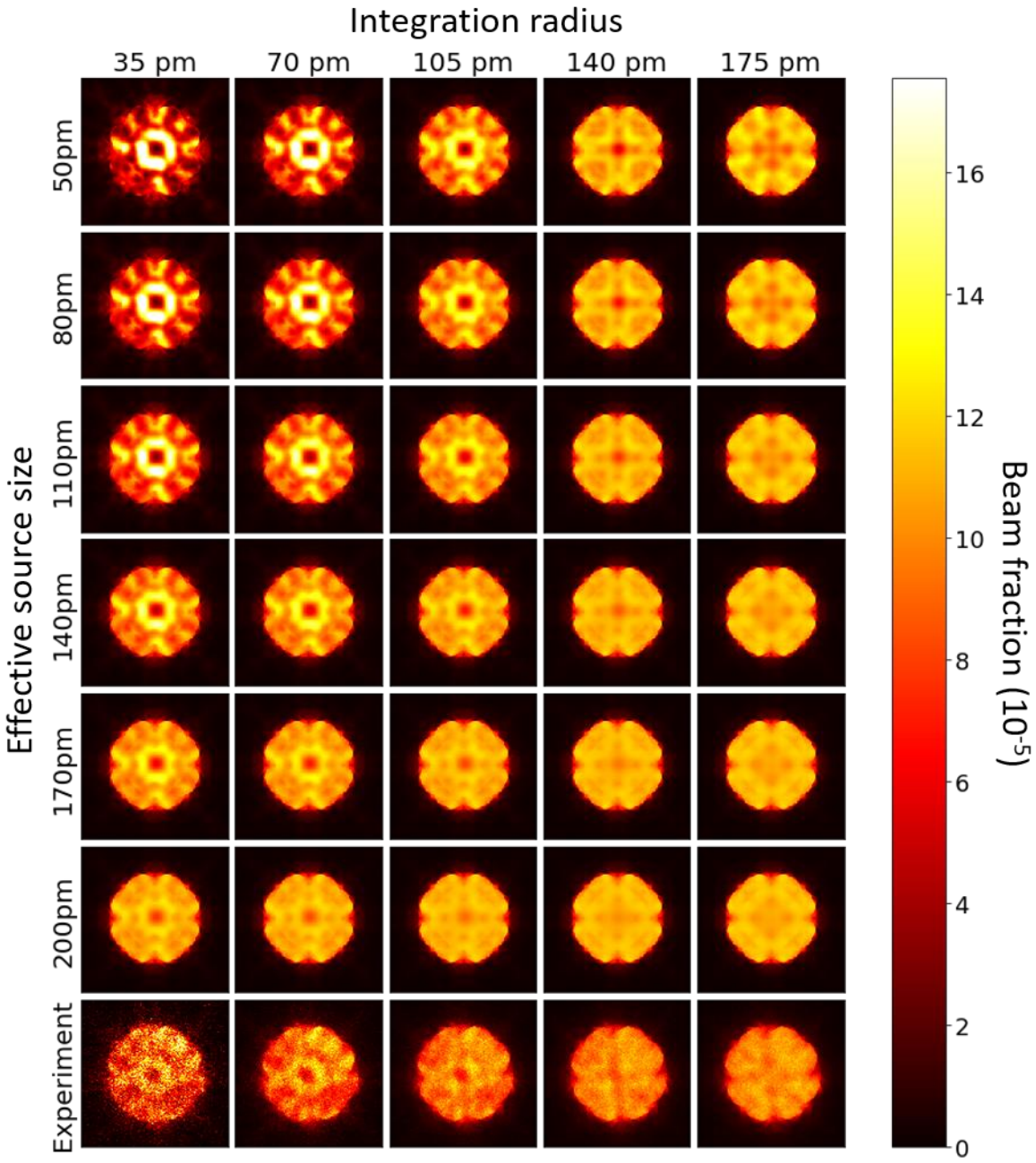

Figure 6.2 Simulated PACBED under different combinations of effective source size and integration radius for $SrTiO_3$ [100] with a thickness of 23 nm. The last row shows an example of experimental data acquired from 23 nm STO [100] sample and integrated by different radii. All PACBED patterns are in the unit of beam fraction.

To cover all possible situations in a real experiment, we have simulated 4D STEM data as a function of thickness for different off-zone axis tilt between 0 and 3 mrad. Considered that scan distortion and sample drift make it extremely difficult to precisely determine the position of a set of CBED patterns in the experiments, 36 different spatial integration configurations were used, including 9 slightly different integration center positions and 4 different integration radii (105.0 pm, 122.5 pm, 140.0 pm, 157.5 pm). In the end, each PACBED pattern passed through the data augmentation process before being fed to the model, which applies random distortion effect, including feature-wise centering, vertical and horizontal flips, vertical and horizontal translations, rotations, shears, and zooms.

In the training step, we adopted the transfer learning scheme instead of training the network from blank. In transfer learning[144], both the structure and some weights are adapted from a pre-trained VGG16 network using the (large) ImageNet dataset. The pre-trained weights on the convolutional layers have already effectively learned features that are important in image classification tasks. Although these adapted weights and convolutional kernels are trained on natural images instead of grayscale diffraction patterns, some basic features are in common such as corners, edges, and arcs, making it a good start point to train the network compared to training from scratch.

The training hyper-parameters are listed in Table 6.2 for two CNNs, with parameters optimized by hand one at a time. CNNs are known to be sensitive to the parameters used during training, and there is no effective way to select these parameters so far. Several trials were performed for each set of parameters to determine the most successful parameter values, which is necessary due to the stochastic processes during training, include image augmentation, random seeds for minimization problems, *etc.*

Cross-validation was used to monitor CNN's prediction accuracy during training. The whole simulation dataset was randomly divided into 80% training data and 20% validation data[145]. Within each training epoch, CNN was trained with the back-propagation method[146] using training data, then the trained network was tested on validation data to get a validation accuracy. CNN training was stopped if no improvement in validation accuracy is observed within a tolerance determined by the patience parameter. The training of C-CNN ended with 85% validation accuracy, i.e., 85% of the prediction fall into the right category, and the R-CNN ended with an RMSE of 1 nm.

Parameter name	C-CNN	R-CNN
Batch size[147]	32	
Fully connected layer size	256	
Dropout rate[148]	0.3	
Decay	1e-6	
Learning rate[149]	0.001	0.0005
Momentum[150]	0.9	0.9
Maximum learning epochs	20 + 50	20 + 50
Training library size	11,232	3,888
Loss function	Categorial accuracy	MSE
Optimizer	Stochastic gradient descent (SGD)	

Table 6.2 Training parameters for two CNNs

6.3 Thickness Estimation on Experimental Data

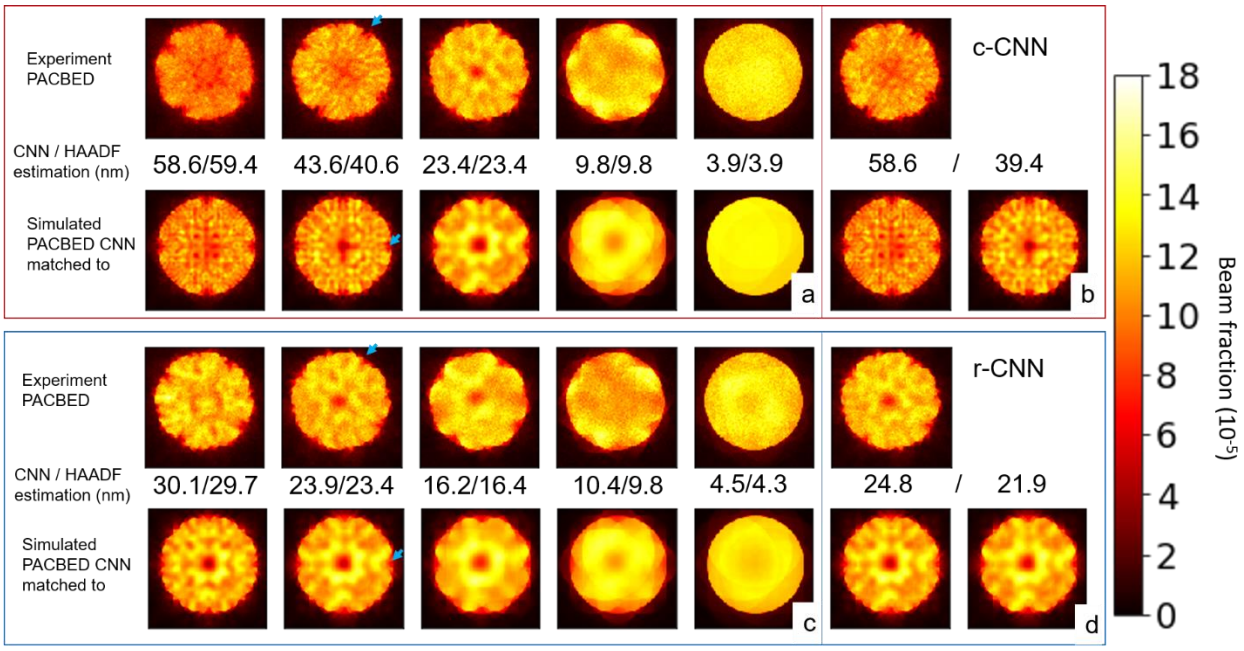

Figure 6.1 Examples of matching between experiment Sr-PACBED and simulated Sr-PACBED from CNN predicted thickness for a,b) C-CNN, and c,d) R-CNN. a) and c) show examples of good agreement between HAADF estimation and CNN measurement, b) and d) show examples of mismatch between two measurements. The sample is [100] single crystal $SrTiO_3$.

Figure 6.3 shows examples of experimental Sr-column PACBED patterns, thickness estimated by CNNs and HAADF intensity, and the corresponding simulated PACBED patterns from the thickness estimated by CNNs. The PACBED pairs shown in Figure 6.3 (a) and (c) covers a wide range of different sample thicknesses, and the experimental PACBED pattern is visually similar to the simulated PACBED. All the experimental PACBED patterns have a fixed rotation from the simulation, as marked by the blue arrows. CNNs are not confused by the rotation for the thickness measurement, as rotation was part of the image augmentation at the training stage. However, this rotation could be critical for LSF methods. Figure 6.3 (b) and (d) show two examples where CNN and HAADF make different estimations. The CNNs over-

estimate the sample thickness in both cases. In both cases, the difference between two simulated PACBED patterns are small and can be hard to distinguish by human eyes. Overall, the matching between experiment and simulation suggests that the trained CNNs are making the PACBED patterns matching at a level that is similar to human eyes.

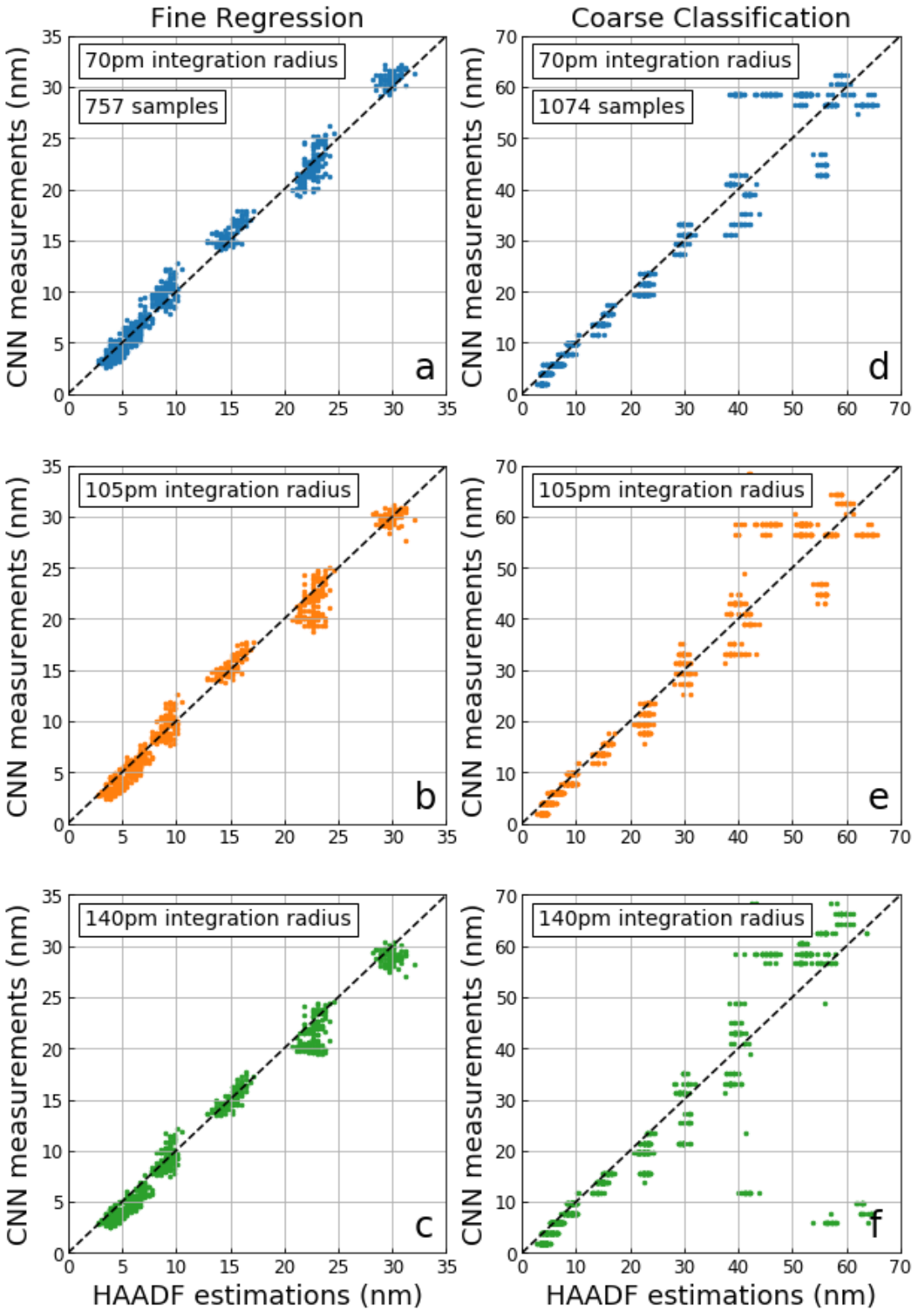
Fine Regression
Coarse Classification
70pm integration radius
757 samples
1074 samples
105pm integration radius
140pm integration radius
CNN measurements (nm)
HAADF estimations (nm)
a
b
c
d
e
f

Figure 6.4 CNN with coarse and fine step size measurements compared to quantitative STEM measurements of thickness from the same atomic columns. *r* is the integration radius about the atomic column position in the CBED data set

Figure 6.4 shows the CNN thickness measurement as a function of the HAADF measurement for all the Sr columns in a total of 34 4D STEM datasets, of which 25 have a thickness within 35 nm and can be used to test the R-CNN. Three different integration radii were used when generating the experimental Sr-PACBED patterns, where a larger integration radius improves the SNR by averaging over more CBED patterns, but potentially brings in more information that is not related to the Sr atomic column. Quantified performance of two CNNs in terms of RMSE and fraction of estimation that fall within ± 1 nm and 2 nm is shown in Table 6.3.

Integration radius (pm)	RMSE (nm)	±1nm	±2nm		Integration radius (pm)	RMSE (nm)	±2nm
70	1.05	68%	94%		70	5.16	66%
105	1.09	70%	94%		105	4.44	61%
140	1.14	65%	93%		140	13.10	55%

Table 6.3. Performance of coarse classification CNN under different integration radii for fine R-CNN (left) and coarse C-CNN (right)

For R-CNN, the thickness measurement works well throughout the whole prediction range, with about 1 nm RMSE between the CNN measurement and HAADF measurement. For C-CNN, the measurement works well on thin samples, and some of the measurements on thick samples get deviated from HAADF measurement, which also dominates the RMSE calculated for C-CNN. Generally speaking, feature-based measurement is more challenging on thick samples, as

the PACBED patterns show denser features than thin samples, and the features have smaller variance with different thicknesses.

6.4 Conclusions

Two CNNs were trained using multislice simulated libraries to estimate local sample thickness at an atomic resolution on the $SrTiO_3$ [100] single crystal sample using Sr-column PACBED patterns generated from 4D STEM data. Both C-CNN and R-CNN are explored for different estimation ranges. 70% and 94% of r-CNN estimations fall within ±1 nm and ±2 nm of the corresponding HAADF estimations, 66% of c-CNN estimations fall within ±2 nm of the HAADF estimations. The CNN training scheme described in this work can be easily adapted to estimate other aspects of the samples with properly generated training data and the correct network architecture. R-CNN, which shows high accuracy on fine estimation, can be helpful for the measurement of local misorientation, octahedral tilt, and doping fraction. C-CNN, which can be applied to the classification type of task and can recognize widely different patterns, can be helpful in identifying different types of lattice symmetry, atom, and defects.

7 Denoising Atomic Resolution 4D STEM Data with Tensor Singular Value Decomposition

7.1 Introduction to Different Types of Denoising Methods

Denoising 4D STEM data is challenging for its massive size[151] and complex noise distribution. A 4D STEM scan with 16-bit 512-by-512 pixel CBED patterns and 100-by-100 beam positions occupies 5 GB, which is challenging, and the recent advent of ultrafast detectors could possibly lead to data size that is multiple times of the current size. The multiple GB data

size imposes a serious challenge for denoising methods to be both time- and memory-efficient. The noise pattern on the collected CBED pattern is also complex and deviates from the ideal Poisson noise[28],[152]. A back-thinned active pixel type detector generates a signal that is proportional to the energy deposited into the active layer by a high-energy electron, which obeys a Landau distribution[153],[154]. The trace of the high energy electron makes it possible to generate signal across multiple adjacent pixels randomly, which further complicate the noise pattern. Altogether, it becomes challenging to directly recover the ground truth from a combination of all the noise sources mentioned above.

