Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 2400

Deep learning assisted phase retrieval and computational methods in coherent diffractive imaging

ALFREDO BELLISARIO

ACTA UNIVERSITATIS
UPSALIENSIS
2024

ISSN 1651-6214
ISBN 978-91-513-2120-2
urn:nbn:se:uu:diva-526527

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in A1:107a, BMC, Biomedical Centre, Husargatan 3, Uppsala, Wednesday, 5 June 2024 at 09:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Dr. Anton Barty (Deutsches Elektronen-Synchrotron (DESY) Notkestrasse 85, 22607 Hamburg, Germany).

Abstract
Bellisario, A. 2024. Deep learning assisted phase retrieval and computational methods in coherent diffractive imaging. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2400. 94 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-2120-2.

In recent years, advances in Artificial Intelligence and experimental techniques have revolutionized the field of structural biology. X-ray crystallography and Cryo-EM have provided unprecedented insights into the structures of biomolecules, while the unexpected success of AlphaFold has opened up new avenues of investigation. However, studying the dynamics of proteins at high resolution remains a significant obstacle, especially for fast dynamics. Single-particle imaging (SPI) or Flash X-ray Imaging (FXI) is an emerging technique that may enable the mapping of the conformational landscape of biological molecules at high resolution and fast time scale. This thesis discusses the potential of SPI/FXI, its challenges, recent experimental successes, and the advancements driving its development. In particular, machine learning and neural networks could play a vital role in fostering data analysis and improving SPI/FXI data processing. In Paper I, we discuss the problem of noise and detector masks in collecting FXI data. I simulated a dataset of diffraction patterns and used it to train a Convolutional Neural Network (U-Net) to restore data by denoising and filling in detector masks. As a natural continuation of this work, I trained another machine learning model in Paper II to estimate 2D protein densities from diffraction intensities. In the final chapter, corresponding to Paper III, we discuss another experimental method, time-resolved Small Angle X-ray Scattering (SAXS), and a new algorithm recently developed for SAXS data, the DENsity from Solution Scattering (DENSS) algorithm. I discuss the potential of DENSS in time-resolved SAXS and its application for structural fitting of AsLOV2, a Light-Oxygen-Voltage (LOV) protein domain from *Avena sativa*.

Keywords: coherent diffractive X-ray imaging, lensless imaging, coherent X-ray diffractive imaging, flash diffractive imaging, single particle imaging, phase retrieval, X-ray free-electron laser, XFEL, FEL, CXI, CDI, CXDI, FXI, Machine learning, Deep learning, Reinforcement learning, Neural networks, Convolutional Neural Networks

Alfredo Bellisario, Department of Cell and Molecular Biology, Molecular biophysics, Box 596, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-2120-2
URN urn:nbn:se:uu:diva-526527 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-526527)

Dedicated to Lina

Charlie Brown: *Do you ever think much about the future, Linus?*
Linus: *Oh, yes... all the time.*
Charlie Brown: *What do you think you'd like to be when you grow up?*
Linus: *Outrageously happy!*
Charles Schultz, *Peanuts*, 1960

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Alfredo Bellisario**, Filipe RNC Maia, and Tomas Ekeberg. Noise reduction and mask removal neural network for X-ray single-particle imaging. *Journal of Applied Crystallography* 55.1: 122-132, 2022.

II **Alfredo Bellisario** and Tomas Ekeberg. Deep learning phase retrieval in X-ray single-particle imaging and object support from autocorrelations. *Manuscript*, 2024.

III Patrick Konold*, Leonardo Monrroy*, **Alfredo Bellisario**, Diogo Filipe, Patrick Adams, Roberto Alvarez, Richard Bean, Johan Bielecki, Szabolcs Bódizs, Gabriel Ducroq, Helmut Grubmueller, Richard Kirian, Marco Kloos, Jayanat Koliyadu, Faisal Koua, Taru Larkiala, Romain Letrun, Fredrik Lindsten, Michael Maihöfer, Andrew Martin, Petra Mészáros, Jennifer Mutisya, Amke Nimmrich, Kenta Okamoto, Adam Round, Tokushi Sato, Joana Valerio, Daniel Westphal, August Wolter, Tej Yenupuri, Tong You, Westenhoff, S., Maia, F. Microsecond time-resolved X-ray scattering by utilizing MHz repetition rate at second-generation XFELs. *Submitted to Nature Methods*, 2023.

List of additional papers

The following are the publications to which I have contributed but not treated in this thesis:

IV Patrick E Konold, Tong You, Johan Bielecki, Joana Valerio, Marco Kloos, Daniel Westphal, **Alfredo Bellisario**, Tej Varma, August Wollter, Jayanath CP Koliyadu, Faisal HM Koua, Romain Letrun, Adam Round, Tokushi Sato, Petra Mészárs, Leonardo Monrroy, Jennifer Mutisya, Szabolcs Bódizs, Taru Larkiala, Amke Nimmrich, Roberto Alvarez, Patrick Adams, Richard Bean, Tomas Ekeberg, Richard A Kirian, Andrew V. Martin, Sebastian Westenhoff, Filipe RNC Maia. 3D-Printed Sheet Jet for Stable Megahertz Liquid Sample Delivery at X-ray Free Electron Lasers. *IUCrJ*. 10, 662-670, 2023.

Contents

Abbreviations

AGIPD	Adaptive Gain Integrating Pixel Detector
AI	Artificial Intelligence
ANN	Artificial Neural Network
AsLOV2	*Avena Sativa* Light-Oxygen-Voltage 2
CDI	Coherent Diffractive Imaging
CNN	Convolutional Neural Network
CXDI	Coherent X-ray Diffractive Imaging
DL	Deep Learning
Cryo-EM	Cryogenic Electron Microscopy
ESI	Electrospray Ionization
ER	Error Reduction
FEL	Free Electron Laser
FSC	Fourier Shell Correlation
FXI	Flash X-ray Imaging
GAN	Generative Adversarial Network
GDVN	Gas Dynamic Virtual Nozzle
HIO	Hybrid Input-Output
ML	Machine Learning
PRTF	Phase Retrieval Transfer Function
RAAR	Relaxed Averaging Alternating Reflection
RL	Reinforcement Learning
SASE	Self Amplified Spontaneous Emission
SPB/SFX	Single Particles, Clusters, and Biomolecules & Serial Femtosecond Crystallography
SPI	Single-Particle Imaging
SQS	Small Quantum Systems
XFEL	X-ray Free Electron Laser

Part I:
Introduction

Motivation

X-RAY crystallography revolutionalized biology, enabling the structural characterization of the most important molecules that characterize life processes. More recently, Cryo-EM achieved similarly stunning results, bridging the gap to crystallography and becoming a powerful new tool for structural biology down to atomic resolution.

These fascinating experimental tools have recently been complemented by the stunning achievements of AI in the field. The unexpected success of AlphaFold [1] provided a new method for investigating proteins and their role in biochemical processes in just a couple of years. At the time of writing, the Protein Data Bank has 215,684 files with experimentally solved structures and 1,068,577 Computed Structure Models. It is impressive that until 2021, no computational structure was available as computational methods for structure solving were limited and their reliability doubtful.

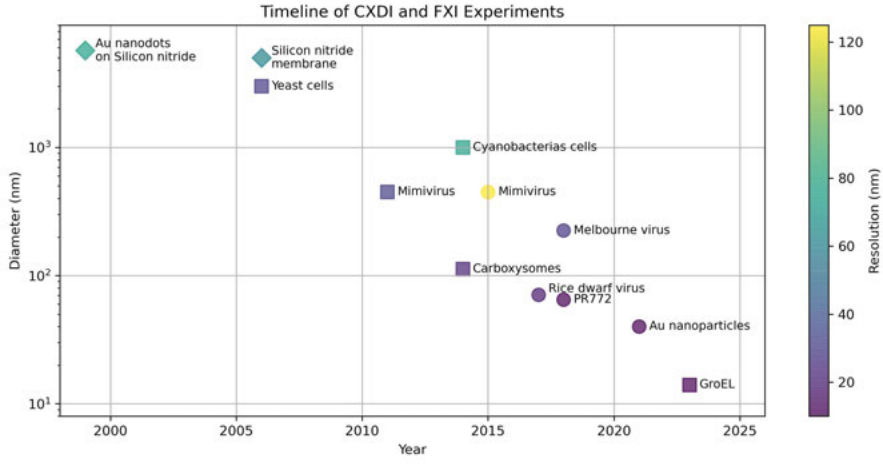

Figure 1. Timeline of few relevant experimental achievements in Coherent X-ray Diffractive Imaging (CXDI) and Flash X-ray Imaging (FXI). Diamond markers refer to 2D proof-of-concept experiments [2, 3], square markers refer to 2D reconstructions of biological particles [4, 5, 6, 7, 8], and circles refer to 3D reconstructions [9, 10, 11, 12, 13]. The state of art resolution for 3D reconstructions is 2.1 nanometers for 17-nanometer gold cubes [13] and, for biological particles, 7.8 nanometers for PR772, a 70 nm diameter icosahedral protein viral capsid [11].

In the age of AI, why should we develop new experimental methods? One possibly trivial answer is that AI still requires data to be trained on or that science, of course, is based on experimental work - but a more interesting answer is that today, one of the open challenges in structural biology is studying the dynamics of proteins at high resolution, a vast field still unexplored. Cryo-EM allows for the study of slow molecule dynamics. However, it is still limited to the order of milliseconds due to the vitrification of the sample [14], and time-resolved crystallography is limited in precision by the synchronization and uniformity of the reaction initiation across the crystal [15]. While X-ray crystallography and Cryo-EM are great tools, they require crystallization and freezing of the sample, making dynamic studies difficult. In the words of Richard Feynman, "The world is a dynamic mess of jiggling things..." and although structural knowledge of proteins plays a fundamental role in science, most intriguing natural phenomena happen at temperatures and time scales, that are not probed by current experimental methods. Even AI tools still struggle to explore such a diverse and complicated landscape as the one of life molecules.

Single-particle imaging (SPI) or Flash X-ray Imaging (FXI) is an emerging technique that allows the probing of nanometer-sized objects in the gas phase. Envisioned to achieve high resolution on single proteins, its development is based on the diffraction and destruction principle [16, 3]. The possibility of using ultrashort and high brilliance pulses to probe matter was one of the leading reasons for building new X-ray sources: X-ray Free Electron Lasers (XFELs).

By outrunning radiation damage, the highly intense femtosecond pulses of XFELs open the possibility of imaging the structure and dynamics of uncrystallized single-bioparticles down to the femtoseconds timescale [17]. They thus may enable the mapping of the conformational landscape of molecules. However, this method has faced challenges mainly due to the low signal of the scattering from the bio-particles compared with background sources and the limitations in the current sample delivery techniques. A simple historical outline of publications in the field is summarized in Figure 1. At the European X-ray Free Electron Laser (EuXFEL) facility, recent experiments reached a nanometer resolution on the 3D reconstruction of several kinds of nanoparticles.

While the resolution is still not comparable to Cryo-EM, these ideas have flourished in applying XFELs in other experimental techniques, such as serial femtosecond crystallography and solution scattering, already providing biologically relevant results. While SPI/FXI is still in development, other methods can already probe the conformational states and dynamics of proteins in their native state many orders of magnitude faster than what is possible with Cryo-EM. Today's main hardware advancements in the field include the development of low-background injection methods, higher repetition rates, and nanofocusing of the X-ray pulses.

In parallel with experimental successes, machine learning methods could foster data analysis. Several terabytes of data are commonly collected within a few days of beamtime. In 2023, the SPB/SFX instrument of the European XFEL in Hamburg collected around 10 PB of data [18]. This data deluge poses new challenges in data processing, and machine learning could be a valuable addition to further improving SPI/FXI.

Thesis outline

This thesis addresses the scientific challenges of XFEL imaging experiments and how new computational methods, such as deep learning, may enhance existing algorithms. Chapter 1 deals with diffraction and the formalism required to describe coherent diffractive imaging. Imaging methods at XFELs and the phase retrieval algorithms are discussed in Chapter 2 and Chapter 3, respectively. Section 2.5 briefly introduces solution scattering; this section will be useful to introduce the context for discussing results from **Paper III**. Chapter 4 introduces neural networks before discussing a few neural network architectures used in the papers in Chapter 5 (**Paper I** and **Paper II**). Chapter 6 concerns Reinforcement Learning, which will be further discussed in Section 7.4. The last two chapters focus on the work done for the papers on which this thesis is based. In particular, Chapter 7 concerns neural networks and their applications in Flash X-ray Imaging and Chapter 8 summarizes computational studies for the structural fitting of Time-resolved Small-Angle Scattering data (**Paper III**).

Part II:
Lensless Imaging

1. X-ray scattering

Following Wilhelm Conrad Röntgen's discovery of X-rays in 1895, the scientific community immediately realized X-ray the potential of radiation for imaging purposes, finding immediate medical applications within just a year [19]. When, in 1912, Max von Laue was for the first time able for the first time to study crystal structures through the diffraction of monochromatic X-rays [20], it was finally possible to see and understand molecular bonds. The application of X-ray crystallography on protein crystals opened the door for the field of structural biology and the investigation of protein molecules down to the atomic scale. The wavelengths suitable for these studies at a sub-nanometer resolution fall within 0.1 to 1 nanometers, part of the X-ray spectrum (0.01 to 10 nanometers). The choice of wavelength for a diffraction experiment depends on many factors, ranging from the desired resolution to many other phenomena like photoabsorption and fluorescence, which could play a significant role.

1.1 Diffraction

Diffraction occurs when scattered waves interfere with each other. During photon scattering, some photons are absorbed by the sample and reemitted in several ways. Many processes, such as Auger decay, fluorescence, and phosphorescence, can change photon energy and momentum. Energy and wavelength remain constant for elastically scattered photons, and only the direction changes. Elastically scattered photons interfere with each other and create a diffraction signal that can be collected and distinguished from inelastically scattered photons. When collecting light far from the radiation source, such that we can assume the incident waves are parallel, we say that the experimental setup satisfies the *Fraunhofer condition*, or that we are observing light in the *far field* [21]. Given a distance, o, between two scattering points, the *Fresnel number* is useful in optics to estimate whether we are in the far or near field and is defined as follows:

$$F = \frac{o^2}{\lambda D} \tag{1.1}$$

where λ is the wavelength of the incident photons and D is the distance between the interaction point and the observation plane. The Fraunhofer condition is satisfied when we collect photons at small angles or $F \ll 1$.

The scattering intensity is measured as a function of the *momentum transfer*, which depends on the scattering angle and the wavelength of the incident radiation. Assuming $\mathbf{k}_i$ is the wavevector, which describes the direction of the upcoming photons, and has an amplitude of $2\pi/\lambda$, we can define the scattered wavevector $\mathbf{k}_f$ as a vector of the same amplitude and different direction. The difference in direction is called *scattering vector* $\mathbf{q}$ or *momentum transfer*

$$\mathbf{q} = \mathbf{k}_f - \mathbf{k}_i \tag{1.2}$$

and its magnitude depends only on the scattering angle

$$|\mathbf{q}| = \frac{4\pi}{\lambda} \sin\left(\frac{\theta}{2}\right) \tag{1.3}$$

The interference depends on the spatial arrangement of the scatterer, which can be an object or an aperture, and in some simple cases, even calculated analytically. When scattering occurs with two pointlike objects, such as electrons, the elastically scattered light wave at each point does not change phase. Still, a phase difference exists due to the difference in path length. If the distance between two objects is $\mathbf{r}$, the phase difference can be calculated as

$$\Delta\phi = (\mathbf{k}_i \cdot \mathbf{r} - \mathbf{k}_\mathbf{f} \cdot \mathbf{r}) = \mathbf{q} \cdot \mathbf{r} \tag{1.4}$$

When considering diffraction from N scatterers, assuming the incoming wave is perfectly coherent and that the diffraction is observed at a high distance, D, the observed interference can be modeled by considering the superposition of waves generated by each scatterer

$$u(\mathbf{q}) = \frac{A}{D} \sum_{j=1}^{N} e^{-i\mathbf{q}\cdot\mathbf{r}_j} \tag{1.5}$$

where S is the scattering vector, A is the wave amplitude, and $\mathbf{r}_j$ the corresponding position in space of each scatterer.

The amplitude part of the wave superposition is the amplitude of the electromagnetic wave, and the phase part represents the wave's phase change. We can interpret the light-matter interaction of the incident radiation as a scattering process of light with an average electronic cloud due to electronic orbitals characterized by an electron density of charge. The formalism for a discrete number of scatterers can be reformulated for a continuous density ρ. We can then calculate the scattering from a density cloud as an integral

$$u(\mathbf{q}) = \frac{A}{D} \int \rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} dr \tag{1.6}$$

over the electronic density. The *structure factor*, $F(\mathbf{q})$ is the ratio between the scattered field from the sample density ρ and the scattered field that would be produced by a single electron. It depends only on the Fourier transform of the density. In the far field, therefore, the scattering of a plane wave is directly proportional to the Fourier transform of the densities.

1.2 Ewald Sphere

The *Ewald sphere* describes the wavefront spanned by the scattering wavevector at all possible angles [22]. One diffraction pattern gives information along one Ewald sphere through the sample's 3D Fourier space. In other words, the diffraction pattern samples the Fourier space. The Ewald sphere can be considered flat if we consider angles sufficiently small when the approximation to the Fraunhofer diffraction is valid.

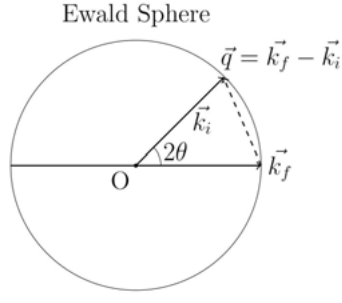

Figure 1.1. Ewald sphere construction from the scattering wavevector. The sphere radius is equal to $2\pi/\lambda$ and centered on the origin, O, of the object, ρ.

When the scattering angle is small, a condition known as *planar Ewald geometry*, the projected image of an object can be obtained by performing an inverse Fourier transform on the scattered wave. However, we cannot directly measure the phase of a wavefront because detectors only collect the photon intensity, which can be calculated as the square of the wave amplitude. This is known as the *phase problem*. The diffraction signal collected by a detector at a distance D can, therefore, be calculated as

$$I(\mathbf{q}) = |\mathfrak{F}(\rho(\mathbf{r})|^2 \tag{1.7}$$

where $\mathfrak{F}(\rho(\mathbf{r}))$ is the Fourier transform of the densities. The object's Fourier transform generates a 3D space relative to the Fourier transform electron density distribution, from which the detector captures a slice of the transform of the projected densities as a 2D diffraction pattern. By collecting diffraction patterns from many angles, we can collect the contribution from different regions of Fourier space. If combined, we can get complete 3D information about the sample.

2. Imaging with X-ray Lasers

Throughout the past century, physicists have developed brighter X-ray sources to improve the resolution of several X-ray imaging techniques. Synchrotrons are brilliant X-ray sources that generate radiation through *bremsstrahlung* by employing magnetic fields in a storage ring. Free Electron Lasers (FELs) are the latest X-ray sources, producing X-ray pulses billions of times brighter than synchrotron radiation. FELs further increased brilliance and coherence, providing short pulses thanks to these linear electron accelerators' micro bunched electron beams. Today, FELs can produce pulses of the highest brilliance (e.g., for comparison, the peak brilliance of the European XFEL is $5 \cdot 10^{33}$ photons/s/mm2/mrad2/(0.1% BW) [23] compared to PETRA III synchrotron $2 \cdot 10^{21}$ photons/s/mm2/mrad2/(0.1% BW) [24]) and very short pulse length, in the order of tens of femtoseconds. In this chapter, we discuss how FELs generate radiation and how we can use FEL radiation to image samples in gas and solution.

2.1 Undulator radiation

Undulators are regular-spaced sequences of magnets with alternating polarity that induce transversal oscillatory motion into the charged particles of an accelerator due to the Lorentz force. Undulators are critical for generating radiation in third-generation synchrotrons and free-electron lasers, producing highly coherent, monochromatic, and pulsed radiation. Synchrotron radiation is broad-spectrum, meaning that the wavelength of the emitted X-rays is not monochromatic. In FELs, electrons are accelerated linearly and passed to a series of undulators without using storage rings, hence avoiding the generation of broadband X-rays.

Due to the linear acceleration through undulators, electrons emit radiation and, more importantly, organize into microbunches. In each microbunch, electrons are in phase with each other and coherently emit radiation at the same wavelength, further increasing the coherence of the pulses, the total number of emitted photons, and the phase-locked motion of electrons. This phenomenon is called *self-amplified spontaneous emission* (SASE) [25]. SASE allows the generation of more intense and coherent X-ray pulses in XFELs.

