Universiteit Antwerpen
Faculteit Wetenschappen

Department Fysica

Quantifying atomic structures using neural networks from 4D scanning transmission electron microscopy datasets

Kwantificering van atomaire structuren met behulp van neurale netwerken uit 4D raster transmissie elektronenmicroscopie datasets

Proefschrift voorgelegd tot het behalen van de graad van

Doctor in de Wetenschappen

aan de Universiteit Antwerpen, te verdedigen door

Thomas Friedrich

Antwerpen, 05/2023

Author contact information

Thomas Friedrich
Department of Physics - University of Antwerp
EMAT - Electron Microscopy for Materials Scinece
Groenenborgerlaan 171
B-2020 Antwerp, Belgium
thomas.friedrich@uantwerpen.be

Doctoral commitee

Promotor
Prof. Dr. Sandra Van Aert
Department of Physics (EMAT) - University of Antwerp

Chairman
Prof. Dr. Jan Sijbers
Department of Physics (imec) - University of Antwerp

Members
Prof. Dr. Michiel Wouters
Department of Physics (TQC) - University of Antwerp

Prof. Dr. Joanne Etheridge
Monash Centre for Electron Microscopy - Monash University

Prof. Dr. Knut Müller-Caspary
Department of Chemistry - Ludwig-Maximilians-University Munich

Funding

This thesis is part of a project that has received funding from the European Research Council (ERC) under the Horizon 2020 research and innovation program of the European Union (Grant Agreement No. 770887 PICOMETRICS)

European Research Council
Established by the European Commission

Acknowledgements

Growing up in a rural east-Germany, that was struggling to adapt to a new system during and shortly after reunification at the time, the chances of me (or anyone else in this position) being able to pursue an academic career were slim. Both, educational and economic circumstances were unfavourable for such a path. The fact that I have gotten as far as pursuing a doctoral degree was hard work on one hand, but it would have been impossible without the support of the people of the European Union (EU) and the institutions that represent them. I have been benefitting from different financial support mechanisms and scholarships from the federal republic of Germany and the EU on every step of my academic career until the PhD (Certified technician, Bachelor of Engineering, Master of Science). I am sincerely grateful for all this support over the years by these institutions and to the people of Europe who ultimately paid for it through their taxes. I acknowledge this collective investment into my education and I am committed to use this privilege, as much as possible, in service of the society that enabled it. I am also grateful to all the teachers and mentors who supported and inspired me along the way leading up to my PhD program, in particular: Dipl.-Ing. Frank Uras (FST Mühlhausen), Prof. Allison McMillan (Glyndŵr University), Prof. Steffen Teichert, Prof. Maik Kunert, Dr. Annett Rechtenbach and Arne Bochmann (EAH Jena).

The journey towards completing my PhD during my time in Antwerp was marked by numerous challenges, both academic and non-academic in nature, which in summation had me face arguably one of the most trying times of my life. The support I experienced from my advisor Prof. Dr. Sandra Van Aert was absolutely vital to the completion of my PhD. She has been supportive way beyond the scope of a thesis advisor. I am more than grateful for all the practical advice, constructive feedback and creative input to my work, for the opportunities you enabled for me, for all the freedom I enjoyed developing my scientific interests and career the way I seemed fit and the unwavering support I received every step of the way. The fact that you made my personal success (by my own definition) your priority in a competitive academic environment that rewards peculiar (not necessarily compatible) performance metrics, I do not take for granted. I also want to specifically thank you for the open dialogues, the empathy and the invaluable help while I navigated mental health issues. You are an amazing advisor and a wonderful person. Thank you so much for your guidance!

If I wanted to thank Chu-Ping Yu explicitly for everything I am grateful for, the completion of this section would require (another) deadline extension, so I will restrain myself to a small selection. As a colleague, your creativity, intellect, enthusiasm and love for science have been incredibly inspiring and motivating. Working with you was certainly among the most exciting experiences of my professional life. You are an amazing scientist, in both intellect and integrity. On a personal level, I want to thank you for your invaluable friendship, for all the trips and hikes, the evenings we shared conversation, food and biertjes. Thank you for all the wonderful music and artistic inspiration. Learning from and with you made me grow not only as a scientist, but also as a person. I would not have been able to complete this PhD without you as a friend and colleague. I am forever grateful!

Chu-Ping is also part of larger group of fantastic people I am very happy and proud to be a part of: The PICO band. That is: Chu-Ping Yu (Vocals, beat-boxing, percussion, fake-trumpeteering - crazy multitalent), Christoph Hofer (Drums, Cajon, Percussion - rocks and swings even in 5/4 if needed), Tereza Špičáková (Keyboard/Piano, Vocals - Virtuoso on keys, also probably the only one among us who truly knows what she is doing), Francisco Vega Ibañez (Harmonica, Vocals - his "voice is fucking powerful" [1]), Anna Katia Russi Millán (Management - also sings very beautifully but doesn't want everyone to know just yet) and myself (Guitar). You people are the most incredible group of friends I have the privilege to be a part of. The happiest, funniest, most beautiful and wholesome moments in these last years I spent with and because of you. You are like a family to me.

I am beyond grateful to Anna Trimborn for the many years in which I could, every single day, 100% rely on her support throughout the countless ups and downs. You are the most important person in my life. Beyond the depth of our personal connection, you have also been an inspiration to me in many ways that had an impact on my development as a person and scientist. Your moral integrity and the way you relentlessly keep fighting for social justice and equity has certainly shaped how I think and reflect about privilege, society and my role in it as a scientist. You keep inspiring me (and others) to become a better person. Thank you so much for all the inspiration and for having my back for all these years. It means the world to me.

I am grateful to the many great colleagues (present and past) of EMAT for sharing exciting science, stimulating collaborations and discussions. In particular I would like to acknowledge (in no particular order):

- Dr. Nadine Schrenker for the wonderful collaboration in her study on PbBrFA-NCs. It has been great working with you.
- Dr. Ivan Pedro Hoyos Lobato for the invaluable support, especially during the first year of my PhD. I have learned a lot from you. I appreciate your mentoring efforts and your friendship. Thank you so much.
- Dr. Annick De Backer for all the help with counting atoms, statistics, the many nice discussions in group-meetings and the help with translating the summary of this thesis into Dutch (Dank u wel!).
- Dr. Zezhong Zhang for the many stimulating scientific conversations and your friendship. You are a great scholar and wonderful person.
- Dr. Lukas Grünewald for the scientific discourses and enthusiastic exchanges about open-source software and productivity tools. Also thank you for sharing data and your contributions to [2] and chapter 5 and for pushing the reset button on my PC a thousand times when I crashed it while working remotely.
- Dr. Annelies De wael, Dr. Ece Arslan Irmak, Dr. Yansong Hao, Duygu Gizem Sentürk, Dr. Noopur Jain, Safiyye Kavak and Jinhui Guo for the years of companionship through this PhD-journey, both professionally and personally. I consider myself lucky having colleagues and friends like you.

Acknowledgements

- Prof. Dr. Johan Verbeeck for the collaborations in the riCOM and AIRPI projects. Your creativity, scientific rigor, vision and enthusiasm have been very inspiring.
- Lydia Cassiers, Miek van Look and Koen De Cauwsemaecker for substantial administrative and IT support.
- Gert De Bont for the support with graphic designs and the production of the print of this thesis.
- The newer members in the group I didn't get to work with so much, but truly appreciated: Dr. Milena Hugenschmidt, Dr. Mohammed Noorul Hussain and Tom Stoops
- The commitee members Jan Sijbers, Michiel Wouters, Joanne Etheridge and Knut Müller-Caspary for the interesting discussions and suggestions regarding my PhD work and for the time and efforts they invested.

Samenvatting

Nanowetenschap en nanotechnologie zijn van enorm belang in vele wetenschappelijke gebieden en voor talrijke praktische toepassingen, zoals energieopslag, halfgeleidertechnologie, precisiegeneeskunde, katalyse, milieutechnologie en nog veel meer. De impact van dit vakgebied op de samenleving mag niet onderschat worden. De karakterisering van structuren op nanometer- en atomaire schaal is van essentieel belang om onze kennis van de kenmerken van nanostructuren en de processen die ze beheersen te verbeteren. In deze context speelt transmissie-elektronenmicroscopie een cruciale rol omdat het wetenschappers in staat stelt 3-dimensionale structuren op atomaire resolutie af te beelden (in tegenstelling tot atomaire resolutie 2D-beeldvorming van oppervlakken in scanning probe microscopie bijvoorbeeld). Met name ringvormige donkerveld rastertransmissie-elektronenmicroscopie (HAADF STEM) is de afgelopen decennia uitgebreid gebruikt voor de studie van nanostructuren met hoge resolutie. De contrastmechanismen van deze methode zijn goed begrepen en kunnen worden gebruikt om 3D-structuren te reconstrueren uit een enkele projectie of uit een tomografieserie van beelden opgenomen bij verschillende kijkrichtingen met behulp van speciale algoritmen. Het grootste nadeel van HAADF is misschien wel de relatief slechte dosisefficiëntie. De invallende elektronenbundel interageert met het specimen, maar slechts enkele elektronen, verstrooid naar grotere hoeken, dragen daadwerkelijk bij aan de contrastvorming van het verkregen beeld. Dit is problematisch aangezien het specimen beschadigd kan worden onder de elektronenbundel en de elektronendosis daarom zo laag mogelijk moet zijn om de structuur niet te veranderen door de meting. Dit veelvoorkomend probleem beperkt de signaal-ruisverhouding bij HAADF-beeldvorming vaak ernstig. Een manier om zoveel mogelijk informatie uit deze lagedosisbeelden te halen, is door gebruik te maken van restauratiemethoden om ruis te verwijderen of verschillende vervormingen te corrigeren. Machine learning-algoritmen zijn de facto standaard geworden voor dit soort computervisietaken. In het bijzonder zijn convolutional neural networks (CNN) gebruikt in talloze denoising- en generieke beeldhersteltaken in de afgelopen tien jaar. Het is duidelijk dat het louter creëren van "mooiere" afbeeldingen niet het doel kan zijn als we geïnteresseerd zijn in het verkrijgen van kwantitatief nauwkeurige gegevens. In plaats daarvan moeten de fysische processen van de beeldvorming geanalyseerd en begrepen worden, inclusief de technische complexiteit van de elektronenmicroscoop, de scanmotoren en de detectiesystemen, om fysisch betrouwbare reconstructies te verkrijgen.

Hoofdstuk 2 "STEM Distortion correction" biedt een analyse en modellen om de ruiskenmerken en scanlijnvervormingen in HAADF-STEM-beelden te beschrijven. Deze modellen werden gebruikt om een groot aantal synthetische, onvervormde-vervormde STEM-beeldparen te genereren, die op hun beurt werden gebruikt om een CNN te trainen om deze beelddegradaties te corrigeren. Verder wordt onderzocht en aangetoond dat deze beeldherstelling niet alleen de signaal-ruisverhouding, maar ook de precisie en nauwkeurigheid waarmee atoomkolommen kunnen worden gelokaliseerd, aanzienlijk kan verbeteren. Bovendien wordt aangetoond dat ook het aantal atomen met hogere nauwkeurigheid geteld kan worden. Deze conclusies zijn te veralgemenen, wat betekent dat ook andere analysemethoden profiteren van het gebruik van de reconstructie, bijvoorbeeld de kwaliteit van tomografische reconstructies zou aanzienlijk verbeteren bij

het corrigeren van alle projectiebeelden in een tomografieserie. Er wordt op gewezen dat vooral bij lage dosiswaarden de onderliggende modellen om de trainingsgegevens te genereren, cruciaal zijn om een betrouwbare reconstructie te verkrijgen, wat de implementatie in dit proefschrift onderscheidt van generieke denoising-algoritmen.

Ondanks de voordelen die dit reconstructie-algoritme biedt, blijft er echter een fundamentele nadeel, namelijk dat lage hoekverstrooiing op geen enkele manier bijdraagt aan het beeldcontrast terwijl dit de meerderheid van de invallende elektronen vertegenwoordigt en het specimen dus kunnen beschadigen. Lage hoekverstrooiing lijdt aan coherentie-effecten en is gevoeliger voor lichte elementen, wat ook betekent dat het signaal meer wordt beïnvloed door bijvoorbeeld koolstofverontreiniging. Idealiter zouden we alle elektronen die met het specimen interageren willen gebruiken voor beeldgeneratie, aangezien elk elektron waardevolle informatie bevat. Dit is het grote voordeel van 4D-STEM, waarbij in plaats van alleen een geïntegreerd signaal voor een gegeven hoekbereik te verzamelen, volledige diffractiepatronen worden opgenomen met gepixelde elektronencamera's. Op deze manier kan het volledige signaal binnen het bereik van de camera worden gebruikt, terwijl ook informatie over de verstrooiingsrichting behouden blijft. Dit betekent dat in plaats van te weten "Hoe sterk verstrooit het specimen?", het mogelijk is om de vraag te beantwoorden: "Hoe sterk verstrooit het specimen in een bepaalde richting?". 4D-STEM is daarom een zeer krachtige techniek, maar het lijdt onder het feit dat de datasets zeer groot zijn en daarom langzaam te verwerken zijn. Het live gebruik tijdens het bedienen van de microscoop is daarom moeilijk of onmogelijk.

Hoofdstuk 3 "Real-time integration centre of mass (riCOM)" introduceert een conventioneel algoritme gebaseerd op het zwaartepunt (COM) van de diffractiepatronen. De berekening van het COM kan zeer snel zijn en de radiale integraal van het COM-signaal (riCOM) geeft aanwijzingen over de geprojecteerde atomaire potentiaal. Het hoofdstuk presenteert een zeer efficiënte implementatie van het algoritme, dat ook veel flexibiliteit biedt wat betreft de frequentietransfer door aanpassing van het integratiebereik en het opnemen van instelbare filters in het liveproces. De methode is daarom zeer geschikt voor livebeeldvorming, maar ook voor snelle en gemakkelijke nabewerkingen van opgenomen 4D-STEM-gegevens. Dit algoritme is geïmplementeerd in een performante, gebruiksvriendelijke GUI-applicatie, inclusief interfaces voor enkele camera's en gegevenstypen en open source gepubliceerd.

Hoofdstuk 4 "AI-assisted rapid phase imaging (AIRPI)" behandelt hetzelfde fundamentele probleem op een andere manier. Geïnspireerd door ptychografische fase reconstructiemethoden, proberen we in dit project, genaamd "AIRPI - AI ondersteunde snelle fasebeeldvorming", fase informatie op een scanpunt te reconstrueren op basis van een set diffractiepatronen van overlappende probe-posities door middel van conventionele neurale netwerken. De ontwikkeling van de trainingsdataset, het machine learning-model en het reconstructie-algoritme worden gedocumenteerd en een aantal reconstructies wordt gepresenteerd. De resultaten suggereren een vrij unieke contrastkarakteristiek die hoge resolutie mogelijk maakt, die in principe alleen wordt beperkt door de opnamehoek van de 4D-STEM-camera, en tegelijkertijd lage frequentiecomponenten ophaalt die indicatief zijn voor de dikte van het specimen. De gevoeligheid voor het atoomnummer Z werd ook onderzocht en lijkt gunstig te zijn in die zin dat zeer lichte en zeer

zware atomen tegelijkertijd en in zekere mate kwantitatief kunnen worden afgebeeld. De prestaties van deze methode worden gedemonstreerd in simulatiestudies en op een aantal experimentele datasets. Vergeleken met riCOM is het nog steeds aanzienlijk meer rekenintensief en het is beperkt door het bereik van trainingsgegevens en sommige modelaannames, die de prestaties vooral beïnvloeden in de aanwezigheid van ongecorrigeerde aberraties.

Het **hoofdstuk 5 "Conclusions"** vat de bevindingen van de drie hoofdstukken samen en plaatst deze in de context van de state-of-the-art. In het bijzonder worden zowel riCOM als AIRPI vergeleken met andere reconstructiemethoden en met elkaar. De sterke en zwakke punten van elke methode worden vanuit een praktisch perspectief geschetst (voor experimentatoren, in plaats van voor theoretici). Ten slotte worden enkele ideeën gepresenteerd over mogelijke verdere ontwikkelingen om het onderzoek in deze scriptie verder te verbeteren.

Summary

Nanoscience and nanotechnologies are of immense importance across many fields of science and for numerous practical applications, such as energy storage, semiconductor technology, precision medicine, catalysis, environmental technology and many more. The impact that advances in this field will have on societies across the world can hardly be overstated. The characterisation of structures at the nanometre- and atomic scale is a cornerstone to improve our understanding of the characteristics of nanostructures and the processes that govern them. In this context transmission electron microscopy plays a crucial role as it allows scientists to image 3-dimensional structures at atomic resolution (as opposed to atomic resolution 2D imaging of surfaces in scanning probe microscopy for example). Especially scanning transmission electron microscopy (STEM) with high-angle-annular-dark-field-detectors (HAADF) has been used extensively over the last decades for the study of nanostructures with high resolution. The contrast mechanisms of this method are well understood and can be used to reconstruct 3D structures from a single projection or from a tomography series of images at different viewing directions using dedicated algorithms. The perhaps biggest drawback of HAADF is its comparatively poor dose efficiency. The entire beam current interacts with the sample but only few electrons, scattered to larger angles, actually contribute to the contrast formation of the image. This is troubling in the sense that samples may be damaged under the electron beam and the electron dose should therefore be as low as possible to avoid altering the structure by the measurement. This is in fact a very common problem that often limits the signal to noise ratio in HAADF imaging severely. One way to extract as much information as possible from these low-dose images is the use of restoration methods, to remove noise or/and correct for various distortions. Machine learning algorithms have become the de-facto standard for this type of computer-vision task. In particular convolutional neural networks (CNN) were used in countless denoising and generic image restoration tasks over the last decade or so. It is clear that merely creating "nicer" images cannot be the goal if we are interested in obtaining quantitative structure information. Instead, the physical processes involved in the image generation need to be analysed and understood, including the technical intricacies of the electron microscope, scan engines and detector systems if we hope to achieve physically sound reconstructions.

Chapter 2 "STEM Distortion correction" provides an analysis and models to describe the noise characteristics and scan line distortions in HAADF-STEM images. These models were used to generate a large number of synthetic, undistorted-distorted STEM image-pairs, which were in turn used to train a CNN to correct for these image degradations. It is further investigated and shown that this image restoration can significantly enhance, not only the signal-to-noise ratio, but also the precision and accuracy with which atomic columns can be located. Further, it is demonstrated that also atom counting studies benefit from improved accuracy. These conclusions are transferable, meaning that also other analysis approaches will benefit from the use of the reconstruction, e.g. the quality of tomography reconstructions would be expected to improve significantly when restoring all projection images in a tomography-series. It is pointed out, that especially at low doses the underlying models used to generate the training data are crucial for a

faithful restoration, which sets the implementation presented in this thesis apart from generic denoising algorithms.

No restoration changes the fundamental drawback, that low angle scattering does not contribute in any way to the image contrast and the majority of probing electrons do nothing for us but damaging the sample. Low angle scattering suffers from coherency effects and is more sensitive to light elements, which also means the signal is more influenced by carbon contamination for example. Ideally, we would want to use all the electrons interacting with the sample for image generation since each one contains some valuable information. This is the big advantage of 4D-STEM, where instead of collecting only an integrated signal for a given annular range, entire diffraction patterns are recorded with pixelated electron cameras. In this way the entire signal within the collection range of the camera can be used while also retaining information about the scattering direction. This means, instead of knowing "How strongly does the sample scatter?", it is possible to answer the question: "How strongly does the sample scatter into which direction?". 4D-STEM is therefore a very powerful technique but it suffers from the fact that the datasets are very large and slow to process accordingly. Using it live while operating the microscope has been difficult or impossible for reconstruction methods beyond virtual STEM imaging. This forms the basis for the chapters 3 and 4, where two new ways are proposed that enable very dose-efficient live imaging.

Chapter 3 "Real-time integration centre of mass (riCOM)" introduces a conventional algorithm based on the centre of mass (COM) of the diffraction patterns. The computation of the COM can be very fast and the radial integral of the COM signal is indicative of the projected atomic potential. The chapter presents a very efficient implementation of the algorithm, that also allows much flexibility regarding the frequency transfer by adjusting the integration range, as well as including tuneable filters into the live process. The method is therefore very well suited for live imaging, but also for fast and easy post processing of recorded 4D-STEM data. This algorithm was implemented in a performant, user-friendly GUI-application, including interfaces for some cameras and data types and published open source.

Chapter 4 "AI-assisted rapid phase imaging (AIRPI)" treats the same fundamental problem in a very different way. Inspired by ptychographic phase-retrieval methods, in this project, coined "AIRPI - AI assisted rapid phase imaging", we try to reconstruct phase information at a scan point based on a set of diffraction patterns of overlapping probe positions by means of CNNs. The development of the training dataset, the machine learning model and reconstruction algorithm are documented and a number of reconstructions are presented. The results suggest a fairly unique contrast characteristic that enables a high resolution, that is in principle only limited by the collection angle of the 4D-STEM camera and at the same time retrieves low frequency components which are indicative of the sample thickness. The sensitivity to atomic number Z was investigated too and appears to be favourable in the sense that very light and very heavy atoms can be imaged at the same time and to some extent in a quantitative manner. The performance of this method is demonstrated in simulation studies and on a number of experimental datasets as well. Compared to riCOM it is still considerably more computationally expensive and it is limited by the range of training data and some

model assumptions, which affect the performance mainly in the presence of uncorrected aberrations.

The **chapter 5 "Conclusions"** summarises the findings of the three main chapters and establishes a context to the state of the art. In particular, both riCOM and AIRPI are compared with other reconstruction methods and against each other as well. The strengths and weaknesses of each respective method are outlined from a more practical perspective (for the experimentalists, rather than for the theoreticians). Finally some thoughts on possible further developments are presented to improve upon the research presented in this thesis.

Contents

Nomenclature

Abbreviations

ADF	Annular dark field
AI	Artificial intelligence
AIRPI	AI-assisted rapid phase imaging
API	Application programming interface
BF	Bright field
BM3D	Block-matching and 3D filtering
BN	Batch normalisation
CBED	Convergent beam electron diffraction pattern
CCD	Charge-coupled device
CI	Confidence interval
CLI	Command line interface
CNN	Convolutional neural network
CNT	Carbon nanotube
COM	Centre of mass
CPU	Central processing unit
CRLB	Cramér-Rao lower bound
CTEM	Conventional transmission electron microscope/microscopy
CTF	Contrast transfer function
CVP	Chemical vapour depsoition
DF	Dark field
DL	Deep learning
DPC	Differential phase contrast
DWF	Debye Waller Factor
EBSD	Electron backscatter diffraction
EBSP	Electron backscatter diffraction pattern
ED	Electron diffraction
EDX	Energy dispersive X-Ray spectroscopy

EELS	Electron energy-loss spectroscopy
ELU	Exponential linear unit
ePIE	Extended ptychographic iterative engine
EU	European union
FA	Formamidinium
FEG	Field emission gun
FFT	Fast Fourier transform
FOLZ	First order Laue zone
FWHM	Full width at half maximum
GD	Gradient descent
GPL	General public license
GPU	Graphics processing unit
GT	Ground truth
GUI	Graphical user interface
HAADF	High angle annular dark field
HOLZ	Higher order Laue zone
iCOM	Integrated centre of mass
iDPC	Integrated differential phase contrast
LGPL	Lesser general public license
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MIT	Massachusetts institute of technology
ML	Machine learning
N2V	Noise to void (specific neural network)
NBED	Nanobeam electron diffraction
NC	Nanocrystal
NN	Neural network
NP	Nanoparticle
PACBED	Position averaged convergent beam electron diffraction pattern
PDF	Probability density function
PMT	Photo multiplier tube

POA	Phase object approximation
PSNR	Peak signal to noise ratio
ReLU	Rectified linear unit
riCOM	Real-time integrated centre of mass
rmsd	Root mean squared displacement
SCS	(probe integrated) scattering cross section
SEM	Scanning electron microscope/microscopy
SGD	Stochastic gradient descent
SNR	Signal to noise ratio
SOTA	State of the art
SSB	Single sideband ptychography
STEM	Scanning transmission electron microscope/microscopy
TEM	Transmission electron microscope/microscopy
vSTEM	Virtual scanning transmission electrom microscopy
WDD	Wigner distribution deconvolution
WDX	Wavelength dispersive X-Ray spectroscopy
WSL	Windows subsystem for linux
XC	Cross correlation
XRD	X-Ray diffraction
YAP	Yttrium aluminum perovskite

Common operators and symbols

$\circledast$	2D Convolution
$\mathbb{E}(x)$	Expectation value of random variable x
$\mathbb{N}$	Normal distribution
$\mathbb{P}$	Poisson distribution
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	Inverse Fourier transform
∇	Gradient operator
$\odot$	Hadamard product (element-wise product)
$\mathrm{Std}(x)$	Standard deviation of random variable x
$\mathrm{Var}(x)$	Variance of random variable x

$\star$	2D Cross correlation
$\vec{k}$	2D Reciprocal space vector (kx,ky)
$\vec{r}$	2D Real space vector (x,y)

List of Figures

List of Tables

CHAPTER

1 Introduction

Contents

1.1 The bigger picture (of tiny things)

1.1.1 Nanoscience

Nanoscience is a field that investigates the behaviour of materials and processes at the nanometre scale, which ranges from about 1 to 100 nanometres. This scale is so tiny that it is difficult to comprehend. To put it into perspective, figure 1.1 illustrates a series of objects commonly known to people. The tip of a push pin has a size of about one millimetre, which is equivalent to 1,000,000 nanometres. A human hair has a diameter of about 100,000 nanometres, while a red blood cell measures around 7,000 nanometres. A typical bacteria cell has a size of 1-10 micrometres, which is between 10,000 and 100,000 nanometres. These kind of structures have been observable for some time already by means of regular light microscopes, but researchers in both, the life sciences and materials science work with objects and phenomena still orders of magnitude smaller, such as the infamous Covid-19 virus depicted in figure 1.1 at 100nm. Materials science research these days is often concerned even with the sub-nanometre range, studying nanoparticles (NP) and individual atoms. The study of matter at this scale is fascinating because nanomaterials possess unique physical and chemical properties that are not evident in macroscopic systems. When the size of a material is confined to the nanoscale, the surface area to volume ratio increases, which means that surface atoms start to control the properties of nanomaterials. This distinctive feature results in extraordinary properties like increased surface energy, enhanced reactivity, and unique optical and magnetic properties. For instance, nanomaterials such as quantum dots, nanowires, and nanotubes have a high surface-area to volume ratio that enables them to act as excellent sensors or catalysts, making them suitable for applications in electronic devices and medical diagnostics. Carbon nanotubes (CNT), for example, are known for their high tensile strength and thermal conductivity, making them promising materials for building nanoelectronics and energy storage devices. Moreover, nanomaterials can be engineered to have specific functionalities based on their unique surface properties, such as high porosity or selectivity in separating different chemicals. The shape and size of nanoparticles can significantly affect their properties. For instance, small gold particles, which are typically inert, can become effective catalysts when their size is reduced to the nanometre scale. This is because the surface atoms of the nanoparticles become more exposed, allowing them to participate in chemical reactions [3]. Consequently, nanoscience and nanotechnology have been instrumental in driving innovation and progress in the last decades, across numerous fields of technology, enabling the development of new materials, energy storage, and generation systems, medical treatments, and environmental technologies, as well as advanced electronics and data storage devices. Nanoscience therefore plays a crucial role in tackling the challenges of our time and the future. Most notably, climate change and global warming and all its perilous consequences and dangers dictate the urgent need for a transition from fossil-fuelled to carbon-neutral energy sources and mobility technologies. The development of more efficient solar cells [4], and batteries has already, and will continue to substantially support the growth of renewable energy production and plays a crucial part on the way to green electromobility. Nanomaterials are also being used in numerous environmental applications [5], such as water treatment and desalination applications

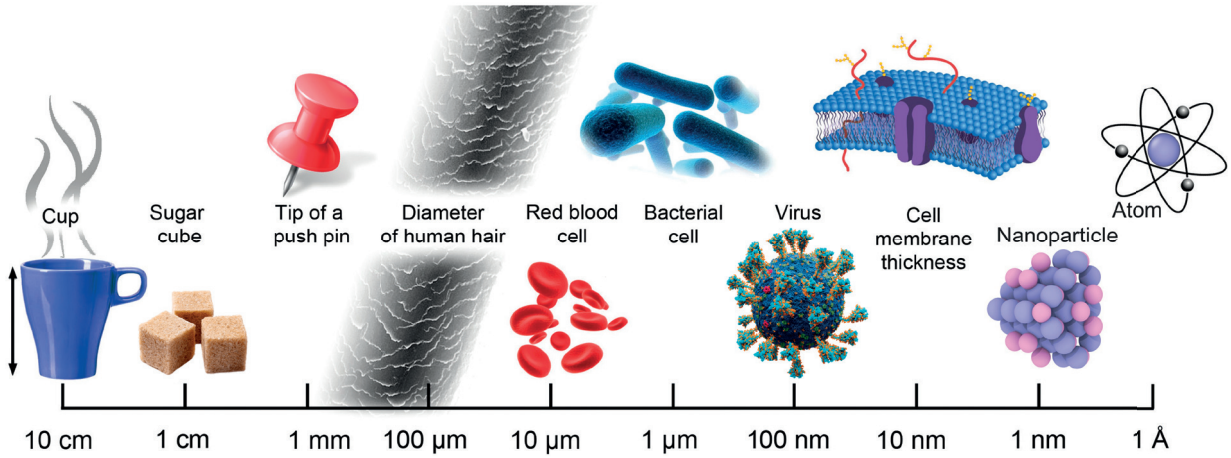


Figure 1.1 | Overview of the scale of different objects from the decimeter to the sub-nanometre scale. [1]

with Zeolites [6, 7], or as catalysts in oxidation and hydrogenation reactions [8]. An arguably less obvious field of nanoscience applications is agricultural science. For example, carbon nanotubes [9] can be used to "engineer" plants and selectively change their characteristics [10, 11]. This holds a large potential to help tackle global food shortages and supply insecurities, which are likely to worsen in the near future in the face or rising global temperatures, as well as political turmoil and conflict. Furthermore, in biomedical applications nano-particles and structures are being used for drug delivery and precision medicine [12]. This can for example be used to selectively target tumors in cancer treatments. All these examples only scratch the surface of the vast amount of applications but serve to show just how immensely important this branch of science is.

1.1.2 Electron microscopy

Being able to characterise these nano-structures is a fundamental cornerstone in the development of novel nano-materials and their synthesis. Optical microscopes are inherently limited by fundamental resolution restrictions arising from the wavelength of the light used. The theoretical diffraction limit on a microscopes resolution was discovered by Ernst Abbe and the popular equation relating the minimum resolvable distance d to a wavelength λ and aperture size Θ is hence known as the Abbe equation:

$$d = \frac{\lambda}{2n\sin\Theta}, \tag{1.1}$$

where n is the refractive index. If we consider a practical limit for Θ of about 70° and the shortest wavelength of visible light to be at $\approx 400nm$ then $d \approx 188nm$. While there are multiple ways to increase the resolution of optical microscopes beyond this example calculation, the fundamental wavelength dependency holds. The wavelength

[1] Illustration adapted/modified from reference [13] with elements from https://upload.wikimedia.org/wikipedia/commons/9/94/Coronavirus._SARS-CoV-2.png, which is published under a CC BY-SA 4.0 license

of electrons on the other hand is determined essentially by their momentum p as established by de Broglie:

$$\lambda = \frac{h}{p} = \frac{h}{m_0 v}, \tag{1.2}$$

where h is Planck's constant and m_0 and v are the electron rest mass and velocity respectively. This implies that the wavelength of an electron can be reduced by increasing its velocity v, wherein lies the basis for superior resolution capabilities of electron microscopes. Since electrons in a TEM are travelling very fast, relativistic effects should be taken into account[14].

$$\lambda = \frac{hc}{\sqrt{eV(2m_0c^2 + eV)}}, \tag{1.3}$$

where eV is the kinetic energy of the electron. This further reduces the wavelength as compared to the description of classical mechanics in equation 1.2. Applying equations 1.1 and 1.2 would suggest that ≈5 kilo volt (kV) of acceleration voltage would be required to reach nanometre resolution. Modern transmission electron microscopes typically work with acceleration voltages of up to 300 kV which would put the resolution according to Abbes equation almost at the Ångstrom (Å) scale, which is 1/10 of a nanometre, which would make objects on the right side of figure 1.1 observable. This is an overly simplistic treatment of the resolution capabilities of course, but it serves to show why electron microscopy is so fundamentally important to nanoscience.

Transmission electron microscopy takes in a unique role in this regard, since structures may be imaged with atomic resolution using high resolution TEM (HRTEM) or scanning transmission electron microscopy (STEM). However, these sophisticated instruments are very versatile and may be used for a variety of atomic scale studies. Beyond creating projection images to gain insight into the arrangements of atomic structures in 2D, it is possible to retrieve 3D information as well, for example by means of tomography reconstructions, which are based on a number of images taken for different projections of the same object [15]. 3D shapes may also be retrieved from single projection images utilising the dependence of the electron scattering intensities at high angles to the sample thickness and atomic numbers [16, 17]. Further, electron microscopes may also be used with spectroscopic methods, such as Electron Energy-Loss Spectroscopy (EELS) or energy-dispersive X-ray analysis (EDX), which may provide access to additional information such as elemental compositions, electron densities, plasmon excitations and bondings [18]. Since electrons interact with magnetic and electric fields they can further be used to probe these fields, which is particularly interesting for the investigation of magnetic textures and skyrmions in materials [19]. Also the collection of full diffraction patterns over a 2D scan grid in STEM, commonly referred to as 4D-STEM, has become more practical and widely available in recent years. All the mentioned modes are increasingly being used also in in-situ experiments, which are conducted in a controlled environment, such as temperature, pressure, radiation exposure, or varying gaseous environments [20]. This allows not only to make observations of a material in a given state, but also to observe dynamic processes that may be triggered by changing the sample environment. This last point in particular highlights the need for very fast imaging processes, as time resolution becomes a critical factor for the observation of dynamic processes. This poses a number of challenges for the development of the

hardware, e.g. scan coils and detector systems, but also calls for highly efficient analysis methods that can actually process these large, multi-dimensional datasets or data-streams and extract the relevant information in a reasonable amount of time and at reasonable amount of computational cost. Spectroscopic in-situ experiments may easily produce tens or hundreds of gigabytes worth of data in a single experiment. Given these challenges and the plethora of imaging modalities and applications mentioned, the data-science behind electron microscopy has become an active field of research in its own.

1.1.3 Thesis Motivation & Outline

The focus of this thesis is on scanning transmission electron microscopy (STEM) and four-dimensional STEM (4D-STEM) techniques. These methods provide a wealth of information, but can be challenging to analyse due to issues such as noise, distortions, and the sheer volume of data. In this work, a set of novel approaches are developed that aim to leverage this information while also considering performance, scalability, and usability. Machine learning models, specifically convolutional neural networks (CNNs), have been shown to be highly efficient in numerous imaging and computer vision tasks. Therefore, this work explores the use of CNNs in STEM image analysis workflows and 4D-STEM data processing pipelines, in addition to developing highly efficient conventional algorithms.

The introduction chapter provides a general overview of scanning electron microscopes, their basic components, and the fundamental physics involved. Additionally, it introduces neural networks and CNNs, the latter being one of the key elements of this thesis.

The second chapter describes prominent distortions and sources of noise in atomic resolution STEM, which can impact atomic structure quantifications. It is shown that these distortions can largely be corrected using a CNN. The benefit of this restoration for atomic column detection and atom counting is investigated, demonstrating that CNN-restoration provides an effective method for improving the accuracy and precision of such studies.

The third and fourth chapters present two novel methods for extracting phase images from 4D-STEM datasets. The first is a highly efficient conventional algorithm, coined "riCOM", which is based on the centre of mass (COM) of diffraction patterns. It works on individual diffraction patterns, granting the method live imaging capability. The fourth chapter presents a different approach to the same problem, where a CNN is used to retrieve the phase image one scan step at a time. This approach is computationally more expensive and slower, but has superior imaging characteristics, including higher resolution and the reconstruction of thickness information, which is a unique feature among phase imaging methods.

The final chapter summarizes the main contributions of this work and compares the developed methods with state-of-the-art (SOTA) techniques for some example datasets. An outlook is provided for how the work presented in this thesis may be useful to the wider electron microscopy and nanoscience research community.

1.2 Scanning transmission electron microscopy

The scanning transmission electron microscope is among the most versatile and powerful instruments for atomic scale analysis at the disposal of the nanoscience community. Crewe et al. were among the first to demonstrate the potential of the STEM by using a field-emission gun (FEG) and an annular-dark-field (ADF) detector to record the first clear image of heavy atoms in 1970 [18, 21]. It has since become a mainstream microscopy mode alongside conventional TEM (CTEM). Most modern instruments are combined systems and are capable of operating in both modes these days. The main difference between the two modalities is that CTEM uses a parallel electron beam to illuminate the sample area, while in STEM a convergent electron beam, forming a small probe in the sample plane, is scanned in a 2D grid over the sample area instead. The basic components of a STEM as outlined in figure 1.2 include an electron gun, a condenser lens, an aperture, an objective lens, and a detector. The electron gun is responsible for producing a beam of electrons, which is focused and directed towards the sample. The most common types of electron guns found in SEMs and STEMs are thermionic, FEG and CFEG (cold FEG) and allow a variety of voltage settings ranging from a few to a few hundred kV. The condenser lens helps shape the electron beam by effectively demagnifying the the electron source, which has a finite size [22]. Apertures control the size of the beam that reaches the sample, by blocking parts of the beam, which also may reduce the beam current. The scan coils, or scanning lenses, are used to deflect the focused electron beam in a controlled way to enable the scanning over the sample, typically in a raster pattern in the x-y plane. However, scanning lenses can be controlled by a computer to generate any desired scanning pattern over the sample, which can be useful to e.g. reduce beam damage [23]. The objective lens is one of the most critical

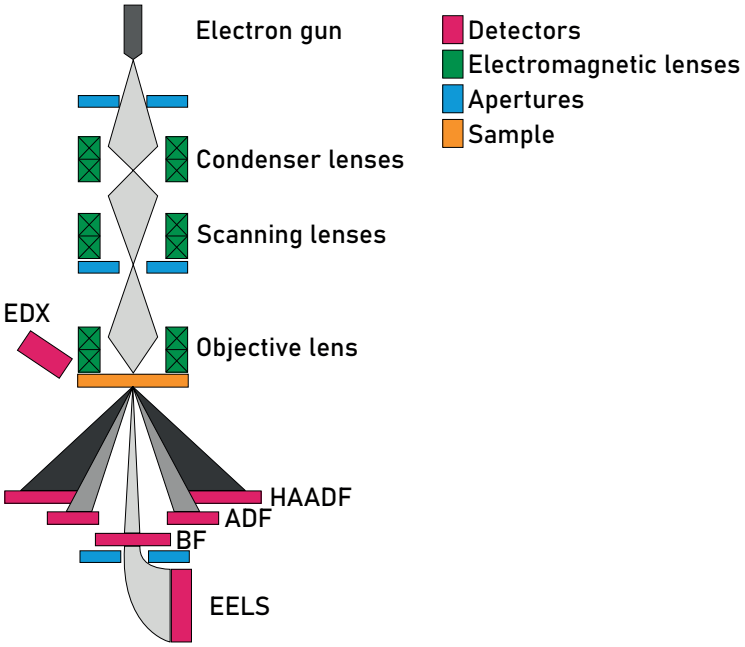


Figure 1.2 | Simplistic schematic of a typical TEM in STEM configuration [2]

[2] Illustration recreated with permission from reference [24]

components of a STEM. It focuses the beam onto the sample, forming a small probe that can be scanned over the sample surface. The size of the probe is determined by the convergence angle of the electron beam, which can be adjusted using the objective lens. The smaller the probe, the higher the resolution of the resulting image in conventional STEM experiments but it will become apparent in following chapters that this is not strictly true for phase retrieval approaches. In addition to these basic components, many modern STEM instruments are equipped with additional features to enhance their imaging capabilities. For example, aberration correction technology can be used to correct for distortions in the electron beam, resulting in higher resolution and better image quality [25]. A variety of detector systems is commonly found in STEM instruments, including integrating detectors, such as annular dark-field (ADF), low-angle annular dark-field (LAADF), annular-bright-field (ABF), bright-field (BF), and high-angle annular dark-field (HAADF) detectors. They span a certain angular range in the diffraction plane and integrate the collected signal over their respective surface area and a given time interval. Energy dispersive X-Ray (EDX) detectors are able to detect X-rays, which may be emitted as a result of excitations caused by the electron beam-sample interaction. The energy of the x-ray photons can be measured and linked to atom-specific excitation energies, which allows elemental mapping. Similarly EELS detectors may be used to determine sample compositions or electronic structures.

1.2.1 (HA)ADF Imaging & 4D-STEM

The angular integration range of signal integrating detectors has decisive influence on the contrast formation. Figure 1.3 shows a comparison illustrating the effect on a $SrTiO_3$ simulation at 300 kV and 20 mrad convergence angle. In the BF region all the signal that underwent no significant scattering is collected and the integrated signal is $\approx$1 in vacuum and attenuated at atomic column positions accordingly. BF-STEM imaging can be considered a coherent imaging mode by reciprocity [22]. Selecting a subset of the BF disk, excluding the centre of the BF, yields an ABF image, which still has significant coherent contributions, but is oftentimes easier to interpret. When the inner collection angle of an annular detector is larger than the convergence angle of the beam the contrast flips into the (LA)ADF regime. If the inner collection angle is $\approx$3 times larger than the convergence angle the resulting image is usually considered HAADF. The (HA)ADF detector is used to collect electrons that are scattered at large angles by heavy atoms in the sample, providing high-contrast images of the atomic structure, while the BF is more sensitive to atoms with lower atomic numbers as can be seen by comparing the oxygen contrast in figure 1.3. These different detectors may be used in combination as well to obtain a more thorough understanding of the sample structure. However, in practice the BF and ABF contrast and its interpretability often suffer from coherency and deteriorating effects such as carbon contamination, which is one of the reasons why (HA)ADF is often preferred and has become one of the- if not the most prominent imaging mode in materials science, in particular for quantitative studies. The contrast formation is well understood and generally scales with the average atomic number in an atomic column in zone axis orientation, as well as with its thickness along the beam direction. This gives rise to the possibility of reconstructing 3D nanostructures from single shot HAADF images [16, 17, 26]. HAADF

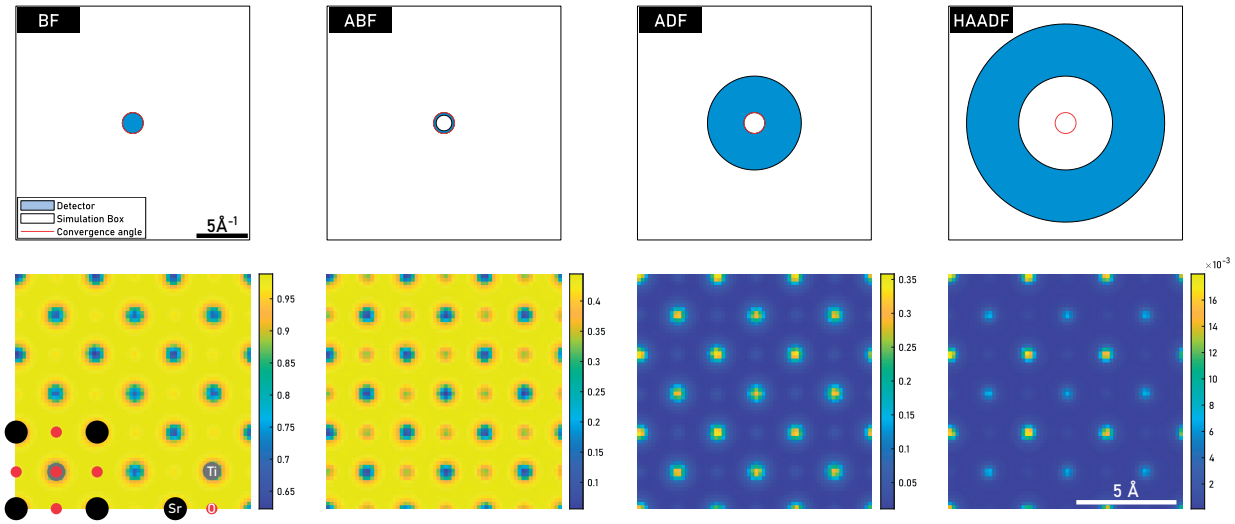


Figure 1.3 | Simulation of $SrTiO_3$ illustrating the influence of BF(0-20 mrad), ABF(15-20 mrad), ADF(20-90 mrad), HAADF(90-190 mrad) detector angles at 20 mrad convergence angle and 300 keV.

images further satisfy the projection requirement of tomography and can therefore be used for atomic resolution 3D-tomographic reconstructions [15].

However, the required dose for HAADF imaging is typically higher compared to other imaging modalities. This is because although all electrons interact with the sample and may cause damage to it, only a fraction of the electrons is actually collected by an HAADF detector to form contrast. This is very apparent when considering the actual numbers on the colour-scales in image 1.3, which represent the fraction of collected electrons per pixel. HAADF therefore requires a larger number of electrons to be scattered by the sample in order to produce a clear image. For many experiments this is problematic as a high dose may damage the sample under study, limiting the dose to an extent that creates poor contrast and noisy images or even makes some studies entirely impossible. One strategy to maximise the value obtained from a low-dose HAADF image is to restore as much information from it as possible, by taking into account the noise formation and other distortions and correcting for these effects [2, 27–32]. This thesis explores this approach in more detail in chapter 2. However, generally it would be preferred to make good use of the lower angle scattered electron instead of just discarding them. This may be achieved by combining different detector angle regimes, which was shown conceptually to improve the accuracy of atomic structure quantifications considerably for angles beyond the convergence angle [33]. However, this approach still neglects the significant signal in the BF-disc. Further, all of the detectors mentioned so far are circularly symmetric and can therefore not provide any directional information, i.e. into which direction the electrons scatter with respect to the beam position. However, this is a valuable information because the electron beam is subject to Coulomb deflection, which is proportional to the electric fields the electrons interact with [34]. The considerations that all angular regions contain useful information and that the direction and magnitude of the beam shift should be captured as well leads to the family of segmented detectors as shown in figures 1.4c-e. The solid quadrant detector (figure 1.4d) was among the

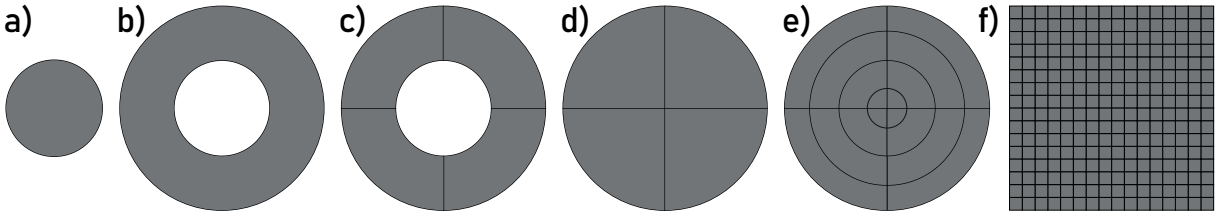


Figure 1.4 | Schematic of common electron detector geometries a) Bright field detector, b) Annular detector, c) Annular 4-Segment detector, d) 4-Segment detector, e) 16-Segment detector, f) pixelated (4D-STEM) detector

first attempts to record the centre-of-mass shift in the BF. Annular quadrant detectors (figure 1.4c) were later found to increase the signal to noise ratio [35] and allow for more flexibility in tuning the contrast transfer for Lorentz microscopy [36]. Further extensions of this idea lead to increased segmentation in detector geometries with 16 detector segments [37] (figure 1.4e) and even up to 120 [38] leading to improved flexibility and precision. Obviously, the ideal scenario would be to increase the segmentation to record actual images of the diffraction patterns, which has become technologically viable only in recent years due to the strides made in the development of direct electron cameras [39–44](figure 1.4f). With this setup a full CBED can be recorded, which for a STEM experiment results in one 2D-diffraction pattern per scan position on a 2D raster, yielding a so-called 4D-STEM dataset. Figure 1.5 illustrates a set of simulated CBEDs that may be recorded with a 4D-STEM camera, representing varying typical scenarios. The first row shows a setup with a small convergence angle, which leads to small BF-disc in reciprocal space, while the probe is relatively large in real space. As the thickness increases it can be seen that additional, distinct diffraction discs appear, which are spaced according to the crystals lattice spacings. This is analogous the conventional electron diffraction, where the illumination is parallel (i.e. convergence angle=0) and the diffraction spots very sharp. A convergence angle>0 increases the size of the diffraction spots in reciprocal space accordingly until they eventually overlap and are no longer distinct (figure 1.5e). Larger probe sizes (as in a) through c)) are commonly used for nanobeam electron diffraction (NBED) studies to e.g. measure strain fields [45, 46]. In the second row of images in figure 1.5, CBEDs with a larger convergence angle are shown. The electron probe is much sharper accordingly, allowing to limit the beam interaction to a very small sample volume. This is a typical microscope setting that would also be used in (HA)ADF experiments. It is easy to see, that simply integrating the signal beyond the BF disc would yield the same effect as using a dedicated ADF detector. The use of such virtual detectors is often referred to as virtual STEM (vSTEM) and allows in principle to construct any number of integrated signals up to the collection angle [33]. The overlap of the diffracted BF-discs and the resulting interference effects can be exploited to reconstruct phase information, as well as the shift of the COM. These concepts are essential for the chapters 3 and 4 of this thesis and are discussed in more detail there. As the thickness increases (figure 1.5c,f) enough so multiple scattering can occur, additionally features like Kikuchi lines and First (or Higher) order Laue zone rings (FOLZ & HOLZ) may become visible as well. HOLZ rings are sensitive to periodicity along the beam direction, while Kikuchi lines can be used to conclude on crystal orientations.