One common approach to denoise data is to perform dimensionality reduction (sometimes referred to as rank reduction) on the noisy data. By treating the noisy data as a group of vectors in a high-dimensional space, dimensionality reduction denoise the data under the assumption that the noiseless ground truth lies in a subspace with lower dimensionality. The denoising procedure typically involves finding a set of new basis vectors for the subspace and use the set of new basis vectors to describe the noisy data. Similar ideas have also been used for spectral unmixing[155],[110] to gain a better understanding of the data. With different constraints to the basis vectors, different denoising methods have been proposed. The most popular dimensionality reduction method in the electron microscopy field is PCA, which finds a set of orthogonal basis vectors to describe the noisy data without significant information loss[110],[156], with adapted targeting Poisson noise distribution, namely Poisson PCA or exponential PCA[157]. PCA has been widely applied to denoise EELS and EDS data collected in STEM, with multiple implementations including DigitalMicrograph plugin[158], temDM[159], and Hyperspy[160]. A less common method is independent component analysis (ICA)[161], which seeks basis vectors without the orthogonal constraint. ICA has also been used to denoise data which contains

additive noise that is independent of the singal[162]. Another dimensionality reduction method is non-negative matrix factorization (NMF)[163], which can only be applied to non-negative data and applies non-negative constraint to the basis vectors and their corresponding coefficients. NMF is useful when looking for basis vectors that have certain physical meanings and thus required to be non-negative, and has been applied to both denoising[164],[165],[166] and data unmixing[167]. These dimensionality reduction methods can only be directly applied to 2D matrices, and data with higher dimensions will have to be unfolded to 2D before being processed, which leads to possible loss of structure and information.

Another category of state-of-the-art denoising methods is the non-local methods[168], which explores the self-similarity within the data and combines information from similar regions to remove random noise. These methods first divide the whole data into small patches, often overlap ones, then group similar patches together and denoise each group of similar patches. Being applied to electron microscopy, Mevenkamp et al. applied non-local means[113] and adapted BM3D[114] to denoise a single frame of HAADF-STEM image corrupted by Poisson noise. Yankovich et al. applied NLPCA, which is optimized for extremely noisy data and extended to 3D data, to denoise hyperspectral STEM EDS data cube[169],[170]. Maggioni et al. proposed the BM4D method[115] as an extension of BM3D to denoise hyperspectral data, with a modified version targeting time sequences[171]. As a result of dividing the data into small patches, these methods are typically memory-intensive and occupy memory tens or hundreds of times larger than the raw data. And there is no straightforward application of non-local methods for data unmixing.

Tensor SVD is a low-rank denoising method for high-order data in three or more dimensions.[172],[173],[174],[175] Here, we apply the tensor SVD method of Zhang and

Xia[176] to denoise atomic-resolution STEM data, focusing on 4D STEM but also testing EDS spectrum images. Compare to matrix-based rank reduction methods, tensor SVD can maintain the full structure of data, giving it the potential to exploit, for example, structure in both spatial dimensions. Compare to non-local methods, tensor SVD is much more memory-efficient and time-efficient as it directly works on the raw data without dividing it into patches. Based on our implementation, single-digit GB 4D STEM data can be denoised in a few minutes on a desktop computer with moderate computing power.

7.2 Tensor Singular Value Decomposition

The concept of the tensor SVD method is shown in figure 7.1. In our case of atomic-resolution 4D STEM data, we first unfold the two dimensions in reciprocal space with less redundancy into one single dimension and get 3D tensors from the 4D data, as shown in figure 7.1 (a). As a result of the repeating structure, the ground truth $\boldsymbol{X}$ is a tensor with Tucker low-rank[177] (r_1, r_2,r_3), which means it can be decomposed into a product of core tensor $\boldsymbol{S}$ with reduced size, and three unitary matrices:

$$\boldsymbol{X} = \boldsymbol{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}, \boldsymbol{S} \in \mathbb{R}^{r_1 \times r_2 \times r_3}, U^{(k)} \in \mathbb{O}^{p_k \times r_k} , \qquad [10]$$

as shown in Figure 1(d). Here, the tensor-matrix product is defined as:

$$\boldsymbol{X}_{ijk} = \sum_{a,b,c} \boldsymbol{S}_{abc} U^{(1)}_{ia} U^{(2)}_{ib} U^{(3)}_{ic} \qquad [11].$$

The problem of finding the low-rank decomposition shown on the right side of eq [10] from the noisy observation $\boldsymbol{Y}$ is what we called a tensor SVD problem.

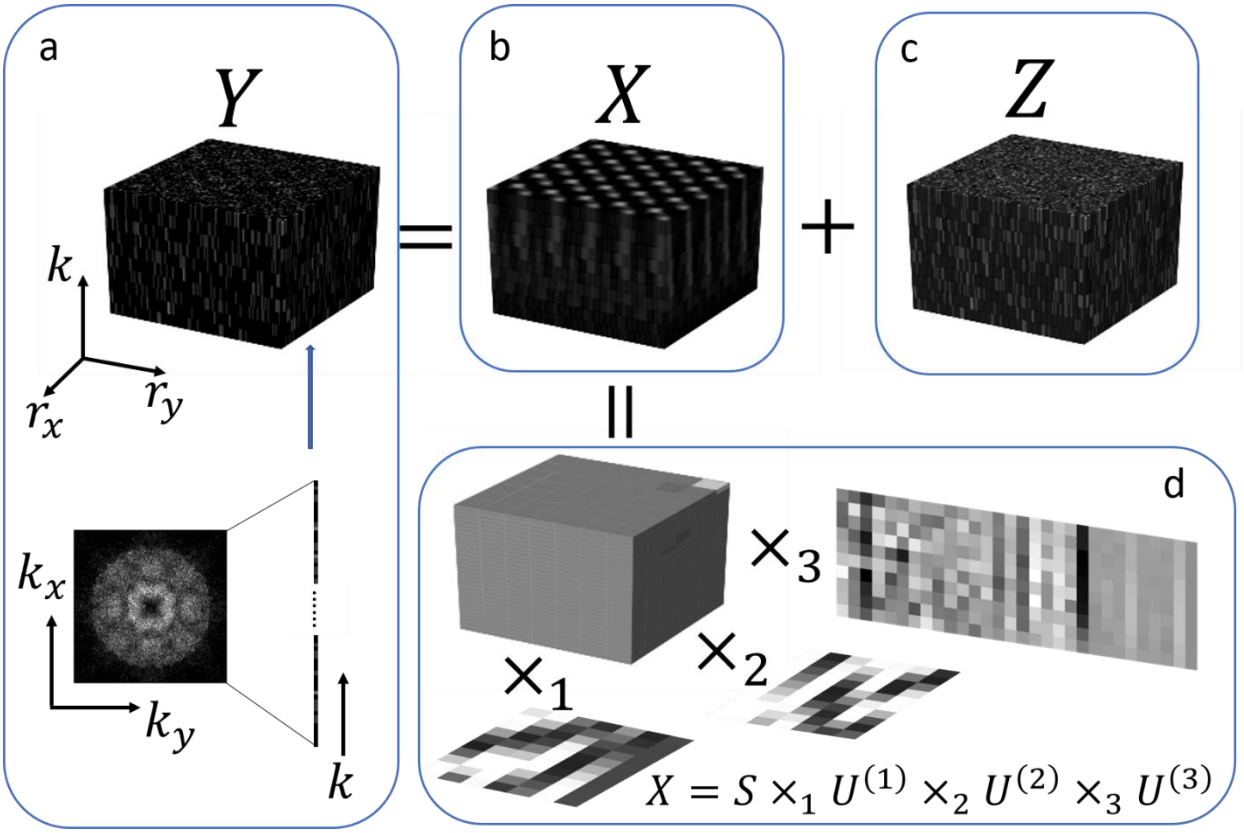

Figure 7.1 Illustration of the tensor SVD problem. (a) noisy observation tensor ***Y*** constructed from CBED patterns by unfolding the reciprocal space into one single dimension, (b) noiseless ground truth tensor ***X*** with low rank, (c) additive noise tensor ***Z***, (d) Tucker decomposition of the low-rank tensor into a core tensor and three unitary matrices.

To solve tensor SVD problems, unlike their counterparts in 2D matrices, the best rank-(r_1, r_2, r_3) approximation cannot be obtained from the truncation of a full decomposition[178], and the ranks have to be determined first before solving the problem. In this work, we set the denoising rank along each dimension to be the number of principal components derived from 2D matrices formed by unfolding the tensor along that dimension. For example, the unfolded 2D matrix from 3D tensor with size $p_1 \times p_2 \times p_3$ along the first dimension gives a 2D matrix with size $p_1 \times (p_2 p_3)$ with the elements arranged as

$$[\mathcal{M}_1(\boldsymbol{Y})]_{i,(j-1)p_3+k} = \boldsymbol{Y}_{ijk},\ \boldsymbol{Y} \in \mathbb{R}^{p_1 \times p_2 \times p_3}, \mathcal{M}_1(\boldsymbol{Y}) \in \mathbb{R}^{p_1 \times p_2 p_3}. \quad [12]$$

The number of principal components on 2D matrices is determined by calculating the eigenvalues and make a scree plot, then find the elbow point[179],[180],[181],[112]. This rank determination method is justified using simulation data, where the denoising result using the set of ranks determined from scree plots are very close to the denoising ranks that best recovers the ground truth.

To find the best rank -(r_1, r_2, r_3) approximation from the noisy input, we applied the higher-order orthogonal iteration (HOOI) method, which was first proposed in ref[178]. In our implementation, the HOOI method seeks to minimize the difference between noisy data and the low-rank tensor in Frobenius norm,

$$\widehat{\boldsymbol{X}} = argmin_{rank(\boldsymbol{X})=r_1,r_2,r_3} \|\boldsymbol{Y} - \boldsymbol{X}\|_{\mathrm{F}}, \quad [13]$$

by iteratively refines the unitary matrix *b* along each dimension with 2D matrix SVD. Additional details of this implementation can be found in ref[176]. The Frobenius norm measurement does not assume any specific noise statistics, and Poisson's negative log-likelihood could be used instead as a similarity measurement for data with pure Poisson statistics[33],[182]. This implementation also assumes that the low-rank ground truth represents a large fraction of the total variance of the noisy data, i.e., the signal should be significant and the noise should be trivial. This assumption may not be satisfied if the noise contributes more significantly to the total variance, for example, if the data contains some high-intensity noise points.

7.3 Denoising Performance

Tensor SVD's performance is first tested on simulated 4D STEM data with known ground truth. Figure 7.2 shows the comparison between noisy data, tensor SVD denoised data, ground

truth, and the difference between the ground truth and denoised data in terms of reconstructed ADF image and a single CBED pattern on simulated $SrTiO_3$ [001] structure under 1000 FPS frame rate. 8, 8, and 30 were used as the denoising ranks for x, y, and k dimensions for the data. In both cases, tensor SVD is able to almost perfectly recover the ground truth, with small residual and mostly noise shown in figure 7.2 (g) and (h). Figure 7.2 (i) shows the data quality measured in PSNR for pairs of noisy and denoised data, and tensor SVD is shown to be able to improve the data quality by about 20 dB under different input noise levels.

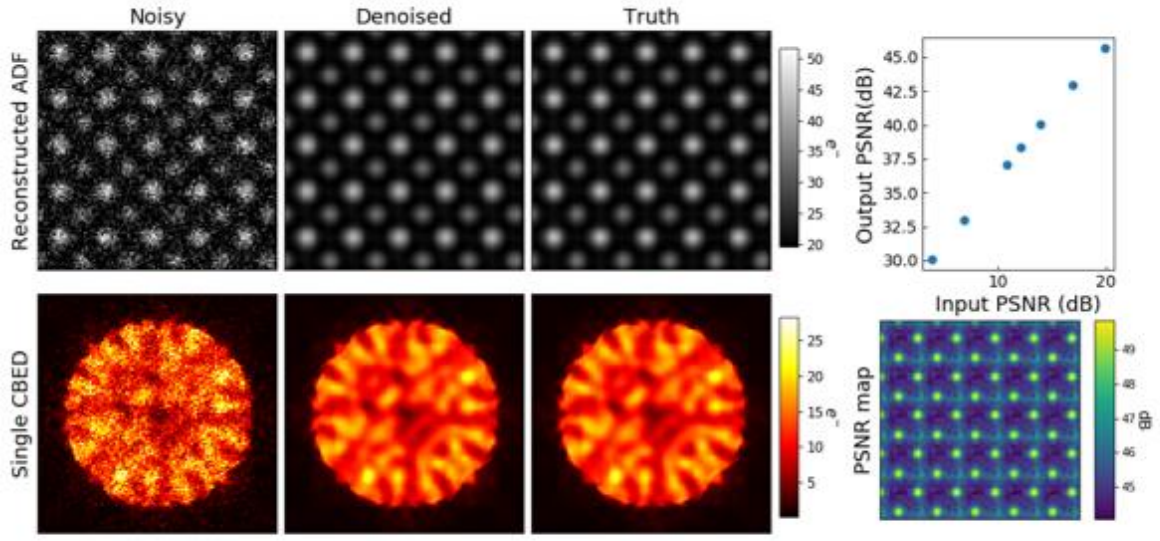

Figure 7.2. Denoising 4D STEM simulated from a perfect STO. crystal: virtual ADF image and a single CBED pattern pairs from (a)-(b) simulated 1000 FPS noisy data, (c)-(d) tensor SVD denoised data, and (e)-(f) ground truth. (g) Output PSNR from SVD denoised data vs. input PSNR from noisy data calculated under different noise levels. (h) Spatial distribution of denoising PSNR for 1000 FPS data.

A more complicated Si edge dislocation model with less periodicity was used to test the tensor SVD method, and the denoising results are shown in Figure 7.3. The denoising ranks for this structure was 32, 32, and 180 for r_x, r_y, and the unfolded k dimension, significantly higher ranks are used for the real space dimensions as a result of the irregular structure around the dislocation core. The comparison between reconstructed ADF images and single CBED patterns extracted from periodic Si lattice and the dislocation core shows almost no visible difference between the denoised image and the ground truth. From the difference image shown in figure 7.3 (j)-(l) and the PSNR map shown in figure 7.3 (n), the denoised CBED pattern around the dislocation core has about 7 dB lower PSNR than the periodic lattice structure, and a higher pixel-wise difference between the denoised image and truth image. The PSNR measured from noisy-denoised pairs from different noise levels shows that tensor SVD is able to improve the PSNR by about 15 dB on the data with irregular structures.

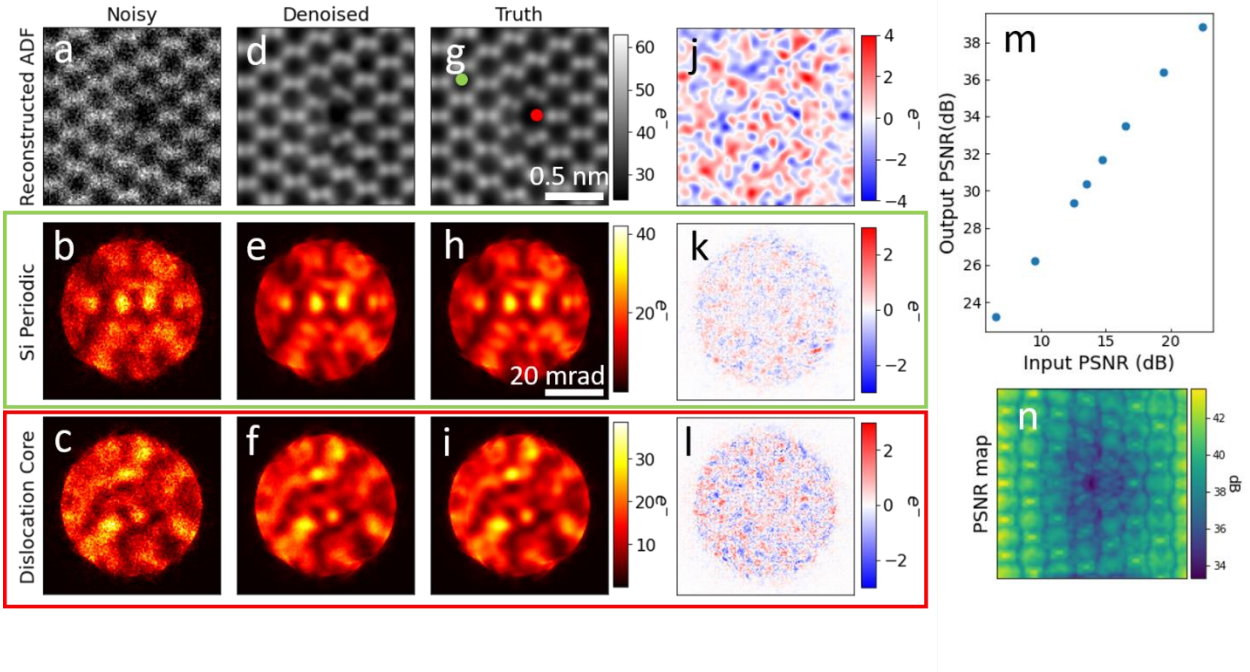

Figure 7.3. Denoising simulated 4D STEM data from a Si dislocation core: The virtual ADF image, a single CBED pattern from the periodic region of the Si structure at the position of the green dot in (g), and a single CBED pattern from the dislocation core region at the position marked by the red dot in (g) extracted from: (a)-(c) simulated 1000 FPS noisy data, (d)-(f) tensor SVD denoised data, and (g)-(i) ground truth. (j) Output PSNR from SVD denoised data vs. input PSNR from noisy data at different noise levels. (k) Spatial distribution of denoising PSNR for 1000 FPS data.

With the denoising performance confirmed on simulated data, tensor SVD was applied to denoise experimentally collected 4D STEM data without known ground truth. A $SrTiO_3$ [001] single crystal sample and a LiZnSb [$11\bar{2}0$] hexagonal Heusler alloy thin film grown on top of GaSb [110] substrate was used to collect 4D STEM data and test tensor SVD. Figure 7.4 shows the comparison between noisy data and denoised data on the two samples in terms of reconstructed ADF image and single CBED patterns. For both samples, the SNR gets much improved after tensor SVD, and the denoised ADF image and CBED patterns show much less random noise point. The denoised CBED pattern from the $SrTiO_3$ sample shows clear arc features in the dark field region from Bragg diffractions, while the denoised CBED pattern from the thicker LiZnSb/GaSb sample shows signs of Kikuchi bands.

The tensor ranks for STO data are determined to be 30, 32, and 180 for r_x, r_y, and k, respectively, and 35, 38, and 180 for r_x, r_y, and k for LiZnSb/GaSb interface. The ranks determined from scree plots are very similar for the two samples, despite a clear difference in the complexity of material structures, confirming that the ranks are heavily influenced by non-sample factors like scan distortions. It took 575.0 seconds for tensor SVD to denoise the 2.7 GB LiZnSb/GaSb data and 538.9 seconds for the 1.6 GB $SrTiO_3$ data on a personal desktop. The processing times are

mostly determined by the denoising ranks, so the processing times for the two datasets were similar despite the $SrTiO_3$ data is about 30% smaller than the LiZnSb data.