The *undulator equation* [26] predicts the wavelength of the radiation pulse emitted by an undulator and can be written as follows

$$\lambda = \frac{\lambda_u}{2\gamma^2}\left(1 + \frac{K^2}{2} + \gamma^2\theta^2\right) \tag{2.1}$$

where γ is the relativistic Lorentz factor, λ_u is the spatial period of the undulator magnets, θ the angle of observation, and K a parameter known as the *undulator strength parameter* and that depends on the spacing of the undulator magnets, and the magnetic field strength. The photon pulse spans a narrow cone of angle $1/\gamma$ along the radiation direction, while the equation's quadratic term considers the off-axis observation.

The wavelength of emitted radiation can be adjusted by modifying the magnetic field strength instead of altering the magnets' period. This is done by changing the dimensionless undulator strength parameter by tuning the magnetic field B

$$K = \frac{eB\lambda_u}{2\pi m_e c} = 0.9337 \cdot B \cdot \lambda_u \tag{2.2}$$

effectively altering the wavelength without physically adjusting the undulator structure.

2.2 Diffraction-before-destruction

Inelastically scattered X-rays cause radiation damage and reduce the resolution of many synchrotron imaging experiments. The tolerable radiation limit depends on many factors, including the exposure length and beam fluence [27]. Pulses of a few femtoseconds can outrun the radiation damage of the absorbed X-rays, allowing structural studies, ideally, without observing radiation damage.

The *diffraction before destruction* principle is one of the milestones of FELs' application [16, 3].Using highly coherent and brilliant pulses of approximately ten femtoseconds ($\sim 10 \cdot 10^{-15}$ s), this phenomenon in principle enables, in principle, diffraction measurements at room temperature, at which the structure better resembles its native state. The intense radiation causes the sample to turn into hot plasma; consequently, the plasma ions repel each other. In the span of 100 femtoseconds, the sample is destroyed. This approach promises to overcome radiation damage in diffraction measurements, obtaining high-resolution images of molecules "frozen" in time.

Understanding the structure of biomolecules -such as proteins, cells, or membranes- in their native state and analyzing their conformational landscape through the realization of "molecular movies" (i.e., images of the same structure probed at different time points) is one of the many scientific causes proposed by the XFEL community thanks to the unique opportunities posed by the diffraction-before-destruction principle. In this sense, XFELs allowed the application of new imaging methods, including serial femtosecond crystallography and flash X-ray imaging.

2.3 Coherent Diffractive Imaging (CDI)

Among the many developments of imaging methods in the last decades, Coherent diffractive imaging (CDI) plays a special role, as it can probe high-resolution noncrystalline structures without lenses. Coherent X-ray diffractive imaging (CXDI), in particular, enables the reconstruction of objects the size of a few nanometers by collecting diffraction signals from scattering X-rays. In a microscope, the light diffracted by the object goes through a lens, and the numerical aperture of the lens and aberration effects from it determine the formation of an image and the resolution. In CXDI, the resolution is, in theory, determined solely by the maximum angle for the collection of photons and, of course, by the number of photons that can be collected. The information content for image reconstruction is available through the collected photons, and different imaging techniques (crystallography, ptychography, holography, etc.) solve the phase problem differently. In CXDI, real-space images are reconstructed thanks to iterative algorithms [28] that act as computational lenses, retrieving the missing phases required to transform diffraction data into images. The specifics of this method and the algorithms will be further discussed in Chapter 3.

2.4 Flash X-ray Imaging (FXI)

Flash X-ray Imaging (FXI) is a type of CXDI, but its experimental challenges are unique. FXI specifically tackles nanoscale imaging in gas with femtosecond pulses, ranging from viral capsids to Xenon clusters. The high brilliance of the XFELs allows enough scattered signal even from one single biological molecule. As discussed, the beam pulse's energy deposited on the sample immediately turns the sample into plasma; however, short pulses allow the diffraction signal to outrun the radiation damage. By spraying samples into a vacuum chamber and analyzing several 2D diffraction patterns of identical objects, it should be possible to study the 3D structure of the sample in its native form. So far, this technique has been applied to reconstruct 3D volumes of viral capsids and metallic nanoparticles with a state-of-the-art resolution of 2.1 nanometers for 17-nanometer gold cubes [13].

2.4.1 Sample delivery

In FXI, the sample particles in the solution are aerosolized in small droplets. Due to the pressure difference between the injector and the chamber, the buffer (e.g., ammonium acetate) evaporates, and the beam probes only the naked sample. In practice, a few molecules of water and buffer remain attached to the sample, and experimentally, the main challenges of sample delivery in FXI include (i) decreasing the droplet size, (ii) reducing gas injection, and (iii) lowering the sample consumption.

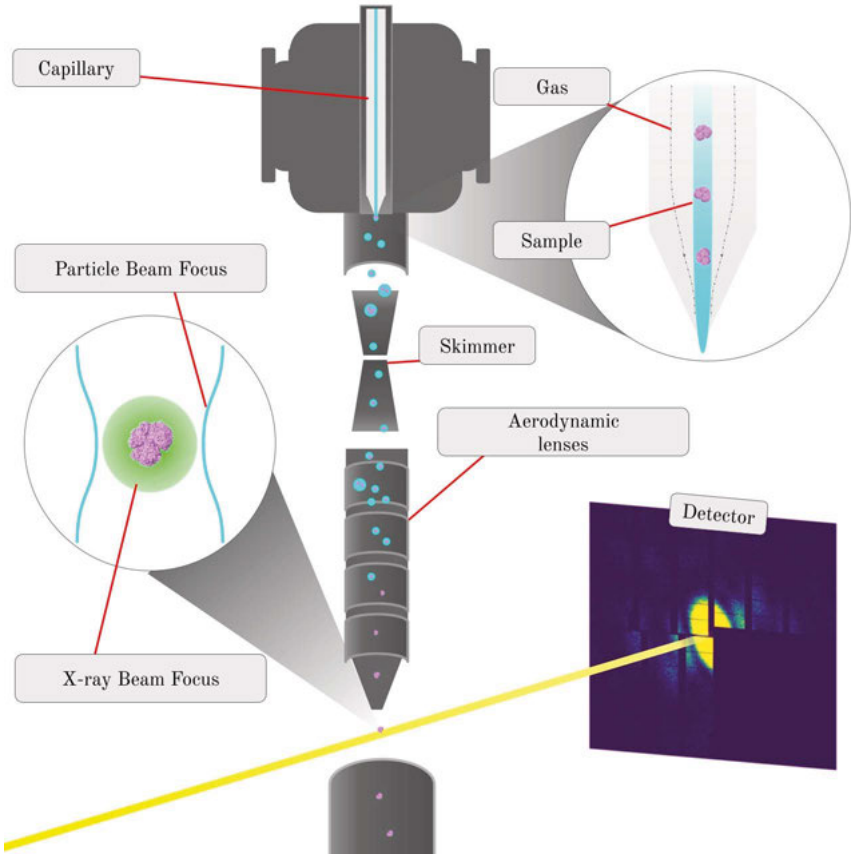

Figure 2.1. Scheme of the FXI experimental setup. The injection step can be driven by GDVN or ESI. When ESI is used, other elements in the experimental setup, such as the neutralizer, are needed to neutralize the charge of the injected particles. After aerosolizing, the particle beam goes through a skimmer and aerodynamic lenses to reach the vacuum chamber. The skimmer removes the excess gas and reduces the pressure to a few mbar, while the aerodynamic lenses are used to focus the particle beam. The pressure difference evaporates the buffer droplets; ideally, only the sample particles are injected. The diffraction signal is collected downstream by a high-speed X-ray detector. The modular detector does not collect the forward beam to avoid damage. The vacuum chamber where the X-ray pulses interact with the sample is kept at a pressure between 10^{-5} to 10^{-6} mbar.

Gas Dynamic Virtual Nozzles (GDVN) were one of the first sample delivery methods used in FXI. In GDVNs, the sample is injected through a nozzle, which propels the solution by a coaxial flow of the gas. The gas also scatters X-rays, one of the major noise sources in FXI; this is why GDVN uses light gases such as Helium for injection. The pressure difference aerosolizes the liquid flow, creating droplets of various sizes ranging from 400 to 2000 nm [29]. On average, the sample flow rate is estimated to be 2μl/min.

For small and volatile samples, aerosols from micron-sized droplets can be a significant problem, as they cause the deposition of nonvolatile solvent components on the sample [29, 30]. One solution to this problem is replacing the GDVN nozzle with an electrospray. The droplets produced by electrospray are significantly smaller (on average 150 nm [29]) and require less sample (0.06μl/min) but higher sample concentration ($5 \cdot 10^{14}$ particles/ml instead of $2 \cdot 10^{12}$ particles/ml for GDVN). In Electrospray ionization (ESI), an external electric field charges and squeezes the liquid jet, deforming the jet into a conical geometry called *Taylor cone* [31]. The electric charge drives the droplet formation by stripping droplets generated at the tip of the cone. A radioactive source or UV light then neutralizes droplets.

After injection, a skimmer reduces the excess gas in the aerosolization chamber. The aerosol stream is focused by stacked aerodynamic lenses, collimating the particle beam and gradually decreasing the pressure. The particles then reach the vacuum chamber, where they are intersected by the X-ray beam.

2.4.2 Orientation in 3D Fourier space

Due to the nature of FXI experiments, measuring the aerosol particles' orientation is impossible. Ways to control the orientation are under development, for example, by exploiting the electric dipole of proteins with an electric field [32, 33, 34]. Still, the orientation of the sample is generally unknown, and several algorithms have been introduced to calculate the relative orientation of each pattern [35, 36, 37]. If the curvature of the Ewald sphere is flat, then the Fourier transform of the projection of an object in real space corresponds to a slice in Fourier space. We can sample the 3D Fourier space through several diffraction patterns of projections at different orientations. The expansion maximization compression (EMC) [37] algorithm is an iterative algorithm that aligns 2D diffraction patterns in a 3D Fourier space by images, ensuring an accurate overlay of the intensities in the patterns in 3D. The algorithm can be divided into three steps as follows:

1. **Expansion** - each diffraction pattern, initially in random orientation, is used to calculate a model of the 3D space. Initially, the model is a starting guess of the 3D intensity space and is sampled into N-slices. The sampling happens along the Ewald-sphere, corresponding to every possible orientation.
2. **Maximization** - each slice of the 3D model is compared to each measurement. The best fit for each pattern in the current model is calculated. This step is the most computationally expensive one. In this step, the best fit is modeled according to the expected distribution function of the photons in the pattern. For a Poisson sampled measurement, we can, for

example, consider

$$d_{\text{poisson}}(K,W) = \prod_i \frac{e^{W_i}(W_i)^{K_i}}{K_i!} \tag{2.3}$$

where K_i and W_i are, respectively, the i-th 2D diffraction patterns and the i-th slice of the 3D volume. This distribution can be changed to consider other kind of photon distributions. This was modified for **Paper I** to account for continuous and Gaussian-distributed intensities. In this step, we can also consider the symmetries of the sample, if known. The maximize step also updates the expanded model based on the probabilities calculated. Each slice in the expanded model is set to a weighted average of all the patterns, where the weights are the calculated probabilities. This will ensure that the slices in the next expanded model are internally consistent.

3. **Compression** - In the final step, a new model is generated from the best fit of each pattern.

These steps are repeated iteratively to achieve an increasingly better fit.

2.5 Small Angle X-ray Scattering (SAXS)

A different kind of low-resolution X-ray diffractive imaging technique that is particularly useful to probe proteins' dynamics in their native state is Small-angle X-ray Scattering. SAXS is a scattering imaging technique involving X-rays' scattering by an ensemble of particles in solution. Collecting the average scattering at small angles (up to 1-2Å^{-1}) provides information about the sample's size and shape while keeping molecules in solution closer to their actual biological environment. In SAXS, detector images are azimuthally integrated because of the expected signal symmetry resulting from collectively imaging multiple randomly oriented samples. Structural information without scattering contributions from the bulk solvent [38] can be obtained from the scattering profile. The isotropic scattering intensities of a protein sample can be calculated with respect to the difference $\Delta\rho$ between the density of the protein ρ and the bulk solvent density ρ_S also known as *contrast* as

$$I(\mathbf{q}) = \frac{1}{V}\left|\int_V \Delta\rho(\mathbf{r})e^{i\mathbf{q}\cdot r}\mathrm{dr}\right|^2 \tag{2.4}$$

over a scattering volume V. Removing the solvent contribution (i.e., the scattering from the solution) is essential for isolating the scattering signal solely due to the particle ensemble. The general approach to achieve this involves measuring the scattering from the solvent alone and then subtracting this from the scattering measured from the sample suspended in the solvent.

$$I_{\text{sample, corrected}}(\mathbf{q}) = I_{\text{sample + solvent}}(\mathbf{q}) - I_{\text{solvent}}(\mathbf{q}) \tag{2.5}$$

This step is necessary before structural fitting and is important to estimate the solvent contribution to achieve a good contrast. The rotational symmetry of SAXS diffraction allows retrieving only low-resolution maps. SAXS only provides the shape of the sample, and often, only bead model reconstructions are possible. The structural fitting is achieved via iterative comparisons of the bead model to experimental data.

2.5.1 Time-resolved SAXS

While conventional SAXS provides static structural information, time-resolved SAXS (TR-SAXS) adds the temporal dimension to the analysis. TR-SAXS enables the capture of structural changes over time by collecting SAXS data at different time points of a triggered reaction, making it an invaluable tool for studying dynamic processes in materials and biological systems. In TR-SAXS experiments [39], rapid events or reactions are triggered within a sample, usually via photon excitation or pH changes. X-ray scattering profiles are collected at multiple time points during the process. Scattering curves corresponding to the same X-ray-laser delay time are averaged together, as they correspond to the same molecular state, to investigate dynamics. We can analyze the difference scattering from the structure probed at different times to gain insight into the structural changes:

$$\Delta I(\mathbf{q},t) = I_{\text{light}}(\mathbf{q},t) - I_{\text{dark}}(\mathbf{q},t_0) \tag{2.6}$$

This time-resolved data can then be analyzed to track changes in structure, conformation, or assembly with high temporal resolution. TR-SAXS has far-reaching applications, including monitoring protein folding, studying chemical reactions, exploring materials' phase transitions, and observing the kinetics of self-assembly processes.

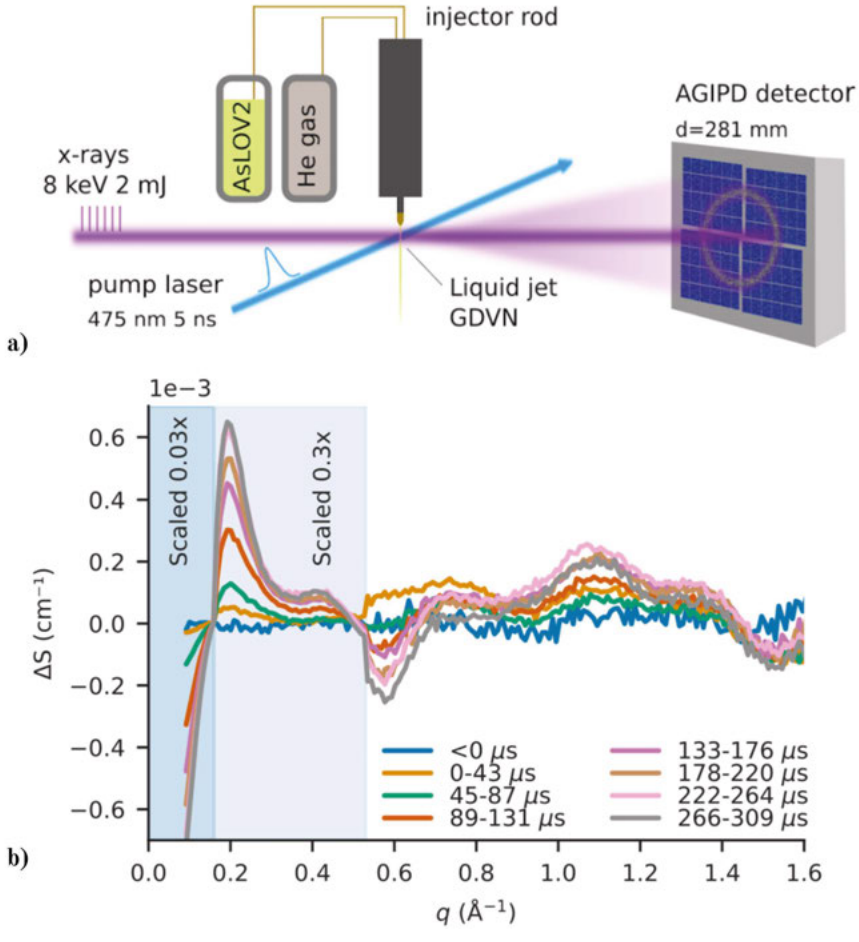

Figure 2.2. Time-resolved SAXS scheme and data, as in **Paper III**. a) Experimental setup, showing the X-ray and optical laser path, the sample delivery system, and the detector. b) The X-ray scattering signal, measured at the European XFEL, adjusted for different sample concentrations and scaled to theoretical data simulated with Pepsi-SAXS [40]. Low-angle data was reduced tenfold, and high-angle data tripled for clarity.

3. Iterative Phase Retrieval

By assuming knowledge of the object's overall shape (i.e., assuming a compact support), iterative algorithms can be applied to retrieve phases in CDI, solving the phase problem. After many iterations, the updated phases will be very close to the real solution, allowing us to calculate a final image reconstruction. Phase-retrieval algorithms are one of the cornerstones of CDI, and their correct application is pivotal to achieving good reconstructions.

3.1 The phase problem

The phase problem is a common challenge in crystallography and, in general, all diffractive imaging methods, e.g., holography, ptychography, and CDI. As noted for crystallography by Sayre [41], in 1952, the Bragg peaks sampled the molecular transform at a rate sufficient to retrieve the electron density in the lattice. This critical rate is the *Nyquist rate* and defines the minimum sampling rate required to retrieve a signal. According to the Shannon sampling theorem [42], the minimum sampling frequency required to reconstruct a signal with the highest frequency B, the sampling rate (or *Shannon pixel size*, the smallest pixel size at which an image can be sampled) of the detector, should be at least twice the highest spatial frequency present in the object. In CDI, the additional information provided by the finite dimensions of the sample helps resolve phase issues. Patterns that satisfy the oversampling condition provide redundant information, which can be leveraged to retrieve the phases.

3.2 The algorithms

In CDI, the physical lenses used in a traditional microscope are replaced with computational algorithms and software that can convert diffraction data into densities. Multiple solutions are possible for each set of amplitudes as phase reconstructions are invariant under translation, centrosymmetric symmetry, and global phase shift. Several phasing algorithms can be used to face different challenges posed by phasing, but most of them follow the same schematic iterative process as represented in Figure 3.1.

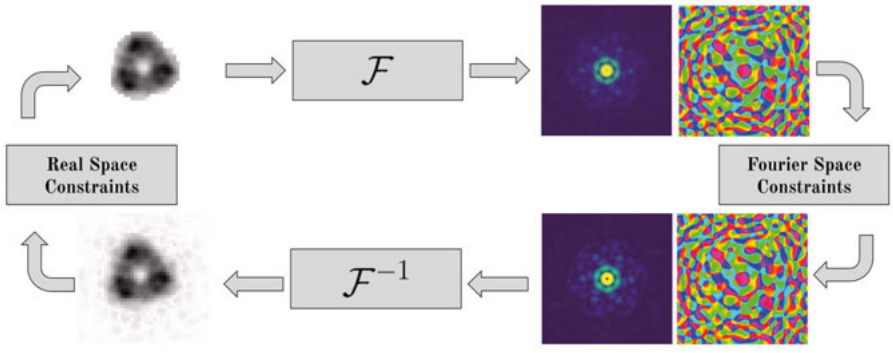

Figure 3.1. Phase Retrieval scheme. In the first iteration of the phasing process of CDI data, the densities and the phases are usually unknown. The algorithm calculates the real space densities starting from a random initialization of the phases. Densities are then projected within the support before projecting back to Fourier space. Other common constraints of real space include requiring positivity and connectivity for the densities. In Fourier space, phases are updated at each iteration while the experimental amplitudes are enforced. Several thousands of iterations are usually required to achieve convergence to a solution.