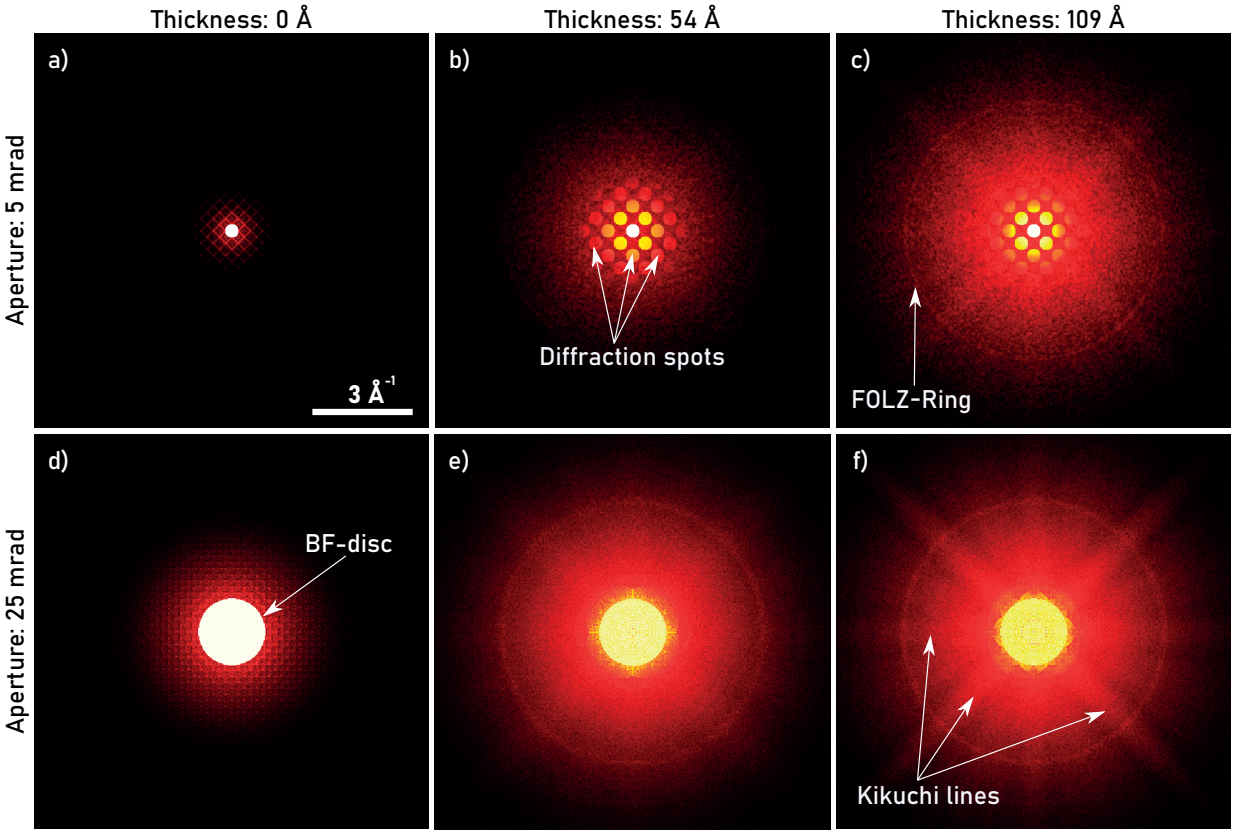


Figure 1.5 | Simulations of a Si crystal in (001) zone axis orientation, illustrating and highlighting prominent features commonly seen in CBEDs. The acceleration voltage is 200 kV and the convergence angle 5 mrad for a)-c) and 25 mrad for d)-f). The columns separate varying thicknesses from 0 (only first layer of atoms) to 109 Å. All figures are in logarithmic scale.

All these details and the many types of studies possible with 4D-STEM data again highlight why 4D-STEM is such a powerful and versatile tool. From a scientific point of view there is no argument for not using 4D-STEM over ADF so far. However, so far the acquisition speed of 4D-STEM is still significantly lower than integrating detectors, which can be problematic when time-resolution is a concern (e.g. for in-situ experiments). Also, the sheer amount of data resulting from 4D-STEM presents challenges in terms of data storage, but also in terms of fast processing, such that microscopists can actually work with the technique live, without relying on additional ("fast") detectors, while setting up an experiment, finding a suitable region of interest or adjusting the microscope. The chapters 3 and 4 both present approaches and implementations to meet these challenges. Scan line distortions, as well as drift, also apply to the collection of 4D-STEM data. Therefore, restoration approaches as outlined in chapter 2 or any of the references [2, 27–32] may also be applicable to 4D-data to some extent.

1.2.2 Physics of image formation & Image Simulations

Accurate simulations of TEM, STEM and ED images are an essential prerequisite for a variety of studies and methods in the field of electron microscopy. They may be used for qualitative comparisons to enable the interpretation of observed experimental images,

but may also be utilised in quantitative ways, for example for atom counting [47–49] or for the evaluation of imaging characteristics [50] and models [51] just to name a few. With the emergence of machine learning methods in electron microscopy, image simulations are also increasingly used for the generation of training datasets [52–55]. This is also true for chapter 4.2.3 in this thesis. For such applications that may require a very large number of synthetic images the computational efficiency of the simulation method is critical.

Multiple approaches for the quantum-mechanical description of electron scattering in thick specimen exist with the two most common concepts being the Bloch wave approach and the multislice algorithm. In the **Bloch wave approach** the electron wavefunction $\psi(\vec{r})$ is represented as a linear combination of j Bloch waves $b\left(\vec{k_j}, \vec{r}\right)$:

$$\psi(\vec{r}) = \sum_j \alpha_j b_j\left(\vec{k_j}, \vec{r}\right), \tag{1.4}$$

where $\vec{k_j}$ denotes scattering wave vectors on the Ewald sphere. A Bloch wave is a plane wave with a periodicity according to the crystal, which can be expressed as a Fourier series:

$$b_j\left(\vec{k_j}, \vec{r}\right) = \sum_G C_{Gj} \exp\left[2\pi i\left(\vec{k_j} + \vec{G}\right) \cdot \vec{r}\right] \tag{1.5}$$

Additionally, each Bloch must satisfy the time independent Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - eV(\vec{r})\right] b_j\left(\vec{k_j}, \vec{r}\right) = E b_j\left(\vec{k_j}, \vec{r}\right) \tag{1.6}$$

where m is the electron mass, $\hbar$ is the reduced Plank's constant ($\hbar = h/2\pi$), E the energy and $V(\vec{r})$ the potential, following the derivation of Kirkland [14, p.120]. While any set of coefficients α_j is allowed in equation 1.4, there is only one that matches also the incident wave function. Combining these requirements and making a few simplifying assumptions yields a set of equations which can be stated in matrix form and solved as Eigenvalue problem or using algorithms to solve the set of differential equations (e.g Runge-Kutta method) [14]. This is a complex task that gets computationally, prohibitively expensive as the number of considered scattering vectors G in equation 1.5 is increased. For G beams considered the CPU time scales as $\mathcal{O}(G^3)$ [14]. The Bloch wave approach is therefore practically limited to periodic structures that can be accurately described with relatively few scattering vectors. The computational cost and memory requirements scale poorly with the number of Fourier components considered, but are, in contrast to the multislice method, independent of the specimen thickness. For the simulation of diffraction patterns and in particular electron backscatter diffraction patterns Bloch waves are still regularly used [56, 57]. However, in TEM and STEM the **multislice method** is way more popular, since it is less restrictive and faster for most cases in transmission mode. The multislice method has no general periodicity requirement. The CPU time scales linearly with thickness, which is often acceptable considering that there is naturally an upper limit to the specimen thickness in transmission electron microscopy. The multislice method at its core uses the phase object approximation (POA), which states that the electron wave function $\psi(\vec{r})$ can be described by a simple multiplication

of the incident wave $\psi_0(\vec{r})$ with a specimen transmission function $T(\vec{r})$, neglecting the effect of wave propagation for thin samples:

$$\psi(\vec{r}) = T(\vec{r})\psi_0(\vec{r}) \tag{1.7}$$

The transmission function depends on an interaction parameter σ, which is constant for a given electron energy and the projected atomic potential $v_z(\vec{r})$:

$$\sigma = \frac{2\pi m e \lambda}{h^2}, \tag{1.8}$$

where m is the relativistic electron mass, e the elementary charge, λ the wavelength of the electron and h Planck's constant. The 3D electrostatic potential of the atoms is integrated along the beam direction into a 2D projected potential $v_z(\vec{r})$:

$$v_z(\vec{r}) = v_z(x, y) = \int V(x, y, z)\mathrm{d}z \tag{1.9}$$

The transmission function is then given by:

$$T(\vec{r}) = \exp\left[i\sigma v_z(\vec{r})\right] \tag{1.10}$$

For a simulation of a very thin specimen (e.g. 2D materials) this is sufficient to compute the exit wave for a given incident wave. If the sample is thicker, this principle can be extended to the standard multislice formalism, by firstly, recognising that ψ_0 in equation 1.7 can be any wave function, including the exit wave of a previous interaction. To make use of this, the 3D sample potential can be split into slices along the beam direction (z) with a given spacing Δz. Applying equation 1.9 to each slice yields a number of projected potentials $v_{z,n}$, which are Δz apart from another as illustrated in figure 1.6a). The potential between two consecutive slices is considered to be zero and the propagation of the electron wave within the slice is approximated by the Fresnel propagator $p(\vec{r}, \Delta z)$:

$$\psi_{n+1}(\vec{r}) = p_n(\vec{r}, \Delta z_n) \circledast \left[T_n(\vec{r})\psi_n(\vec{r})\right], \tag{1.11}$$

where $\circledast$ denotes a 2D-convolution operation. For computational efficiency this convolution in real space is usually implemented as a multiplication in Fourier space, where the propagator $P(\vec{k}, \Delta z)$ is given as:

$$\mathcal{F}\left[p(\vec{r}, \Delta z)\right] = P\left(\vec{k}, \Delta z\right) = exp\left(-i\pi\lambda k^2 \Delta z\right), \tag{1.12}$$

where $\mathcal{F}$ denotes a 2D Fourier transform, such that equation 1.11 can be restated as:

$$\psi(\vec{r}, z + \Delta z) = \mathcal{F}^{-1}\left[P(\vec{k}, \Delta z)\mathcal{F}\left[T(\vec{r}, z)\psi(\vec{r}, z)\right]\right] \tag{1.13}$$

With equation 1.13 the wave function can be determined at any z-position in the sample. This is illustrated with a very simple example in figure 1.6 for an incident plane wave $\psi_0(\vec{r})$ (CTEM mode) in panel b). Panel c) shows the phase of the intermediate wave after interacting with the potential $v_{z,1}$ and propagating a Δz of 5 Å according to equation 1.13. The exit wave (panel d) is again computed using equation 1.13, but the incident

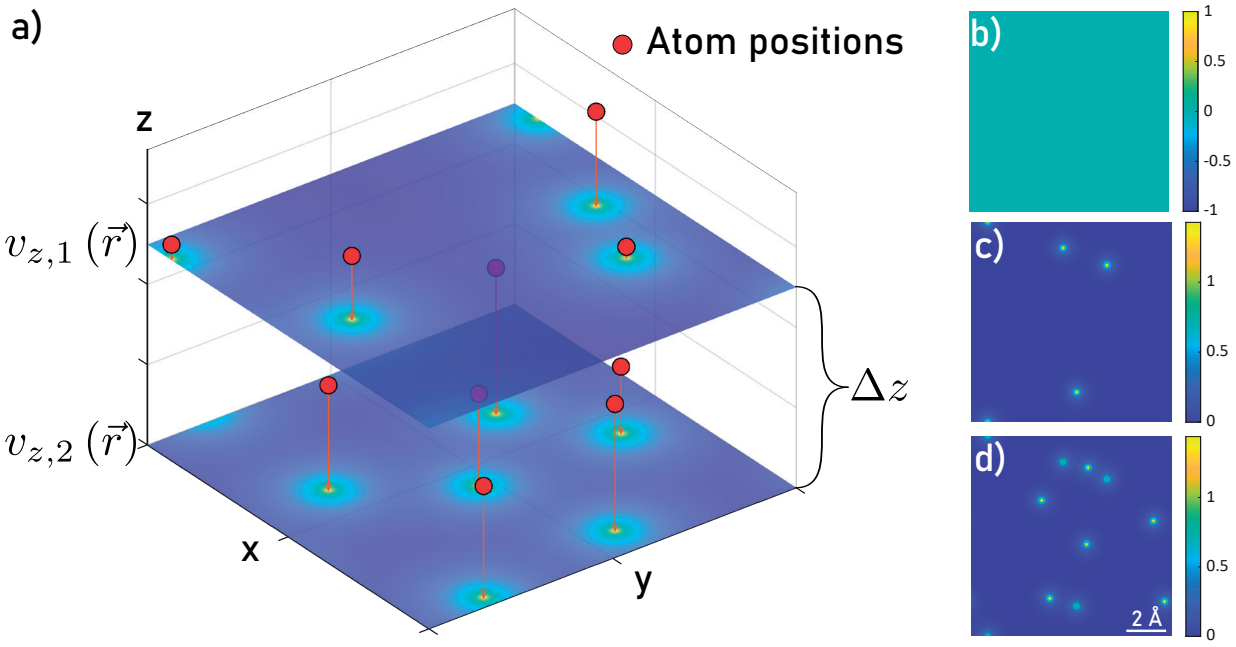
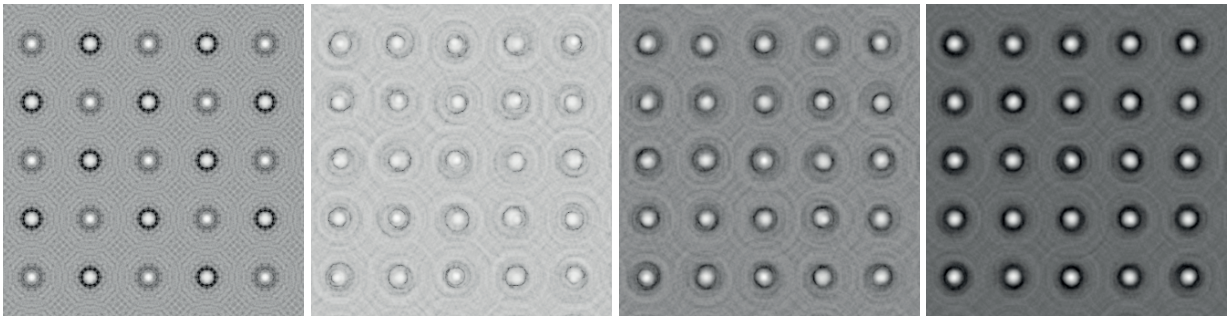


Figure 1.6 | a) An amorphous sample sliced into two segments of height Δz with their atomic electrostatic potentials projected into two slices $v_{z,1}$ and $v_{z,2}$. The simulation box is 10x10x10 Å in size. b) Shows the phase of the incident plane wave ψ_0, c) the phase of the intermediate wave after interacting with the potential $v_{z,1}$ and propagating 5 Å. d) The exit wave after interacting with the potential $v_{z,1}$ and $v_{z,2}$ and propagating 2x5 Å.

wave is now the intermediate wave function depicted in panel c). This is a somewhat simplistic description but captures the essence of the method. In practice multislice codes are considerably more sophisticated and more accurate versions of the multislice formalism exist [58–64]. A further important consideration is the inclusion of thermal diffuse scattering effects. One of the more accurate and most popular models to include this electron-phonon interaction is the frozen phonon model, which is based on the fact that the velocity of the electron passing through the sample is much higher than the atomic vibrations, such that the simplification can be made, that the electron wave interacts with a static atomic structure while traversing the sample [65]. Every atom is considered a harmonic oscillator with a single energy (one frequency) according to the root-mean-squared-displacement (rmsd) of the atomic site, which is related to the Debye-Waller-Factor (DWF) by: $rmsd = \sqrt{\frac{DWF}{8\pi^2}}$. As a result, all atoms are slightly shifted from their reference positions and the final exit wave can be computed as the average of multiple electron waves which interacted with different phonon configurations. Phonon configurations may be simply considered uncorrelated as in the Einstein approximation [61], or with the phonon band structure and correlated atomic motions taken into account [66]. The effect of this model is illustrated by CTEM simulations of a Ag crystal in (001) zone axis orientation in figure 1.7. Neglecting atomic displacements (still atom approximation), results in Figure 1.7a. The interference is strongly pronounced because all atoms are exactly stacked over another within the columns. Introducing a random atomic position offset causes a blurring as seen in Figure 1.7b. Averaging over ever

more phonon configurations creates increasingly smooth images and more realistic simulations (figures 1.7b to 1.7d).



(a) Still atom approximation | (b) Frozen phonon approximation (1 Phonon) | (c) Frozen phonon approximation (5 Phonons) | (d) Frozen phonon approximation (20 Phonons)

Figure 1.7 | Comparison of electron phonon interaction approximations (log-scale)

For STEM and 4D-STEM the electron-specimen interaction is exactly the same as for the presented CTEM case. The only difference is the choice of the incident wave, which is no longer just a plane wave with unity intensity, but a convergent probe.

Incident waves & aberrations

Shaping the electron beam requires the use of magnetic lenses, which cause aberrations, similarly as glass lenses cause aberrations in light microscopes. However, the quality of magnetic lenses is significantly worse. These technological difficulties have long been the limiting factor for the resolution in STEM until the development of probe correctors in 1998 [25]. Aberrations are modelled by means of an aberration function χ. The imaging wave function ψ_0 is a product of the aperture function $A(\vec{k})$ and the transfer function of the objective lens $H_0(\vec{k})$.

$$\psi_0(\vec{r}) = \mathcal{F}^{-1}\left[A(\vec{k})H_0(\vec{k})\right] \tag{1.14}$$

$$A(\vec{k}) = \begin{cases} 1 & \lambda|k| = \alpha < \alpha_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{1.15}$$

$$H_0(\vec{k}) = e^{-i\chi(\vec{k})}, \tag{1.16}$$

where k is the spatial frequency in the image plane, $\mathcal{F}^{-1}$ denotes inverse Fourier transforms and $\alpha_{\max}$ the maximum semiangle allowed by the objective aperture [14]. In principle, the largest possible outer aperture radius would be desirable (very large $\alpha_{\max}$), as the probe becomes smaller and smaller in real space. However, a larger outer aperture radius also means that higher scattering angles contribute to the signal, which are influenced by aberrations a lot stronger. For the example of spherical aberrations this is because the magnetic field further away from the optic axis is stronger than would be required to focus the electrons at these large angles (α) [14], which results in a position error in the electron trajectories, as shown in figure 1.8. In practice, aberrations cannot be fully omitted or corrected, which means that increasing the outer aperture is

always a trade off for an accentuation of aberration effects. The aberration function

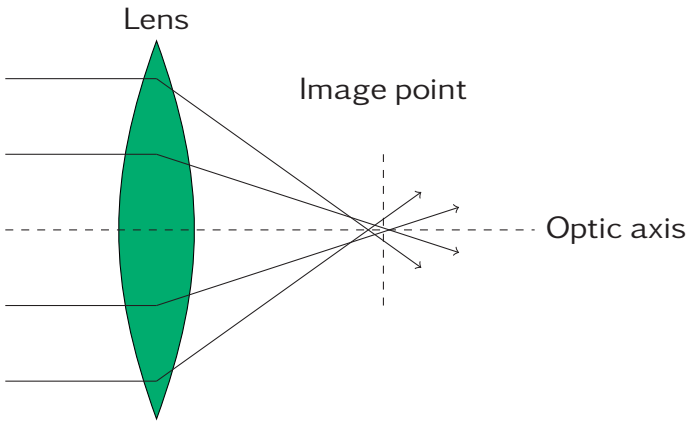


Figure 1.8 | Ray diagram illustrating spherical aberrations.

$\chi(\vec{k})$ can be constructed as a sum of the contributions arising from the different types of aberrations. The defocus Δf and the spherical aberration C_s are rotationally symmetric and its influence on the aberration function $\chi(k)$ is shown in equation (1.17):

$$\chi(k) = \pi\lambda k^2 \left(0.5 C_s \lambda^2 k^2 - \Delta f\right) \tag{1.17}$$

Generally, smaller aberrations allow for a smaller probe and therefore, a better resolution. The spherical aberration C_s is essentially constant for a given beam energy and cannot be fully, but partially compensated by adjusting an appropriate defocus value (Scherzer defocus), which can be found using equation 1.18

$$\Delta f = \sqrt{1.5 C_S \lambda} \tag{1.18}$$

Nonsymmetric aberration terms, for e.g. coma and astigmatism may be added to χ for a full description of the aberration function. However, in practice these contributions are often considered negligible for modern, well-adjusted instruments. They are at no point in this thesis considered and are therefore omitted in this chapter as well. The interested reader is referred to reference [14], chapter 2.6.

1.3 Neural Networks

1.3.1 Introduction

A neural network (NN) is a type of machine learning model that is inspired by the structure and function of the human brain. It consists of a large number of interconnected processing nodes, called artificial neurons, which work together to process input data and make predictions or decisions.

The structure of a simple neural network is depicted in Figure 1.9 and consists of layers of interconnected neurons, where each layer receives input from the previous layer and produces output that is passed to the next layer. The input layer receives the

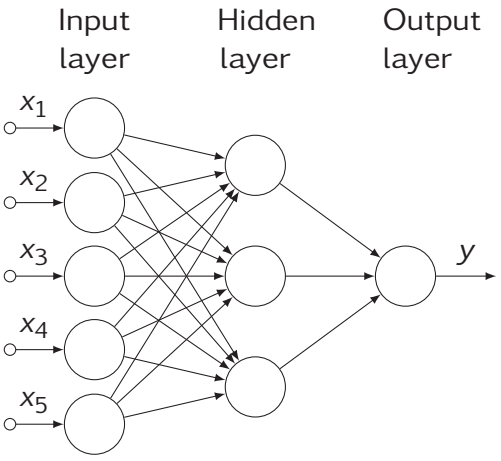


Figure 1.9 | Principal layout of a neural network.

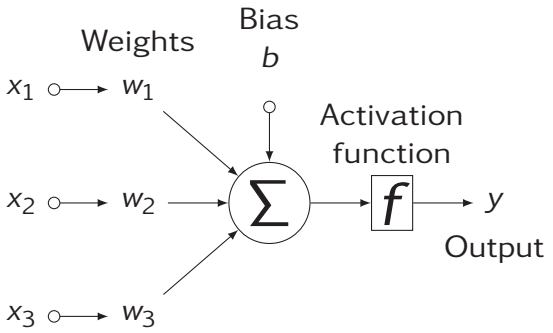


Figure 1.10 | Schematic depiction of an artificial neuron.

raw input data, and the output layer produces the final prediction or decision made by the network. Layers in between the input and output layers are called hidden layers, and the number of hidden layers and the number of neurons in each layer can be varied to control the complexity and capacity of the network. Artificial neurons are simple computational units that receive input, transform the input using a mathematical operation, and produce an output. The structure and function of a single neuron as depicted in Figure 1.10 can be described mathematically as follows: Let $x_1, x_2, ..., x_n$ be the input features for a single sample, and $w_1, w_2, ..., w_n$ be the corresponding weights. The weighted sum of the inputs is calculated as follows:

$$z = \sum_{i=1}^{n} w_i x_i + b, \tag{1.19}$$

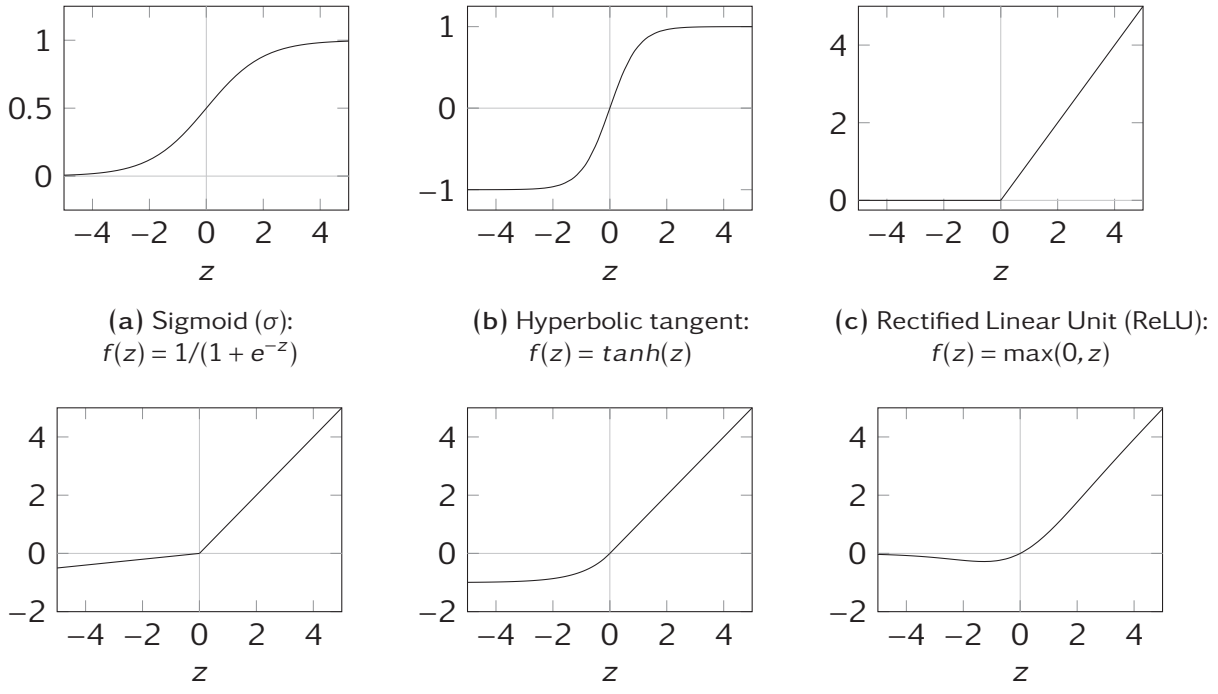
where b is a bias term. The output of the neuron is then computed by applying an activation function $f(z)$ to the weighted sum:

$$y = f(z) \tag{1.20}$$

Activation functions

Activation functions are essential to introduce non-linear behaviour. There are many different types of activation functions that can be used, including a small selection depicted in Figure 1.11. Sigmoid and *tanh*-functions were popular in the past but are hardly seen in modern, deep learning systems anymore, because they are relatively slow to compute and suffer from vanishing gradients for large (positive or negative) input values [67]. The de-facto standard activation function, especially in deep convolutional neural networks has become the rectified linear unit (ReLU), which is simple and fast to compute and has shown to accelerate the convergence of model trainings using gradient descent optimizers. However, ReLUs are prone to the so-called "dying-ReLU

problem", which is a special case of the vanishing gradient problem, referring to the common observation that neurons with ReLU activation become inactive and output only zeros for any input [68]. Since the derivative of the ReLU function is zero for negative inputs "dead" neurons usually do not recover during training, which can lead to large parts of a NN to become inactive. To mitigate the problems arising from the ReLU numerous other functions were introduced altering the behaviour of the ReLU for negative inputs [69] as illustrated in Figure 1.11(d-f). The ELU and Swish functions have the additional benefit that they are continuous around zero, which can help to improve the training and accuracy somewhat. Particularly in deep NN these modifications of the ReLU have shown to improve results considerably [69, 70].



(a) Sigmoid (σ): $f(z) = 1/(1 + e^{-z})$

(b) Hyperbolic tangent: $f(z) = tanh(z)$

(c) Rectified Linear Unit (ReLU): $f(z) = \max(0, z)$

(d) Leaky ReLU: $f(z) = \max(\alpha z, z)$

(e) Exp. Linear Unit (ELU): $f(z) = z$ if $z \geq 0$, $f(z) = \alpha(e^z - 1)$ if $z < 0$

(f) Swish: $f(z) = z/(1 + e^{-z})$

Figure 1.11 | Commonly used activation functions used to include (a) the sigmoid $\sigma(z)$ and (b) the hyperbolic tangent $tanh(z)$ functions. More recently mostly (c) ReLU and derived functions like (d) leaky ReLU, (e) ELU and (f) the Swish function [70], are being used in practical deep learning applications.

Training

The process of training a neural network involves adjusting the weights and biases of the connections between neurons in order to minimize the error between the predicted output and the ground truth for a given set of input data. The goal of the training is to find a set of weights and biases for a model that can accurately predict the output for a given input. The match between the prediction and ground truth label is computed by some scalar-producing loss function $\mathcal{L}$. Most commonly, training algorithms are using the first order derivative of this multivariate objective function (or loss function) with

respect to the model parameters θ to update the weights and biases in an iterative manner. This concept is easily demonstrated using the example of the batch gradient descent method.

$$\theta_{t+1} \leftarrow \theta_t - \eta \nabla_\theta \mathcal{L}(\theta; x; y) \tag{1.21}$$

In this equation, θ_t represents the model parameters (including the weights and biases) at time step t, and η is the learning rate, which determines the step size of the update. The gradient of the loss function with respect to the model parameters is denoted by $\nabla_\theta \mathcal{L}(\theta)$ and is computed using the backpropagation algorithm [71]. Since this implementation requires to the entire model and dataset to fit into computer memory it is not very practical in modern applications anymore. Adaptions to the concept are the stochastic gradient descent (SGD) and minibatch-gradient descent and optimization algorithms. The model weights and biases are updated after each training sample or batch of samples, respectively. This changes the update rules accordingly for the SGD:

$$\theta_{t+1} \leftarrow \theta_t - \eta \nabla_\theta \mathcal{L}(\theta; x_i; y_i) \tag{1.22}$$

and for the minibatch-GD with a batch size n:

$$\theta_{t+1} \leftarrow \theta_t - \eta \nabla_\theta \mathcal{L}(\theta; x_{i:i+n}; y_{i:i+n}) \tag{1.23}$$

The process is repeated until the loss function reaches a minimum or until a predefined number of iterations has been reached. An example of such a process is given in Figure 1.12 for a single arbitrary parameter of the NN. Each parameter has a random initialization[3]. The gradient at this point is determined and the parameter updated accordingly, where the learning rate η determines the scale of the update. This hyperparameter is crucial. Examining Figure 1.12 one can easily imagine how a learning rate too large would cause the parameter to never actually reach the minimum, but jump back and forth between the sides of the valley. On the other hand, if we imagine the inital parameter being on the right side of the graph, one can see that a learning rate too small could lead to the optimization getting stuck in the local minimum. While Figure 1.12 depicts this for one parameter, it is important to consider that training a NN potentially means optimizing thousands, millions or even billions of parameters simultaneously, resulting in an extremely complex and difficult optimization task. A number of improvements and algorithms were put forward to account for these difficulties, including strategies of adding a momentum term to the update rule [72, 73], the introduction of parameter-individual, adaptive learning rates, and combinations thereof. Notable developments on gradient-descent optimizers include "Adagrad" [74], "Adadelta" [75], "RMSProp"[4] and "Adam", using Adaptive Moment Estimation [76]. The latter in particular has shown to be reliable in many scenarios and hence emerged as a default choice for many applications [77]. The Adam algorithm uses individual adaptive learning rates for all parameters, computed from estimates of the first and second moments of the gradients as outlined in algorithm 1.

[3] The appropriate initialization method is a topic in itself and may depend on the architecture and activation functions used.

[4] This algorithm was developed by Geoffrey Hinton. The work is available in his lecture notes at http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf but was never formally published.

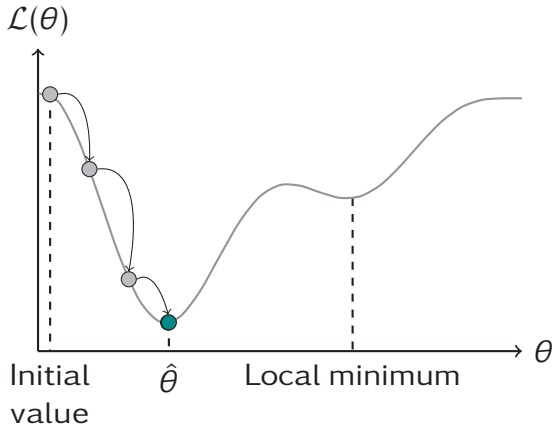


Figure 1.12 | Illustration of gradient descent on a single variable.

1.3.2 Convolutional neural networks

A convolutional neural network (CNN) is a type of neural network specifically designed to process data that has a grid-like topology, such as an image. The main concept behind the operation of a CNN is the idea of convolutional layers. In a convolutional layer, the network learns a set of filters, which are small matrices of weights. These filters are used to scan over the input data, such as an image, and extract features from it. The output of a convolutional layer is called a feature map, which represents the presence of certain features in the input data. By stacking multiple convolutional layers on top of each other, a CNN can learn increasingly complex features of the input data, such as edges, shapes, and textures.

A CNN is trained using a labelled dataset of images, where the network learns to associate the features it extracts with a particular class or label. Once trained, the network can be used on new images by extracting their features and using them to make predictions based on what it has learned during training. CNNs gained a lot of relevance in the 1990s with one of the most significant early contributions being the LeNET-5 [78, 79] for handwritten digit and character classification. This network was among the first to contain all fundamental building blocks of most modern CNNs, as shown in figure 1.13. In the illustrated example the network takes a greyscale image (one channel) as input, applies a convolution with a 5x5 kernel and a 2x2 pooling operation. This convolution-pooling sequence is repeated once more. The last convolution has 120 filters with a kernel size of 5x5 pixels, which reduces the 5x5 feature maps to scalar values. This amounts to a full connection. After another full connection follows the final classifier step [79]. The individual components will be explained in a general manner in the following sections.

Dense connections

Dense connections, also called "Fully connected layers", are often used in CNNs towards the output of the network after feature maps were "flattened", i.e. the maps were reduced to one dimension and concatenated into a vector. This allows to map spatial data (e.g. images) to vectors or scalars, which is necessary for example in classification tasks.

Algorithm 1 Adam Optimization as per reference [76]. g_t^2 indicates the elementwise square $g_t \odot g_t$. All vector operations are element-wise too.

Require: : $\mathcal{L}(\theta)$ ▷ Objective function
Require: : θ_0 ▷ Initial parameter vector
Require: α ▷ learning rate
Require: $\beta_1 \in [0,1)$ ▷ Exponential decay rate for m
Require: $\beta_2 \in [0,1)$ ▷ Exponential decay rate for v
Require: ϵ ▷ Small value to avoid division by zero
$t \leftarrow 0$ ▷ Initialize time step variable
$m_0 \leftarrow 0$ ▷ initialize 1^{st} moment vector
$v_0 \leftarrow 0$ ▷ initialize 2^{nd} moment vector
while θ_t not converged **do**
  $t \leftarrow t+1$
  $g_t \leftarrow \nabla_\theta f_t(\theta_t - 1)$ ▷ Get gradients w.r.t. stochastic objective at timestep t
  $m_t \leftarrow \beta_1 m_{t-1} + (1-\beta_1) g_t$ ▷ Update biased first moment estimate
  $v_t \leftarrow \beta_2 v_{t-1} + (1-\beta_2) g_t^2$ ▷ Update biased second raw moment estimate
  $\hat{m}_t \leftarrow m_t / (1-\beta_1^t)$ ▷ Compute bias-corrected first moment estimate
  $\hat{v}_t \leftarrow v_t / (1-\beta_2^t)$ ▷ Compute bias-corrected second raw moment estimate
  $\theta_t \leftarrow \theta_{t-1} - \alpha \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)$ ▷ Update parameters
end while
return θ_t ▷ Optimized parameters

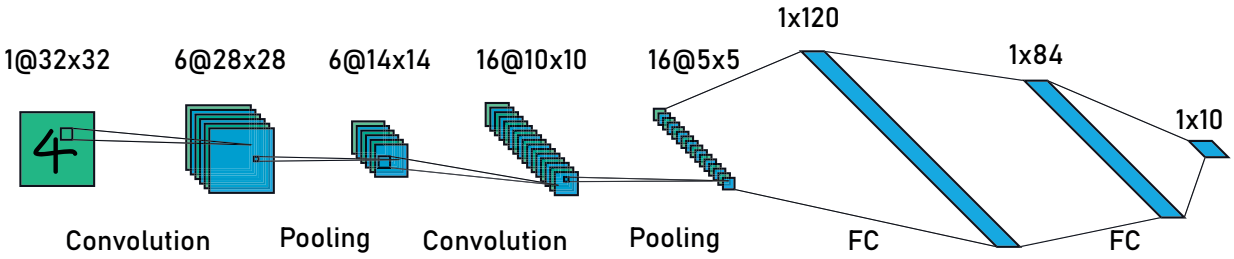


Figure 1.13 | Minimalistic example of a CNN, based on the LeNET [79], including the fundamental building blocks: convolutional layers, pooling layers and dense layers. [5]

These layers are identical to the ones described for generic neural networks (figures 1.9 & 1.10).

Pooling layers

Pooling layers introduce a down-sampling to feature maps individually, usually applied after convolutional layers. The most common pooling operations are the max- and average(or mean)-pooling. The former maps the maximum value of a feature map within a given kernel to a corresponding pixel value of the output map, whereas the latter computes the mean over the kernel area. The effect is demonstrated in figure 1.14 for the common setting with a pooling size 2×2 pixels and a stride of 2 pixels. The size of the feature map is reduced by a factor of 0.5 for both operations. One of the reasons these

[5] Illustration generated with http://alexlenail.me/NN-SVG/LeNet.html [80]

layers are so commonly employed can be intuitively understood for Max-pooling. Albeit some information is lost, the resulting feature map would not change for a shift of the maximum value within the pooling kernel. This helps to make the CNN invariant to small translations in the input maps [71]. Since layers in CNNs tend to encode the presence of patterns and features it is often beneficial to retrieve the strongest response for a given pattern. Particularly in computer vision tasks, such as image classification and object detection, max-pooling frequently showed better results and is therefore more commonly used [81]. A special case of pooling layer is the global pooling layer, which

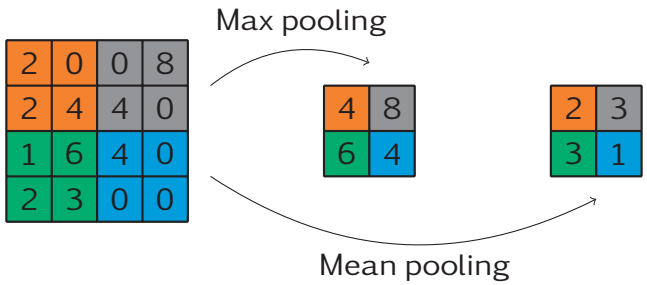


Figure 1.14 | Visualisation of max- and mean-pooling operations with a kernel size of 2x2 and strides of 2 in both, x and y directions. The pooling kernels and their corresponding results are color-coded.

has a kernel size equal to the feature map size and hence reduces each feature map to a single value. This may be used at the end of a CNN before fully connected layers.

Convolutional layers

In a CNN, each convolutional layer consists of a set of "filters" that are used to detect certain patterns or features in the input data by. This process of convolving the input data with a set of learned filters (also called kernels) allows a CNN to automatically learn spatial hierarchies of features from the raw input data, which is particularly useful for tasks such as image classification and object detection.

The terms *convolution* and *cross-correlation* are often used interchangeably in the machine learning community. Both operations are equivalent except for a flip of the kernel which can be avoided in the case of cross-correlation. Most AI-libraries actually use cross-correlations, but often call it convolution. The formal definition of the "convolution"[6] operation for a 2D image I with pixel coordinates i, j and a Kernel K of size $m \times n$ is therefore given as [71]:

$$S(i,j) = (K * I)(i,j) = \sum_m \sum_n I(i+m, j+n) K(m,n) \tag{1.24}$$

, which is essentially the dot product. The size of S would be $(i-(m-1)) \times (j-(n-1))$ correspondingly, which implies that feature maps would be shrinking in each convolutional layer. The architecture shown in figure 1.13 for example shows this behaviour. This can

[6] for consistency with other literature the text will keep referring to the term convolution, albeit it is technically a cross-correlation

be altered by zero-padding the input to account for the kernel size. For example for an output to maintain the input size would require a padding p of one pixel around the input for a 3×3 kernel. Further, the output size can be reduced by skipping steps along each dimension, which is referred to as a strided convolution with a stride parameter s. A stride of $s = 2$ would mean to compute only every other convolution along the given dimension. The size of the output with a given padding p, input size i and stride s is thus given by [82]:

$$o = \left\lfloor \frac{i + 2p - k}{s} \right\rfloor + 1 \tag{1.25}$$

Figure 1.15 illustrates the discrete convolution for the first two steps along x and y with numerical examples, including a zero-padding with $p = 1$, a stride of $s = 2$ for an input size $i = 5$ and kernel (figure 1.15-e) of size $k = 3$.

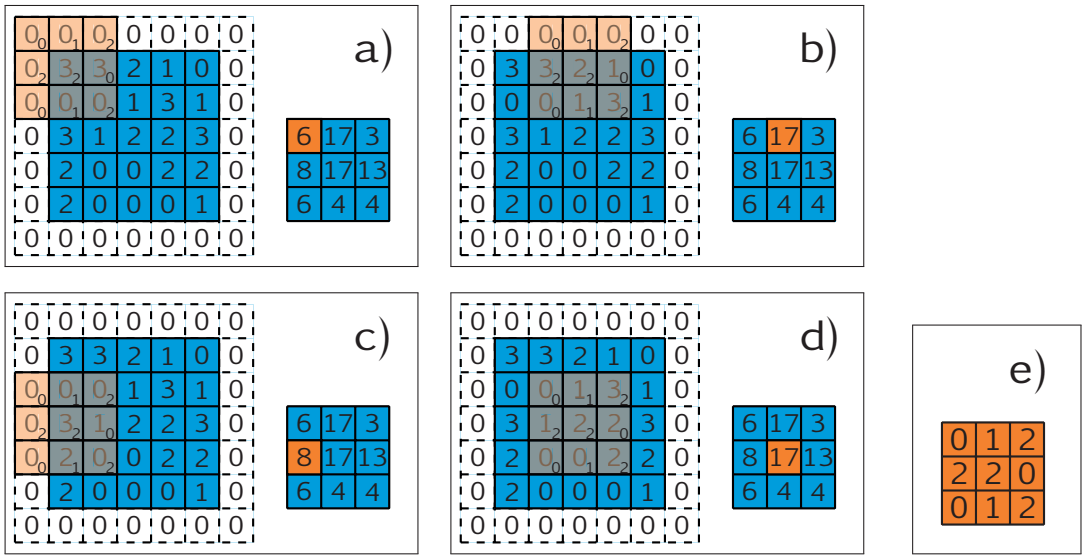


Figure 1.15 | Numerical example of a strided, discrete 2D convolution with $i = 5$, $k = 3$, $s = 2$, and $p = 1$. The highlighted pixel in the smaller output maps in figures a-d are computed by the dot product of the kernel (e) and the highlighted pixels in the corresponding input arrays. [7]

However, in practice the input tensor and output tensor are typically three-dimensional: width, height, and depth. The convolutional kernel is typically much smaller than the input tensor in the spatial dimensions and has a depth that matches the depth of the input tensor. Taking this into account and considering a number of c filters modifies equation 1.24 accordingly.

$$S_{c,j,k} = \sum_{l} \sum_{m} \sum_{n} I_{l,j+m,k+n} K_{c,l,m,n} \tag{1.26}$$

This means that the number of output feature maps equals the number of filters in the layer. For example, consider a 5x5x3 input tensor and five 3x3x3 filters as illustrated in figure 1.16. The convolution operation would involve iterating over the input tensor and the kernel, computing the dot product of the elements in the kernel and the input tensor at each position, and adding the result to the output tensor. This process would be repeated for every position in the input tensor, producing a new output tensor with a

depth of 5. As described in section 1.3.1, analogous to the classical artificial neuron, also convolutional layers use trainable biases and activation functions, which are applied to the output tensor. In summary the process of computing a convolution can be broken

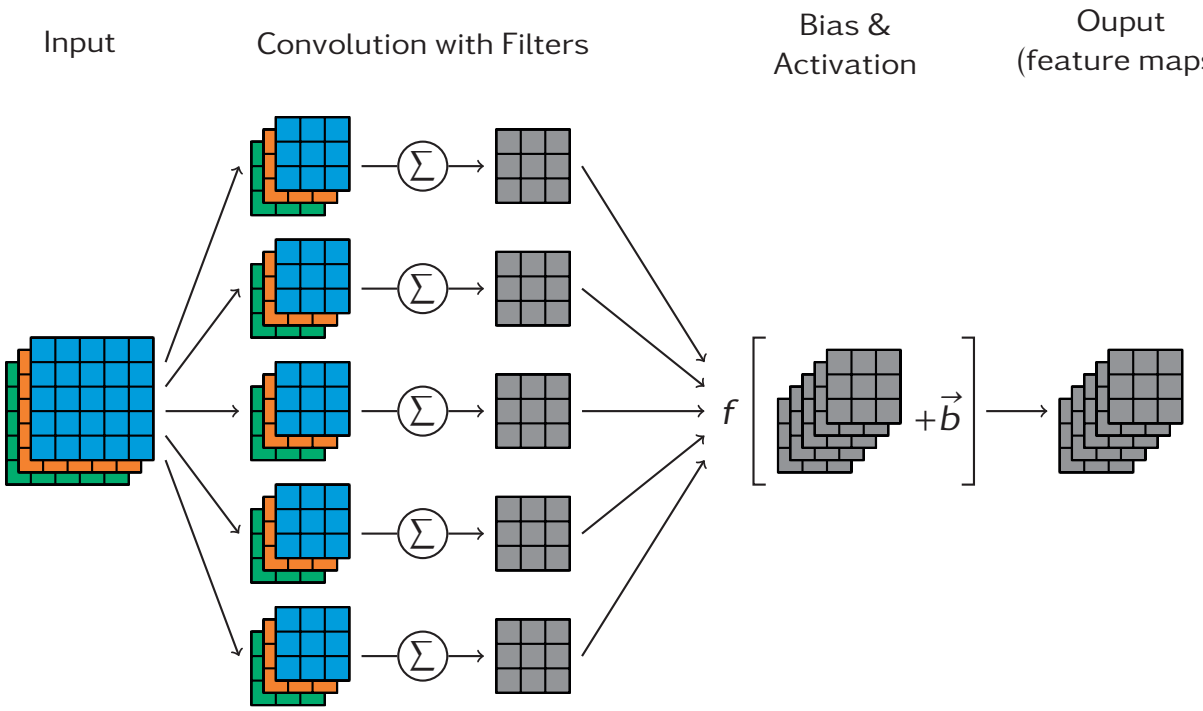


Figure 1.16 | Schematic illustration of a convolutional layer in a CNN with a input of size 5x5x3. The input is convolved with with 5 3x3x3 filters resulting in 5 3x3 feature maps (output size is reduced because no padding was applied). Trainable biases $\vec{b}$ are added to each map and an activation function f is applied to the output tensor.

down into the following steps:

1. Initialize the output tensor with zeros.
2. Iterate over the input tensor and the convolutional kernel, using their respective width, height, and depth dimensions.
3. For each iteration, compute the dot product of the elements in the convolutional kernel and the input tensor that are being "convolved" (i.e., multiplied and summed).
4. Add the dot product to the corresponding element in the output tensor.
5. Repeat this process until the entire input tensor has been convolved with the kernel, producing the output tensor.
6. Add a bias to each feature map and apply the activation function to the output tensor.

[7] Images for figures 1.15 & 1.17 were generated based on TikZ code given in reference [82], available at https://github.com/vdumoulin/conv_arithmetic

Another important type of convolutional layer is the so-called **transposed convolutional layer**, which is often used to upsample feature maps. To achieve this, the kernel(array) is multiplied with a given pixel in the input map (scalar) and added to the corresponding positions in the output map. The output size naturally increases correspondingly $(k-1)/2$ pixels. Striding and padding may be applied similarly as in the standard convolution case. A numerical example for a 4x4 input map, a 3x3 kernel, a stride of 1 and no padding is given in figure 1.17.

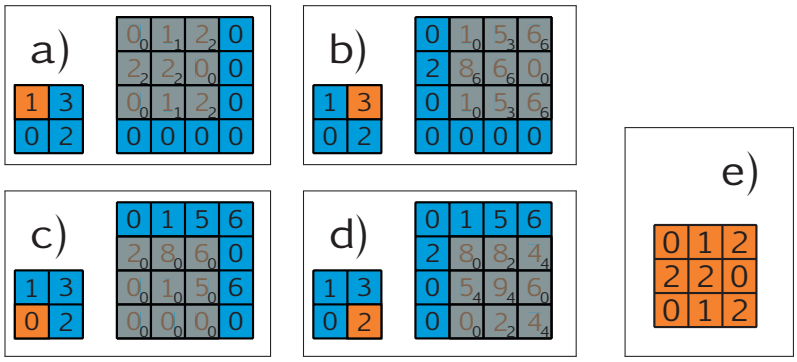


Figure 1.17 | Numerical example of a strided, discrete 2D transposed convolution with $i = 1$, $k = 3$, $s = 2$, and $p = 0$. The highlighted pixel in the input maps in figures a-d are multiplied with the kernel (e) and added to the output array.

Modifications for deep convolutional neural networks

The use of the chain rule in the backpropagation algorithm for the computations of gradients presents a problem for the construction of deep networks. If convolutional layers are just arranged sequentially as in figure 1.18-a), the gradients keep decreasing towards the beginning of the CNN and eventually vanish. Multiple strategies were developed to avoid this problem and to allow for the design of deep CNNs. Most of the approaches are based on the idea of **creating short paths from earlier to later layers** in the network. This can be done by addition as shown in figure 1.18-b), which is a residual block of the popular ResNet [83]. Figure 1.18-c) illustrates the use of so-called dense connections. In this case the short path is established by concatenating feature maps from earlier to later layers in the network. This idea may be used to construct densely connected architectures as in reference [84]. Also, both concepts may be employed simultaneously to construct residual dense blocks as in references [85, 86]. Skip connections may also be used to enable global residual learning which is a useful concept for image-to-image tasks, e.g. denoising [86] or image super-resolution [85].