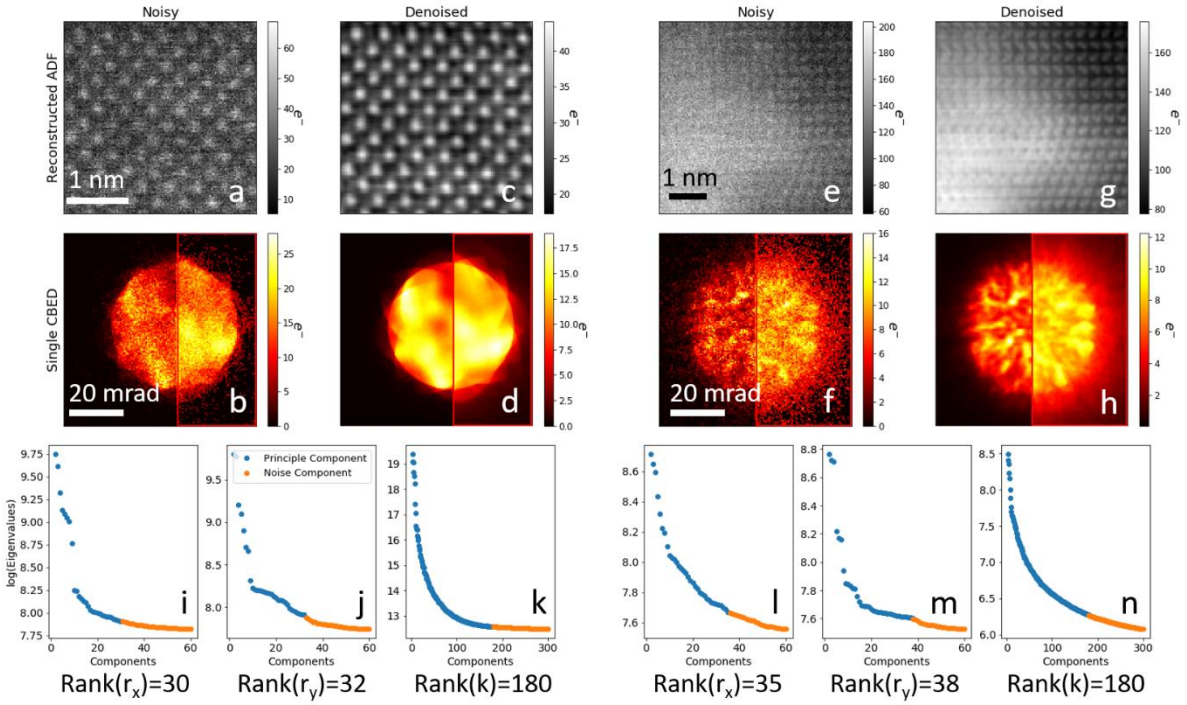

Figure 7.4 Denoising experimental 4D STEM data: (a) – (h) Virtual ADF images and single CBED patterns pairs. The right half of each CBED is shown as the square root of intensity to emphasize weak features. The color scales are for the linear left half of each image. (a) – (d) Data from STO [100]: (a), (b) as acquired, (c), (d) after tensor SVD denoising. (e) – (f) Data from a LiZnSb [11$\bar{2}$0] / GaSb [110] interface: (e), (f) as acquired, (g), (h) after tensor SVD denoising. (i) – (n) Scree plots used to determine the tensor rank for denoising. Components in blue are retained and components in orange are discarded in the denoised reconstruction of the data. The 0 component which represents the mean of the data is omitted from the graphs. (i) – (k) STO. scree plots. (l) – (n) LiZnSb / GaSb scree plots.

Tensor SVD has also been applied to denoise hyperspectral EDS data collected from Ca-stabilized $NdTiO_3$ [001] sample, the denoised results from tensor SVD and a few other denoising methods are shown in figure 7.5. The EDS spectrum images from 1000 energy channels were integrated into seven elemental maps corresponding to seven x-ray bands, which increase the mean signal to 1.8 counts/px from 0.015 counts/px in the raw data. However, this data still has relatively low counts compare to the data collected in 4D STEM. Scree test determined denoising ranks were 10, 10, and 7 along x, y, and the energy dimension, respectively. The comparison between noisy data, and denoised data from tensor SVD, NLPCA, and conventional PCA shows that tensor SVD was able to remove noise and improve SNR from the noisy data, but the denoise is not as effective as NLPCA, especially for the Ca map which has a more limited count.

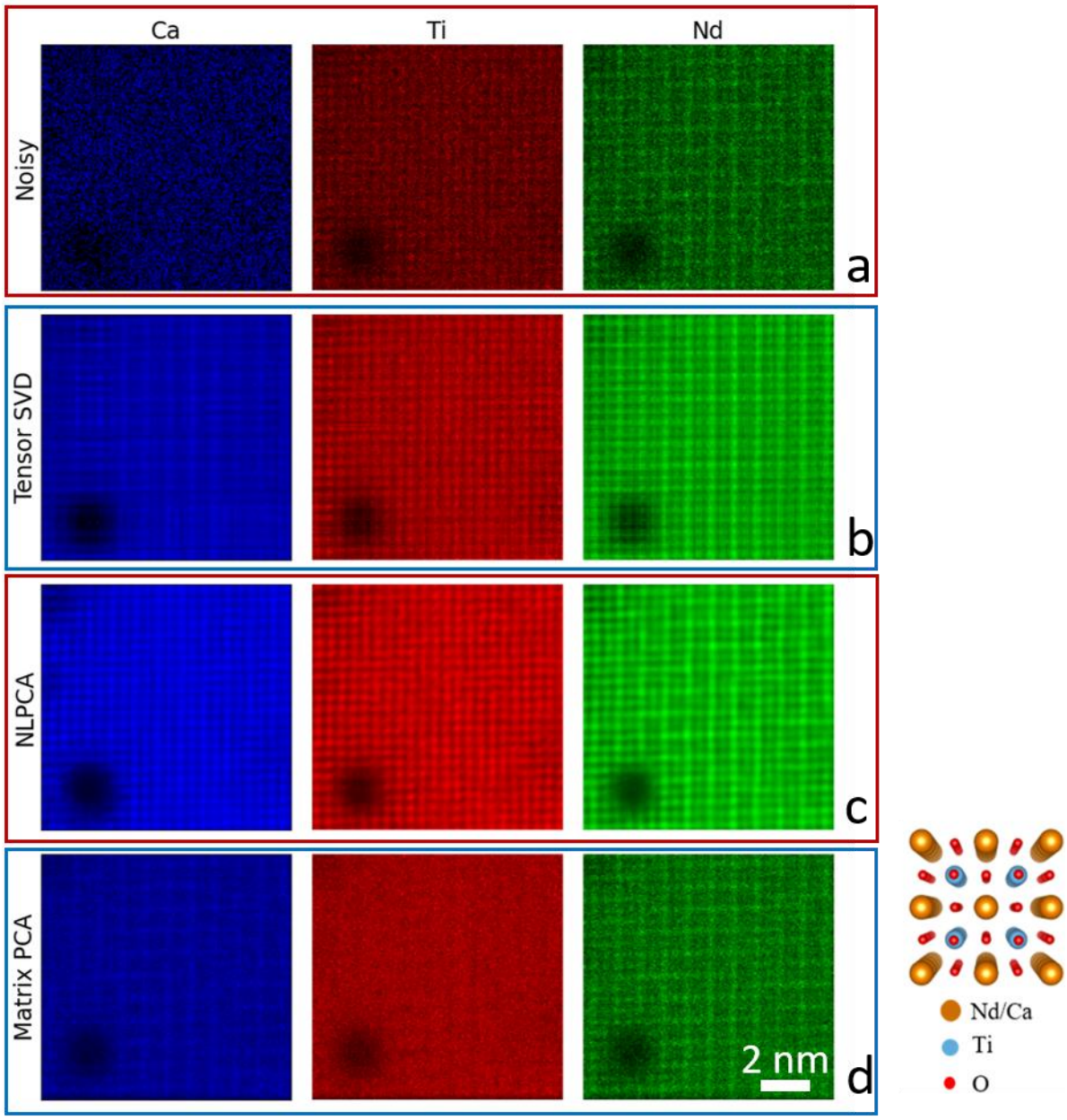

Figure 7.5, EDS elemental maps from (a) noisy data, (b) tensor SVD denoised data, (c) NLPCA denoised data, and (d) matrix PCA denoised data displayed in the order of Ca, Ti, Nd, from left to right. (a), (c) and (d) are replotted from the data in ref [169].

7.4 Comparison to Other Denoising Methods and Data Arrangements

Subsets of the simulated 4D STEM with reduced size in reciprocal space were constructed and used to compare the performance of tensor SVD, NLPCA, BM4D, and matrix PCA. Reduced data size was required in order to fit the memory requirement of the non-local style method NLPCA and BM4D. The final 3D tensor for the test with reciprocal space unfolded into one single dimension has 114×114×100 pixels on r_x, r_y, and k dimensions.

Figure 7.6 shows the comparison between noisy data, denoised data from four different methods, and the ground truth for both $SrTiO_3$ data and the Si edge dislocation data in terms of reconstructed real space image (Figure 7.6 (a) and (c)), and the CBED pattern subset (Figure 7.6 (b) and (d)). The reconstructed real space image lacks details such as the contrast between Sr and Ti sites, and the dumbbell structure in Si, as a result of limited coverage in reciprocal space. From the denoising quality quantified in PSNR shown in figure 7.6 (e) and (f), tensor SVD performs either the best or close to the best among the four methods on both structures. In terms of processing times, matrix PCA is the fastest one, closely followed by tensor SVD, and both are much faster than the non-local methods NLPCA and BM4D. Overall, with both denoising performance and processing time considered, tensor SVD is the optimal method for denoising massive 4D STEM data with good denoising performance and fast processing speed.

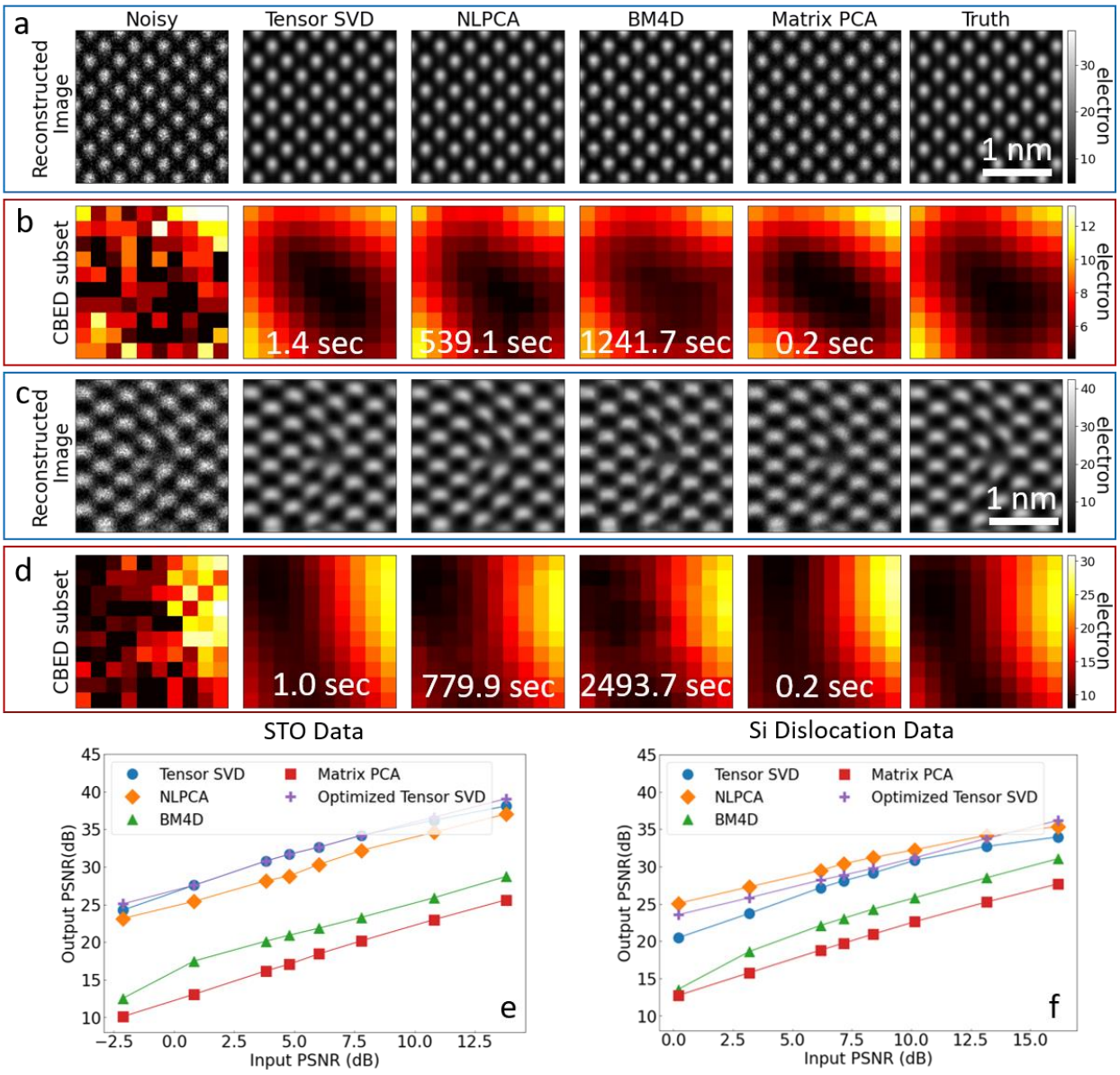

Figure 7.6. Comparison of tensor SVD to other denoising methods: (a) virtual images from simulated STO data, (b) CBED pattern patches from STO. data, (c) virtual images from simulated Si dislocation data, (d) CBED pattern subset from Si dislocation data extracted. (a) – (d) use the signal level of 1000 fps experimental data, and the numbers in (b) and (d) are the computing time required for denoising. (e)-(f) PSNR of the denoised data as a function of input noise level for (e) STO data and (f) Si dislocation data.

7.5 Conclusions

Tensor SVD is a method to find a low-rank representation of tensor data. It is analogous to component analysis of matrix data using PCA, but it preserves the full three or four (or more) dimensional structure of the data set. We have used tensor SVD to denoise atomic-resolution 4D STEM data. The method can be directly applied to multiple GB 4D STEM data sets with the adjustable parameters determined by scree tests. Computing times on a typical desktop computer is a few minutes, and the required memory is twice the size of the data. Tests on simulated data show that tensor SVD has the best or close to the best performance compared to other state-of-the-art methods which require substantially more computation time and memory. Tests on experimental data show that tensor SVD significantly improves the signal to noise ratio, and the denoised data shows clearer structures in both real space images and CBED patterns. Tensor SVD can also be applied to other high dimensional data sets, such as EDS spectrum images.

8 Strain Mapping of Nb-rich Precipitate in ZrNb Alloy

8.1 Strain Measurement Using 4D STEM

Strain engineering of materials has shown itself to be a powerful tool in materials research. The change in lattice constant has been used to tune different properties of materials, including charge mobility[183], polarization[184], superconductivity[185], and in this work, serving as the driving force of Nb-rich platelets in the ZrNb alloy[36].

TEM and STEM have long been used to characterize local strain in materials. One popular method for strain measurement via Fourier transform of the atomic-resolution image is geometric phase analysis (GPA)[186], which has been implemented in Digital Micrograph[187] and is

shown to be able to measure lattice mismatch to 1 pm precision[188]. The strain has also been directly measured from the real space image after precisely fitting the atomic column positions[35]. Both methods are limited in the size of their field of view, as a sufficiently high enough magnification has to be used to get an image of lattice and derive the strain information. 4D STEM solved this problem by calculating the strain information directly from diffraction spots[34]. Using a small convergence angle and collecting well separated diffracted beams, the distances between diffracted beams with different Bragg index are directly related to certain lattice constants. Then, the strain information could be calculated from the relative change of distances between Bragg diffraction spots at each probe position. The 4D STEM measurement of strain has shown a strain measurement precision of 0.1%, with the possibility of measuring μm-sized area in a single scan and a spatial resolution of about 1 nm.

For the 4D STEM strain measurement, the spatial resolution is determined by the size of the electron probe, which is typically about nm size as a result of the diffraction limit from a small convergence angle. And the strain measurement precision depends on the ability to accurately measuring the position of each diffraction spot. To precisely measure the position of each diffraction spot, a thin sample is often used to suppress the patterns within each diffraction spot generated by dynamical scattering. Some more complicated registration tools have been used to fit the position of diffraction disks[38], and patterned probes have been used to deliberately introduce features to the probe to enhance position registration[40].

8.2 Strain Map of Zr Precipitates

In this work, we focused on the strain states around the Zr-rich platelets precipitates in ZrNb alloy. The hypothesis has been made that the Zr-rich platelets are mainly driven by the dislocation-induced strain fields, and accelerated by the irradiation[189]. Although there is

currently no way to directly confirm the growth mechanism via in-situ methods, the residual strain states around the precipitates could give useful information, as strain driven growth would result in obvious residual strain in the final precipitates.

ZrNb alloy with 1% Nb and irradiated by 1.0 dpa proton irradiation, which contains a high density of Zr-rich precipitates, has been used for the 4D STEM study. The sample has been prepared by FIB lift out using a Zeiss Auriga Ga FIB, followed by final thinning in a Fishione 1040 nanomill to get an electron transparent sample and a smooth surface. A 1.75 mrad convergence angle was used to get well-separated diffraction spots, and 100 FPS frame rate was used on the DE-16 camera to integrate a high enough intensity on each frame under 20 pA beam current.

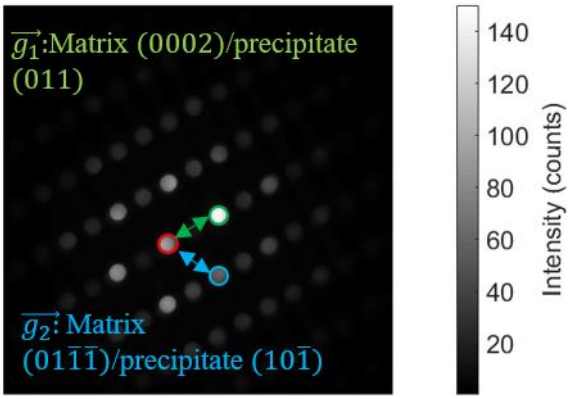

Figure 8.1 PACBED pattern from the 4D STEM scan around an Nb-rich precipitate, with two vectors corresponding to the lattice planes in strain maps marked by blue and green arrows.

