UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1589

Algorithms for Coherent Diffractive Imaging with X-ray Lasers

BENEDIKT J. DAURER

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017

ISSN 1651-6214
ISBN 978-91-513-0129-7
urn:nbn:se:uu:diva-329012

Dissertation presented at Uppsala University to be publicly examined in Room B7:101a, Biomedicinska Centrum (BMC), Husargatan 3, Uppsala, Friday, 15 December 2017 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Richard Neutze (University of Gothenburg).

Abstract

Daurer, B. J. 2017. Algorithms for Coherent Diffractive Imaging with X-ray Lasers. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1589. 64 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0129-7.

Coherent diffractive imaging (CDI) has become a very popular technique over the past two decades. CDI is a "lensless" imaging method which replaces the objective lens of a conventional microscope by a computational image reconstruction procedure. Its increase in popularity came together with the development of X-ray free-electron lasers (XFELs) which produce extremely bright and coherent X-rays. By facilitating these unique properties, CDI enables structure determination of non-crystalline samples at nanometre resolution and has many applications in structural biology, material science and X-ray optics among others. This work focuses on two specific CDI techniques, flash X-ray diffractive imaging (FXI) on biological samples and X-ray ptychography.

While the first FXI demonstrations using soft X-rays have been quite promising, they also revealed remaining technical challenges. FXI becomes even more demanding when approaching shorter wavelengths to allow subnanometre resolution imaging. We described one of the first FXI experiments using hard X-rays and characterized the most critical components of such an experiment, namely the properties of X-ray focus, sample delivery and detectors. Based on our findings, we discussed experimental and computational strategies for FXI to overcome its current difficulties and reach its full potential. We deposited the data in the Coherent X-ray Database (CXIDB) and made our data analysis code available in a public repository. We developed algorithms targeted towards the needs of FXI experiments and implemented a software package which enables the analysis of diffraction data in real time.

X-ray ptychography has developed into a very useful tool for quantitative imaging of complex materials and has found applications in many areas. However, it involves a computational reconstruction step which can be slow. Therefore, we developed a fast GPU-based ptychographic solver and combined it with a framework for real-time data processing which already starts the ptychographic reconstruction process while data is still being collected. This provides immediate feedback to the user and allows high-throughput ptychographic imaging.

Finally, we have used ptychographic imaging as a method to study the wavefront of a focused XFEL beam under typical FXI conditions.

We are convinced that this work on developing strategies and algorithms for FXI and ptychography is a valuable contribution to the development of coherent diffractive imaging.

Keywords: X-ray lasers, coherent diffractive imaging, algorithms, lensless imaging, flash diffractive imaging, flash X-ray imaging, aerosol injection, FEL, XFEL, CXI, CDI, FXI

Benedikt J. Daurer, Department of Cell and Molecular Biology, Molecular biophysics, Box 596, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-0129-7
urn:nbn:se:uu:diva-329012 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-329012)

Dedicated To Niko, Gabi and Hans

List of papers

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

I **B. J. Daurer**[†], K. Okamoto[†] *et al.*
Experimental strategies for imaging bioparticles with femtosecond hard X-ray pulses
IUCrJ **4**, 251-262 (2017).

II **B. J. Daurer**, M. F. Hantke, C. Nettelblad and F. R. N. C. Maia
Hummingbird: monitoring and analyzing flash X-ray imaging experiments in real time
Journal of Applied Crystallography **49**, 1042-1047 (2016).

III S. Marchesini, H. Krishnan, **B. J. Daurer**, D. A. Shapiro, T. Perciano, J. A. Sethian and F. R. N. C. Maia
SHARP: a distributed GPU-based ptychographic solver
Journal of Applied Crystallography **49**, 1245-1252 (2016).

IV **B. J. Daurer**, H. Krishnan, T. Perciano, F. R. N. C. Maia, D. A. Shapiro, J. A. Sethian and S. Marchesini
Nanosurveyor: a framework for real-time data processing
Advanced Structural and Chemical Imaging **3**:7 (2017).

V **B. J. Daurer**[†], S. Sala[†] *et al.*
Wavefront sensing of individual XFEL pulses using ptychography
In preparation.

Reprints were made with permission from the publishers.

[†]These authors contributed equally to this study

List of additional papers

VI A. Munke *et al.*
Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source
Scientific Data **3** 160064 (2016).

VII H. K. N. Reddy *et al.*
Coherent soft X-ray diffraction imaging of coliphage PR772 at the Linac coherent light source
Scientific Data **4** 170079 (2017).

VIII R. P. Kurta *et al.*
Correlations in scattered x-ray laser pulses reveal nanoscale structural features of viruses
Physical Review Letters **119**:158102 (2017).

Contents

List of abbreviations

ADU	Analogue-to-Digital Unit
ALS	Advanced Light Source
AMO	Beamline for Atomic, Molecular and Optical experiments
ASIC	Application-Specific Integration Circuit
CDI	Coherent Diffractive Imaging
cryoEM	Cryo-Electron Microscopy
CSPAD	Cornell-SLAC Pixel-Array Detector
CXI	Beamline for Coherent X-ray Imaging experiments
CXIDB	Coherent X-ray Imaging Data Base
DFT	Discrete Fourier Transform
DM	Difference Map
EMC	Expansion Maximization Compression
ePIE	extended Ptychographic Iterative Engine
ER	Error Reduction
ESI	Electrospray Ionization
EuXFEL	European X-ray Free-Electron Laser
FFT	Fast Fourier Transform
FLASH	Free-electron LASer in Hamburg
FSC	Fourier Shell Correlation
FXI	Flash X-ray Diffractive Imaging
FZP	Fresnel Zone Plate
GDVN	Gas Dynamic Virtual Nozzle
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HDF5	Hierarchical Data Format Version 5
HIO	Hybrid Input-Output
KB	Kirkpatrick-Baez mirror
LCLS	Linac Coherent Light Source

LINAC LINear ACcelerator

ML Maximum Likelihood

OmRV Omono River Virus

PRTF Phase Retrieval Transfer Function

RAAR Relaxed Averaged Alternating Reflectors

SASE Self Amplified Stimulated Emission
SLAC Stanford LINear ACcelerator
SNR Signal-to-Noise Ratio

TCP Transmission Control Protocol

UDP User Datagram Protocol

XFEL X-ray Free-Electron Laser

1. Introduction

When studying the function of complex systems such as the life cycle of a battery, a virus or a human cell, it seems logical to first take images in order to understand the building blocks of the system and identify its structural components. So it should come as no surprise that imaging has become a major field of study in many areas of modern science.

The invention of the first light microscope in the late 16th century made it possible to investigate transparent specimen beyond the macroscopic level. With the discovery of X-rays in 1895 by Wilhelm Röntgen, the limits were pushed even further and it became possible to also investigate non-transparent (or opaque) specimen at the microscopic level. Soon after X-rays, in the 1930's, even electrons were used to investigate microstructures and started "competing" with light. In conventional light, electron and X-ray microscopes, lenses are used to form images. Although fundamentally limited only by the wavelength of the source, in practice the achievable resolution is determined by the quality of the lenses that can be manufactured, which is a particular challenge for X-ray lenses. An alternative approach able to overcome these limitations is Coherent Diffractive Imaging (CDI), which replaces the objective lens by a computational image reconstruction procedure and is therefore often referred to as "lensless" imaging.