3.2.1 Error Reduction

Error Reduction (ER) [43] was introduced to calculate the phase of an object from the measurement of an oversampled diffraction pattern. Given the intensity signal of a density ρ, the algorithm alternates between real space and Fourier space through successive iterations; the support is used as a restraint in the real space and provides a region where the object is supposed to have non-zero density while setting the rest of the real space to zero. By keeping the amplitudes equal to the measured ones and iterating the process, it is possible to demonstrate a monotonic convergence of the algorithm. The measured amplitudes are kept constant throughout the iterations of the algorithms while the phases are iteratively updated. The ER algorithm can be summarized in the following steps.

1. Start with an initial guess for the phase of the object, ϕ_0, and combine it with the measured amplitude to construct a complex field in the Fourier domain.
$$F_0 = \sqrt{I_{\text{meas}}} \cdot \exp(i\phi_0) \tag{3.1}$$
2. Inverse Fourier Transform this field to get to real space
$$\hat{\rho}_1 = \mathscr{F}^{-1}[F_0] \tag{3.2}$$
where Π_S is the projection map over the known support S.
3. Apply real-space constraints and Fourier Transform back
$$F_1 = \mathscr{F}[\Pi_S \hat{\rho}_1] \tag{3.3}$$

where Π_S corresponds to applying the following constraint

$$\hat{\rho} = \begin{cases} \hat{\rho} & \text{if } \hat{\rho} \text{ is inside the support} \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

4. Replace the amplitude with the measured amplitude and the phase obtained from the Fourier Transform in the previous step

$$F_2 = \sqrt{I_{\text{meas}}} \cdot \exp(i\phi_1) \tag{3.5}$$

5. Iterate steps 2-4 until convergence.

Loose supports can cause stagnation into local minima. It is common to obtain multiple non-global solutions when running the same algorithm in parallel several times. Applying real-space constraints like non-negativity and support boundaries helps break invariant symmetries and enables convergence in CXDI. The convergence of ER can be monitored by calculating the mean squared error in Fourier space

$$E_f = \sum_i (|\mathscr{F}^{-1}[\Pi_S \hat{\rho}]| - |\sqrt{I_{\text{meas}}}|)^2 \tag{3.6}$$

where $\hat{\rho}$ is the real space solution of ER and the sum over all the points. In the same way, in real space, we can track the convergence of the algorithm by calculating the MSE metric

$$E_r = \sum_i (|\Pi_S \hat{\rho}| - |\mathscr{F}[\sqrt{I_{\text{meas}}} \cdot \exp(i\hat{\phi})]|)^2 \tag{3.7}$$

where $\hat{\phi}$ are the phases obtained by ER. The algorithm reduces the error after every step but cannot escape local minima once it enters that solution. For this reason, this algorithm cannot solve the phase problem except for very particular cases. Instead, it's usually only used as a refinement algorithm when we are already at the correct minimum.

3.2.2 Hybrid Input-Output

The Hybrid Input-Output (HIO) [44] controls the step size between the support constraint projection and the modulus constraint projection, escaping local minima and enabling HIO to explore diverse solution spaces and solving the stagnation problem of ER. In phase retrieval algorithms, the constraints applied in each iteration are always determined by combining two projections, one onto the support constraint and the other onto the modulus constraint. If the separation between the two projections is large, the step length will be large, and the algorithm will probably explore other areas. For the HIO algorithm, the β parameter is a relaxation factor used to scale the changes made in each iteration.

As for ER, starting from an initial guess ϕ_0 the algorithm operates as follows for each n-th iteration:

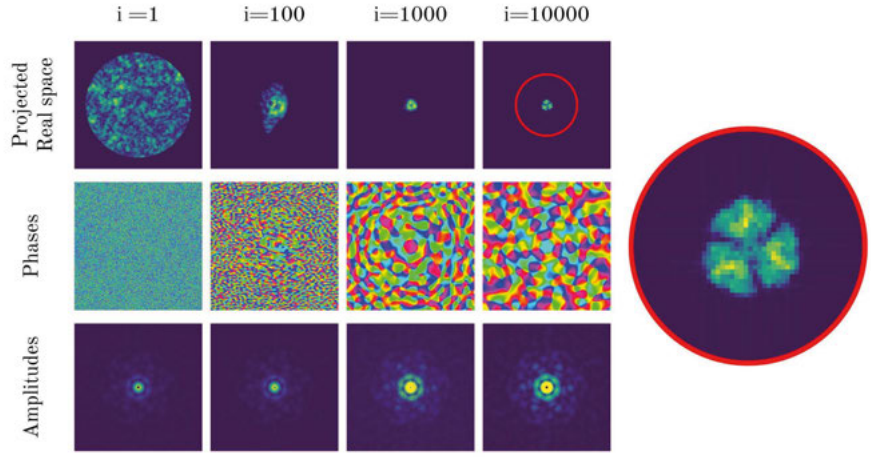

Figure 3.2. Reconstruction example using Shrinkwrap after i iterations. Top row, real space after support projection. Middle row, phases. Bottom row, amplitudes. On the right, zoom in on the densities from photosystem I, PDB file 1JB0 [45].

1. Calculate the diffraction pattern:

$$I_n = |\mathscr{F}[\hat{\rho}_{n-1}|]$$

2. Apply the measured magnitude and keep the calculated phase:

$$F_n = |\sqrt{I_{\text{meas}}}| \cdot \exp(i\hat{\phi}_n)| \tag{3.8}$$

3. Inverse Fourier Transform:

$$\hat{\rho}_n = \mathscr{F}[F_n]$$

4. Apply the object domain constraints:

$$\hat{\rho}_n = \begin{cases} \hat{\rho}_n & \text{if } \hat{\rho}_n \text{ is inside the support} \\ \hat{\rho}_{n-1} - \beta\hat{\rho}_n & \text{otherwise} \end{cases} \tag{3.9}$$

HIO performance could be compromised when dealing with very noisy data. HIO does not have a unique solution, and depending on the strength of the negative feedback, there might be no point that fulfills both density constraints.

3.2.3 Shrinkwrap

The shrinkwrap algorithm [46, 47] refines the support for every N-th iteration; it was introduced to improve HIO without relying on accurate support functions, allowing it to converge for loose supports. Shrinkwrap applies a Gaussian filter on the current best guess and keeps only those pixels above

a certain threshold in the support. For example, the autocorrelation A of the signal

$$A(\mathbf{r}) = \mathscr{F}^{-1}\{I(\mathbf{q})\}$$

is sometimes used to estimate a starting support. If the object's overall shape is known, a pragmatic solution consists of using it as an initial guess. When this is unfeasible, the autocorrelation function can be used as loose support. Typically, phasing algorithms are run in parallel several times to retrieve many independent solutions from different initial conditions. The reconstruction can be obtained by averaging the solutions with the lowest error.

A simple example of phase retrieval is reported in Figure 3.2.

3.2.4 Other algorithms

Other similar algorithms introduce other constraints, such as sparsity or non-negative densities. Other algorithms for phase retrieval can be developed. Some examples include Relaxed Averaged Alternating Reflections (RAAR) [48] for iterative refinement, Difference Map (DM) [49] for error reduction, saddle point optimization techniques [50] for fine-tuning, noise-aware phasing methods [51, 52], and memetic algorithms [53] for exploring solution space. Other algorithms, such as the Convex Optimization of Autocorrelation with Constrained Support (COACS) [54], were introduced to recover the autocorrelation when missing data is present or to estimate tighter constraints.

3.2.5 Triple intersection method

The triple-intersection method [56] was introduced to derive tight supports from the signal's autocorrelation. If we assume no knowledge of the object shape, using the autocorrelation function of the diffraction intensities as a loose support for the algorithms is possible. However, the autocorrelation is insufficient for phase retrieval in many cases, especially for noisy patterns, and tighter supports are often necessary. The method works by calculating a binary mask from the autocorrelation function A of the signal and looking for its extremal points. Copies of this autocorrelation mask are then shifted along vectors defined by the extremal points, and the overall intersection is calculated, obtaining a smaller support mask of S. This method requires a good approximation of the autocorrelation, which is often experimentally unavailable. This makes the technique unfeasible on most occasions. Yet, when this technique can be applied, it returns reasonable estimates of the optimal object support and possibly the best *a priori* support without relying on previous low-resolution masks.

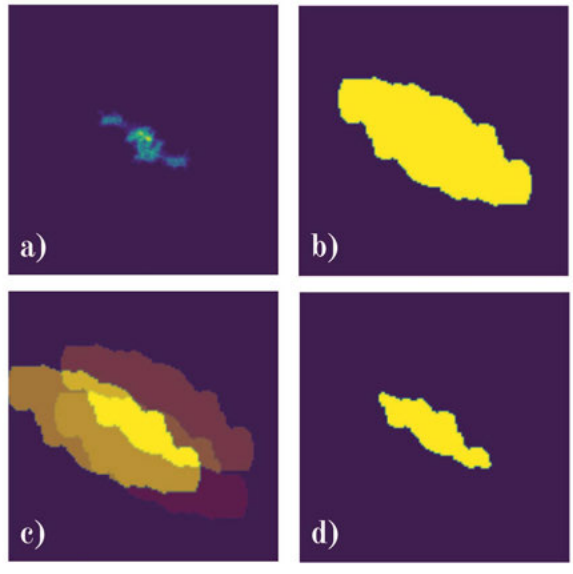

Figure 3.3. Example of the application of the triple intersection method to retrieve support estimates reproduced from **Paper II**. a) Protein densities (PDB-ID 1K3E [55]). b) Autocorrelation mask. c) Three copies of the autocorrelation mask intersecting. d) Support estimate.

3.2.6 DENsity from Solution Scattering (DENSS)

DENsity from Solution Scattering (DENSS) [57] is a new computational algorithm that reconstructs a molecule's three-dimensional electron density distribution based on the SAXS profile. This provides a big step forward in the structural fitting of SAXS data, for which only low-resolution shape reconstructions are usually possible. In some cases, another low-parameter model, such as an existing structure, can be used while enabling the free movement of a few molecule domains. Still, standard SAXS reconstructions generally are bead models and cannot provide direct information on the electron density.

The algorithm utilizes the structure factor derived from the SAXS profile to compute three-dimensional intensities that are then partitioned into concentric shells according to their wavevector values. Spherical averages are obtained and compared against experimental scattering data by averaging the intensities within each shell. The 3D structure factors are adjusted to match the experimental data within each shell, and an inverse Fast Fourier Transform (FFT) is applied to generate a new electron density map in real space. Density values beyond the defined support are set to zero to constrain the reconstruction. Figure 3.4 shows a schematic summary of these steps.

The process iterates by calculating a forward FFT from the flattened density to produce updated structure factors, and the cycle continues until convergence is achieved. This iterative approach allows DENSS to refine the electron

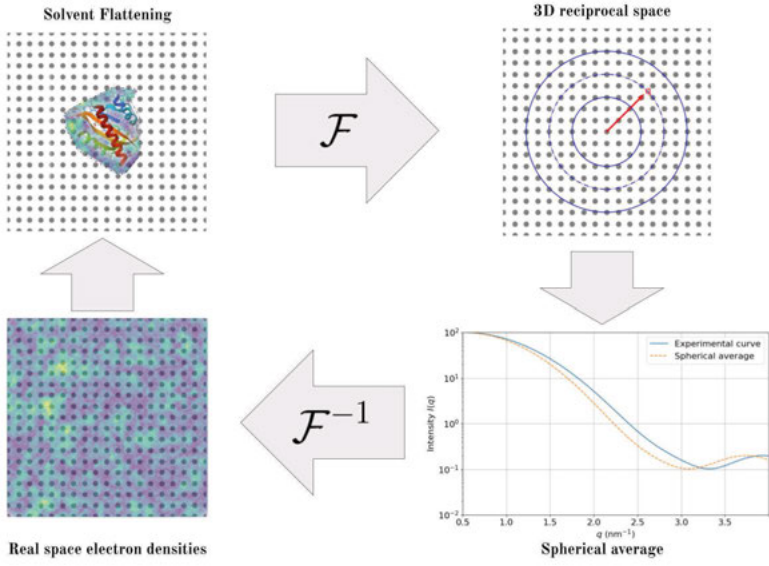

Figure 3.4. DENSS scheme as described in [57]. DENSS calculates 3D structure factors from scattering intensities, scales them to match experimental data, and repeats the cycle until convergence. Intensities are divided into concentric shells and compared to the experimental data.

density map, gradually improving its agreement with the experimental SAXS data. Many solutions fulfill the constraints; in principle, one global solution is not ensured.

3.3 Reconstruction quality

The presence of noise and artifacts in the data collection worsens the structural fitting's performance in SAXS and CXDI. In this section, I introduce the concept of resolution and the metrics used for estimating it.

3.3.1 Resolution

The resolution in an experiment is determined by how far from the origin in Fourier space we collect data. As discovered by Ernst Abbe [58] in mi-

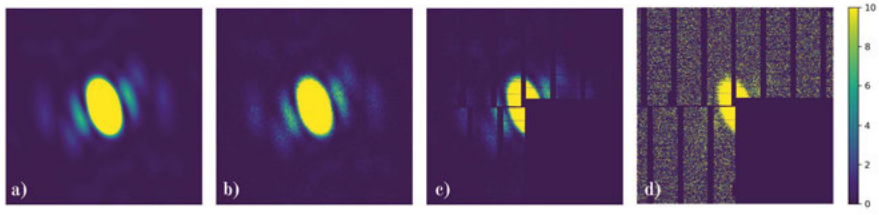

Figure 3.5. Examples of simulated diffraction data of photosystem I from PDB file 1JB0 [45]. Simulations were performed in *Condor* [61] with a photon energy of 1.2 keV, pulse energy 10 mJ, pulse diameter 1 μm and setting the detector distance at 30 cm from the interaction region for a pixel size of 200 μm. a) Continuous intensities. b) Poisson sampled intensities. c) Masked and Poisson sampled intensities, simulating the diffraction signal recorded with an AGIPD detector assuming no background signal. d) Simulated background noise, simulated as a uniform noise with a FWHI (full width at half maximum) of the Gaussian distribution of 10 pixels, added on pattern c).

croscopy, this corresponds to the experiment's numerical aperture or maximum scattering angle. We can write the following formula for resolution δr:

$$\delta r = \frac{\lambda}{2 \cdot NA} = \frac{\lambda}{2 \cdot \sin \theta_{max}} \tag{3.10}$$

Therefore, the resolution in CDI is determined by the incident radiation's wavelength and the maximum scattering angle. The higher the resolution, the more detailed the structural information. In practice in CDI, error metrics calculated from the reconstructions, such as the *phase retrieval transfer function* or the *Fourier shell correlation*, are used to estimate the resolution of the experiment.

3.3.2 Background signal and missing data

High-speed X-ray detectors, such as the AGIPD detector [59] of the SPB/SFX instrument or the pnCCD [60] detector of SQS at the EuXFEL, are amazing cameras, capable of detecting and counting photons over a huge range (from one to thousands) per pixel. They are also required to keep up with the high repetition of pulses from the beamline. These cameras are extremely sensitive and have areas (called *masks*) where data is not collected to avoid damaging the detector with the intense forward scattering from the beam. Dealing with missing data is a significant challenge during image reconstruction, as the lack of low-angle data makes phasing increasingly difficult. Strategies for managing missing data typically involve using interpolation and statistical methods to estimate the absent information [54]. Scattering from the gas injection and the X-ray optics is another significant noise source.

In particular, the background signal from the injection gas can be ten-fold higher than the strength of the naked protein signal for small low scatterers.

The background is usually modeled by collecting dark runs where only gas is injected into the chamber. This allows us to estimate the contribution of the signal coming from gas, which can then be subtracted from the measurements. Dark runs with no sample nor gas injection are also useful for studying the detector's response and creating binary masks of missing data and saturated pixels due to stray lights or misbehaving pixels. Pixel gain correction [62] and hit-finding routines [63] are other essential steps in preprocessing raw data to correct measurements before orientation retrieval and phasing. To evaluate the reproducibility and quality of reconstruction, creating new metrics for FXI to correctly assess the signal-to-noise ratio and the resolution of reconstruction post-phasing becomes essential. The phase retrieval transfer function and the Fourier shell correlation are two common error metrics used to estimate an experiment's resolution.

3.3.3 Phase Retrieval Transfer Function

The Phase Retrieval Transfer Function (PRTF) [3, 17] is a metric used to evaluate the performance of phase retrieval algorithms. It measures the reproducibility with which a phase retrieval algorithm can recover the phase of an object from its diffraction pattern or experimental data. In other words, it estimates how independent the reconstruction is from the random start used for phase retrieval. Many reconstructions are performed, and the PRTF compares them. Mathematically, the PRTF is defined as the ratio of the Fourier transform of the estimated object phase to the measured intensities.

$$PRTF(\mathbf{q}) = \frac{|(\sum_{k=1}^{N} e^{(i\cdot\phi_k(\mathbf{q}))})|}{N} \tag{3.11}$$

Then, the radial average of the PRTF is calculated to estimate the achieved resolution. The spatial frequency at which the correlation drops below a certain threshold (e.g., $1/e$) is often used to define this resolution limit for the given dataset.

3.3.4 Fourier Shell Correlation

The Fourier Shell Correlation (FSC) [64] is a metric used to assess the resolution and reliability of reconstructed images, particularly in Cryo-EM and FXI. It calculates the Fourier transforms of the corresponding spatial frequency shells in two reconstructions

$$FSC(\mathbf{q}) = \frac{\sum_{\mathbf{q}_i \in \mathbf{q}} F_1(\mathbf{q}_i) \cdot F_2(\mathbf{q}_i)^*}{\sqrt[2]{\sum_{\mathbf{q}_i \in \mathbf{q}} |F_1(\mathbf{q}_i)|^2 \cdot \sum_{\mathbf{q}_i \in \mathbf{q}} |F_2(\mathbf{q}_i)|^2}} \tag{3.12}$$

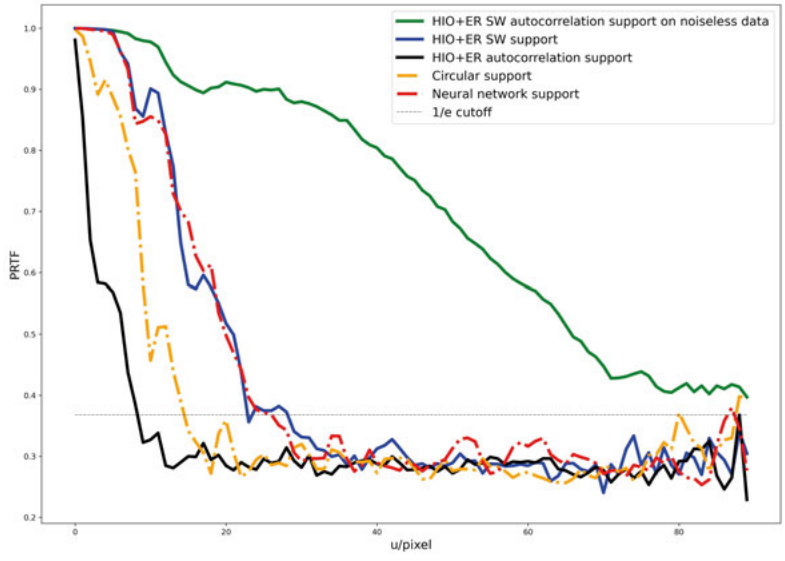

Figure 3.6. Phase retrieval transfer function example from reconstructions using different supports as in **Paper II**. Values closer to 1 correspond to reproducible phases; below the cutoff value, phases are not reproducible. If the PRTF curve is always higher than the cutoff, we can estimate the resolution from the maximum angle of collected data.

The FSC curve represents the correlation between the two reconstructions as a function of the spatial frequency. As for the PRTF, the spatial frequency at which the correlation drops below a certain threshold (e.g., for FSC, the 1/2-bit threshold is more commonly used, which changes with wavelength to account for the increasing number of pixels contributing to each shell) is often used to define the resolution limit. A higher FSC value at a given spatial frequency indicates better agreement between the two reconstructions and, thus, confidence in the resolved features at that resolution.

Part III:
Deep Learning

4. Machine Learning and Neural Networks

Consider categorizing a dataset of images into two groups: one containing photos of dogs and the other of cats. One might initially think of designing an algorithm to identify specific features in an image, such as the muzzle's shape or the animal's size. However, accounting for every possible angle, breed, and detail distinguishing dogs from cats would be incredibly daunting. The challenge lies in the fundamental difference in how humans and machines perceive images. Yet, the scientific objective remains: to enable machines to differentiate between the two, recognizing which image feature corresponds to which class. *Machine Learning* (ML) approaches this challenge by allowing the data to lead the problem-solving process. In particular, *Neural Networks* are powerful ML algorithms that are extremely popular and common in Artificial Intelligence (AI).