The PACBED pattern from the whole 4D STEM scan is generated after compensating for the diffraction pattern drift during the scan and is shown in figure 8.1, where different Bragg diffraction spots are separated from each other. $\boldsymbol{G_1}$ is defined to be the vector corresponding to the (0002) lattice planes of the Zr matrix and the (011) planes of the precipitate, and $\boldsymbol{g_2}$ is defined

to be the one corresponding to the $(01\bar{1}\bar{1})$ of the matrix and $(10\bar{1})$ of the precipitate. These two vectors are used to calculate the four components of the strain tensor around the precipitate, namely ε_{xx}, ε_{yy}, ε_{xy}, and ε_{rot}, where x and y correspond to the directions of $\boldsymbol{g_1}$ and $\boldsymbol{g_2}$, respectively. Only four components of the strain tensor could be derived out of a total of nine, as we can only collect the 2D projection of the sample from STEM, and the out-of-plane information is missing. The detailed derivation of the four components can be found in the supplementary materials of ref[37].

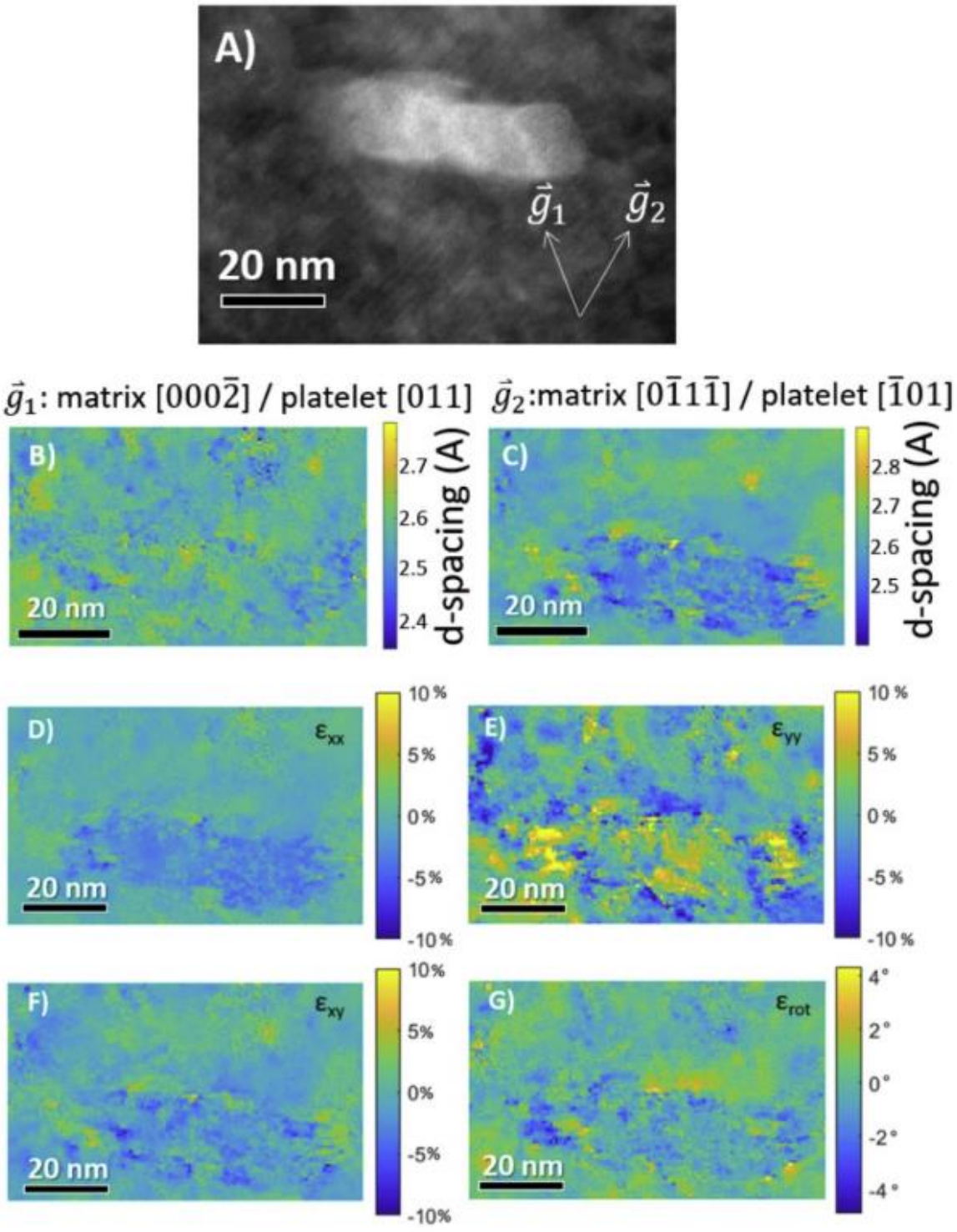
A)
$\vec{g}_1$
$\vec{g}_2$
20 nm
$\vec{g}_1$: matrix [000$\bar{2}$] / platelet [011]
$\vec{g}_2$:matrix [0$\bar{1}$1$\bar{1}$] / platelet [$\bar{1}$01]
B)
20 nm
2.7
2.6
2.5
2.4
d-spacing (A)
C)
20 nm
2.8
2.7
2.6
2.5
d-spacing (A)
D)
ε_{xx}
20 nm
10 %
5 %
0 %
-5 %
-10 %
E)
ε_{yy}
20 nm
10 %
5 %
0 %
-5 %
-10 %
F)
ε_{xy}
20 nm
10%
5%
0%
-5%
-10%
G)
ε_{rot}
20 nm
4°
2°
0°
-2°
-4°

Figure 8.2 4D STEM analysis of Nb-rich precipitates in the ZrNb alloy. (a) simultaneously acquired HAADF-STEM image showing the precipitate, (b) – (c) lattice spacing map along $\boldsymbol{g_1}$ and $\boldsymbol{g_2}$, (d) – (g), components of strain tensor calculated from 4D STEM data.

The strain components, together with simultaneously acquired HAADF-STEM image and the lattice spacing maps, are shown in figure 8.2. Figure 8.2 (d) and (e) show that the precipitate is coherent with almost no strain along the *x* direction, while a strong strain can be observed along the *y* direction, especially at the two ends of the precipitate. The strong residual strain at the two ends of the long side of the precipitate agrees with the hypothesis of strain driven growth of the precipitate. On the rotation map shown in figure 8.2 (g), a few bright spots can be seen along the edge of the precipitate, which corresponds to the origin of edge dislocations.

8.3 Conclusions

4D STEM has been used to study the strain states around Nb-rich precipitates grown in ZrNb alloy under irradiation. Areas with positive strain in the ε_{yy} map indicate platelet has high tensile strain along the thickness direction. The ε_{rot} map clearly shows the residual coherent strain field at the platelet's length interfaces, and the dislocation induced strain fields nearby the platelet's two ends. The observed strain states agree with the strain driving growth mechanism of the platelet-shaped precipitates.

9 Polarization Measurement in Hexagonal Heusler Alloys

9.1 Polar Metals and Polarization Measurements

Polar metal refers to the category of materials that simultaneously have free carriers and hold polarization in the structure. Polar metals show promising properties in nonlinear optics, nonreciprocal charge transport[190], and can be potentially used as electrode materials for

ferroelectric capacitors[191]. A few different types of polar metals have been proposed, as can be seen in the black curves in Figure 9.1 (a). Most of them are oxides with perovskite structure[192], with one exception being a metal dichalcogenide[193].

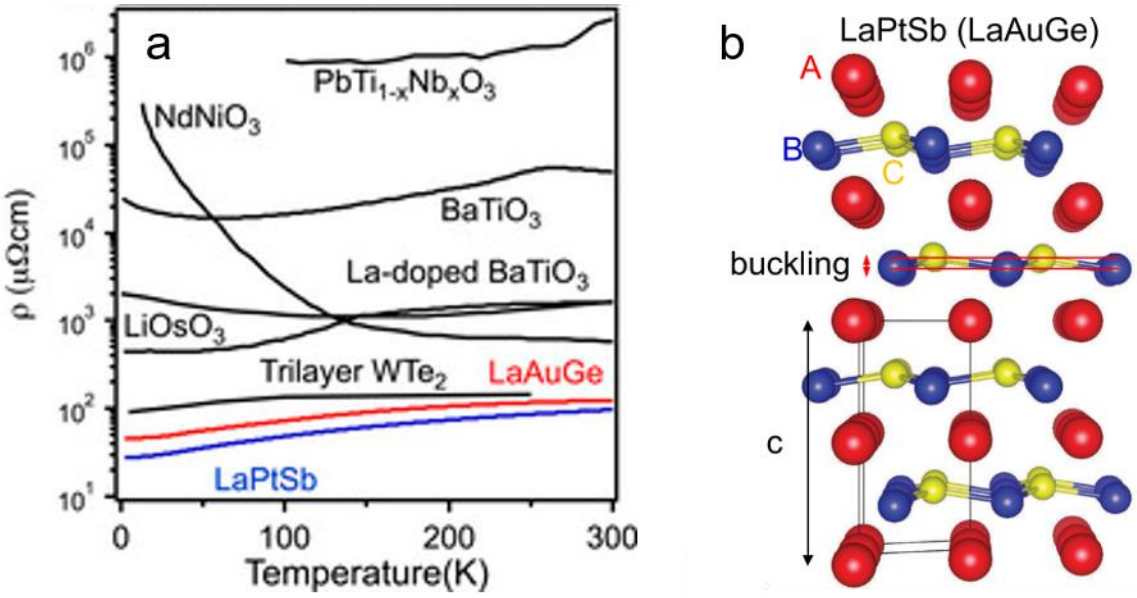

Figure 9.1 (a) Temperature dependence of resistivity for LaAuGe, LaPtSb, and other materials that have been proposed to be polar metals. (b) Atomic structure model of LaAuGe and LaPtSb with a single unit cell marked in black line and the buckling definition marked by the red arrow.

Here, we propose a new type of polar metal based on hexagonal Heusler alloy, include LaAuGe and LaPtSb[194]. Figure 9.1 (a) shows the temperature dependence of resistivity for LaAuGe and LaPtSb, both of them show a higher resistivity with increased temperature, which is a characteristic of conductors with free carriers. Compared with other polar metals from previous reports, both hexagonal Heusler alloys show orders of magnitude lower resistivity. The atomic structure of the Heusler alloys is shown in figure 9.1 (b). La atoms form a hexagonal sub-lattice (A sites), with planes of AuGe/PtSb fills in between (B and C sites). A constant non-zero

displacement between the planes formed by B and C sites is what we called a buckling displacement[86] and result in the polar structure.

The characterization of the polarization in the hexagonal Heusler alloy includes structural polarization and electrical polarization. Structural polarization refers to the fact that the lattice has no inversion symmetry, which in this case, is equal to the existence of a non-zero constant buckling displacement. This can be measured by directly visualize the lattice structure using HR-STEM, which will be discussed in detail in section 9.2. Electrical polarization refers to the fact that an internal electric field is generated by the polar structure. This is common in ferroelectric materials and is typically measured by switching the internal electric field and measure the hysteresis loop[89]. However, this is not possible for polar metals, as the external electric field would directly generate a current in the material instead of switching the internal electric field. Here, we propose to measure the internal electric field using 4D STEM, which would be discussed in detail in section 9.3.

9.2 Structural Polarization Measurement with HRSTEM

To characterize the polar structure of LaPtSb and LaAuGe, we applied high-resolution HAADF-STEM to directly visualize the lattice structures of the two materials. LaPtSb and LaAuGe thin film were grown using molecular beam epitaxy (MBE) on Al_2O_3 substrates, and TEM samples were prepared by FIB liftout using a Zeiss Auriga Ga FIB and final thinning using a Fishione 1040 nanomill to get an electron-transparent sample with a smooth surface. HAADF-STEM image series were acquired on a 200 kV Titan STEM with a 24.5 mrad semi-convergence angle and a HAADF detector covering 54-329 mrad. Image series were collected and post-processed with NRR to obtain high SNR for Gaussian fitting of the atom positions.

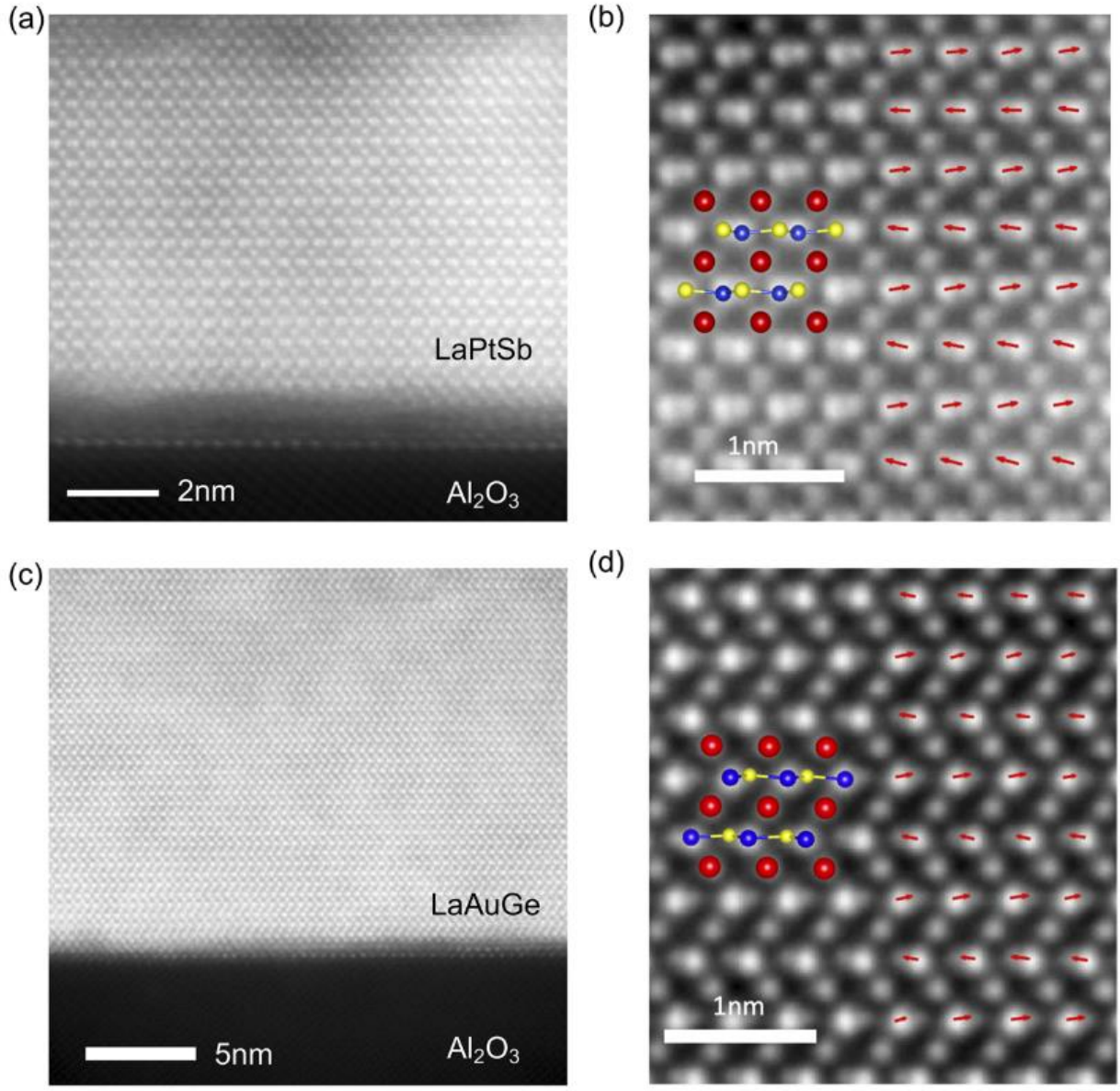

Figure 9.2 HAADF-STEM image of (a) and (b) LaPtSb and (c) and (d) LaAuGe measured along a [11$\bar{2}$0] zone axis with [0001] growth direction points up. The atomic structure is overlapped in (c) and (d) with red arrows marking the displacement between Pt/Au and Sb/Ge sites.

Figure 9.2 (a)-(d) shows the HAADF-STEM characterization of the LaPtSb and LaAuGe thin film, with a clear buckling structure shown in Figure 9.2 (c) and (d). The position of the Pt/Au and Sb/Ge sites are determined by fit each pair of PtSb/AuGe atoms to a sum of two 2D

Gaussian functions. The atomic positions were then used to calculate the buckling displacements in the two materials, and the result suggests the buckling displacement along the polar direction ([0001] direction, which is also the growth direction) are 0.222 Å for LaPtSb and 0.176 Å for LaAuGe, confirming the structural polarization in both structures.

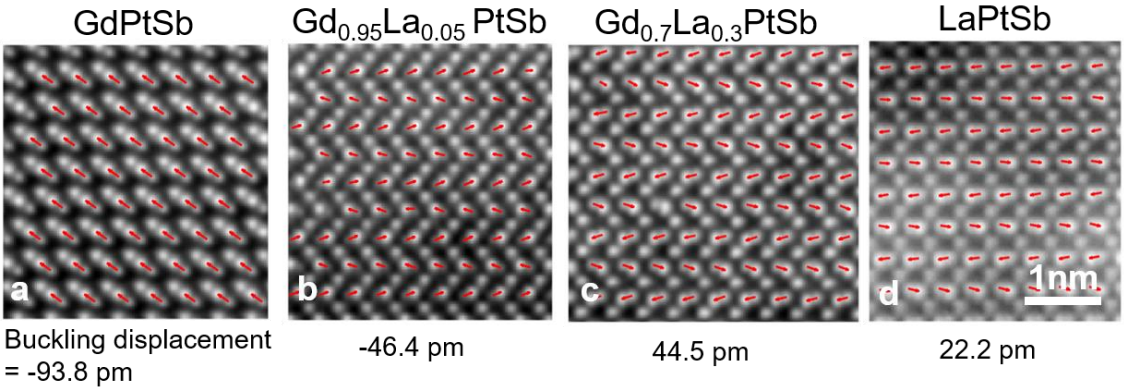

Figure 9.3 HAADF-STEM image of $La_xGd_{1-x}PtSb$ with buckling analysis and the buckling displacement along polar direction for (a) x=0, (b) x=0.05, (c) x=0.3, and (d) x=1.