The idea of CDI dates back to David Sayre who referred to Shannon's sampling theorem [1] and realized that direct structure determination based on intensities recorded in diffraction space can be accomplished under sufficient sampling conditions [2]. In a CDI measurement, only intensities can be recorded and the phase information is lost. This imposes the so called *phase problem* which can be solved using iterative phase retrieval techniques in combination with sufficiently sampled intensity data. The first algorithms of this kind were developed in the 1970's [3, 4] (mostly for electron microscopy) and became more and more advanced over the years [5]. Since its successful experimental demonstration [6], CDI has become a popular choice for high-resolution imaging with applications across many disciplines.

The increased popularity of CDI based methods came together with the development of X-ray free-election lasers (XFELs) which produce extremely bright and coherent X-rays. One such method, which makes use of these unique properties, is flash X-ray diffractive imaging (FXI). In FXI, single non-crystalline bio-particles are injected as aerosols and intersected with the intense XFEL pulses which are short enough to outrun key radiation damage processes [7]. This idea of "diffraction before destruction" enables structural studies of

bio-molecules and viruses. First successfully demonstrated in 2006 at the Free-electron LASer in Hamburg (FLASH) [8] and followed up by a number of successful experiments at the Linac Coherent Light Source (LCLS) [9--15], FXI has not yet reached its full potential. In 2014, the LCLS has invited the entire FXI community and started an initiative to tackle current challenges [16]. The ultimate goal of this technique is to obtain three-dimensional structures of small molecules and viruses at atomic resolution. This requires the use of hard X-rays (wavelengths below a nanometre) and makes this experiment even more challenging as signal levels rapidly decrease with increasing photon energy. This leads to the core question: How do we get there? In this thesis, we tried to find answers to that question.

Another popular CDI technique is X-ray ptychography, in which a focused X-ray beam (typically produced by a synchrotron) is scanned across a fixed target such that individual illuminated areas overlap on the sample. This idea goes back to Walter Hoppe [17] and permits quantitative imaging of extended objects [18] and is also capable to reconstruct the illumination profile alongside the object [19]. This makes ptychography a valuable tool for wavefront characterization of X-ray beams, which can help to interpret the data of FXI experiments.

After giving a short introduction to bright X-ray sources (chapter 2), the theory behind CDI is described in chapter 3 with a special focus on FXI and ptychography. Chapter 4 summarizes the current status of FXI and proposes strategies towards the goal of imaging small molecules and viruses at atomic resolution using hard X-rays (**Paper I**). Chapter 5 is dedicated to an experiment which uses ptychography to learn more about the wavefront of focused XFEL beams (**Paper V**) relevant for FXI experiments. Finally, in chapter 6, we describe algorithms and software that have been developed for the analysis of data from FXI (**Paper I**) and ptychography (**Paper III**) experiments. These methods are implemented both for robust offline and fast real-time analysis, the latter giving feedback on data quality and experimental conditions already during data collection (**Paper II**,**IV**).

With this work, we hope to bring the development of both FXI and ptychography closer to their potential and are convinced that CDI based methods will further improve and become standard tools for high-resolution imaging in many disciplines such as Structural Biology and Material Science.

2. Bright X-ray sources

The past decades have given us access to very bright and tunable accelerator based X-ray sources, which in turn have found applications in many areas of science. The most common of those X-ray sources are synchrotrons and X-ray free-electron lasers.

2.1 Undulator radiation

If relativistic electrons are sent through an undulator, which is a pair of periodic magnetic structures, they follow a sinusoidal trajectory due to the Lorentz force acting on the charges. Since the electrons travel at relativistic speeds, they emit light in form of undulator radiation with increasing intensity over the length of the undulator. The fundamental wavelength of this emitted radiation is given by

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

where λ_u is the period of the undulator, γ the energy of a relativistic electron with speed v, charge e and mass m_e in units of its rest energy

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}\,, \tag{2.2}$$

and K the so called *undulator parameter*

$$K = \frac{e\lambda_u B_0}{2\pi m_e c}\,, \tag{2.3}$$

which scales linearly with the strength of the magnetic field B_0. By increasing/decreasing the gap size between the two periodic arrays, the magnetic field gets weaker/stronger allowing the wavelength of the emitted X-rays to be tuned.

2.2 Synchrotron

In a modern synchrotron, electrons are injected into a storage ring. In such a ring, so called bending magnets keep the electrons on their trajectory while tangentially emitting synchrotron radiation. On straight section of the storage ring, insertion devices such as an undulator emit undulator radiation. Both bending magnets and undulators are used as the source point for experimental X-ray end stations (beamlines). In a typical synchrotron, γ can be estimated to be of order 10^4 (5 GeV electron energy, 0.511 MeV mass at rest) giving rise to an X-ray beam with wavelength $\lambda \sim 1$ Å at the exit of an undulator with a period of 1 cm. These X-rays usually come at a rate of $\sim$ 100 MHz with pulse lengths on the order of 100 picoseconds. They have a high degree of transversal coherence and are about 10 orders of magnitudes brighter than the ones produced by conventional X-ray sources (e.g. rotating anode).

2.3 X-ray free-electron laser

With an electron gun and a linear accelarator (LINAC) instead of a storage ring, it is possible to produce and compress electrons into very short and dense bunches on the order of $\sim$ 100 fs. At an X-ray free-electron laser such high electron density clouds are sent into long unduluators which build up a radiation field increasing with distance (undulator radiation). As the field strength rises, the electron bunches start to interact with radiation field and form compressed micro-bunches with a separation equal to the fundamental undulator wavelength λ and its intensity growing exponentially with distance until the intensity saturates. This effect is called Self Amplified Stimulated Emission (SASE) and is able to produce X-ray pulses as short as a few femtoseconds which are about 10 orders of magnitude brighter than the ones produced by a synchrotron. Much like for synchrotrons, the X-ray wavelength is tunable with the *undulator parameter*, with even better transversal coherence properties. The first X-ray laser reaching hard X-rays was the LINAC coherent light source (LCLS) which started to operate in 2009 [20]. Since then, it delivers short and intense X-ray pulses at a rate of 120 Hz. The European X-ray free-electron laser (EuXFEL), which started to operate in mid 2017 with limited capabilities, is expected to deliver X-rays at a repetition rate of 27 kHz once it reaches full operation mode. Also, the LCLS is currently planning an upgrade to a MHz rate for 2020 (LCLS-II).

3. Coherent diffractive imaging

X-rays, as produced by a synchrotron or an XFEL, can be approximated[1] as a monochromatic scalar wave

$$\Psi(x, y, z, t) = \tilde{\psi}(x, y, z) \exp(-i\omega t) \tag{3.1}$$

with frequency ω traveling along the z direction, x and y span the transversal plane. The wave is characterized by the wave number

$$k = \omega/c = \frac{2\pi}{\lambda} \tag{3.2}$$

where c is the speed of light in vacuum. The wavelength λ is related to the energy of a photon by

$$\varepsilon_{h\nu} = \frac{hc}{\lambda} \tag{3.3}$$

with h being Planck's constant.