4.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) are mathematical models historically created to simulate the behavior of the brain [65]. This was done by first attempting to identify the essential characteristics of neurons and their interconnections. Neurons are made up of dendrites, synapses, and axons. A neuron collects signals from other cells through its dendrites and emits electrical impulses through its axon, which is divided into several branches at the ends of which the synapse can be found. When a neuron receives a signal, it sends an electrical activity impulse along the axon through the synapse. Translating all this to our context, artificial networks are composed of interconnected units, which correspond to neurons and are called *nodes*. At the same time, the function of the synapse is simulated by a modifiable weight associated with each connection. Each unit transforms the received signals into a single output signal, transmitted to the next units. This transformation is done by multiplying the incoming signal by the weight of the connection, and then by summing up all the obtained results. By tuning the weight parameters, the node thus specializes in recognizing certain stimuli and "learning" them.

4.1.1 Perceptron

In ML, the *perceptron* was the first model of ANN ever introduced. This algorithm, also known as the *McCulloch-Pitts neuron* after its inventors, who developed it in 1943 [66], is often used as a pedagogical introduction for ANNs.

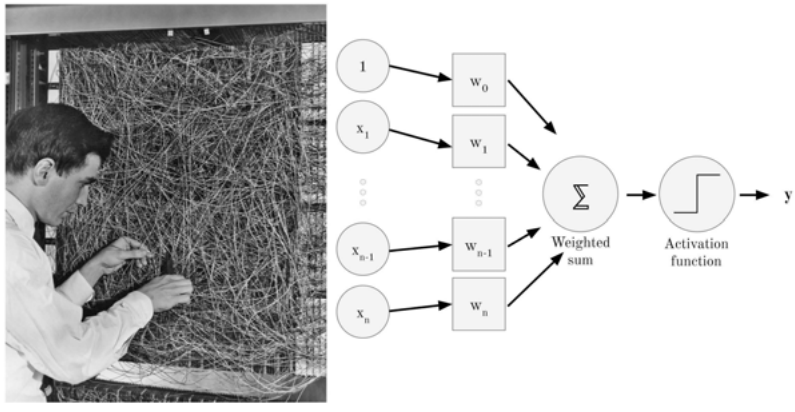

Figure 4.1. On the left, Frank Rosenblatt is working on the first Perceptron machine, Mark I. On the right, a simplified scheme of the perceptron.

The first model, the Mark I Perceptron machine, was built in 1957 by Frank Rosenblatt [67] at the Cornell Aeronautical Laboratory. The perceptron is a binary classifier determining whether a given input belongs to a particular category.

Given an input array $x_1, x_2, ..., x_n$, the perceptron produces a single binary output. The weights w_i corresponding to each input x_i must be combined to evaluate this output. As mentioned, this output is binary, thus it is either 0 or 1, depending on whether the weighted sum of the inputs is less than or greater than a threshold value that depends on the neuron.

The output of a perceptron can thus be calculated using the following equation:

$$y = f\left(\sum_{i=1}^{n} w_i x_i - b\right) \tag{4.1}$$

where y is the perceptron's output, f is the activation function, which is a step function in the case of a perceptron, w_i represents the weight corresponding to the i-th input, x_i represents the i-th input, b is a scalar free parameter, called *bias*.

The step function f can be explicitly defined as:

$$f(u) = \begin{cases} 1 & \text{if } u \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

The response of each neuron, or node, contributes to the output. The parameters are adjusted to converge towards increasingly accurate and precise prediction. The set of weights and the bias value represent the information the node learns during training and retains for subsequent reuse.

The perceptron is historically an important model, but even then, it was very limited; a famous book, *Perceptrons: an introduction to computational geometry* [68] by Marvin Minsky and Seymour Papert, showed that these networks couldn't learn the XOR function (or exclusive disjunction). It was already known that multilayered perceptrons, perceptrons for which the output is fed as an input to another perceptron, were not subject to this criticism. Still, no one knew how to train multilayered networks at that time.

4.1.2 Loss function

The *loss function* typically represents a measure of the error made by the ML model, in other words, the difference between the model's predictions and a known expected update, often called the *ground-truth*. Training can thus be seen as the process through which the model progressively adjusts its parameters to minimize the error as much as possible.

An example is the case of a mean squared error:

$$\mathrm{MSE}(y,\hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.3}$$

Where $\hat{y}_i$ represents the *prediction* of the network, and y_i is the ground-truth. This measure is intuitive and allows for identifying positive and negative deviations. Beyond Mean Squared Error (MSE), numerous other loss functions are tailored for various machine learning tasks, such as binary cross-entropy, mean absolute error, and Kullback-Leibler divergence, to name a few, each with its specific applications. The optimization of the loss function happens in two steps: *forward propagation* and *backward propagation* (or *backpropagation*).

4.1.3 Forward and Backward Propagation

Forward propagation is the process by which input data is passed through a neural network layer by layer, calculating the output of each layer using the associated weights, biases, and activation functions. During each iteration or *epoch* of the training phase, the ANN sees *batches* of the training dataset and, for each batch, calculates a prediction $\hat{y}$ for which the loss function is evaluated. Every time the network goes through the entire dataset, we say one epoch has passed.

The minimization process of the loss function is not straightforward, especially for deep networks with multiple layers. The backpropagation algorithm allows modern neural networks to calculate the gradient of the loss function and update the weights accordingly. Backpropagation propagates errors backward through the network—from the output layer to the input layer evaluating the contribution of each parameter to the overall error. The gradient of the loss

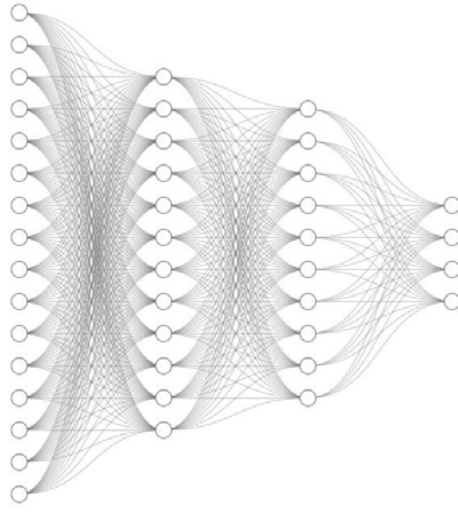

Figure 4.2. Schematic drawing of a deep neural network with two hidden layers. The input layer with 16 nodes is mapped to an output of dimension 4. The total number of parameters, weights, and biases that can be trained represent the network complexity. The number of trainable parameters (weights and biases) is the sum of the product of the number of nodes in each layer multiplied by the number of nodes in the next layer, plus one bias parameter per node. For example, in this network, we have 378 parameters.

L is calculated with respect to weights $W^{(l)}$, the biases $b^{(l)}$, and the activation functions $f^{(l)}$. This process is crucial for multilayer perceptrons because it allows for tuning weights across multiple layers. Stacking multiple perceptrons enables neural networks to approximate more complex functions. The backpropagation algorithm involves computing the partial derivatives of the loss with respect to the weights, which indicates how weight changes would impact the loss. This is often achieved with optimization methods such as *stochastic gradient descent*.

4.1.4 Deep neural networks

Multilayered neural networks can, in principle, approximate very complex functions. In computer science, several theorems that prove which class of functions can be efficiently estimated are known as *universal approximation theorems* [69]. The key ingredients that enable neural networks to represent various functions are non-linear activation functions and "depth". Depth refers to using multiple layers in the network, where one layer is used as input to the next layer; this is where the term *Deep Learning* comes from. Layers other than the input and the output layer are often called *hidden layers*. Given an input vector $\mathbf{x}$, each layer l of a deep neural network acts as follows:

$$\mathbf{h}^{(l)} = f^{(l)}\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{4.4}$$

where $\mathbf{h}^{(l)}$ is the output vector of layer l, $f^{(l)}$ is the activation function for layer l, $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the weight matrix and bias vector for layer l, respectively, and where

$$\mathbf{h}^{(0)} = \mathbf{x}$$

is the input to the network. The output of the network, $\mathbf{y}$, is given by the final layer's output:

$$\mathbf{y} = \mathbf{h}^{(L)} \tag{4.5}$$

where L is the total number of layers. These layers, where each node is connected to the node of the previous and following layers, are called *dense layers*.

4.1.5 Activation Functions

In each node, inputs are weighted and summed, then passed through an activation function. This function maps the input to the output and helps the neural network learn complex relationships in the data. Some common activation functions include:

- the sigmoid function

$$\sigma(h) = \frac{1}{1+e^{-h}} \tag{4.6}$$

- the hyperbolic tangent function

$$\tanh(h) = \frac{e^{h}-e^{-h}}{e^{h}+e^{-h}} \tag{4.7}$$

- the Rectified Linear Unit (ReLU) function

$$\mathrm{ReLU}(h) = \max(0,h). \tag{4.8}$$

4.1.6 Learning paradigms

During training, feedback algorithms calculate the loss function at the end of each iteration. There are two major categories of feedback for the algorithms:

- *Supervised learning algorithms* require pre-existing labeled data as those that the model will encounter in its applications. The model is given examples that pair the input data with the expected outcome and its task is to find a way to map the input to the correct prediction, minimizing a loss function calculated over the prediction $\hat{y}$ and the ground truth y. An

example is training a network for classifying pictures into N different classes while providing the network with the expected label.

- *Unsupervised learning algorithms* do not use labeled data. The loss function is defined only over the prediction $\hat{y}$ and the input without including labeled data. A common application of unsupervised learning is for dimensionality reduction.

4.2 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are one of the most commonly used neural networks in Computer Vision [70]. Convolutional networks employ *convolutions* to process images; their introduction was fundamental for applying neural networks in computer vision, where the inputs can be huge images or even videos, and processing each pixel as an independent node in the network would be extremely inefficient. The convolutional layer is, in fact, much less computationally demanding than mapping each pixel to a different neuron. Furthermore, the Convolutional layer extracts *feature maps*, which are significant characteristics of images used to calculate matches between key points during testing.

4.2.1 Convolution Operation

A convolution operation reduces the quantity of weights a model is required to learn. This is achieved by the systematic traversal application of a *kernel* across the image. The kernel is by an $n \times n$ matrix. In a CNN, each kernel represents a node, and we optimize the weights for these output nodes. Each layer of a CNN consists of a set of kernels, which are applied to the image to produce feature maps. In CNNs, a convolution operation is represented as:

$$(I * K)(i,j,k) = \sum_m \sum_n I(i-m, j-n) K(m,n,k) \tag{4.9}$$

where I is the input, K is the kernel, i, j are the spatial coordinates, and k is the number of feature channels. The convolution layer aims to identify accurately geometrical patterns in an image. The kernels in each layer are learned through backpropagation during training. In a purely practical but simplistic summary, it can be said that a CNN learns the values of these kernels during training, and the more kernels used, the more features will be extracted. The same kernel can be applied in different ways and several parameters can be adjusted to process feature maps differently. The most common free parameters are the *stride*, the *kernel size*, and the *padding*. The stride corresponds to the number of pixels by which we slide the kernel over the input image

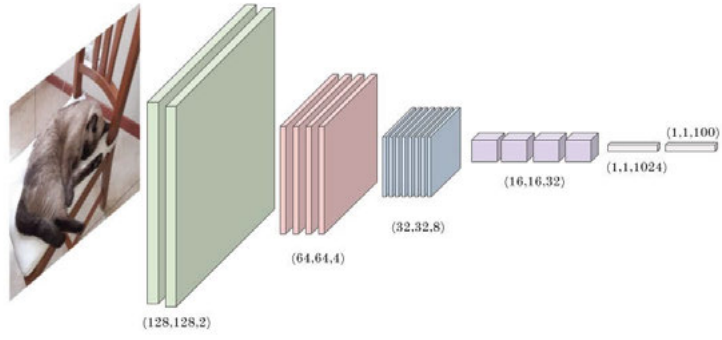

Figure 4.3. Example of the convolutional neural network, showing convolutional layers, max-pooling, and fully connected layers. Each convolution processes the input image according to the kernel, creating feature maps. The number of feature maps is reported in the third dimension. Pooling decreases the dimension of the input layer; in this schematic figure, the application of a pooling layer is highlighted using a different color for the feature maps. In this scheme, the feature maps are then passed to two dense layers with 1024 and 100 nodes each, respectively. This CNN could be trained, for example, for classifying pictures into 100 different classes (i.e., *dogs*, *cats*, *birds*, etc.), each corresponding to one dimension of the output layer.

before calculating the output. A stride equal to 1 moves the kernel of one pixel over the input image. The kernel size is simply the dimension of the kernel matrix; big kernels reduce the computational complexity of the networks. The *padding* involves adding extra pixels around the edge of the input image before applying the convolution operation. The number of convolutional layers can vary depending on the network architecture; the greater their number, the more complex the features they can identify. Changing the kernel with the same input image will result in different feature maps; an example of the effect of different kernels is given in Figure 4.4.

4.2.2 Pooling Operation

Pooling layers downsample the dimensions of the input layer. Different pooling methods exist, among which the most famous and used is *max-pooling*, which returns the largest element of the map within a chosen "window". This layer makes the input of the features smaller and more manageable by reducing their dimension, hence their complexity and consequently speeding computation; it also makes the network invariant to small transformations such as distortion or translation relative to the initial image. Max-pooling can be

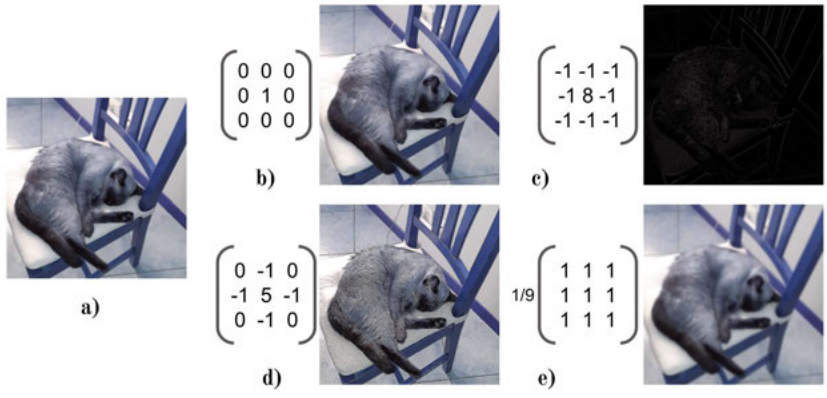

Figure 4.4. Example of the effect of different kernels on an input image, each matrix refers to a different kernel. Each of these kernels was applied on the input image using the following parameters: a stride equal to 1, a kernel size of 3×3, and no padding. a) Ground-truth. b) Effect of the "identity" kernel applied to the figure a). c) The kernel used here can be helpful for edge detection. d) This kernel sharpens the image. e) This kernel blurs the image.

mathematically expressed as follows:

$$O(i,j,k) = \max_{m,n,k} I(i \cdot s + m, j \cdot s + n, k) \tag{4.10}$$

where O is the output, I is the input, s is the stride, and m,n iterate over the spatial extent of the pooling window, and k is the number of feature channels.

4.2.3 Dropout

Dropout refers to the process of ignoring certain nodes during the training phase by excluding different nodes at each epoch. It acts as a form of regularization by disconnecting some nodes from the network to reduce or prevent overfitting. This method is especially common in the hidden layers furthest from the input and output layers, where interpreting the weight parameters is particularly challenging. The application of dropout results in a different, reduced version of the starting model at each epoch.

5. Scientific applications of CNN Models

This section discusses some models used or tested for **Paper I** and **Paper II**. All the models were developed in *Keras* [71] using Tensorflow [72] in the backend.

5.1 Encoder-decoder

A generative *encoder-decoder* network (**Paper II**) with 14 trainable layers was trained to predict protein electron densities from diffraction data. Encoder-decoder models are network models that use pooling layers to shrink the information content into a bottleneck of dense layers. The model can be divided into two sections, encoder and decoder, connected by the bottleneck. The encoder squeezes with pooling layers the input into a collection of feature maps obtained from convolutional layers. These maps are subsequently upsampled and processed by the convolutional layers of the decoder. Our architecture, detailed in Table 5.1, uses 3×3 convolutional kernels with a stride of 3 and applies 2×2 max-pooling after every two convolutional layers. Starting from a 128×128 input, the bottleneck reduces dimensions to 16×16 with 256 feature maps. The bottleneck comprises two dense layers (128 nodes and 65536 nodes). ReLU activation functions are used for all layers, while the output layer employs Tanh to constrain the output range.

5.2 U-Net

U-Net is a convolutional neural network architecture introduced by Olaf Ronneberger in 2015 [73] for image segmentation. U-Net is a CNN encoder-decoder model in which connections that skip layers between the encoder and decoder are added to the network; supplying the decoder with features from the encoder improves the training. Providing the decoder with features from the encoder that can bypass the bottleneck helps the convergence during network training. The network is effective in various image segmentation tasks and was invented explicitly for medical imaging and the segmentation of cells, brain tumors, and retinal blood vessels. It is also often employed as an image processing method and as a generator in *generative adversarial networks*.

Layer (type)	Output Shape	Param #	Description
InputLayer	(128, 128, 1)	0	Input Pattern
Conv2D	(128, 128, 64)	640	64 feature maps, with ReLU activation function
Conv2D	(128, 128, 64)	36928	64 feature maps, ReLU
MaxPooling2D	(64, 64, 64)	0	2x2 Pooling
Conv2D	(64, 64, 128)	73856	128 feature maps, ReLU
Conv2D	(64, 64, 128)	147584	128 feature maps, ReLU
MaxPooling2D	(32, 32, 128)	0	2x2 Pooling
Conv2D	(32, 32, 256)	295168	256 feature maps, ReLU
Conv2D	(32, 32, 256)	590080	256 feature maps, ReLU
MaxPooling2D	(16, 16, 256)	0	2x2 Pooling
Flatten	(65536)	0	Flatten Layer
Dense	(128)	8388736	Fully Connected Layer, ReLU
Dense	(65536)	8454144	Fully Connected Layer, ReLU
Reshape	(16, 16, 256)	0	Reshape Layer
Conv2DTranspose	(16, 16, 256)	590080	256 feature maps, ReLU
Conv2DTranspose	(16, 16, 256)	590080	256 feature maps, ReLU
UpSampling2D	(32, 32, 256)	0	2x Upsampling
Conv2DTranspose	(32, 32, 128)	295040	128 feature maps, ReLU
Conv2DTranspose	(32, 32, 128)	147584	128 feature maps, ReLU
UpSampling2D	(64, 64, 128)	0	2x Upsampling
Conv2DTranspose	(64, 64, 64)	73792	64 feature maps, ReLU
Conv2DTranspose	(64, 64, 64)	36928	64 feature maps, ReLU
UpSampling2D	(128, 128, 64)	0	2x Upsampling
Conv2DTranspose	(None, 128, 128, 1)	577	Output layer, Tanh activation function

Table 5.1. *Summary of the encoder-decoder network.*

5.3 Generative Adversarial Network

Generative Adversarial Networks (GANs), developed by Ian Goodfellow in 2014 [74], are neural network models that train two parallel models in competition: the *generator* and the *discriminator*. The generator generates data that resemble the ground truth from the dataset. Usually, the input of this model is noise for purely generative tasks, but it can use other images to sample and create a new output. The discriminator, on the other hand, acts as a classifier. Its role is to differentiate between genuine examples from the dataset and fake examples generated by the generator. It is trained on "real", coming from the training dataset, and "fake" data, produced by the generator, and learns to distinguish between them. Initially, each model is trained independently over a different loss function, and then the pre-trained models are trained together by minimizing a shared loss function. The shared loss function in GANs facilitates effective training by enabling adversarial learning, stabilization, and convergence of both the generator and discriminator networks. The goal is to improve the generator to produce data that is indistinguishable from real batches from the dataset for the discriminator. This balance is delicate, and the overfitting of the model can cause the GAN to generate significantly worse outputs.