On LaPtSb, we have further demonstrated that by doping the La sites with Gd, we are able to control both the buckling direction and magnitude. Figure 9.3 shows the high-precision HAADF-STEM characterization on a series of LaPtSb with different Gd doping fractions. Here, the substrate is on the bottom side, and with buckling defined as the displacement from Sb to Pt, we define the cases where Pt is closer to the substrate as positive buckling displacement. From the Gaussian fitting of Sb and Pt positions, the four different doping fractions show different buckling magnitude as well as directions. This tuning over the buckling displacement means that we will be able to grow a more complex polarization structure with possible polar domains by controlling the Gd doping fraction in different layers.

9.3 Electrical Polarization Measurement with 4D STEM

With the structural polarization confirmed from HAADF-STEM images, a more interesting question is whether this structure holds an internal electric field from the polarization. To answer this question, we applied 4D STEM to directly probe the inner electric field using an electron probe. In 4D STEM, the negatively charged electron probe will move under the Columb force from the electric field, and this movement will be captured in the diffraction pattern, and by analyzing the change in the diffraction pattern at each probe position, one is able to map the electric field distribution in the sample.

Currently, there exist two different flavors of 4D STEM to measure the electric field. The first setup uses a small convergence angle to obtain separated the undiffracted beam from the diffracted beam, and then derive the electric field from the undiffracted beam itself. This method has been successful in resolving the negative capacitance in polar vortices[44] and resolving room-temperature polar skyrmions[43]. The disadvantage of this setup is the limited spatial resolution, as the beam size is typically nm at a small convergence angle and thus won't be able to resolve the fine electric field distribution inside a single unit cell or around a single atomic site. On the other hand, this method has the advantage of being able to eliminate the effect from sample tilt, as the tilt would only change the intensity of the diffracted beams.

The other setup uses a large convergence angle to get a well-focused beam, typically with sub-Å size, then perform a fine scan with sub-unit cell step size on the sample. The same setup has been used in differential phase contrast (DPC) STEM for a long time and is capable of resolving the electric field at high spatial resolution. Recently, this setup has been used to study the electric field generated by polarization in GaN/AlN multilayer structure[60], with another work observing a similar electric field in the $BiFeO_3$ thin film with polar structure[61]. In this

setup, it is important to eliminate the effect of sample tilt, as the sample tilt generates a similar signal as the electric field from polarization, even with a small tilt. A typical way to eliminate the effect of sample tilt is to find a reference within the sample itself that has the same tilt but a different polarization state, such as the difference between GaN and AlN in ref[60], and the difference between $BiFeO_3$ and $SrTiO_3$ in ref[61].

In this work, we followed the setup used in ref[60]. A small electron probe with sub-Å size was used to scan over the sample, center of mass (COM) of the CBED pattern is calculated at each probe position, followed by an average over Voroni unit cells covering half unit cell to cancel out the contribution from the strong electric field around each atom to reveal the weak electric field generated by polarization. Figure 9.4 (a) and (b) show the electric field signal from polarization after averaging over Voronoi cells along the [1100] direction (non-polar direction) and the [0001] direction (polar direction). It can be clearly seen that the polar direction shows a strong electric field signal compared to the non-polar direction, which shows a signal that is very close to zero. This preliminary result suggests that there might exist an internal long-range electric field generated by the polar structure in polar metals.

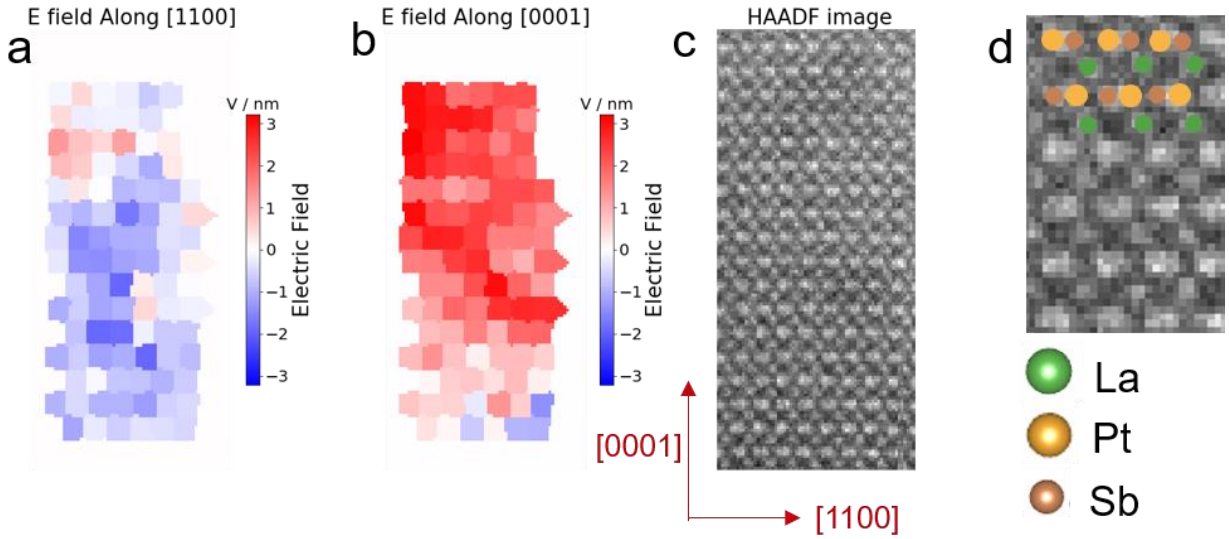

Figure 9.4 (a) and (b) electric field measurement after average over Voronoi cells in LaPtSb along [1100] direction and [0001] direction. (c) simultaneously acquired HAADF image showing the lattice structure of LaPtSb and (d) a zoomed-in patch from the HAADF image with atomic sites marked on the image.

9.4 Conclusions

STEM and 4D STEM has been used to study the polarization in thin film polar metal LaPtSb and LaAuGe synthesized using MBE. High-resolution STEM was used to confirm the polar structure in LaPtSb and LaAuGe, with buckling displacement quantified in both structures as well as Gd doped LaPtSb thin films. 4D STEM has been used to directly map the long-range electric field in LaPtSb generated by the polar structure. Measurement shows a clear electric field signal along the polar direction compare to the non-polar direction, but the results still need to be carefully inspected to eliminate the effect from sample tilt and other sources that could generate a similar signal as the electric field.

10 Study of Twist Pattern in Supertwist WS_2 Structure with 4D STEM

10.1 Super Twist Structure from Non-Euclidean Surface

Twisted multilayer of 2D materials has drawn increasing interest over the years as the Morie superlattice shows emerging phenomena such as unconventional superconductivity[195] and exciton behaviors[196]. However, there have been limited ways to twist the lattice. Most works prepare twisted 2D stacks using mechanical stacking, which requires dedicate control and is very inefficient. Dislocation driven Eshelby twist has been reported to twist the lattice[197],[198], but the twist angle is negatively correlated to the lateral size of the structure and thus does not fit into 2D materials.

In this work, a radically new way is proposed to grow layered 2D material with a constant twist angle between consecutive layers using a substrate with a non-Euclidean surface. On a non-Euclidean surface, a full circle is either larger than 360° on a hyperbolic surface, or less than 360° on an elliptic surface. Then if we grow a layered 2D material spiral with a screw dislocation at the center on these non-Euclidean surface, the lattice itself still covers a full 360°, then the difference between 360° and the angle of a full-circle on the surface will generate a twist angle and lead to two different types of twisted patterns on an elliptic surface or hyperbolic surface.

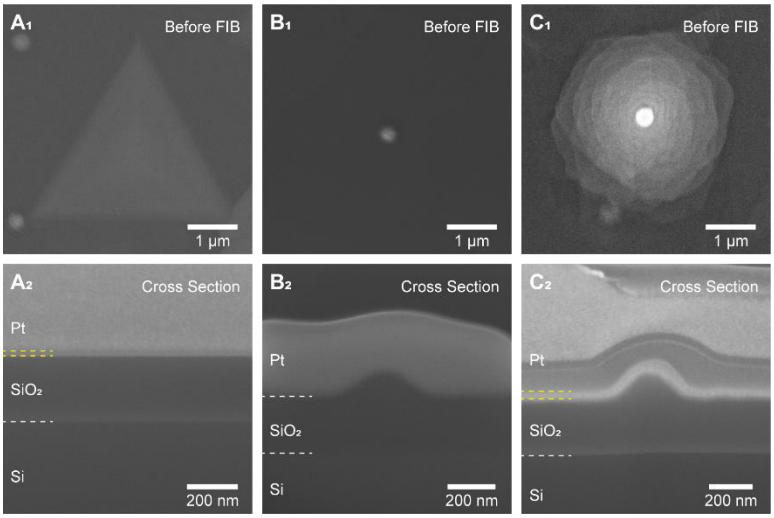

Figure 10.1 (A) An aligned WS_2 spiral. (B) A deformed SiO_2 particle. (C) A super-twisted WS_2 spiral with a deformed SiO_2 particle at its center. The first row (A_1, B_1, C_1) is the top-down SEM images of the objects before FIB. The second row (A_2, B_2, C_2) is the corresponding cross-section images with boundaries highlighted by dashed lines. The areas between the two yellow dashed lines are WS_2. Fig. B_2 and C_2 clearly show that the deformed SiO_2 nanoparticles formed cone-shaped protrusions on Si/SiO_2 substrates after the growth process.

The non-Euclidean surface was generated by dispersing SiO_2 nanoparticles on a flat SiO_2 substrate, then heat the surface to deform the SiO_2 nanoparticles to form a protruded substrate.

Then the top of the cone is an elliptic surface, while the side of the deformed nanoparticle is a hyperbolic surface. The image from the top and side of the deformed nanoparticle can be found in Figure 10.1 (b). These deformed nanoparticles then serve as the growth core for the multilayer WS_2 structure in the vapor assisted chemical vapor deposition synthesis[199]. An example of a multilayer WS_2 structure grow on top of the deformed nanoparticle is shown in Figure 10.1 (c), where a twist pattern can be clearly seen in the top-down view.

10.2 STEM and 4D STEM Characterization of the Twist Pattern

Although the geometric twist pattern can be easily characterized using SEM or atom force microscopy (AFM) as the structure is pretty large with μm size, none of these methods could resolve the lattice twist in the structure, which is much more important than the geometric twist. To study the lattice twist pattern, we applied 4D STEM with the same setup as described in section 2.3.2 and section 8.1 for strain mapping, and we use the twist angle between different sets of diffraction spots to directly measure the twist angles of different layers.

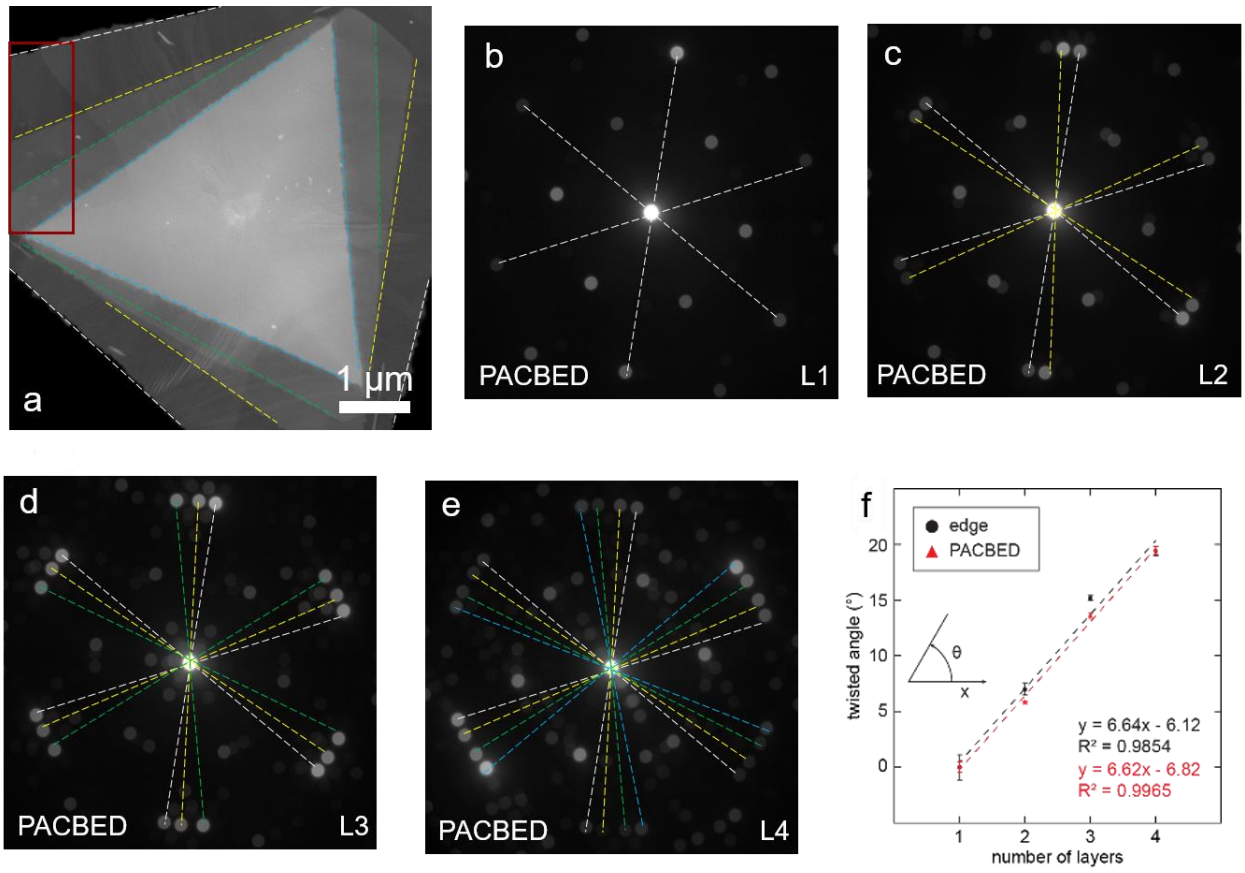

Figure 10.2 (a) low magnification HAADF-STEM image from the whole structure, (b) – (e) PACBED pattern generated from areas with 1-, 2-, 3-, and 4- layers of WS_2 structure respectively, (e) relationship between layer index and geometric twist angle and lattice twist angle measured using the shape and PACBED pattern.

The twist angle was measured on a simple four layers structure with a constant twist pattern, as shown in Figure 10.2 (a). 4D STEM data was collected within the area marked by the red box, and PACBED patterns were generated from areas with 1-, 2-, 3-, and 4- layers of WS_2 structure, as shown in Figure 10.2 (b) – (e), respectively. The edge of each layer is marked with dashed lines with different colors, and the set of diffraction spots from that layer is marked by the dashed lines with the same color in the PACBED patterns. Using the shape of each layer in the low magnification HAADF-STEM image, and the diffraction spots in the PACBED patterns,

we are able to plot the relationship between geometric twist (black dots) and lattice twist (red dots) and the layer index, as shown in Figure 10.2 (f). The results suggest that the lattice twist closely matches to the geometric twist, and shows a linear twist pattern with a constant twist angle between consecutive layers.

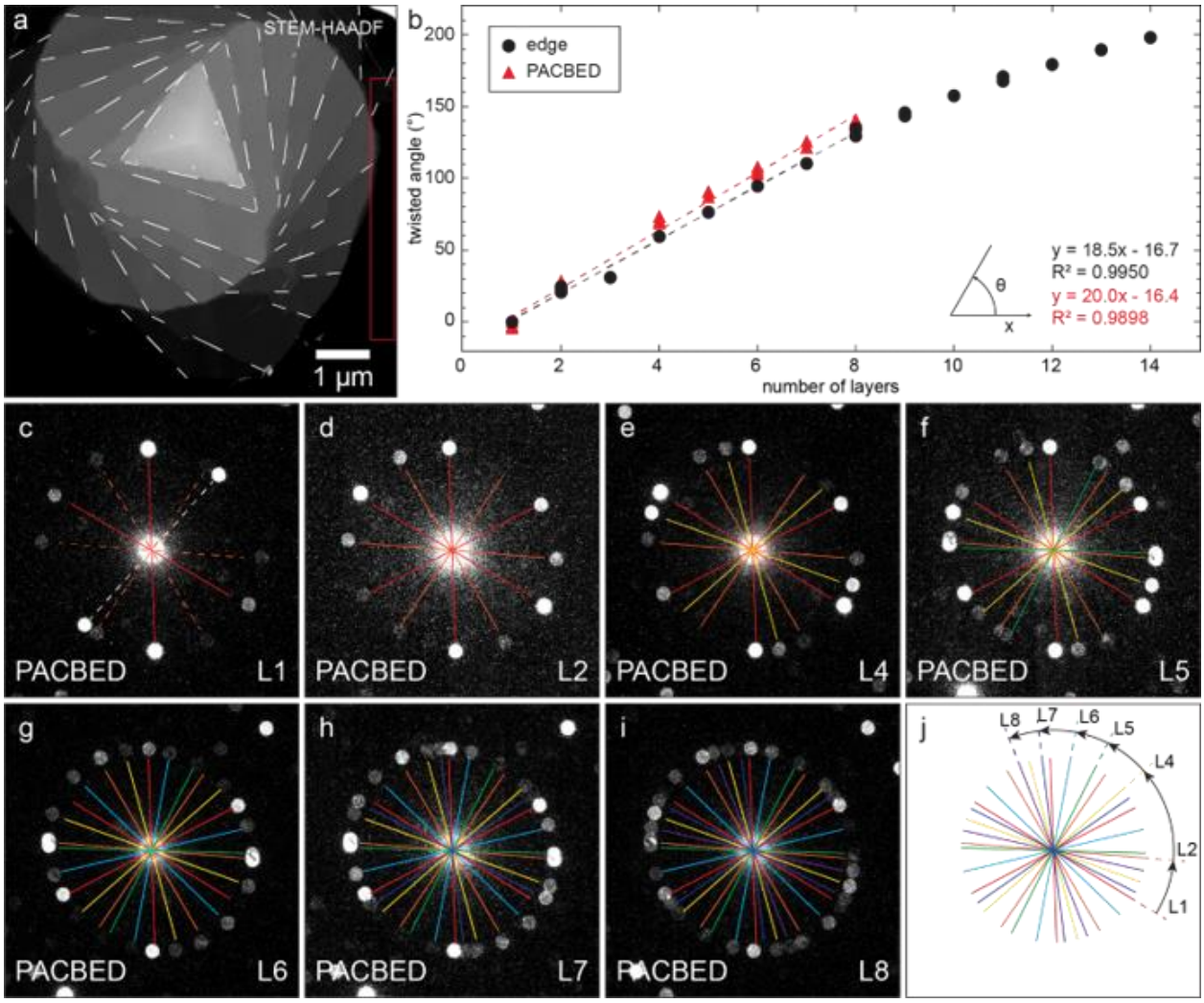

Figure 10.3 (a) low magnification HAADF-STEM image from the whole structure, (b) relationship between layer index and geometric twist angle and lattice twist angle measured using the shape and PACBED pattern, (c) – (i) PACBED pattern generated from areas with 1-8

layers of WS_2 structure except three layers (j) relationship between layer index and the position of each set of the diffraction pattern.