To understand the interaction of this wave with an inhomogeneous medium described by the refractive index $n(x, y, z)$, the time-independent scalar Helmholtz equation[2]

$$\left[\nabla^2 + k^2 n^2(x, y, z)\right] \tilde{\psi}(x, y, z) = 0 \tag{3.4}$$

needs to be solved. The time dependence is well described by equation (3.1).

With equation (3.4) as a starting point, the following paragraphs build up the formalism necessary to describe FXI and ptychography experiments and the image reconstruction steps involved.

[1] The natural bandwidth of SASE XFEL pulses with $\Delta\omega/\omega \sim 10^{-3}$ is sufficiently monochromatic in the context of imaging, at synchrotrons it is common to use monochromators to achieve similar bandwidths.

[2] For a derivation, see for example chapter 2.1 of *Coherent X-ray Optics* [21].

3.1 X-ray interaction with matter

Let us consider an experiment in which incoming X-rays interact with a sample volume as shown in Fig. 3.1. If we describe the incoming X-ray beam as a

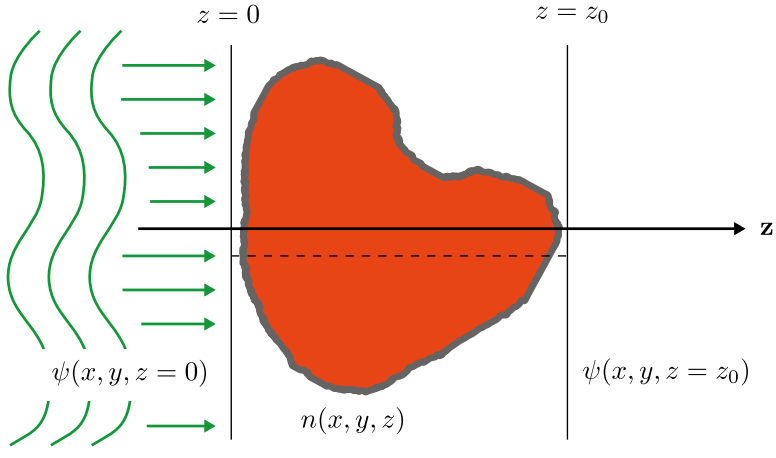

Figure 3.1. An incoming scalar wave field $\psi(x, y, z = 0)$ interacts weakly with a sample volume as defined by the refractive index $n(x, y, z)$. The volume is confined between the planes $z = 0$ and $z = z_0$ where the outgoing wave field is described as $\psi(x, y, z = z_0)$.

plane wave traveling along z with an envelope $\psi(x, y, z)$, we can write

$$\tilde{\psi}(x, y, z) = \psi(x, y, z) \exp(ikz) \tag{3.5}$$

and use this expression as an ansatz to solve equation (3.1), which yields[3]

$$\left\{2ik\partial_z + \partial_x^2 + \partial_y^2 + \partial_z^2 + k^2 \left[n^2(x, y, z) - 1\right]\right\} \psi(x, y, z) = 0 \tag{3.6}$$

where ∂_z, ∂_x^2, ∂_y^2 and ∂_z^2 denote first and second spatial derivatives. Assuming that ψ is "beam-like", meaning that it varies much stronger along x and y than in the direction of z, we can neglect the ∂_z^2 term (*paraxial approximation*). Furthermore, assuming that individual rays are not coupled to their neighbors as they pass through the sample, we can also neglect the terms ∂_x^2 and ∂_y^2 (*projection approximation*) and rewrite (3.6) to

$$\partial_z \psi(x, y, z) = \frac{k}{2i} \left[1 - n^2(x, y, z)\right] \psi(x, y, z)\,. \tag{3.7}$$

[3] See e.g. chapter 2.2 of *Coherent X-ray Optics*[21] for the intermediate steps.

Solving this partial differential equation for the case shown in *Figure 3.1*, we can define the wave at the exit of the sample volume ($z = z_0$)

$$\psi(x, y, z_0) = \exp\left\{\frac{ki}{2}\int_{z=0}^{z=z_0}\left[n^2(x, y, z) - 1\right]\mathrm{d}z\right\}\psi(x, y, 0) \tag{3.8}$$

in relation to the wave at the entrance ($z = 0$). For X-rays, the refractive index is typically written in the form

$$n(x, y, z) = 1 - \delta(x, y, z) + i\beta(x, y, z) \tag{3.9}$$

where δ and β are real numbers close to zero. We can thus approximate the expression $n^2 - 1 \approx 2(n - 1)$ and write equation (3.8) as

$$\psi(x, y, z_0) = \exp\left\{-ik\int_0^{z_0}\left[\delta(x, y, z) - i\beta(x, y, z)\right]\mathrm{d}z\right\}\psi(x, y, 0)\,. \tag{3.10}$$

We can directly see that $\int_z \delta(x, y, z)\mathrm{d}z$ is associated with phase shifts induced by the sample material and by taking the squared modulus of equation (3.10), we obtain Beer's law

$$I(x, y, z_0) = |\psi(x, y, z_0)|^2 = \exp\left[-\int_0^{z_0}\mu\mathrm{d}z\right] I(x, y, 0) \tag{3.11}$$

which describes the absorption properties $\mu = 2k\beta(x, y, z)$ of the material.

We have so far described the wave field as one entity that interacts with the sample as it propagates through the medium. Defining the coordinate vector as $\mathbf{x} = (x, y, z)$ and the wave vector $\mathbf{k}_0$ with $|\mathbf{k}_0| = k$, we can formulate the incoming wave as a multiplication of a "beam-like" profile with a plane wave

$$\psi_0(\mathbf{x}) = \psi(x, y, 0)\exp(i\mathbf{k}_0 \cdot \mathbf{x}) \tag{3.12}$$

Choosing a different perspective to understand X-ray interaction, we can now describe the wave behind the sample volume at location $\mathbf{x}$ as the superposition of the incoming wave $\psi_0(\mathbf{x})$ and the coherent sum of all spherical waves that have originated from single-scattering events located at position $\mathbf{x}'$ within the

sample, namely

$$\psi(\mathbf{x}) = \psi_0(\mathbf{x}) + \frac{k^2}{4\pi} \iiint \frac{\exp(ik|\mathbf{x}-\mathbf{x}'|)}{|\mathbf{x}-\mathbf{x}'|} \left[n^2(\mathbf{x}') - 1\right] \psi_0(\mathbf{x}')\mathrm{d}\mathbf{x}' \,. \quad (3.13)$$

This expression is valid under the assumption that the incident wave is equal to $\psi_0(\mathbf{x}')$ for all scattering events (*first Born approximation*) which applies for weak X-ray interactions. It can be shown that equations (3.10) and (3.13) are equivalent for the case of weakly interacting X-rays and optically thin materials[4].