6. Reinforcement Learning

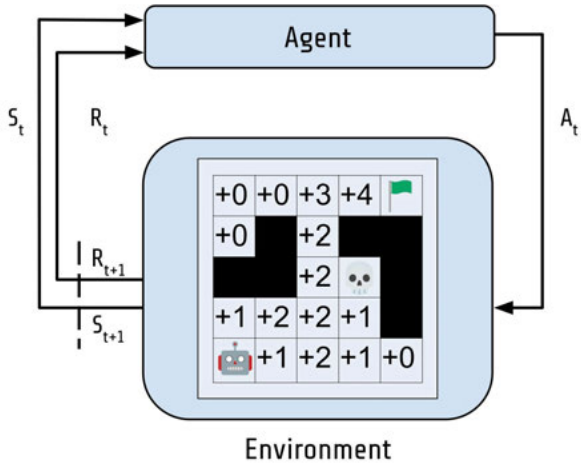

Figure 6.1. Reinforcement learning scheme. An agent is tasked to reach the green flag without ending on the skull. The agent can take action at each time point, by choosing, in this scenario, where to move the robot emoji, for which a reward will be calculated, as shown in the depicted values on the table. If the agent lands on the green flag or the skull, the algorithm stops, and the cumulative reward is calculated. Part of the reward system can also include returning a high reward for reaching the flag in the least amount of actions. Each action corresponds to a movement to the next tile, and the black tiles are not accessible states. At each state, corresponding to the position of the robot emoji on the board, the agent is provided with some information (*observation*) on the environment; for example, it can be aware of only the next accessible states or of all the next accessible states and even the position of the flag. The model is trained to maximize the cumulative reward.

Reinforcement learning is a subfield of ML, where an *agent* learns dynamically how to operate in an *environment* by maximizing a cumulative reward. The agent is usually software, which can, by trial and error, learn to optimize actions in a given state. Very famous applications of RL include optimizing neural network models to play games, such as chess or Go. The learning process occurs through the agent's ongoing interaction with its environment, during which it receives information about its current state. Based on the agent's actions on the environment, the agent takes a reward value that is provided. The word environment refers to everything the agent can interact with. At any

time step t, the agent and the environment find themselves in a particular state s; see an example in Figure 6.1. From this state, the agent can execute an action, leading to a state-action pair, which can be a discrete action, such as a position on a board, or continuous, such as a parameter value up to some decimal point. Before advancing to the next time step $t+1$, the agent is rewarded and carried over to the following state. The decision-making strategy of the agent, or its "policy", enables it to select actions based on the current state. In reinforcement learning, the agent is trained dynamically, such that thanks to a reward system, it can learn to pick an action based on past and future state-action pairs. A reward is computed for each possible action, and the next one is chosen stochastically by prioritizing the maximum reward. During the training, the agent is often allowed to explore and see the reward outcomes of future action-state pairs. One way to tackle agent training includes keeping track of state-action pairs and their reward. In real-world applications, maintaining a table mapping each state-action pair to its value becomes impractical due to the many possibilities if the state space is high-dimensional. To overcome this, neural networks can be used to estimate the value of actions in specific states. In this context of Deep Learning, Deep Q-Networks (DQN) are neural networks trained to maximize the reward for taking an action in a given state. The seminal *Playing Atari with Deep Reinforcement Learning* showcased the potential of DQNs by applying them to playing Atari 2600 games, where the model learned to play these games at a level comparable to, and sometimes surpassing, human experts. This approach allows for handling environments with high-dimensional state spaces, which are challenging for traditional RL algorithms. An attempt to use DQNs for solving the phase problem is discussed in Section 7.4.

Part IV:
Results

7. Neural networks for FXI

The development and success of neural networks in solving many Computer Vision tasks have a cascade effect on all imaging methods, and their implementation can assist experiments. Problems such as data classification, denoising, and post-processing can be tackled with neural networks. This chapter elaborates on a few topics where neural networks could assist in analyzing FXI diffraction data. For training the models in **Paper II** and **Paper III**, I simulated a dataset of amplitudes, phases, diffraction signals, and densities for proteins from the Protein Data Bank (PDB).

7.1 The dataset

My dataset of 10,900 proteins was simulated on PDB files using *Condor* [61] and *Chimera* [75]. To generalize my deep learning models as much as possible, I simulated data from various molecules of different shapes and sizes, selecting PDB structures with molecular weight within the range of 10 kDa - 100 kDa and simulated diffraction data under different conditions of noise and masks.

7.2 Data healing

In an FXI experiment, each diffraction pattern is typically very noisy, and photons are often not collected over the entire field of view due to detector masks. Denoising and filling in the missing data can potentially help in real-time data analysis of experiments, such as hit finding, diffraction pattern classification, and other experimental diagnostics. The study presented in **Paper I** aimed to train a U-Net model to denoise and demask diffraction images from X-ray single particle imaging data. U-Net was trained over 9900 structures to perform denoising and demasking. The remaining 1000 structures have been used to test the neural network's performance. Further tests using simulated patterns from PDB structures of viruses and geometrical objects outside the molecular range of the dataset were used to prove the model's reliability. The U-Net was benchmarked to a low-pass filtering algorithm based on autocorrelation constraints. By denoising and demasking patterns from simulated data of a phytochrome protein (PDB-ID 4O01 [76]), U-Net also proved successful in achieving a faster and more reliable convergence for orientation recovery with the Expansion-Maximization-Compression algorithm (EMC).

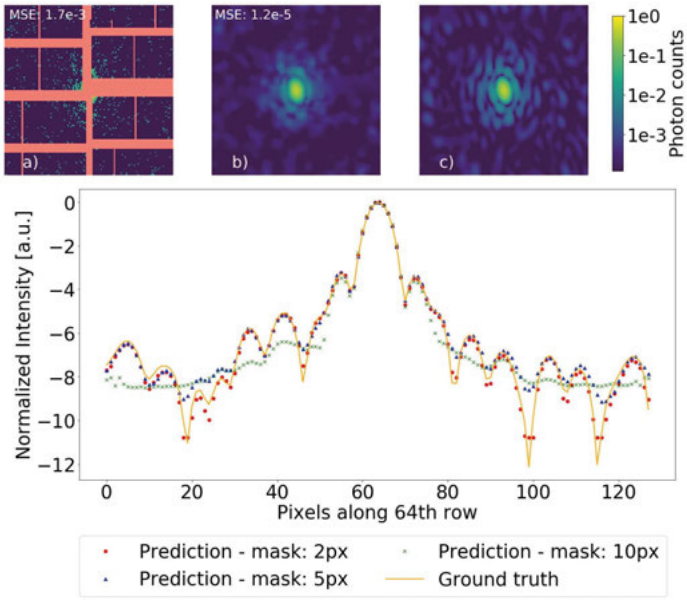

Figure 7.1. Top figure: a) simulated diffraction from PDB file 1HJG [77] with AGIPD geometry mask, b) U-Net denoised version of the diffraction pattern, c) ground-truth. Bottom figure: profile of recovered intensities along the central row for PDB file 4ZS0 [78] in logarithmic scale, with various mask sizes. As mask size increases, the output's quality worsens, struggling for most patterns when the entire central speckle is missing.

Table 7.1. *Mean squared errors for diffraction intensities of geometrical shapes and viral particles as in **Paper I**. The dataset was originally simulated using a detector distance of 0.3m. For viral particles, it was necessary to adjust the detector distance (reported near the object name in the table when changed) to fit the central speckle of the pattern within the field of view of the detector. The network performance is still within the expected values; for reference, the average MSE in our test dataset is* $1.2 \cdot 10^{-5}$*, and the PSNR is 49.2.*

Object	MSE	PSNR
Sphere	5.69e-06	52.4
Cube	1.04e-05	49.8
Icosahedron	7.34e-06	51.3
PDB:1WCD (2.67 m)	1.71e-05	47.7
PDB:1WCD (2.00 m)	5.10e-06	52.9
PDB:3CJI (2.67 m)	7.06e-06	51.1
PDB:1P58 (2.67 m)	5.28e-06	52.8

The key achievements of this work include:

1. **Data healing**
   The denoising task becomes more challenging as the signal decreases, and the autocorrelation filter outperforms U-Net when the signal is strong but does not work with missing data at a low angle. With U-Net, filling missing data from detector masks is possible and was also proved using an AGIPD [59] detector mask. Experience shows that to achieve a good recovery, we require a mask size smaller than half of the central speckles, as shown in 7.1.
2. **Robustness to new data**
   Additional diffraction intensities were simulated from geometric objects and virus-like nanoparticles, each producing distinct speckle structures, different from the training dataset. These simulations were performed under different experimental setups, preserving oversampling by adjusting detector distances. The network performed well as it did for the test dataset, showcasing its robustness to new data. The network's performance was evaluated across various scenarios by calculating the peak signal-to-noise ratio (PSNR) and the MSE for each case, as shown in Table 7.2.
3. **Enhancing orientation recovery**
   Restoring diffraction intensities effectively reduces the number of patterns needed for EMC convergence. As more restored data becomes available, EMC converges reliably, see Figure 7.2. As mentioned in Section 2.4.2, another version of EMC was used, here called *Blurred EMC*. This EMC variant uses a 3D Gaussian kernel to distribute each pixel value in the patterns over many voxels after every iteration of EMC. Blurred EMC was tested on noisy masked data and required fewer patterns to work reliably, yet twice as many as EMC when applied to restored patterns. Furthermore, Blurred EMC tends to slightly increase rotation error, unlike standard EMC, which does not exhibit this issue.

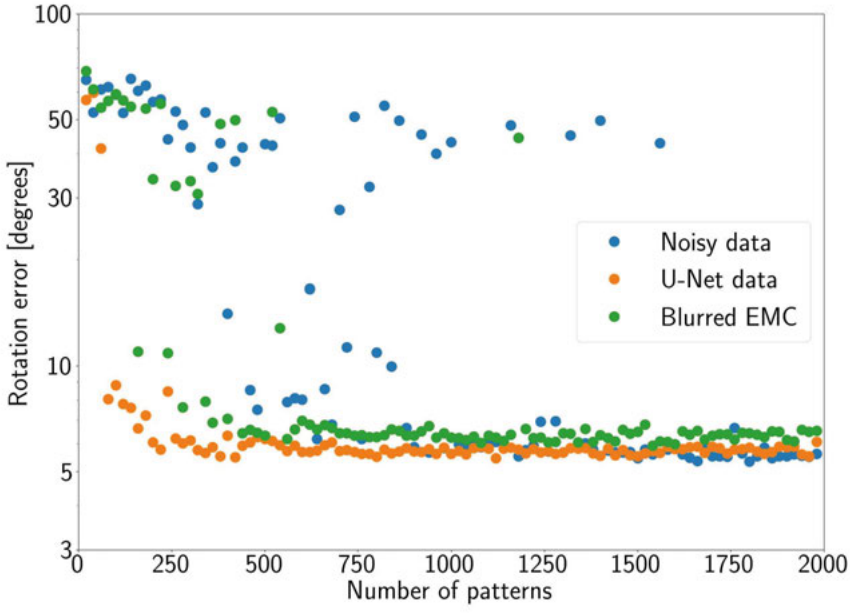

Figure 7.2. Mean rotational error calculated from EMC, EMC applied to restored intensities, and blurred EMC. Blue, orange, and green markers refer, respectively, to: EMC applied to masked and noisy data, EMC on restored intensites, Blurred EMC. While Blurred EMC converges to a good 3D Fourier space model with just 500 patterns, it also introduces a higher mean rotational error. On the other hand, EMC applied to restored data requires only 250 patterns to achieve a lower mean rotational error.

These achievements prove that our neural network model can effectively restore diffraction intensities from unknown structures. Orientation recovery with healed data is more dependable, particularly on smaller datasets. This approach is adaptable to different experimental conditions, has a short training period, and is robust to new data. Further improving U-Net, including background signal in our simulation, may prove to be a valuable tool for real-time data analysis during beam times.

7.3 Phase retrieval and support finding

All phase retrieval algorithms necessitate real space constraints to work. This means starting from an initial guess, usually given by the autocorrelation function or a low-resolution map obtained in a previous experiment. Several recent results reported the successful application of machine learning algorithms for phasing uniform shapes from convex nanocrystals [79, 80, 81]. In **Paper II**, I

trained the convolutional neural network (see Section 5.1) to map diffraction patterns to normalized protein densities. We also used these guesses as support constraints during phasing. Theoretically, estimating a set of tight supports is possible by intersecting copies of the autocorrelation function thanks to the triple intersection method. We benchmarked our neural network guesses using the triple intersection method while testing the neural network guesses to improve and speed up the phasing process. These achievements extend previous results from DL models for phasing nanocrystals' convex data to protein data, as shown in Figure 7.3 and Figure 7.4.

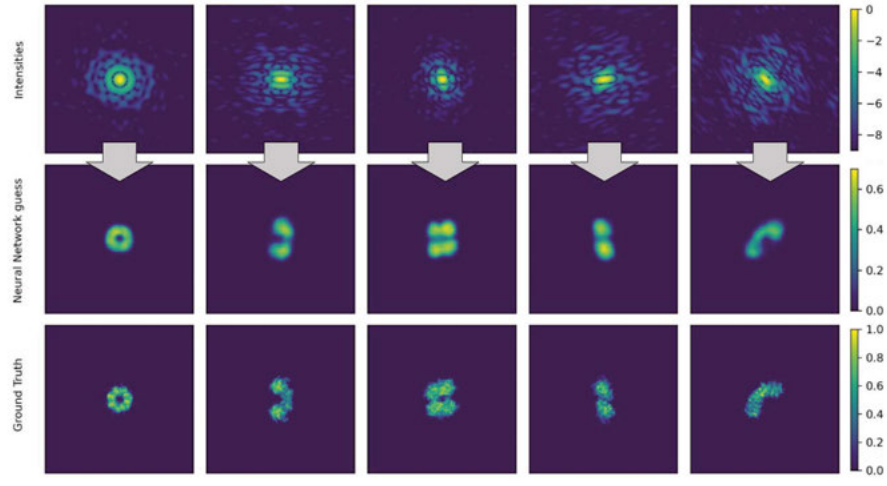

Figure 7.3. Representative reconstructions with the encoder-decoder model. The first row from the top displays diffraction patterns, the middle row shows density estimates generated by the neural network model, and the last row shows the actual protein densities. PDB-ID in order from left to right: 1SS8 [82], 3ETA [83], 5IRC [84], 3EPP [85], 5KLT [86]. It is interesting to notice how the neural network model can estimate non-uniform densities, correctly guessing the hollow barrel shape of GroEL (protein from file 1SS8, in the first column).

The key achievements of this work include:

1. **Support finding**
   Both the DL and the triple-intersection methods allow a break of symmetry invariance in some cases, which can help achieve higher resolution when phasing. Comparing the number of pixels that do not belong to the ideal supports, our evaluation indicates that the encoder-decoder network offers reliable results an.
2. **Improved phase retrieval transfer function**
   Restoring diffraction intensities significantly reduced the number of patterns required for orientation recovery. We, therefore, used U-Net from the previous paper [87] to calculate "healed" data and fed the result to the

encoder-decoder model. We then used the density estimate to calculate a support mask for phasing. The PRTF (Phase Retrieval Transfer Function) behavior favored the HIO+ER method when a good guess is provided, showcasing its stability and performance on par with shrinkwrap starting from a much bigger support. This indicates the potential of neural networks for phasing even by just providing a size estimate of the object, as it serves as a valuable starting point for, for example, breaking the centrosymmetry and translation invariance of the solution, reducing the number phasing minima.

3. **Phasing speed**
   Remarkably, using the output estimates, our model demonstrated a notable speed improvement in phasing, reducing the convergence as in Figure 7.5. This development hints at exciting possibilities for tackling challenging phasing patterns.

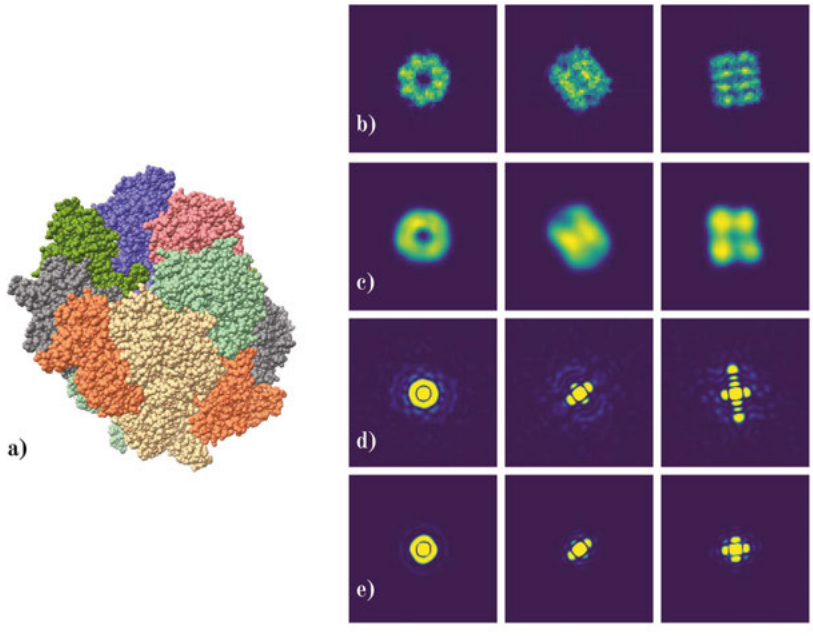

Figure 7.4. More examples of reconstructions from different orientations of GroEL. a) 3D GroEL model from PDB file 1SS8 [82]. b) GroEL densities at different orientations were simulated using the PDB file. c) Electron densities were estimated using the encoder-decoder neural network. d) Input diffraction intensities. e) Diffraction intensities calculated from the estimates shown in c).

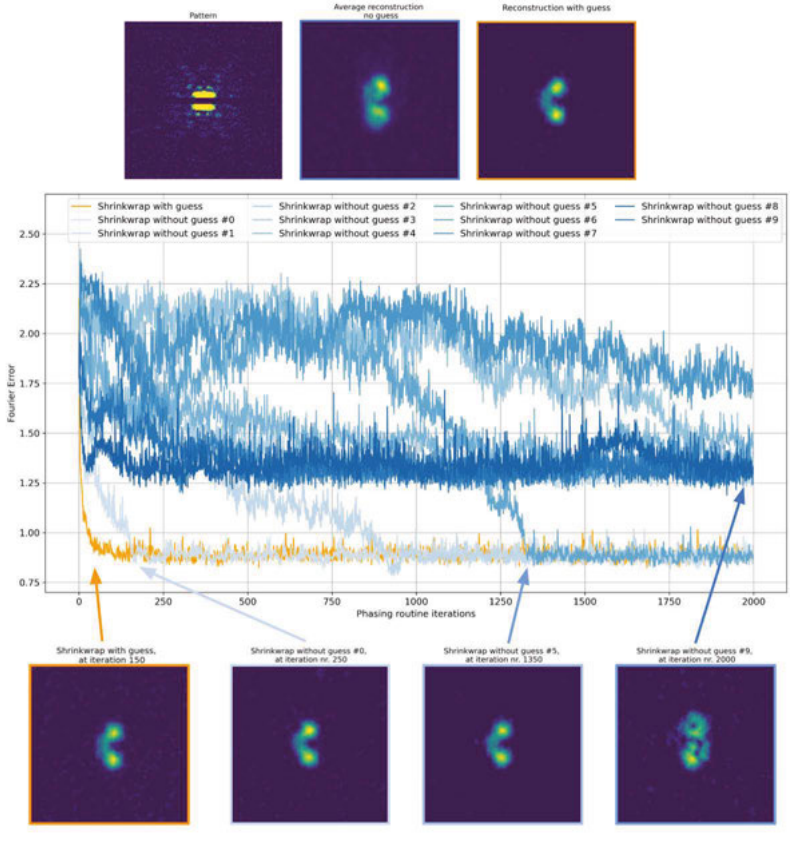

Figure 7.5. Fourier error for 10 reconstructions of PDB-ID 3ETA [83] and for one reconstruction initialized with an initial guess from the neural network. By seeding the HIO+ER routine with a density estimate, phasing achieves stable convergence faster. On the other hand, random initialization of densities in many cases remains stuck to local minima.