Another general approach to support the training of deeper networks is the **normalisation of feature maps throughout the network**. The most commonly used normalisation method is the batch normalisation (BN) as introduced by Ioffe and Szegedy [87]. The core idea is to normalize each feature map in the CNN over a mini-batch by subtracting the mean and dividing by the variance, such that the expectation $\mathrm{E}(\mathrm{x})$ is zero and the variance unity.

$$\hat{x} = \frac{x - \mathbb{E}[x]}{\sqrt{\mathrm{Var}[x]}}, \tag{1.27}$$

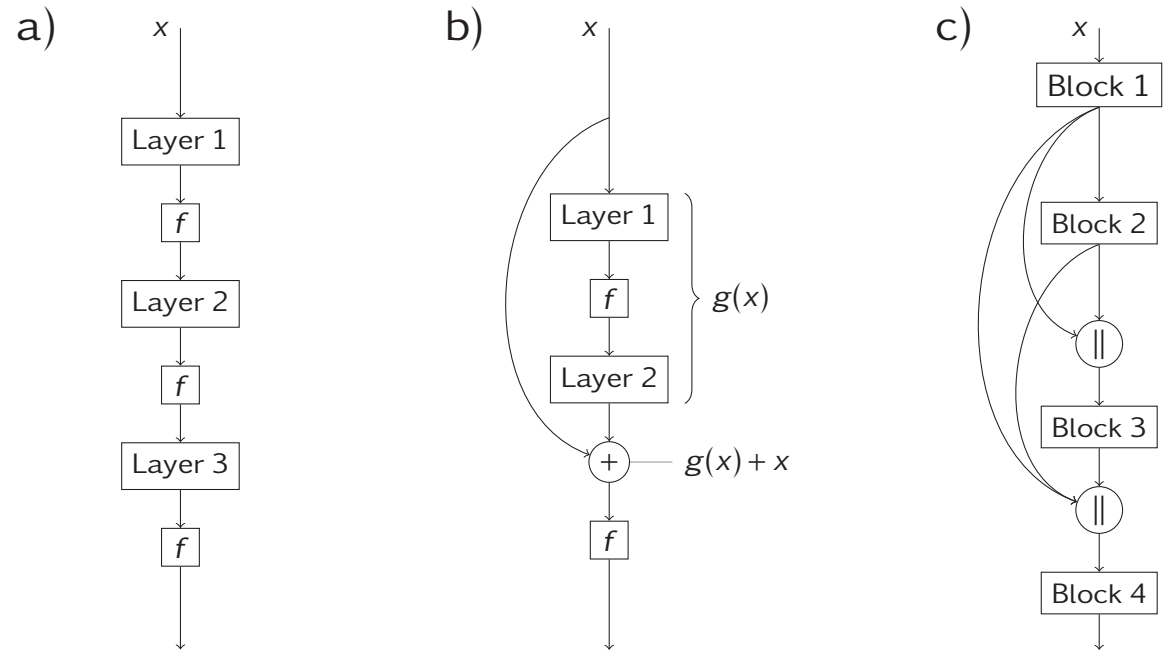


Figure 1.18 | Connections of convolutional layers: a) depicts a standard feed-forward connection, b) show a residual block with a skip connection as in reference [83], c) shows a dense block with dense connections (concatenations). ⊕ is an addition operator, ‖ symbolises a concatenation of feature maps along the channel dimension, f is the activation function. The blocks in c) include a convolution layer, activation function and batch normalization as per reference [84]

where the tensor x has the size of the feature map multiplied by the mini-batch size. The output of a BN layer $\hat{x}$ is then computed by scaling and shifting $\hat{x}$ by trained parameters γ and β respectively.

$$y = \gamma\hat{x} + \beta \tag{1.28}$$

More generally we may consider a 3D tensor as depicted in figure 1.19, that spans c channels, of a feature map of size $(h \times w)$ and n samples in a mini-batch. Slicing this cube into normalised sections in different ways yields different normalisation methods, with the three most common approaches depicted in figure 1.19, In figure 1.19-a) the tensor is normalised for each feature map over the batch size, which corresponds to batch normalisation. Figure 1.19-b highlights the dimensions to be normalised in the so-called layer normalisation and c) shows the instance normalisation. Further, a compromise between instance- and layer normalization is the group normalisation [88] where the layer is split into multiple groups.

[8] Figure design inspired by reference [88]

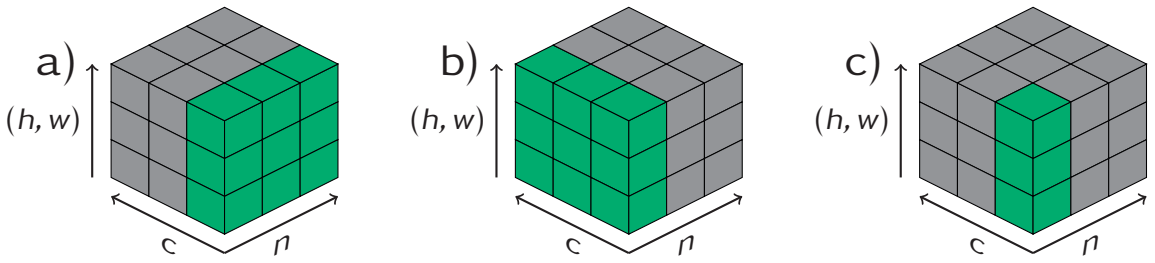


Figure 1.19 | Common normalisation methods used in CNNs include a) Batch normalisation, b) Layer normalisation and c) Instance normalisation. The dimensions n correspond to samples in a mini-batch, c to the channels in a feature map and (h, w) corresponds to a combined dimension representing the spatial dimensions of a feature map. The normalised section of the tensor is highlighted in green.[8]

CHAPTER 2

STEM Distortion correction

Contents

This chapter is based on- and uses content form the preprint:
"Lobato, I., T. Friedrich, and S. Van Aert. „Deep convolutional neural networks to restore single-shot electron microscopy images“. arXiv, 29. March 2023. https://doi.org/10.48550/arXiv.2303.17025",
which is currently under review for publication at npj Computational materials.

2.1 Introduction

The image resolution of modern scanning transmission electron microscopes (STEM) has reached the sub-Angstrom range, enabling microscopists to directly image individual atoms or atomic columns. Beyond merely visual interpretations, atomic-resolution STEM images can be used to extract quantitative measures, such as atomic column positions, scattering cross sections (SCS[1]) and a number of derived quantities, such as the number of atoms in a nanoparticle or strain distributions [89] Reliable determination of these measures is an important cornerstone for materials engineering at the nanoscale. Statistical parameter estimation theory was successfully employed to obtain accurate and precise estimates of atomic column positions and SCS [89] enabling one to locate atomic columns and to count atoms with single-atom sensitivity [90]. The reliability of these estimates depends on the quality of the underlying STEM image, which is naturally limited by systematic and statistical distortions. In the case of STEM images, the most prominent systematic distortions include X- and Y-jitter and fast scan distortion and unavoidable statistical distortions include a mixture of counting- and thermal noise. These distortions are particularly severe, when investigating beam sensitive materials and light elements and may make it very difficult to draw quantitative conclusions from STEM data [91].

This reasoning has been motivating the development of image post-processing and restoration approaches for both TEM [92] and STEM data [93–95]. Methods dedicated to atomic-resolution STEM images often make use of lattice periodicity, which imposes an important limitation to crystalline materials [96]. Other methods require a pair of STEM images of the same area but rotated by 90° to each other around the optical axis [97], or a multiframe series of the same region [98]. Jones et al. made considerable progress on the correction of scan line distortions on single images by essentially refining each pixel position based on its surroundings and imposing the assumption that image features be continuous [27]. While this approach reliably conserves the integrated intensity of the entire image without assuming any periodicity, it has some notable limitations, especially for data of low-dose experiments. The referenced examples show that in order to correct these image distortions by means of classical image processing techniques, they need to be individually identified and corrected, resulting in a variety of specialised, complex and computationally expensive algorithms that often have specific requirements on the data. The introduction of machine learning techniques led to a paradigm shift in image processing in the 70s and 80s [99–102] heralding the more recent surge in developments of deep convolutional neural networks (CNN) for numerous applications, especially for image reconstruction, image restoration, image enhancement, feature extraction, segmentation, object recognition and classification problems [102]. We have shown that deep convolutional neural networks can be used to denoise and correct image distortions in single-shot STEM data without any particular constraint, prior knowledge or additional user input using a blind restoration approach. Especially the correction of scan-line distortions can be very significant, also beyond standard (HA)ADF imaging scenarios. If the CNN can faithfully recover the correct, undistorted image

[1] formally defined in section 2.3.2

it can be used as a ground truth image to determine a shift mapping that transforms the distorted to the undistorted scan grid. With the knowledge of this transformation also multidimensional data, such as spectroscopy data (e.g. EELS and EDX) and 4D-STEM data can be corrected as well. This is because all spectra or CBEDs in a dataset are linked to real-space positions on a scan grid and can be shifted, similarly as in references [28, 98]. Figure 2.1 illustrates that also 4D data can simply be reshaped into a cube which can then be corrected by interpolating each data column to their correct real-space location. Besides reviewing the most relevant distortion and noise

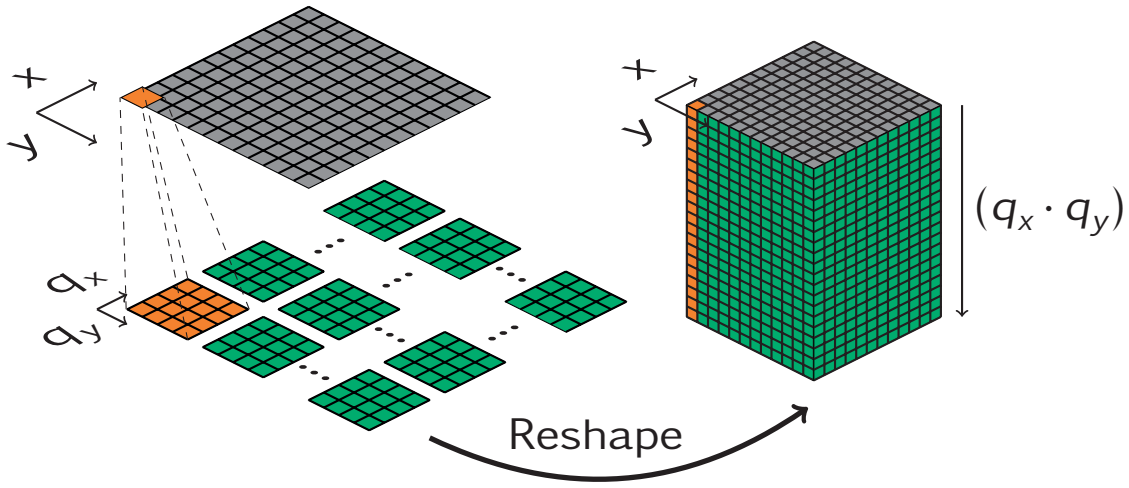


Figure 2.1 | Correcting 4D STEM data based on 2D images is possible due to their shared real-space locations (x and y). Distortions in the scan grid may be recovered from the STEM image. After reshaping each CBED with reciprocal dimensions q_x and q_y into a single vector, the resulting cube can be interpolated over the real-space dimensions with the corresponding shift mapping.[2]

contributions for atomic resolution STEM, as well as the fundamentals of parametric model estimation for STEM data, this chapter is dedicated to the determination of the improvements in accuracy and precision for quantitative data analysis resulting from this image restoration over a wide range of signal to noise ratios (SNR). We confirm that this procedure does neither introduce any significant bias nor eliminates specimen features such as dislocations or grain boundaries.

2.2 Distortions in (HA)ADF STEM

2.2.1 Detector Noise

Because electron detection is in principle a counting of events (electrons detected per pixel area) it is naturally subject to Poisson noise (or shot noise). In statistical terms a

[2] Image was recreated based on reference [98]

measured pixel intensity is a random variable following the Poisson distribution. So the probability P to observe k electrons is:

$$P(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \quad , k = 0, 1, 2, ... \tag{2.1}$$

where the λ-parameter is both the sample mean and variance. To model Poisson noise in a STEM image, each pixel is assigned a random number following this distribution (equation 2.1), with a value of $\lambda = I \cdot d \cdot dx \cdot dy$, where I is the simulated pixel intensity, d is the dose in $e^-/Å^2$ and dx and dy the pixel dimensions in Å. This simple model is a good approximation for high dose (HA)ADF imaging, when the image noise is dominated by Poisson noise. However, at lower doses the influence of the detector system becomes more pronounced, as the intensity distribution of imaged noise is no longer purely Poissonian. An accurate model of the detector system is necessary to adequately describe and understand the noise in ADF-STEM images.

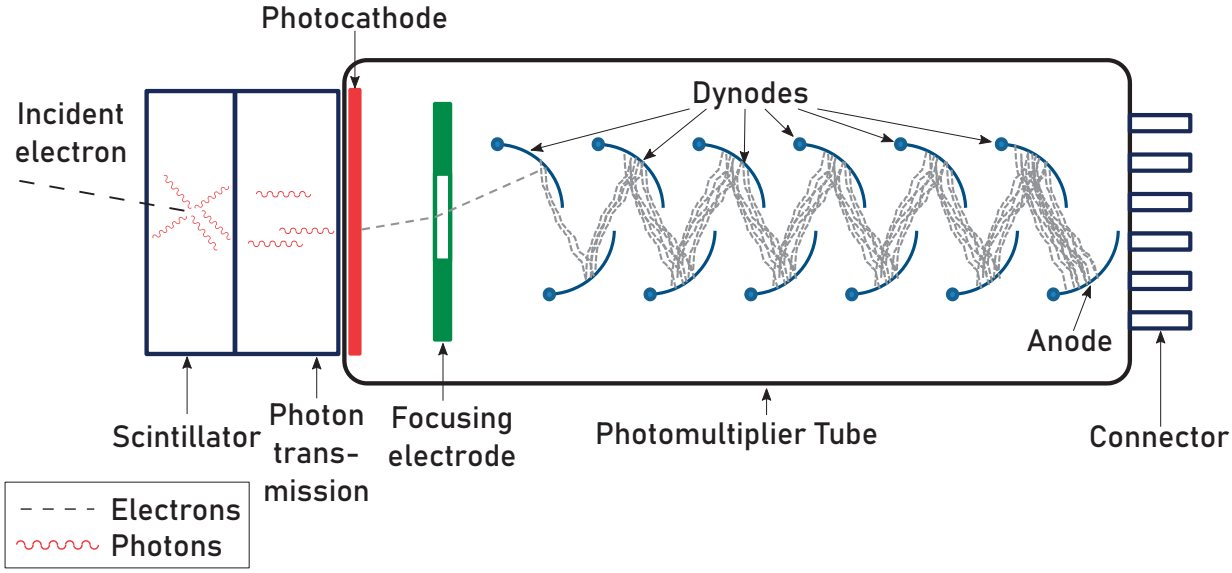


Figure 2.2 | Schematic of a typical (HA)ADF detector setup, consisting of a scinitllator and photmultiplier tube. [3]

A typical (HA)ADF detector consist of a scintillator coupled to a photomultiplier tube (PMT) as illustrated in figure 2.2. An electron incident on the scintillator produces photons, which are then transmitted to a photomultiplier tube (PMT) through e.g. a parabolic mirror [103] or a light pipe [104, p. 139]. The photons are converted back into electrons by a photocathode upon entering the PMT with a quantum efficiency η [105]. The focusing electrode accelerates the photoelectrons towards the first of a set of dynodes, which are electrodes with a high secondary-electron yield. At each dynode electrons are multiplied by a secondary emission factor δ with an average value and a variance. The successive acceleration and secondary-electron emission over n dynodes with a secondary-electron yield δ amplifies the signal in a process that can be understood as a cascade of independent, sequential events. In general for a cascade

[3] Illustration adapted/modified from https://commons.wikimedia.org/wiki/File:PhotoMultiplierTubeAndScintillator.svg, which is published under a CC BY-SA 3.0 license

of two events with generating functions $\mathcal{A}(x)$ and $\mathcal{B}(x)$, the generating function of the cascade $Q_{\mathcal{AB}}(x)$ is

$$Q_{\mathcal{AB}}(x) = \mathcal{A}[\mathcal{B}(x)] \tag{2.2}$$

and the average product and variance of the product are given by:

$$\overline{n}_{\mathcal{AB}} = \overline{n}_{\mathcal{A}}\overline{n}_{\mathcal{B}} \tag{2.3}$$

$$\sigma^2_{\mathcal{AB}} = \overline{n}^2_{\mathcal{B}}\sigma^2_{\mathcal{A}} + \overline{n}_{\mathcal{A}}\sigma^2_{\mathcal{B}} \tag{2.4}$$

[106, p.161]. These relationships hold for the **scintillator-photocathode stage**, as well as for the subsequent processes of the multiplier chain and can be used to estimate the characteristics of the detector. Firstly, the photon yield of a single electron impacting the scintillator is energy dependent and also varies depending on the type of the scintillator. Ishikawa et. al reported ≈ 20 electrons per kV for YAP (Yttrium Aluminium Perovskite) scintillators and an efficiency of the transmission setup guiding the photons to the photocathode of $\approx 5\%$. This yields an approximation for the number of collected photons m_c as about 1 per kV acceleration voltage. Reference [106] reports up to 8 photoelectrons per kV for NaI(Tl) (Thallium-doped Sodium Iodide) scintillators. The photocathode is quoted to have a detective quantum efficiency of $\eta \approx 25\%$. Using the photon count at the photocathode m_c and its quantum efficiency η as a starting point, the average number of photoelectrons per pulse can be stated using equation 2.3:

$$n_p = m_c\eta \tag{2.5}$$

Given that the variance of the photoemission process is [106]

$$\sigma^2 = \eta - \eta^2 \tag{2.6}$$

the variance in the number of photoelectrons, using equation 2.4 is given as:

$$\sigma^2_p = \eta^2\sigma^2_c + m_c(\eta - \eta^2) \tag{2.7}$$

It follows the **multiplier chain** of the PMT. The gain of the photomultiplier G_{PM} over n dynodes, would be given by generalising equation 2.3 as:

$$G_{PM} = \delta_1\delta_2...\delta_n = \prod_{k=1}^{n}\delta_k \tag{2.8}$$

Under the assumption that all dynodes have equal gain, the expected amount of electrons on the anode of the PMT per incident electron may be estimated as the detector gain G.

$$G = m_c\eta\delta^n \tag{2.9}$$

The distribution of a single electron response of a dynode can be generally described by a Pólya distribution.

$$P(k|\lambda) = \frac{\lambda^k}{k!}(1+b\lambda)^{-k-1/b}\prod_{j=1}^{k-1}(1+jb), k = 0, 1, 2, ... \tag{2.10}$$

Setting the parameter $b = 0$ in equation 2.10 leads to the special case of the Poisson distribution in equation 2.1 [106, p. 164]. This is a reasonable simplification, provided that inhomogeneities across the dynodes are small [106, 107]. The amplification can thus be seen as a cascade of Poisson processes. Further, using again the assumption that all dynode gains are equal, following the derivation given in reference [106] eventually yields an expression for the variance in the response of a single electron. Since the the number of photons reaching the photocathode per incoming electron is large ($\approx$ 100) [103], and the gain is also typically in the order $G = \delta^k \approx 4^{10}$ [105] a gaussian approximation is suitable to describe this distribution [108] with a standard deviation of:

$$\sigma = m_c \eta G \sqrt{\frac{1 - \eta + \frac{1}{\delta - 1}}{m_c \eta} + \frac{\delta_c^2}{m_c^2}} \tag{2.11}$$

Therefore, in order to model the noise accurately for given simulation x one would need a two-step process. Firstly, x is scaled by a dose-dependent parameter α and subjected to Poisson noise $\mathbb{P}$

$$x \leftarrow \mathbb{P}(\alpha x) \tag{2.12}$$

Secondly, noise from a gaussian distribution $\mathbb{N}$ is added, where the standard deviation σ is dependent on the updated x, because m_c in equation 2.11 must be scaled accordingly.

$$x \leftarrow (x + \mathbb{N}(0, \sigma(x)))/\alpha \tag{2.13}$$

Dividing by α in the latter equation brings x back to approximately its original range. The most obvious effect of this accurate modelling of the detector is that the values produced by single electrons are actually subject to distribution spreads, which can clearly be seen in figure 2.3, which shows a simulated low-dose background noise in panel a) and its histogram in panel b). In the histogram the peaks are evenly spaced and each peak represents a given electron count. Although this is a simulation according to equations 2.11, 2.12 and 2.13 this can be experimentally verified as for example in references [108] and [105]

2.2.2 Jitter

X- and Y-jitter are caused by beam instabilities while scanning line by line over the sample. X-jitter causes a shift Δ_j^x of the entire j-th horizontal scan line. While an image is recorded on a regular grid as shown in figure 2.4-a), the actual probe positions are displaced as a result of the these shifts along x (figure 2.4-b) and y (figure 2.4-c). X- and Y-jitter are caused by beam instabilities while scanning line by line over the sample. X-jitter causes a shift Δ_j^x of the entire j-th horizontal scan line. Y-jitter appears as a stretching/squishing of scan lines, or line interchanges, caused by an offset Δ_j^y of the vertical position of each scan line j. Both effects occur not fully random, but are dependent on the previous state (e.g. the position of the previous scan line).

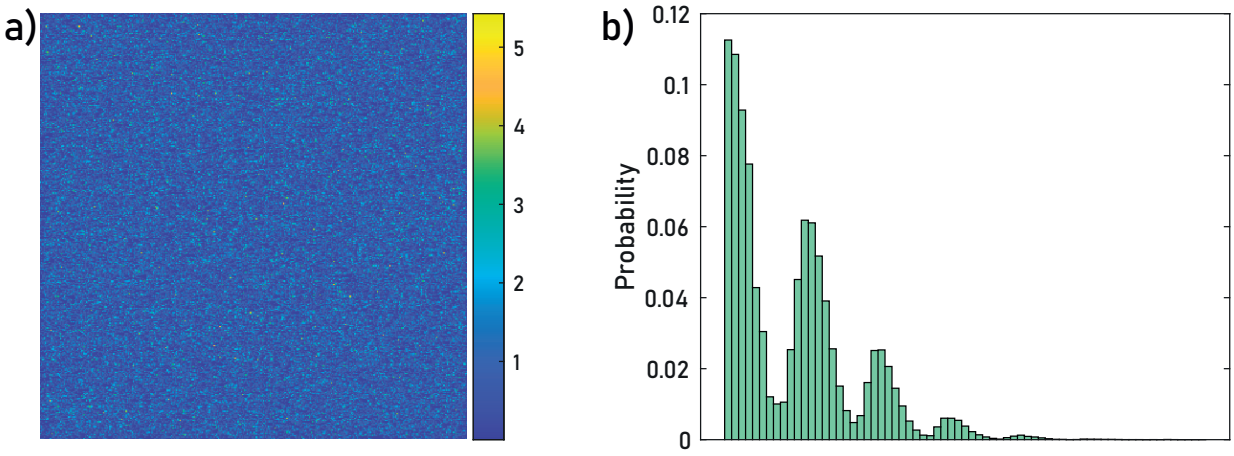


Figure 2.3 | HAADF detector noise distribution for a low dose background image. a) shows a simulated noise image according to equations 2.11, 2.12 and 2.13. b) shows the histogram of a), where distinct peaks can be seen, which correspond to the number detected electrons (e.g. 1st peak=1 electron, 2nd peak=2 electrons, etc.)

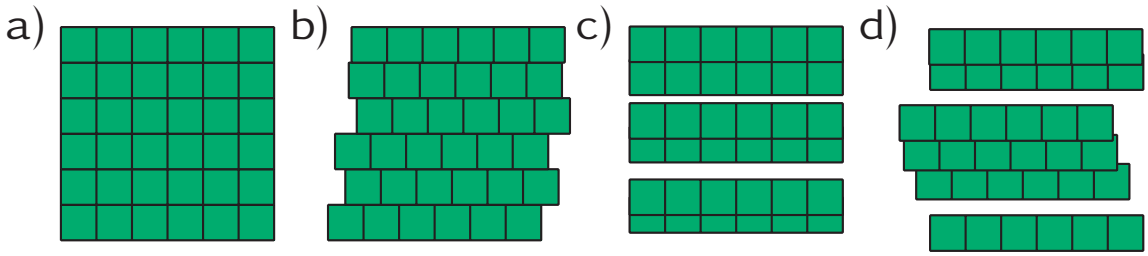


Figure 2.4 | Schematic illustration of distorted scan grids. a) perfect grid, b) x-jitter, c) y-jitter, d) combination of x-and y-jitter.

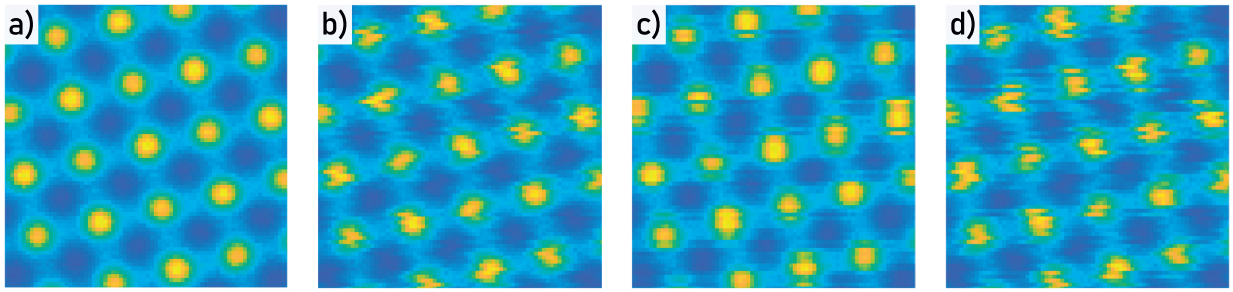


Figure 2.5 | Simulated examples of a Pt lattice, illustrating how jitter affects a given undistorted image (a). b) shows the effect of x-jitter only, c) the effect of y-jitter only and d) a combination of both.

Mathematically the series of j shifts along either direction may be described by series of related terms [109]:

$$\Delta_j^x = \begin{cases} \frac{a_j^x}{\sqrt{1-\phi_x^2}} & j = 1 \\ \phi_x \Delta_{j-1}^x + a_j^x & j > 1 \end{cases} \tag{2.14}$$

$$\Delta_j^y = \begin{cases} \frac{a_j^y}{\sqrt{1-\phi_y^2}} & j = 1 \\ \phi_y \Delta_{j-1}^y + a_j^y & j > 1 \end{cases}, \tag{2.15}$$

where a_j^x and a_j^y are random numbers following a normal distribution with zero mean and a standard deviations σ_x and σ_y. ϕ_x and ϕ_y are factors, describing how strongly the series value is coupled to the previous value. The scan line positions are shifted accordingly and the distorted image is created by linearly interpolating over the newly arranged scan lines.

2.2.3 Fast scan distortion

Fast scan distortion arises from a finite detector response time. At short dwell times, the beam scans too fast for the detector to process, resulting in a smearing along the scan line. The response function may be modelled as a decay function, adding a tail to objects in the image along the scan direction [110]. This effect is illustrated on single electron responses in figure 2.6-a). Figure 2.6-b) shows the characteristic damping of high frequencies in horizontal direction [108]. Mathematically, this may be expressed as a convolution of the image I with some point spread function $h(x)$ to compute a distorted model I_d.

$$I_d(x, y) = I(x, y) \circledast h(x) \tag{2.16}$$

The point spread function may be described in various ways, with one being a simple exponential decay [110], given in equation 2.17

$$h(x) = \begin{cases} e^{-\beta|x|} & x <= 0 \\ 0 & x > 0 \end{cases} \tag{2.17}$$

However, more recent studies suggest that the point spread may be more accurately described by a half-Lorentzian profile as given in equation 2.18 [2]. At extremely fast scan speeds this model may needs to be extended to also include a point-spread originating from the readout system [2].

$$h(x) = \begin{cases} \frac{\alpha}{4\pi^2 x^2 + \alpha^2} & x <= 0 \\ 0 & x > 0 \end{cases} \tag{2.18}$$

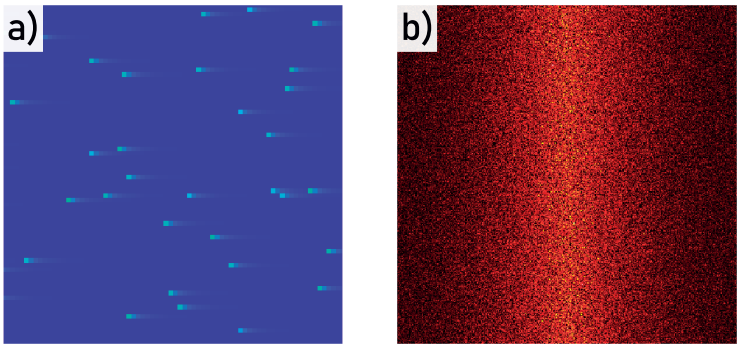


Figure 2.6 | HAADF detector noise for a low dose background image. a) shows a simulated noise image according to equations 2.11, 2.12 and 2.13 with an applied fast-scan distortion as given in equations 2.18 and 2.16. The smearing along the scan direction can be clearly seen on single electron signals. Panel b) shows the Fourier transform of a) with the fast-scan distortion-typical fingerprint in Fourier space.

2.3 Distortion correction & Atomic structure quantification

While the CNN from reference [2] was trained to restore images of a wide variety of imaging modes, STEM is of particular interest as it is routinely used for the quantification of atomic structures [89, 111, 112] in terms of atomic column positions and their corresponding SCS, which allows us to study the impact of the image restoration method quantitatively. The evaluation of the effect image restoration has on the quantitative assessments of STEM images is done in three complementary approaches, using MULTEM [60, 61] to create multislice simulations and the StatSTEM software for all model fittings [89]. All evaluations are based on 100 distortion/noise realisations for each dose setting.

1. We demonstrate the effect of image denoising with an idealised setup in analogy to the study conducted in reference [89], where the precision of the determination of the location and scattering cross section of an atomic column was determined over a wide range of signal-to-noise-ratios (SNRs). This setting allows the comparison to the theoretical limits of variance, the Cramér–Rao-Lower Bounds(CRLBs). The simulated STEM dataset is of a bulk Pt crystal in [001] orientation and contains STEM images over 75 depth sections with unit cell spacing in z-direction.

2. A more practical example, that includes crystal irregularities, is chosen to determine the impact of a combination of noise, scan-line-distortions and fast-scan distortion. In this case, we evaluate the mean absolute error (MAE) for atomic column positions and the mean absolute percentage error (MPE) for the scattering cross sections of atomic columns, as well as the variance of these measurements. This serves to show in particular the independence of the approach on the structural periodicity for atomic-resolution STEM images.

3. For a simulated Pt-nanoparticle it is demonstrated that distortion correction yields not only a more accurate localisation of atomic columns but also enables more reliable atom counting.

2.3.1 Simulations

All specimens were modelled and simulated with the MULTEM software [61], using the microscope- and aberration settings, as summarised in table 2.1. The electrostatic potentials were modelled using the parameterisation method by Lobato and Van Dyck [113]. Incoherent scattering effects were included employing the frozen phonon model with 200 phonon configurations. The effect of source size broadening of the electron probe was accounted for by a subsequent convolution of the simulated images with a two-dimensional, normalised, symmetric Gaussian function with a width given as a full-width-half-maximum (FWHM) $FWHM_{SB}$.

Table 2.1 | Simulation parameters for all multislice simulations used in section "Results - Atomic structure quantification".

Description	Parameter	Bulk	Dislocation	Particle
Acceleration voltage	E_0	300 kV	300 kV	300 kV
Detector inner angle	Θ_{di}	60 $mrad$	60 $mrad$	42 $mrad$
Detector outer angle	Θ_{do}	190 $mrad$	190 $mrad$	220 $mrad$
Outer aperture angle	Θ_{ai}	21.1 $mrad$	21.1 $mrad$	21.0 $mrad$
3rd order spherical ab.	C_{30}	0.04 mm	0.04 mm	0.04 mm
5th order spherical ab.	C_{50}	10 mm	10 mm	10 mm
Defocus	C_{10}	–8.874 nm	–8.874 nm	–10.868 nm
Source size broadening	$FWHM_{SB}$	0.7Å	0.7Å	0.75Å
Pixel size	Δ_x	0.15Å	0.16Å	0.12Å
# Phonon configurations	n_{FP}	200	200	200

Table 2.2 | Distortion modelling parameters for simulations of the edge-dislocation and nanoparticle samples used in section "Results - Atomic structure quantification"

Distortion	Parameter	Value
Poisson noise (Dose)	$d[e^-/\text{Å}^{-2}]$	$5e2, 1e3, 5e3, 1e4, 5e4, 1e5, 5e5$
X-Jitter	ϕ_x	0.5
	σ_x	0.5
Y-Jitter	ϕ_y	0.5
	σ_y	0.25
Dwell time	t_{fs}	$0.5\mu s$

2.3.2 Parameter estimation & metrics

Model fitting

The StatSTEM [89] software was used to determine atomic column positions and SCS. Each atomic column is modelled as a two-dimensional, symmetrical gaussian function, and the entire image is modelled as a superposition of those gaussian peaks. The expectation model $f_{kl}(\Theta)$ is therefore given for any pixel over a discrete grid (k, l) with position (x_k, y_l):

$$f_{kl}(\Theta) = \zeta + \sum_{n=1}^{N} \eta_n \exp\left(-\frac{\left(x_k - \beta_{x_n}\right)^2 + \left(y_l - \beta_{y_n}\right)^2}{2\rho_n^2}\right) \tag{2.19}$$

where ζ is a uniform background, the summation runs over N atomic columns in the model with index n. Each peak has the parameters η as the height, ρ the width and β_x and β_y as position coordinates. This model is fitted to the observation with a least-squares optimisation algorithm to determine the model parameters $\Theta = (\beta_{x_1}, \ldots, \beta_{x_N}, \beta_{y_1}, \ldots, \beta_{y_N}, \rho_1, \ldots, \rho_N, \eta_1, \ldots, \eta_N, \zeta)^T$ for all columns in the image. The SCS are calculated by integrating their corresponding gaussian functions:

$$SCS_n = 2\pi\eta_n\rho_n^2 \tag{2.20}$$

When all the estimated SCS in an image are visualised in a histogram, ideally one would see distinct peaks where each peak relates to a certain number of atoms. This enables atom counting in principle. However, in practice peaks are usually smeared out and overlap, such that they can neither be distinguished easily nor is the correct number of distinct peaks obvious or known. To conclude on atom counts in a reliable manner the distribution of SCS is modelled as a gaussian mixture model, where the probability distribution and the number of significant components need to be estimated [114]. Alternatively the fitted SCSs may also be compared to simulated library values and assigned to the closest match.

Metrics

The metrics used for evaluating the errors between the models fitted to raw- and restored images in this section are given in the following overview, where the variable N denotes the number of samples considered (i.e. noise realisations) and C the number of atomic columns per sample. We define the mean absolute error (MAE) for position measurements as the mean error of the euclidean norm for two-dimensional positions with x and y coordinates. This metric represents the mean of the magnitude of the displacement between the sets of atomic columns a and b, which corresponds to the fitted trial and known ground-truth data in this context.

$$MAE_{xy} = \frac{1}{NC}\sum_{n}^{N}\sum_{c}^{C}\sqrt{(x_{a,n,c} - x_{b,c})^2 + (y_{a,n,c} - y_{b,c})^2} \tag{2.21}$$

For SCS measurements, the relative error is a more meaningful description because atoms in each column should be counted correctly, independent of the actual number of atoms. Therefore we relate the measurements to the ground truth and define the mean absolute percentage error (MAPE) for SCS measurements:

$$MAPE_{SCS} = \frac{100}{NC}\sum_{n}^{N}\sum_{c}^{C}\frac{1}{SCS_{b,c}}\left|SCS_{a,n,c} - SCS_{b,c}\right| \tag{2.22}$$

In accordance with reference [89], errorbars on MAE and MAPE are given by the $\alpha = 0.95$ confidence intervals, which are computed for any given set of values x of size N and confidence interval α by equation 2.23.

$$CI_{\bar{x}}(\alpha, x) = \left[\bar{x} - t_{1-\frac{\alpha}{2},N-1}\frac{s_x}{\sqrt{N}},\quad \bar{x} + t_{1-\frac{\alpha}{2},N-1}\frac{s_x}{\sqrt{N}}\right], \tag{2.23}$$

where $\bar{x}$ and s_x denote the sample mean and standard deviation respectively. $t_{1-\frac{\alpha}{2},N-1}$ corresponds to the $1-\frac{\alpha}{2}$ quantile of the student's t-distribution with $N-1$ degrees of freedom. Similarly, the confidence intervals on variances s_x^2 are computed using the $1-\frac{\alpha}{2}$ and $\frac{\alpha}{2}$ quantiles of the χ^2 distribution with $N-1$ degrees of freedom:

$$CI_{s_x^2}(\alpha, x) = \left[\frac{(N-1)s_x^2}{\chi^2_{\frac{\alpha}{2},N-1}}, \frac{(N-1)s_x^2}{\chi^2_{1-\frac{a}{2},N-1}}\right] \tag{2.24}$$

Cramér-Rao lower bound

The Cramér-Rao lower bound (CRLB) is a fundamental result in statistics that provides a lower bound on the variance of an unbiased estimator of a population parameter. Within the context of this study it is used to compare the obtained precisions of the model parameters, as outlined in section 2.3.2, against a theoretical lower bound. We assume a joint probability function $p(\omega)$ for statistically independent observations, modelled by a stochastic variable ω following a Poisson distribution:

$$p(\omega) = \prod_{k=1}^{K}\prod_{l=1}^{L}\frac{(\lambda_{kl})^{\omega_{kl}}}{\omega_{kl}!}\exp(-\lambda_{kl}) \tag{2.25}$$

where the expectation values λ_{kl} are obtained by an appropriate parametric model f_{kl} (in this case given by equation 2.19). The dependence on ω can be used to create a covariance matrix, known as the Fisher information matrix F_θ:

$$F_\theta = -\mathbb{E}\left[\frac{\partial^2 \ln p(\omega;\theta)}{\partial\theta^2}\right] \tag{2.26}$$

The individual elements of F_θ, with indices (r, s) can be computed according to the joint probability function given in equation 2.25:

$$F_{rs} = \sum_{k=1}^{K} \sum_{l=1}^{L} \frac{1}{\lambda_{kl}} \frac{\partial \lambda_{kl}}{\partial \theta_r} \frac{\partial \lambda_{kl}}{\partial \theta_s} \tag{2.27}$$

The diagonal elements of this matrix define lower bounds on the model parameters θ. To obtain the lower bound on the variance of the derived parameter SCS_n (equation 2.20), the general rule for the CRLB of a function should be used[89, 115]:

$$\sigma^2_{SCS_n} = \frac{\partial SCS_n}{\partial \theta} F_\theta^{-1} \left(\frac{\partial SCS_n}{\partial \theta} \right)^T \tag{2.28}$$

Since SCS_n depends only on the parameters ρ_n and η_n, the vector $\frac{\partial SCS_n}{\partial \theta}$ will have only 2 non-zero elements accordingly:

$$\begin{aligned} \frac{\partial SCS_n}{\partial \rho_n} &= 4\pi\eta_n\rho_n \\ \frac{\partial SCS_n}{\partial \eta_n} &= 2\pi\rho_n^2 \end{aligned} \tag{2.29}$$

2.3.3 Results & Discussion

The results of the first study are shown in figure 2.7. Examples of the underlying STEM images are given for the extremes of signal-to-noise-ratios (i.e. smallest thickness and lowest dose and largest thickness and highest dose) for raw and restored images in panels (e) and (f) and (g) and (h) respectively. Comparing figure 2.7-(e) and (f) it can be seen visually, that even at a very low dose, the CNN can recover the underlying structure faithfully. This effect can be quantified in terms of the precision with which both, the atomic column positions and SCSs can be measured, and is particularly pronounced in the low dose range as illustrated in figure 2.7-(a) and (b). As the dose increases the precision of the structural measurements converge eventually (figure 2.7-(c)-(d)). An interesting observation is that the theoretical precision limit given by the CRLB, can be overcome employing image restoration. This makes a strong point for using image restoration for quantitative studies, like atom counting or strain measurements in general. The restoration results in the first example arguably benefit from the underlying perfect crystal symmetry, which is why we test the CNN also for imperfect structures. The Pt-bulk model depicted in figure 2.8-(a) is in [112] zone axis orientation, six unit cells thick and contains a unit edge dislocation of Burgers vector $b = 1/2[110]$ in the (111) glide plane; a dislocation commonly observed in fcc metals [116]. The structure was created using the Atomsk software, which determines atom positions corresponding to the displacement fields predicted by the elastic theory of dislocations [117]. The simulated HAADF images were subjected to varying noise levels from 5e2 $e/Å^2$ to 5e4 $e/Å^2$, and further corrupted by scan-line distortions as outlined in section "S(T)EM noise model". Example reconstructions for raw images at doses of 5e2 $e/Å^2$ and 5e4 $e/Å^2$

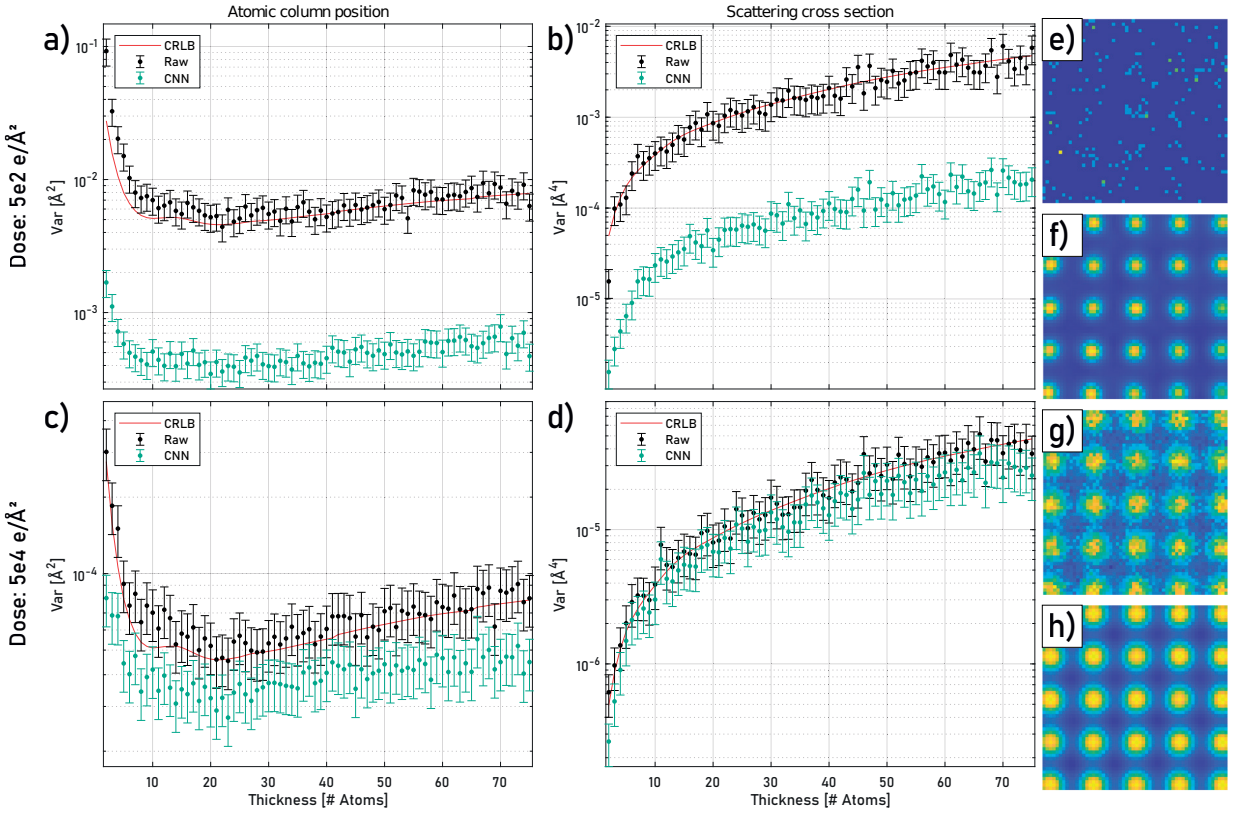


Figure 2.7 | Precision of atomic column position and SCS-measurements over a series of Pt-bulk samples with a thickness varying from 2-75 unit cells. Error-bars are given as the 95% confidence intervals. (a) Position precision for dose 5e2 $e/Å^2$ (b) SCS precision for dose 5e2 $e/Å^2$ (c) Position precision for dose 5e4 $e/Å^2$ (d) SCS precision for dose 5e4 $e/Å^2$ (e) Example of a raw STEM image at thickness=2 atoms and dose=5e2 $e/Å^2$ (f) example of a raw STEM image at thickness=75 atoms and dose=5e4 $e/Å^2$ (g) example of a restored STEM image at thickness=2 atoms and dose=5e2 $e/Å^2$ (h) example of a restored STEM image at thickness=75 atoms and dose=5e4 $e/Å^2$

(figure 2.8-(b) and (c)) are shown in figure 2.8-(d) and (e) respectively. In the low-dose raw image individual atomic columns are hardly recognisable. Without the prior knowledge of the atomic column positions any attempt of model fitting would have to overcome the challenge of performing reliable peak finding first, which is a factor not considered here. The reconstruction of this image (figure 2.8-(d)) on the other hand shows very clear peaks. A Burgers circuit is superimposed on the image to highlight that despite the poor separation of columns in the raw image, the dislocation with its correct Burgers vector *b* is maintained, which means that the structure at large is correctly retrieved, albeit the individual column positions may not be fully accurate as seen in the mean absolute position error of the columns in around the centre of the dislocation (columns within the red circle in figure 2.8-(a)) for low doses shown in figure 2.8-(f). However, the error drops rapidly with increasing dose and shows a clear improvement against raw images. The position accuracy is therefore not only a result of denoising but also the result of the accurate correction of scan-line and fast-scan distortions. The comparatively high accuracy for the raw image fitting at low doses can be attributed to the fact that

correct initial column positions are given for the fitting procedure. Since the column can hardly be located in the noisy images, the fitting algorithm does not actually move the position away from this initial position much on average. The CNN on the other hand reconstructs a clearly visible atomic column, but the available information in the underlying image is insufficient for an accurate positioning. However, the proper retrieval of the dislocated atomic column at higher doses shows that the CNN is not by default just picking up on periodicity, but faithfully recovers the atomic structure also in the presence of non-periodic features in atomic resolution STEM images.

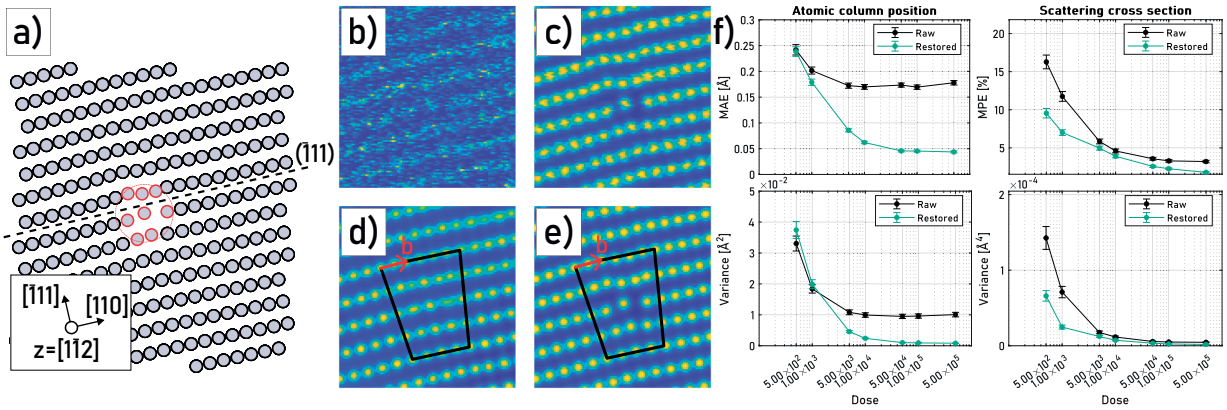


Figure 2.8 | (a) Schematic of the Pt structure in [112] zone axis with a unit edge dislocation of Burgers vector $b = 1/2[110]$ in the (111) glide plane. (b) Corrupted raw image with a dose of $500e/Å^2$. (c) Corrupted raw image with a dose of $5e5e/Å^2$. (d) Restored image with a dose of $500e/Å^2$. (e) Restored image with a dose of $5e5e/Å^2$. (f) Quantification results for the atomic column positions and scattering cross sections of the atomic columns around the centre of the edge dislocation (marked with a red circle in panel (a)).

Also the SCS measurements improve in accuracy by the restoration, which would translate directly into improvements for atom counting studies. An example for such an atom counting scenario is presented in figure 2.9. These results were obtained from a simulated spherical Pt nanoparticle with 11 unit cells in diameter in [100] zone axis orientation under the same distortion and noise parameters as given in the previous example. Atom counts were obtained by matching retrieved SCS values against simulated library values. The improvement in column position measurements over all dose settings again indicates the proper correction of scan-line and fast-scan distortions. The improvement of SCS measurement accuracies, especially at low-dose conditions greatly decreases the chance of miscounting atoms in the structure, which in turn may be very beneficial for e.g. the reconstruction of 3D information from atom-counts [16, 17].