3.2 Free-space propagation

We are now aiming to describe the propagation of the wave $\psi(x, y, z_0)$ downstream of the sample in a scenario like the one shown in *Figure 3.1*. For this geometry, we define a wave vector $\mathbf{k} = (k_x, k_y, k_z)$ of length $|\mathbf{k}| = k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ pointing in the outgoing propagation direction. With this setup, we can write the propagated wave at a distance $r > z_0$ as

$$\psi(x, y, z = r) = \mathcal{D}_r \psi(x, y, z = z_0) \,, r \geq z_0 \quad (3.14)$$

using the free-space propagator defined as

$$\mathcal{D}_r = \mathcal{F}^{-1} \exp\left[ir\sqrt{k^2 - k_x^2 - k_y^2}\right] \mathcal{F} \quad (3.15)$$

where

$$\mathcal{F} : f(x, y) \mapsto F(k_x, k_y) = \frac{1}{2\pi} \iint f(x, y) \exp\left[-i(k_x x + k_y y)\right] \mathrm{d}x\mathrm{d}y \quad (3.16)$$

is the two-dimensional Fourier transform and

$$\mathcal{F}^{-1} : F(k_x, k_y) \mapsto f(x, y) = \frac{1}{2\pi} \iint F(k_x, k_y) \exp\left[i(k_x x + k_y y)\right] \mathrm{d}k_x\mathrm{d}k_y \quad (3.17)$$

its inverse. Equation (3.14) is known as free-space propagation and describes a solution to the Helmholtz equation (3.4) for vacuum ($n = 1$) or air ($n \approx 1$).

[4] see e.g. chapter 2.9 of *Coherent X-ray Optics* [21].

Assuming that all non-zero plane wave components of the propagating wave describe a small angle with respect to the optical axis (*paraxial approximation*), we can approximate

$$\sqrt{k^2 - k_x^2 - k_y^2} \approx k - \frac{k_x^2 + k_y^2}{2k} \tag{3.18}$$

and rewrite equation (3.2) to

$$D_r \approx \exp(ikr)\, \mathcal{F}^{-1} \exp\left[\frac{-ir(k_x^2 + k_y^2)}{2k}\right] \mathcal{F} \tag{3.19}$$

which is known as the Fresnel propagator. Making use of the convolution theorem in combination with (3.2) and (3.19), it is possible to derive an alternate form of the Fresnel diffraction integral[5]

$$\begin{aligned}\psi(x, y, z = r) = -\frac{ik \exp(ikr)}{r} \iint \psi(x', y', z = z_0) \\ \exp\left\{\frac{ik}{2r}\left[(x - x')^2 + (y - y')^2\right]\right\} \mathrm{d}x' \mathrm{d}y' \,.\end{aligned} \tag{3.20}$$

3.3 Far-field diffraction

For most diffraction experiments, the propagation distance is much larger than the characteristic length scale b of the unpropagated wave. This condition is satisfied if the dimensionless Fresnel number

$$F_N = \frac{kb^2}{2\pi} \tag{3.21}$$

is much smaller than unity, and is often called Fraunhofer or far-field regime. Equation (3.20) rewrites to

$$\psi(x, y, z = r) = -\frac{ik \exp(ikr)}{r} \exp\left[\frac{ik}{2r}\left(x^2 + y^2\right)\right] (\mathcal{F}\psi)\left(\frac{kx}{r}, \frac{ky}{r}, z = z_0\right) \tag{3.22}$$

[5] see e.g. chapter 1.4 of *Coherent X-ray Optics* [21].

which again makes use of the Fourier transform $\mathcal{F}$, evaluated at reciprocal coordinates $\mathbf{q}_\perp = (q_x, q_y) = (kx/r, ky/r)$. By taking the squared modulus, it gives a simple expression for the relation between the exit wave $\psi(x, y, z_0)$ and the diffracted intensities

$$I(\mathbf{q}_\perp) = |\psi(x, y, z = r)|^2 = \frac{k^2}{r^2} \left|\mathcal{F}\psi(x, y, z = z_0)\right|^2 \tag{3.23}$$

in the far-field.

We have derived a formulation of diffraction under the assumptions of the *projection approximation*. It will also be useful to obtain a similar expression for the consideration of single-scattering events (*first Born approximation*). We start from equation (3.13) which defines the exit wave behind the sample volume and consider a situation where the observer is at a distance $r = |\mathbf{x}| \gg |\mathbf{x}'|$ far from the sample, allowing for the approximation[6]

$$\begin{aligned} \frac{\exp(ik|\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|} &\approx \frac{\exp\left[ikr\sqrt{1 - 2r^{-2}\mathbf{x} \cdot \mathbf{x}'}\right]}{r} \\ &\approx \frac{\exp\left[ikr(1 - 2r^{-2}\mathbf{x} \cdot \mathbf{x}')\right]}{r} . \end{aligned} \tag{3.24}$$

We identify the wave vector $\mathbf{k} = k\mathbf{x}/|\mathbf{x}| = k\mathbf{x}/r$ to point in the same direction as $\mathbf{x}$ and find an expression for the wave in the far-field

$$\begin{aligned} \psi(\mathbf{x}) &= \psi_0(\mathbf{x}) + \psi_1(\mathbf{q}) \\ &= \psi_0(\mathbf{x}) + \frac{k^2 \exp(ikr)}{4\pi r} \iiint \left[n^2(\mathbf{x}') - 1\right] \psi(x', y', 0) \exp(-i\mathbf{q} \cdot \mathbf{x}') \mathrm{d}\mathbf{x}' \end{aligned} \tag{3.25}$$

as a sum of the unscattered and the scattered wave, where $\mathbf{q} = \mathbf{k} - \mathbf{k}_0$. Here we can see that the scattered wave is proportional to the three-dimensional Fourier transform[7] of the scattering potential

$$\varphi(\mathbf{x}) = \frac{k^2}{4\pi} \left[n^2(\mathbf{x}) - 1\right] \approx \frac{k^2}{2\pi}[n(\mathbf{x}) - 1] \tag{3.26}$$

[6] see e.g. chapter 2.5.1 of *Coherent X-ray Optics* [21].

[7] similar to the definitions (3.16) and (3.17) but extended to three dimensions and with a prefactor $(2\pi)^{-3/2}$ instead of $(2\pi)^{-1}$.

times the incident beam profile $\psi(x, y, z = 0)$[8], and is evaluated on a series of points $\mathbf{q}$ which form a spherical surface in Fourier space, known as the *Ewald sphere*.