Another example of lattice twist angle measurement on a more complex structure with more layers is shown in Figure 10.3. 4D STEM data was collected from the area marked by the red box in Figure 10.3 (a), and the PACBED patterns are generated from areas with 1, 2, and 4-8 layers of twisted WS_2. The area with 3 layers of twisted WS_2 was missing in the structure, probably damaged during sample transfer. With more than 8 layers of WS_2, the sample gets thick with a strong background generated by inelastic scattering and lots of overlapping between diffraction spots from different layers. Both factors make it hard to identify the diffraction spots on the PACBED pattern. A plot of geometric twist angle and lattice twist angle for each layer is shown in Figure 10.3 (b), which shows a similar linear twist pattern and the match between geometric twist and lattice twist.

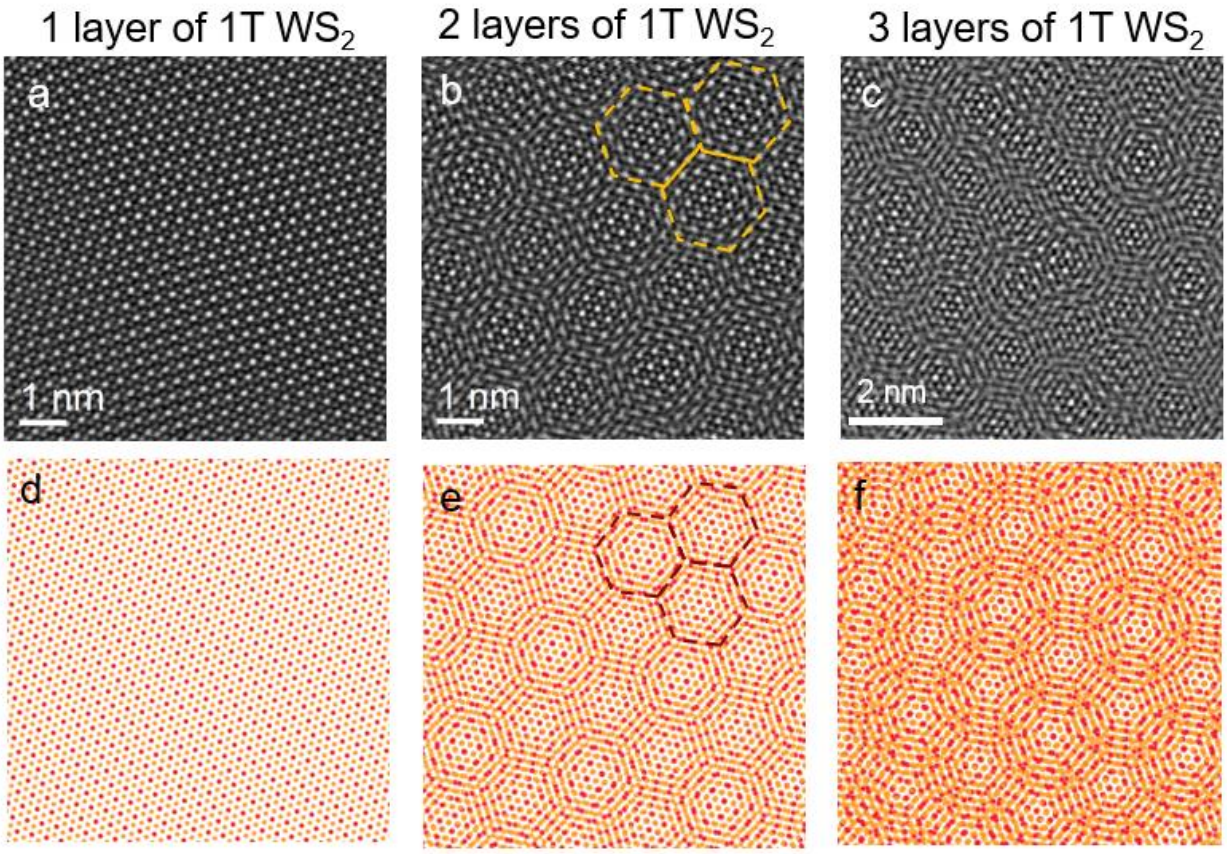

Figure 10.4 (a)-(c) Morie patterns acquired from HAADF-STEM on 1-, 2-, and 3- layers of twisted WS_2, (d)-(f) simulated Morie pattern by overlapping 1, 2, and 3 layers of 1T WS_2 models with the twist angle measured using PACBED patterns.

Using high-resolution HAADF-STEM, we were able to directly visualize the Moire patterns from 1-, 2-, and 3- layers of twisted WS_2, which is shown in Figure 10.4 (a) – (c). The HAADF-STEM images are filtered using an average background subtraction filter[200] based on the work in ref[124] to remove the non-uniform background generated by carbon contamination. Expected Moire patterns from the twist angles measured using PACBED patterns are generated by overlapping multiple layers of 1T WS_2 models and is shown in Figure 10.4 (d) – (f). The

experimental Moire patterns closely match the expected ones, again proofing that the lattice twist angles measured using PACBED patterns are accurate.

10.3 Conclusions

A novel way to grow multilayer twisted 2D materials using a non-Euclidean substrate has been proposed and tested on WS_2. 4D STEM has been used to study the lattice twist angles by generating PACBED patterns from areas with different numbers of twisted layers. Result suggests that the lattice twist angles closely match the geometric twist angle and is almost constant between consecutive layers, which agrees with the twist model we are proposing. The twist angles are directly visualized using HR-STEM by directly visualize the Moire pattern, and the results confirmed the measurement using PACBED patterns.

11 Outlook

This chapter will focus on possible future researches based on 4D STEM, and most of them are just simple ideas without very careful considerations. The possible researches are divided into three parts. The first part discusses possible new image modes of 4D STEM by using a faster detector or combining with in-situ techniques and spectroscopies. The second part discusses possible ways to extract more information from 4D STEM data sets. And the third part focuses on the combination of machine learning methods and 4D STEM.

There is little doubt that 4D STEM could benefit from the usage of a faster detector with acquisition frequency closer to the typical STEM scan frequencies. A faster detector and shorter collection time could reduce scan distortions caused by sample drift during acquisition and make it possible to apply distortion-correction methods with serial 4D STEM. Another benefit of using

a faster detector is to simplify the noise statistic by running in counting mode[201] and make it possible to apply Poisson denoise methods to directly solve the ground truth.

Besides the usage of a faster detector, 4D STEM can also be combined with other imaging techniques in STEM. One possibility is to apply 4D STEM to in-situ experiments. So far, 4D STEM has been successfully applied to study the in-situ deformation process and measure the strain change in the process[202]. Other than that, 4D STEM data can also be used to study other processes such as charge transfer under electrical bias[203], polar-nonpolar phase transition during heating[204], polar domain switching under external strain[205] or bias[206], and oxygen vacancy segregation and ordering during heating[207]. The relatively slow collection speed could be problematic when combining with in-situ experiments. A 4D STEM scan with fewer beam positions and lower resolution could be beneficial in in-situ experiments.

Another possibility is to combine 4D STEM with the EDS and EELS spectroscopy data collection in STEM. A combination of these techniques simultaneously collects the energy-dispersive data, which is sensitive to composition and bonding change, and the diffraction data, which is sensitive to structural change. The combination of 4D STEM and EELS would require a customized detector that allows part of the electron beam to pass through, and it has been shown that ptychography problems could be solved with the center of the diffraction pattern missing[208].

Other than new experiment settings, more information about the material structure could be extracted from the 4D STEM data with new ways to analyze the data. One possible improvement is to recover the whole 3D atomic structure instead of the 2D projection of the structure from the data. Efforts have been made to solve the whole 3D scattering matrix using 4D STEM focal series[57] and the multislice ptychography problem[209]. The ability to retrieve the full 3D

structure could be used to study the full 3D strain states, atom ordering in alloys, and position of point defects inside the material. Other than 3D reconstruction, fine structure change in the material could be resolved by interpreting 4D STEM data quantitatively and match to multislice simulations. It has been shown that by including the inelastic scattering in the simulation, the difference between experiment and simulation could be reduced to below 10% over a wide annular range and below 5% for some narrower annular range[77], which means any structural change generating a diffraction intensity difference larger than these residual difference could be resolved by quantitatively matching experiment and simulation.

4D STEM can also be combined with machine learning methods to facilitate data acquisition and analysis. First, machine learning methods like CNN can be used to understand the experiment environment, like aberration status[210], sample tilt, thickness from collected CBED patterns. With this information and a control system to tune the microscope, the 4D STEM acquisition could be automatically performed with minimum human interaction. Besides automated data collection, compressive sensing could be applied to both the real space scan and the diffraction pattern acquisition to speed up the acquisition.

With data collected, machine learning methods could be applied to pre-process the collected data. One possible application is to use the U-net[211] as a counting filter, which performs image segmentation on the collected image and classifies each pixel as an electron hit or no electron hit. CNNs can also be trained to compress 4D STEM data[212] for faster saving and transport of the massive 4D STEM data by learning characteristic features in 4D STEM data and then compress the data by saving a map of features instead of the raw data. Machine learning methods can also be used for denoising. Besides the tensor SVD method described in Chapter 7, certain types of

generative adversarial networks [213] and U-nets[214] are also powerful tools that can be trained to learn certain noise pattern and denoise 4D STEM data.

Machine learning methods can also help to understand the collected 4D STEM data. The data processing scheme described in Chapter 6 can be adapted to analyze other aspects of the sample, with possible performance improvement by adapting novel network structures. Another possible model that can be used to understand materials science properties from the collected data is graph neural network (GNN)[215][216]. Unlike CNN that uses a convolutional base to detect features, GNN generates a graph from the input data, performs graph convolutions, then extracts information from the final graph. GNNs are much less computationally intensive than CNNs, as it has much less tunable parameters. The graph structure can hold prior knowledge about the data and embed physical meaning in the network, instead of processing the data simply as images. Neural networks can also be applied to solve ptychography problems. Work has been reported using neural network to mimic the electron propagation in the sample[217], then solve the reverse propagation problem using back-propagation. Although this would be a powerful tool if working properly, the solving process would be slow. A possible implementation of neural networks, possibly U-nets, is to directly learn the relationship between the collected 4D STEM data and the phase information of the sample, then directly generate the phase image from input data using forward propagation. Although this may not be as accurate as the iterative solvers like ePIE, it could significantly reduce the time consumption.

[1] P. M. Voyles, D. A. Muller, J. L. Grazul, P. H. Citrin, and H.-J. L. Gossmann, “Atomic-scale imaging of individual dopant atoms and clusters in highly n-type bulk Si.,” *Nature*, vol. 416, no. 6883, pp. 826–9, Apr. 2002.

[2] R. Erni, M. Rossell, C. Kisielowski, and U. Dahmen, “Atomic-Resolution Imaging with a Sub-50-pm Electron Probe,” *Phys. Rev. Lett.*, vol. 102, no. 9, p. 096101, Mar. 2009.

[3] S. Weiner *et al.*, “Atomic-Scale Chemical Imaging of Composition and Bonding by Aberration-Corrected Microscopy,” no. February, pp. 1073–1077, 2008.

[4] C. Ophus, “Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond,” *Microsc. Microanal.*, vol. 25, no. 3, pp. 563–582, Jun. 2019.

[5] A. R. Lupini, D. Peckys, N. de Jonge, R. Sougrat, and S. Pennycook, *Three-Dimensional Aberration-Corrected Scanning Transmission Electron Microscopy for Biology*, no. January. 2007.

[6] O. L. Krivanek *et al.*, “Progress in ultrahigh energy resolution EELS,” *Ultramicroscopy*, vol. 203, no. November 2018, pp. 60–67, 2019.

[7] C. J. R. Duncan, D. A. Muller, and J. M. Maxson, “Lossless monochromation for electron microscopy with pulsed photoemission sources and rf cavities,” pp. 1–17, 2020.

[8] P. E. Batson, N. Dellby, and O. L. Krivanek, “Sub-angstrom resolution using aberration corrected electron optics,” *Nature*, vol. 418, no. 6898, pp. 617–620, 2002.

[9] J. R. Jinschek, “Achieve atomic resolution in in situ S/TEM experiments to examine complex interface structures in nanomaterials,” *Curr. Opin. Solid State Mater. Sci.*, vol.

21, no. 2, pp. 77–91, 2017.

[10] Y. Jiang *et al.*, “Electron ptychography of 2D materials to deep sub-ångström resolution,” *Nature*, 2018.

[11] S. Morishita, R. Ishikawa, Y. Kohno, H. Sawada, N. Shibata, and Y. Ikuhara, “Resolution Achievement of 40.5 pm in Scanning Transmission Electron Microscopy using 300 kV Microscope with Delta Corrector,” *Microsc. Microanal.*, vol. 24, no. S1, pp. 120–121, 2018.

[12] P. M. Voyles, J. L. Grazul, and D. A. Muller, “Imaging individual atoms inside crystals with ADF-STEM,” *Ultramicroscopy*, vol. 96, no. 3–4, pp. 251–273, Sep. 2003.

[13] P. Zhang, J. J. Maldonis, M. F. Besser, M. J. Kramer, and P. M. Voyles, “Medium-range structure and glass forming ability in Zr–Cu–Al bulk metallic glasses,” *Acta Mater.*, vol. 109, pp. 103–114, May 2016.

[14] J. P. McCaffrey, “Improved TEM samples of semiconductors prepared by a smallangle cleavage technique,” *Microsc. Res. Tech.*, vol. 24, no. 2, pp. 180–184, 1993.

[15] M. Bornhöfft, J. Benke, J. Pries, P. M. Voyles, M. Wuttig, and J. Mayer, “TEM/STEM investigation of the crystallization kinetics of GeTe and Ag4In3Sb67Te26,” *Eur. Microsc. Congr. 2016 Proc.*, no. 1, pp. 1002–1003, 2016.

[16] R. Ishikawa, A. R. Lupini, S. D. Findlay, T. Taniguchi, and S. J. Pennycook, “Three-dimensional location of a single dopant with atomic precision by aberration-corrected scanning transmission electron microscopy,” *Nano Lett.*, vol. 14, no. 4, pp. 1903–1908, 2014.

[17] M. Homer, P. Cappillino, F. El Gabaly, H. Gnaegi, D. Robinson, and J. Sugar, "Preparation of Electron and X-Ray Transparent Inorganic Particles for Analytical Microscopy Using the Ultramicrotome," *Microsc. Microanal.*, vol. 21, no. S3, pp. 1815–1816, 2015.

[18] A. E. Ennos, "The origin of specimen contamination in the electron microscope," *Br. J. Appl. Phys.*, vol. 4, no. 4, pp. 101–106, 1953.

[19] A. E. Ennos, "The sources of electron-induced contamination in kinetic vacuum systems," *Br. J. Appl. Phys.*, vol. 5, no. 1, pp. 27–31, 1954.

[20] A. J. V. Griffiths and T. Walther, "Quantification of carbon contamination under electron beam irradiation in a scanning transmission electron microscope and its suppression by plasma cleaning," *J. Phys. Conf. Ser.*, vol. 241, 2010.

[21] S. Horiuchi, T. Hanada, M. Ebisawa, Y. Matsuda, M. Kobayashi, and A. Takahara, "Contamination-free transmission electron microscopy for high-resolution carbon elemental mapping of polymers," *ACS Nano*, vol. 3, no. 5, pp. 1297–1304, 2009.

[22] J. M. Zuo, "Electron detection characteristics of a slow-scan CCD camera, imaging plates and film, and electron image restoration," *Microsc. Res. Tech.*, vol. 49, no. 3, pp. 245–268, 2000.

[23] M. W. Tate *et al.*, "High Dynamic Range Pixel Array Detector for Scanning Transmission Electron Microscopy," *Microsc. Microanal.*, no. 2014, pp. 1–13, 2016.

[24] H. Ryll *et al.*, "A pnCCD-based, fast direct single electron imaging camera for TEM and STEM," *J. Instrum.*, vol. 11, no. 04, pp. P04006–P04006, 2016.

[25] “QUADRO - Dectris.” [Online]. Available: https://www.dectris.com/products/quadro/. [Accessed: 28-May-2020].

[26] S. Tem, “K2 Direct Detection Camera,” *Gatan*, pp. 3–4, 2015.

[27] R. N. Clough, G. Moldovan, and A. I. Kirkland, “Direct Detectors for Electron Microscopy,” *J. Phys. Conf. Ser.*, vol. 522, p. 012046, 2014.

[28] X. Li *et al.*, “Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM,” *Nat. Methods*, vol. 10, no. 6, pp. 584–590, 2013.

[29] “K3 Camera | Gatan, Inc.” [Online]. Available: https://www.gatan.com/k3-camera. [Accessed: 28-May-2020].

[30] “DE-16 – Direct Electron.” [Online]. Available: https://www.directelectron.com/products/de-16/. [Accessed: 28-May-2020].

[31] R. Plackett, I. Horswell, E. N. Gimenez, J. Marchal, D. Omar, and N. Tartoni, “Merlin: A fast versatile readout system for Medipix3,” *J. Instrum.*, vol. 8, no. 1, 2013.

[32] G. W. Paterson, R. J. Lamb, R. Ballabriga, D. Maneuski, V. O’Shea, and D. McGrouther, “Sub-100 nanosecond temporally resolved imaging with the Medipix3 direct electron detector,” *Ultramicroscopy*, vol. 210, no. May 2019, p. 112917, 2020.

[33] A. Phys and E. Liberti, “Phase reconstruction using fast binary 4D STEM data Phase reconstruction using fast binary 4D STEM data,” vol. 124101, no. March, 2020.

[34] V. B. Ozdol *et al.*, “Strain mapping at nanometer resolution using advanced nano-beam electron diffraction,” *Appl. Phys. Lett.*, vol. 106, no. 25, 2015.

[35] A. A. Oni *et al.*, “Large area strain analysis using scanning transmission electron

microscopy across multiple images," *Appl. Phys. Lett.*, vol. 106, no. 1, 2015.