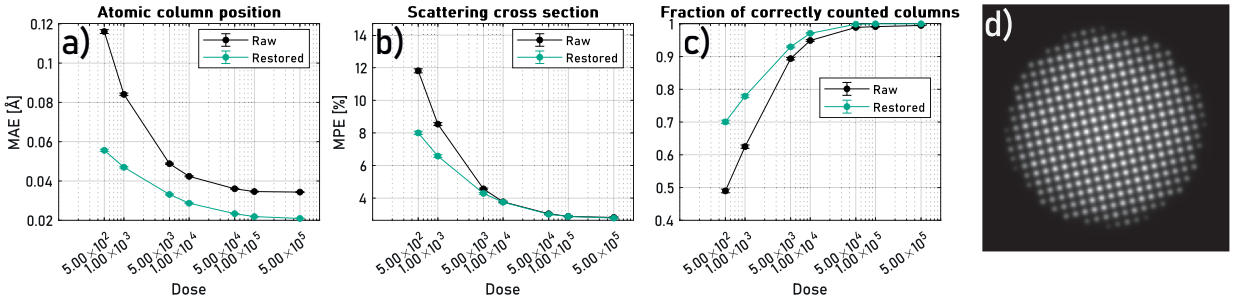


Figure 2.9 | Quantification results for a spherical Pt nanoparticle with a diameter of 11 unit cells in [100] orientation. The values are based on all 333 atomic columns.

2.4 Conclusions

In this chapter the most prominent distortions in STEM are reviewed and mathematical models to describe them presented. While the description of noise characteristics is specific to scintillator-PMT type detectors, scan line distortions apply to spectral- and 4D-STEM data just the same. It was shown that a convolutional neural network can be used to correct these distortions in a single shot [2]. This approach is not only very practical and creates nicer images, but also aids and improves the entire process of atomic structure determinations, starting with the mere identification of column positions(e.g. for the purpose of peak finding in noisy images). Also the quantification of structures was shown to improve from the restoration. We could show that the errors and variances in the measured atomic column positions and SCSs are significantly reduced in three separate simulation studies, which indicates that image restoration may be a very easy way to improve structure quantifications and due to its ease of use may be well suited as a general first step in structure determination studies.

2.5 Additional information

CNN implementation

The model and training designs were developed by Ivan Lobato alone. The relevant implementation details here are provided for context from reference [2].

In single-shot EM image restoration, the goal is to estimate an undistorted image y from a distorted image x. To achieve this, we train a generator G using a deep neural network approach, which learns to estimate the corresponding undistorted image y for a given input x. During the training procedure, a loss function is minimised to evaluate the quality of the results.

Traditionally, pixel-wise losses such as $\mathcal{L}_1$ or $\mathcal{L}_2$ have been used to obtain quantitative results for the image restoration problem [118]. However, these losses often lead to blurred images that do not look realistic. To address this, we propose a conditional generative adversarial network (cGAN) that trains both a generator and a discriminator. The generator G maps the distorted image x to the undistorted image $y_g = G(x)$, and the discriminator is trained to differentiate between real and generated images [119].

We use pixel-wise losses to ensure quantitative results while restricting the GAN discriminator to model high-frequency details, resulting in sharper and more realistic restored images.

Our training is supervised, which requires input pairs of distorted and undistorted EM images. However, in practice, we only have access to distorted EM data. To overcome this, we can partially address the problem by collecting time-series EM images and using an average procedure based on rigid and non-rigid registration to generate an undistorted image. However, the combination of high-speed scans, jitter, and low-dose leads to highly correlated distortions [120]. Furthermore, long exposure to the electron beam can result in charging, beam damage, atom hopping and rotation of the specimen under study, which can further hamper the average procedure. Therefore, the only solution is to train the GAN using synthetic pairs of undistorted/distorted EM images.

Network architecture

A GAN [71] is a powerful framework that encourages predictions to be realistic and thus to be close to the undistorted data distribution. A GAN consists of a generator (G) and discriminator (D) playing an adversarial game. A generator learns to produce output that looks realistic to the discriminator, while a discriminator learns to distinguish between real and generated data. The models are trained together in an adversarial manner such that improvements in the discriminator come at the cost of a reduced capability of the generator and vice versa. The GAN involves the generation of conditional data, which is fed to the generator and/or the discriminator [121]. The generator and discriminator architectures proposed here are adapted from those described in [122] and [119], respectively. The details of these architectures are discussed in the following sections.

Generator architecture

Our generator architecture, called Concatenated Grouped Residual Dense Network (CGRDN), is shown in Fig. 2.10. This network architecture is an extension of the GRDN for image denoising [123], which was ranked first for real image denoising in terms of the PSNR and the structure similarity index measure in the NTIRE2019 Image Denoising Challenge [124]. The GRDB architecture is shown in Fig. 2.10(a). The building module of this architecture is the residual dense block (RDB) [122], which is shown in Fig. 2.10(b). The original GRDN architecture can be conceptually divided into three parts. The first part consists of a convolutional layer followed by a downsampling layer based on a convolutional stride, the middle part is built by cascading GRDBs and the last part consists of an upsampling layer based on transposed convolution followed by a convolutional block attention module (CBAM) [125] and a convolutional layer. The GRDN also includes the global residual connection between the input and the last convolutional layer. In the original version of the GRDN [123], residual connections are applied in three different levels (global residual connection, semi-global residual connection in GRDB, and local residual connection in each RDB). However, in the version submitted for the NTIRE2019 Image Denoising Challenge [124], residual connections for every 2 GRDBs were included.

Although it has been demonstrated that one architecture developed for a certain image restoration task also performs well for other restoration tasks [119, 122, 126, 127], an

architecture for a given task will be data dependent. When applied to EM data, we found out that 2 modifications of GRDN are necessary in order to best handle the nature of our data, which involves different types and levels of distortions with high correlation between pixels:

1. The cascading of the GRDN is replaced by feature concatenation, feature fusion, and a semiglobal residual connection. This allows us to exploit hierarchical features in a global way, which is important for highly correlated pixels that extend over a large area of the image.

2. The CBAM, which is included in [122] is removed from our network. The reason for this is the use of large image sizes (256x256) for training, which reduces its gain [123].

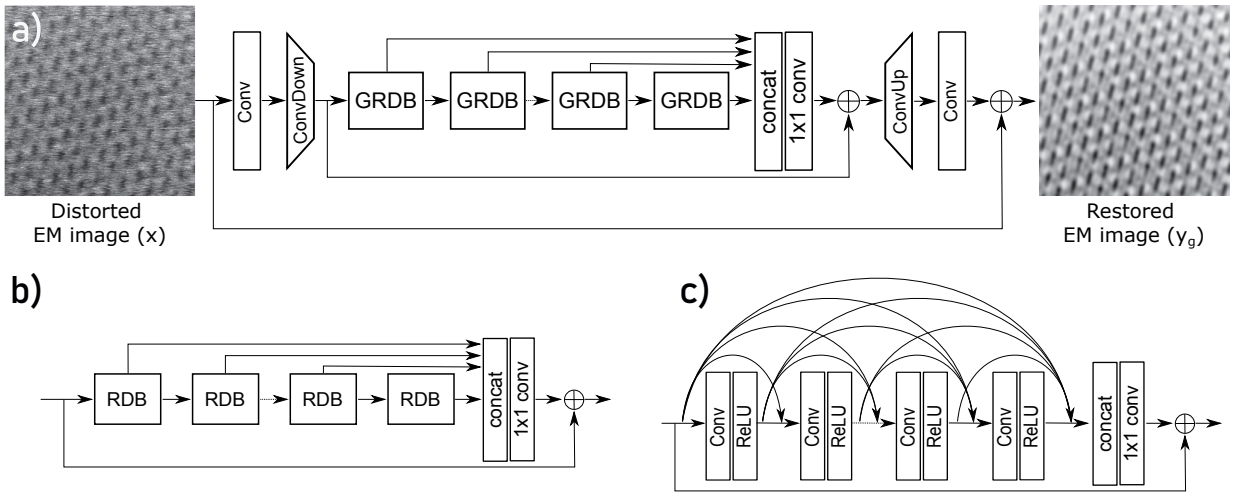


Figure 2.10 | Concatenated Grouped Residual Dense Network (CGRDN) architecture for EM image restoration. (a) Overall architecture, (b) GRDB architecture used in (a), (c) RDB architecture used in (b).

Discriminator architecture

The purpose of the discriminator network is to judge the quality of the output data resulting from the generator network. For our discriminator, we use the 70x70 convolutional patch discriminator described in [119] with some minor modifications. The zero-padding layers were removed and batch normalization layers [87] were replaced by instance normalization layers (IN) [128]. Figure 2.11 shows the structure of the discriminator network. The result of the network is the non-transformed output $C(y)$ or $C(y_g)$ of dimensions 32x32. Some benefits of the discriminator architecture shown in Fig. 2.11 include that it is fully convolutional and it only penalizes structure at the scale of image patches. Furthermore, we enhance our discriminator based on the relativistic GAN, which has been shown to improve the data quality and stability of GANs at no computational cost [129]. Different from the standard discriminator, which estimates the probability that input data is real, a relativistic discriminator predicts the probability that real data y is relatively more realistic than generated data $y_g = G(x)$. If we denote our relativistic

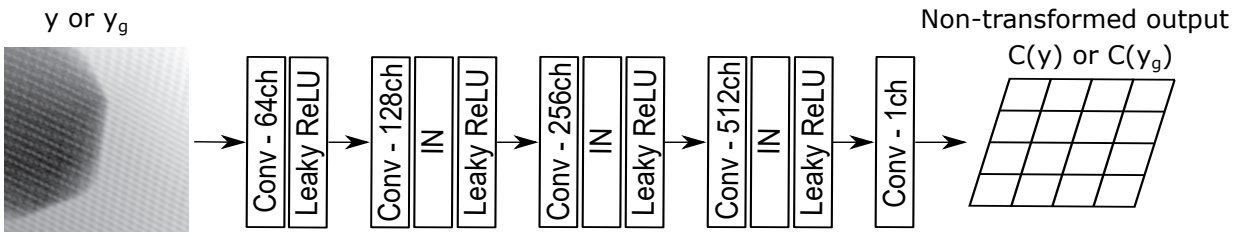


Figure 2.11 | Patch discriminator architecture.

average patch discriminator as $D_{Rap}(x)$, then the output of the discriminator can be written as:

$$D_{Rap}\left(y, y_g\right) = \sigma\left(C(y) - \mathbb{E}[y_g]C(y_g)\right) \tag{2.30}$$

$$D_{Rap}\left(y_g, y\right) = \sigma\left(C(y_g) - \mathbb{E}[y]C(y)\right) \tag{2.31}$$

where σ is the sigmoid function and $\mathbb{E}[x_1, \ldots x_n]$. is an operator representing the expectation value computed on the variables $x_1, \ldots x_n$. In the next section, these functions will be used in the definition of the loss functions.

Loss function

The loss function is the effective driver of the network's learning. Its goal is to map a set of parameter values of the network onto a scalar value, which allows candidate solutions to be ranked and compared. In our case, the discriminator and adversarial losses are based on the relativistic average GAN loss defined in [129]. We design our generator loss function as a sum of different contributions in such a manner that it keeps the quantitative information of the image at the pixel level and produces perceptually correct and realistic images. The different contributions of these loss functions are described in the following sections.

$\mathcal{L}_1$ loss

Pixel-wise losses are advantageous to keep quantitative information of the ground truth image. In this work, we used the $\mathcal{L}_1$ loss, which as compared to the $\mathcal{L}_2$ loss yields less blurred results [118]. The $\mathcal{L}_1$ loss can be written as:

$$\mathcal{L}_1 = \mathbb{E}_{y,y_g}\left\{w_y \| y - y_g \|\right\}, \tag{2.32}$$

$$w_y = 1/\max\left(\sigma_{\min}, \mathrm{Std}_y\{y\}\right) \tag{2.33}$$

where w_y is a weighting factor that gives equal importance to each example regardless of its contrast, $\sigma_{\min}$ is a small value to limit the maximum scaling factor, and $\mathrm{Std}_{x_1,\ldots x_n}\{.\}$ is an operator that represents the standard deviation calculated on the variables $x_1, \ldots x_n$.

$\mathcal{L}_2$ loss

Due to the design of our architecture, which is learning the residual difference between the distorted and undistorted image and based on the fact that distorted images can

have few outliers in the distribution of pixel intensities (i.e. X-rays hitting the EM detector, saturation of the detector, low dose and dead-pixels), the output of the generator will show a strong correlation at those pixel positions. For this reason, we also used the $\mathcal{L}_2$ loss which strongly penalized the outliers:

$$\mathcal{L}_2 = \mathbb{E}_{y,y_g}\left\{w_y \|y - y_g\|^2\right\} \tag{2.34}$$

Multi-local whitening transform loss

Local contrast normalisation (LCN) is a method that normalises the image on local patches on a pixel basis [130]. A special case of this method is the whitening transform which is obtained by subtracting the mean and dividing by the standard deviation of a neighborhood from a particular pixel:

$$y_{ij}^S = \left(y_{ij} - \mathbb{E}_{\hat{S}}\left\{y_{i,j}\right\}\right) / \max\left(\sigma_{\min}, \mathrm{Std}_{\hat{S}}\left\{y_{i,j}\right\}\right), \tag{2.35}$$

where $\hat{S}$ is a local neighbourhood around the pixel i, j of window size S. The whitening transform makes the image patches less correlated with each other and can highlight image features that were hidden in the raw image due to its low local contrast. This effect can be seen in Fig. 2.12a), which shows a simulated ADF-STEM image of a random nanoparticle on a carbon support. The edge of the nanoparticle shows low contrast due to its reduced thickness, resulting in lower intensity values. Based on this observation, we introduce a multi-local whitening transform (MLWT) loss which pays more attention to fine details independent of the intensity value. Specifically, the generated and the ground truth image are local whitening transforms corresponding to different window sizes of $2x2$, $4x4$, $8x8$, and $16x16$ pixels. Then, we calculate the average $\mathcal{L}_1$ loss for these 4 images:

$$\mathcal{L}_{mlwt} = \frac{1}{4} \sum_{S=2,4,8,16} \mathbb{E}_{y^S, y_g^S}\left\{\|y^S - y_g^S\|\right\}. \tag{2.36}$$

Using different windows sizes for the calculation of the whitening transform, we ensure that the relevant features present in the image are highlighted independently of its pixel size. Figs. 2.12(b)-(e) show an enhancement of the edge of the nanoparticle as well as the carbon support after applying the whitening transform to Fig. 2.12(a) by using different window sizes.

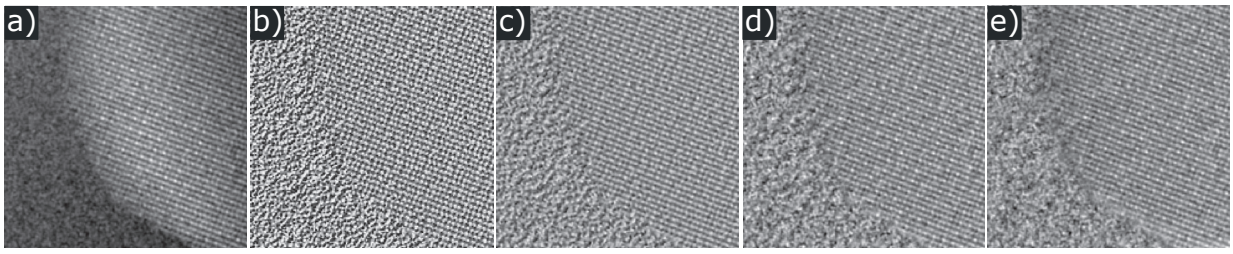


Figure 2.12 | a) Undistorted ADF STEM image of a nanoparticle on a carbon support. Images are generated by applying the whitening transform to (a) by using different window sizes of (b) 2, (c) 4, (d) 8 and (e) 16.

Fourier space loss

In electron microscopy, the Fourier space contains crucial information about the sample and any distortions that may be difficult to discern in real space. To address this issue, we introduce the $\mathcal{L}_\gamma$ loss in the 2D Fourier transform of the difference between the generated data y_g and the ground truth image y. The loss is defined as follows:

$$\mathcal{L}_{fs-\gamma} = \mathbb{E}_{y,y_g}\left\{\|\mathcal{F}(y-y_g)\|^\gamma\right\}, \tag{2.37}$$

Here, $\mathcal{F}$ represents the 2D Fourier transform, and γ is a parameter in the range $(0.0, 1.0]$. In our study, we use $\gamma = 0.125$.

Constraint losses

Some important parameters for EM quantification are the total intensity and the standard deviation of the images. The reason for this is that they carry information about physical quantities of the sample or microscope, such as the number of atoms, defocus and spatial and temporal incoherence [131, 132]. Therefore, we encourage that the restored images have to minimize the above quantities, resulting in the following two loss functions:

$$\mathcal{L}_{mean} = \|\mathbb{E}_y\{y\} - \mathbb{E}_{y_g}\left\{y_g\right\}\|, \tag{2.38}$$

$$\mathcal{L}_{std} = \|\mathrm{Std}_y\{y\} - \mathrm{Std}_{y_g}\left\{y_g\right\}\|. \tag{2.39}$$

Adversarial loss

The job of the relativistic adversarial loss is to fool the discriminator which can be expressed as:

$$\mathcal{L}_{Adv} = -\mathbb{E}_{x,y}\left\{\log\left(1 - D_{Rap}(y, y_g)\right)\right\} - \mathbb{E}_{y_g}\left\{\log\left(D_{Rap}(y_g, y)\right)\right\}, \tag{2.40}$$

with $D_{Rap}(y, y_g)$ and $D_{Rap}(y_g, y)$ defined in equations 2.30 and 2.31, respectively. This definition is based on the binary cross entropy between the ground truth and the generated images. Different from the conventional adversarial loss, in which y is not used, our generator benefits from y and y_g in the adversarial training.

Generator loss

Our total generator loss function can be written as:

$$\mathcal{L}_G = \mathcal{L}_{pixel-wise} + \lambda_{Adv}\mathcal{L}_{Adv}, \tag{2.41}$$

$$\mathcal{L}_{pixel-wise} = \lambda_1\mathcal{L}_1 + \lambda_2\mathcal{L}_2 + \lambda_{mlwt}\mathcal{L}_{mlwt} + \lambda_{fs-\gamma}\mathcal{L}_{fs-\gamma} + \lambda_{mean}\mathcal{L}_{mean} + \lambda_{std}\mathcal{L}_{std}, \tag{2.42}$$

where $\mathcal{L}_{pixel-wise}$ is our pixel-wise loss function, λ_1, λ_2, λ_{mlwt}, $\lambda_{fs-\gamma}$, λ_{mean}, λ_{std} and λ_{Adv} are the weighting parameters to balance the different loss terms.

Discriminator loss

Symmetrically to the relativistic adversarial loss, the relativistic discriminator is trying to predict the probability that real data is relatively more realistic than generated data, and it can be expressed as:

$$\mathcal{L}_D = -\mathbb{E}_{x,y}\left\{\log\left(D_{cRap}(x, y, y_g)\right)\right\} - \mathbb{E}_{x,y_g}\left\{\log\left(1 - D_{cRap}(x, y_g, y)\right)\right\}. \tag{2.43}$$

CHAPTER 3

Real-time integration centre of mass (riCOM)

Contents

This chapter was published as: "C.-P. Yu, T. Friedrich, D. Jannis, S. Van Aert, and J. Verbeeck, "Real-Time Integration centre of Mass (riCOM) Reconstruction for 4D STEM," Microscopy and Microanalysis, vol. 28, no. 5, pp. 1526–1537, 2022. https://doi.org/10.1017/S1431927622000617"

3.1 Introduction

Scanning transmission electron microscopy (STEM) is one of the most powerful tools for inspecting materials with sub-nanometre or even sub-angstrom level resolution. By scanning with a sharp electron probe, information of the sample from each scan position is collected and images that contain features at the atomic level are generated. There are several methods to form images using the data collected from such experiments. Traditionally, detectors that capture electrons from certain ranges of scattering angles are used in the microscope. They generate a value based on the sum of received electrons at each probe position and result in 2D images. Images formed by detectors that collect signals at high scattering angles are even capable of reflecting the scattering power experienced by the electron probe at the corresponding probe position [22].

A pixelated detector does not generate a single value, but instead records a convergent beam electron diffraction pattern (CBED) for each probe position by using a large number of pixels, where each pixel can be seen as an individual detector. This results in a 4D dataset (2D CBEDs on a 2D scan grid). More importantly, these advanced direct electron detectors [41, 133, 134] record CBEDs at a much higher rate than traditional charge-coupled device detectors, and allow collecting a regular-sized 4D dataset before serious sample drifting happens.

To process 4D datasets, one is the filter function and the symbol can define virtual detectors by selecting specific groups of pixels on the detector plane for summation, which result in similar 2D images as traditional detectors would produce, or seek solutions from more advanced and complex methods. Most of these methods take into account the distribution of the electrons on the detector plane, as well as the relationship between CBEDs and their corresponding probe positions, allowing extra information to be extracted from the dataset. This enables reconstructions with resolution beyond the limitation imposed by the optical system [135] and can reduce the dose needed for microscopists to obtain the necessary information to analyse their samples. Within the category of 4D dataset processing methods, iterative optimization approaches [136–140] reconstruct subsets of the full dataset one region at a time. The process repeats and reprocesses each subset until the algorithm converges to an estimated version of electric potential distribution. Other methods that handle 4D datasets without an iterative process, for example single sideband ptychography (SSB) [141, 142], or integrated centre of mass (iCOM) or integrated differential phase contrast (iDPC) [142–145] reconstruction methods, have also proven to be much more dose efficient than traditional imaging methods. Compared to iterative processes, they are less computationally demanding and guarantee unique solutions since they do not depend on optimization algorithms. Also, some prior knowledge, such as the prediction of a phase distribution that may arise from astigmatism and defocus, can be provided to this post process for acquiring more detailed information [146]. However, the ability to achieve fast reconstructions, regardless whether they are iterative or not, usually relies on accelerators (e.g. GPU) as well as large amounts of computer memory in order to fit in the whole dataset, or some reduced version of it. With an exception of iCOM, most of these post processing methods are thus limited by the hardware to a certain number of probe positions.

Even though the reconstruction methods may be further optimized to reduce the processing time, users still need to wait for the recording of the dataset to be completed before a resulting image can be generated. This waiting time varies, but for datasets composed of a large number of scan points or in situations where the detector has a slow frame rate, this delay would hinder the process of searching for features of interest, as well as adjusting the optical system based on the observations. Some rather simple approaches such as traditional imaging methods, centre of mass (COM) shift, or some of its derivatives such as COM shift divergence [147] can effectively reduce or eliminate this delay. However, these methods also require a higher number of electrons to generate images with adequate quality, compared to more complex methods such as SSB, iCOM, and iDPC. As proposed by A. Strauch et al. [148], a dose efficient reconstruction with live image update can be done by first allocating memory for the dataset, and then gradually filling it with collected and processed data during the scanning process. An update of the reconstructed image can be generated anytime by SSB reconstruction, even before the dataset is complete. However, it also indicates that the number of probe positions in a dataset is limited by the GPU memory, as it needs to store data for later processing. At the current state of technology, this approach is limited in terms of processing rate to about 1000 probe positions per second in the implementation of Strauch et al. [148], while the collection frame rate of direct electron detectors is approaching 100 kHz [146] and even the MHz range for event driven cameras at suitable conditions [149].

To overcome these hardware and speed limitations, we hereby propose a new live reconstruction method based on iCOM, which does not rely on storing the entire 4D dataset in memory, does not require accelerators of any kind and thus greatly reduces the computational requirements, as well as allowing reconstructions of images of a larger scale. In this chapter, the physical formulation of real-time iCOM (riCOM) is first derived, and details of the software implementation of the reconstruction algorithm are discussed. This software implements a direct interface to the electron camera, and several real-time reconstructed results are recorded, from which one can see that the tuning of the imaging conditions are immediately reflected in live-updated images. RiCOM reconstruction from existing experimental datasets are also shown. These datasets are recorded frame-by-frame or per-event [149, 150]. Both formats can be processed with the riCOM method with little alteration of the algorithm. Reconstruction results with different range of integration and integrated filters are also displayed. They are compared with each other and with other reconstruction methods to put the proposed method into context.

3.2 Methods

3.2.1 Physical Formulation

In 4D STEM, the distribution of the electron intensity at each probe position is recorded. The centre of mass of this distribution can then be calculated, resulting in a vector image $\vec{I}^{COM}(\vec{r}_p)$ or two scalar images describing its x component $I^{COMx}(\vec{r}_p)$ and y component $I^{COMy}(\vec{r}_p)$. For the x-component,

$$I^{COMx}(\vec{r}_p) = \int \int_{-\infty}^{\infty} k_x I(\vec{k}, \vec{r}_p) d^2\vec{k}, \tag{3.1}$$

where $\vec{r}_p$ is the probe position, $\vec{k}$ indicates a point on the detector plane with components k_x and k_y, and $I(\vec{k}, \vec{r}_p)$ is the intensity at $\vec{k}$ while the probe is situated at $\vec{r}_p$. From previous work [144] it follows that (derivation in supplementary)

$$\begin{aligned} \vec{I}^{COM}(\vec{r}_p) &= \frac{1}{2}|\psi_{in}(\vec{r}, \vec{r}_p)|^2 \star \nabla\phi(\vec{r}) \\ &= \frac{1}{2}\nabla(|\psi_{in}(\vec{r}, \vec{r}_p)|^2 \star \phi(\vec{r})) \\ &= \nabla O(\vec{r}_p). \end{aligned} \tag{3.2}$$

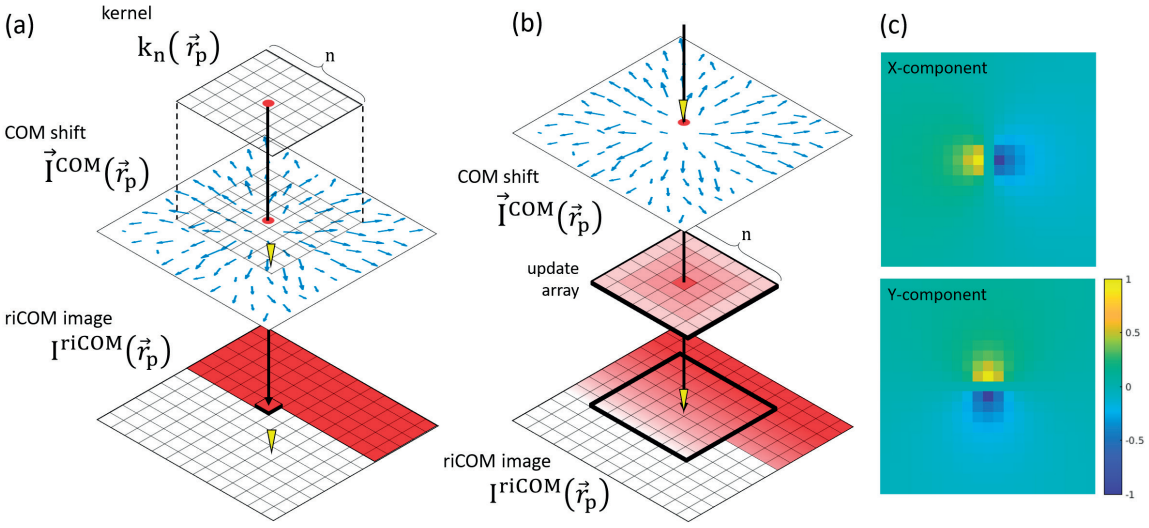


Figure 3.1 | (a) The kernel takes multiple data points from the COM shift map to calculate the value for one pixel in the riCOM image. (b) The riCOM image is being updated based on the contributions of the COM shift at one probe position. The yellow triangles indicate the scanning probe position. (c) X and Y components of a kernel of size 21×21.

In equation 3.2, the COM shift signal is understood as the gradient (∇) of a function $O(\vec{r}_p)$, which is the local projected potential $\phi(\vec{r})$ cross-correlated ($\star$) with the intensity distribution of the incoming electron beam at a given probe position $|\psi_{in}(\vec{r}, \vec{r}_p)|^2$. Note that this result is achieved under the phase object approximation, which assumes that the electron probe remains unmodified while passing through the object. With this approximation the 3D potential established by the material is simplified to a projected potential in a 2D plane. It clearly fails for thicker objects, but it allows a simple derivation and easy understanding of how experimental conditions can affect reconstructed images.

To solve for a scalar function describing the object, path integration is performed on the COM shift signal to remove the gradient from the right-hand side of equation 3.2. For an ideal case the path of the integration can be taken arbitrarily, since the integral

is only dependent on the end point of the path integration. However, in realistic cases, the measurement of COM shift contains noise, and thus it would give better estimation of the noise-free result by taking the average of all possible path integrals. By the assumption that equipotential can be found at infinity, this can be achieved by averaging path integrals at all possible azimuthal angles, from infinity towards the probe position. In order for this concept to work with a 2D grid of probe positions, the averaged integral can be expressed in a discretized form:

$$\begin{aligned} & O(\vec{r}_p) \\ &= \int_{\infty}^{\vec{r}_p} \vec{I}^{COM}(\vec{r}) \cdot d\vec{r} \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_{\infty}^{r_p} \vec{I}^{COM}(\vec{r}) \cdot \hat{n} dr d\theta \\ &= \frac{a}{2\pi} \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} \frac{\vec{r}_p - \vec{r}_{xy}}{|\vec{r}_p - \vec{r}_{xy}|^2} \cdot \vec{I}^{COM}(\vec{r}_{xy}). \end{aligned} \tag{3.3}$$

In the continuous representation of the radial averaged path integral (third line of equation 3.3), $\hat{n}$ is a unit vector pointing towards $\vec{r}_p$. In the discrete representation (fourth line of equation 3.3), $\vec{r}_{xy}$ describes a vector pointing at each probe position that composes the 2D array, and a describes a factor proportional to the square of the step size taken to discretize the integration.

The discrete representation in equation 3.3 states that the summation has to go over an infinite amount of points, or at least all probe positions in the dataset (as for iCOM reconstructions) in order to acquire or to approximate the desired object function. This would require the full dataset to be collected first, and rendering a live update of the partially reconstructed dataset is therefore impossible. However, it is found that by limiting the spatial range of the summation, the algorithm results in similar reconstructions as iCOM, but with more emphasis on local variations of the object function. This behaviour can be understood qualitatively. The term $(\vec{r}_p - \vec{r}_{xy})/|\vec{r}_p - \vec{r}_{xy}|^2$ describes an odd function since the vector distribution on both sides of the probe position $\vec{r}_p$ is the same in magnitude but opposite in direction as the sign changes for $\vec{r}_p - \vec{r}_{xy}$. For a global homogeneous COM shift, or for cases where the variation is negligible within the range of the kernel size, it results in an even function for $\vec{I}^{COM}(\vec{r}_{xy})$, and thus the sum of the product of the two will always be zero. But for short range variations of the object function, which results in local fluctuations of the $\vec{I}^{COM}$ distribution, it would generate non-zero contribution to the summation result. By replacing the infinite sum in equation 3.3 with a finite sum considering a kernel of $n \times n$ pixels, it results in:

$$
\begin{aligned}
&I^{riCOM}(\vec{r}_p) \\
&= \frac{a}{2\pi} \sum_{x=r_{p,x}-\frac{n-1}{2}}^{r_{p,x}+\frac{n-1}{2}} \sum_{y=r_{p,y}-\frac{n-1}{2}}^{r_{p,y}+\frac{n-1}{2}} \frac{\vec{r}_p - \vec{r}_{xy}}{|\vec{r}_p - \vec{r}_{xy}|^2} \cdot \vec{I}^{COM}(\vec{r}_{xy}) \\
&= k_n(\vec{r}_p) \star \vec{I}^{COM}(\vec{r}_p) \\
&k_n(\vec{r}_p) = \frac{a}{2\pi} \sum_{x=r_{p,x}-\frac{n-1}{2}}^{r_{p,x}+\frac{n-1}{2}} \sum_{y=r_{p,y}-\frac{n-1}{2}}^{r_{p,y}+\frac{n-1}{2}} \frac{\vec{r}_p - \vec{r}_{xy}}{|\vec{r}_p - \vec{r}_{xy}|^2}
\end{aligned}
\tag{3.4}
$$

Equation 3.4 shows the summation with a range n centred at probe position $\vec{r}_p = (r_{p,x}, r_{p,y})$. With this constraint on the range, the integrated centre of mass at one point can be found by only processing COM data from its limited surrounding (Fig. 3.1-a), allowing data processing to begin and results to be generated during the scanning session. This reconstruction method is thus given the name "real-time iCOM" or "riCOM", as indicated in the same equation by $I^{riCOM}(\vec{r}_p)$. This process is equivalent to a cross correlation between an array $k_n(\vec{r}_p)$ of size $n \times n$, that stores vectors $(\vec{r}_p - \vec{r}_{xy})/|\vec{r}_p - \vec{r}_{xy}|^2$, and the COM shift map $\vec{I}^{COM}(\vec{r}_{xy})$. This array will be referred to as the "kernel" throughout this manuscript, and images generated by processing COM shift maps with such kernels will be denoted as "riCOM results" or "riCOM images".

Since the kernel processes a group of data points and outputs a value corresponding to the probe position at the centre of the kernel, the collection of data has to lead the reconstruction by $(n-1)/2$ scan lines to fill up the kernel (when scanning in a traditional line by line fashion). This delay between the data collection progress and reconstruction result can be troublesome for operations that highly rely on real time feedback from the scanning process. Since the summation in equation 3.4 describes a linearly independent process, the contribution from multiple probe positions to a common pixel in the riCOM array can be separately calculated. Furthermore, by collecting the contribution from the COM shift at a specific probe position to its vicinity, an update to the riCOM image can be generated in the form of an array of the same size as the kernel. Since this reconstruction scheme depends on one CBED at a time, it leads to a live update of the riCOM result without any delay (Fig. 3.1-b). Although this does not reduce the time differences between the latest scanning point and the fully updated riCOM pixel, the partially reconstructed fraction of the riCOM image can already show atomic features[1], and therefore valuable information at the newly scanned probe position appears with minimal delay. This way the user can also get a quick feedback of their operation. Another advantage is that once the contribution from one probe position is calculated and the corresponding update to the riCOM array is made, the CBED pattern can be discarded, freeing up memory. This effectively removes any memory imposed restriction on scan size if the user is only interested in the resulting riCOM image.

[1] See supplementary data at [151] for example videos

3.2.2 Kernel Design

As mentioned in the previous section, summation carried out by a smaller kernel emphasizes local object function variations. In other words, it gives more weights to components of higher spatial frequency. To show the relationship between this effect and the kernel size, we start with the Fourier transform of the function $O(\vec{r}_p)$ for the case of a perfect COM shift measurement (second line of equation 3.3).

$$\begin{aligned}\mathcal{F}\left\{O(\vec{r}_p)\right\} &= \mathcal{F}\left\{\int_{-\infty}^{\vec{r}_p} \vec{I}^{COM}(\vec{r})\cdot d\vec{r}\right\} \\ &= \mathcal{F}\left\{\vec{I}^{COM}(\vec{r}_p)\right\}\cdot\frac{1}{i\vec{k}_p}.\end{aligned} \tag{3.5}$$

Here the symbol $\mathcal{F}$ indicates Fourier transform and $\vec{k}_p$ is a vector in the Fourier domain. As seen in equation 3.5, each of the Fourier components of the COM shift map is transferred to the final image with a weight $1/i\vec{k}_p$ after the path integration. This transfer function decays fast with the frequency, and thus low frequency components are attenuated much less than high frequency ones. By integrating over a finite range, an analytical expression for the riCOM result can be obtained as follows:

$$\begin{aligned}&\mathcal{F}\left\{I^{riCOM}(\vec{r}_p)\right\} \\ &= \mathcal{F}\left\{\int_{\vec{r}_p-\Delta\vec{r}}^{\vec{r}_p} \vec{I}^{COM}(\vec{r})\cdot d\vec{r} - \int_{\vec{r}_p}^{\vec{r}_p+\Delta\vec{r}} \vec{I}^{COM}(\vec{r})\cdot d\vec{r}\right\}/2 \\ &= \mathcal{F}\left\{\vec{I}^{COM}(\vec{r}_p)\right\}\cdot\frac{1}{i\vec{k}_p}\times[1-cos(\Delta\vec{r}\cdot\vec{k}_p)] \\ &= \mathcal{F}\left\{O(\vec{r}_p)\right\}\times[1-cos(\Delta\vec{r}\cdot\vec{k}_p)].\end{aligned} \tag{3.6}$$

In equation 3.6, the riCOM result is approximated by the contribution from both sides of the probe position $\vec{r}_p$ in a single line, within the range of $2\Delta\vec{r}$. The result shows that by limiting the integration range, it reproduces the function $O(\vec{r}_p)$ with an extra weighting function $1 - cos(\Delta\vec{r}\cdot\vec{k}_p)$. This function is close to zero when $\Delta\vec{r}\cdot\vec{k}_p$ is small, and thus strongly suppresses the low frequency signal in the retrieved object function. Also, it peaks at $\vec{k}_p = \frac{\pi}{\Delta\vec{r}}$, which implies that by choosing smaller $\Delta\vec{r}$, or shorter integration range, one can put more weight to the high frequency components. By using kernels with sizes smaller than the real space dimension of the dataset, this effect of limiting integration range can be achieved. Although the actual frequency spectrum of a 2D kernel deviates, the weight of a kernel of size n at each frequency k can be well approximated with the formula derived from line-integration:

$$\frac{N}{2k\pi} \times [1 - cos(\frac{n-1}{2} \times \frac{2k\pi}{N})]. \tag{3.7}$$

Here N is the number of pixels of the image in one direction, and the extra factor $2\pi/N$ scales with the pixel size in the Fourier transformed result.

This effect is not equal to but can be compared to a high-pass filter as it emphasizes high frequency details in the reconstructed image. However, other filtering effects, such as low-pass or band-pass filtering, cannot be created simply by altering the kernel size. A filter can be seen as a mask that reduces or eliminates a certain range of frequency signals of an image. This is done by a piece-wise multiplication between the filter and the image in the frequency domain, which is equivalent to a convolution between their real space counterparts. For riCOM images, which can be seen as a cross correlation between a COM shift map and a kernel, the application of such a filter can be included to the design of the kernel:

$$\begin{aligned} I^{riCOM}(\vec{r_p}) \circledast f(\vec{r_p}) &= [\vec{I}^{COM}(\vec{r_p}) \star k_n(\vec{r_p})] \circledast f(\vec{r_p}) \\ &= \vec{I}^{COM}(\vec{r_p}) \star [k_n(\vec{r_p}) \circledast f(\vec{r_p})] \end{aligned} \tag{3.8}$$

With

$$\begin{aligned} f(\vec{r_p}) &= \mathcal{F}^{-1}\left\{F(\vec{k_p})\right\} \\ F(\vec{k_p}) &= \begin{cases} 1, & k_{max} \geq |k_p| \geq k_{min} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{3.9}$$

In equation 3.8, $f(\vec{r_p})$ is the filter function and the symbol $\circledast$ indicates convolution. Equation 3.9 writes one of the possible ways to design such a filter, with a hard cutoff at two frequency limits k_{max} and k_{min}, i.e. a band-pass filter. The real space counterpart of the filter can be found by performing an inverse Fourier transform $\mathcal{F}^{-1}$ to the filter function in the Fourier domain. This real space filtering effect can be incorporated to the kernel due to the associative property of cross correlation and convolution. It is worth noting that the last part of equation 3.8 only holds for centrally symmetric filters that treat frequency components at different azimuthal angles equally, which is indeed the case for the filter shown in equation 3.9. We also want to point out that to create a sharp cutoff at the frequency domain, one would need a filter matching the size of the COM shift array. But in order to keep the size of the kernel, the outcome of the convolution is reduced in size. In other words, the outcome of $k(\vec{r_p}) \circledast f(\vec{r_p})$ is kept at the same size as $k(\vec{r_p})$. This would make the cutoff appear in the fashion of a slope and also distorts the rest of the frequency spectrum.

In Fig. 3.2 the frequency components of different kernel designs are illustrated. From bottom to top, the curves correspond to the template kernel with size of 101×101, a smaller kernel with size of 41×41, template kernel with high-pass filter ($k_{min} = \frac{12.19}{2}\ px^{-1}$), low-pass filter ($k_{max} = 12.19\ px^{-1}$), and band-pass filter (k_{min}, k_{max} same as before). For the bottom two curves, the result of the corresponding line integration approximation

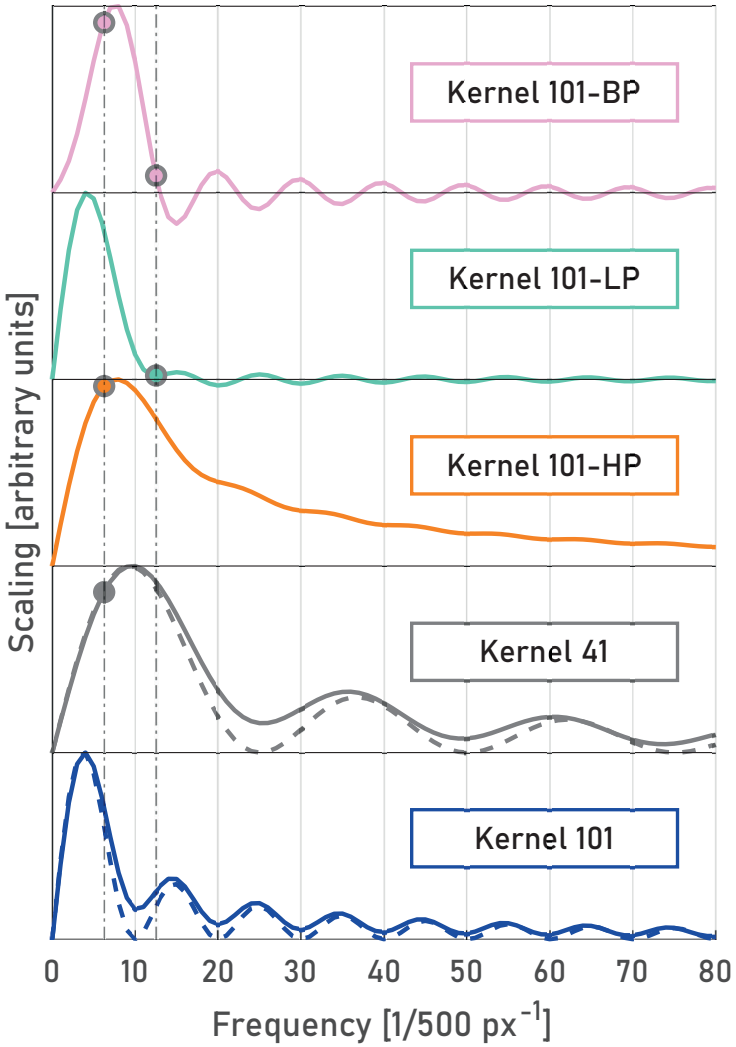


Figure 3.2 | Frequency components of a set of kernels acting on a COM shift map of size 500×500. The presented examples include, from bottom to top, the template kernel with size of 101×101, a smaller kernel with size of 41×41, template kernel with high-pass filter, low-pass filter, and band-pass filter. The dashed line shows the predicted transfer function with line-integration approximation. The two vertical lines indicate the cutoff frequency of the filter or the inverse of the kernel size, and the circles at the intersection of the vertical lines and integral indicate whether a cutoff frequency is applied to the specific design.

(dashed lines), with $\Delta\vec{r}$ chosen to be half of the kernel size, is also drawn to show their similarity in oscillation frequency and magnitude.

By comparing the blue and grey curves in Fig. 3.2, it is clear that Kernel 41 peaks at a higher frequency than Kernel 101, as predicted by the analytical formula, and that the cutoff of the lower frequency due to a smaller kernel happens approximately at the inverse of the kernel size, indicated by the grey circle. This value is then used for k_{min} of the high-pass filter incorporated to Kernel 101-HP (orange curve), which indeed

shows a similar overall frequency spectrum as the one of Kernel 41. Notice that the size reduction after convolution between kernel and the decorating filter causes a smooth decrease of frequency components below k_{min}, and spectrum differences beyond k_{min} compared to Kernel 101. Similarly, a kernel with low-pass filter *Kernel 101-LP* (green curve) and a kernel incorporating a band-pass filter *Kernel 101-BP* (pink curve) is created. Both of them are showing a suppression of the higher frequency ranges. Kernel 101-BP also shows a shift of the spectrum peak to a higher frequency because of its high-pass characteristic.

Despite the fact that it is not always possible to recreate the exact characteristics of common post processing filters, the incorporation of filters into the kernel, as well as the choice of kernel sizes allows for a great flexibility for frequency tuning and yields consistent and predictable solutions. Combining the kernel and the filter in real space also enables these image processing functions to be applied before the complete riCOM image is rendered, and thus compatible with the live update algorithm.

3.2.3 Data Processing

Due to the simplicity of the algorithm, the processing can be carried out completely by CPU with very limited usage of memory. However, in order to reach real time reconstruction that is limited only by the frame rate of the camera, an efficient implementation of the algorithm is crucial. The benchmark shown in figure 3.3 shows that an optimized implementation using C++ can easily achieve the maximum speed of ≈14 kHz of a MerlinEM camera. Additionally, the pre-processing of binary live data benefits from the low-level features of C++ (e.g. adapting endianness and efficient conversion of binary into numerical arrays). An implementation of the algorithm tailored to event-driven cameras and their corresponding sparse datasets is even significantly faster. Depending on the dose, >100 kHz have been obtained. The live visualizations at such rates also benefit from using C++ through the possibility of directly accessing and modifying OpenGL textures across threads.

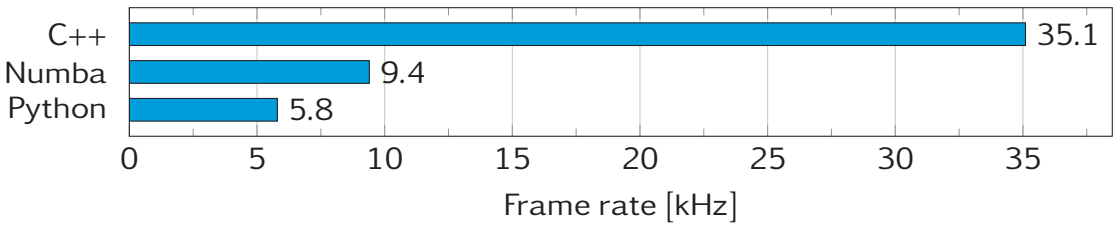


Figure 3.3 | Speed (Frame Rate in kHz) vs. Implementation benchmark for the computation of riCOM signal with Kernel size of 61x61, data type uint16 and camera size of 256x256 pixels, run on a single thread of an Intel i5-10210U @ 4.2GHz processor. Comparison of a simple implementation in Python, a just-in-time compiled optimisation of the same code, using Numba and a version written completely in C++ (compiled with GNU gcc-11).

The program was developed as a cross-platform application that can be run through a command-line-interface (CLI) or interactively through a graphical user interface (GUI) as shown in figure 3.4. The core functionality of the algorithm is implemented in a single C++ class object. Visual interfaces interact with an instance of that class across

threads through pointers, which allows live updates to be displayed immediately while maintaining a responsive interface without interrupting the reconstruction process. Furthermore, kernel settings for riCOM reconstruction and virtual STEM (vSTEM) settings, such as rotation angle due to Lorentz force, kernel size, filter, and the virtual detector size, can be changed during the process without interruption, which is helpful to find suitable settings interactively while spending the lowest amount of dose on the precious sample area.

The riCOM base class is independent of specific camera models and data types, while additional dedicated classes provide live- and file interfaces for given camera types/file formats. This allows for easy extendibility of the program by simply including further interface-classes. The current implementation includes a live- and data interface for the MerlinEM as an example for frame-based data and a filetype interface for the event-based Timepix3 camera and is available on GitHub under a GPL license.

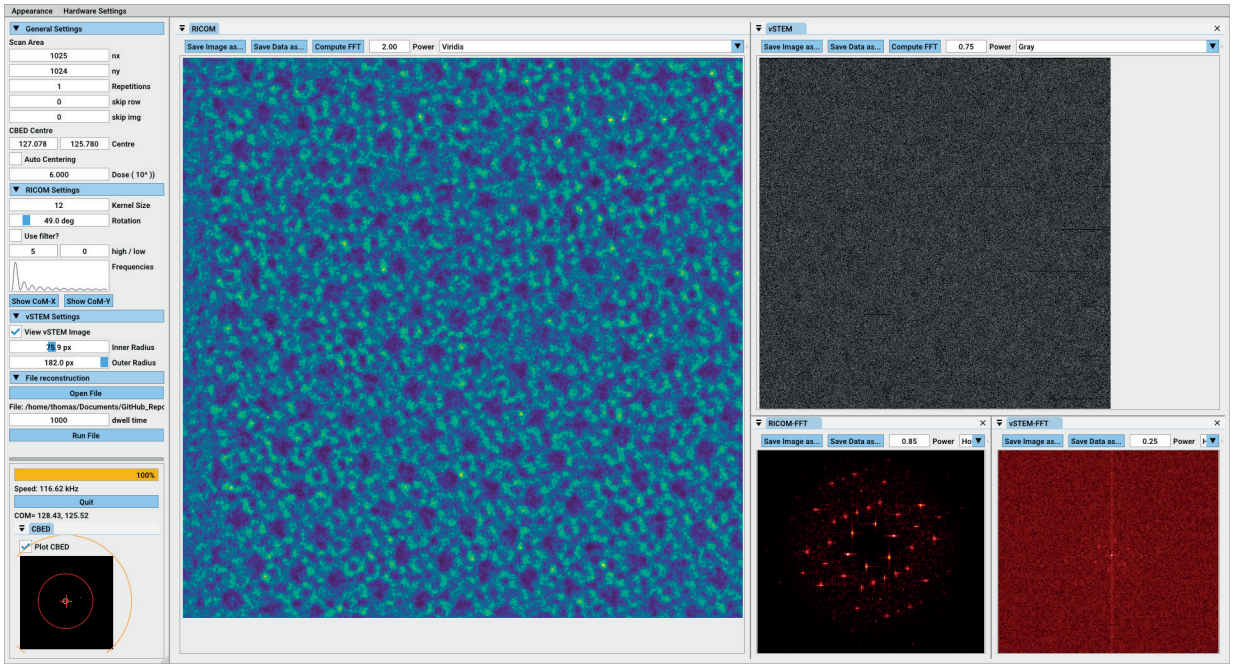


Figure 3.4 | Layout of the GUI. The Menu column on the left allows the user to change various settings, such as scan size, riCOM Kernel and filter settings, virtual STEM settings and interfaces for live mode and file dialogues. During a running reconstruction, a CBED is plotted at the bottom of this menu to visually assist interactive tuning of pattern centre and integration area for vSTEM. All other windows are floating panels and can be moved and resized.