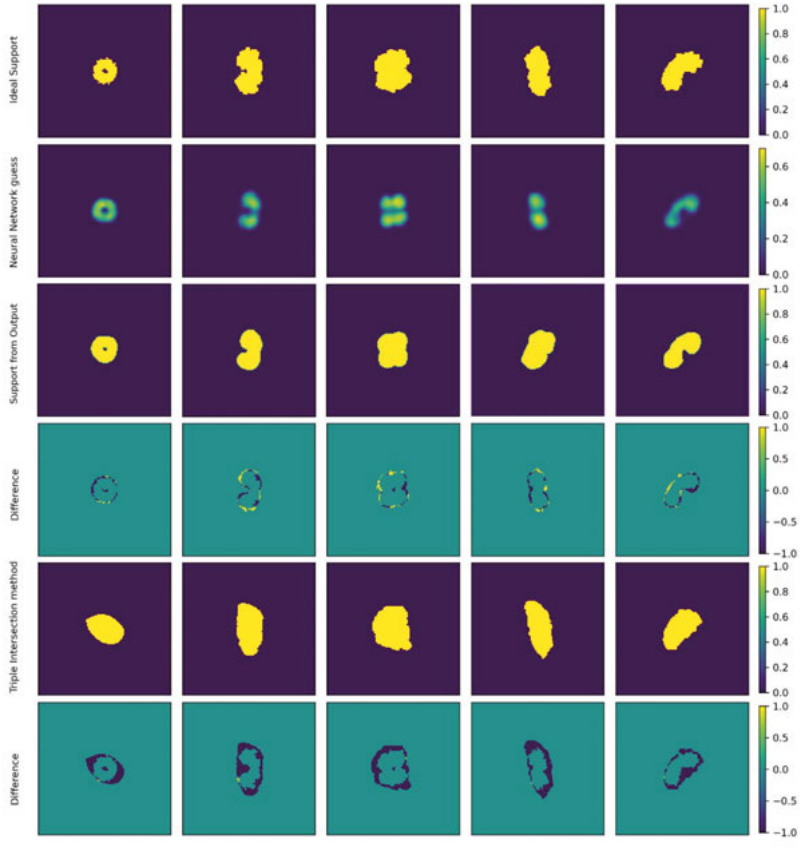

Figure 7.6. Representative reconstructions as in Figure 7.3. The first row from the top displays the ideal support calculated by thresholding the ground truth. The second row shows density estimates generated by the neural network model, while the third row shows the support masks generated by thresholding the corresponding estimates. The fourth row shows the difference calculated by subtracting the support masks from the ideal support. The fifth row shows support masks estimated using the triple-intersection method. The last row shows the difference calculated by subtracting the estimates calculated with the triple-intersection method from the ideal support.

7.4 Phase retrieval optimization

Phase retrieval is a complex process that requires careful selection of algorithm parameters. Ideally, the error decreases with time as it reaches convergence. A typical issue with these algorithms is that they often stagnate into a local minimum. Although HIO and other phasing algorithms exist to moderate this issue, the stagnation problem is still the major cause of failed reconstructions. The shrinkwrap algorithm refines the support for every N-th iteration and the final result strongly depends on the initial conditions. Typically, phasing algorithms

are run in parallel several times to retrieve many independent solutions from different initial conditions; the final solution can then be obtained by averaging across the solutions with the lowest error. Nonetheless, phase retrieval is still very sensitive to the choice of algorithm parameters. The pixel threshold and the Gaussian standard deviation parameters used by Shrinkwrap are usually set a priori at the beginning of the phasing process and kept static during each iteration. I integrated Reinforcement Learning (RL) into the Shrinkwrap algorithm to explore the possibility of parameter optimization during the phasing process. This project is somewhat inspired by *memetic algorithms* [53], where the best reconstruction is chosen over several instances of the same phasing routine run over different parameters. I used the open-source libraries *Gymnasium* [88] to develop the environment and agent-environment interaction and *Keras-RL* [89] to optimize the agent. I aimed to independently optimize the choice of the negative feedback of HIO β, and the Gaussian standard deviation parameter. The agent was provided phases at each iteration of the phasing algorithm while the state was given amplitudes. The error function of ER determined how the reward was calculated. The β and Gaussian parameters were treated as actions in the RL framework, and the RL agent learned to adjust these parameters to maximize the reward. By doing so, the accuracy of the phase retrieval process was expected to be improved by dynamically adjusting the algorithm parameters. The phasing process consisted of a HIO+ER Shrinkwrap routine, selecting a new parameter every 100th iteration. Each HIO+ER routine consisted of 100 iterations of HIO followed by 5 ER. The phasing routine would run 50 times, allowing the agent to pick a new parameter 50 times during the phasing process. Then, the phase retrieval would restart from scratch. The DQN agent is a short CNN, as depicted in Figure 7.7. Although this application was promising, the implementation proved to be more challenging than expected. The DQN model manages to return only one static parameter and does not follow any interesting strategy. Several factors may have played a role.

1. **Impact of the parameters on the reconstruction**
   While using good parameters is essential for the convergence of HIO and ER, the impact of the parameters is smaller than expected (especially for β, the negative feedback of HIO). The algorithms can still end in local minima when a good parameter is chosen, and the reward system based solely on the Fourier error of ER might be insufficient for training the model.
2. **Restricted observation space**
   The only information given to the agent is the state of amplitudes and phases during the training. Due to several local minima, determining an interesting strategy for which action to take for the next 100 iterations of HIO and 5 of ER, other than a static choice of the parameters for the entire phasing process, could be too challenging.

3. **Insufficient training time and Neural Network depth**
   RL training requires several epochs or, in the RL jargon, *episodes* to find an optimal strategy. The neural network used as an RL agent was shallow to speed up the training time, but the bottleneck was given by the time required for the phasing routine. A shallow network might be responsible for the lack of complexity required to distinguish the states of different phases to solve the phasing problem effectively. On the other hand, a deeper network would have been very time-consuming to train.

Despite the outcomes, this approach to phase retrieval remains intriguing. As shown in recent publications, where memetic algorithms [53] enabled to probe the solutions landscape of phasing algorithm, new developments in this direction could open new ways to tackle the phase problem. More investigation on the use of RL in CDI might lead to interesting results, as it has already found applications in optimizing the beamline parameters of XFELs [90].

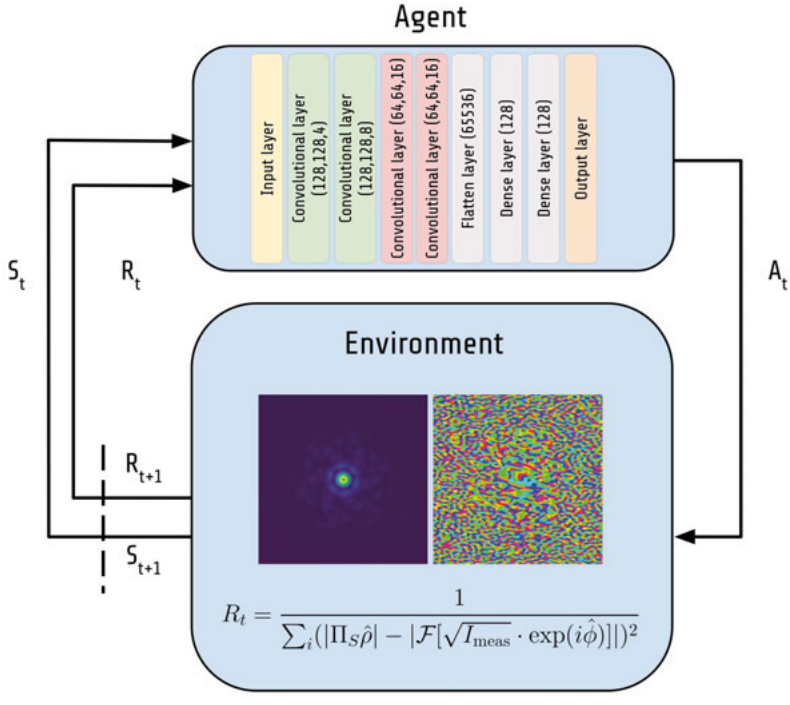

Figure 7.7. Scheme of the phasing environment with a CNN-agent. When the model was trained to optimize the blur, the action was restricted within a value of 0 to 5 pixels of standard deviation of the Gaussian distribution. For the β parameters, the action was chosen within the range 0 to 1. The reward was calculated as the inverse of the error metric, and each state corresponded to a pair of amplitudes-phases.

8. Ab-Initio modeling of AsLOV2

In the last chapter, I discuss my application of DENSS on TR-SAXS data. **Paper III** introduced an innovative application of time-resolved solution X-ray scattering using an XFEL. Thanks to XFELs' unique internal bunch structure, which allows us to achieve microsecond time resolution, capturing data with exceptionally low noise levels was possible. These developments enabled the study of the conformational pathways of the Light-Oxygen-Voltage (LOV) domain from *Avena sativa*, elucidating the structural dynamics of photodissociation and unfolding the J-α helix. Structural fitting was achieved by comparing several structures obtained from AlphaFold.

8.1 AsLOV2

Photosensory proteins underpin how plants react to changing environmental stimuli. Among these, phototropins, a type of blue-light photoreceptors, play a pivotal role in controlling vital processes such as the movement of chloroplasts, the direction of plant growth towards light (phototropism), and stomata opening. They achieve this by conveying light-induced molecular signals through a protein structure that undergoes a sequence of conformational changes.

AsLOV2 is a domain of phototropins which is involved in plant phototropism and is being studied for its application in optogenetics. Irradiation unfolds part (J helix) of AsLOV2 in 240 μs, but recent findings imply Aα helix shift crucial for Jα unfolding.

8.2 DENSS with enforced densities

So far, DENSS has never been used for molecules in excited states. In principle, DENSS can process data obtained from time-resolved experiments, but as the difference between the structure of the dark state and the excited structure is often small, it requires extra constraints. Using the scattering data of the excited structure as a direct input to DENSS does not take into account the fact that the protein is into a mixture of states. During an experiment, an optical laser probes the molecules, but only a fraction of molecules (the *yield*) will be excited. For example, in **Paper III**, we estimated the yield to

be around $15 \pm 5\%$ for AsLOV2. For experimental data application, as originally envisioned for retrieving the densities of the unfolding AsLOV2, it is crucial to estimate the yield correctly before structural fitting. Enforcing the density of the dark structure, which is known from crystallography in the case of AsLOV2, could allow us to gain insights into the structural dynamics of molecules under photoexcitation. Although the application of DENSS was ultimately unsuccessful in reconstructing the unfolded structure of AsLOV2 from experimental data, it was possible to gain some useful insights into the feasibility of this idea. To characterize DENSS in a more controlled environment, its application was tested on simulated data from AsLOV2. In particular, I simulated a scattering SAXS profile from the PDB file 7PGX [91] of AsLOV2, up to 2Å^{-1}. Then, I simulated another scattering profile from the same protein with two missing amino acids from the amino-terminal Aα helix. My procedure to adapt DENSS to retrieve the densities of the two missing amino acids is described in the following steps:

1. Start with the electron densities map from the dark state.
2. Initialize random densities in the support within the area with missing/evolving structure.
3. Calculate the theoretical scattering of the excited state from the dark state structure.
$$I^{\text{excited}} = \frac{1}{\gamma}(I^{\text{dark}}_{\text{calc}} + \Delta I_{\text{exp}})$$
where γ is the photo yield of the protein (in our simulated case, we set $\gamma = 1$), $I^{\text{dark}}_{\text{calc}}$ is the simulated scattering from the dark state, and ΔI_{exp}) is the experimental difference scattering curve.
4. Apply restraints in real space - connectivity, positivity, and apply flat solvent
5. Compare to experimental data.
6. Iterate steps as in DENSS for densities within the support.

In our simulated case, the "excited" structure corresponds to 7PGX, and the dark structure corresponds to 7PGX without a glycine and a glutamic acid from the N-terminal Aα helix. Other forms of constraints should be possible, such as global and local density conservation $\sum(\rho_{out} - \rho_{dark}) = 0$.

A predefined or known "dark" support structure is essential, serving as a reference for fitting the scattering data. While DENSS is intrinsically a low-resolution computational method, enforcing the densities from 7PGX seems to allow the reconstruction of missing glutamic acid and glycine within the support, see Figure 8.1. Using DENSS on experimental AsLOV2 data is complicated because the densities in the excited state derive from an unfolding structure. Therefore, the SAXS intensities of the excited state likely originate from a mixture of several unfolded states, raising questions about the feasibility of accurate reconstruction under these conditions. Moreover, 22 amino acids out of 146 of the entire structure of AsLOV2 belong to the Jα helix. The density volume that needs to be reconstructed is quite large compared to

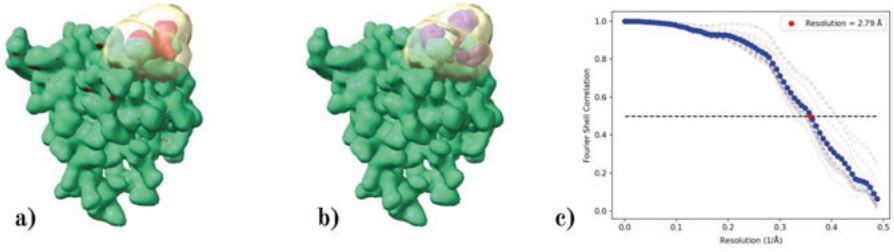

Figure 8.1. Example of AsLOV2 reconstruction using DENSS from simulated data. a) AsLOV2 densities in green, where the end of the Aα helix is emphasized in red. The support area is depicted in yellow. The region, which is kept static and enforced throughout the iterations, is illustrated in green. The two amino acids absent from the model are the terminal glycine and one glutamic acid. b) AsLOV2 with the density of the missing amino acid reconstructed in purple. c) Fourier shell correlation calculated after 20 reconstructions.

the enforced densities, and given that DENSS is inherently a low-resolution method, this exacerbates the difficulty of the reconstruction.

Outlook

It is fascinating to see how far we have come since the debut of the first soft X-ray laser, FLASH, in Hamburg in 2005. The XFEL community has been growing rapidly, with facilities being opened worldwide, such as LCLS in the United States, FERMI in Italy, SACLA in Japan, SwissFEL in Switzerland, and more. However, the European XFEL in Germany stands out with its unprecedented high repetition rate and brilliance. This massive leap has significantly improved our ability to collect experimental data but has posed new challenges, such as managing the massive amounts of data collected. Of course, there are also other challenges to overcome. For instance, much effort has recently been put into improving the X-ray beam focus. Also, new developments in sample delivery could significantly decrease the background signal, as current estimates envision the possibility of reducing the background signal tenfold by using Helium in the ESI injection. The future of X-ray Free-Electron Laser (XFEL) imaging is poised for significant advancements, with the potential to revolutionize our understanding of molecular and atomic structures. I want to conclude this thesis with a few perspectives on future scientific work. Following are a few possible future developments in the field.

Nanofocusing

XFEL nanofocusing systems utilize total reflection from Kirkpatrick-Baez mirrors to produce tightly focused beams. By focusing the beam, the number of photons per unit area increases, and so does our collected signal. Current research at EuXFEL promises a beam focus of roughly 200 nm, but in practice, a focal spot of roughly 1μm is still common during most beamtimes. Still, future work aims to reach focusing below 10 nanometers - a very recent publication just showcased a stunning 7 nm focus [92].

Helium ESI sample delivery

Recent developments in helium-ESI can enhance particle delivery by 10-fold and reduce gas scattering by 80%, promising higher quality and quantity SPI diffraction patterns in future FXI experiments [93, 94]. A lower background injection system can potentially be the missing ingredient to achieve sub-nanometer resolution on samples of the size of protein complexes.

FXI with in solution injection

Demonstrated at the European XFEL SPB/SFX nano focus beamline [95], liquid sheet jet exhibits superior performance over conventional liquid jets, offering advantages such as enhanced background stability, favorable radiation-induced explosion dynamics, and the potential for ultrafast single-particle solution scattering, making it a promising alternative injection system for single particle imaging.

Sub-nanometer resolution and dynamics

Currently, SAXS/WAXS and serial femtosecond crystallography are the only methods available to probe the dynamics of biological particles. Achieving sub-nanometer resolution, especially on dynamical systems, will finally allow FXI to compete with these experiments, providing a new tool to integrate into the current arsenal of structural biology imaging tools. A recent beamtime held in November 2023 has successfully collected data for photosystem I using FXI, marking a new milestone in the field. Although the data analysis of this beamtime is still ongoing, the preliminary results are promising.

Artificial Intelligence and Machine Learning

The integration of XFEL technology with advanced computational methods is set to help make the most out of the current data processing routines. Current achievements of ML in XFEL imaging include: efficient data classification for online analysis [96, 97], preprocessing to improve signal quality [98, 87], offline analysis and structural fitting in 2D and 3D for geometrical metallic nanoparticles [99, 100, 98, 101, 81] and for nanodroplets [102], and the optimization of beamline parameters [103, 90, 104]. The field of AI is vast and evergrowing, and the research landscape is developing too quickly to keep track of all technological applications. Yet, I speculate that in the future, the development of AI-integrated software into the beamlines might allow the processing of experiments on the fly, leading, for example, to real-time reconstructions in FXI. As described in this thesis, ML models can enhance several established algorithms. They could foster the analysis of challenging experiments for which standard analysis is too difficult due, for example, to low signal. Developing advanced ML software will also be essential in data reduction to handle the data deluge in XFEL imaging techniques, facilitating faster and more accurate interpretations of the results.

Author's Contributions

Paper I
Writing and implementation. The project idea was discussed and developed by me, Tomas, and Filipe.

Paper II
Project idea conceptualization, writing, and analysis.

Paper III
Online and offline analysis of beamtime data, scientific discussion, and structure modelling with DENSS.

Popular Science Summary

Life is curious - the more we learn, the more it seems to elude us with its complexity. Scientists have tried to tackle it in several ways; some looked at the stars and their motion to make sense of our role in the universe, some looked into plants and animals to keep us grounded, a few sought comfort in mathematical abstractions and had everything figured it out in their minds.

Robert Hooke was a strange man, by all means. His peers described him as cynical with a caustic tongue and quite irritable - but we must give him credit for something modern scientists cannot afford: he was not a specialist. He was a thinker and a tinkerer. He knew of astronomy and optics, and in his seminal work *Micrographia*, published in 1665, thanks to his new invention, a perfected microscope, he describes plants, insects, and even the first cells from cork in great detail. He wrote, "*By the means of Telescopes, there is nothing so far distant but may be represented to our view; and by the help of Microscopes, there is nothing so small, as to escape our inquiry; hence there is a new visible World discovered to the understanding. By this means the Heavens are open'd, and a vast number of new Stars, and new Motions, and new Productions appear in them, to which all the ancient Astronomers were utterly Strangers*". Nowadays, he is remembered mostly as a physicist, but if you think about it, what would be biology without the microscope? He really opened the door to a new microscopic world. Since his invention, biologists have been interested in looking at smaller and smaller things! We can see tiny insects as well as plants and their cells better.

Could we even see within a cell? Certainly so! It may not have been that obvious in the 17th century, but if you know physics, you know this dream could not last forever. In 1873, Ernst Abbe published his famous formula, which puts a limit to what we can see with a microscope. You see, two things limit a microscope. First, the dimension of the lens. Second, microscopes back then were only optical. This might sound confusing, but in brief, that means that visible light is not enough - it limits how small we can go. You might think this is the end of the story, and yet Life is complicated.

Wilhelm Röntgen was a shy, silent man tinkering with a new toy for his laboratory. It was the 8th of November 1895 when he was applying a voltage to this glass tube where the air had been removed. He expected to see the so-called cathode rays, which electrical charges (nowadays we know they are simply electrons) stripped by the metallic wire on one extremity and projected to the other side by the voltage. A real technological marvel! Like a rainbow in the dark - he saw a faint soft glow at the end of the room. Little did he

know, that he was going to become the first Nobel laureate in Physics. You see, accelerated charges emit light. But these high-speed electrons accelerated by very high voltage were colliding with the other metallic extremity of the tube were emitting very energetic photons. They were a new, unknown source of radiation. He soon discovered a peculiarity of this radiation - it could penetrate glass, go through paper, pass through flesh. He exposed his wife's hand to this radiation and placed a photographic plate behind it. To his wife's horror, the developed photo revealed her hand's bone structure and ironically, her wedding ring - this picture still exists, and I find it so romantically macabre! This mysterious source of radiation was named appropriately: *the X-rays*! The world changed in the blink of a few months - as he discovered the X-rays and published his results, this discovery went mouth to mouth all around the world - if you allow me a joke, faster than light. By New Year's Day 1896, many esteemed colleagues had reproduced his result, sending him enthusiastic letters. Within the end of the year, the first medical applications of X-rays were widespread worldwide.

How do X-rays play a role in my story? X-rays push the resolution limit defined by Ernst Abbe down to the atomic level. An X-ray microscope would be a real technological marvel that would thrill Hooke and scare even Röntgen's wife. Indeed, X-ray microscopes are a reality today, but they have a few limitations. Creating good lenses for microscopes is extremely challenging; as I said, they penetrate almost anything. You can record them, but redirecting them is difficult. And yet scientists are resourceful and rarely give up.

In 1912, Max Von Laue was a young *Privatdozent* when he realized that by using X-rays, he could probe crystals and collect images on the photographic plate. The curious part is that the pattern the X-rays create on the photographic plate is related to the structure of the crystal. This required quite a level of abstraction that is difficult to explain in a few lines, but as I said, some scientists prefer abstraction, and they can be very resourceful! The world could now look at crystal structures down to the atomic and molecular levels.