[36] Z. Yu *et al.*, "Microstructure and microchemistry study of irradiation-induced precipitates in proton irradiated ZrNb alloys," *Acta Mater.*, vol. 178, pp. 228–240, 2019.

[37] Y. Han *et al.*, "Strain Mapping of Two-Dimensional Heterostructures with Subpicometer Precision," *Nano Lett.*, vol. 18, no. 6, pp. 3746–3751, 2018.

[38] T. C. Pekin, C. Gammer, J. Ciston, A. M. Minor, and C. Ophus, "Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping," *Ultramicroscopy*, vol. 176, no. August 2016, pp. 170–176, 2017.

[39] T. Grieb *et al.*, "Strain analysis from nano-beam electron diffraction: influence of specimen tilt and beam convergence," *Ultramicroscopy*, 2018.

[40] S. E. Zeltmann *et al.*, "Patterned probes for high precision 4D-STEM bragg measurements," *Ultramicroscopy*, vol. 209, no. November 2019, p. 112890, 2020.

[41] E. Padgett *et al.*, "The exit-wave power-cepstrum transform for scanning nanobeam electron diffraction: robust strain mapping at subnanometer resolution and subpicometer precision," *Ultramicroscopy*, vol. 214, no. May, p. 112994, 2020.

[42] T. Latychevskaia *et al.*, "Convergent beam electron di ff raction of multilayer Van der Waals structures," *Ultramicroscopy*, vol. 212, no. October 2019, p. 112976, 2020.

[43] S. Das *et al.*, "Observation of room-temperature polar skyrmions," *Nature*, vol. 568, no. 7752, pp. 368–372, 2019.

[44] A. K. Yadav *et al.*, "Spatially resolved steady-state negative capacitance," *Nature*, 2019.

[45] S. Im *et al.*, "Direct determination of structural heterogeneity in metallic glasses using

four-dimensional scanning transmission electron microscopy," *Ultramicroscopy*, vol. 195, no. September, pp. 189–193, 2018.

[46] R. H. T. Bates and J. M. Rodenburg, "Sub-ångström transmission microscopy: A fourier transform algorithm for microdiffraction plane intensity information," *Ultramicroscopy*, vol. 31, no. 3, pp. 303–307, 1989.

[47] M. Road and N. Zealand, "The theory of super-resolution electron microscopy via Wigner-distribution deconvolution," *Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci.*, vol. 339, no. 1655, pp. 521–553, 1992.

[48] J.M.Rodenburg and R.H.T.Bates, "The Theory of Super-Resolution Electron Microscopy via Wigner-Distribution Deconvolution." pp. 521–553, 1992.

[49] J. M. Rodenburg, B. . Mc Callum, and P. D. Nellist, "Experimental via STEM tests on double resolution coherent imaging," *Ultramicroscopy*, vol. 48, pp. 304–314, 1993.

[50] P. D. Nellist, B. C. McCallum, and J. M. Rodenburg, "Resolution beyond the `information limit' in transmission electron microscopy," *Nature*, vol. 374, no. 6523. pp. 630–632, 1995.

[51] H. Yang *et al.*, "Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures," *Nat. Commun.*, vol. 7, p. 12532, 2016.

[52] M. J. J. Humphry, B. Kraus, a. C. C. Hurst, a. M. M. Maiden, and J. M. Rodenburg, "Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging," *Nat. Commun.*, vol. 3, p. 730, 2012.

[53] A. M. Maiden and J. M. Rodenburg, “An improved ptychographical phase retrieval algorithm for diffractive imaging,” *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009.

[54] S. Marchesini *et al.*, “SHARP: a distributed, GPU-based ptychographic solver,” pp. 57–63, 2016.

[55] T. J. Pennycook, A. R. Lupini, H. Yang, M. F. Murfitt, L. Jones, and P. D. Nellist, “Efficient phase contrast imaging in STEM using a pixelated detector. Part 1: Experimental demonstration at atomic resolution,” *Ultramicroscopy*, vol. 151, pp. 160–167, 2015.

[56] H. Yang, T. J. Pennycook, and P. D. Nellist, “Efficient phase contrast imaging in STEM using a pixelated detector. Part II: Optimisation of imaging conditions,” *Ultramicroscopy*, vol. 151, pp. 232–239, 2015.

[57] H. G. Brown *et al.*, “Structure Retrieval at Atomic Resolution in the Presence of Multiple Scattering of the Electron Probe,” *Phys. Rev. Lett.*, vol. 121, no. 26, p. 266102, Dec. 2018.

[58] K. Müller-Caspary *et al.*, “Atomic electric fields revealed by a quantum mechanical approach to electron picodiffraction,” *Nat. Commun.*, vol. 5, p. 5653, 2014.

[59] K. Müller-Caspary *et al.*, “Measurement of atomic electric fields and charge densities from average momentum transfers using scanning transmission electron microscopy,” *Ultramicroscopy*, vol. 178, pp. 62–80, Jul. 2017.

[60] K. Müller-Caspary *et al.*, “Electrical Polarization in AlN/GaN Nanodisks Measured by Momentum-Resolved 4D Scanning Transmission Electron Microscopy,” *Phys. Rev. Lett.*,

vol. 122, no. 10, p. 106102, 2019.

[61] W. Gao *et al.*, “Real-space charge-density imaging with sub-ångström resolution by four-dimensional electron microscopy,” *Nature*, vol. 575, no. November, 2019.

[62] T. Matsumoto *et al.*, “Jointed magnetic skyrmion lattices at a small-angle grain boundary directly visualized by advanced electron microscopy,” *Sci. Rep.*, vol. 6, no. April, p. 35880, 2016.

[63] N. Shibata *et al.*, “Direct Visualization of Local Electromagnetic Field Structures by Scanning Transmission Electron Microscopy,” *Acc. Chem. Res.*, vol. 50, no. 7, pp. 1502–1512, 2017.

[64] N. Shibata *et al.*, “Atomic resolution electron microscopy in a magnetic field free environment,” *Nat. Commun.*, vol. 10, no. 1, pp. 1–5, 2019.

[65] “FEI Titan 80-200 Aberation Corrected (S)TEM/EDS/EELS – Wisconsin Centers for Nanoscale Technology – UW–Madison.” [Online]. Available: https://wcnt.wisc.edu/microscopy/transmission-electron-microscopy/fei-titan-aberration-corrected-stem/. [Accessed: 31-Jul-2020].

[66] C. K. Liang, L. W. Chang, and H. H. Chen, “Analysis and compensation of rolling shutter effect,” *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1323–1330, 2008.

[67] B. Bammes, C. Zhang, R. Bilhorn, and P. M. Voyles, “Effectively Synchronizing 4D-STEM Detectors with Probe Movement,” *Microsc. Microanal.*, vol. 25, no. S2, pp. 68–69, 2019.

[68] E. J. Kirkland, *Advanced computing in electron microscopy: Second edition*. 2010.

[69] J. M. Lebeau, S. D. Findlay, L. J. Allen, and S. Stemmer, "Quantitative atomic resolution scanning transmission electron microscopy," *Phys. Rev. Lett.*, vol. 100, no. 20, pp. 1–4, 2008.

[70] A. Pryor, C. Ophus, and J. Miao, "A Streaming Multi-GPU Implementation of Image Simulation Algorithms for Scanning Transmission Electron Microscopy," *Adv. Struct. Chem. Imaging*, 2017.

[71] C. Ophus, "A Fast Image Simulation Algorithm for Scanning Transmission Electron Microscopy," *Adv. Struct. Chem. Imaging*, pp. 1–11, 2017.

[72] M. L. Odlyzko, B. Himmetoglu, M. Cococcioni, and K. A. Mkhoyan, "Atomic bonding effects in annular dark field scanning transmission electron microscopy. I. Computational predictions," *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.*, vol. 34, no. 4, p. 041602, 2016.

[73] M. L. Odlyzko, J. T. Held, and K. A. Mkhoyan, "Atomic bonding effects in annular dark field scanning transmission electron microscopy. II. Experiments," *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.*, vol. 34, no. 4, p. 041603, 2016.

[74] H. Kohl and H. Rose, "Theory of Image Formation by Inelastically Scattered Electrons in the Electron Microscope," 1985, pp. 173–227.

[75] L. J. Allen, A. J. D'Alfonso, and S. D. Findlay, "Modelling the inelastic scattering of fast electrons," *Ultramicroscopy*, vol. 151, pp. 11–22, 2015.

[76] K. A. Mkhoyan, S. E. MacCagnano-Zacher, M. G. Thomas, and J. Silcox, "Critical role of inelastic interactions in quantitative electron microscopy," *Phys. Rev. Lett.*, vol. 100, no.

2, pp. 1–4, 2008.

[77] J. Barthel, M. Cattaneo, B. G. Mendis, S. D. Findlay, and L. J. Allen, “Angular dependence of fast-electron scattering from materials,” *Phys. Rev. B*, vol. 101, no. 18, p. 184109, May 2020.

[78] Y. Yu *et al.*, “Metastable Intermediates in Amorphous Titanium Oxide: A Hidden Role Leading to Ultra-Stable Photoanode Protection,” *Nano Lett.*, vol. 18, no. 8, pp. 5335–5342, 2018.

[79] K. Zhao *et al.*, “Surface Gradient Ti-Doped MnO 2 Nanowires for High-Rate and Long-Life Lithium Battery,” *ACS Appl. Mater. Interfaces*, vol. 10, no. 51, pp. 44376–44384, 2018.

[80] C. Sun *et al.*, “Interconnected Vertically Stacked 2D-MoS2 for Ultrastable Cycling of Rechargeable Li-Ion Battery,” *ACS Appl. Mater. Interfaces*, vol. 11, no. 23, pp. 20762–20769, 2019.

[81] J. Li, S. Han, C. Zhang, W. Wei, M. Gu, and L. Meng, “High-Performance and Reactivation Characteristics of High-Quality, Graphene-Supported SnS2 Heterojunctions for a Lithium-Ion Battery Anode,” *ACS Appl. Mater. Interfaces*, pp. 3–11, 2019.

[82] L. Liu *et al.*, “Examination of B in the Mo solid solution (Mo ss) in Mo ss + Mo 5 SiB 2 + Mo 2 B alloys,” *Scr. Mater.*, vol. 163, pp. 62–65, 2019.

[83] L. Liu, C. Shi, C. Zhang, P. M. Voyles, J. H. Fournelle, and J. H. Perepezko, “Microstructure, microhardness and oxidation behavior of Mo-Si-B alloys in the Moss+Mo2B+Mo5SiB2 three phase region,” *Intermetallics*, vol. 116, no. August 2019, p.

106618, 2020.

[84] Y. Wang *et al.*, “Bioinspired Synthesis of Quasi-Two-Dimensional Monocrystalline Oxides,” *Chem. Mater.*, 2019.

[85] G. Yan *et al.*, “Enhanced Ferromagnetism from Organic-Cerium Oxide Hybrid Ultrathin Nanosheets,” *ACS Appl. Mater. Interfaces*, vol. 11, no. 47, pp. 44601–44608, 2019.

[86] P. J. Strohbeen *et al.*, “Electronically enhanced layer buckling and Au-Au dimerization in epitaxial LaAuSb films,” *Phys. Rev. Mater.*, vol. 3, no. 2, pp. 1–6, 2019.

[87] D. Du *et al.*, “Control of polymorphism during epitaxial growth of hyperferroelectric candidate LiZnSb on GaSb (111)B,” no. 111, pp. 1–6, 2020.

[88] E. H. Shourov *et al.*, “Semi adsorption-controlled growth window for half Heusler FeVSb epitaxial films,” no. 001, pp. 1–7, 2020.

[89] D. Lee *et al.*, “Emergence of room-temperature ferroelectricity at reduced dimensions,” *Science (80-.).*, vol. 349, no. 6254, pp. 1314–1317, Sep. 2015.

[90] Y. L. Tang, Y. L. Zhu, and X. L. Ma, “On the benefit of aberration-corrected HAADF-STEM for strain determination and its application to tailoring ferroelectric domain patterns,” *Ultramicroscopy*, vol. 160, pp. 57–63, 2015.

[91] A. Y. Borisevich *et al.*, “Mapping octahedral tilts and polarization across a domain wall in BiFeO3 from Z-contrast scanning transmission electron microscopy image atomic column shape analysis,” *ACS Nano*, vol. 4, no. 10, pp. 6071–6079, 2010.

[92] Y. Wang, U. Salzberger, W. Sigle, Y. Eren Suyolcu, and P. A. van Aken, “Oxygen octahedra picker: A software tool to extract quantitative information from STEM images,”

Ultramicroscopy, vol. 168, pp. 46–52, 2016.

[93] T. Nilsson Pingel, M. Jørgensen, A. B. Yankovich, H. Grönbeck, and E. Olsson, "Influence of atomic site-specific strain on catalytic activity of supported nanoparticles," *Nat. Commun.*, vol. 9, no. 1, 2018.

[94] J. Hwang, J. Y. Zhang, A. J. D'Alfonso, L. J. Allen, and S. Stemmer, "Three-Dimensional Imaging of Individual Dopant Atoms in <math display="inline"> <msub> <mi>SrTiO</mi> <mn>3</mn> </msub> </math>," *Phys. Rev. Lett.*, vol. 111, no. 26, p. 266101, 2013.

[95] J. Feng *et al.*, "Imaging of Single La Vacancies in LaMnO$_{3}$," *Microsc. Microanal.*, vol. 22, no. S3, pp. 902–903, Nov. 2017.

[96] J. Feng, A. V. Kvit, A. B. Yankovich, C. Zhang, D. Morgan, and M. Paul, "Prospects for Detecting Single Vacancies by Quantitative Scanning Transmission Electron Microscopy," vol. 21, no. 0942, pp. 1887–1888, 2015.

[97] J. Feng, A. V. Kvit, C. Zhang, D. Morgan, and P. M. Voyles, "Three-Dimensional Imaging of Single La Vacancies in LaMnO3," *Microsc. Microanal.*, vol. 22, no. Suppl 3, pp. 902–903, 2016.

[98] K. Kimoto, T. Asaka, X. Yu, T. Nagai, Y. Matsui, and K. Ishizuka, "Local crystal structure analysis with several picometer precision using scanning transmission electron microscopy," *Ultramicroscopy*, vol. 110, pp. 778–782, Jun. 2010.

[99] B. Berkels, P. Binev, D. a. Blom, W. Dahmen, R. C. Sharpley, and T. Vogt, "Optimized imaging using non-rigid registration," *Ultramicroscopy*, vol. 138, pp. 46–56, Nov. 2013.

[100] A. B. Yankovich *et al.*, “Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts.,” *Nat. Commun.*, vol. 5, no. May, p. 4155, Jan. 2014.

[101] N. Braidy, Y. Le Bouar, S. Lazar, and C. Ricolleau, “Correcting scanning instabilities from images of periodic structures,” *Ultramicroscopy*, vol. 118, pp. 67–76, 2012.

[102] L. Jones and P. D. Nellist, “Identifying and correcting scan noise and drift in the scanning transmission electron microscope.,” *Microsc. Microanal.*, vol. 19, no. 4, pp. 1050–60, 2013.

[103] L. Jones *et al.*, “Smart Align—a new tool for robust non-rigid registration of scanning microscope data,” *Adv. Struct. Chem. Imaging*, vol. 1, no. 1, p. 8, 2015.

[104] X. Sang and J. M. LeBeau, “Revolving scanning transmission electron microscopy: Correcting sample drift distortion without prior knowledge,” *Ultramicroscopy*, vol. 138, pp. 28–35, 2014.

[105] C. Ophus, J. Ciston, and C. T. Nelson, “Correcting nonlinear drift distortion of scanning probe and scanning transmission electron microscopies from image pairs with orthogonal scan directions,” *Ultramicroscopy*, vol. 162, pp. 1–9, Mar. 2016.

[106] S. Ning *et al.*, “Scanning distortion correction in STEM images,” *Ultramicroscopy*, vol. 184, pp. 274–283, 2018.

[107] L. Jones *et al.*, “Optimising multi-frame ADF-STEM for high-precision atomic-resolution strain mapping,” *Ultramicroscopy*, vol. 179, pp. 57–62, Aug. 2017.

[108] O. L. Krivanek, N. Dellby, and A. R. Lupini, “Towards sub-Å electron beams,”

Ultramicroscopy, vol. 78, pp. 1–11, Jun. 1999.

[109] H. S. Von Harrach, P. Dona, B. Freitag, H. Soltau, A. Niculae, and M. Rohde, "An integrated Silicon Drift Detector System for FEI Schottky Field," *Microsc. Microanal.*, vol. 15, no. Suppl 2, pp. 8–9, 2009.

[110] P. Wold, S; Esbensen, K; Geladi, "Principle component analysis," *Chemom. Intell. Lab. Syst.*, vol. 2, pp. 37–52, 1987.

[111] N. Bonnet, N. Bonnet, N. Brun, N. Brun, C. Colliex, and C. Colliex, "Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis," *Ultramicroscopy*, vol. 77, pp. 97–112, 1999.

[112] T. Yaguchi, M. Konno, T. Kamino, and M. Watanabe, "Observation of three-dimensional elemental distributions of a Si device using a 360 degree-tilt FIB and the cold field-emission STEM system," *Ultramicroscopy*, vol. 108, no. 12, pp. 1603–1615, 2008.

[113] N. Mevenkamp, A. B. Yankovich, P. M. Voyles, and B. Berkels, "Non-local Means for Scanning Transmission Electron Microscopy Images and Poisson Noise based on Adaptive Periodic Similarity Search and Patch Regularization," in *Vision, Modeling, and Visualization*, 2014, pp. 63–70.

[114] N. Mevenkamp, P. Binev, W. Dahmen, P. M. Voyles, A. B. Yankovich, and B. Berkels, "Poisson noise removal from high-resolution STEM images based on periodic block matching," *Adv. Struct. Chem. Imaging*, vol. 1, no. 1, p. 3, Dec. 2015.

[115] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal Transform-Domain Filter for Volumetric Data Denoising and Reconstruction," *IEEE Trans. Image Process.*,

vol. 22, no. 1, pp. 119–133, Jan. 2013.

[116] J. Salmon, Z. Harmany, C. A. Deledalle, and R. Willett, “Poisson noise reduction with non-local PCA,” *J. Math. Imaging Vis.*, vol. 48, no. 2, pp. 279–294, 2014.

[117] T. Kanungo *et al.*, “efficient k-means clustering algorithm: Analysis and implementation, IEEE Trans,” *Patt. Anal. Mach. In*, vol. tell24, no. 7, pp. 881–892, 2002.

[118] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, “Poisson Noise Reduction with Non-local PCA,” *J. Math. Imaging Vis.*, vol. 48, no. 2, pp. 279–294, Feb. 2014.