3.3 Experimental Details

The results presented in this chapter are produced from data collected in two experiments. In the first experiment, a $SrTiO_3$ focused ion beam (FIB) lamella is examined with a probe corrected Thermo Fisher Titan^3 (X-Ant-TEM) operated in STEM mode. The resulting CBEDs are collected with a MerlinEM direct electron detector [42] and form 4D datasets for further analysis, as well as movies demonstrating the real-time processing

power of the method. The experiment is performed with a beam energy of 300 keV and a convergence angle of 20 mrad.

The second STEM experiment is performed on a silicalite-1 zeolite sample with a Thermo Fisher Themis Z (Advan-TEM). The data is collected with a custom made Timepix3 detector [44] based on an Advapix TP3 camera unit, and is recorded in event-based format. The beam energy and convergence angle used in the second experiment are 200 keV and 12 mrad, respectively.

All the datasets and movies recorded in both experiments, including necessary parameters for the reconstruction, can be found in the online repository [151].

3.4 Results and Discussion

3.4.1 Real Time Reconstruction

To demonstrate riCOM imaging, the software for real time reconstruction is run directly on incoming data during live experiments. The computer receives frames of CBEDs from the detector, and the software reads the data through a TCP socket. Throughout the process, the only extra prior knowledge to be provided to the algorithm is the COM of an undiffracted pattern in vacuum, so that the relative shift of COM at each probe position can be computed. Alternatively, it can also be approximated by averaging the COM from multiple probe positions, thereby omitting any calibration steps, making this method equivalent to more traditional imaging methods regarding ease of use. This step also inherently corrects for systematic shifts of the CBED away from the centre of the detector. While scanning, some of the most basic parameters of the microscopic imaging system are tuned, for example changing the defocus, astigmatism, and magnification, as shown in figure 3.5a-c.

Defocus broadens the intensity distribution of the electron probe, and astigmatism has the effect of creating two focal points, making the beam to be first focused in one direction and then the other when travelling down along the optical axis. This would reduce the electron beam sharpness and make the beam elliptical if out of focus, resulting in stretched atomic features in the images, as can be seen in figure 3.5-b in the region scanned before achieving right focus. According to equation 3.3, the intensity in the iCOM image equals the cross correlation between the projected electric potential of the material and the probe function, and therefore the reduction in contrast as well as distortions of the atomic features in the riCOM reconstruction is directly related to these beam aberrations. Hence, microscopists can tune optical conditions intuitively to maximize contrast and produce circular atoms with the live updated results. By changing the magnification during the scanning process the step size is changed accordingly. The live process can still continue, although the intensity needs to be adjusted since a is changed as the scan step size is changed, as shown in equation 3.4. Besides, the optimal kernel size changes with the magnification, as the spatial frequency of the desired features will be shifted when the step size is changed. However, since the kernel size can be adjusted during the process, a suitable choice can always be found by tuning the kernel size according to the quality of the live updated reconstruction image.

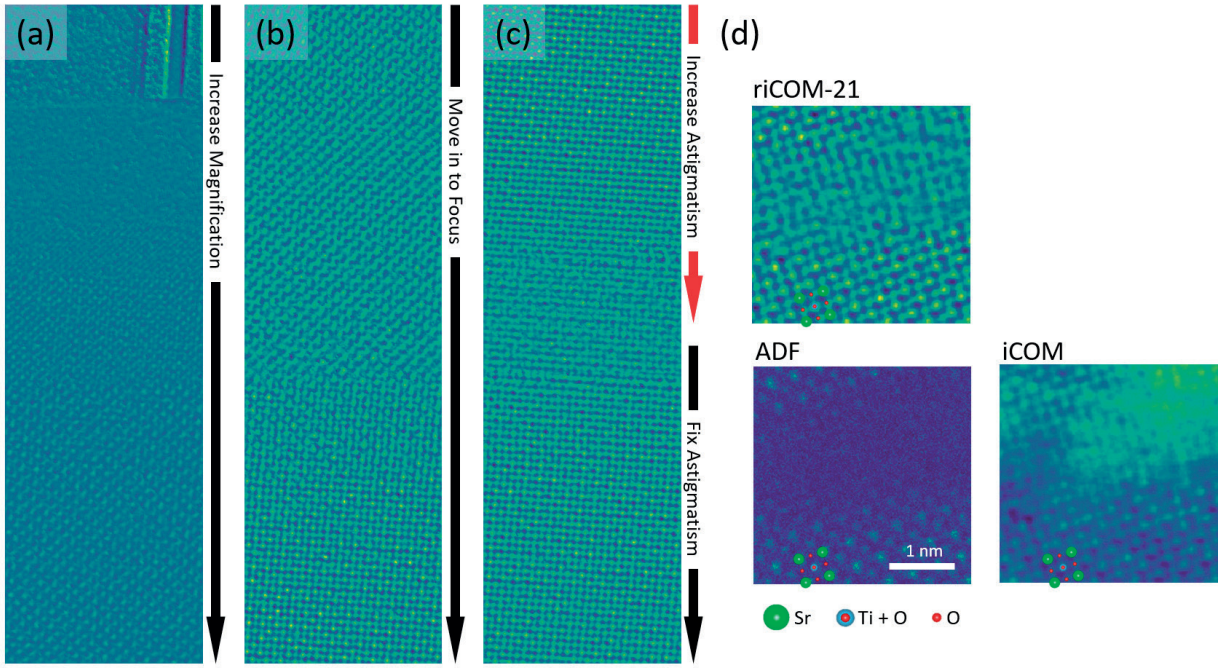


Figure 3.5 | Real-time reconstruction of $SrTiO_3$ while tuning the magnification, defocus, and stigmator. (a) The magnification is increased during the scanning. In the top most part the contrast reveals a layered structure of the FIB lamella, and with increasing magnification the atoms can be captured in the image. (b) Tuning focus is reflected by the change of the shape of atomic columns. (c) Tuning the stigmator affects the electron probe sharpness and also the contrast between atom and vacuum. (d) Simultaneous imaging using riCOM, ADF, and iCOM. RiCOM successfully images the crystalline structure in the centre of the image and the O columns, which is missing in the ADF image. The small kernel size used in riCOM reconstruction reduces long range intensity distribution shown in iCOM.

In figure 3.5-d, a riCOM image rendered with a kernel size of 21 is compared to the ADF image and the iCOM result. Apparent differences can be found in the centre of the images, which appears to have a hole according to ADF result but shows some crystalline structure in the riCOM and iCOM images, indicating possible extension of the crystalline material with lower thickness. ADF gives more significant contrast for differences in scattering ability, making it easier to distinguish Sr columns from Ti + O columns, but also reduces intensity of weak scatterers, such as thin regions and the pure O columns, to a level that is completely invisible, while riCOM and iCOM successfully image all three types of columns with a trade-off of less distinction between the columns. On the other hand, atomic structures are blurred by the long range intensity variation in the iCOM result. The origin of this variation could be local strain, misorientation, contamination, charge accumulation, etc., but it is very difficult to pinpoint the actual cause. RiCOM with an appropriate kernel size suppresses these low frequency signals and shows a clear image of atomic columns.

The examples shown in figure 3.5 show how riCOM images can be used to fine tune optical systems in a similar manner as using ADF. Moreover, the method is superior to ADF imaging in terms of required electron dose and provides contrast also for the

weak scatterers in the object, including thinner regions or atomic columns composed of lighter atoms. The high-pass characteristic of the suitable kernel size has shown to be helpful in highlighting features of higher spatial frequency and reduce low frequency components, but it also means that the contrast interpretation has to be evaluated carefully, especially for quantitative analysis, as they can be affected by multiple factors unknowingly.

3.4.2 Comparison of Reconstruction Methods

In this section, results from the riCOM reconstruction are compared with other reconstruction methods that have the potential to provide real time imaging. For 4D datasets, ADF images can be computed using a virtual detector which integrates all electrons in a specified region of the detector. The summing process is independent of the probe position, and does not require information beyond the scope of a single diffraction pattern, thus making virtual ADF reconstruction possible for real-time visualization of the dataset. To showcase the performance of riCOM reconstruction, it is compared to both ADF as a traditional imaging mode, and SSB, which is generally considered as a highly dose efficient and quantitative ptychography method. For riCOM reconstruction, three results generated using different kernels are put into comparison, including two kernel sizes and one kernel incorporating a band pass filter.

The dataset used for the comparison is a 4D dataset recorded from a silicalite-1 zeolite specimen. The dataset is recorded in a sparse array, in which the location where electrons hit the detector and the arrival time is recorded. This type of data format has several advantages over more commonly seen frame-by-frame types at suitable experiment conditions. For instance, in the case of low dose imaging, sparse arrays result in datasets many times smaller than full-frame arrays, since only the pixels of the detector that successfully capture an electron generate data, while other inactive pixels remain silent. For riCOM reconstruction, this format also shows its strength in terms of processing speed. Yet another important feature of this format is that the arrival time can be used to adjust the dose in the post reconstruction stage. Since the arrival time of each electron is recorded, the amount of dose put into the reconstruction algorithm can be post-adjusted by reducing the acceptance time from each probe position. For example, with a dataset recorded with a beam dwell time of 6000 ns, the dose for the post reconstruction can be reduced to 1/3 of the original dose if the acceptance time is set to be 2000 ns since any electrons that arrive to the detector after the acceptance time for each probe position will be discarded.

Accordingly, five data treatment algorithms/setups are used for the experimental data at 3 different dose levels. The results are presented in figure 3.6. Comparing the images generated by a virtual ADF detector with other reconstruction methods, it is obvious that even with the maximum dose, it is not enough to generate an interpretable ADF image. The vertical lines in the ADF image is a result of the camera being inactive for unknown reason, which is discussed in previous work [149]. This however makes almost unnoticeable difference to other reconstruction methods, since the value of each pixel in the reconstructed image not only depends on the corresponding probe position but also on its surroundings. For SSB reconstruction, it includes a process to integrate

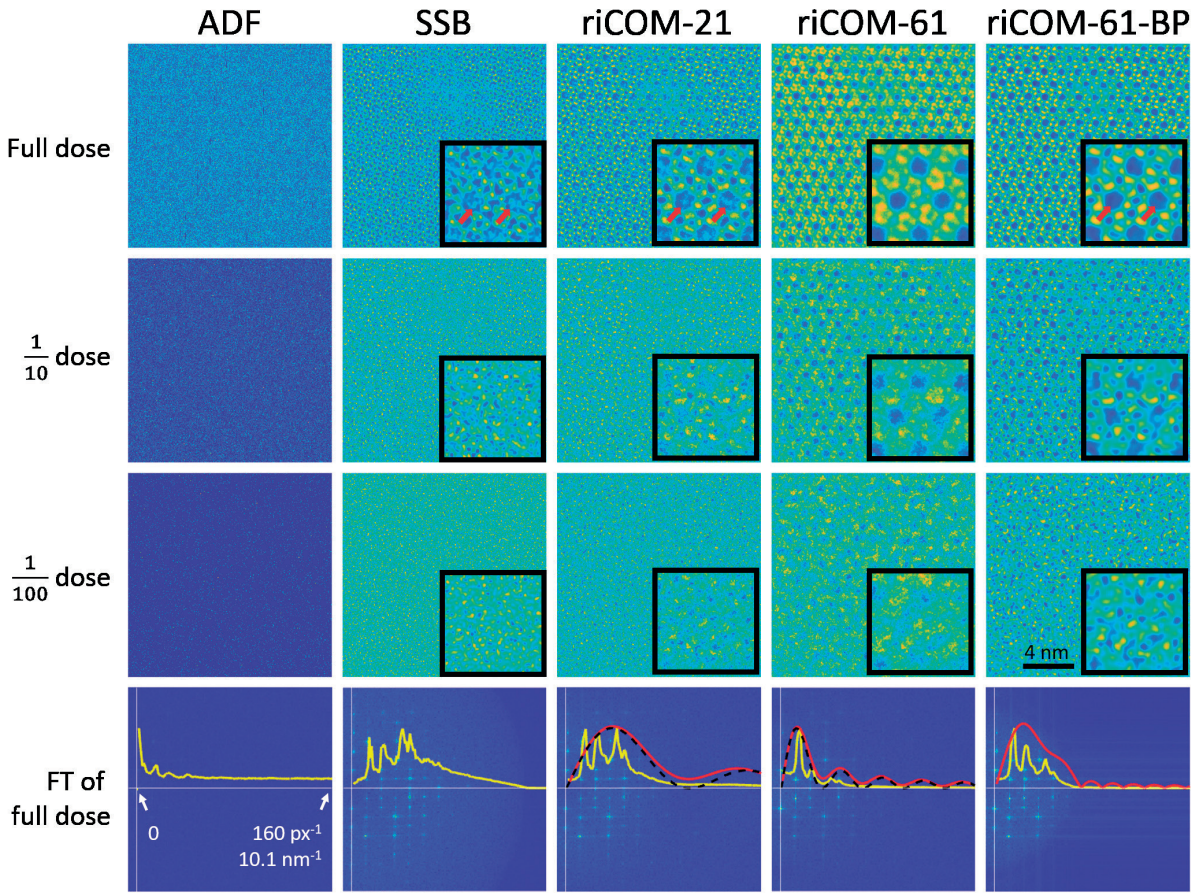


Figure 3.6 | Reconstructed images from an experimental zeolite dataset with different doses (Full dose: 1.27e+4 e/Å^2). ADF images are generated by integrating the intensities in the detector area beyond the convergence angle at each probe position. For SSB reconstruction, a frame-based dataset is first generated from the event array, with the detector space binned down to 32×32 (8 times smaller). For riCOM reconstruction, three different kernels are used: 21-by-21, 61-by-61, and 61-by-61 with a band-pass filter. The effect is however much less significant in other reconstruction method. The insets show magnified versions of the centre of their respective images, and the red arrows point out intensity fluctuations within the holes. The last row shows the Fourier transform of each reconstructed result. The radial averaged frequency spectra are represented with yellow curves, the frequency components of each kernel in red, and the line-integration approximation in black dashed curve.

specific regions in the CBEDs according to their spatial frequency by performing Fourier transformation with respect to probe position. Certain spatial frequencies are weighted more strongly from larger integration area, and thus creating a band-pass filtering effect [50, 142]. The riCOM images of smaller kernel size (riCOM-21) are shown to be similar to the SSB results, also manifested by the similarity of their frequency spectra, as low frequency signal is suppressed. For riCOM-61 result, by using a larger kernel size, more components at lower spatial frequencies can be found in the image. These components greatly increase the contrast for the long range structure in the material, such as the pores and framework of the zeolite crystal, but reduces high frequency

components, making the short range structures such as atomic columns less clear. This is especially highlighted in the result of 1/10 dose. However, by integrating a band pass filter to the big kernel (riCOM-61-BP), noise from the high frequency parts are removed and weights are redistributed to mid-range components from the low frequency end. It results in a much clearer image of the atomic structure even at 1/10 dose. The filter used for the last column is designed to remove signals from 3.8 nm^{-1} to 1.14 nm^{-1}, with $k_{max} = 60\ px^{-1}$ and $k_{min} = 18\ px^{-1}$.

In the third row, only 1/100 of the electrons in the dataset is used for imaging. The insufficient amount of electrons introduces a large amount of noise and hides the atomic structure in the images. Yet for the reconstruction result of riCOM-61, the pores within the zeolite framework are preserved in the image. This is possibly due to the fact that features of larger scale are reconstructed from more data points and is thus a result averaged over more possible integration paths. This kind of low frequency components are only supported by kernels of larger size, explaining why other reconstruction methods shown here do not benefit from them and fail to present any meaningful information in the images.

Imaging of zeolites at atomic resolution with iDPC, a similar method as iCOM, has been demonstrated to be successful at low dose between 100 and 1000 electrons per $Å^2$ [152, 153]. In a similar dose range, riCOM is capable of presenting structural features of the sample at different spatial frequencies, showing that the dose efficiency of the method is not sacrificed to enable real-time reconstruction.

While riCOM benefits from amplifying signals at specific frequencies so that clearer images of the lattice structures and atomic features can be captured, one has to bear in mind that the same effect is also applied to the statistical noise present in the experimental data. To study how noise affects the reconstructed images, one could compare results from ideal data with results from data with noise. However, for many reconstruction methods, it does not mean that the effect of noise can be simply acquired by subtracting one from the other since noise is not additive. Luckily, due to the linear independent nature, it is indeed the case for riCOM. In other words, the reconstructed image from a COM shift map with noise is exactly the same as the combination of the reconstructed image from a noise-free COM shift map and the one from pure noise. The latter is thus a suitable candidate for further noise analysis.

To demonstrate how noise is transferred to the reconstructed image at each frequency, a 4D dataset of a 20 nm thick zeolite sample is simulated according to the condition used in the second experiment (see Experimental Details). The noise is separated from the dataset to reconstruct an ADF image of pure noise, and the noise-induced COM deviation is calculated by subtracting the COM shift map from the noise-included dataset from the one without noise. The COM deviation map is then used for riCOM reconstruction with kernel size of 21 and 61. The components at different radial frequency of these images are plotted in figure 3.7. Two major differences between ADF and riCOM images can be found. First, the noise amplitudes of ADF images are higher when the dose is higher but the opposite for riCOM reconstruction is observed. It is due to the fact that the ADF intensity values follow a Poisson distribution, where the noise increases with the square root of the dose while the signal scales linearly with the dose. The COM shift on the

other hand is based on the spatial distribution of electrons, rather than the cumulative intensity and thus is not directly linked to this kind of shot noise. However, the error of the COM estimation still decreases when more electrons are used. Therefore, despite different noise behaviours, the signal to noise ratios of both methods increase with dose. The second difference lies in the distribution of noise at different frequencies. For the ADF noise image, the noise is distributed equally at different frequencies, yet for riCOM, the noise is amplified according to the approximated weighting function based on the kernel size (eq. 3.7). Through this analysis, it is clear that not only the signal from the examined object but also the noise is affected by the weighting in frequency domain. This greatly changes how noise appears in the reconstructed images compared to traditional imaging methods such as ADF and is worthy of the attention of the microscopists in order not to misinterpret features created by noise.

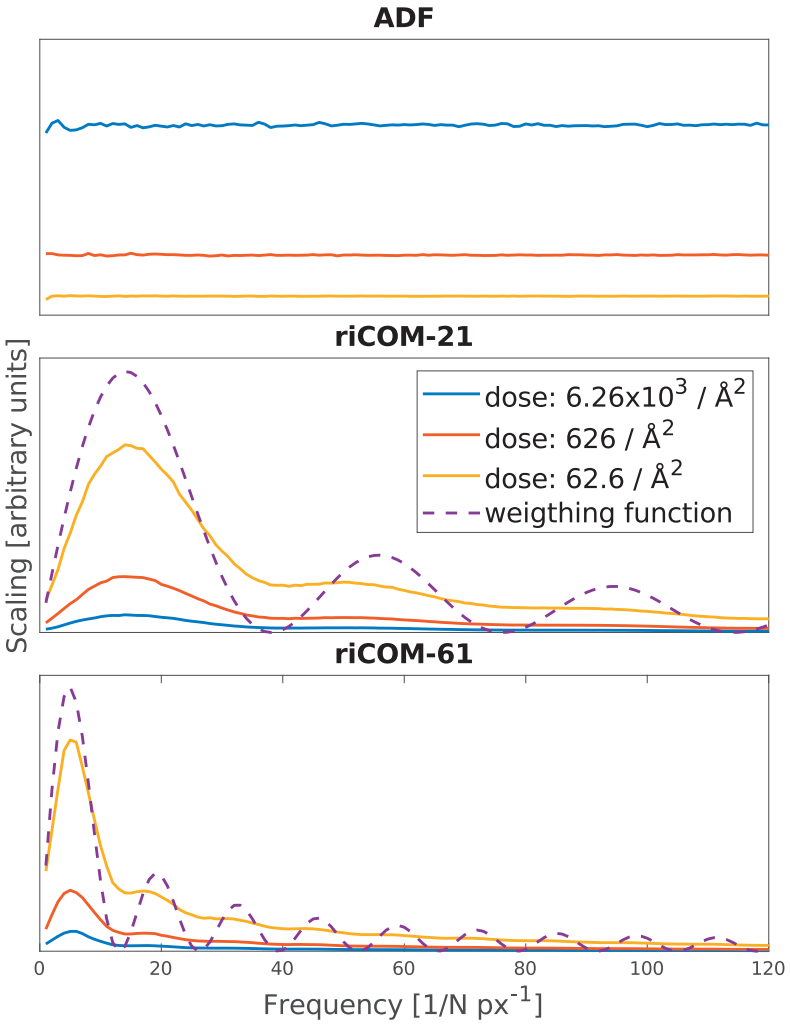


Figure 3.7 | Components of the noise images at different frequencies. The noise images are rendered by applying virtual ADF detector, riCOM with kernel size 21, and riCOM with kernel size 61. Three levels of dose for noise realization are chosen, and the curves are drawn by averaging 30 random noise configurations at each dose. The weighting functions given by the line-integration approximation are also presented for the riCOM results in dashed lines.

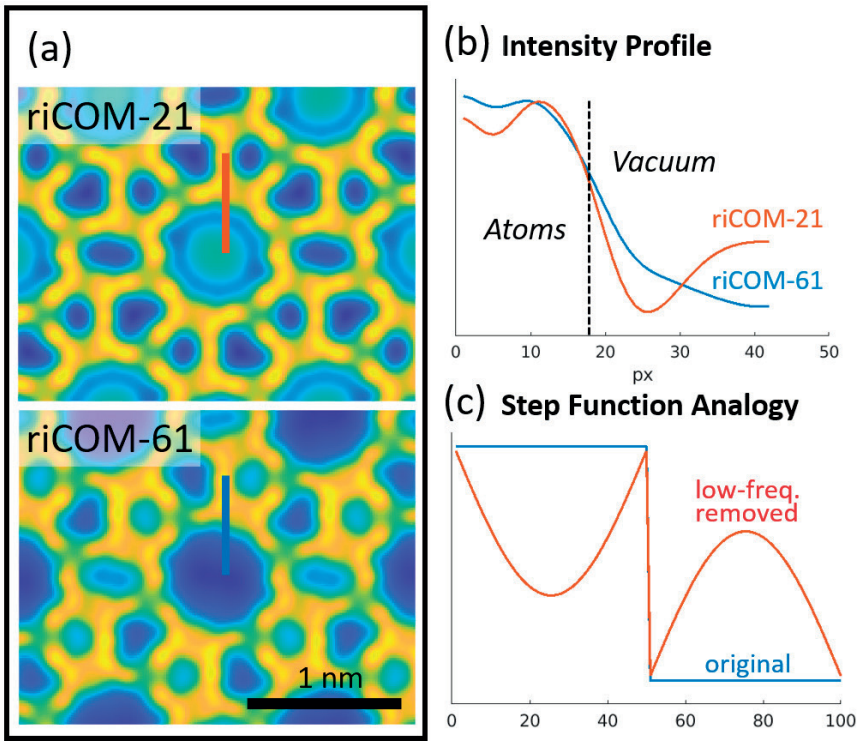


Figure 3.8 | (a) Reconstruction results of a simulated zeolite dataset with different kernel sizes. The red and blue lines indicate the locations of intensity line profiles drawn in subplots (b). (b) The intensity profile shows that the intensity inside the hole area increases in riCOM-21 results but decays in riCOM-61 towards the centre. (c) Step function for analogy shows that removing low frequency components may cause imaging artefacts similar to the ones seen in reconstruction results from smaller kernel sizes.

The different reconstruction results in figure 3.6 show a disagreement about the content inside of the pores that exist in the zeolite framework. Results from methods that give more weight to the high-frequency components, such as SSB and riCOM-21, show some intensity fluctuation inside of the pores, indicating the possible existence of dopants, yet these do not appear in the riCOM-61 image. In order to understand the cause of the difference, another simulation is run with the same condition to compare the reconstructed results with different kernel sizes in figure 3.8. To eliminate the possibility that this difference originates from the presence of noise, the reconstruction is done without adding noise to the dataset. From each reconstructed image an intensity profile is drawn over the atom framework into the pore (Fig. 3.8-a), which is indeed vacuum as designed for the simulation. The profile reveals that for riCOM-21, the intensity increases, while riCOM-61 shows a monotonic decay towards the centre of the pore (Fig. 3.8-b). The intensity increase for riCOM-21 cannot be explained by the projected atomic potential, since it can only decay when moving further away from the atoms.

To investigate the origin of this false intensity, the Fourier transformed riCOM images are analysed (Fig. 3.6). The bright spots at the lowest frequency correspond to the periodic structure of the pores and framework. The intensity of these spots are greatly reduced in riCOM-21 but supported in riCOM-61, indicated by the approximated weighting function as red curves. This causes major differences to features that necessarily rely on such

low frequency signals. To illustrate the principle, we simplify the atom framework and the pore using a step function (Fig. 3.8-c). By removing the low frequency components, the step becomes a curve with a concave and a convex segment in the regions of the high and the low step, respectively. This step function analogy conceptually captures the differences between the zeolite framework and the holes and explains the protruding intensity in the hole for riCOM-21 as the effect of reduced low-frequency components. For riCOM-61 such components are included by the larger kernel size, so that no such phantom intensity can be found in the same area.

These examples show that the proposed method, like many other reconstruction methods, is capable of providing extra information compared to traditional imaging methods. RiCOM also shows great dose efficiency, allowing high quality reconstruction results under low dose conditions. The freedom to use different kernel sizes grants users the ability to tune the desired spatial frequency range, which is very important in order to avoid misinterpretation of details in the image. Including more low-frequency components has shown to enable the reconstruction of long range structures of the object with even lower amounts of electrons. This could be very useful for microscope operators when imaging objects of larger scale.

3.5 Conclusion

In this chapter we propose and demonstrate a reconstruction method for real time STEM based on the integrated centre of mass that is applicable to any kind of segmented detector dataset, including but not limited to 4D-STEM. Through derivation of the physical formulation, we illustrate the physical relevance and the benefits for numerically efficient implementations of this approach, motivating the application particularly in real-time imaging scenarios. The freedom to change the size of the kernel or incorporating filters are also discussed, with examples showing their effect.

It is shown that riCOM can effectively reproduce iCOM results, but allows for more flexibility in terms of selecting contributing spatial frequencies. The method, including frequency band pass filtering depends only on the individual intensity distribution (or CBED) at its corresponding real space location, which in combination with a rather simple algorithm, creates a uniquely flexible and fast reconstruction method that requires very little user input. We further present a well optimized, interactive GUI implementation, developed in standard C++ and published open source on GitHub.

Demonstrations of the method on an operating microscope shows that firstly, the process is fast enough to keep up with the highest frame rate supported by currently available detectors, and secondly, providing dynamic feedback to the microscope operator when tuning and optimizing the microscope parameters. This ability enables swift search of the sample, or region of interest, as well as adjustments of the imaging conditions, at potentially very low dose conditions. The algorithm can run on any kind of data from which the centre of mass of the electron diffraction pattern, or derivatives of COM such as DPC signals, can be calculated, and therefore it is by no means limited to the hardware demonstrated in this chapter.

Comparisons with results of other non-iterative reconstruction methods show that riCOM renders high quality images on par with established methods, even at very low doses. The pros and cons of using different frequency components are discussed. Users can accordingly choose the most suitable designs of kernels, and run simultaneously other imaging forming methods, in order to reach the highest dose efficiency or extract the most amount of knowledge from the investigated sample in real time.

CHAPTER

4 AI-assisted rapid phase imaging (AIRPI)

Contents

This chapter was published as: "Thomas Friedrich, Chu-Ping Yu, Johan Verbeeck, Sandra Van Aert, Phase Object Reconstruction for 4D-STEM using Deep Learning, Microscopy and Microanalysis, 2023;, ozac002, https://doi.org/10.1093/micmic/ozac002"

4.1 Introduction

In the previous chapter one approach was presented that makes use of fast pixelated electron detectors [39, 41–43, 147, 149], which extends the set of modern imaging modalities for 4D-STEM (e.g. ptychography, the calculation of phase contrast [144] and true centre of mass (COM) imaging [143]). These imaging modalities can all be considered phase retrieval algorithms in a wider sense [154], as they all aim to retrieve the projected electrostatic potential of a sample, which directly affects the phase of the transmitted electron wave. While riCOM is attractive for its simplicity, ease of use and flexibility, its resolution is limited by the probe size. Ptychographic methods have been of particular interest for their super-resolution capabilities and the possibility to determine/correct for microscope aberrations as well [155–157]. However, the computational cost and memory requirements for these algorithms are considerable. Iterative algorithms like ePIE [137, 139] use optimization algorithms to fit an object to a given dataset such that the estimated phase corresponds to the intensity observations across the entire dataset. This is a computationally intensive task and the result is influenced by optimization parameters and convergence criteria. Other non-iterative methods typically use only the bright-field disc for phase reconstructions which imposes limitations on the maximum achievable resolution [157]. They rely on taking Fourier transforms with respect to the probe positions in real space, which means that for conventional Fast Fourier Transform algorithms (FFT), entire datasets (or at least substantial parts of them) have to fit into computer memory, which is becoming increasingly restrictive considering the growing size of 4D-datasets. For these reasons, ptychography has found many useful applications, mainly as a post-experiment data processing analysis step in specialized studies, but has not become a mainstream imaging modality so far. However, there is an increasing interest in using these algorithms interactively during experiments. To that end, live imaging using SSB was recently implemented and demonstrated by [148] and [146], as well as live centre of mass imaging by [158].

In the present work we explore the possibility to use machine learning (ML) for dose-efficient phase object (PO) reconstructions with super-resolution in (near) real-time. We show that using a convolutional neural network (CNN) enables fast exit wave retrieval for a given CBED, by using only a 3x3 kernel of adjacent diffraction patterns at a time. The method allows the retrieval of exit waves, with a resolution theoretically only limited by the Nyquist frequency of the detector and thus enables super-resolution imaging at sufficiently high doses. Using only nine CBEDs per probe position in a 4D-STEM dataset implies that the dataset can practically be of arbitrary size and the reconstruction can be performed live during the experiment with appropriate accelerator hardware, such as a modern, single GPU.

In this chapter, the character and capability of the proposed method is discussed in detail and demonstrated on both simulated and experimental data. Comparisons are also made with other possible live processing methods. To the best of our knowledge the only reconstruction methods that go beyond utilizing the traditional annular detector and enable live imaging are single sideband ptychography (SSB) [146, 148], integrated differential phase contrast (iDPC) and integrated centre of mass (iCOM) [158]. This is why we focus our analysis on comparing the results of our proposed method to those methods.

4.2 Materials & Methods

4.2.1 Theoretical Framework

The interaction of fast electrons with thin specimens can be conveniently described with the phase object approximation (POA). As an electron passes through a positive electrostatic potential its wavelength λ is temporarily altered which is equivalent to shifting the phase of the electron [14]. For cases where the specimen is extremely thin, the propagation of the wave as it goes through the material can be neglected as a reasonable approximation and the real space 3D electrostatic potential of the atomic structure $V_s(\vec{r},z)$ can be expressed as its integral along the optical axis z, resulting in the projected electrostatic potential $v_z(\vec{r}) = \int V_s(\vec{r},z)dz$, which is a function of the vector $\vec{r}$ that spans the remaining two dimensions. With this approximation the exit wave $\psi_{out}(\vec{r})$ is simply the product of the incident wave $\psi_{in}(\vec{r})$ and the object, which can be described by the transmission function $T(\vec{r})$:

$$\psi_{out}(\vec{r}) = \psi_{in}(\vec{r})T(\vec{r}) \tag{4.1}$$

where $T(\vec{r}) = \exp(i\sigma v_z(\vec{r}))$ and σ is an interaction parameter (See [14] for a more detailed derivation). However, a direct solution of the transmission function according to equation 4.1 is only possible if both incident and exit wave are known, while in practice neither of them are known a priori. The incident wave $\psi_{in}(\vec{r})$, can be fairly well approximated as the Fourier transform ($\mathcal{F}$) of the product of the aperture function $A(\vec{k})$ and an aberration-function($\chi(\vec{k})$)-dependent phase shift [159].

$$\psi_{in}(\vec{r}) = \mathcal{F}^{-1}\left[A(\vec{k})\exp[i\chi(\vec{k})]\right] \tag{4.2}$$

Here, $\vec{r}$ describes a 2D space at the object plane and $\vec{k}$ describes the reciprocal space. The function $\chi(\vec{k})$, considering only the spherical aberration C_s and defocus Δf, is given by:

$$\chi\left(\vec{k}\right) = \pi\lambda k^2\left(0.5C_s\lambda^2k^2 - \Delta f\right) \tag{4.3}$$

Assuming that at least the low order aberration parameters of $\chi(\vec{k})$ are known, equations 4.2 and 4.3 can be used to estimate $\psi_{in}(\vec{k})$. The other piece of missing information is then the exit wave $\psi_{out}(\vec{r})$. From a 4D STEM experiment, only the intensity $|\psi_{out}(\vec{k})|^2$ can be measured (figure 4.1-②), and thus a method to retrieve the exit waveform on the sample plane based on the information accessible from the experiment is required to solve equation 4.1 to get the transmission function $T(\vec{r})$. Retrieving the phase of $\psi_{out}(\vec{k})$ is a common inverse problem, but is severely complicated in 4D-STEM by the presence of noise to a level that makes even the estimation of $|\psi_{out}(\vec{k})|$ a challenging task in its own. The idea in this study is to leverage the multislice formalism, incorporating the calculation of electrostatic atomic potentials [113] and the frozen phonon approximation [65], as a forward model to generate a large synthetic dataset. This dataset can then be used to train a convolutional neural network (CNN) to retrieve an estimate of ψ_{out} for any given experiment within the boundaries of the validity of the used forward model

and within the parameter space of the training data, which will be discussed in section "Training data generation" and is given concisely in table 4.1.

4.2.2 General Workflow

The general concept of the proposed reconstruction method is schematically illustrated in figure 4.1. The workflow to retrieve the object $T(\vec{r})$ (figure 4.1-①) of a 4D-STEM dataset (figure 4.1-②) can be divided into two main steps: firstly, a neural network, trained to solve the inverse problem as outlined in section "Neural Network implementation", reconstructs the phase ϕ_{out} and amplitude $|\psi_{out}|$ of the exit wave, based on the intensity measurements of a 3x3 kernel of adjacent diffraction patterns, as depicted in figure 4.1-(③-⑤). Secondly, a patch of the object is retrieved from the previously reconstructed exit wave according to equation 4.1, as shown in figure 4.1-(⑥). In order to take the

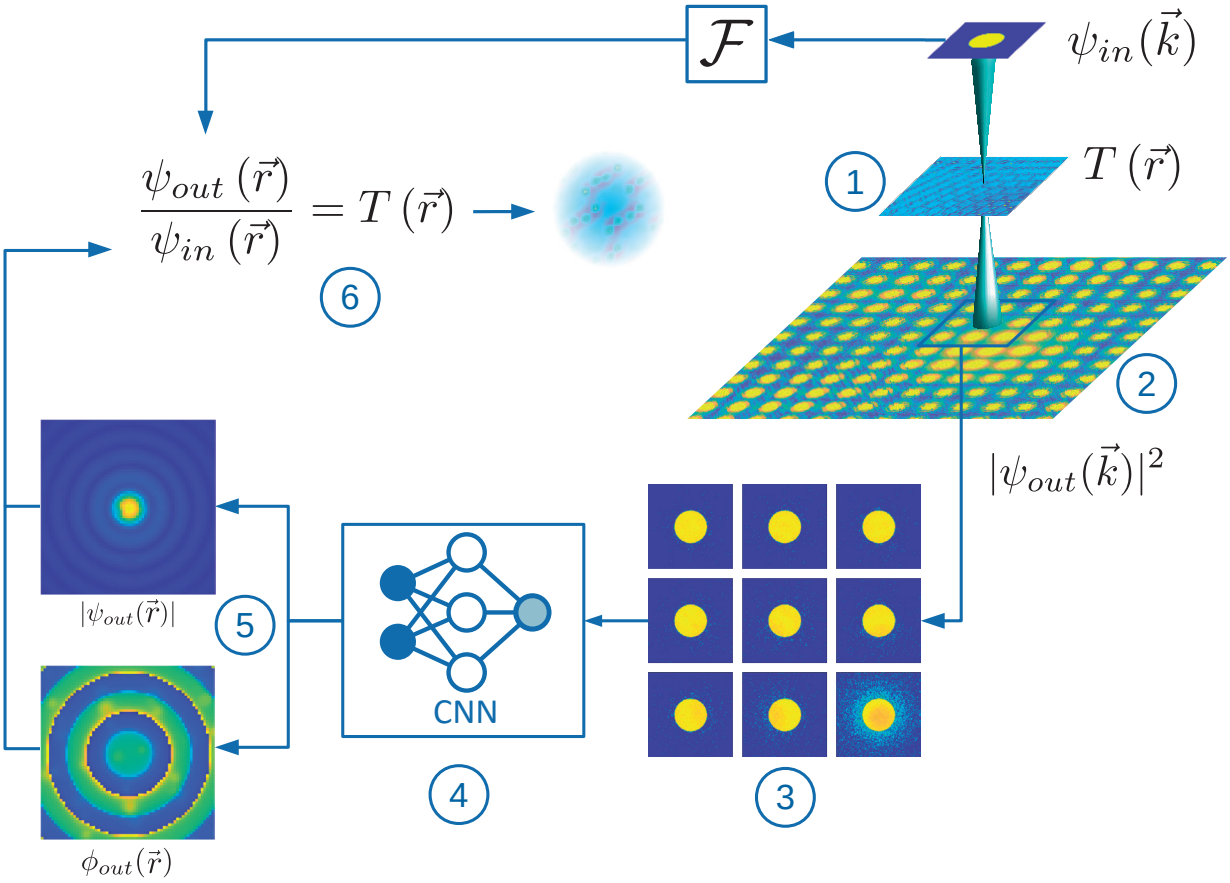


Figure 4.1 | General workflow: A patch of the phase object ① of a 4D-STEM dataset ② is reconstructed by extracting a 3x3 kernel of adjacent CBEDs ③, using a CNN ④ to reconstruct the amplitude ($|\psi_{out}(\vec{r})|$) and phase ($\phi_{out}(\vec{r})$) of the exit wave of the central CBED ⑤ and using the phase object approximation to reconstruct the object patch ⑥ from the reconstructed exit wave and an estimated probe function $\psi_{in}(\vec{r})$. Patches are then stitched together by complex addition to yield a reconstruction of the full phase object.

relative position of each individual object patch into account, a phase factor is included to the approximated $\psi_{in}(\vec{k})$ and the predicted $\psi_{out}(\vec{k})$ so that their real space counterpart $\psi_{in}(\vec{r})$ and $\psi_{out}(\vec{r})$ are found at the right position. This phase factor ϕ_{shift} is a function of

the displacement of the probe position in the x and y directions (Δx and Δy) with regard to the centre of an array of size $N_x \times N_y$.

$$\phi_{shift}(\Delta x, \Delta y) = \exp\left[2\pi i\left(\frac{\Delta x}{N_x} + \frac{\Delta y}{N_y}\right)\right] \tag{4.4}$$

The real space wave functions considering the phase factor are:

$$\begin{aligned}\psi_{in,n}(\vec{r}) &= \mathcal{F}\{\psi_{in}(\vec{k}) \times \phi_{shift,n}\} \\ \psi_{out,n}(\vec{r}) &= \mathcal{F}\{\psi_{out}(\vec{k}) \times \phi_{shift,n}\}\end{aligned} \tag{4.5}$$

In equation 4.5, $\mathcal{F}$ is Fourier transformation, and the amount of displacement Δx and Δy is absorbed into n, which specifies the n^{th} probe position.
The phase object approximation assumes that the retrieved object patch should have a homogeneous amplitude of 1 with a phase distribution reflecting the projected potential of the imaged material. However, since the transmitting electron probe carries information mostly from a specific region of the examined sample at the probe position, the retrieved object patches are given a weighting function ω_n according to the n^{th} incident probe intensity distribution, and the accordingly weighted object patch $\psi_{patch,n}$ is expressed as:

$$T_{patch,n}(\vec{r}) = \frac{\psi_{out,n}(\vec{r})/\psi_{in,n}(\vec{r})}{|\psi_{out,n}(\vec{r})/\psi_{in,n}(\vec{r})|} \times \omega_n(\vec{r}) \tag{4.6}$$

with the weighting function ω_n as:

$$\omega_n(\vec{r}) = \begin{cases} \frac{|\psi_{in,n}(\vec{r})|^2}{\sum_{\vec{r}}|\psi_{in,n}(\vec{r})|^2} & \text{if} \quad |\psi_{in,n}(\vec{r})|^2 > \frac{1}{10} max(|\psi_{in,n}(\vec{r})|^2) \\ 0 & \text{if} \quad otherwise \end{cases} \tag{4.7}$$

This procedure is repeated for all real space coordinates in the 4D-STEM dataset and the object patches are combined into the final phase object T by complex addition over n probe positions.

$$T(\vec{r}) = \sum_n T_{patch,n}(\vec{r}) \tag{4.8}$$

The object patch estimations coming from the CNN are not perfect but carry some errors. However, the full object is the combined result of the predictions made at multiple probe positions (equation 4.8). Even if one particular prediction is very inaccurate, its impact on the final result is limited as long as it is outweighed by the contributions from neighbouring probe positions, which is the case when a significant probe overlap is established by a dense scanning raster.

4.2.3 Training data generation

We created a large synthetic dataset, using atomic structures extracted from the materials project [160] database. Based on the unit cell definitions we created bulk specimens in one of the low-index zone axis orientations given in table 4.1. Each sample consists of a 3x3 kernel of simulated CBED patterns as features and the corresponding exit wave

in real-space as label. The simulations were performed using the multislice formalism and the frozen phonon approximation. In the implementation given by [60, 61], following the derivation of [65] the CBED intensity can be expressed as the sum of coherent and incoherent intensities.

$$\left\langle \left|\psi_{out}\left(\vec{k}\right)\right|^2 \right\rangle = \left|\left\langle \psi_{out}\left(\vec{k}\right)\right\rangle\right|^2 + \left\langle \left|\delta\left(\vec{k},t\right)\right|^2 \right\rangle, \tag{4.9}$$

where $\psi_{out}(\vec{k})$ is the exit wave, δ is a phonon configuration(t)-dependent difference and $\langle\rangle$ denotes the average over t. This formalism gives us access to the average, coherent, complex wave function. We use this wave function of the central CBED of the 3x3 kernel as labels (i.e. the ground truth training target) and the CBED intensities of all patterns in the kernel, as given in equation 4.9 as features (i.e. the CNN input). Only low order aberration parameters Δf and C_s of $\chi(\vec{k})$ (equation 4.3) are considered as they are unavoidable and typically have the strongest influence. We assume that including these effects phenomenologically with a constant, small C_s and corresponding Scherzer defocus is sufficient. Temporal and spatial incoherence are also not taken into account. This reduces the parameter space considerably but also implies that the method is (for now) limited to aberration corrected, well adjusted microscopes.

Further, the dataset includes the CBED-size in Å^{-1}, the aperture size and the acceleration voltage, which allows the computation of the probe function $\psi_{in}(\vec{k})$ within the data preprocessing pipeline during the training using equations 4.2 and 4.3. Also the effect of finite electron dose is applied as a data augmentation step during the training, assuming only Poisson noise. To accommodate the possibility that there may be no specimen interacting with the beam, another augmentation step replaces the CBEDs with the probe function intensity with a 3% chance. An example of the resulting training sample inputs

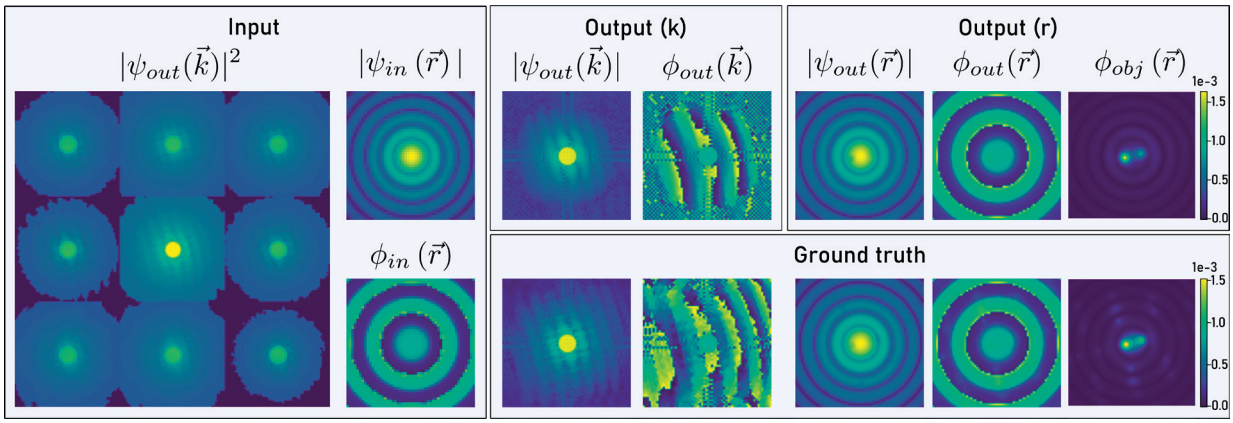


Figure 4.2 | Example of an exit wave reconstruction taken from the validation dataset, illustrating the inputs and outputs of the CNN, as well as the Fourier transforms of the (real space) exit waves. Intensities and amplitudes are depicted in log scale.

and labels is illustrated in figure 4.2 in the "input" and "ground truth" panels respectively. All parameters describing the dataset are summarized in table 4.1. Visualizations of the parameter distributions are shown in figure 4.3. The data generation code was published open source under https://github.com/ThFriedrich/ap_data_generation, as well as the training dataset used [161].

Description	Value
Acceleration voltage	$\in$ {30, 40, 50, 60, 80, 100, 120, 140, 160, 180, 200, 300}kV
Step size ($\vec{r}$)	(0.0167...≈2.0) Å
Convergence angle	5...30 mrad
Spherical aberration	0.001 mm
Defocus	Scherzer defocus
# Frozen phonons	30
Atom rmsd	0.08 Å
Orientation	$\in$ {[1 1 0], [0 1 1], [1 0 1], [0 0 1], [1 0 0], [0 1 0], [1 1 1]}
Thickness	<30 Å
Dose	3...3e9 e/CBED
# Structures	126,335
# Samples	742,688

Table 4.1 | Simulation parameters for the training dataset.

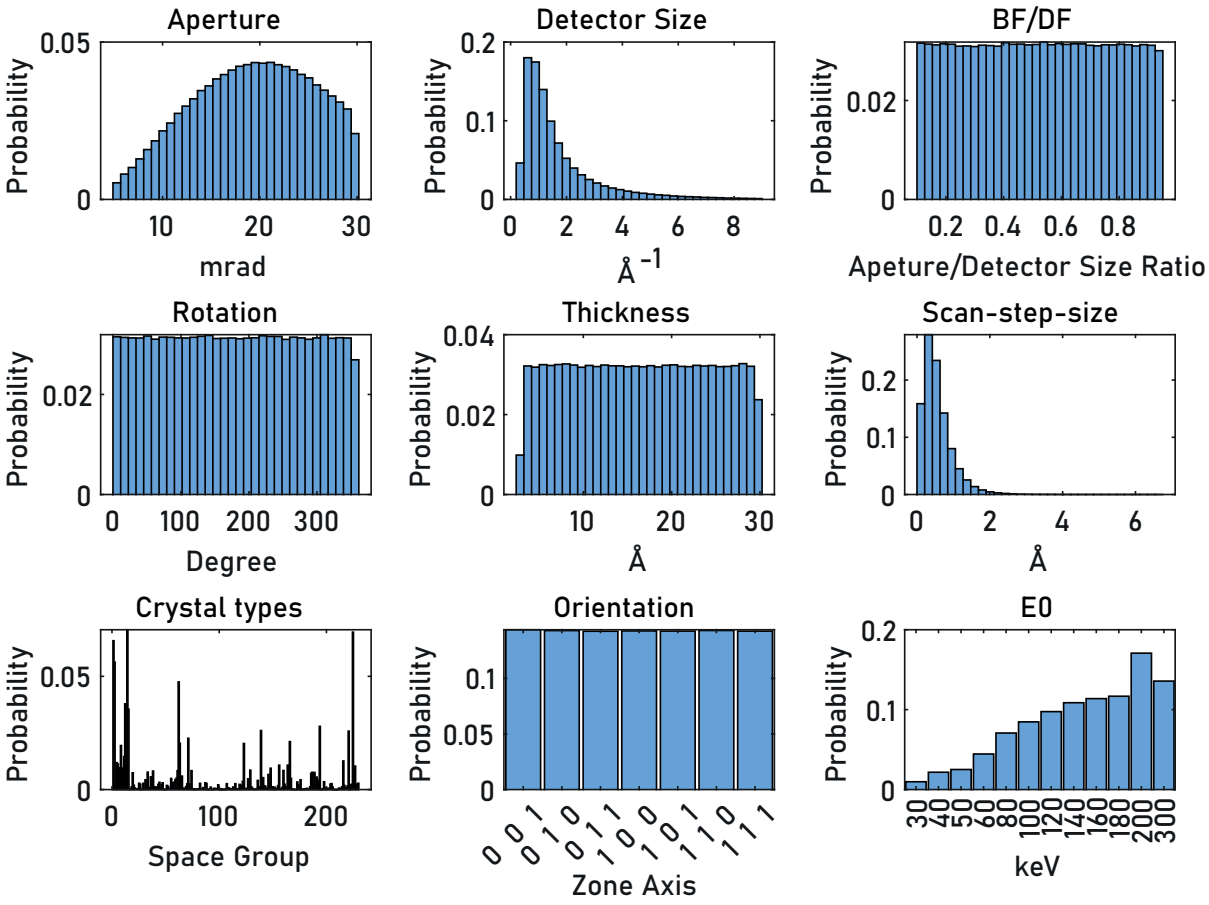


Figure 4.3 | Parameter distributions of most important simulation parameters of the training dataset, consisting of 742,688 samples. The "rotation" refers to sample rotation around the viewing direction, relative to the projected, crystallographic a-axis.