Dorothy Hodgkin was a chemist and a genius in her own right. She studied crystals already at the young age of 10, and she is dearly remembered as a lovable, admirable person, a quality sometimes rare among geniuses. She was the first to study crystallized biological molecules and admire the building blocks of Life itself, starting with sterols, continuing with vitamin B-12, and finally, proteins such as pepsin and hormones like insulin (a gargantuan 34-year-long task). She pioneered the fields of X-ray crystallography and molecular biology.

Almost 100 years have passed since the experiments of Dorothy Hodgkin and many other techniques exist nowadays that allow scientists worldwide to look at Life under new and sometimes missing lenses. Too many to describe here. Now that we can see so far and at such a level of detail, what is left to accomplish?

In **Part I** of my thesis, I explain why we need to develop new imaging methods. The dreams are now two. One, seeing single isolated molecules that are not crystallized. Two, see and understand how they move. In the year 2000, a new idea prompted the biological application of new and brighter sources of X-rays. The diffraction-before-destruction principle. By shining extremely bright pulses of X-ray light produced by machines called X-ray Free Electron Lasers (XFELs), physicists can now look at small nanoparticles like viruses as if they were frozen in time! This method is now known as *Flash X-ray Imaging* (FXI). The radiation is so strong that it destroys the particle, but the light scattered by the particle is faster than the time required to damage it. We can now see single nanoscopic objects with no crystals. The goal is to eventually study proteins. By capturing frames at different time points, we can create molecular movies! Consider that the X-rays can be tuned to capture frames down to the femtoseconds, a time-scale billions of millions of times shorter than a second. Molecular movies would be a marvel with infinite potential. Currently, several experimental methods compete to achieve dynamical studies, and in this thesis, we will briefly discuss a method called time-resolved Small Angle X-ray Scattering (TR-SAXS). In **Part II**, I try my best to describe the key concepts, the experimental setup, and the computational methods that allow the development of such methods.

Talking about computational methods, in **Part III**, I describe something completely different: the basics of Neural Networks. In 1957, a young psychologist called Frank Rosenblatt was sweating and trying to ravel and unravel several wires connected to several electric motors. He was building a dream, a machine that, as he excessively boldly claimed, was "the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence". This machine, known as *Perceptron*, was the prototype of what we now call Neural Networks. Neural Networks moved from the abstract world to the world of wires and motors till landing nowadays pretty much in several of the software in your electronic devices. Neural networks enable solving complicated tasks that would otherwise be very difficult to program. One stunning achievement was recently attained in Molecular Biology. AlphaFold is a software with roots in Neural Networks that enables computationally derived molecular structure models of proteins by providing a list of the molecules (the amino acids) that compose them. To understand why this is so difficult, just consider that proteins are composed of hundreds or even thousands (and more) of amino acids folded together into a cumbersome structure.

Part IV discusses my application of neural networks for processing Flash X-ray Imaging data. Most of the results of this thesis aim to illustrate how this method could be helpful during experiments and are based on simulated data. I also discuss applying a computational method called DENSS (DENsity from Solution Scattering), which allows retrieval of protein densities from X-ray imaging experiments in solution to TR-SAXS data. I hope this can be a small

but inspiring contribution to developing these methods. I am convinced these methods will foster a new age for biology where we will understand Life a bit better. And yet Life is complicated...

Populärvetenskaplig Sammanfattning

Livet är intressant – desto mer vi lär oss, desto mer tycks det gäcka oss med sin komplexitet. Forskare har försökt ta itu med det på flera sätt; några tittade på stjärnorna och deras rörelser för att förstå vår roll i universum, några tittade på växter och djur för att hålla oss förankrade, några sökte tröst i matematiska abstraktioner och hade allt uträknat i huvudet.

Robert Hooke var, för all del, en konstig man. Hans jämlikar beskrev honom som cynisk med en vass tunga och ganska irriterad - men vi måste ge honom äran för något som moderna vetenskapsmän inte har råd med: han var ingen specialist. Han var en tänkare och en pysslare. Han kände till astronomi och optik, och i sitt centrala verk *Micrographia*, publicerat 1665, beskriver han med hjälp av sin nya uppfinning mikroskopet växter, insekter, och till och med de första cellerna från korkträd i detalj. Han skrev "*Med hjälp av teleskop finns det ingenting så långt borta men som kan representeras enligt vår uppfattning; och med hjälp av mikroskop finns det inget så litet som kan undgå vår undersökning; därmed finns det en ny synlig värld att upptäckas. På detta sätt öppnas himlen, och ett väldigt antal nya stjärnor och nya rörelser och nya produktioner dyker upp i dem, för vilka alla forntida astronomer var totala främlingar*". Nuförtiden blir han mest ihågkommen som fysiker, men om man tänker efter så vad vore biologin utan mikroskopet? Han öppnade verkligen dörren till en ny mikroskopisk värld. Sedan han uppfann mikroskopet har biologer varit intresserade av att titta på allt mindre saker! Vi kan se pyttesmå insekter samt växter och deras celler bättre.

Kunde vi ens se inuti en cell? Visst! Det kanske inte var så självklart på 1600-talet, men om man kan fysik så vet man att denna dröm inte kunde vara för evigt. I 1873 publicerade Ernst Abbe sin berömda formel, som sätter en gräns för vad vi kan se med ett mikroskop. Du förstår, två saker begränsar ett mikroskop. Det första är objektivets dimension. Det andra var att mikroskop på den tiden var bara optiska. Detta kanske låter förvirrande, men i korthet betyder det att synligt ljus inte räcker till - det begränsar hur smått vi kan gå. Detta kanske verkar som historians slut, men livet är komplicerat.

Wilhelm Röntgen var en blyg, tyst man som pysslade med en ny leksak till sitt laboratorium. Det var den 8 november 1895 när han applicerade spänning på detta glasrör där luften hade avlägsnats. Han förväntade sig att se de så kallade katodstrålarna, vilka elektriska laddningar (numera vet vi att de helt enkelt är elektroner) skalas av metalltråden på ena änden och projiceras till andra sidan av spänningen. Ett riktigt tekniskt underverk! Som en regnbåge i mörkret såg han ett svagt mjukt sken i ändan av rummet. Föga anade

han att han skulle motta det första Nobelpriset i fysik. Du förstår, accelererade laddningar avger ljus. Men dessa höghastighetselektroner accelererade av mycket hög spänning kolliderade med den andra metalliska ändan av röret och avgav mycket energiska fotoner. De var en ny, okänd strålningskälla. Han upptäckte snart en egenhet med denna strålning - den kunde penetrera glas, gå genom papper, passera genom kött och ben. Han exponerade sin frus hand för denna strålning och placerade en fotografisk platta bakom den. Till hans frus fasa avslöjade det framkallade fotot hennes hands benstruktur och ironiskt nog hennes vigselring - den här bilden finns fortfarande, och jag tycker den är så romantiskt makaber! Denna mystiska strålningskälla fick ett passande namn: röntgenstrålning! Världen förändrades på några månader – efter att han upptäckte röntgenstrålningen och publicerade sina resultat spreds denna upptäckt mun till mun över hela världen - om du tillåter mig ett skämt, snabbare än ljuset. Vid nyårsdagen 1896 hade många högt aktade kollegor reproducerat hans resultat och skickat honom entusiastiska brev. Inom slutet av året var de första medicinska tillämpningarna av röntgenstrålar utbredda över hela världen.

Hur spelar röntgenstrålning en roll i min berättelse? Röntgenstrålar ökar upplösningsgränsen som definierades av Ernst Abbe ner till atomnivå. Ett röntgenmikroskop skulle vara ett verkligt tekniskt underverk som skulle exaltera Hooke och skrämma till och med Röntgens fru. Visserligen är röntgenmikroskop en realitet idag, men de har några begränsningar. Att skapa bra linser för mikroskop är extremt utmanande; som sagt, de penetrerar nästan vad som helst. Man kan mäta dem, men det är svårt att omdirigera dem. Ändå är forskare fyndiga och ger sällan upp.

1912 var Max Von Laue en ung "Privatdozent" när han insåg att han genom att använda röntgenstrålar kunde undersöka kristaller och samla bilder på den fotografiska plattan. Det märkliga är att mönstret som röntgenstrålarna skapade på den fotografiska plattan är relaterat till kristallens struktur. Detta krävde en ganska hög abstraktionsnivå som är svår att förklara på några rader, men som sagt, vissa vetenskapsmän föredrar abstraktion, och de kan vara väldigt fyndiga! Världen kunde nu se på kristallstrukturer ner till atom- och molekylnivå.

Dorothy Hodgkin var en kemist och ett geni i sig. Hon studerade kristaller redan vid 10 års ålder, och hon är ihågkommen som en älskvärd och beundransvärd person, en egenskap som ibland är sällsynt bland genier. Hon var den första att studera kristalliserade biologiska molekyler och beundra byggstenarna i livet självt, börjande med steroler, fortsättande med vitamin B-12, och, slutligen, proteiner som pepsin och hormoner som insulin (en gigantisk 34 år lång uppgift). Hon var en pionjär inom röntgenkristallografi och molekylärbiologi.

Nästan 100 år har gått sedan Dorothy Hodgkins första experiment, och många andra tekniker finns nuförtiden som gör det möjligt för forskare över hela världen att titta på livet under nya och ibland saknade linser. För många

för att beskriva här. Nu när vi kan se så långt och på en sådan detaljnivå, vad återstår att åstadkomma?

I **del I** av min uppsats förklarar jag varför vi behöver utveckla nya avbildningsmetoder. Drömmarna är nu två. En, att se enstaka isolerade molekyler som inte är kristalliserade. Två, se och förstå hur de rör sig. År 2000 ledde en ny idé till den biologiska tillämpningen av nya och ljusare källor till röntgenstrålar. Diffraktion-före-destruktion-principen. Genom att skina extremt starka pulser av röntgenljus som produceras av maskiner som kallas X-ray Free Electron Lasers (XFELs) kan fysiker nu titta på små nanopartiklar såsom virus som om de frysts i tid! Denna metod är nu känd som Flash X-ray Imaging (FXI). Strålningen är så stark att den förstör partikeln, men ljuset som sprids av partikeln är snabbare än den tid som krävs för att skada den. Vi kan nu se enstaka nanoskopiska föremål utan kristaller. Målet är att så småningom studera proteiner. Genom att fånga bildrutor vid olika tidpunkter kan vi skapa molekylära filmer! Tänk på att röntgenstrålarna kan anpassas för att fånga bilder ner till femtosekunder, en tidsskala miljarder miljoner gånger kortare än en sekund. Molekylära filmer skulle vara ett under med oändlig potential. För närvarande tävlar flera experimentella metoder om att uppnå dynamiska studier, och i denna avhandling kommer vi kort att diskutera en metod som kallas time-resolved Small Angle X-ray Scattering (TR-SAXS). I **del II** gör jag mitt bästa för att beskriva nyckelbegreppen, experimentupplägget och de beräkningsmetoder som möjliggör utveckling av sådana metoder.

På tal om beräkningsmetoder beskriver jag i **del III** något helt annat: grunderna av neurala nätverk. 1957 svettades en ung psykolog vid namn Frank Rosenblatt och försökte reda upp flera kablar kopplade till flera elmotorer. Han byggde en dröm, en maskin som, som han överdrivet djärvt hävdade, var "*embryot till en elektronisk dator som [marinen] förväntar sig kommer kunna gå, prata, se, skriva, reproducera sig själv och vara medveten om dess existens*". Denna maskin, känd som *Perceptron*, var prototypen till vad vi nu kallar Neural Networks. Neurala nätverk flyttade från den abstrakta världen till en värld av ledningar och motorer tills de landade numera i stort sett i flera av mjukvarorna i våra elektroniska enheter. Neurala nätverk möjliggör lösningen av komplicerade uppgifter som annars skulle vara mycket svåra att programmera. En fantastisk prestation uppnåddes nyligen inom molekylärbiologi. AlphaFold är en programvara med rötter i neurala nätverk som möjliggör beräkningsmässigt härledda molekylstrukturmodeller av proteiner genom att tillhandahålla en lista över molekylerna (aminosyrorna) som utgör dem. För att förstå varför detta är så svårt, tänk bara på att proteiner består av hundratals eller till och med tusentals (och fler) aminosyror hopvikta till en besvärlig struktur.

Del IV diskuterar min tillämpning av neurala nätverk för bearbetning av Flash X-ray Imaging data. De flesta av resultaten i denna avhandling syftar till att illustrera hur denna metod kan vara till hjälp för experiment och är baserade på simulerad data. Jag diskuterar också tillämpningen av en beräkningsmetod som kallas DENSS (DENsity from Solution Scattering), som möjliggör hämt-

ning av proteindensiteter från röntgenavbildningsexperiment i lösning till TR-SAXS-data. Jag hoppas att detta kan vara ett litet men inspirerande bidrag till att utveckla dessa metoder. Jag är övertygad om att dessa metoder kommer att främja en ny tid för biologi där vi kommer att förstå livet lite bättre. Dock så är livet komplicerat...

Acknowledgements

C'è un'ape che se posa
su un bottone de rosa:
lo succhia e se ne va...
Tutto sommato, la felicità
è una piccola cosa.
Trilussa, *La felicità*, 1937

This work was possible only through several people's support - pun intended... got it? No? What do you mean, no? You have not read the thesis!

First, I want to start by thanking my supervisor, **Tomas**, who encouraged me and believed in me more than I do. To a brighter future! I thank my co-supervisor, **Filipe**, for the help and opportunities throughout the years.

I thank **Tong** and **August** for the friendship, for the music (some jazz, some classical, some shitty-fluting), the chaos, the rants, the board games, and the work done and more often not done together. **Tej**, **Therese**, and **Mayra** for being the warmest and loveliest people I know. **Emiliano**, **Stavros**, and **Laura** thank you for always checking on me when you were around and sometimes even when you were not around. **Lena**, **Steinar**, and **Emilia**; thanks for all the fun and many board games. To all the **LMB people** - new arrivals, current members, and past people - thanks for all the chats and shenanigans throughout the years. In particular, thank you **Ana**, **Daniel**, and **Kenta** for the best ski trips. I think **Anna** needs to be thanked twice for being a past and current member of the lab: thank you, thank you! To the **Biophysics Network**, because in the ups and downs, they always had a minute to spare and listen to me rant about something - **Calle**, in particular. I want especially to thank **Sebastian** and **Pamela** for a great trip to California! And to **Nic**, who taught me that chestnuts can propel dreams and rockets; it was always a pleasure chatting with you - mulțumesc! Thanks also to **Ibbe** and to **Han**, and congratulations again for your defenses!

To **Gabriele**, **Jelena**, **Simone**, **Alessandro**, **Marius**, and **Vita**: for a long friendship, a short trip to Lofoten, and for pushing me to write most of this thesis before it. And also to **Patta**; you were with us in spirit! To **Matilda** and **Johan** for their constant encouragement and all the gluten-free sweets. They are great people who know the meaning of "depth perception". I also want to thank my **complex family** - my two real parents and my two imaginary ones. Thank you - I love you a lot! Thanks also to my sister **Alice**, good luck with

university! It is your turn now! To **Lina**, for being the best human being I know, for tolerating my quirks, for our adventures (past and future), and for making me feel very loved.

I want to end the acknowledgments by thanking **Stefano**, because he has always believed in me... and also because he knows why I must end this thesis by acknowledging **B. Sammartino**.

References

[1] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[2] Jianwei Miao. *Extending the methodology of X-ray crystallography to noncrystalline specimens*. State University of New York at Stony Brook, 1999.

[3] Henry N Chapman, Anton Barty, Michael J Bogan, Sébastien Boutet, Matthias Frank, Stefan P Hau-Riege, Stefano Marchesini, Bruce W Woods, Saša Bajt, W Henry Benner, et al. Femtosecond diffractive imaging with a soft-x-ray free-electron laser. *Nature Physics*, 2(12):839–843, 2006.

[4] Pierre Thibault, Veit Elser, Chris Jacobsen, David Shapiro, and David Sayre. Reconstruction of a yeast cell from x-ray diffraction data. *Acta Crystallographica Section A: Foundations of Crystallography*, 62(4):248–261, 2006.

[5] M Marvin Seibert, Tomas Ekeberg, Filipe RNC Maia, Martin Svenda, Jakob Andreasson, Olof Jönsson, Duško Odić, Bianca Iwan, Andrea Rocker, Daniel Westphal, et al. Single mimivirus particles intercepted and imaged with an x-ray laser. *Nature*, 470(7332):78–81, 2011.

[6] Max F Hantke, Dirk Hasse, Filipe RNC Maia, Tomas Ekeberg, Katja John, Martin Svenda, N Duane Loh, Andrew V Martin, Nicusor Timneanu, Daniel SD Larsson, et al. High-throughput imaging of heterogeneous cell organelles with an x-ray laser. *Nature Photonics*, 8(12):943–949, 2014.

[7] Gijs Van Der Schot, Martin Svenda, Filipe RNC Maia, Max Hantke, Daniel P DePonte, M Marvin Seibert, Andrew Aquila, Joachim Schulz, Richard Kirian, Mengning Liang, et al. Imaging single cells in a beam of live cyanobacteria with an x-ray laser. *Nature communications*, 6(1):1–9, 2015.

[8] Tomas Ekeberg, Dameli Assalauova, Johan Bielecki, Rebecca Boll, Benedikt J Daurer, Lutz A Eichacker, Linda E Franken, Davide E Galli, Luca Gelisio, Lars Gumprecht, et al. Observation of a single protein by ultrafast x-ray diffraction. *Light: Science & Applications*, 13(1):15, 2024.

[9] Tomas Ekeberg, Martin Svenda, Chantal Abergel, Filipe RNC Maia, Virginie Seltzer, Jean-Michel Claverie, Max Hantke, Olof Jönsson, Carl Nettelblad, Gijs Van Der Schot, et al. Three-dimensional reconstruction of the giant mimivirus particle with an x-ray free-electron laser. *Physical review letters*, 114(9):098102, 2015.

[10] Ruslan P Kurta, Jeffrey J Donatelli, Chun Hong Yoon, Peter Berntsen, Johan Bielecki, Benedikt J Daurer, Hasan DeMirci, Petra Fromme, Max Felix Hantke, Filipe RNC Maia, et al. Correlations in scattered x-ray laser pulses reveal nanoscale structural features of viruses. *Physical review letters*, 119 (15):158102, 2017.

[11] Max Rose, Sergey Bobkov, Kartik Ayyer, Ruslan P Kurta, Dmitry Dzhigaev, Young Yong Kim, Andrew J Morgan, Chun Hong Yoon, Daniel Westphal, Johan Bielecki, et al. Single-particle imaging without symmetry constraints at an x-ray free-electron laser. *IUCrJ*, 5(6):727–736, 2018.

[12] Ida V Lundholm, Jonas A Sellberg, Tomas Ekeberg, Max F Hantke, Kenta Okamoto, Gijs van der Schot, Jakob Andreasson, Anton Barty, Johan Bielecki, Petr Bruza, et al. Considerations for three-dimensional image reconstruction from experimental data in coherent diffractive imaging. *IUCrJ*, 5(5):531–541, 2018.

[13] Kartik Ayyer, P Lourdu Xavier, Johan Bielecki, Zhou Shen, Benedikt J Daurer, Amit K Samanta, Salah Awel, Richard Bean, Anton Barty, Martin Bergemann, et al. 3d diffractive imaging of nanoparticle ensembles using an x-ray laser. *Optica*, 8(1):15–23, 2021.

[14] Eugene YD Chua, Joshua H Mendez, Micah Rapp, Serban L Ilca, Yong Zi Tan, Kashyap Maruthi, Huihui Kuang, Christina M Zimanyi, Anchi Cheng, Edward T Eng, et al. Better, faster, cheaper: recent advances in cryo–electron microscopy. *Annual review of biochemistry*, 91:1–32, 2022.

[15] Arwen R Pearson and Pedram Mehrabi. Serial synchrotron crystallography for time-resolved structural biology. *Current Opinion in Structural Biology*, 65: 168–174, 2020.

[16] Richard Neutze, Remco Wouts, David Van der Spoel, Edgar Weckert, and Janos Hajdu. Potential for biomolecular imaging with femtosecond x-ray pulses. *Nature*, 406(6797):752–757, 2000.

[17] Henry N Chapman, Anton Barty, Stefano Marchesini, Aleksandr Noy, Stefan P Hau-Riege, Congwu Cui, Malcolm R Howells, Rachel Rosen, Haifeng He, John CH Spence, et al. High-resolution ab initio three-dimensional x-ray diffraction microscopy. *JOSA A*, 23(5):1179–1200, 2006.