[119] A. Bagnall and G. Janacek, “Clustering time series with clipped data,” *Mach. Learn.*, vol. 58, no. 2–3, pp. 151–178, 2005.

[120] A. A. Markowicz and R. E. Van Grieken, “Composition dependence of bremsstrahlung background in electron-probe x-ray microanalysis,” *Anal. Chem.*, vol. 56, no. 12, pp. 2049–2051, 1984.

[121] M. R. Keenan and P. G. Kotula, “Accounting for Poisson noise in the multivariate analysis of ToF-SIMS spectrum images,” *Surf. Interface Anal.*, vol. 36, no. 3, pp. 203–212, 2004.

[122] G. Lucas, P. Burdet, M. Cantoni, and C. H??bert, “Multivariate statistical analysis as a tool for the segmentation of 3D spectral data,” *Micron*, vol. 52–53, pp. 49–56, 2013.

[123] K. Dabov, A. Foi, and V. Katkovnik, “Image denoising by sparse 3D transformation-domain collaborative filtering,” *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 1–16, 2007.

[124] R. Kilaas, “Optimal and near-optimal filters in high-resolution electron microscopy,” *J.*

Microsc., vol. 190, no. 1–2, pp. 45–51, 1998.

[125] H. Du, "A nonlinear filtering algorithm for denoising HR(S)TEM micrographs," *Ultramicroscopy*, vol. 151, pp. 62–67, 2015.

[126] B. Berkels and B. Wirth, "Joint denoising and distortion correction of atomic scale scanning transmission electron microscopy images," *Inverse Probl.*, vol. 33, no. 9, p. 095002, Sep. 2017.

[127] S. Bals, S. Van Aert, G. Van Tendeloo, and D. Ávila-Brande, "Statistical estimation of atomic positions from exit wave reconstruction with a precision in the picometer range," *Phys. Rev. Lett.*, vol. 96, no. 9, p. 096106, Mar. 2006.

[128] J. P. Buban, Q. Ramasse, B. Gipson, N. D. Browning, and H. Stahlberg, "High-resolution low-dose scanning transmission electron microscopy," *J. Electron Microsc. (Tokyo).*, vol. 59, no. 2, pp. 103–112, 2010.

[129] J. M. LeBeau, S. D. Findlay, L. J. Allen, and S. Stemmer, "Position averaged convergent beam electron diffraction: theory and applications.," *Ultramicroscopy*, vol. 110, no. 2, pp. 118–25, 2010.

[130] J. M. LeBeau, A. J. D'Alfonso, N. J. Wright, L. J. Allen, and S. Stemmer, "Determining ferroelectric polarity at the nanoscale," *Appl. Phys. Lett.*, vol. 98, no. 5, p. 052904, Jan. 2011.

[131] J. Hwang, J. Y. Zhang, J. Son, and S. Stemmer, "Nanoscale quantification of octahedral tilts in perovskite films," *Appl. Phys. Lett.*, vol. 100, no. 19, pp. 2–5, 2012.

[132] C. Ophus, P. Ercius, M. Huijben, and J. Ciston, "Non-spectroscopic composition

measurements of SrTiO3-La0.7Sr0.3MnO3multilayers using scanning convergent beam electron diffraction," *Appl. Phys. Lett.*, vol. 110, no. 6, pp. 0–4, 2017.

[133] J. A. Pollock, M. Weyland, D. J. Taplin, L. J. Allen, and S. D. Findlay, "Accuracy and precision of thickness determination from position-averaged convergent beam electron diffraction patterns using a single-parameter metric," *Ultramicroscopy*, vol. 181, pp. 86–96, 2017.

[134] C. Ophus *et al.*, "Non-spectroscopic composition measurements of SrTiO3-La0.7Sr0.3MnO3 multilayers using scanning convergent beam electron diffraction," vol. 063102, pp. 0–4, 2017.

[135] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012.

[136] Y. Kim, "Convolutional Neural Networks for Sentence Classification," 2014.

[137] M. Ziatdinov *et al.*, "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations," *ACS Nano*, vol. 11, no. 12, pp. 12742–12752, Dec. 2017.

[138] T. Xie and J. C. Grossman, "Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties," *Phys. Rev. Lett.*, vol. 120, no. 14, p. 145301, 2018.

[139] D. M. Wei Li, Kevin Field, "Automated defect analysis in electron microscopy images," *Submitt. Rev.*, no. February, pp. 1–9, 2018.

[140] W. Xu and J. M. LeBeau, “A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns,” *Ultramicroscopy*, vol. 188, pp. 59–69, May 2018.

[141] C. Dwyer, R. Erni, and J. Etheridge, “Measurement of effective source distribution and its importance for quantitative interpretation of STEM images,” *Ultramicroscopy*, vol. 110, no. 8, pp. 952–957, 2010.

[142] C. Dwyer, R. Erni, and J. Etheridge, “Method to measure spatial coherence of subangstrom electron beams,” *Appl. Phys. Lett.*, vol. 93, no. 2, p. 021115, 2008.

[143] E. J. Kirkland, “Computation in electron microscopy,” *Acta Crystallogr. Sect. A Found. Adv.*, vol. 72, no. 1, pp. 1–27, 2016.

[144] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1717–1724, 2014.

[145] L. K. A. I. Hansen and P. Salamon, “Neural Network Ensembles,” vol. 12, no. October, pp. 993–1001, 1990.

[146] L. D. Le Cun Jackel, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, B. Le Cun, J. Denker, and D. Henderson, “l,” *Adv. Neural Inf. Process. Syst.*, pp. 396–404, 1990.

[147] G. E. G. Ap, S. H. M. Inima, J. Nocedal, P. Tak, and P. Tang, “ON LARGE-BATCH TRAINING FOR DEEP LEARNING: GENERALIZATION GAP AND SHARP MINIMA,” pp. 1–16, 2017.

[148] S. Maeda, “A Bayesian encourages dropout,” pp. 1–9, 2015.

[149] M. Zeiler, “ADADELTA: AN ADAPTIVE LEARNING RATE METHOD.”

[150] J. Martens and G. Hinton, “On the importance of initialization and momentum in deep learning,” no. 2010, 2012.

[151] M. Nord, R. W. H. Webster, K. A. Paton, S. Mcvitie, D. Mcgrouther, and G. W. Paterson, “Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part I: Data Acquisition, Live Processing and Storage,” no. Ccd, pp. 1–14, 2019.

[152] M. W. Tate *et al.*, “High Dynamic Range Pixel Array Detector for Scanning Transmission Electron Microscopy,” *Microsc. Microanal.*, no. 2014, pp. 1–13, 2016.

[153] G. McMullan, A. R. Faruqi, and R. Henderson, *Direct Electron Detectors*, 1st ed., vol. 579. Elsevier Inc., 2016.

[154] G. McMullan, A. R. Faruqi, D. Clare, and R. Henderson, “Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy,” *Ultramicroscopy*, vol. 147, pp. 156–163, Dec. 2014.

[155] P. M. Voyles, “Informatics and data science in materials microscopy,” *Curr. Opin. Solid State Mater. Sci.*, vol. 21, no. 3, pp. 141–158, Jun. 2017.

[156] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 2002.

[157] M. Collins and R. E. Schapire, “A Generalization of Principal Component Analysis to the Exponential Family,” *Adv. Neural Inf. Process. Syst.*, no. 1, 2001.

[158] M. Watanabe, E. Okunishi, and K. Ishizuka, “Analysis of Spectrum-Imaging Datasets in Atomic-Resolution Electron Microscopy,” *Microsc. Anal.*, vol. 23, no. 7, pp. 5–7, 2009.

[159] “MSA.” [Online]. Available: http://temdm.com/web/msa/. [Accessed: 02-Jun-2020].

[160] F. de la Peña *et al.*, “hyperspy/hyperspy: HyperSpy v1.5.2,” Sep. 2019.

[161] A. Hyvarinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural Networks*, vol. 13, pp. 411–430, Feb. 2000.

[162] G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, and S. M. Smith, “Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers,” *Neuroimage*, vol. 90, pp. 449–468, 2014.

[163] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, “Algorithms and applications for approximate nonnegative matrix factorization,” *Comput. Stat. Data Anal.*, vol. 52, no. 1, pp. 155–173, 2007.

[164] Y. Chen, T. Z. Huang, X. Le Zhao, and L. J. Deng, “Hyperspectral image restoration using framelet-regularized low-rank nonnegative matrix factorization,” *Appl. Math. Model.*, vol. 63, pp. 128–147, 2018.

[165] M. Ye, Y. Qian, and J. Zhou, “Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2621–2639, 2015.

[166] A. Karami, M. Yazdi, and A. Zolghadre Asli, “Noise reduction of hyperspectral images using kernel non-negative tucker decomposition,” *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 3, pp. 487–493, 2011.

[167] W. He, H. Zhang, and L. Zhang, “Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing,” *IEEE Trans. Geosci.*

Remote Sens., vol. 55, no. 7, pp. 3909–3921, 2017.

[168] a. Buades and B. Coll, “A non-local algorithm for image denoising,” *Comput. Vis. Pattern*, vol. 2, no. 0, pp. 60–65, 2005.

[169] A. B. Yankovich *et al.*, “Non-rigid registration and non-local principle component analysis to improve electron microscopy spectrum images,” *Nanotechnology*, vol. 27, no. 36, p. 364001, 2016.

[170] J. Salmon, C. Deledalle, R. Willett, and Z. Harmany, “Poisson noise reduction with non-local PCA,” *2012 IEEE Int. Conf. Acoust. Speech Signal Process.*, no. 5, pp. 1109–1112, Mar. 2012.

[171] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, “Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms,” *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, 2012.

[172] J. Chen and Y. Saad, “On the tensor svd and the optimal low rank orthogonal approximation of tensors,” *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 4, pp. 1709–1734, 2008.

[173] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, “Novel methods for multilinear data completion and de-noising based on tensor-SVD,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3842–3849, 2014.

[174] A. Zhang and R. Han, “Optimal Sparse Singular Value Decomposition for High-Dimensional High-Order Data,” *J. Am. Stat. Assoc.*, vol. 114, no. 528, pp. 1708–1725, 2019.

[175] R. Han, R. Willett, and A. Zhang, “An Optimal Statistical and Computational Framework for Generalized Tensor Estimation,” 2020.

[176] A. Zhang and D. Xia, “Tensor SVD: Statistical and Computational Limits,” *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7311–7338, 2018.

[177] L. Tucker, “SOME MATHEMATICAL NOTES ON THREE-MODE FACTOR ANALYSIS,” *Z. Psychol.*, vol. 31, no. 3, pp. 279–311, 1966.

[178] L. D. E. Lathauwer, B. D. E. Moor, and J. Vandewalle, “On The Best Rank-1 and Rank-(R1, R2, ..., Rn) Approximation of Higher-Order Tensors,” vol. 21, no. 4, pp. 1324–1342, 2000.

[179] F. I. Allen, M. Watanabe, Z. Lee, N. P. Balsara, and A. M. Minor, “Chemical mapping of a block copolymer electrolyte by low-loss EFTEM spectrum-imaging and principal component analysis,” *Ultramicroscopy*, vol. 111, no. 3, pp. 239–244, 2011.

[180] P. Cueva, R. Hovden, J. A. Mundy, H. L. Xin, and D. A. Muller, “M icroscopy M icroanalysis Data Processing for Atomic Resolution Electron Energy Loss Spectroscopy,” *Microsc. Microanal.*, vol. 18, pp. 667–675, 2012.

[181] J. M. Titchmarsh and S. Dumbill, “Multivariate statistical analysis of FEG-STEM EDX spectra,” *J. Microsc.*, vol. 184, no. 3, pp. 195–207, 1996.

[182] X. Li *et al.*, “Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM,” *Nat. Methods*, vol. 10, no. 6, pp. 584–590, 2013.

[183] M. Chu, Y. Sun, U. Aghoram, and S. E. Thompson, “Strain: A solution for higher carrier mobility in nanoscale mosfets,” *Annu. Rev. Mater. Res.*, vol. 39, pp. 203–229, 2009.

[184] A. E. Romanov, T. J. Baker, S. Nakamura, and J. S. Speck, “Strain-induced polarization in wurtzite III-nitride semipolar layers,” *J. Appl. Phys.*, vol. 100, no. 2, 2006.

[185] A. V. Pogrebnyakov *et al.*, “Enhancement of the superconducting transition temperature of MgB2 by a strain-induced bond-stretching mode softening,” *Phys. Rev. Lett.*, vol. 93, no. 14, pp. 1–4, 2004.

[186] M. J. Hÿtch, E. Snoeck, and R. Kilaas, “Quantitative measurement of displacement and strain fields from HREM micrographs,” *Ultramicroscopy*, vol. 74, no. 3, pp. 131–146, 1998.

[187] “GPA for DigitalMicrograph.” [Online]. Available: https://www.hremresearch.com/Eng/plugin/GPAEng.html. [Accessed: 30-Aug-2020].

[188] Y. Zhu, C. Ophus, J. Ciston, and H. Wang, “Interface lattice displacement measurement to 1 pm by geometric phase analysis on aberration-corrected HAADF STEM images,” *Acta Mater.*, vol. 61, no. 15, pp. 5646–5663, 2013.

[189] A. A. Turkin, A. V. Buts, and A. S. Bakai, “Construction of radiation-modified phase diagrams under cascade-producing irradiation: Application to Zr-Nb alloy,” *J. Nucl. Mater.*, vol. 305, no. 2–3, pp. 134–152, 2002.

[190] Y. Tokura and N. Nagaosa, “Nonreciprocal responses from non-centrosymmetric quantum materials,” *Nat. Commun.*, vol. 9, no. 1, 2018.

[191] D. Puggioni, G. Giovannetti, and J. M. Rondinelli, “Polar metals as electrodes to suppress the critical-thickness limit in ferroelectric nanocapacitors,” *J. Appl. Phys.*, vol. 124, no. 17, 2018.

[192] Y. Shi *et al.*, “A ferroelectric-like structural transition in a metal,” *Nat. Mater.*, vol. 12, no. 11, pp. 1024–1027, 2013.

[193] Z. Fei *et al.*, “Ferroelectric switching of a two-dimensional metal,” *Nature*, vol. 560, no. 7718, pp. 336–339, 2018.

[194] D. Du *et al.*, “High electrical conductivity in the epitaxial polar metals LaAuGe and LaPtSb,” *Appl. Phys. Rev.*, vol. 7, no. 12, 2019.

[195] Y. Cao *et al.*, “Unconventional superconductivity in magic-angle graphene superlattices,” *Nature*, vol. 556, no. 7699, pp. 43–50, 2018.

[196] K. Tran *et al.*, “Evidence for moiré excitons in van der Waals heterostructures,” *Nature*, vol. 567, no. 7746, pp. 71–75, 2019.

[197] Y. Liu *et al.*, “Helical van der Waals crystals with discretized Eshelby twist,” *Nature*, vol. 570, no. 7761, pp. 358–362, 2019.

[198] M. J. Bierman, Y. K. A. Lau, A. V. Kvit, A. L. Schmitt, and S. Jin, “Dislocation-driven nanowire growth and Eshelby twist,” *Science (80-.).*, vol. 320, no. 5879, pp. 1060–1063, 2008.

[199] Y. Zhao and S. Jin, “ Controllable Water Vapor Assisted Chemical Vapor Transport Synthesis of WS 2 –MoS 2 Heterostructure ,” *ACS Mater. Lett.*, vol. 2, no. 1, pp. 42–48, 2020.

[200] “HRTEM Filter.” [Online]. Available: http://www.dmscripting.com/hrtem_filter.html. [Accessed: 31-Aug-2020].

[201] J. Ciston *et al.*, “The 4D Camera: Very High Speed Electron Counting for 4D-STEM,”

Microsc. Microanal., vol. 25, no. S2, pp. 1930–1931, 2019.

[202] T. C. Pekin, C. Gammer, J. Ciston, C. Ophus, and A. M. Minor, "In situ nanobeam electron diffraction strain mapping of planar slip in stainless steel," *Scr. Mater.*, vol. 146, pp. 87–90, 2018.

[203] Q. Jeangros *et al.*, "In Situ TEM Analysis of Organic−Inorganic Metal-Halide Perovskite Solar Cells under Electrical Bias," 2016.

[204] S. M. Selbach, T. Tybell, M.-A. Einarsrud, and T. Grande, "The Ferroic Phase Transitions of BiFeO 3," *Adv. Mater.*, vol. 20, no. 19, pp. 3692–3696, Oct. 2008.

[205] S. H. Baek *et al.*, "Ferroelastic switching for nanoscale non-volatile magnetoelectric devices," *Nat. Mater.*, p. 28, 2010.

[206] C. T. Nelson *et al.*, "Domain Dynamics During Ferroelectric Switching," *Science (80-.).*, vol. 334, no. 6058, pp. 968–971, Nov. 2011.

[207] R. F. Klie, Y. Ito, S. Stemmer, and N. D. Browning, "Observation of oxygen vacancy ordering and segregation in Perovskite oxides," *Ultramicroscopy*, vol. 86, no. 3–4, pp. 289–302, 2001.

[208] B. Song *et al.*, "Hollow Electron Ptychographic Diffractive Imaging," *Phys. Rev. Lett.*, vol. 121, no. 14, p. 146101, 2018.

[209] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg, "Ptychographic transmission microscopy in three dimensions using a multi-slice approach," *J. Opt. Soc. Am. A*, vol. 29, no. 8, p. 1606, 2012.

[210] A. R. Lupini, M. Chi, and S. Jesse, "Rapid aberration measurement with pixelated

detectors," *J. Microsc.*, vol. 263, no. 1, pp. 43–50, 2016.

[211] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9351, pp. 234–241.

[212] L. Cavigelli, P. Hager, and L. Benini, *CAS-CNN: A Deep Convolutional Neural Network for Image Compression Artifact Suppression*. 2017.

[213] J. Chen, J. Chen, H. Chao, and M. Yang, "Image Blind Denoising with Generative Adversarial Network Based Noise Modeling," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, vol. 9, no. 2, pp. 3155–3164.

[214] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, 2017.

[215] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.

[216] J. Zhou *et al.*, "Graph Neural Networks: A Review of Methods and Applications," Dec. 2018.

[217] X. Jiang, W. Van Den Broek, and C. T. Koch, "Inverse dynamical photon scattering (IDPS): an artificial neural network based algorithm for three-dimensional quantitative imaging in optical microscopy Abstract :," vol. 24, no. 7, pp. 609–619, 2016.