3.4 Discrete intensity measurements

In the previous chapters we have described an idealistic diffraction experiment. In a real experiment, diffraction data is recorded on a pixelated area detector. To account for this reality, we discretize $\mathbf{q}_\perp$ into a regular grid of $M \times M$ pixels and map each pixel (μ, ν) onto its reciprocal coordinate

$$\mathbf{q}_{\mu\nu} = (q_\mu, q_\nu) = (\mu\Delta q, \nu\Delta q) \quad \mu, \nu = 0...M-1 \tag{3.27}$$

where Δq is defined as pk/r and p is the size of a pixel. Most common X-ray detectors register the arrival of photons which follows a Poisson process. For each detector pixel, the probability of detecting $n_\mathbf{q}$ photons given intensity measurements $I_\mathbf{q}\Delta A$ is

$$p\left(n_\mathbf{q}|I_\mathbf{q}\Delta A\right) = \frac{(I_\mathbf{q}\Delta A)^{n_\mathbf{q}}}{n_\mathbf{q}!}\exp(-I_\mathbf{q}\Delta A)\,, \tag{3.28}$$

where $I_\mathbf{q} = I(\mathbf{q}_\perp)$ is defined as formulated in equation (3.23) and $\Delta A = p^2$ is the area of a pixel. This intensity definition has been derived using continuous Fourier transforms, as defined in (3.16) and (3.17) for two dimensions. Since we are dealing with discrete signals, we may instead use discrete Fourier transforms (DFTs). Given a discrete function $h(\mathbf{x}_{mn})$, with discrete real space coordinates

$$\mathbf{x}_{mn} = (x_m, x_n) = (m\Delta x, n\Delta x) \quad m, n = 0...M-1 \tag{3.29}$$

where $\Delta x = \lambda r/(pM)$ is the half-period resolution, we define the two-dimensional DFT as

$$\text{DFT} : h(\mathbf{x}_{mn}) \mapsto H\left[\mathbf{q}_{\mu\nu}\right] = M^{-1}\sum_{m=0}^{M-1}\sum_{n=0}^{M-1} h(\mathbf{x}_{mn})\exp\left[-i\mathbf{q}_{\mu\nu}\cdot\mathbf{x}_{mn}\right] \tag{3.30}$$

[8] can be interpreted as the same two-dimensional profile in each transveral slice of the scattering volume.

and

$$\text{IDFT} : H(\mathbf{q}_{\mu\nu}) \mapsto h\left[\mathbf{x}_{mn}\right] = M^{-1} \sum_{\mu=0}^{M-1} \sum_{\nu=0}^{M-1} H(\mathbf{q}_{\mu\nu}) \exp\left[i\mathbf{q}_{\mu\nu} \cdot \mathbf{x}_{mn}\right] \tag{3.31}$$

as its inverse[9]. Since the computational complexity of equations (3.30) and (3.31) is $\mathcal{O}\left((M^2)^2\right)$, we can most often use the implementation[10] of fast Fourier transforms FFTs which scale with a complexity of $\mathcal{O}(M^2 \log(M^2))$ [23].

With the correct choice of sampling, the result of the DFT is identical to the one obtained by the continuous Fourier transform. An upper limit for this choice of sampling is defined by the Shannon[11] sampling theorem [1] which states that $\Delta q = 2\pi/s$ for a band-limited signal $H(\mathbf{q})$ within $[0, s/2]$, meaning that $h(|\mathbf{x}| > s/2) = 0$. For CDI, this condition is satisfied since the sample volume has a finite size s. It is useful to define the linear sampling ratio

$$\kappa = \frac{2\pi}{s\Delta q} \tag{3.32}$$

where $\kappa = 1$ denotes the case of "critical sampling" at the Shannon limit.

The goal of any CDI experiment is to recover structural information based on intensity measurements which means that phase information is lost and needs to be restored. This is possible if the intensity measurements are sufficiently "oversampled" ($\kappa > 1$), a realization which goes back to David Sayre who proposed such a strategy already in 1952 in the context of crystallography [2]. With this core concept of CDI in mind, we can close the general description of diffraction theory and move on to the more specific cases of FXI and X-ray ptychography.

3.5 Flash X-ray diffractive imaging

In a common FXI experiment, X-ray pulses produced by an XFEL are intersected with a stream of biological particles as shown in *Figure 3.2*. These X-rays can reach power densities of $10^{17}\,\mathrm{W\,cm^{-2}}$ or more in a single pulse and typically have photon energies between 1 and 10 keV. These pulses have enough power to destroy the sample in a single shot. However, with very short pulses it is possible to outrun key damage processes and capture structural information by means of recording a diffraction pattern, a concept known

[9] In three dimensions, the DFT and its inverse have a similar form, but with a prefactor $M^{-3/2}$ instead of M^{-1}.

[10] a definition can be found in *Numerical recipes* [22].

[11] Also known as the Nyquist-Shannon sampling theorem.

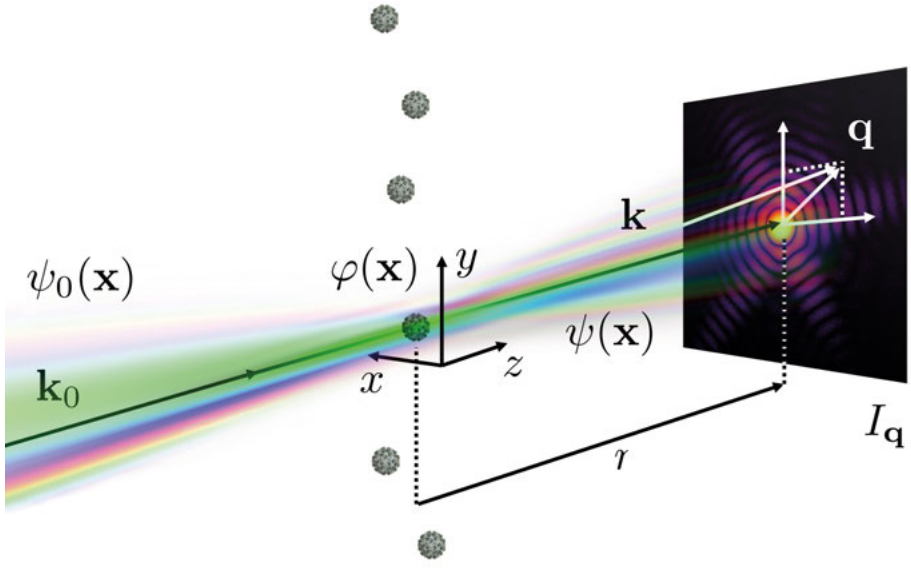

Figure 3.2. Geometry for FXI experiments. A stream of biological particles is brought into the focus of a strong and short X-ray pulse. The incident X-ray wave $\psi_0(\mathbf{x})$ travels in the direction of $\mathbf{k}_0$ and interacts (elastic scattering) with a particle described by its scattering potential $\varphi(\mathbf{x})$. Before the key damage processes start to act on the particle, the outgoing wave $\psi(\mathbf{x})$ propagates into the far-field where a pixel area detector is placed at distance r from the interaction region. At reciprocal location $\mathbf{q}$, the detector records the scattered intensity in the direction of $\mathbf{k}$ forming a diffraction pattern $I_{\mathbf{q}}$. Based on this pattern, the structure of the particle can be reconstructed.

as "diffraction before destruction" [7]. At the given photon energies, the most relevant interaction processes are elastic scattering, photon absorption and (inelastic) compton scattering. The latter two are attributed to radiation damage by transferring energy to atoms and molecules which causes the ejection of electrons and subsequent disorder of the structure (Auger decay and secondary processes) [24]. The process which proves to be most useful for FXI is elastic scattering on electrons as it leaves the structure unchanged. Considering a plane wave being incident on a cloud of free electrons with number density $\rho(\mathbf{x})$, we can write

$$\varphi(\mathbf{x}) = r_0 \rho(\mathbf{x}) \tag{3.33}$$

for the scattering potential, where r_0 is the classical electron radius. To also account for the fact that electrons in a molecule are bound to atoms of different

species a, we can write

$$\varphi(\mathbf{x}) = r_0 \sum_a \rho_a(\mathbf{x}) f_a(\lambda)\,, \tag{3.34}$$

where $f_a(\lambda)$ are the tabulated wavelength-dependent atom scattering factors [25] relative to the scattering strength of a free electron. Since this expression for the scattering potential is equivalent to (3.26), we can use (3.25) and write for the scattered wave from a particle (e.g. a bio-molecule)[12]

$$\psi_1(\mathbf{q}) = \frac{\psi_0}{r}\mathrm{DFT}\left[r_0 \sum_a \rho_a(\mathbf{x}) f_a(\lambda)\right] \tag{3.35}$$

making the assumption that the X-ray beam as seen by the particle has a "flat" profile with constant intensity $|\psi_0|^2$.