[18] Egor Sobolev, Philipp Schmidt, Janusz Malka, David Hammer, Djelloul Boukhelef, Johannes Möller, Karim Ahmed, Richard Bean, Ivette Jazmín Bermúdez Macías, Johan Bielecki, et al. Data reduction activities at european xfel: early results. *Frontiers in Physics*, 12:1331329, 2024.

[19] David J DiSantis. Early american radiology: the pioneer years. *American Journal of Roentgenology*, 147(4):850–853, 1986.

[20] Mariusz Jaskolski, Zbigniew Dauter, and Alexander Wlodawer. A brief history of macromolecular crystallography, illustrated by a family tree and its n obel fruits. *The FEBS journal*, 281(18):3985–4009, 2014.

[21] Francis Arthur Jenkins and Harvey Elliott White. Fundamentals of optics. *Indian Journal of Physics*, 25:265–266, 1957.

[22] PP Ewald. Introduction to the dynamical theory of x-ray diffraction. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 25(1):103–108, 1969.

[23] Massimo Altarelli, Reinhard Brinkmann, and Majed Chergui. The european x-ray free-electron laser. technical design report. 2007.

[24] K Balewski, W Brefeld, W Decking, H Franz, R Röhlsberger, and Edgar Weckert. Petra iii: A low emittance synchrotron radiation source. *Technical Design Report, DESY*, 35:2004, 2004.

[25] C Pellegrini and A Marinelli. The physics of x-ray free-electron lasers. *Reviews of Modern Physics*, 88(1):015006, 2016.
[26] David Attwood. *Soft x-rays and extreme ultraviolet radiation: principles and applications*. Cambridge university press, 2000.
[27] Stefan P Hau-Riege, Richard A London, Gosta Huldt, and Henry N Chapman. Pulse requirements for x-ray diffraction imaging of single biological molecules. *Physical Review E*, 71(6):061919, 2005.
[28] Jianwei Miao, David Sayre, and HN Chapman. Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects. *JOSA A*, 15(6): 1662–1669, 1998.
[29] Johan Bielecki, Max F Hantke, Benedikt J Daurer, Hemanth KN Reddy, Dirk Hasse, Daniel SD Larsson, Laura H Gunn, Martin Svenda, Anna Munke, Jonas A Sellberg, et al. Electrospray sample injection for single-particle imaging with x-ray lasers. *Science Advances*, 5(5):eaav8801, 2019.
[30] Benedikt J Daurer, Kenta Okamoto, Johan Bielecki, Filipe RNC Maia, Kerstin Mühlig, M Marvin Seibert, Max F Hantke, Carl Nettelblad, W Henry Benner, Martin Svenda, et al. Experimental strategies for imaging bioparticles with femtosecond hard x-ray pulses. *IUCrJ*, 4(3):251–262, 2017.
[31] Geoffrey Ingram Taylor. Disintegration of water drops in an electric field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 280(1382):383–397, 1964.
[32] Erik G Marklund, Tomas Ekeberg, Mathieu Moog, Justin LP Benesch, and Carl Caleman. Controlling protein orientation in vacuum using electric fields. *The journal of physical chemistry letters*, 8(18):4540–4544, 2017.
[33] Anna Sinelnikova, Thomas Mandl, Harald Agelii, Oscar Grånäs, Erik G Marklund, Carl Caleman, and Emiliano De Santis. Protein orientation in time-dependent electric fields: orientation before destruction. *Biophysical Journal*, 120(17):3709–3717, 2021.
[34] August Wollter, Emiliano De Santis, Tomas Ekeberg, Erik G Marklund, and Carl Caleman. Enhanced emc—advantages of partially known orientations in x-ray single particle imaging. *The Journal of Chemical Physics*, 160(11), 2024.
[35] Gábor Bortel and Miklós Tegze. Common arc method for diffraction pattern orientation. *Acta Crystallographica Section A: Foundations of Crystallography*, 67(6):533–543, 2011.
[36] Russell Fung, Valentin Shneerson, Dilano K Saldin, and Abbas Ourmazd. Structure from fleeting illumination of faint spinning objects in flight. *Nature Physics*, 5(1):64–67, 2009.
[37] Ne-Te Duane Loh and Veit Elser. Reconstruction algorithm for single-particle diffraction imaging experiments. *Physical Review E*, 80(2):026705, 2009.
[38] Cy M Jeffries, Jan Ilavsky, Anne Martel, Stephan Hinrichs, Andreas Meyer, Jan Skov Pedersen, Anna V Sokolova, and Dmitri I Svergun. Small-angle x-ray and neutron scattering. *Nature Reviews Methods Primers*, 1(1):70, 2021.
[39] Hyun Sun Cho, Friedrich Schotte, Valentyn Stadnytskyi, and Philip Anfinrud. Time-resolved x-ray scattering studies of proteins. *Current opinion in structural biology*, 70:99–107, 2021.
[40] Sergei Grudinin, Maria Garkavenko, and Andrei Kazennov. Pepsi-saxs: an adaptive method for rapid and accurate computation of small-angle x-ray

scattering profiles. *Acta Crystallographica Section D: Structural Biology*, 73 (5):449–464, 2017.
[41] David Sayre. Some implications of a theorem due to shannon. *Acta Crystallographica*, 5(6):843–843, 1952.
[42] Claude E Shannon. Communication theory of secrecy systems. *The Bell system technical journal*, 28(4):656–715, 1949.
[43] James R Fienup. Reconstruction of an object from the modulus of its fourier transform. *Optics letters*, 3(1):27–29, 1978.
[44] James R Fienup, Thomas R Crimmins, and Wlodzimierz Holsztynski. Reconstruction of the support of an object from the support of its autocorrelation. *JOSA*, 72(5):610–624, 1982.
[45] Patrick Jordan, Petra Fromme, Horst Tobias Witt, Olaf Klukas, Wolfram Saenger, and Norbert Krauß. Three-dimensional structure of cyanobacterial photosystem i at 2.5 å resolution. *Nature*, 411(6840):909–917, 2001.
[46] S Marchesini, HN Chapman, SP Hau-Riege, RA London, A Szoke, H He, MR Howells, H Padmore, R Rosen, JCH Spence, et al. Coherent x-ray diffractive imaging: applications and limitations. *Optics Express*, 11(19): 2344–2353, 2003.
[47] Stefano Marchesini, H He, Henry N Chapman, Stefan P Hau-Riege, Aleksandr Noy, Malcolm R Howells, Uwe Weierstall, and John CH Spence. X-ray image reconstruction from a diffraction pattern alone. *Physical Review B*, 68(14): 140101, 2003.
[48] D Russell Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse problems*, 21(1):37, 2004.
[49] Veit Elser. Phase retrieval by iterated projections. *JOSA A*, 20(1):40–55, 2003.
[50] Stefano Marchesini. Phase retrieval and saddle-point optimization. *JOSA A*, 24(10):3289–3296, 2007.
[51] A. V. Martin, F. Wang, N. D. Loh, T. Ekeberg, F. R. N. C. Maia, M. Hantke, and G. van der Schot et al. Noise-robust coherent diffractive imaging with a single diffraction pattern. *Opt. Express*, 20(15):16650–16661, Jul 2012. doi: 10.1364/OE.20.016650. URL `https://opg.optica.org/oe/abstract.cfm?URI=oe-20-15-16650`.
[52] Kartik Ayyer, Andrew J Morgan, Andrew Aquila, Hasan DeMirci, Brenda G Hogue, Richard A Kirian, P Lourdu Xavier, Chun Hong Yoon, Henry N Chapman, and Anton Barty. Low-signal limit of x-ray single particle diffractive imaging. *Optics express*, 27(26):37816–37833, 2019.
[53] Alessandro Colombo, Davide Emilio Galli, Liberato De Caro, Francesco Scattarella, and Elvio Carlino. Facing the phase problem in coherent diffractive imaging via memetic algorithms. *Scientific reports*, 7(1):42236, 2017.
[54] Alberto Pietrini and Carl Nettelblad. Using convex optimization of autocorrelation with constrained support and windowing for improved phase retrieval accuracy. *Optics express*, 26(19):24422–24443, 2018.
[55] Yu Luo, Michela G Bertero, Elizabeth A Frey, Richard A Pfuetzner, Markus R Wenk, Louise Creagh, Sandra L Marcus, Daniel Lim, Frank Sicheri, Cyril Kay, et al. Structural and biochemical characterization of the type iii secretion chaperones cest and sige. *Nature structural biology*, 8(12):1031–1036, 2001.

[56] TR Crimmins, JR Fienup, and BJ Thelen. Improved bounds on object support from autocorrelation support and application to phase retrieval. *JOSA A*, 7(1): 3–13, 1990.
[57] Thomas D Grant. Ab initio electron density determination directly from solution scattering data. *Nature methods*, 15(3):191–193, 2018.
[58] Abbe. The relation of aperture and power in the microscope. *Journal of the royal microscopical society*, 2(3):300–309, 1882.
[59] A Allahgholi, J Becker, L Bianco, A Delfs, R Dinapoli, P Goettlicher, D Greiffenberg, H Hirsemann, S Jack, et al. Agipd, a high dynamic range fast detector for the european xfel. *Journal of Instrumentation*, 10(01):C01023, 2015.
[60] Markus Kuster, Karim Ahmed, K-E Ballak, Cyril Danilevski, Marko Ekmedžić, Bruno Fernandes, Patrick Gessler, Robert Hartmann, Steffen Hauf, Peter Holl, et al. The 1-megapixel pnccd detector for the small quantum systems instrument at the european xfel: system and operation aspects. *Journal of Synchrotron Radiation*, 28(2):576–587, 2021.
[61] Max F Hantke, Tomas Ekeberg, and Filipe RNC Maia. Condor: a simulation tool for flash x-ray imaging. *Journal of applied crystallography*, 49(4): 1356–1362, 2016.
[62] Alberto Pietrini and Carl Nettelblad. Artifact reduction in the cspad detectors used for lcls experiments. *Journal of Synchrotron Radiation*, 24(5): 1092–1097, 2017.
[63] Alberto Pietrini, Johan Bielecki, Nicusor Timneanu, Max F Hantke, Jakob Andreasson, N Duane Loh, Daniel SD Larsson, Sébastien Boutet, Janos Hajdu, Filipe RNC Maia, et al. A statistical approach to detect protein complexes at x-ray free electron laser facilities. *Communications Physics*, 1(1):92, 2018.
[64] Marin Van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of structural biology*, 151(3):250–262, 2005.
[65] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.
[66] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5: 115–133, 1943.
[67] Frank Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.
[68] Marvin Minsky and Seymour Papert. An introduction to computational geometry. *Cambridge tiass., HIT*, 479(480):104, 1969.
[69] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5): 359–366, 1989.
[70] Yoshua Bengio, Ian Goodfellow, and Aaron Courville. *Deep learning*, volume 1. MIT press Cambridge, MA, USA, 2017.
[71] François Chollet et al. Keras, 2015. URL `https://keras.io`.
[72] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh

Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

[73] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[74] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[75] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. Ucsf chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

[76] Heikki Takala, Alexander Björling, Oskar Berntsson, Heli Lehtivuori, Stephan Niebling, Maria Hoernke, Irina Kosheleva, Robert Henning, Andreas Menzel, Janne A Ihalainen, et al. Signal amplification and transduction in phytochrome photosensors. *Nature*, 509(7499):245–248, 2014.

[77] Hwei-Jen Lee, Matthew D Lloyd, Ian J Clifton, Karl Harlos, Alain Dubus, Jack E Baldwin, Jean-Marie Frere, and Christopher J Schofield. Alteration of the co-substrate selectivity of deacetoxycephalosporin c synthase: The role of arginine 258. *Journal of Biological Chemistry*, 276(21):18290–18295, 2001.

[78] Falco Kilchmann, Maria J Marcaida, Sachin Kotak, Thomas Schick, Silvan D Boss, Mahendra Awale, Pierre Gonczy, and Jean-Louis Reymond. Discovery of a selective aurora a kinase inhibitor by virtual screening. *Journal of medicinal chemistry*, 59(15):7188–7211, 2016.

[79] Mathew J Cherukara, Youssef SG Nashed, and Ross J Harder. Real-time coherent diffraction inversion using deep generative networks. *Scientific reports*, 8(1):1–8, 2018.

[80] Alexander Scheinker and Reeju Pokharel. Adaptive 3d convolutional neural network-based reconstruction method for 3d coherent diffraction imaging. *Journal of Applied Physics*, 128(18):184901, 2020.

[81] Longlong Wu, Shinjae Yoo, Ana F Suzana, Tadesse A Assefa, Jiecheng Diao, Ross J Harder, Wonsuk Cha, and Ian K Robinson. Three-dimensional coherent x-ray diffraction imaging via deep convolutional neural networks. *npj Computational Materials*, 7(1):1–8, 2021.

[82] Charu Chaudhry, Arthur L Horwich, Axel T Brunger, and Paul D Adams. Exploring the structural dynamics of the e. coli chaperonin groel using translation-libration-screw crystallographic refinement of intermediate states. *Journal of molecular biology*, 342(1):229–245, 2004.

[83] Samarjit Patnaik, Kirk L Stevens, Roseanne Gerding, Felix Deanda, J Brad Shotwell, Jun Tang, Toshihiro Hamajima, Hiroko Nakamura, M Anthony Leesnitzer, Anne M Hassell, et al. Discovery of 3, 5-disubstituted-1h-pyrrolo

[2, 3-b] pyridines as potent inhibitors of the insulin-like growth factor-1 receptor (igf-1r) tyrosine kinase. *Bioorganic & medicinal chemistry letters*, 19 (11):3136–3140, 2009.

[84] Ehsan Amin, Mamta Jaiswal, Urszula Derewenda, Katarina Reis, Kazem Nouri, Katja T Koessmeier, Pontus Aspenström, Avril V Somlyo, Radovan Dvorsky, and Mohammad R Ahmadian. Deciphering the molecular and functional basis of rhogap family proteins: a systematic approach toward selective inactivation of rho family proteins. *Journal of Biological Chemistry*, 291(39):20353–20371, 2016.

[85] De-Ping Wang, Rong Zhao, Hao-Feng Wang, Mei-Yue Wang, Wen-Shu Hu, Meng-Meng Lin, Wen Shu, Yao-Jun Sun, Ji-Min Cao, Wen Cui, et al. Crystal structure of mrna cap (guanine-n7) methyltransferase e12 subunit from monkeypox virus and discovery of its inhibitors. *International Journal of Biological Macromolecules*, 253:127565, 2023.

[86] Kate M Smith, Veronica Di Antonio, Luca Bellucci, David R Thomas, Fabiana Caporuscio, Francesco Ciccarese, Hanieh Ghassabian, Kylie M Wagstaff, Jade K Forwood, David A Jans, et al. Contribution of the residue at position 4 within classical nuclear localization signals to modulating interaction with importins and nuclear targeting. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1865(8):1114–1129, 2018.

[87] Alfredo Bellisario, Filipe RNC Maia, and Tomas Ekeberg. Noise reduction and mask removal neural network for x-ray single-particle imaging. *Journal of applied crystallography*, 55(1):122–132, 2022.

[88] Mark Towers, Jordan K. Terry, Ariel Kwiatkowski, John U. Balis, Gianluca de Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Arjun KG, Markus Krimmel, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Andrew Tan Jin Shen, and Omar G. Younis. Gymnasium, March 2023. URL `https://zenodo.org/record/8127025`.

[89] Matthias Plappert. keras-rl. `https://github.com/keras-rl/keras-rl`, 2016.

[90] Phillip M Maffettone, Joshua K Lynch, Thomas A Caswell, Clara E Cook, Stuart I Campbell, and Daniel Olds. Gaming the beamlines—employing reinforcement learning to maximize scientific outcomes at large-scale user facilities. *Machine Learning: Science and Technology*, 2(2):025025, 2021.

[91] Julia Dietler, Renate Gelfert, Jennifer Kaiser, Veniamin Borin, Christian Renzl, Sebastian Pilsl, Américo Tavares Ranzani, Andrés García de Fuentes, Tobias Gleichmann, Ralph P Diensthuber, et al. Signal transduction in light-oxygen-voltage receptors lacking the active-site glutamine. *Nature Communications*, 13(1):2618, 2022.

[92] Jumpei Yamada, Satoshi Matsuyama, Ichiro Inoue, Taito Osaka, Takato Inoue, Nami Nakamura, Yuto Tanaka, Yuichi Inubushi, Toshinori Yabuuchi, Kensuke Tono, et al. Extreme focusing of hard x-ray free-electron laser pulses enables 7 nm focus width and 1022 w cm- 2 intensity. *Nature Photonics*, pages 1–6, 2024.

[93] Tej Varma Yenupuri, Safi Rafie-Zinedine, Lena Worbs, Michael Heymann, Joachim Schulz, Johan Bielecki, and Filipe RNC Maia. Helium-electrospray improves sample delivery in x-ray single-particle imaging experiments.

Scientific Reports, 14(1):4401, 2024.

[94] Safi Rafie-Zinedine, Tej Varma Yenupuri, Lena Worbs, Filipe R. N. C. Maia, Michael Heymann, Joachim Schulz, and Johan Bielecki. Enhancing electrospray ionization efficiency for particle transmission through an aerodynamic lens stack. *Journal of Synchrotron Radiation*, 31(2):222–232, Mar 2024. doi: 10.1107/S1600577524000158. URL `https://doi.org/10.1107/S1600577524000158`.

[95] Patrick E Konold, Tong You, Johan Bielecki, Joana Valerio, Marco Kloos, Daniel Westphal, Alfredo Bellisario, Tej Varma Yenupuri, August Wollter, Jayanath CP Koliyadu, et al. 3d-printed sheet jet for stable megahertz liquid sample delivery at x-ray free-electron lasers. *IUCrJ*, 10(6), 2023.

[96] Dameli Assalauova, Young Yong Kim, Sergey Bobkov, Ruslan Khubbutdinov, Max Rose, Roberto Alvarez, Jakob Andreasson, Eugeniu Balaur, Alice Contreras, Hasan DeMirci, et al. An advanced workflow for single-particle imaging with the limited data at an x-ray free-electron laser. *IUCrJ*, 7(6): 1102–1113, 2020.

[97] Dameli Assalauova, Alexandr Ignatenko, Fabian Isensee, Darya Trofimova, and Ivan A Vartanyants. Classification of diffraction patterns using a convolutional neural network in single-particle-imaging experiments performed at x-ray free-electron lasers. *Journal of applied crystallography*, 55 (3):444–454, 2022.

[98] Sung Yun Lee, Do Hyung Cho, Chulho Jung, Daeho Sung, Daewoong Nam, Sangsoo Kim, and Changyong Song. Denoising low-intensity diffraction signals using k-space deep learning: Applications to phase recovery. *Physical Review Research*, 3(4):043066, 2021.

[99] Thomas Stielow, Robin Schmidt, Christian Peltz, Thomas Fennel, and Stefan Scheel. Fast reconstruction of single-shot wide-angle diffraction images through deep learning. *Machine Learning: Science and Technology*, 1(4): 045007, 2020.

[100] Thomas Stielow and Stefan Scheel. Reconstruction of nanoscale particles from single-shot wide-angle free-electron-laser diffraction patterns with physics-informed neural networks. *Physical Review E*, 103(5):053312, 2021.

[101] Longlong Wu, Pavol Juhas, Shinjae Yoo, and Ian Robinson. Complex imaging of phase domains by deep neural networks. *IUCrJ*, 8(1), 2021.

[102] Julian Zimmermann, Bruno Langbehn, Riccardo Cucini, Michele Di Fraia, Paola Finetti, Aaron C LaForge, Toshiyuki Nishiyama, Yevheniy Ovcharenko, Paolo Piseri, Oksana Plekan, et al. Deep neural networks for classifying complex features in diffraction images. *Physical Review E*, 99(6):063309, 2019.

[103] Niky Bruchon, Gianfranco Fenu, Giulio Gaio, Marco Lonza, Finn Henry O'Shea, Felice Andrea Pellegrino, and Erica Salvato. Basic reinforcement learning techniques to control the intensity of a seeded free-electron laser. *Electronics*, 9(5):781, 2020.

[104] Simon Hirlaender and Niky Bruchon. Model-free and bayesian ensembling model-based deep reinforcement learning for particle accelerator control demonstrated on the fermi fel. *arXiv preprint arXiv:2012.09737*, 2020.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2400

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-526527

ACTA UNIVERSITATIS
UPSALIENSIS
2024