4.2.4 Neural Network implementation

CNN architecture

The complexity of the inverse problem practically dictates the use of deep neural networks in this study. The U-Net architecture [162] is one of the most popular choices for deep learning applications, since it allows to expand the number of parameters considerably, while maintaining strong backpropagation. Each depth level in the "U-shape" reduces the filter map size which accommodates the retrieval of features of different sizes. This makes the U-NET an extremely versatile and easy to train CNN and hence suitable as a starting point for this project as well [54]. However, since the training target is naturally complex valued, a generic pixel-to-pixel mapping, as commonly employed for CNNs in image processing, cannot be used. Two popular ways of dealing with this exist. Firstly, both, phase and amplitude retrieval problems can essentially be treated separately by defining two outputs and optimizing for dedicated loss functions on the phase and amplitude components of the wave function [163]. This treats the complex wave as two real valued images, which in practice has the advantage that common, highly optimized AI tools can be readily employed. Another approach that naturally lends itself to this kind of problem is using complex valued neural networks [164]; an approach that has found relatively few applications so far. However, the theoretical framework for complex CNNs is established [164] and implementations have been showing promising results [165, 166]. We implemented the U-NET architecture for both CNN types to test the main ideas. The complex networks delivered reconstruction results and accuracies of predicted phases, which are, for all intents and purposes, equivalent to the real-valued CNN, while decreasing the speed performance considerably. Since live imaging is an envisioned application, inference speed is a critical concern and the faster real-valued CNN was chosen in the study accordingly.

The structure of the U-NET is modified to account for physical considerations. The aim of the neural network is to model the electron-specimen interaction. Adding skip connections from the input probe function to the output exit waves (essentially enabling global residual learning) isolates the specimen interaction contributions from the probe function contributions to the exit wave. The CNN therefore does not need to learn to actually model the electron probe. The skip connections have the additional benefit of providing a common template during inference. The training is done on isolated scan points of only 9 CBEDs per specimen, while during inference the probe function should be consistent for the entire dataset, which is a requirement that cannot be captured by any metric during the training. Providing an estimated probe function greatly promotes this consistency during inference. Global residual learning also enhances noise robustness because the probe function serves as template, which is hardly altered if the input is merely noise.

On the input the dynamic range of the CBEDs are being scaled by taking them to the power of 0.1 in a preprocessing step, which puts more relative weight on the dark field scattered electrons to support exit wave reconstructions beyond the convergence angle. Subsequently each pattern is scaled, according to equation 4.10 depending on its distance from the central beam position, where ζ_{xy} and ζ_d correspond to the CBED

weights adjacent to the central CBED along x and y and on the diagonal, respectively. This is a straightforward way to include the step size Δs into the workflow.

$$\zeta_{xy} = \frac{1}{\Delta s * 50} \qquad \zeta_d = \frac{1}{\sqrt{2 * \Delta s^2} * 50} \tag{4.10}$$

The constant factor of 50 accounts for the range of step sizes in the training datasets such that all ζ are between 0 and 1. The effect of this weighting can be seen in the "Input" panel of figure 4.2, evident by the higher mean intensity of the central CBED.

Other notable differences as compared to the original U-NET implementation [162] are different map sizes, the use of Swish-Activation functions [70] and the use of strided convolutional layers instead of max-pooling layers to avoid any loss of information when feature map sizes are reduced. Trainable scaling factors on the two output layers for the phase and amplitude were added to scale between the batch-normalized feature maps (with standard deviations of 1 and means of zero) in the CNN and the relatively small target value distributions, which correspond to the difference between exit waves and probe functions. The variables were initialized accordingly with small values of 1e-4 and optimized during training. The amplitude output layer includes a regularization which penalizes integrated exit wave intensities > 1. The phase output layer penalizes values larger than π and smaller than $-\pi$. These penalties are added to the loss during training. The resulting overall architecture is depicted in figure 4.4. The tensorflow implementations of the models and individual layers are available open source on Github at `https://github.com/ThFriedrich/airpi`. During the training the CNN can process >750,000 sample points in about 6 minutes on a single Nvidia RTX 3080 GPU, indicating that the model itself may be well suited for live processing at rates >2kHz, if the pre- and post-processing pipelines are well optimized.

Loss function

To facilitate the learning of a general representation for both the phase and amplitude we designed a multi-objective loss function as a linear combination of $\mathcal{L}_2$-losses on the phase and the amplitude in Fourier space and in real space. Enforcing the correspondence between $\vec{r}$ and $\vec{k}$-space during the training encourages the CNN to abide to physical constraints. It was also observed that the decomposition of the phase into its sin and cos components facilitates better convergence, compared to just optimizing for the phase directly. This is presumably related to the decompositions being smooth so the CNN does not have to account for phase wrapping effects. Since the probe function is an input to the CNN the object can be solved directly and used in the loss function too. The optimization of the $\mathcal{L}_2$ error of the phase of the object $T(\vec{r})$ directly promotes an agreement of the object with the transmission function, which is practically the most meaningful metric. However, a good quantitative match may be impossible to achieve in certain scenarios (e.g. very low dose). To promote at least a visual match, the object phase is also optimized for its cross correlation xc as given in equation 4.14. Further it was empirically determined that a higher weight on the phase in $(\vec{k})$-space leads to

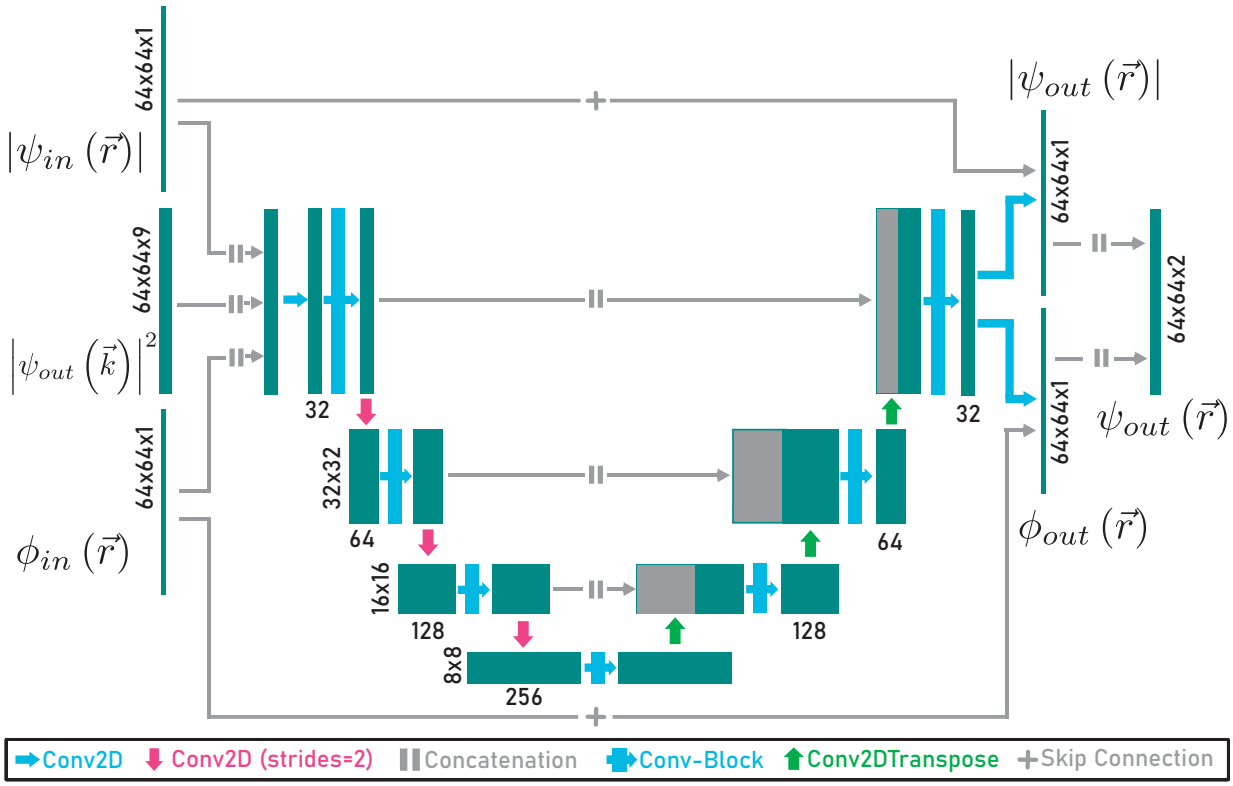


Figure 4.4 | The CNN architecture used in this study is a modified U-NET with separate, real valued phase and amplitude outputs. The model leverages global residual learning through added skip connections of the probe function to the output. Each "convolution layer" is composed of a 2D convolution layer, batch normalization and a Swish [70] activation function. Each "convolution block" consists of 3 consecutive convolution layers.

faster convergence and overall better results. Putting all terms into a sum, for an exit wave ψ with phase ϕ and an object phase ϕ_{obj} the loss function is given by:

$$\begin{aligned}\mathcal{L} = \\ &\mathcal{L}_2(|\psi(\vec{k})|^2) + (\mathcal{L}_2(\sin\phi(\vec{k})) + \mathcal{L}_2(\cos\phi(\vec{k}))) * 3 \\ &+\mathcal{L}_2(|\psi(\vec{r})|^2) + \mathcal{L}_2(\sin\phi(\vec{r})) + \mathcal{L}_2(\cos\phi(\vec{r})) \\ &+\mathcal{L}_2(\phi_{obj}(\vec{r})) + \mathcal{L}_{xc}(\phi_{obj}(\vec{r}))\end{aligned} \tag{4.11}$$

with:

$$\mathcal{L}_2(x) = \left(x_{true} - x_{predicted}\right)^2 \tag{4.12}$$

$$\mathcal{L}_{xc}(x) = \left(1 - xc\left(x_{true}, x_{predicted}\right)\right)/2, \tag{4.13}$$

where:

$$xc(x,y) = \frac{\sum_{i,j}[(y(i,j)-\overline{y})(x(i,j)-\overline{x})]}{\sqrt{\sum_{i,j}[y(i,j)-\overline{y}]^2}\sqrt{\sum_{i,j}[x(i,j)-\overline{x}]^2}}, \tag{4.14}$$

where x and y correspond to pixel values at the locations (i, j) and $\overline{x}$ and $\overline{y}$ are the mean values respectively.

Training

The training was performed using the Adam optimizer with a learning rate of 5e-4, a batch size of 256 and a momentum setting of 0.9. The learning rate was decreased by a factor of 0.5 when the validation loss did not decrease for 3 epochs. Convergence was reached after ≈50 epochs. After convergence the training was resumed for another 10 epochs with a weighting factor of 10 applied to the $\mathcal{L}_2(\phi_{obj}(\vec{r}))$ term of equation 4.11, which leads to a further small decrease on the object error. This step does not alter the reconstruction results considerably but improves the quantitative match between reconstructed objects and transmission functions somewhat.

4.2.5 Experiments & Simulations

The demonstrations of the reconstruction methods are performed on both experimental and simulated datasets. For the experiments, probe corrected Thermo Fisher Titan (X-Ant-TEM) and Themis (Advan-TEM) were used. The former is equipped with a MerlinEM direct electron detector [42] and the latter with a custom-made Timepix3 detector [44]. For the experimental datasets of an Au crystal and a $SrTiO_3$ focused ion beam lamella, which can be found in the the online repository [158], the acceleration voltage is set at 300 kV, the semi convergence angle of the electron beam is 20 mrad, and the scanning step size 0.2 Å and 0.185 Å, respectively. The USY zeolite dataset, which can be found in [167], is collected at 200 keV, with 12 mrad convergence angle and 0.15 Å step size. The simulated twisted bilayer graphene dataset is generated with an acceleration voltage 200 kV, a convergence angle of 25 mrad, and a scan step size 0.2 Å. The twisted bilayer MoS_2 dataset was simulated with the settings: acceleration voltage 300 kV, convergence angle 20 mrad, and scan step size 0.1 Å. The MgO dataset was created with an acceleration voltage of 300 kV, a convergence angle of 20 mrad, and a scan step size 0.05 Å. All of the simulated datasets are generated with the MULTEM software [61].

4.3 Results & Discussion

4.3.1 Super-Resolution

The reconstruction of the proposed method is based on solving equation 4.1 for the object using the incident and the exit wave functions, and therefore the resolution of the method is not explicitly limited by neither the optical conditions of the imaging system nor the sampling density of the electron probe. By reciprocity, the object plane is sampled with a maximum resolution determined by the maximum scattering angle the detector can reach, or the highest angle at which the exit wave can be accurately retrieved, which can potentially result in a higher resolution than permitted by the former two limitations. This super-resolution granted by the knowledge of the exit wave at higher scattering angles is demonstrated by the reconstruction of a simulated dataset of a twisted bilayer

graphene sample at infinite dose [54]. The result from the CNN reconstruction is shown in figure 4.5-a, and compared with a SSB reconstruction in figure 4.5-b.

To analyse the spatial frequency achieved by each method, the Fourier transformed (FT) images are presented as well (figure 4.5-d, e). The circle in the FT SSB image indicates twice the range of the convergence angle (α), which is the upper limit of the spatial frequency of this reconstruction method [157], and therefore all the frequency components beyond are eliminated. The reconstruction of the CNN successfully retrieves components beyond this limitation, also reflected in the ability to distinguish atoms with very short spacing in between, as can be seen in the atom pairs profile (figure 4.5-c)

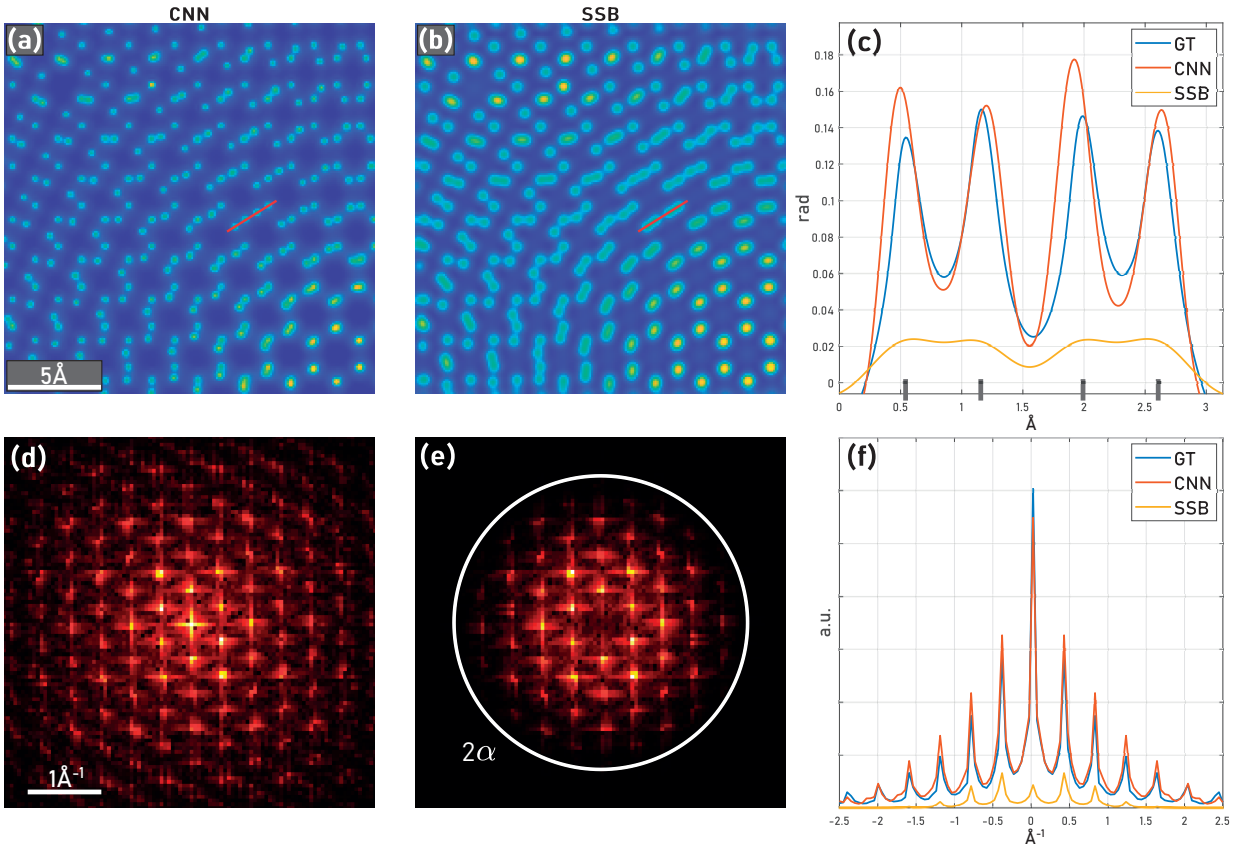


Figure 4.5 | Demonstration of super resolution capabilities on simulated datasets with infinite dose. Compared are CNN reconstructions and standard SSB ptychography (b). Their corresponding Fourier transformed (FT) intensities show the maximal spatial frequency achieved by each method (d, e). For the FT image of the SSB result, a circle indicating twice the convergence angle is added, which corresponds to the maximal spatial frequency of the method. (c) shows the intensity along the line profile drawn in each image. Markers on the x-axis indicate atom positions. (f) depicts the integrated intensity of the FT images along y-axis.

The improved resolution capability of the method, as well as its dependency on the individual per-CBED-dose (more thoroughly discussed in section "Step Size") are also demonstrated on an experimental USY-zeolite dataset [167]. This dataset has a fairly low dose and small step size. To increase the dose of each individual diffraction pattern, the CBED at each probe position is replaced by a summation of CBEDs within a 5×5 box, while the reconstruction is done with a step size twice as large as originally. This

repetition in data usage increases the effective dose in the dataset, as each individual CBED now contains 25/4-times more electrons and greatly increases the accuracy of the neural network prediction. The actual dose that inflicts damage while interacting with the material, on the other hand, remains the same. For comparison, SSB is performed on the original dataset and a dataset with the same data repetition strategy applied. In figure 4.6, it is shown that the last three reconstructions successfully build a clear image of the zeolite crystal structure with atomic level resolution. The CNN reconstruction based on the original dataset (figure 4.6-a) does not showcase a similar quality, since the dose for individual CBEDs is too low to make accurate predictions of the exit waves, but after data repetition is applied, the neural network gives results that capture details of the material (figure 4.6-b). SSB, on the other hand, does neither benefit nor suffer from this repetition, at least not at a noticeable level. From the Fourier transform of the three images one can estimate the resolution limits of the methods by comparing the most distant frequency component. The neural network reconstruction shows a maximum frequency component at 0.78 $Å^{-1}$, which according to the Rayleigh criterion:

$$d = \frac{0.61\lambda}{\sin\alpha} \tag{4.15}$$

is equivalent to the resolving power of an un-aberrated perfect optical system of convergence angle 12 mrad, which is the same as the aperture size used in the experiment, at the electron wavelength (λ) of 0.02 Å. As most of the microscopes, even ones equipped with probe corrector, cannot achieve the resolving power given by the Rayleigh criterion, the presented method shows the ability to overcome the effect of remaining aberrations, shot noise, and other imperfection in the system to reach a higher resolution.

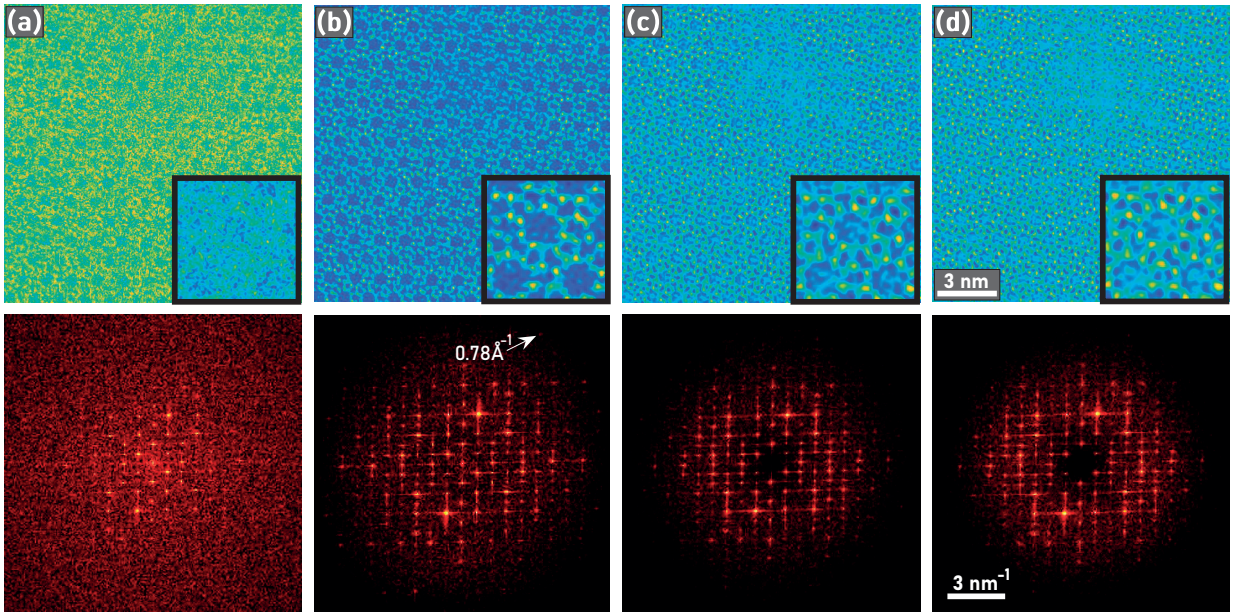


Figure 4.6 | Reconstruction results of three different approaches. (a, b) Neural Network performed on datasets without and with data repetition, respectively. (c, d) SSB reconstruction done on datasets without and with data repetition, respectively. The Fourier transforms of the reconstructed images are shown below. Notice that in (a) vertical streaks can be found, which originate from an unknown defect of the detector, also reported in [149].

4.3.2 Step Size

Since the proposed reconstruction method is based on retrieving individual object patches, which are commonly sampled finer than the step the electron beam takes to scan the sample, a rather coarse scanning grid can produce high quality images, as long as a good prediction of the exit wave, and hence the object patch, can be made. This character of the proposed reconstruction method is demonstrated on a simulated MgO particle with different step sizes of 0.1, 0.4, 0.8 and 1.6 Å. The "ratio" values shown in figure 4.7 refer to the ratio of the diameter of the incident probe function (1.2 Å) and the step size, where the probe size for a given convergence angle α in $Å^{-1}$ is defined by the first root of the Bessel function of the first kind and first order:

$$d = 2 * \frac{3.8317}{\alpha\pi} \tag{4.16}$$

Figures in the left column are generated with infinite dose and therefore the neural network has very detailed knowledge of the amplitude of the exit wave to make accurate predictions. In this case the difference between results of overlap ratio of 12 and overlap ratio of 3 is barely noticeable. By further reducing the scan density, probe positions reach a distance where the weighting function forbids any overlap, as shown in equation 4.7. Despite the weighting function cutoff, which creates blank spaces between the object patches, the actual probe positions used to generate the data overlap with each other just enough, making exit wave predictions possible to maintain the crystal structure to a certain level in the reconstructed image. As the step size reaches 1.6 Å and the ratio drops below 1, the retrieved object patches deteriorate severely and no longer reflect any crystal periodicity. This failure shows that the neural network follows certain physical and mathematical constraints, such as necessary probe overlap for accurate exit wave retrieval, and that it would fail rather then making false predictions that continue to resemble atoms or the crystal. This failure can be identified by the user not only based on the deviation of the resulting image from the expected appearance of the object, but also by the wide blanks left between the object patches, indicating insufficient probe overlap.

The images in the right column of figure 4.7 were generated with the same dose per area. As mentioned in the previous section, the accuracy of the retrieved object patch is not directly related to the total dose in the dataset, but rather to the dose per CBED. By this consideration it follows that larger step sizes work better for the neural network, since this would mean fewer probe positions in the same area and a higher dose at every individual CBED. On the other hand, a certain level of probe overlap is also required for accurate predictions. Therefore, not only the total dose per area, but also the scanning strategy is an important consideration for the proposed method. A balanced scan density will generate better results as compared to a very fine scan grid, even if the total dose per area would be the same. This behaviour is illustrated in the images in the right column of figure 4.7. The noise level is lowered significantly as the step/probe-width ratio drops from 12 to 3 in figures 4.7-h and 4.7-f. Ptychographic methods in comparison offer more flexibility in this regard as shown in figures 4.6-c and 4.6-d.

The noise created by inaccurate predictions also creates different features as the step size changes. As the training is exclusively done on crystalline materials in zone axis orientations, the predicted object patches may show atomic scale features, even if the input is merely noise. In other words, the frequency transfer function of each object patch is highest at the spatial frequency that would compose an image of an atom. This somewhat dangerous behaviour of the neural network is compensated by the stitching of the patches, since atomic-scale features of the noise would not remain sharp as multiple object patches contribute to the same area, and would thus contribute to a cloudy, low intensity background, as seen in figure 4.7-h. However, when the degree of overlap is reduced, such that the phase value is completely determined by 1 or few object patches, the risk to observe a false, atom-like feature, such as the ones seen in figures 4.7-f and 4.7-d, greatly increases. Awareness of this effect is therefore important when using the method and large step-size scanning patterns should be treated with extra caution. The other effect of noise is that the phase value drops in cases of very few electrons per CBED. This is essentially related to the failure of the CNN at making accurate phase object predictions, but it also shows that as long as sufficient dose is present in the 3×3 CBED input set, the phase value retrieved is not strongly related to step size.

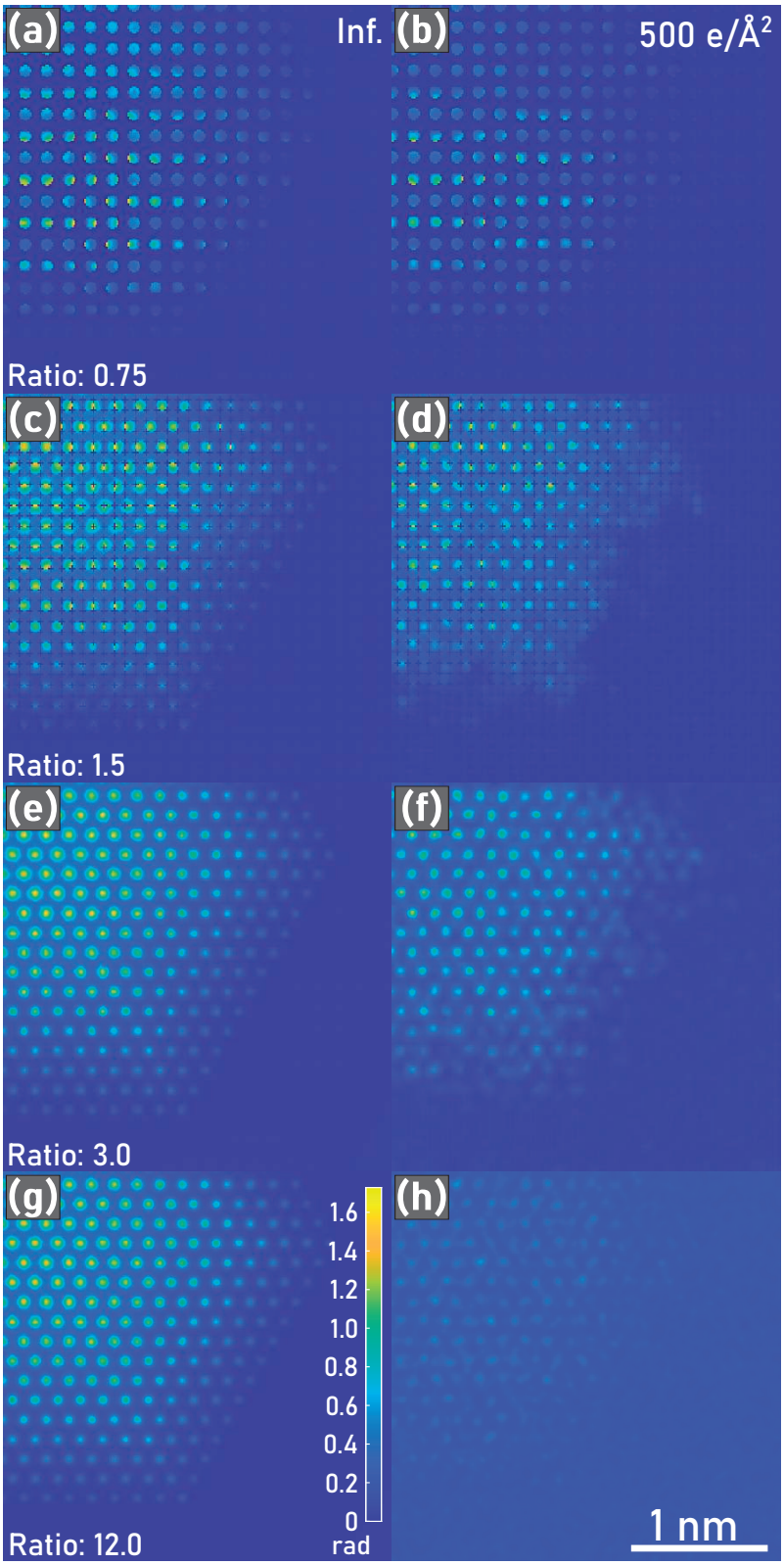


Figure 4.7 | Reconstruction results of simulated MgO particle. In the left column (a, c, e, g) the images are generated with infinite dose, and in the right (b, d, f, h) the dose is set to be 500 electron per $Å^2$. Each row of images is constructed with the same step size, as well as the same step/probe-width ratio. The colorbar in the bottom left panel applies to all images in the figure.

4.3.3 Contrast Analysis

As outlined in section "Theoretical Framework", the phase of the object is proportional to the electrostatic potential. By this relation atomic species should, at least within the boundaries of the phase object approximation, be distinguishable. To verify whether this requirement holds true for the CNN reconstructions, 4D STEM datasets of isolated, single atoms for each species in the periodic table up to Z=103 are simulated individually with a step size of 0.2 Å, a simulation box size of 3×3 Å, an aperture angle of 20 mrad and a collection angle 60 mrad. The frozen phonon approximation was used with a root-mean-square-displacement of 0.08 Å and 100 phonon configurations for each simulation. The retrieved phase objects of these datasets are compared to the ground truth transmission function, which is based on the parameterization by [113]. The comparison is carried out by taking averaged phase values from pixels within various ranges. In figure 4.9, from top to bottom, the curves show phase values at the peak only (figure 4.9-a), averaged phase over 3×3 pixels around the atomic position (figure 4.9-b), and the average over 5×5 pixels (4.9-c). Both, the curves of the ground truth transmission function and the CNN predictions, generally increase against atomic number, with the exception of certain dips at larger averaging ranges. This effect stems from different electron orbital distributions in the radial direction, and thus only the phase value averages at larger ranges are sensitive to this difference. The CNN predictions obviously also preserve these sub-atomic level details to some extent, as the shape of the curves bear strong resemblance to the ground truth curves. Although the reconstructed phase values and the transmission functions do not match exactly, the predictions are accurate enough, such that the phase values of the reconstructed objects are indeed useful as an indicator for different atomic species, potentially even allowing semi-quantitatively predicting the exact atomic species. For thick samples the method is not expected to yield results in quantitative agreement with projected potentials, because even if the neural network would retrieve the correct exit wave, the reconstruction algorithm is still based on the POA and inherits its limitations. The analysis of thicker samples presented here is therefore done in a more qualitative/empirical manner and comparisons are made against ADF imaging and SSB. ADF images are well known for their strong contrast related to the scattering power of the imaged object, and thus are suitable for examining the thickness variation of the sample [89] and local elemental compositions [168]. The contrast of SSB reconstructions is not as strong as ADF [169], yet the method is often used to study crystals containing elements of a wide range of atomic numbers due to its ability to image heavy and light atomic columns at the same time with distinguishable contrast [170]. Albeit a quantitative match can hardly be expected, it is important to verify whether reconstructed phase images still reflect the relative projected potentials of thicker samples to avoid misinterpretations. To that end an experimental dataset of a $SrTiO_3$ FIB-lamella was analysed. The reconstruction results are presented and compared to ADF and SSB images in figure 4.8.

For ADF imaging the contrast difference between the Sr and O columns is too large, making it difficult to locate the O columns without the help of the profiling. On the other hand, the SSB reconstruction does successfully image both atomic column types. While the peak intensities of the columns are ambiguous, they can still be distinguished by their corresponding size. The O columns are sharper than the Sr ones, indicating that

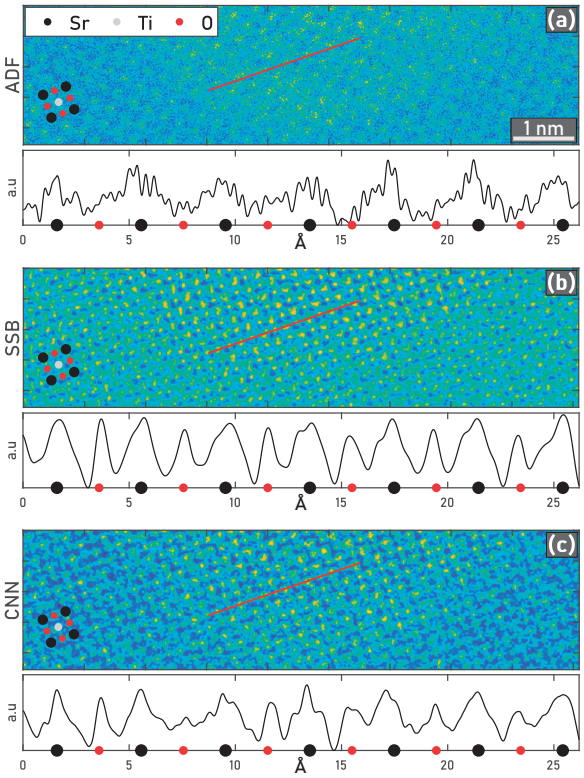


Figure 4.8 | Reconstructed images of $SrTiO_3$ with (a) virtual ADF detector, (b) SSB, and (c) CNN. Line profiles are drawn to illustrate contrast between the heavier Sr columns and the lighter O columns, which are also indicated respectively with black and red dots under the curves.

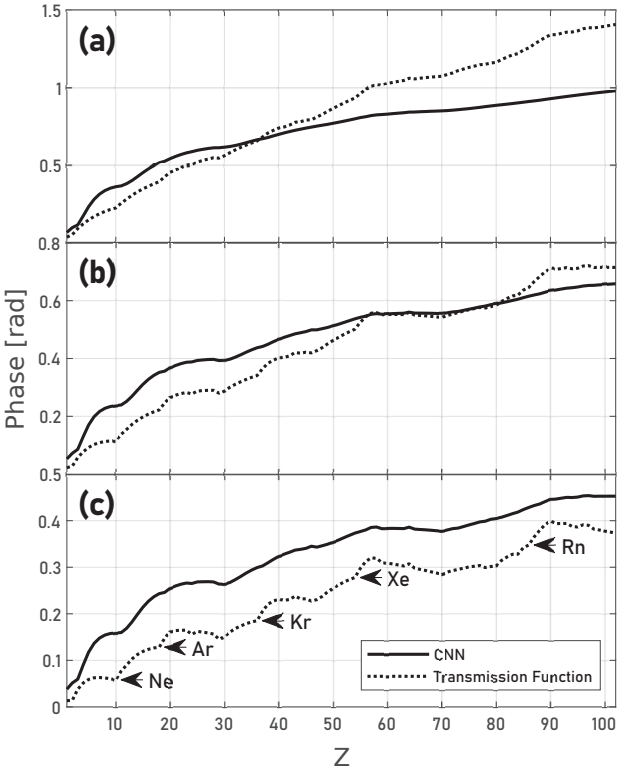


Figure 4.9 | Phase response of the CNN compared to ground truth transmission functions for simulations of single atoms throughout the periodic table at infinite dose for (a) peak intensity, (b) mean over 3x3 pixels around the atomic position, and (c) mean over 5x5 pixels around the atomic position.

an integrated signal from the area of each column could still be used as a reference of the local projected potential. The CNN reconstruction exhibits the advantage of both: while the light atom columns are observable, both the intensity and size differences are large enough to distinguish their types. This confirms, that the Z-contrast sensitivity is preserved for thicker specimen.

To further investigate the thickness dependence of the retrieved phase signal, an experimental dataset of the tip of an Au nanorod was used. As shown in figure 4.10, the intensity recorded by the virtual low angle ADF (LAADF) detector (20 mrad to 30 mrad) and HAADF detector (45 mrad and beyond) increases from the top to the bottom of the image. Based on statistical analysis of the HAADF signal to retrieve atom counts in each column [89], the thickest part in the image is about 9 nm. A line profile is then drawn for each imaging method presented in the figure. For the two ADF imaging methods at different collection angle, the profiles show monotonic increase at different pace against thickness, while the SSB profile only shows locations but the intensity is not correlated to the thickness of the atomic column. Compared to these profiles, the CNN reconstruction appears to be qualitatively most similar to the one of HAADF, and correlates with the estimated thickness accordingly. It should be noted that the maximum thickness of

≈9 nm is well outside the parameter range of the training data (<3 nm). Also the $SrTiO_3$-sample, being a FIB-lamella, should be well beyond 3 nm thickness. This means that these examples also demonstrate extrapolation capabilities of the CNN, which albeit being quantitatively arguably inaccurate, may still provide very useful reconstructions for imaging purposes. The strong resemblance with HAADF images at larger thicknesses may in fact be a very desirable characteristic, as it aligns well with many microscopists' experience and intuition.

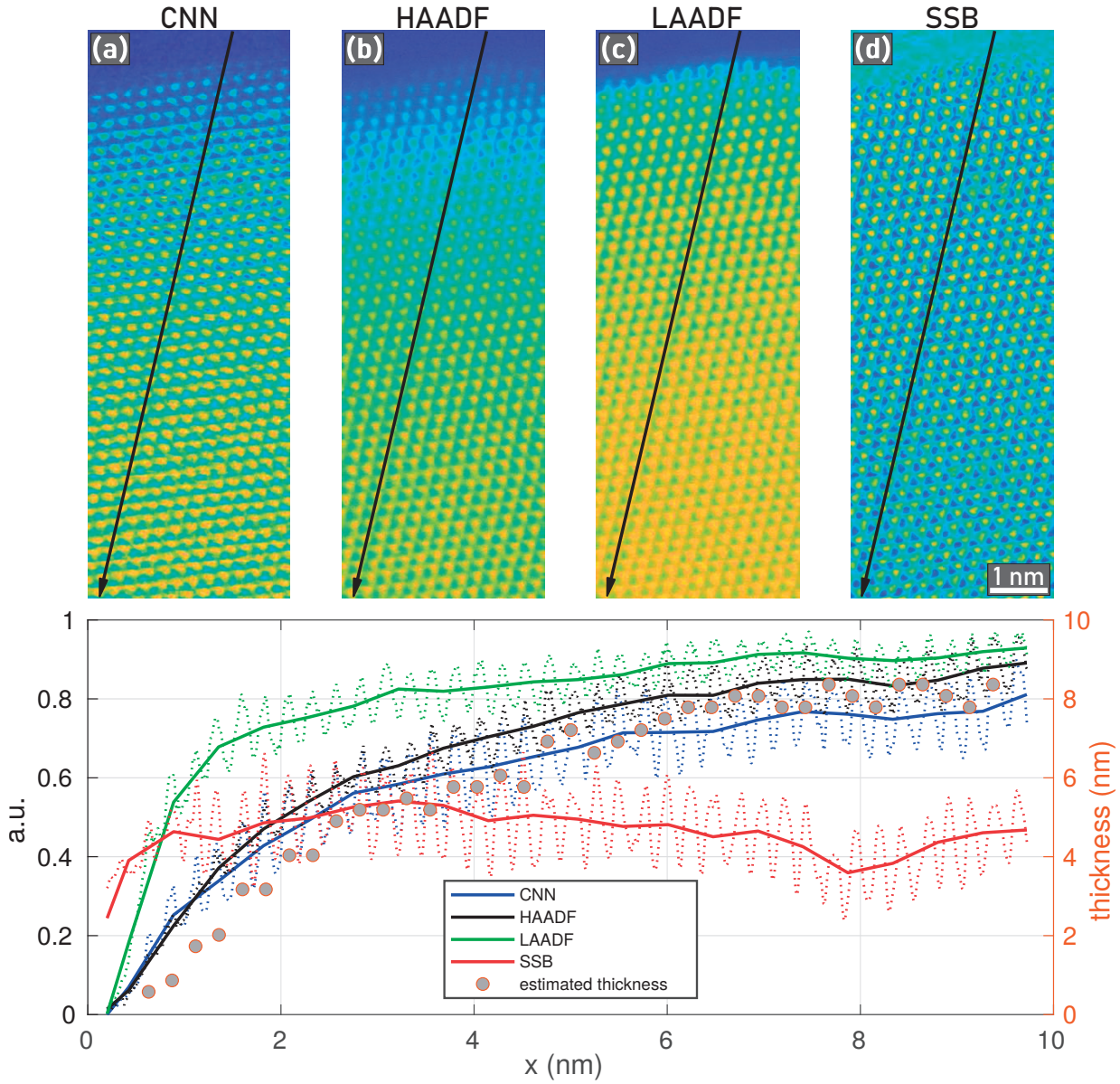


Figure 4.10 | Reconstructed images of an edge of a Au crystal using (a) CNN reconstruction, (b) HAADF, (c) LAADF and (d) SSB. Line profiles across the nanorod illustrate the thickness dependence of the corresponding signals.

By comparing the reconstructed images from the Au crystal and $SrTiO_3$, one would notice that SSB recovers contrast of higher spatial frequencies, such as the intensity and shape of atomic columns, but it does not recover long range features induced by e.g.

thickness variation. This is due to the band-pass characteristics of the method [50, 142], which in practice has a cutoff for high frequencies and a strong tendency to attenuate low frequencies. Long range features are built with low frequency components, and thus for reconstruction methods that filter out, or cannot utilize signals that fall in the low frequency end, these features are lost. For the CNN reconstruction, the object patches are also localized and no information beyond one probe position away are shared among the predictions of the exit waves. Therefore, the existence of thickness-related contrast variations can only be attributed to a reasonably good prediction accuracy of the CNN, also for the low frequency components.

4.3.4 Noise robustness

The performance of the method under different dose conditions is demonstrated and analysed on a simulated dataset of a twisted MoS_2 bilayer. The dose used for the reconstruction ranges from 500 to 1e5 e/Å^2, and the dataset is processed by the proposed method, SSB, and iDPC. The methods reconstructions are illustrated in figure 4.11 and compared against the ground truth transmission function in terms of their normalized cross correlations (equation 4.14).

It is evident that the method is strong in all three dose conditions, as can be seen visually in figure 4.11, and is confirmed by a cross correlations higher than those of the corresponding SSB and iDPC reconstructions. The higher cross correlation is not only due to a lower noise level, but also the generally better matched atom shape and phase value with respect to the ground truth, as indicated in the line profiles drawn under each dose condition (figures 4.11j-l). The sharper atom shape is due to the superior resolution, which is confirmed by the better distinction of close atom pairs, as well as by the higher frequency component found in the Fourier transform of the image (figure 4.12). In figure 4.11-l, the atom profile of the CNN reconstruction at high dose almost strikes a perfect match with the transmission function. Note that the CNN- and SSB-lines are not normalized or shifted, indicating that a very accurate wave retrieval is achieved by the CNN. Additionally, the contrast of the CNN results allows to distinguish Mo and S_2, which is more difficult with other methods, as their difference is much smaller. At low dose, the line profiles suggest a stronger low-dose robustness may be found in iDPC, as the two peaks are preserved in the reconstructed image. However, the signal almost completely falls into the noise level, as confirmed by the Fourier transform, and thus this seemingly better low-dose performance could very likely be a coincidental noise distribution. The given example highlights the potential of the proposed method for low dose imaging. As pointed out in sections "Super-Resolution" and Step Size" and illustrated in figures 4.6 and 4.7 respectively, the noise robustness may depend substantially on the step size and the effective dose per CBED. To gain an advantage over other methods in this regard this context needs to be taken into account and scanning strategies adapted accordingly.

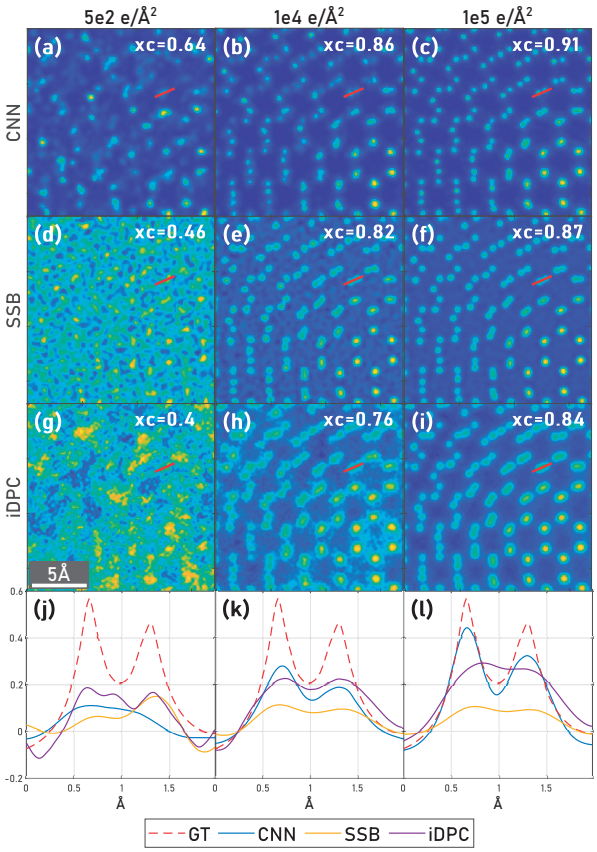


Figure 4.11 | The dose dependency of the proposed reconstruction method is demonstrated for a simulated dataset of twisted bilayer MoS_2 (top row panels) and compared to the corresponding SSB (2^{nd} row) and iDPC(3^{rd} row) reconstructions. Cross correlation values *xc* are given with respect to the ground truth phase object in panel of each result image. In (j, k, l) line profiles averaged in the perpendicular direction over 1.6 Å are drawn across a Mo-S_2 pair in all the images, and shown with the ground truth. iDPC values were normalized by the maximum value of the transmission function.

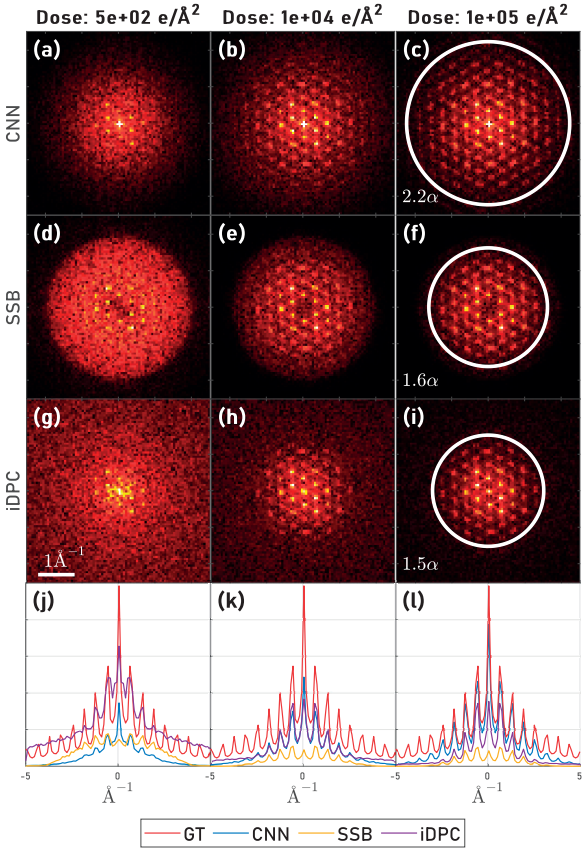


Figure 4.12 | Fourier transforms of the the the MoS2 dataset presented in figure 4.11. All images correspond to their counterparts in figure 4.11 with the same labels. Panels (j), (k) and (l) show compare the integrated intensities of the power spectrum along the y-axis. Circles in panels (c), (f) and (i) depict estimated resolution limits in the given example for the CNN, SSB and iDPC methods respectively.