3.5.1 Sphere diffraction

For the purpose of determining the size of a particle based on its diffraction pattern, it can be useful to model the data with an analytical expression for diffraction from simple objects, such as a sphere (see section 6.4). For a sphere with homogeneous density, we can write the scattering potential as

$$\varphi_{\text{sphere}}(\mathbf{x}) = \begin{cases} \frac{k^2}{2\pi}(n-1), & |\mathbf{x}| < R \\ 0, & |\mathbf{x}| > R \end{cases} \tag{3.36}$$

where R is the radius of the sphere and n the refractive index of the homogeneous material which can also be formulated as[13]

$$n = 1 - \frac{2\pi}{k^2} r_0 \sum_a \rho_a f_a(\lambda)\,. \tag{3.37}$$

[12] We have ignored the phase factor $\exp[ikr]$ since we can only measure intensities in FXI.
[13] combining (3.26) and (3.34)

Using equation (3.25), the scattered wave (or scattering factor) of a sphere at $q = |\mathbf{q}|$ is defined as

$$\begin{aligned}\psi_{1,\text{sphere}}(q) &= \psi_0 \frac{k^2(n-1)}{2\pi r} \iiint \exp[-i\mathbf{q}\cdot\mathbf{x}]\mathrm{d}\mathbf{x} \\ &= \psi_0 \frac{k^2(n-1)}{2\pi r} \frac{4\pi}{q} \int_0^R x^2 \sin qx \mathrm{d}x \\ &= \psi_0 \frac{k^2(n-1)}{2\pi r} \frac{4\pi}{q} \left(\frac{\sin(qR)}{q^2} - R\frac{\cos(qR)}{q} \right)\end{aligned} \tag{3.38}$$

This leads to a simple expression for the diffraction pattern from a homogeneous sphere[14]

$$I_{\text{sphere}}(q) = |\psi_0|^2 \left[\frac{8\pi^2 R^3(n-1)}{\lambda^2 r} \right]^2 \left| \frac{\sin(qR) - qR\cos(qR)}{(qR)^3} \right|^2 . \tag{3.39}$$

3.5.2 2D imaging

A single two-dimensional diffraction pattern $I_\mathbf{q}$ is related to the projected two-dimensional scattering potential $\varphi_\perp(x,y) = \int_z \varphi(\mathbf{x})\mathrm{d}z$ via

$$I_\mathbf{q} = \frac{\psi_0}{r} \left| \text{DFT}\left[\varphi_\perp(x,y)\right] \right|^2 \tag{3.40}$$

provided that there is no "lift off" of the Ewald sphere, which means that $q_z z$ is small and thus $\exp[iq_z z] \approx 1$ for all values of z inside the particle. In 2D FXI, we basically have to solve the inverse problem of equation (3.40), namely

$$\varphi_\perp(x,y) = \frac{r}{\psi_0} \text{IDFT}\left[\sqrt{I_\mathbf{q}} \exp(-i\phi_\mathbf{q}) \right] \tag{3.41}$$

where $\phi_\mathbf{q}$ is the lost phase information to be recovered by means of oversampling the intensity measurements. This requirement is full-filled for linear sampling ratios[15] $\kappa > 2$, which can be seen as the limit where the diffracted intensities are band-limited within $[0, s]$ since the autocorrelation of $\varphi_\perp$ is equal to $\text{IDFT}(\text{I}_\mathbf{q})$ and extends to $2s$.

[14] Note that the same expression for the scattered intensity of a sphere in **Paper I** has two mistakes. The factor of 3 in the denominator should be removed and s_i should be defined as $2\pi(d/2)|\mathbf{q}_i|$

[15] as defined in (3.32)

In FXI, the particles are isolated and we know that the scattering potential $\varphi_\perp(x, y)$ is zero outside a certain area which is denoted the *Support* $\mathcal{S}$. This scenario is described by the function

$$S(x,y) = \begin{cases} 1, & \forall x, y \in \mathcal{S} \text{ and} \\ 0, & \text{otherwise}\,. \end{cases} \tag{3.42}$$

Imagine the following iterative algorithm

1. Start with an initial guess for $\varphi_\perp(x, y)$ (e.g. random noise)
2. Multiply $\varphi_\perp(x, y)$ with the support function $S(x, y)$
3. Compute $\psi_1(\mathbf{q}_\perp)$ using the forward FFT
4. Replace the amplitude of $\psi_1(\mathbf{q}_\perp)$ by $\sqrt{I_\mathbf{q}}$ and keep the phase
5. Compute $\varphi_\perp(x, y)$ using the inverse FFT
6. Repeat steps 2 through 4 until a certain convergence criteria is met.

This is known as the *Error Reduction* (ER) algorithm for phase retrieval with isolated objects [3, 4]. It can be described as iterative projections onto two constraint sets, namely the *support constraint* and the *Fourier intensity constraint*. Since the latter is non-convex, the simple ER scheme is likely to get trapped in local minima. To avoid this problem, many different algorithms for iterative phase retrieval have been proposed[16]. The most popular ones are *Hybrid Input-Output* (HIO) [26], *Relaxed Averaged Alternating Projections* (RAAR) [27] and *Difference Map* (DM) [28]. Another modification of the reconstruction scheme outlined above is the *shrinkwrap* algorithm [29], which also updates or "shrinks" the support $S(x, y)$ while iteratively recovering the phase.

A common strategy for 2D image reconstruction in FXI is to use an algorithm like RAAR in combination with *shrinkwrap* to get close to the global solution in phase space and finish with a few iterations of ER. This is exemplified in *Figure 4.6* which shows a 2D reconstruction of a virus particle, based on experimental data.

3.5.3 Validation

Depending on the choice of algorithm, parameters and the quality of the diffraction data, iterative phase retrieval may or may not converge to the correct solution. Thus, the validation of the obtained outcome is an important step of the reconstruction process. There are a number of metrics that can be used to

[16] A comprehensive overview and comparison of different algorithms can be found in [5].

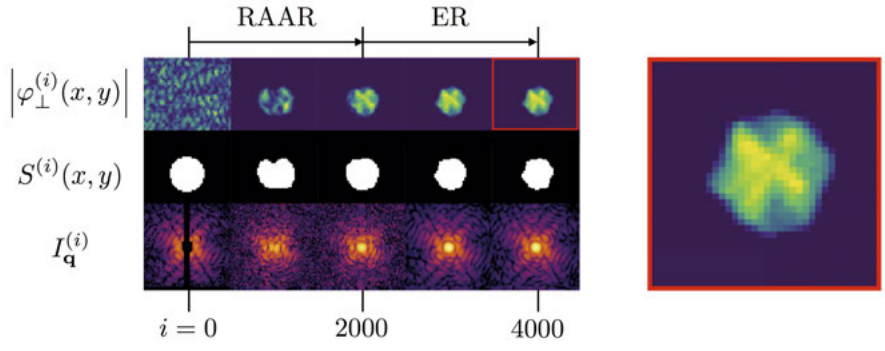

Figure 3.3. Example of a 2D image reconstruction by iterative phase retrieval in combination with *shrinkwrap* using 2000 iterations of RAAR followed by 2000 iterations of ER.

validate a given phase retrieval result. The Fourier space error, defined as

$$\epsilon_F = \sqrt{\frac{\sum_{\mathbf{q}} \left(|\mathrm{DFT}[\varphi_\perp]_{\mathbf{q}}| - \sqrt{I_{\mathbf{q}}}\right)^2}{\sum_{\mathbf{q}} I_{\mathbf{q}}}}\,, \tag{3.43}$$

indicates how well the estimated scattering potential is described by the data. Reconstructions with high ϵ_F should be regarded as failed phase searches. The real space error, defined as

$$\epsilon_R = \sqrt{\frac{\sum_{x,y}[1 - S(x,y)]|\varphi_\perp(x,y)|^2}{\sum_{x,y}|\varphi_\perp(x,y)|^2}}\,, \tag{3.44}$$

measures the integrated power outside of the support. Reconstructions with a high ϵ_R suggest that the support has not been chosen correctly and should be adjusted. Another metric testing the reproducibility of a given algorithm is the *Phase Retrieval Transfer Function* (PRTF) [8, 30], defined as

$$\mathrm{PRTF}(\mathbf{q}_\perp) = \frac{\langle \mathrm{DFT}[\varphi_\perp]_{\mathbf{q}} \rangle_N}{\sqrt{I_{\mathbf{q}}}} \tag{3.45}$$

where $\langle\cdot\rangle_N$ is the average over N reconstructions. For recovered phases identical over all N, the PRTF has a value of 1, while it approaches $1/\sqrt{N}$ for completely random phases. This metric can be used to define the spatial res-

olution up to which a given reconstruction is reproducible, a commonly used resolution cut is $1/e$ [31]. It has been suggested to incorporate a Wiener filter into the PRTF to suppress artifacts due to noise [32]. Another approach to check against overfitting to noise, is the calculation of the *Fourier Shell Correlation* (FSC). This is a common validation technique in electron microscopy which is based on a comparison of two reconstructions from independent measurements. This idea has been adopted to FXI by splitting one diffraction measurement into two measurements [11].

3.5.4 3D imaging

We have so far described how it is possible to recover structural information of a biological particle based on a single diffraction measurement. However, equation (3.35) describes the relation of a three-dimensional electron density and its representative in Fourier space. This already indicates a potential route towards 3D imaging. Imagine we can find a way to combine multiple diffraction patterns originating from identical copies of the same particle in random orientations and sufficiently sample the Fourier intensity volume $I(\mathbf{q})$. Now we can employ the same iterative phase retrieval techniques as described in the previous sections and reconstruct a three-dimensional structure. The problem of orientation recovery based on a set of diffraction patterns has been approached with different types of algorithms. A popular choice is the *Expansion, Maximization and Compression* (EMC) algorithm [13, 33], which is clustering the obtained diffraction intensities into slices through the intensity model $I(\mathbf{q})$. The algorithm includes three main steps. First, an initial model $I^{(i=0)}(\mathbf{q})$ is *expanded* into a number of views, which are then used as the basis for a clustering of the data based on an expectation-*maximization* approach and finally the result is *compressed* into a new intensity model $I^{(i+1)}(\mathbf{q})$. This procedure is repeated iteratively until a convergence criterion is met.

3.6 X-ray ptychography

In a typical X-ray ptychography experiment, a focused X-ray beam as for example produced at a synchrotron by a *Fresnel Zone Plate* (FZP) or an X-ray lens is scanned across a fixed target as shown in *Figure 3.4*. Starting from equation (3.8), we can define the exit wave at scan position $\mathbf{x}_j$ as

$$\psi_j(\mathbf{x}) = P_{\mathbf{x}-\mathbf{x}_j} O_{\mathbf{x}} \tag{3.46}$$

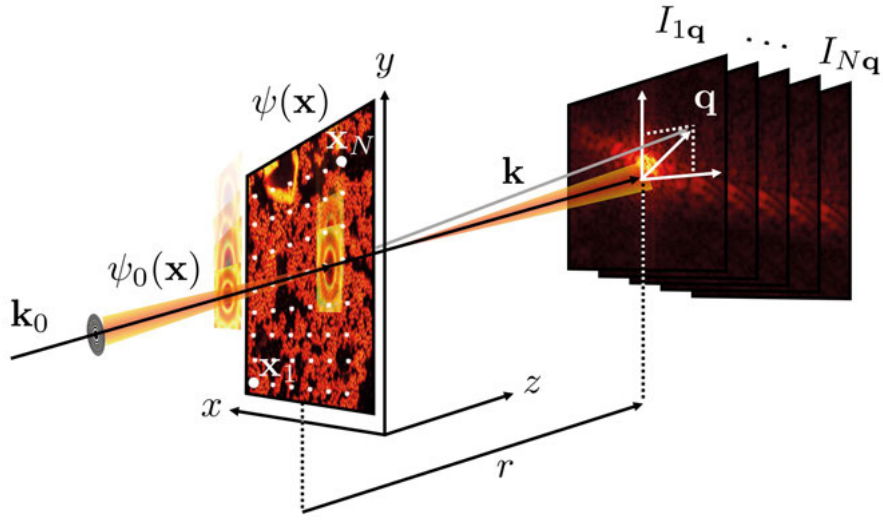

Figure 3.4. Geometry for X-ray ptychography experiments. A focused X-ray beam $\psi_0\mathbf{x}$ travelling in the direction of $\mathbf{k}_0$ is scanned across a specimen from position $\mathbf{x}_1 = (x_1, y_1)$ to $\mathbf{x}_N = (x_N, y_N)$ such that neighboring illuminations overlap. For each scan position with index j, the outgoing wave propagates into the far-field where a pixel area detector is placed at a distance r from the sample. At reciprocal location $\mathbf{q}$, the detector records the diffracted intensity in the direction of $\mathbf{k}$ forming diffraction patterns $I_{j\mathbf{q}}$. The N diffraction patterns can be combined to reconstruct a common exit wave $\psi(\mathbf{x})$ holding structural information on sample and illumination. Figure adapted from Fig. 1 in **Paper III**.