4.4 Conclusion

This chapter presents a new computational imaging method, leveraging a CNN to retrieve complex exit wave functions from CBEDs and an algorithm to reconstruct the phase object from the predictions of the neural network. Since the exit waves are retrieved for each real-space coordinate in a 4D-STEM dataset, based only on a small kernel of adjacent diffraction patterns, the method can be employed in a sequential manner, thus enabling live imaging during an experiment. The machine learning system is based on a well established model but streamlined to the task at hand and adapted

to account for physical constraints and considerations. The model was trained on a large synthetic dataset of multislice simulations. Large and higher order aberrations, as well as CBED distortions, like non-centricity, geometric distortions and hot/dead pixels are not considered in the training data. Therefore, experimental data may require a pre-processing step. The range of practical conditions for which the method works reliably is therefore arguably limited accordingly to aberration corrected, well adjusted instruments at this stage. The trained model, code and training data are publicly available as summarized in the supplementary information. In the discussion multiple unique characteristics and advantages of the method are demonstrated. The CNN-based reconstruction is shown to enable higher resolutions than any other live-imaging-capable method considered, on simulated, as well as on experimental data, provided that a sufficiently high dose-per-CBED is maintained. In correspondence to this consideration the effect of the step size is analysed. While a better estimation of the exit wave is obtained with the electron dose-per-area distributed across fewer probe positions, some probe overlap is necessary to insure the accuracy of the exit wave retrieval. Hence, the method is most suitably applied at a balanced scan density. If these considerations are taken into account the reconstruction method can be very dose efficient.

The Z-contrast was analysed on single atom-simulations across the periodic table. The phase signal of the reconstructions could indeed be linked qualitatively to atomic properties and a semi-quantitative analysis of thin specimen within the limits of the POA, was shown to be possible. We confirmed the contrast sensitivity to atomic species and sample thickness on experimental datasets of a $SrTiO_3$ FIB-lamella and an Au nanorod respectively. The observed monotonic increase of the phase signal with thickness and nearly monotonic increase with atomic number indicates that quantitative analyses based on the reconstruction results may be feasible.

Generally, we believe the proposed method presents an attractive imaging modality for its super-resolution capability, high noise robustness, and the feasibility of qualitative or even quantitative contrast analysis. While further studies would be necessary to obtain a more detailed view on the model performance over the entire parameter space (and beyond), we could already show that the method is robust for a wide range of practically meaningful applications, even exhibiting reasonably good extrapolation behaviour well beyond the maximum sample thickness of the training data. The fact that none of the examples shown in this study exist in those exact configurations in the training data, further indicates that the system generalizes well within the parameter interpolation range as well.

4.5 Supporting Information

Code and Data

Trained Neural Network and Reconstruction implementations:
authors: Thomas Friedrich and Chu-Ping Yu
title: airpi
license: GNU GPL3
url: https://github.com/ThFriedrich/airpi

Training Datsets:
authors: Thomas Friedrich, Chu-Ping Yu, Jo Verbeeck and Sandra van Aert
title: Phase Object Reconstruction for 4D-STEM using Deep Learning, (4D-STEM Training Data)
doi: 10.5281/zenodo.6971200
license: Creative Commons Attribution 4.0 International Public License
url: https://doi.org/10.5281/zenodo.6971200

Data generation code:
authors: Thomas Friedrich and Chu-Ping Yu
title: ap_data_generation
license: GNU GPL3
url: https://github.com/ThFriedrich/ap_data_generation

STO datset:
authors: Chu-Ping Yu, Thomas Friedrich, Daen Jannis, Xie Xiaobin, Sandra van Aert and Jo Verbeeck
title: Real Time Integration centre of Mass (riCOM) Reconstruction for 4D-STEM
doi: 10.5281/zenodo.6971200
license: Creative Commons Attribution 4.0 International Public License
url: https://doi.org/10.5281/zenodo.5572123

Zeolite datset:
authors: Daen Jannis, Christoph Hofer, Chuang Gao, Xiaobin Xie, Armand Béché, Timothy J. Pennycook and Jo Verbeeck
title: Event driven 4D STEM acquisition with a Timepix3 detector: microsecond dwelltime and faster scans for high precision and low dose applications
doi: 10.5281/zenodo.6971200
license: Creative Commons Attribution 4.0 International Public License
url: https://doi.org/10.5281/zenodo.5068510

Example Datsets:
authors: Thomas Friedrich, Chu-Ping Yu, Jo Verbeeck and Sandra van Aert
title: Phase Object Reconstruction for 4D-STEM using Deep Learning, (4D-STEM Examjple Data)
doi: 10.5281/zenodo.7034879
license: Creative Commons Attribution 4.0 International Public License
url: https://doi.org/10.5281/zenodo.7034879

Additional data and figures

Table 4.2 | Table summarizing all occurrences of material structures used in this chapter in the training data with corresponding simulation settings. This shows, that the examples presented in the chapter are not actually part of the training data in those settings. "SG" is short for Space Group. Δs refers to the probe-scan-step-size, "E0" to the microscope acceleration voltage.

SG	Zone Axis	Rotation [°]	Thickness [Å]	E0 [kV]	Aperture [mrad]	Detector Size [$Å^{-1}$]	Δs [Å]
Gold (mp-81)							
225	1 1 1	338.3	15.8	300	20.94	5.216	0.180
225	0 0 1	50.7	22.8	200	8.02	0.522	1.357
225	1 0 0	84.7	9.1	100	22.63	0.955	0.617
225	0 1 1	148.0	18.5	180	24.97	3.073	0.375
Graphene (mp-48)							
194	1 0 1	2.6	14.1	80	15.63	0.398	0.364
194	0 1 0	273.4	13.0	160	25.28	4.242	0.405
194	1 0 0	217.1	7.6	300	20.63	1.871	0.302
194	0 1 0	168.9	24.1	30	28.43	0.473	1.070
MgO (mp-1265)							
225	1 1 0	274.9	6.2	200	22.50	1.031	0.348
MoS_2 (mp-2815)							
194	1 0 1	75.8	8.1	140	10.60	1.355	0.381
194	0 1 1	45.9	7.9	120	26.46	0.855	0.685
194	0 1 1	330.8	23.1	200	29.22	1.370	0.145
194	1 0 1	256.0	9.4	180	18.86	1.945	0.191
194	0 1 1	22.1	29.7	160	20.48	0.760	0.471
$SrTiO_3$ (mp-5229)							
221	1 1 0	300.6	14.2	120	20.42	3.693	0.755
221	0 1 1	64.8	14.7	100	5.57	0.581	1.416
221	0 1 0	356.4	12.1	300	25.83	1.796	0.383
221	1 0 0	122.6	28.0	160	22.50	0.845	0.223
221	1 0 1	346.2	15.4	100	7.88	0.461	0.892
221	0 1 0	194.9	9.1	40	22.55	0.486	0.426
221	0 1 0	1.8	23.5	300	20.42	1.187	0.363
221	0 1 1	139.2	15.3	160	22.61	1.072	0.446

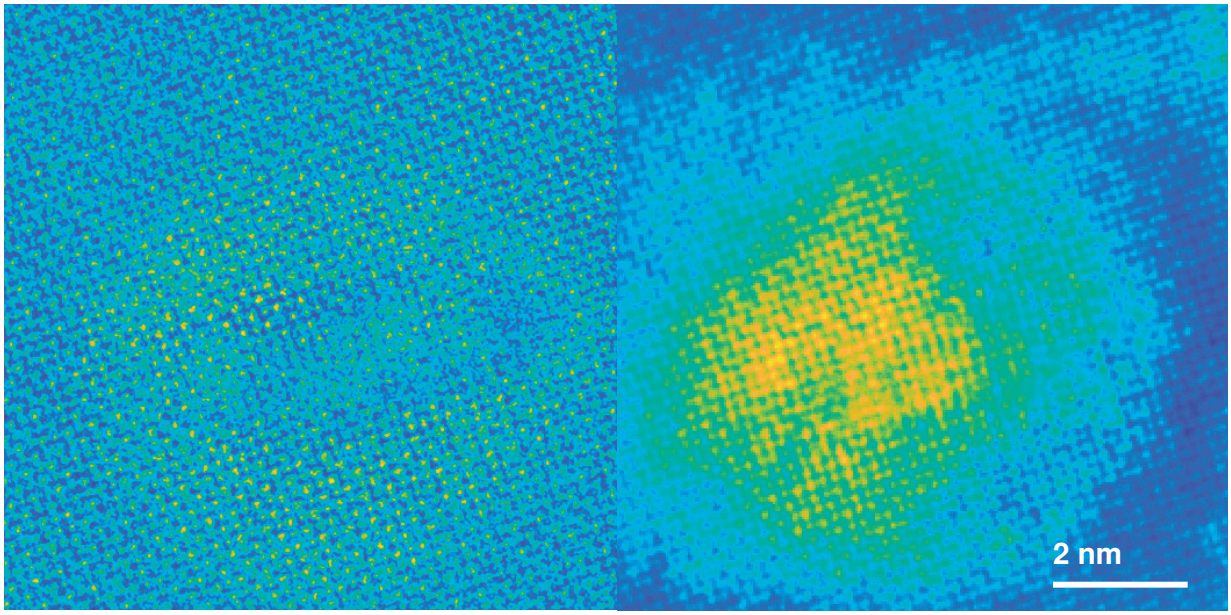


Figure 4.13 | Comparison of iDPC (right) and Neural network reconstructions (left) of a FIB lamella, including a hole in the center. This illustrates reasonable tolerance of the proposed method towards thickness variations and bending.

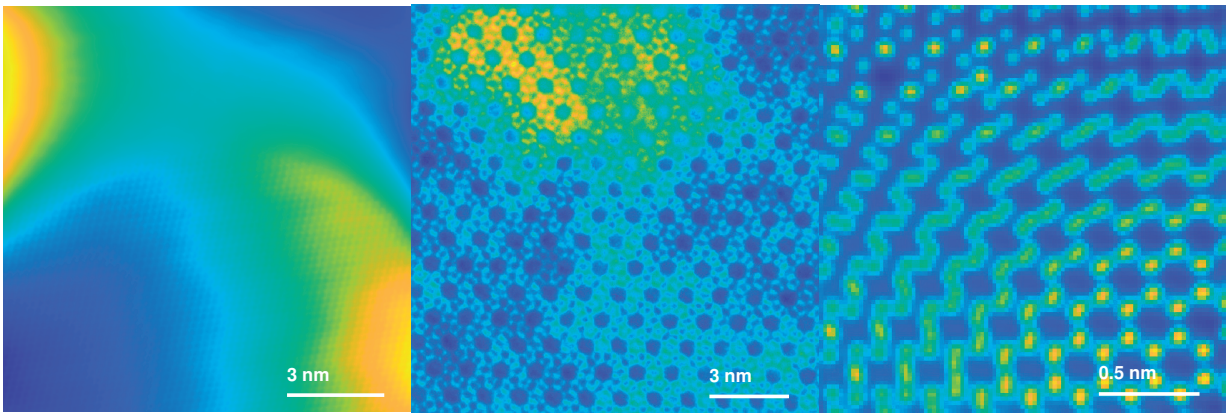


Figure 4.14 | Additional iDPC reconstructions for comparison of samples presented in the chapter, f.l.t.r: experimental Gold nanorod, experimental Zeolite, simulated twisted bilayer graphene with infinite dose.

CHAPTER

5 Conclusions

Contents

This chapter uses a set of examples to compare the developed methods described in this thesis to SOTA algorithms, in order to draw conclusions in a contextualised manner. This allows to summarise the findings, estimate the potential impact and identify possible applications and some future development opportunities.

5.1 Distortion correction

We could show that the use of modern CNNs holds enormous potential for the correction of distortions. While CNNs have been used in generic image processing for denoising purposes for some time, the inclusion of accurate STEM distortion models adds significant capabilities to the system presented here. Figure 5.1 shows a comparison of the CNN performance to SOTA algorithms for image denoising, nameley the Block-matching and 3D filtering algorithm (BM3D) [171] and the application of a Wiener filter in combination with Anscombe and inverse Anscombe transformations [172]. Images are being evaluated by the peak-signal-to-noise-ration (PSNR). It can be seen that generic denoisers can indeed also deliver some improvements in the mid-dose range, but generally fail to account for scan-line distortions. Also, and especially for low doses, generic algorithms cannot infer any underlying structure, while the CNN can clearly recover atomic features. Also comparing the performance to other machine learning/CNN based de-

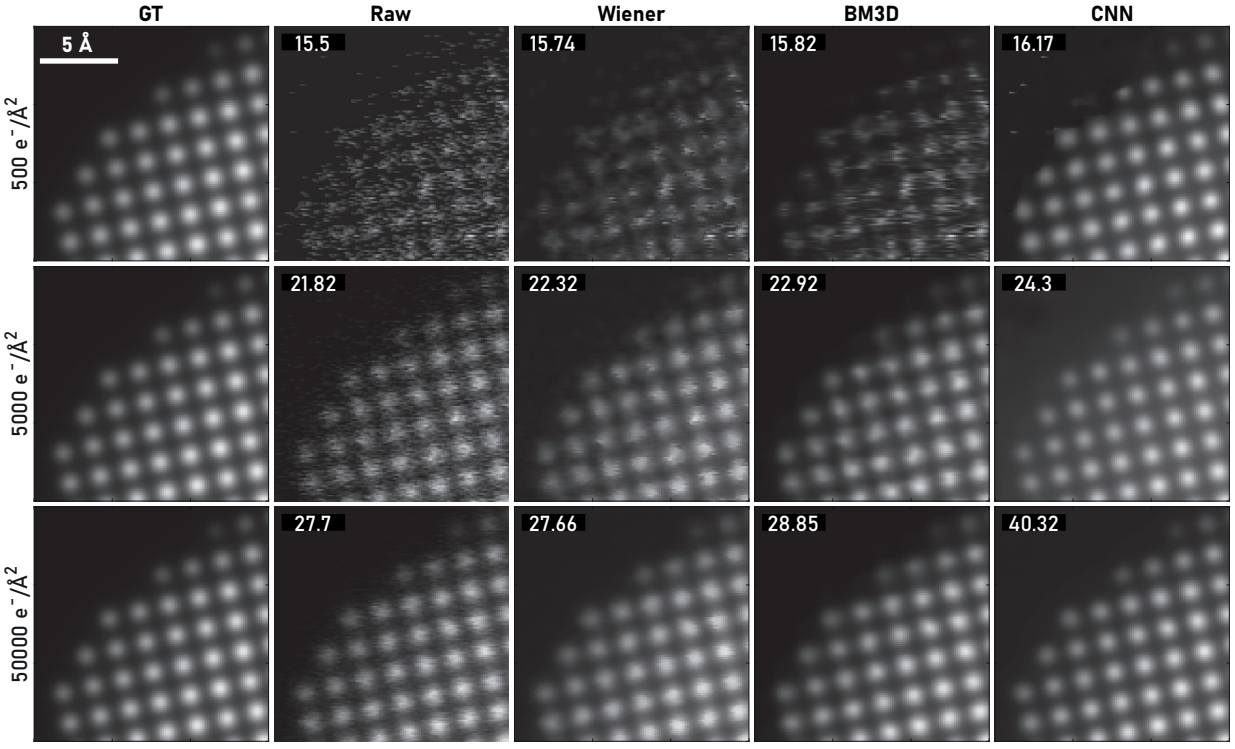


Figure 5.1 | Comparison of CNN-restoration, Wiener filter and BM3D algorithm for image denoising on a simulated Pt-NP. The row labels show the dose applied to the ground truth (GT) image to get the raw image. The numbers given in each reconstruction is the PSNR.

noising approaches illustrates the importance of the proper modelling of the noise and distortion sources as seen in figure 5.2. In this example we compare the reconstruction

performance of our CNN, AtomSegNet [30] and the Noise2Void-NN (N2V) [32], where the N2V was retrained on the presented experimental image itself. The sample is a $BaHfO_3$ nanoparticle (figure 5.2-③) embedded in a superconducting $REBa_2Cu_3O_{7-\delta}$ (REBCO) matrix [173, 174] (figure 5.2-②), which was grown on a $SrTiO_3$ substrate (figure 5.2-①). While all three networks successfully remove the noise from the image, there are notable differences in the reconstruction results. In region ① the N2V reconstruction recovers all the weaker intensities of the Ti + O columns to some degree, which is not the case for the AtomSegNet reconstruction. There, some of the columns blur or even disappear. Our CNN reliably recovers all atomic columns with superior contrast to the other two methods. Similar improvements are evident also in region ② but most notably in region ③, which is an embedded $BaHfO_3$ nanoparticle. This region at the top of the image is also degraded, presumably by either FIB damage or carbon contamination. In both N2V and AtomSegNet reconstructions features tend to blur into diagonal streaks, while our CNN recovers clearly distinguishable atomic columns and, given that the $BaHfO_3$ nanoparticle was grown epitaxially on the $SrTiO_3$ substrate, that is indeed what would be expected. Considering the N2V network is a generic denoising network, the results are quite remarkable, albeit the additional training step is somewhat inconvenient from a user perspective. However, this example illustrates that the CNN presented in this work does not only benefit from the latest advances in deep learning, but also from the development of accurate, physically meaningful models of all distortions specific to HAADF-STEM. This CNN is shown to be accurate, not only in perceived contrast enhancement, but also in a quantitative way which boosts the accuracy and precision of atomic structure determinations in ADF-STEM studies. In some scenarios only the proposed restoration method was able to recover atomic features at all. Such cases should generally be checked quite carefully, but so far we have no evidence of the CNN generating any "dreamed features". In numerous tests, well beyond what is shown in this thesis, the CNN has proven very relaible across many samples and imaging conditions. The method therefore provides a significant contribution to the wider electron microscopy community. We believe that the combination of sound physical models underlying the training data and a rigourous approach to model evaluation throughout the development process, including repeated quantitative studies, make a strong point for the trustworthiness of the reconstructions and we hope that the community will appreciate the contribution.

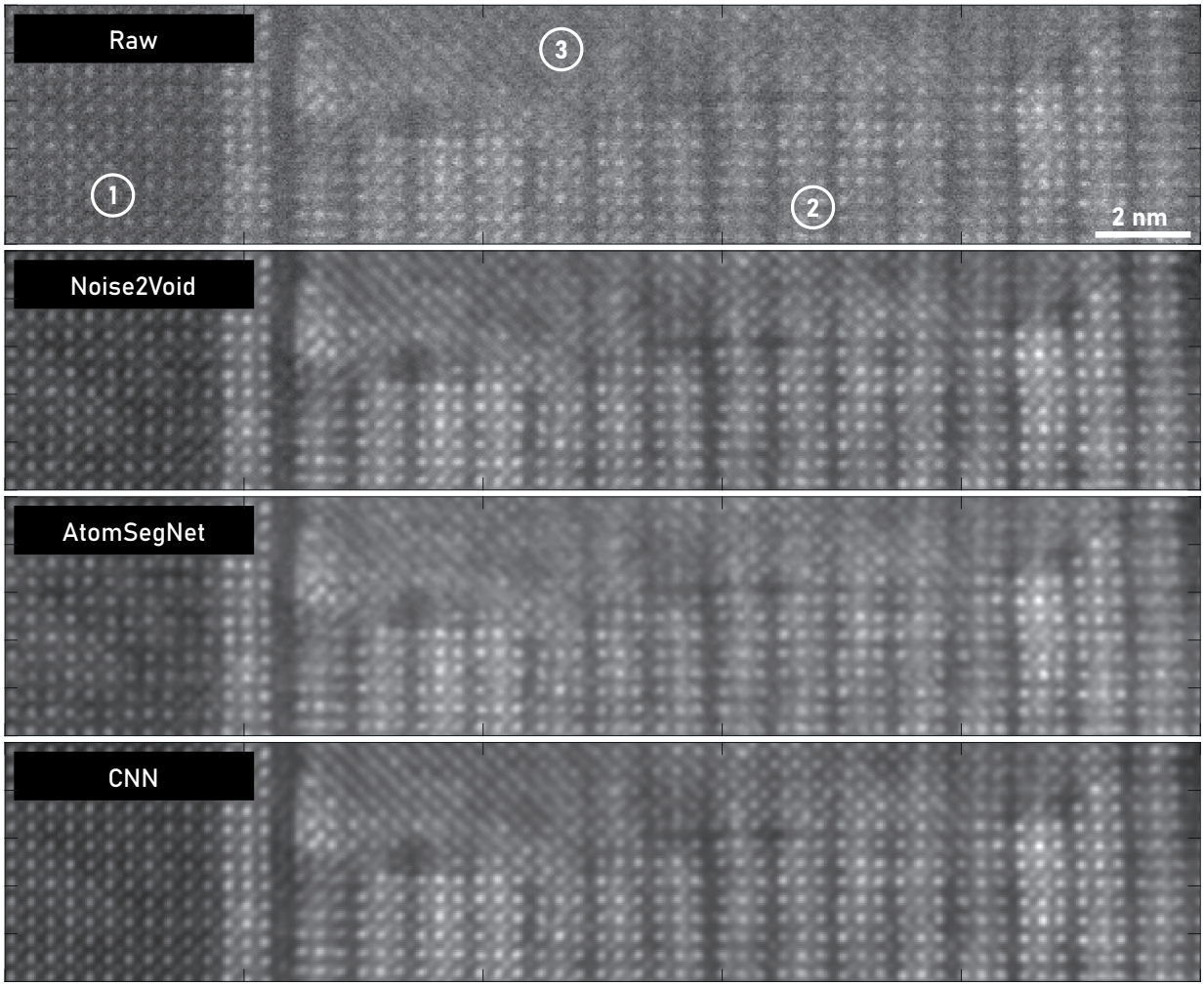


Figure 5.2 | Comparison of different CNN-restoration approaches on an experimental HAADF-STEM dataset of a $BaHfO_3$ nanoparticle (③) embedded in a superconducting $REBa_2Cu_3O_{7-\delta}$ (REBCO) matrix (②), which grew epitaxially on a $SrTiO_3$ substrate(①). Images were acquired on a non-probe-corrected Titan microscope with 300 keV at KIT Karlsruhe. The data is described in detail in references [173] and [174]

5.2 Live 4D-STEM Methods

This thesis describes two new methods to interpret 4D-STEM data, riCOM[1] and AIRPI, both of which can run efficiently and accurately on a very small subset of CBEDs, such that an image of the object is created one scan-point at a time, without iterative optimizations. This not only permits live imaging in principle, but also allows experimentalists to work with datasets of arbitrary size, as the computer memory is no longer dictating the maximum file size. RiCOM benefits from a dedicated and efficient reformulation and implementation of conventional algorithms, based on DPC, while AIRPI is using a CNN at its core to reconstruct electron exit waves. The imaging characteristics of both methods were described in detail in chapters 3.4 and 4.3, but some further examples, including direct comparisons between the two methods and established algorithms are presented here in order to provide a better context as to how and in which circumstances these new approaches may be useful. Figure 5.3 shows images of a Pt nanoparticle, embedded in an amorphous carbon and graphite matrix, taken at 200 keV and a convergence angle of 21 mrad. In this example the probe and step size are relatively small and the

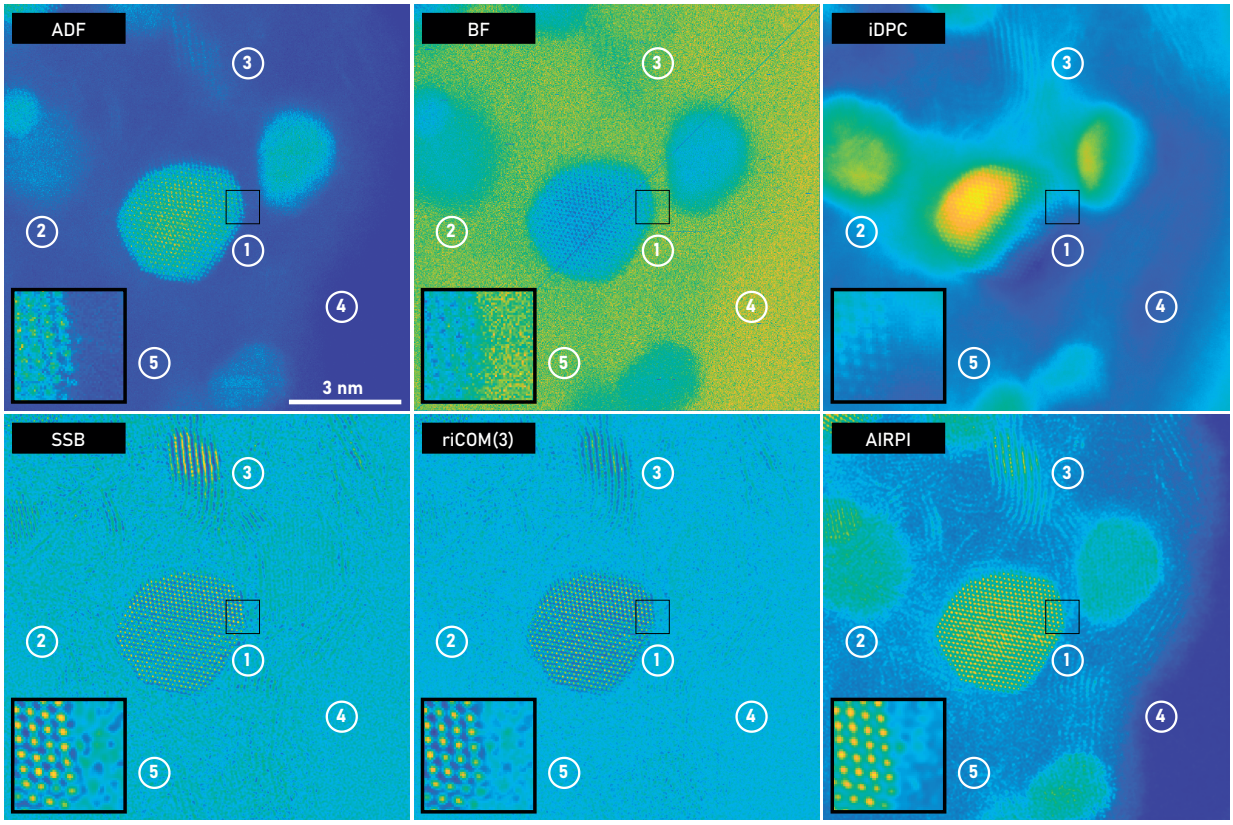


Figure 5.3 | Comparison of different reconstruction methods on a Pt-nanoparticles in a graphite matrix.

dose is actually high enough, such that ADF and BF images can show reasonably good contrast at atomic resolution as seen at the NP at (1). In the ADF image the edge of the

[1] In figure labels the kernel size used for the riCOM reconstruction is given in brackets, e.g. riCOM(3) denotes a kernel radius of 3 pixels. No filtering was applied unless explicitly stated.

C matrix (④), as well as the out-of-zone-axis NPs at ②,⑤ and to the right of ① can be identified due to the Z- and thickness contrast dependency. However, due to the relative insensitivity to light elements, many details of the graphite arrangements are lost (e.g. at ③ or along the edge at ④). That is where phase imaging methods, including iDPC, SSB, riCOM and AIRPI can provide additional information. However, the iDPC result is difficult to interpret due to the large low-frequency obstructions. The flexibility riCOM allows in terms of frequency selection makes it possible to isolate the contributions responsible for the features of interest. This way, the atomic resolution features in the image can be highlighted, at the expense of low frequency features, such as the particle at ② or the edge (④), which become all but invisible. This is also seen for the SSB reconstruction, which has an intrinsic frequency filtering effect due to its contrast transfer function (CTF) [50]. Though the results of SSB and riCOM are very similar, it should be noted that riCOM is orders of magnitude faster, especially in a live imaging scenario and easier to use from a user perspective. AIRPI shows a reconstruction dominated by Z- and thickness contrast, similar to the ADF image. Because it is more sensitive to light elements, the carbon matrix has a stronger effect and the graphite sheets are still clearly visible. At the same time the resolution is also very good, so no details are lost or blurry. This may even cause different interpretations to be made. For example the magnified insets show a supposedly single Pt atom or column to the side of the particle. Only in the AIRPI image the two graphite sheets in between the particles are clearly visible, giving rise to the possibility that this contrast may be caused by the carbon, similar as at point ③. All previously mentioned features are visible in the AIRPI-reconstruction as well, with superior resolution which is arguably a very intuitive and comprehensive representation of the actual sample under study. Particularly in a live imaging context, when exploring the sample, this is likely to be the most favourable CTF among the methods presented here for many microscopists.

A similar point can be made for the example given in figure 5.4, which shows PbBrFA nanocrystals (NCs) in (110) zone axis orientation. This organic-inorganic lead halide perovskite contains a formamidinium (FA) group at the A-cation [175, 176], which consists of light atoms (C,H and N) only.

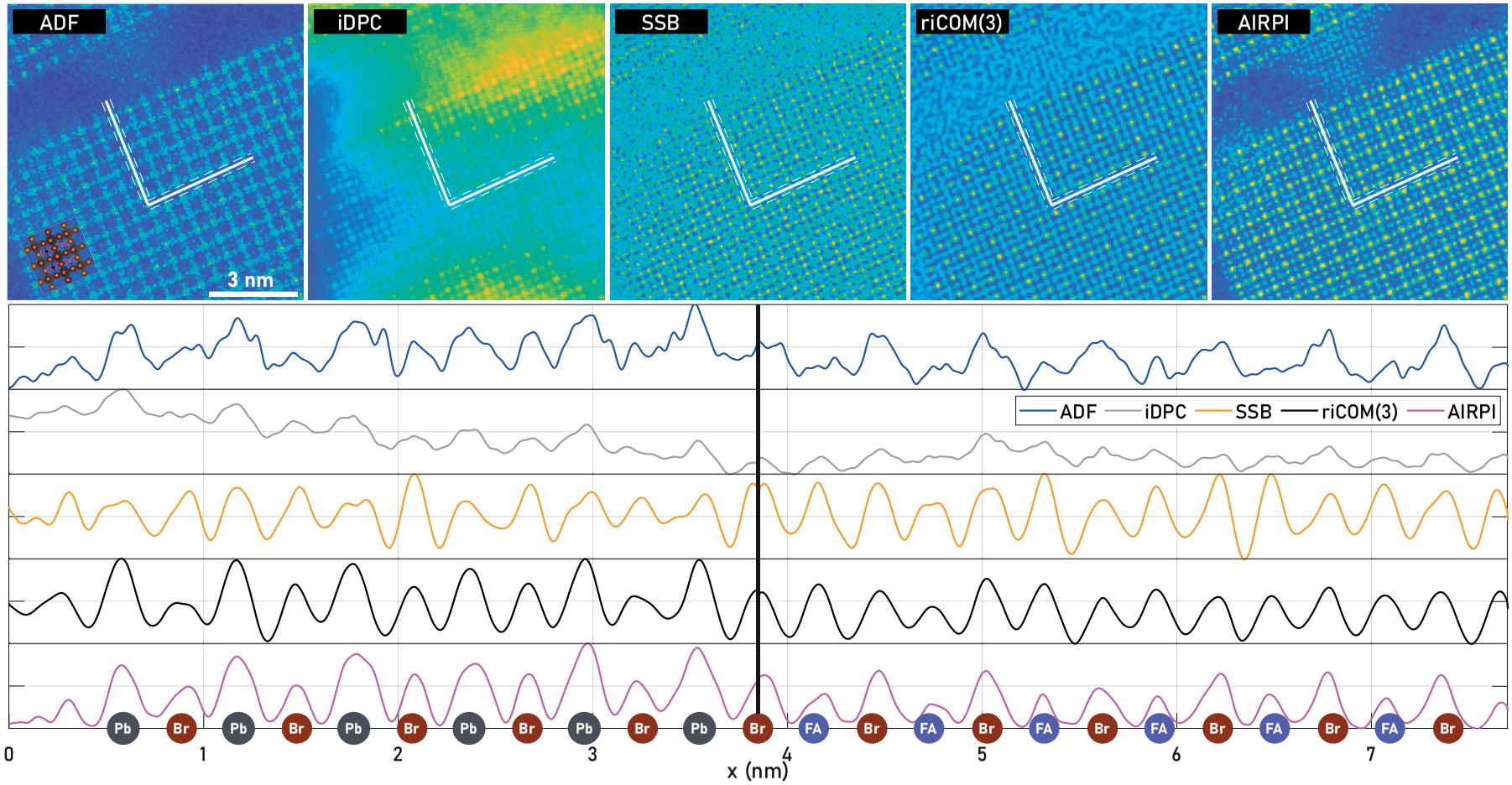


Figure 5.4 | Comparison of element contrast of common reconstruction methods on PbBrFA nanocrystals. This dataset was recorded using 200keV, an convergence angle of 13 mrad and a step size of 0.21 Å. Line profiles are overlayed on the images and plotted out underneath the images. The direction goes from the top left to the bottom right of the lines, around a corner, such that the first half of the line goes across some Pb + Br, and pure Br columns and the second half spans across pure Br and the FA columns. Dashed lines indicate the integration range used for the line profiles.

Accordingly, they are for all intents and purposes invisible in the ADF image. The intensity smudge in between the two NCs shows weak atomic features in iDPC, SSB and riCOM, while in the AIRPI reconstruction this can be clearly identified as a small Pb-NC in zone axis, which may form as a result of beam damage. Line profiles are overlaid on the images and plotted out underneath the images in figure 5.4. The direction goes from the top left to the bottom right of the lines, around a corner, such that the first half of the line goes across some Pb + Br, and pure Br columns and the second half spans across pure Br and the FA columns. Dashed lines indicate the integration range used for the line profiles. Despite the integration, the ADF profile is still very noisy and only the Pb columns can be reliably identified, but Br columns are mostly visible too. The iDPC reconstruction is again obstructed by low frequency components, although all columns show some contrast. The riCOM and SSB reconstructions are qualitatively fairly similar. However, it is easier to distinguish the different column types in riCOM. At least the Pb + Br, and the pure Br columns can be separated by their different peak intensities. However, the FA- and pure Br-columns can be reliably distinguished only by AIRPI, as seen in the right half of the line profiles of figure 5.4. Therefore, also for this example AIRPI shows some advantageous imaging characteristics.

However, this is not always the case. Figure 5.5 shows a reconstruction of a single layer WS_2 [167]. With a step size of only 0.09 Å, an acceleration voltage of 60 kV

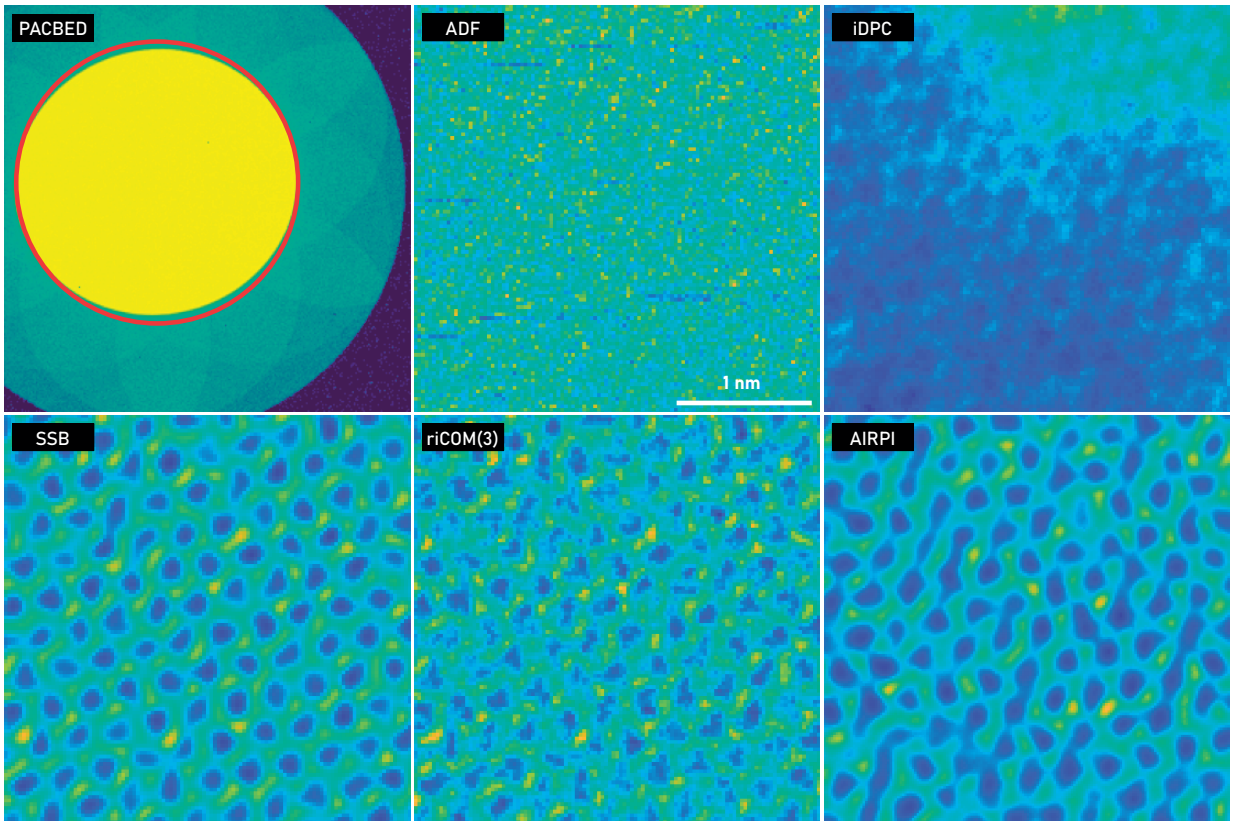


Figure 5.5 | Comparison of common reconstruction methods on 2D-WS_2. The dataset was recorded at 60 keV, an convergence angle of 25 mrad and a step size of 0.09 Å. It includes residual aberrations, most notably 2-fold astigmatism, The dataset was binned by a factor 8 for AIRPI.

and an aperture angle of 25 mrad this data was recorded with a relatively large probe ($d \approx 4.75$Å), dense scan grid and low dose. These settings are most beneficial for ptychographical reconstructions, and likely the data was recorded specifically to suit these algorithms. Methods like ADF, iDPC and riCOM, which depend on a sharp probe for a high resolution obviously are at a disadvantage with such settings. AIRPI can in principle handle large probes and even benefits from it to some extent, provided that the dose per CBED is large enough, as described in section 4.3.2. This can be mitigated by binning down the dataset (e.g. adding 2x2 CBEDs into one while decreasing the step size by 50%), which was done for the reconstruction in figure 5.5 by a factor of 8 to obtain a decent image. The SSB reconstruction is still arguably of better quality. Inspecting the position-averaged-CBED (PACBED) in figure 5.5 reveals some further aspects disadvantageous for AIRPI. Firstly, the pattern is highly off centre and includes shadowing from the HAADF detector. Even though we correct for this COM-shift in a pre-processing step, a large amount of the dark field electrons is lost, even worse so in a non-radially symmetric way. This scenario was not considered in the training and is likely to deteriorate the accuracy of the CNN. Also non-spherically-symmetric aberrations were not considered, but as highlighted by the overlaid red circle in the PACBED, the probe function clearly includes 2-fold astigmatism, which results in an elongation of the probe in one axis. A great advantage of SSB and other ptychographic methods is that they can actually correct for these aberrations (post-acquisition at least), while AIRPI relies on a very well adjusted instrument with probe corrector and a reasonably accurate approximation of the probe function. However, this probe function is hardly known exactly, which clearly presents a limitation to AIRPI in its current implementation. Considering how poorly the WS_2 dataset is suited for AIRPI, it is still encouraging that the reconstruction does not fail completely or introduces unreal artefacts but merely degrades in quality and resolution.

Table 5.1 summarises the most important aspects of the discussed 4D-STEM reconstruction methods. It is fair to state that in terms of dose efficiency and resolution vSTEM methods have a clear disadvantage compared to all phase retrieval approaches, but it allows quantitative imaging (e.g. atom counting), especially when employing techniques like complementary ADF [177]. Due to the missing low frequency components phase retrieval methods hardly allow quantitative structure analysis beyond a thickness of a few atoms. AIRPI can overcome these limitations in part as shown in figure 4.10. Wether it can be used reliably for atom counting or not will have to be determined in a dedicated study, but the results outlined in section 4.3.3 suggest that the contrast is indeed dependent on atomic species and thickness, as it is the case for ADF, albeit the Z-sensitivity and behaviour at larger thicknesses is different. Regarding the usability of the methods ADF and riCOM are the easiest to use as they require only very few parameters that are intuitively understood without too much in-depth knowledge of the method and require no pixel-size calibration. They are also considerably faster and can benefit greatly and easily from event-driven electron cameras. Independent of the implementation, the memory requirements and computational complexity scales with $\mathcal{O}(1)$ and $\mathcal{O}(N)$ respectively for ADF, riCOM and AIRPI, which is an important advantage over direct ptychography methods such as SSB [148] and WDD [178], when scaling to larger datasets. Table 5.1 is a somewhat simplistic summary of these methods, but

Table 5.1 | Table comparing the features of live imaging approaches ADF, SSB, riCOM and AIRPI.

	ADF	SSB	riCOM	AIRPI
Necessary information				
	• COM • integration limits	• conv. angle • step size • energy • rotation	• kernel size • rotation • filter limits (optional)	• conv. angle • step size • energy • rotation
Imaging characteristics				
Contrast determined by	Species and thickness[a]	Frequency (experiment conditions)	Frequency (adjustable)	Species and thickness
Resolution limit	$\lessapprox$ Probe size	$\lessapprox$ Convergence angle	$\lessapprox$ Probe size	$\lessapprox$ Collection angle
Dose robustness	Poor	Very good[b]	Very good	Very good[b]
Quantitative imaging	yes	no	no	limited/possibly
Performance				
Speed (Frames)	$\approx$ 30 kHz	$\approx$ 1 kHz [148]	$\approx$ 30 kHz	$\approx$ 4 kHz
Speed (Events)	$\gtrapprox$ 100 kHz	-	$\gtrapprox$ 100 kHz	-
Memory scalability	$\mathcal{O}(1)$	$\mathcal{O}(N)$	$\mathcal{O}(1)$	$\mathcal{O}(1)$
Computanional scalability	$\mathcal{O}(N)$	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$	$\mathcal{O}(N)$

[a] Light elements are often not visible
[b] At appropriate experimental conditions

captures the most important points and may be useful as a quick-reference overview for experimentalists, looking for the right tool for their image reconstructions.

Our contribution to the microscopy community through the work on riCOM and AIRPI lies not only in developing and publishing these methods, but also in making them accessible. This includes developing performant code beyond the proof-of-concept level, publishing this code open source under permissive licenses and documenting it as

well. Especially the riCOM software has reached a relatively advanced state, with well performing C++ implementations including camera interfaces and file type handling for the MediPix 3 and AdvaCam TimePix 3 cameras, and a live interface for the "CheeTah"-camera from Amsterdam scientific instruments (ASI) currently under development. This allows not only live imaging, but also presents a very convenient tool to quickly inspect datasets, without the need for any kind of application programming interface (API) or python scripting. This also implies that no setup of Java environments, python or virtual environments are necessary (contrary to most other toolboxes available to the community). Instead the entire software compiles into a ≈1 MB executable, that links into standard OS libraries. Since the compilation (for Windows and Linux) is automated in a GitHub continuous integration (CI) pipeline, riCOM requires essentially no setup beyond downloading the pre-compiled software. RiCOM includes visualisation options for riCOM, COM-X and COM-Y maps, vSTEM and electric field mapping. The combination of an intuitive GUI (and CLI for e.g. bash-scripting), out-of-the-box interfaces, no meaningful hardware requirements, easy setup and various imaging modalities makes this a unique tool for microscopists. Further improvements, besides implementing more camera interfaces, could be made by a more efficient multi-threading strategy. This mainly affects the performance of the COM integration, which presents a bottleneck when the kernel size is large in the current implementation. For AIRPI the situation is somewhat more difficult due to the relatively complex dependency environment for GPU-accelerated machine learning frameworks. For this reason the code and trained model were published on GitHub alongside a Docker container, which is comparable to a virtual machine that includes all necessary libraries and dependencies. This makes the use much easier but still requires Docker to be installed (and Windows subsystem for Linux (WSL) for Windows users), which can also be challenging. Eventually, we hope to use Tensorflow Lite for the model deployment, which allows model quantization that could potentially lead to significant performance improvements. Since Tensorflow Lite has language support for C++ this would also open a path towards integrating AIRPI into the riCOM codebase, where all the camera- and file-interfaces, as well as visualisation tools could be readily used. This transition would require a significant amount of technical work however.

References

[1] J. Black and K. Gass. *Master Exploder*. The Pick of Destiny. Epic records, Nov. 14, 2006.

[2] I. Lobato, T. Friedrich, and S. Van Aert. *Deep Convolutional Neural Networks to Restore Single-Shot Electron Microscopy Images*. Version 1. Mar. 29, 2023. DOI: 10.48550/arXiv.2303.17025. arXiv: 2303.17025. URL: http://arxiv.org/abs/2303.17025. preprint.

[3] C. Louis and O. Pluchery. *Gold Nanoparticles For Physics, Chemistry And Biology (Second Edition)*. World Scientific, June 2, 2017. 681 pp. ISBN: 978-1-78634-126-6. Google Books: VPAnDwAAQBAJ.

[4] J. Y. Kim, J.-W. Lee, H. S. Jung, H. Shin, and N.-G. Park. "High-Efficiency Perovskite Solar Cells". In: *Chemical Reviews* 120.15 (Aug. 12, 2020), pp. 7867–7918. ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.0c00107.

[5] A. Aghababai Beni and H. Jabbari. "Nanomaterials for Environmental Applications". In: *Results in Engineering* 15 (Sept. 1, 2022), p. 100467. ISSN: 2590-1230. DOI: 10.1016/j.rineng.2022.100467.

[6] Z. Cao et al. "Ultrathin ZSM-5 Zeolite Nanosheet Laminated Membrane for High-Flux Desalination of Concentrated Brines". In: *Science Advances* 4.11 (Nov. 23, 2018), eaau8634. DOI: 10.1126/sciadv.aau8634.

[7] Y. Wang et al. "Fabrication and Evaluation of Effective Zeolite Membranes for Water Desalination". In: *Desalination* 504 (May 15, 2021), p. 114974. ISSN: 0011-9164. DOI: 10.1016/j.desal.2021.114974.

[8] M. J. Ndolomingo, N. Bingwa, and R. Meijboom. "Review of Supported Metal Nanoparticles: Synthesis Methodologies, Advantages and Application as Catalysts". In: *Journal of Materials Science* 55.15 (May 1, 2020), pp. 6195–6241. ISSN: 1573-4803. DOI: 10.1007/s10853-020-04415-x.

[9] S. Iijima. "Helical Microtubules of Graphitic Carbon". In: *Nature* 354.6348 (Nov. 1991), pp. 56–58. ISSN: 1476-4687. DOI: 10.1038/354056a0.

[10] J. P. Giraldo et al. "Plant Nanobionics Approach to Augment Photosynthesis and Biochemical Sensing". In: *Nature Materials* 13.4 (Apr. 2014), pp. 400–408. ISSN: 1476-4660. DOI: 10.1038/nmat3890.

[11] M. Safdar et al. "Engineering Plants with Carbon Nanotubes: A Sustainable Agriculture Approach". In: *Journal of Nanobiotechnology* 20.1 (June 14, 2022), p. 275. ISSN: 1477-3155. DOI: 10.1186/s12951-022-01483-w.

[12] M. J. Mitchell et al. "Engineering Precision Nanoparticles for Drug Delivery". In: *Nature Reviews Drug Discovery* 20.2 (Feb. 2021), pp. 101–124. ISSN: 1474-1784. DOI: 10.1038/s41573-020-0090-8.

[13] N. Claes. "3D Characterisation of Coated Nanoparticles and Soft-Hard Nanocomposites". PhD thesis. University of Antwerp, 2018.

[14] E. J. Kirkland. *Advanced Computing in Electron Microscopy*. Springer Science & Business Media, 2010.

[15] B. Goris et al. "Atomic-Scale Determination of Surface Facets in Gold Nanorods". In: *Nature Materials* 11.11 (Nov. 2012), pp. 930–935. ISSN: 1476-4660. DOI: 10.1038/nmat3462.

[16] E. Arslan Irmak, P. Liu, S. Bals, and S. Van Aert. "3D Atomic Structure of Supported Metallic Nanoparticles Estimated from 2D ADF STEM Images: A Combination of Atom-Counting and a Local Minima Search Algorithm". In: *Small Methods* 5.12 (2021), p. 2101150. ISSN: 2366-9608. DOI: 10.1002/smtd.202101150.

[17] A. De Backer et al. "Experimental Reconstructions of 3D Atomic Structures from Electron Microscopy Images Using a Bayesian Genetic Algorithm". In: *npj Computational Materials* 8.1 (Oct. 12, 2022), pp. 1–8. ISSN: 2057-3960. DOI: 10.1038/s41524-022-00900-w.

[18] S. Amelinckx, D. Van Dyck, J. Van Landuyt, and G. Van Tendeloo. *Electron Microscopy Principles and Fundamentals*. VCH Verlagsgescllschaft mbH, 1997. ISBN: 3-527-29479-1.

[19] D. Wolf et al. "Unveiling the Three-Dimensional Magnetic Texture of Skyrmion Tubes". In: *Nature Nanotechnology* 17 (Mar. 1, 2022). DOI: 10.1038/s41565-021-01031-x.

[20] T. Altantzis et al. "Three-Dimensional Quantification of the Facet Evolution of Pt Nanoparticles in a Variable Gaseous Environment". In: *Nano Letters* 19.1 (Jan. 9, 2019), pp. 477–481. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.8b04303.

[21] A. V. Crewe, J. Wall, and J. Langmore. "Visibility of Single Atoms". In: *Science (New York, N.Y.)* 168.3937 (June 12, 1970), pp. 1338–1340. DOI: 10.1126/science.168.3937.1338.

[22] 2. J. Pennycook. "Z-Contrast STEM for Materials Science". In: *Ultramicroscopy* 30.1-2 (1989), pp. 58–69.

[23] A. Velazco, M. Nord, A. Béché, and J. Verbeeck. "Evaluation of Different Rectangular Scan Strategies for STEM Imaging". In: *Ultramicroscopy* 215 (Aug. 1, 2020), p. 113021. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2020.113021.

[24] D. Stroppa, L. Zagonel, L. Montoro, E. Leite, and A. Ramirez. "High-Resolution Scanning Transmission Electron Microscopy (HRSTEM) Techniques: High-Resolution Imaging and Spectroscopy Side by Side". In: *Chemphyschem : a European journal of chemical physics and physical chemistry* 13 (Feb. 1, 2012), pp. 437–43. DOI: 10.1002/cphc.201100729.

[25] M. Haider et al. "A Spherical-Aberration-Corrected 200kV Transmission Electron Microscope". In: *Ultramicroscopy* 75.1 (Oct. 1, 1998), pp. 53–60. ISSN: 0304-3991. DOI: 10.1016/S0304-3991(98)00048-5.

[26] K. H. Van Den Bos et al. "Unscrambling Mixed Elements Using High Angle Annular Dark Field Scanning Transmission Electron Microscopy". In: *Physical Review Letters* 116.24 (2016), pp. 1–6. ISSN: 10797114. DOI: 10.1103/PhysRevLett.116.246101. pmid: 22654305.

[27] L. Jones and P. D. Nellist. "Identifying and Correcting Scan Noise and Drift in the Scanning Transmission Electron Microscope". In: *Microscopy and Microanalysis* 19.4 (2013), pp. 1050–1060. DOI: 10.1017/S1431927613001402.

[28] L. Jones et al. "Managing Dose-, Damage- and Data-Rates in Multi-Frame Spectrum-Imaging". In: *Microscopy : the journal of the Quekett Microscopi-*

cal Club 67 (suppl_1 Mar. 1, 2018), pp. i98–i113. ISSN: 2050-5698. DOI: 10.1093/jmicro/dfx125.

[29] L. Gambini, T. Mullarkey, L. Jones, and S. Sanvito. "Machine-Learning Approach for Quantified Resolvability Enhancement of Low-Dose STEM Data". In: *Machine Learning: Science and Technology* (2023). ISSN: 2632-2153. DOI: 10.1088/2632-2153/acbb52.

[30] R. Lin, R. Zhang, C. Wang, X.-Q. Yang, and H. L. Xin. "TEMImageNet Training Library and AtomSegNet Deep-Learning Models for High-Precision Atom Segmentation, Localization, Denoising, and Deblurring of Atomic-Resolution Images". In: *Scientific Reports* 11.1 (Mar. 8, 2021), p. 5386. ISSN: 2045-2322. DOI: 10.1038/s41598-021-84499-w.

[31] F. Wang, T. R. Henninen, D. Keller, and R. Erni. "Noise2Atom: Unsupervised Denoising for Scanning Transmission Electron Microscopy Images". In: *Applied Microscopy* 50.1 (Oct. 20, 2020), p. 23. ISSN: 2287-4445. DOI: 10.1186/s42649-020-00041-8.

[32] A. Krull, T.-O. Buchholz, and F. Jug. "Noise2Void - Learning Denoising From Single Noisy Images". In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019, pp. 2129–2137.

[33] D. G. Sentürk, A. De Backer, T. Friedrich, and S. Van Aert. "Optimal Experiment Design for Element Specific Atom Counting Using Multiple Annular Dark Field Scanning Transmission Electron Microscopy Detectors". In: *Ultramicroscopy* 242 (2022), p. 113626.

[34] S. Toyama et al. "Quantitative Electric Field Mapping of a p–n Junction by DPC STEM". In: *Ultramicroscopy* 216 (Sept. 1, 2020), p. 113033. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2020.113033.

[35] L. Clark et al. "The Effect of Dynamical Scattering on Single-plane Phase Retrieval in Electron Ptychography". In: *Microscopy and Microanalysis* (Dec. 20, 2022), ozac022. ISSN: 1431-9276. DOI: 10.1093/micmic/ozac022.

[36] J. Chapman, I. McFadyen, and S. McVitie. "Modified Differential Phase Contrast Lorentz Microscopy for Improved Imaging of Magnetic Structures". In: *IEEE Transactions on Magnetics* 26.5 (Sept. 1990), pp. 1506–1511. ISSN: 1941-0069. DOI: 10.1109/20.104427.

[37] N. Shibata et al. "New Area Detector for Atomic-Resolution Scanning Transmission Electron Microscopy". In: *Journal of Electron Microscopy* 59.6 (Dec. 1, 2010), pp. 473–479. ISSN: 0022-0744. DOI: 10.1093/jmicro/dfq014.

[38] M. Haider, A. Epstein, P. Jarron, and C. Boulin. "A Versatile, Software Configurable Multichannel STEM Detector for Angle-Resolved Imaging". In: *Ultramicroscopy* 54.1 (May 1, 1994), pp. 41–59. ISSN: 0304-3991. DOI: 10.1016/0304-3991(94)90091-4.

[39] I. MacLaren, T. A. MacGregor, C. S. Allen, and A. I. Kirkland. "Detectors-The Ongoing Revolution in Scanning Transmission Electron Microscopy and Why This Important to Material Characterization". In: *APL Materials* 8.11 (Nov. 2020). ISSN: 2166532X. DOI: 10.1063/5.0026992.

[40] R. N. Clough, G. Moldovan, and A. I. Kirkland. "Direct Detectors for Electron Microscopy". In: *Journal of Physics: Conference Series* 522.1 (June 2014), p. 012046. ISSN: 1742-6596. DOI: 10.1088/1742-6596/522/1/012046.

[41] M. W. Tate et al. "High Dynamic Range Pixel Array Detector for Scanning Transmission Electron Microscopy". In: *Microscopy and Microanalysis* 22.1 (Feb. 2016), pp. 237–249. ISSN: 14358115. DOI: 10.1017/S1431927615015664. pmid: 26750260.

[42] R. Ballabriga et al. "Medipix3: A 64 k Pixel Detector Readout Chip Working in Single Photon Counting Mode with Improved Spectrometric Performance". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 633 (SUPPL. 1 May 2011), S15–S18. ISSN: 0168-9002. DOI: 10.1016/J.NIMA.2010.06.108.

[43] J. Ciston et al. "The 4D Camera: Very High Speed Electron Counting for 4D-STEM". In: *Microscopy and Microanalysis* 25.S2 (Aug. 2019), pp. 1930–1931. ISSN: 1431-9276. DOI: 10.1017/s1431927619010389.

[44] T. Poikela et al. "Timepix3: A 65K Channel Hybrid Pixel Readout Chip with Simultaneous ToA/ToT and Sparse Readout". In: *Journal of instrumentation* 9.05 (2014), p. C05013.

[45] V. B. Ozdol et al. "Strain Mapping at Nanometer Resolution Using Advanced Nano-Beam Electron Diffraction". In: *Applied Physics Letters* 106.25 (June 22, 2015), p. 253107. ISSN: 0003-6951. DOI: 10.1063/1.4922994.

[46] T. C. Pekin, C. Gammer, J. Ciston, A. M. Minor, and C. Ophus. "Optimizing Disk Registration Algorithms for Nanobeam Electron Diffraction Strain Mapping". In: *Ultramicroscopy*. 70th Birthday of Robert Sinclair and 65th Birthday of Nestor J. Zaluzec PICO 2017 – Fourth Conference on Frontiers of Aberration Corrected Electron Microscopy 176 (May 1, 2017), pp. 170–176. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2016.12.021.

[47] J. M. Lebeau, S. D. Findlay, L. J. Allen, and S. Stemmer. "Standardless Atom Counting in Scanning Transmission Electron Microscopy". In: *Nano Letters* 10.11 (Nov. 2010), pp. 4405–4408. ISSN: 15306984. DOI: 10.1021/NL102025S. pmid: 20945926.

[48] H. E et al. "Probe Integrated Scattering Cross Sections in the Analysis of Atomic Resolution HAADF STEM Images". In: *Ultramicroscopy* 133 (Oct. 2013), pp. 109–119. ISSN: 03043991. DOI: 10.1016/j.ultramic.2013.07.002. pmid: 23969066.

[49] A. De wael, A. De Backer, L. Jones, P. D. Nellist, and S. Van Aert. "Hybrid Statistics-Simulations Based Method for Atom-Counting from ADF STEM Images". In: *Ultramicroscopy* 177 (June 2017), pp. 69–77. ISSN: 18792723. DOI: 10.1016/j.ultramic.2017.01.010. pmid: 28292688.

[50] C. M. O'Leary et al. "Contrast Transfer and Noise Considerations in Focused-Probe Electron Ptychography". In: *Ultramicroscopy* 221 (Feb. 2021), p. 113189. ISSN: 18792723. DOI: 10.1016/j.ultramic.2020.113189.

[51] A. De wael, A. De Backer, I. Lobato, and S. Van Aert. "Modelling ADF STEM Images Using Elliptical Gaussian Peaks and Its Effects on the Quantification of Structure Parameters in the Presence of Sample Tilt". In: *Ultramicroscopy* 230 (Nov. 2021), p. 113391. ISSN: 03043991. DOI: 10.1016/j.ultramic.2021.113391.

[52] J. M. Ede, J. J. P. Peters, J. Sloan, and R. Beanland. *Exit Wavefunction Reconstruction from Single Transmission Electron Micrographs with Deep Learning*. Jan. 30, 2020. DOI: 10.48550/arXiv.2001.10938. arXiv: 2001.10938. URL: http://arxiv.org/abs/2001.10938 (visited on 12/29/2022). preprint.

[53] N. Laanait et al. "Exascale Deep Learning for Scientific Inverse Problems". In: *arXiv e-prints* (2019). arXiv: 1909.11150.

[54] T. Friedrich, C.-P. Yu, J. Verbeek, T. Pennycook, and S. V. Aert. "Phase Retrieval From 4-Dimensional Electron Diffraction Datasets". In: *2021 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021, pp. 3453–3457. ISBN: 978-1-66544-115-5. DOI: 10.1109/icip42928.2021.9506709.

[55] A. H. Combs et al. "Fast Approximate STEM Image Simulations from a Machine Learning Model". In: *Advanced Structural and Chemical Imaging* 5.1 (Dec. 2019), p. 2. ISSN: 21980926. DOI: 10.1186/s40679-019-0064-2.

[56] A. Winkelmann, C. Trager-Cowan, F. Sweeney, A. P. Day, and P. Parbrook. "Many-Beam Dynamical Simulation of Electron Backscatter Diffraction Patterns". In: *Ultramicroscopy* 107.4-5 (Apr. 2007), pp. 414–421. DOI: 10.1016/j.ultramic.2006.10.006.

[57] A. Winkelmann. "Dynamical Simulation of Electron Backscatter Diffraction Patterns". In: *Electron Backscatter Diffraction in Materials Science*. 2009, pp. 21–33. ISBN: 978-0-387-88135-5. DOI: 10.1007/978-0-387-88136-2_2.

[58] D. V. Dyck. "Fast Computational Procedures for the Simulation of Structure Images in Complex or Disordered Crystals: A New Approach". In: *Journal of Microscopy* 119.1 (1980), pp. 141–152. ISSN: 1365-2818. DOI: 10.1111/j.1365-2818.1980.tb04084.x.

[59] I. Lobato and D. Van Dyck. "Improved Multislice Calculations for Including Higher-Order Laue Zones Effects". In: *Ultramicroscopy*. Special Issue: Gertrude F. Rempfer 100th Birthday Memorial 119 (Aug. 1, 2012), pp. 63–71. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2012.01.003.

[60] I. Lobato, S. van Aert, and J. Verbeeck. "Progress and New Advances in Simulating Electron Microscopy Datasets Using MULTEM". In: *Ultramicroscopy* 168.2016 (2016), pp. 17–27. ISSN: 18792723. DOI: 10.1016/j.ultramic.2016.06.003. arXiv: 0504469[math].

[61] I. Lobato and D. Van Dyck. "MULTEM: A New Multislice Program to Perform Accurate and Fast Electron Diffraction and Imaging Simulations Using Graphics Processing Units with CUDA". In: *Ultramicroscopy* 156.2015 (2015), pp. 9–17. ISSN: 18792723. DOI: 10.1016/j.ultramic.2015.04.016.

[62] L. J. Allen, A. J. D'Alfonso, and S. D. Findlay. "Modelling the Inelastic Scattering of Fast Electrons". In: *Ultramicroscopy*. Special Issue: 80th Birthday of Harald Rose; PICO 2015 – Third Conference on Frontiers of Aberration Corrected Electron Microscopy 151 (Apr. 1, 2015), pp. 11–22. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2014.10.011.

[63] L. Rangel DaCosta et al. "Prismatic 2.0 – Simulation Software for Scanning and High Resolution Transmission Electron Microscopy (STEM and HRTEM)". In: *Micron (Oxford, England : 1993)* 151 (Dec. 1, 2021), p. 103141. ISSN: 0968-4328. DOI: 10.1016/j.micron.2021.103141.

[64] J. Barthel. "Dr. Probe: A Software for High-Resolution STEM Image Simulation". In: *Ultramicroscopy* 193 (Oct. 1, 2018), pp. 1–11. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2018.06.003.

[65] D. Van Dyck. "Is the Frozen Phonon Model Adequate to Describe Inelastic Phonon Scattering?" In: *Ultramicroscopy* 109.6 (May 2009), pp. 677–682. ISSN: 0304-3991. DOI: 10.1016/J.ULTRAMIC.2009.01.001.

[66] X. Chen, D. S. Kim, and J. M. LeBeau. "A Comparison of Molecular Dynamics Potentials Used to Account for Thermal Diffuse Scattering in Multislice Simulations". In: *Ultramicroscopy* 244 (Feb. 1, 2023), p. 113644. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2022.113644.

[67] J. Hochreiter. "Untersuchungen zu dynamischen neuronalen Netzen". Diplomarbeit Informatik. München: Technische Universität München, June 19, 1991.

[68] L. Lu, Y. Shin, Y. Su, and G. E. Karniadakis. "Dying ReLU and Initialization: Theory and Numerical Examples". In: *Communications in Computational Physics* 28.5 (June 2020), pp. 1671–1706. ISSN: 1815-2406, 1991-7120. DOI: 10.4208/cicp.OA-2020-0165. arXiv: 1903.06733 [cs, math, stat].

[69] B. Xu, N. Wang, T. Chen, and M. Li. *Empirical Evaluation of Rectified Activations in Convolutional Network*. Nov. 27, 2015. DOI: 10.48550/arXiv.1505.00853. arXiv: 1505.00853. URL: http://arxiv.org/abs/1505.00853 (visited on 12/30/2022). preprint.

[70] P. Ramachandran, B. Zoph, and Q. V. Le. "Searching for Activation Functions". 2017. arXiv: 1710.05941.

[71] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[72] Y. E. Nesterov. "A Method for Solving the Convex Programming Problem with Convergence Rate $ Obigl (Frac $\{$1$\}$$\{$k\^{}} 2$\}$ Bigr) $". In: *Dokl. Akad. Nauk SSSR,* vol. 269. 1983, pp. 543–547.

[73] N. Qian. "On the Momentum Term in Gradient Descent Learning Algorithms". In: *Neural Networks* 12.1 (Jan. 1, 1999), pp. 145–151. ISSN: 0893-6080. DOI: 10.1016/S0893-6080(98)00116-6.

[74] J. Duchi, E. Hazan, and Y. Singer. "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization." In: *Journal of machine learning research* 12.7 (2011).

[75] M. D. Zeiler. *ADADELTA: An Adaptive Learning Rate Method*. Dec. 22, 2012. DOI: 10.48550/arXiv.1212.5701. arXiv: 1212.5701. URL: http://arxiv.org/abs/1212.5701 (visited on 01/02/2023). preprint.

[76] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. Jan. 29, 2017. DOI: 10.48550/arXiv.1412.6980. arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980 (visited on 12/30/2022). preprint.

[77] S. Ruder. *An Overview of Gradient Descent Optimization Algorithms*. June 15, 2017. DOI: 10.48550/arXiv.1609.04747. arXiv: 1609.04747. URL: http://arxiv.org/abs/1609.04747 (visited on 12/30/2022). preprint.

[78] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (Dec. 1989), pp. 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541.

[79] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-Based Learning Applied to Document Recognition". In: *Proceedings of the IEEE* 86.11 (Nov. 1998), pp. 2278–2324. ISSN: 1558-2256. DOI: 10.1109/5.726791.

[80] A. LeNail. "NN-SVG: Publication-Ready Neural Network Architecture Schematics". In: *Journal of Open Source Software* 4.33 (Jan. 15, 2019), p. 747. ISSN: 2475-9066. DOI: 10.21105/joss.00747.

[81] F. Chollet. *Deep Learning with Python, Second Edition*. Simon and Schuster, Dec. 7, 2021. 502 pp. ISBN: 978-1-63835-009-5. Google Books: mjVKEAAAQBAJ.

[82] V. Dumoulin and F. Visin. "A Guide to Convolution Arithmetic for Deep Learning". In: *ArXiv e-prints* (Mar. 2016).

[83] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016, pp. 770–778.

[84] G. Huang, Z. Liu, L. vander Maaten, and K. Q. Weinberger. "Densely Connected Convolutional Networks". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017, pp. 4700–4708.

[85] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. "Residual Dense Network for Image Super-Resolution". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018, pp. 2472–2481.

[86] B. Park, S. Yu, and J. Jeong. "Densely Connected Hierarchical Network for Image Denoising". In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Vol. 2019-June. IEEE Computer Society, June 1, 2019, pp. 2104–2113. ISBN: 978-1-72812-506-0. DOI: 10.1109/CVPRW.2019.00263.

[87] S. Ioffe and C. Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *Proceedings of the 32nd International Conference on Machine Learning*. International Conference on Machine Learning. PMLR, June 1, 2015, pp. 448–456.

[88] Y. Wu and K. He. "Group Normalization". In: Proceedings of the European Conference on Computer Vision (ECCV). 2018, pp. 3–19.

[89] A. De Backer, K. vanden Bos, W. Van den Broek, J. Sijbers, and S. Van Aert. "StatSTEM: An Efficient Approach for Accurate and Precise Model-Based Quantification of Atomic Resolution Electron Microscopy Images". In: *Ultramicroscopy* 171.2016 (2016), pp. 104–116. ISSN: 18792723. DOI: 10.1016/j.ultramic.2016.08.018.

[90] S. Van Aert et al. "Procedure to Count Atoms with Trustworthy Single-Atom Sensitivity". In: *Physical Review B - Condensed Matter and Materials Physics* 87.6 (2013), pp. 1–6. ISSN: 10980121. DOI: 10.1103/PhysRevB.87.064107.

[91] J. Fatermans et al. "Single Atom Detection from Low Contrast-to-Noise Ratio Electron Microscopy Images". In: *Physical Review Letters* 121.5 (2018), p. 56101. ISSN: 14745488. DOI: 10.1016/S1470-2045(15)00369-1. pmid: 26522332.

[92] B. Bajic, J. Lindblad, and N. Sladoje. "Blind Restoration of Images Degraded with Mixed Poisson-Gaussian Noise with Application in Transmission Electron Microscopy". In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, Apr. 2016, pp. 123–127. ISBN: 978-1-4799-2349-6. DOI: 10.1109/ISBI.2016.7493226.

[93] N. Braidy, Y. Le Bouar, S. Lazar, and C. Ricolleau. "Correcting Scanning Instabilities from Images of Periodic Structures". In: *Ultramicroscopy* 118 (July 2012), pp. 67–76. ISSN: 03043991. DOI: 10.1016/j.ultramic.2012.04.001.

[94] N. Nakanishi et al. "Retrieval Process of High-Resolution HAADF-STEM Images". In: *Journal of Electron Microscopy* 51.6 (Nov. 2002), pp. 383–390. ISSN: 00220744. DOI: 10.1093/jmicro/51.6.383.

[95] A. M. Sanchez et al. "An Approach to the Systematic Distortion Correction in Aberration-Corrected HAADF Images". In: *Journal of Microscopy* 221.1 (Jan. 2006), pp. 1–7. ISSN: 00222720. DOI: 10.1111/j.1365-2818.2006.01533.x.

[96] Y. Wang et al. "Correcting the Linear and Nonlinear Distortions for Atomically Resolved STEM Spectrum and Diffraction Imaging". In: *Microscopy : the journal of the Quekett Microscopical Club* 67 (suppl_1 Mar. 1, 2018), pp. i114–i122. ISSN: 2050-5698. DOI: 10.1093/jmicro/dfy002.

[97] S. Ning et al. "Scanning Distortion Correction in STEM Images". In: *Ultramicroscopy* 184 (Jan. 1, 2018), pp. 274–283. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2017.09.003.

[98] C. M. O'Leary, B. Haas, C. T. Koch, P. D. Nellist, and L. Jones. "Increasing Spatial Fidelity and SNR of 4D-STEM Using Multi-Frame Data Fusion". In: *Microscopy and Microanalysis* 28.4 (Aug. 2022), pp. 1417–1427. ISSN: 1431-9276, 1435-8115. DOI: 10.1017/S1431927621012587.

[99] S.-I. Amari. "Learning Patterns and Pattern Sequences by Self-Organizing Nets of Threshold Elements". In: *IEEE Transactions on Computers* C-21.11 (1972), pp. 1197–1206. ISSN: 00189340. DOI: 10.1109/T-C.1972.223477.

[100] Y.-T. Zhou, R. Chellappa, A. Vaid, and B. Jenkins. "Image Restoration Using a Neural Network". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.7 (July 1988), pp. 1141–1151. ISSN: 00963518. DOI: 10.1109/29.1641.

[101] N. R. Pal and S. K. Pal. "A Review on Image Segmentation Techniques". In: *Pattern Recognition* 26.9 (Sept. 1993), pp. 1277–1294. ISSN: 0031-3203. DOI: 10.1016/0031-3203(93)90135-J.

[102] M. Egmont-Petersen, D. de Ridder, and H. Handels. "Image Processing with Neural Networks—a Review". In: *Pattern Recognition* 35.10 (Oct. 2002), pp. 2279–2301. ISSN: 0031-3203. DOI: 10.1016/S0031-3203(01)00178-9.

[103] O. L. Krivanek et al. "An Electron Microscope for the Aberration-Corrected Era". In: *Ultramicroscopy* 108.3 (Feb. 2008), pp. 179–195. ISSN: 03043991. DOI: 10.1016/j.ultramic.2007.07.010.

[104] L. Reimer and H. Kohl. *Transmission Electron Microscopy: Physics of Image Formation (Google eBook)*. Springer Science & Business, 2008, p. 587. ISBN: 0-387-40093-1. DOI: 10.1007/978-0-387-34758-5.

[105] R. Ishikawa, A. R. Lupini, S. D. Findlay, and S. J. Pennycook. "Quantitative Annular Dark Field Electron Microscopy Using Single Electron Signals". In: *Microscopy and Microanalysis* 20.1 (2014), pp. 99–110. DOI: 10.1017/S1431927613013664.

[106] Burle Industries. *Photomultiplier Handbook*. Burle Technologies Inc, 1980.

[107] J. R. Prescott. "A Statistical Model for Photomultiplier Single-Electron Statistics". In: *Nuclear Instruments and Methods* 39.1 (Jan. 1966), pp. 173–179. ISSN: 0029554X. DOI: 10.1016/0029-554X(66)90059-0.

[108] A. Mittelberger, C. Kramberger, and J. C. Meyer. "Software Electron Counting for Low-Dose Scanning Transmission Electron Microscopy". In: *Ultramicroscopy*

188 (May 1, 2018), pp. 1–7. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2018.02.005.

[109] G. Walker. "On Periodicity in Series of Related Terms". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 131 (1931), pp. 518–532.

[110] J. P. Buban, Q. Ramasse, B. Gipson, N. D. Browning, and H. Stahlberg. "High-Resolution Low-Dose Scanning Transmission Electron Microscopy". In: *Journal of Electron Microscopy* 59.2 (2010), pp. 103–112. ISSN: 00220744. DOI: 10.1093/jmicro/dfp052.

[111] S. Van Aert et al. "Quantitative Atomic Resolution Mapping Using High-Angle Annular Dark Field Scanning Transmission Electron Microscopy". In: *Ultramicroscopy* 109.10 (Sept. 2009), pp. 1236–1244. ISSN: 0304-3991. DOI: 10.1016/J.ULTRAMIC.2009.05.010.

[112] G. T. Martinez, A. De Backer, A. Rosenauer, J. Verbeeck, and S. Van Aert. "The Effect of Probe Inaccuracies on the Quantitative Model-Based Analysis of High Angle Annular Dark Field Scanning Transmission Electron Microscopy Images". In: *Micron (Oxford, England : 1993)* 63.2014 (2014), pp. 57–63. ISSN: 09684328. DOI: 10.1016/j.micron.2013.12.009.

[113] I. Lobato and D. Van Dyck. "An Accurate Parameterization for Scattering Factors, Electron Densities and Electrostatic Potentials for Neutral Atoms That Obey All Physical Constraints". In: *Acta Crystallographica Section A: Foundations and Advances* 70.6 (2014), pp. 636–649.

[114] A. De Backer, G. T. Martinez, A. Rosenauer, and S. Van Aert. "Atom Counting in HAADF STEM Using a Statistical Model-Based Approach: Methodology, Possibilities, and Inherent Limitations". In: *Ultramicroscopy*. 65th Birthdays of W Owen Saxton, David J Smith and Dirk Van Dyck / PICO 2013 & From Multislice to Big Bang 134 (Nov. 1, 2013), pp. 23–33. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2013.05.003.

[115] A. van den Bos. *Parameter Estimation for Scientists and Engineers*. John Wiley & Sons, Aug. 3, 2007. 291 pp. ISBN: 978-0-470-17385-5. Google Books: lXVWB_Y8GQoC.

[116] D. Hull and D. J. Bacon. "Introduction to Dislocations". In: *Introduction to Dislocations*. Butterworth-Heinemann, 2011. ISBN: 978-0-08-096672-4. DOI: 10.1016/B978-0-08-096672-4.00001-3.

[117] P. Hirel. "Atomsk: A Tool for Manipulating and Converting Atomic Data Files". In: *Computer Physics Communications* 197 (Dec. 2015), pp. 212–219. ISSN: 00104655. DOI: 10.1016/j.cpc.2015.07.012.

[118] H. Zhao, O. Gallo, I. Frosio, and J. Kautz. "Loss Functions for Image Restoration With Neural Networks". In: *IEEE Transactions on Computational Imaging* 3.1 (Mar. 2017), pp. 47–57. ISSN: 2333-9403. DOI: 10.1109/TCI.2016.2644865.

[119] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. "Image-To-Image Translation With Conditional Adversarial Networks". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017, pp. 1125–1134.

[120] J. P. Buban, Q. Ramasse, B. Gipson, N. D. Browning, and H. Stahlberg. "High-Resolution Low-Dose Scanning Transmission Electron Microscopy". In: *Journal*

of Electron Microscopy 59.2 (2010), pp. 103–112. ISSN: 00220744. DOI: 10.1093/jmicro/dfp052.

[121] M. Mirza and S. Osindero. *Conditional Generative Adversarial Nets*. arXiv.org. Nov. 6, 2014. URL: https://arxiv.org/abs/1411.1784v1 (visited on 03/27/2023).

[122] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. "Residual Dense Network for Image Restoration". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.7 (July 2021), pp. 2480–2495. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2020.2968521.

[123] D.-W. Kim, J. R. Chung, and S.-W. Jung. *GRDN:Grouped Residual Dense Network for Real Image Denoising and GAN-based Real-world Noise Modeling*. arXiv.org. May 27, 2019. URL: https://arxiv.org/abs/1905.11172v1 (visited on 03/27/2023).

[124] A. Ignatov et al. "NTIRE 2019 Challenge on Image Enhancement: Methods and Results". In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Vol. 2019-June. 2019, pp. 2224–2232. ISBN: 978-1-72812-506-0. DOI: 10.1109/CVPRW.2019.00275.

[125] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon. "CBAM: Convolutional Block Attention Module". In: Proceedings of the European Conference on Computer Vision (ECCV). 2018, pp. 3–19.

[126] T. Plötz and S. Roth. "Neural Nearest Neighbors Networks". In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. NIPS'18. Red Hook, NY, USA: Curran Associates Inc., Dec. 3, 2018, pp. 1095–1106.

[127] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7 (July 2017), pp. 3142–3155. ISSN: 1941-0042. DOI: 10.1109/TIP.2017.2662206.

[128] Z. Li, Y. Fan, and W. Liu. "The Effect of Whitening Transformation on Pooling Operations in Convolutional Autoencoders". In: *EURASIP Journal on Advances in Signal Processing* 2015.1 (Apr. 14, 2015), p. 37. ISSN: 1687-6180. DOI: 10.1186/s13634-015-0222-1.

[129] A. Jolicoeur-Martineau. *The Relativistic Discriminator: A Key Element Missing from Standard GAN*. arXiv.org. July 2, 2018. URL: https://arxiv.org/abs/1807.00734v3 (visited on 03/27/2023).

[130] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun. "What Is the Best Multi-Stage Architecture for Object Recognition?" In: *2009 IEEE 12th International Conference on Computer Vision*. 2009 IEEE 12th International Conference on Computer Vision. Sept. 2009, pp. 2146–2153. DOI: 10.1109/ICCV.2009.5459469.

[131] S. Van Aert, K. J. Batenburg, M. D. Rossell, R. Erni, and G. Van Tendeloo. "Three-Dimensional Atomic Imaging of Crystalline Nanoparticles". In: *Nature* 470.7334 (Feb. 2011), pp. 374–377. ISSN: 1476-4687. DOI: 10.1038/nature09741.

[132] A. Beyer, R. Straubinger, J. Belz, and K. Volz. "Local sample thickness determination via scanning transmission electron microscopy defocus series". In: *Journal of Microscopy* 262.2 (2016), pp. 171–177. ISSN: 1365-2818. DOI: 10.1111/jmi.12284.

[133] K. Müller et al. "Scanning Transmission Electron Microscopy Strain Measurement from Millisecond Frames of a Direct Electron Charge Coupled Device". In: *Applied Physics Letters* 101.21 (2012), p. 212110.
[134] R. Plackett et al. "Merlin: A Fast Versatile Readout System for Medipix3". In: *Journal of Instrumentation* 8.01 (2013), p. C01038.
[135] P. D. Nellist, B. C. McCallum, and J. M. Rodenburg. "Resolution beyond the'information Limit'in Transmission Electron Microscopy". In: *Nature* 374.6523 (1995), pp. 630–632.
[136] J. M. Rodenburg and H. M. Faulkner. "A Phase Retrieval Algorithm for Shifting Illumination". In: *Applied physics letters* 85.20 (2004), pp. 4795–4797.
[137] A. M. Maiden and J. M. Rodenburg. "An Improved Ptychographical Phase Retrieval Algorithm for Diffractive Imaging". In: *Ultramicroscopy* 109.10 (2009), pp. 1256–1262. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2009.05.012.
[138] M. Odstrčil, A. Menzel, and M. Guizar-Sicairos. "Iterative Least-Squares Solver for Generalized Maximum-Likelihood Ptychography". In: *Optics express* 26.3 (2018), pp. 3108–3123.
[139] Z. Chen et al. "Mixed-State Electron Ptychography Enables Sub-Angstrom Resolution Imaging with Picometer Precision at Low Dose". In: *Nature Communications* 11.1 (Dec. 2020). ISSN: 20411723. DOI: 10.1038/S41467-020-16688-6. pmid: 32533001.
[140] Z. Chen et al. "Electron Ptychography Achieves Atomic-Resolution Limits Set by Lattice Vibrations". 2021. arXiv: 2101.00465.
[141] T. J. Pennycook et al. "Efficient Phase Contrast Imaging in STEM Using a Pixelated Detector. Part 1: Experimental Demonstration at Atomic Resolution". In: *Ultramicroscopy* 151 (2015), pp. 160–167.
[142] 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". In: *Ultramicroscopy* 151 (2015), pp. 232–239.
[143] K. Müller et al. "Atomic Electric Fields Revealed by a Quantum Mechanical Approach to Electron Picodiffraction". In: *Nature communications* 5.1 (2014), pp. 1–8.
[144] I. Lazić, E. G. Bosch, and S. Lazar. "Phase Contrast STEM for Thin Samples: Integrated Differential Phase Contrast". In: *Ultramicroscopy* 160 (2016), pp. 265–280.
[145] E. Yücelen, I. Lazić, and E. G. Bosch. "Phase Contrast Scanning Transmission Electron Microscopy Imaging of Light and Heavy Atoms at the Limit of Contrast and Resolution". In: *Scientific reports* 8.1 (2018), pp. 1–10.
[146] P. M. Pelz, I. Johnson, C. Ophus, P. Ercius, and M. C. Scott. "Real-Time Interactive 4D-STEM Phase-Contrast Imaging from Electron Event Representation Data: Less Computation with the Right Representation". In: *IEEE Signal Processing Magazine* 39.1 (2022), pp. 25–31. DOI: 10.1109/MSP.2021.3120981.
[147] B. Haas et al. "High-Fidelity 4D-STEM Enabled by Live Processing at 15'000 Detector Frames per Second". In: *Microscopy and Microanalysis* 27.S1 (2021), pp. 994–997.

[148] A. Strauch et al. "Live Processing of Momentum-Resolved STEM Data for First Moment Imaging and Ptychography". In: *Microscopy and Microanalysis* (2021), pp. 1–15.

[149] D. Jannis et al. "Event Driven 4D STEM Acquisition with a Timepix3 Detector: Microsecond Dwell Time and Faster Scans for High Precision and Low Dose Applications". In: *Ultramicroscopy* 233 (2022), p. 113423.

[150] H. Guo et al. "Electron-Event Representation Data Enable Efficient cryoEM File Storage with Full Preservation of Spatial and Temporal Resolution". In: *IUCrJ* 7.5 (2020).

[151] C.-P. Yu et al. *Real Time Integration Center of Mass (riCOM) Reconstruction for 4D-STEM*. Dec. 2021. DOI: 10.5281/zenodo.5572123.

[152] D. Liu et al. "Possible Misidentification of Heteroatom Species in Scanning Transmission Electron Microscopy Imaging of Zeolites". In: *The Journal of Physical Chemistry C* 125.34 (Sept. 2021), pp. 18952–18960. DOI: 10.1021/ACS.JPCC.1C05516.

[153] L. Liu et al. "Direct Imaging of Atomically Dispersed Molybdenum That Enables Location of Aluminum in the Framework of Zeolite ZSM-5". In: *Angewandte Chemie International Edition* 59.2 (2020), pp. 819–825.

[154] R. Close, Z. Chen, N. Shibata, and S. D. Findlay. "Towards Quantitative, Atomic-Resolution Reconstruction of the Electrostatic Potential via Differential Phase Contrast Using Electrons". In: *Ultramicroscopy* 159 (2015), pp. 124–137. ISSN: 18792723. DOI: 10.1016/j.ultramic.2015.09.002. pmid: 26381331.

[155] J. M. Rodenburg and R. Bates. "The Theory of Super-Resolution Electron Microscopy via Wigner-distribution Deconvolution". In: *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* 339.1655 (June 1992), pp. 521–553. ISSN: 0962-8428. DOI: 10.1098/rsta.1992.0050.

[156] Y. Jiang et al. "Electron Ptychography of 2D Materials to Deep Sub-Ångström Resolution". In: *Nature 2018 559:7714* 559.7714 (July 2018), pp. 343–349. ISSN: 1476-4687. DOI: 10.1038/s41586-018-0298-5. pmid: 30022131.

[157] J. Rodenburg, B. McCallum, and P. Nellist. "Experimental Tests on Double-Resolution Coherent Imaging via STEM". In: *Ultramicroscopy* 48.3 (1993), pp. 304–314. ISSN: 0304-3991. DOI: 10.1016/0304-3991(93)90105-7.

[158] C.-P. Yu, T. Friedrich, D. Jannis, S. Van Aert, and J. Verbeeck. "Real-Time Integration Center of Mass (riCOM) Reconstruction for 4D STEM". In: *Microscopy and Microanalysis* (2021), pp. 1–12.

[159] J. M. Zuo and J. C. Spence. "Advanced Transmission Electron Microscopy: Imaging and Diffraction in Nanoscience". In: *Advanced Transmission Electron Microscopy: Imaging and Diffraction in Nanoscience*. Springer, 2016, pp. 1–729. ISBN: 978-1-4939-6607-3. DOI: 10.1007/978-1-4939-6607-3.

[160] A. Jain et al. *Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation*. 2013. DOI: 10.1063/1.4812323.

[161] T. Friedrich, C.-P. Yu, J. Verbeeck, and S. Van Aert. *Phase Object Reconstruction for 4D-STEM Using Deep Learning, (4D-STEM Training Data)*. Aug. 2022. DOI: 10.5281/zenodo.6971200.

[162] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[163] L. Wu, P. Juhas, S. Yoo, and I. Robinson. "Complex Imaging of Phase Domains by Deep Neural Networks". In: *IUCrJ* 8.1 (Jan. 2021), pp. 12–21. ISSN: 20522525. DOI: 10.1107/S2052252520013780.

[164] C. Trabelsi et al. "Deep Complex Networks". In: *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*. 2018. arXiv: 1705.09792.

[165] P. Virtue, S. X. Yu, and M. Lustig. "Better than Real: Complex-valued Neural Nets for MRI Fingerprinting". In: *Proceedings - International Conference on Image Processing, ICIP*. Vol. 2017-Septe. 2018, pp. 3953–3957. ISBN: 978-1-5090-2175-8. DOI: 10.1109/ICIP.2017.8297024. arXiv: 1707.00070.

[166] J. Munshi et al. "Disentangling Multiple Scattering with Deep Learning: Application to Strain Mapping from Electron Diffraction Patterns". 2022. arXiv: 2202.00204.

[167] D. Jannis et al. *Event Driven 4D STEM Acquisition with a Timepix3 Detector: Microsecond Dwelltime and Faster Scans for High Precision and Low Dose Applications*. July 2021. DOI: 10.5281/zenodo.5068510.

[168] S. J. Pennycook and L. A. Boatner. "Chemically Sensitive Structure-Imaging with a Scanning Transmission Electron Microscope". In: *Nature* 336.6199 (1988), pp. 565–567.

[169] H. Yang et al. "Simultaneous Atomic-Resolution Electron Ptychography and Z-contrast Imaging of Light and Heavy Elements in Complex Nanostructures". In: *Nature Communications* 7.1 (2016), pp. 1–8.

[170] J. G. Lozano, G. T. Martinez, L. Jin, P. D. Nellist, and P. G. Bruce. "Low-Dose Aberration-Free Imaging of Li-rich Cathode Materials at Various States of Charge Using Electron Ptychography". In: *Nano letters* 18.11 (2018), pp. 6850–6855.

[171] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering". In: *IEEE Transactions on Image Processing* 16.8 (Aug. 2007), pp. 2080–2095. ISSN: 1941-0042. DOI: 10.1109/TIP.2007.901238.

[172] M. Makitalo and A. Foi. "Optimal Inversion of the Anscombe Transformation in Low-Count Poisson Image Denoising". In: *IEEE Transactions on Image Processing* 20.1 (Jan. 2011), pp. 99–109. ISSN: 1057-7149. DOI: 10.1109/TIP.2010.2056693.

[173] P. Cayado et al. "Critical Current Density Improvement in CSD-grown High-Entropy $REBa_{2<\sub>Cu_{3<\sub>O_{7-\delta}}}$ films". In: *RSC Advances* 12.44 (2022), pp. 28831–28842. DOI: 10.1039/D2RA03807B.

[174] L. Grünewald. "Electron Microscopic Investigation of Superconducting Fe- and Cu-based Thin Films". Laboratorium für Elektronenmikroskopie (LEM), Karlsruher Institut für Technologie (KIT), 2022. DOI: 10.5445/IR/1000149357.

[175] H. Bhatia et al. "Deciphering the Role of Water in Promoting the Optoelectronic Performance of Surface-Engineered Lead Halide Perovskite Nanocrystals". In:

ACS Applied Materials & Interfaces 15.5 (Feb. 8, 2023), pp. 7294–7307. ISSN: 1944-8244. DOI: 10.1021/acsami.2c20605.

[176] C. Otero-Martínez et al. "Fast A-Site Cation Cross-Exchange at Room Temperature: Single-to Double- and Triple-Cation Halide Perovskite Nanocrystals". In: *Angewandte Chemie* 134.34 (2022), e202205617. ISSN: 1521-3757. DOI: 10.1002/ange.202205617.

[177] B. D. Esser and J. Etheridge. "Complementary ADF-STEM: A Flexible Approach to Quantitative 4D-STEM". In: *Ultramicroscopy* 243 (Jan. 1, 2023), p. 113627. ISSN: 0304-3991. DOI: 10.1016/j.ultramic.2022.113627.

[178] A. Bangun, P. F. Baumeister, A. Clausen, D. Weber, and R. E. Dunin-Borkowski. "Wigner Distribution Deconvolution Adaptation for Live Ptychography Reconstruction". In: *Microscopy and Microanalysis* (Mar. 8, 2023), ozad021. ISSN: 1431-9276. DOI: 10.1093/micmic/ozad021.

Author's Contributions

Peer-reviewed articles

[1] I. Lobato, T. Friedrich, and S. Van Aert. *Deep Convolutional Neural Networks to Restore Single-Shot Electron Microscopy Images*. Version 1. Mar. 29, 2023. DOI: 10.48550/arXiv.2303.17025. arXiv: 2303.17025. URL: http://arxiv.org/abs/2303.17025. preprint.

[2] T. Friedrich, C.-P. Yu, J. Verbeeck, and S. Van Aert. "Phase Object Reconstruction for 4D-STEM Using Deep Learning". In: *Microscopy and Microanalysis* (Jan. 2023), ozac002. ISSN: 1431-9276. DOI: 10.1093/micmic/ozac002.

[3] J. Dinger, T. Friedrich, T. Reimann, and J. Töpfer. "$NiMn_2O_4$ Revisited: Temperature Dependent Cation Distribution from in Situ Neutron Diffraction and Thermopower Studies". In: *Journal of the American Ceramic Society* 106.3 (2023), pp. 1834–1847.

[4] D. G. Sentürk, A. De Backer, T. Friedrich, and S. Van Aert. "Optimal Experiment Design for Element Specific Atom Counting Using Multiple Annular Dark Field Scanning Transmission Electron Microscopy Detectors". In: *Ultramicroscopy* 242 (2022), p. 113626.

[5] C.-P. Yu, T. Friedrich, D. Jannis, S. Van Aert, and J. Verbeeck. "Real-Time Integration Center of Mass (riCOM) Reconstruction for 4D STEM". In: *Microscopy and Microanalysis* (2021), pp. 1–12.

[6] T. Friedrich, C.-P. Yu, J. Verbeek, T. Pennycook, and S. V. Aert. "Phase Retrieval From 4-Dimensional Electron Diffraction Datasets". In: *2021 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021, pp. 3453–3457. ISBN: 978-1-66544-115-5. DOI: 10.1109/icip42928.2021.9506709.

[7] V. G. Efremenko et al. "Two-Body Abrasion Resistance of High-Carbon High-Silicon Steel: Metastable Austenite vs Nanostructured Bainite". In: *Wear : an international journal on the science and technology of friction lubrication and wear* 418–419 (Jan. 2019), pp. 24–35. ISSN: 00431648. DOI: 10.1016/j.wear.2018.11.003.

[8] T. Friedrich, A. Bochmann, J. Dinger, and S. Teichert. "Application of the Pattern Matching Approach for EBSD Calibration and Orientation Mapping, Utilising Dynamical EBSP Simulations". In: *Ultramicroscopy* 184 (Jan. 2018), pp. 44–51. ISSN: 03043991. DOI: 10.1016/j.ultramic.2017.10.006.

- [1] is submitted and currently under review at the journal npj Computational Materials
- For [2] and [5] the first two authors contributed equally.
- [8], [7] and [3] are not directly linked to the thesis topic

Open source contributions

1. **riCOM_cpp** (riCOM (real time centre of mass) application for 4D STEM)
   Language: C++
   Role: Creator & Owner
   License: GNU GPL3
   URL (Docs): https://thfriedrich.github.io/riCOM_cpp/
   URL (Code): https://github.com/ThFriedrich/riCOM_cpp

2. **AIRPI** (AI-assisted rapid phase imaging for 4D-STEM)
   Language: Python
   Role: Creator & Owner
   License: GNU GPL3
   URL: https://github.com/ThFriedrich/airpi

3. **lammps_vscode** (VSCODE extension for language support of LAMMPS scripts)
   Language: TypeScript, JavaScript, Python
   Role: Creator & Owner
   License: GNU GPL2
   URL (Page): https://thfriedrich.github.io/lammps_vscode/
   URL (Code): https://github.com/ThFriedrich/lammps_vscode

4. **atomic_specimen_creation** (Set of crystallographic tools for bulk atomic specimen generation in any zone axis from cif-files)
   Language: Matlab
   Role: Creator & Owner
   License: GNU GPL3
   URL: https://github.com/ThFriedrich/atomic_specimen_creation

5. **ap_data_generation** (Data generation code for AIRPI)
   Language: Matlab
   Role: Creator & Owner
   License: GNU GPL3
   URL: https://github.com/ThFriedrich/ap_data_generation

6. **rmsd_parameterization** (Matlab implementation of Cao, H. X., & Peng, L. M. (1999). Parameterization of the temperature dependence of the Debye-Waller factors)
   Language: Matlab
   Role: Creator & Owner
   License: GNU GPL3
   URL: https://github.com/ThFriedrich/rmsd_parameterization

7. **fft2d_cpp** (Simple fftw wrapper for 2d Fourier transforms in image processing)
   Language: C++
   Role: Creator & Owner
   License: GNU GPL3
   URL: https://github.com/ThFriedrich/fft2d_cpp

8. **StatSTEM** (Quantification & Atom counting toolbox for STEM images)
   Language: Matlab
   Role: Maintainer & Contributor
   License: GNU GPL3
   URL: https://github.com/quantitativeTEM/StatSTEM

9. **MULTEM** (Multislice simulation software for HRTEM, STEM, ISTEM, ED, PED, CBED, ADF-TEM, ABF-HC, EFTEM, and EELS.)
   Language: C++, CUDA, Matlab
   Role: Contributor
   License: GNU GPL3
   URL: https://github.com/Ivanlh20/multem

10. **tinycolormap** (A header-only, single-file library for colormaps written in C++11)
    Language: C++
    Role: Contributor
    License: MIT
    URL: https://github.com/yuki-koyama/tinycolormap

11. **lammps** (Public development project of the LAMMPS Molecular dynamics software package)
    Language: C++
    Role: Contributor
    License: GNU GPL2
    URL: https://github.com/lammps/lammps

12. **merlin_interface** (Python library for interfacing with a Medipix3 detector through the Merlin software through the TCP/IP command interface.)
    Language: Python
    Role: Contributor
    License: GNU LGPL3
    URL: https://gitlab.com/fast_pixelated_detectors/merlin_interface

13. **Phase Object Reconstruction for 4D-STEM using Deep Learning, (4D-STEM Training Data)** (Zenodo open data repository)
    Role: Creator
    License: Creative Commons Attribution 4.0 International Public License
    URL: https://doi.org/10.5281/zenodo.6971200

14. **Phase Object Reconstruction for 4D-STEM using Deep Learning, (4D-STEM Example Data)** (Zenodo open data repository)
    Role: Creator
    License: Creative Commons Attribution 4.0 International Public License
    URL: https://doi.org/10.5281/zenodo.7034879

Conference talks, papers & poster presentations

1. **Sino-European Early Career Researchers Workshop on Emerging Techniques and Applications in Electron Microscopy (SEEM) 2022**
   Contributions: Invited talk
   At: Online
   Topic: Phase Object Reconstruction of 4D-STEM datasets using Deep Learning

2. **PICO 2022**
   Contributions: Poster, Conference paper
   At: Vaalsbroek, Netherlands
   Topic: Phase Object Reconstruction of 4D-STEM datasets using Deep Learning

3. **IEEE International Conference on Image Processing 2021**
   Contributions: Presentation, Poster, Conference paper
   At: Anchorage, USA (Online)
   Topic: Phase Retrieval From 4-Dimensional Electron Diffraction Datasets
   URL: https://doi.org/10.1109/ICIP42928.2021.9506709

4. **Microscience Microscopy Congress 2021 incorporating EMAG 2021**
   Contributions: Presentation, Conference paper
   At: Online
   Topic: Phase Object Reconstruction of 4D-STEM datasets using Deep Learning
   URL: http://dx.doi.org/10.22443/rms.mmc2021.200
   Award: Poster Competition 2nd Prize
   Electron Microscopy & Engineering and Physical Sciences

5. **European Microscopy Congress 2020**
   Contributions: Presentation, Conference paper
   At: Online
   Topic: STEM image restoration using convolutional neural networks and its impact on the quantification of atomic structures
   URL: http://dx.doi.org/10.22443/rms.emc2020.1249

6. **ATOM 2020**
   Contributions: Presentation
   At: Online
   Topic: STEM image restoration using convolutional neural networks and its impact on the quantification of atomic structures

7. **Microscopy and Microanalysis 2020**
   Contributions: Conference paper (Co-Author)
   At: Milwaukee, USA (Online)
   Topic: 3D Atomic Scale Quantification of Nanostructures and their Dynamics Using Model-based STEM
   URL: http://dx.doi.org/10.1017/S1431927620022163

Other publications

1. **EP4075126A1** (PHASE RETRIEVAL IN ELECTRON MICROSCOPY)
   Type: Patent
   Role: Inventor (40%)
   URL: https://worldwide.espacenet.com

2. **Can AI do your Ptychography?**
   Authors: T. Friedrich, C.-P. Yu, J. Verbeeck and S. van Aert
   Type: Magazine article (invited contribution, no peer review)
   Journal: Imaging & Microscopy
   Status: Accepted, publication pending