where

$$P_{\mathbf{x}-\mathbf{x}_j} = \psi(x - x_j, y - y_j, z = 0) \tag{3.47}$$

is the probe (or illumination) and

$$O_{\mathbf{x}} = \exp\left\{\frac{ki}{2}\int_0^{z_0}\left[n^2(x,y,z) - 1\right]\mathrm{d}z\right\} \tag{3.48}$$

is the object (or sample) after integration in z. Based on (3.23) and (3.30), we can write

$$I_{j\mathbf{q}} = \frac{k^2}{r^2}\left|\mathrm{DFT}[P_{\mathbf{x}-\mathbf{x}_j}O_{\mathbf{x}}]\right|^2 \tag{3.49}$$

for a diffraction pattern at scan position index j. Similar to FXI, this imposes a phase problem which can be solved by means of oversampling the measured intensities. But instead of applying the *support constraint*, ptychography uses the shared information between overlapping regions in real space to compensate for the lack of phase information in Fourier space. This idea goes back to Walter Hoppe, who realized already in 1969, in the context of electron diffraction from crystals, that the phase ambiguity can be removed by translating the lattice and recording at least two diffraction patterns [17]. The condition for sufficient sampling $\kappa \geq 1$ still applies to a single diffraction measurement in ptychography, although the extent s refers to the size of the illumination in that case. However, to state the minimal sampling requirements for a complete ptychography dataset, not only the size and shape of the illumination but also the amount of overlap between adjacent scan points needs to be considered [34]. An example for an iterative reconstruction scheme with ptychography is given in **Paper III** using the RAAR algorithm [27]. Other popular algorithms for ptychographic reconstruction are the *extended Ptychographic Iterative Engine* (ePIE) [35, 36], DM adapted to ptychography [19, 28] and *Maximum Likelihood* (ML) optimization [37]. First introduced for DM [38], most of these algorithms have been adapted such that they independently reconstruct the probe alongside the reconstruction of the object. This allows ptychography to be used as a way to characterize X-ray optics [39] or the wavefront of XFELs [40].

A recent extension to ptychography allows the reconstruction of state mixtures and is able to capture variations in the probe, the object or the detection of diffracted intensities [41]. For the most common case of having an X-ray beam that fluctuates with different exposures during a ptychography scan, the probe can be expanded to describe different modes $P^{(m)}_{\mathbf{x}-\mathbf{x}_j}$ all captured in the diffraction measurements

$$I_{j\mathbf{q}} = \frac{k^2}{r^2} \sum_m \left| \mathrm{DFT} \left[P^{(m)}_{\mathbf{x}-\mathbf{x}_j} O_{\mathbf{x}} \right] \right|^2 \tag{3.50}$$

as an incoherent sum.

4. Experimental strategies for flash X-ray diffractive imaging

Most of the recent FXI experiments have been done on rather large objects ($>$ 100 nm) and in the soft X-ray regime (500 to 2000 eV) [9--13, 42, 43]. But FXI has the potential to image much smaller biological objects ($<$ 50 nm) at high spatial resolution permitted by the use of hard X-rays (5 to 10 keV). However, such experiments have proven to be quite challenging. In **Paper I**, we described one of the first FXI experiments performed on a small virus particle ($\sim$ 40 nm) using hard X-rays. The data was collected at the *Coherent X-ray Imaging* (CXI) end station at the LCLS. By trying to understand the diffraction data for this experiment, we learned about some critical problems for FXI and provided suggestions on how these problems can be overcome.

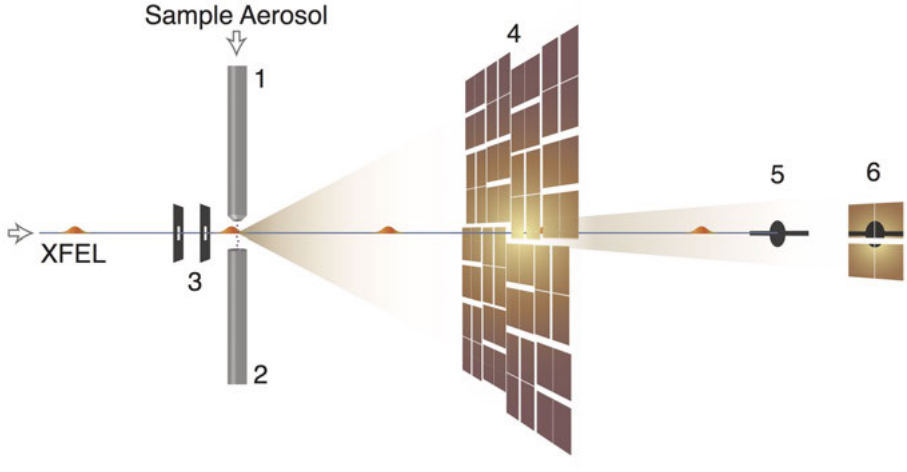

Figure 4.1. Experimental setup. Virus particles carried by aerosols are focused with an aerodynamic lens stack (1) into the XFEL focus where they either interact with the X-ray beam or travel through to the catcher (2). A pair of Si apertures (3) is placed close to interaction region to clean up the beam from stray light. The diffracted signal is recorded on a pair of pixel area detectors. The front detector (4) captures large diffraction angles, while the back detector (6) detects low resolution features of the virus. The direct beam is guided into a beam stop (5). Figure adopted from Fig. 1b in **Paper I**.

The experimental setup is sketched in *Figure 4.1*. The XFEL beam was focused by a pair of Kirkpatrick-Baez (KB) mirrors to form a nominal spot

size of $\sim$ 100 nm. The photon energy was 5.5 keV and the beam was delivered at the LCLS repetition rate of 120 Hz with an average pulse energy of 3.29 mJ, as measured upstream of the KBs. We used a pair of Si apertures to reduce unwanted stray light scattering, adding background to our diffraction measurements. The sample, 40 nm *Omono River Virus* (OmRV) [44], was aerosolized using a gas dynamic virtual nozzle (GDVN) [45, 46] with helium as carrier gas and focused into the interaction region by a aerodynamic lens stack [11] producing a $\sim$ 20 µm wide particle stream. The diffracted intensity signal was recorded by two Cornell-SLAC pixel-array detectors (CSPAD) [47] placed 497 mm and 2.4 m downstream of the interaction region, and both read out at the same rate as the LCLS repetition rate.

4.1 Data collection and processing

The CSPAD detectors consist of a number of application-specific integrated circuits (ASICs). The raw data values, measured in analogue-to-digital units (ADUs), were corrected by their common mode across these ASICs in addition to a subtraction of the average dark pedestals. The corrected signal on the small 140k back detector was used to discriminate "hits" from "misses" based on counting the number of lit pixels (see algorithm described in section 6.2). Within 4 hours of data collection, we recorded about 700 000 diffraction events under stable injection conditions, from which 5771 were identified as hits, equivalent to an average hit ratio of 0.83 %. Using a fitting-based classification approach (see algorithm described in 6.4), we assigned particle sizes and intensities to 4555 of those events. We assembled the ASICs on the 2.3M front detector and combined the diffraction patterns with the resampled diffraction on the back, as shown in *Figure 4.2* for a selection of diffraction hits.

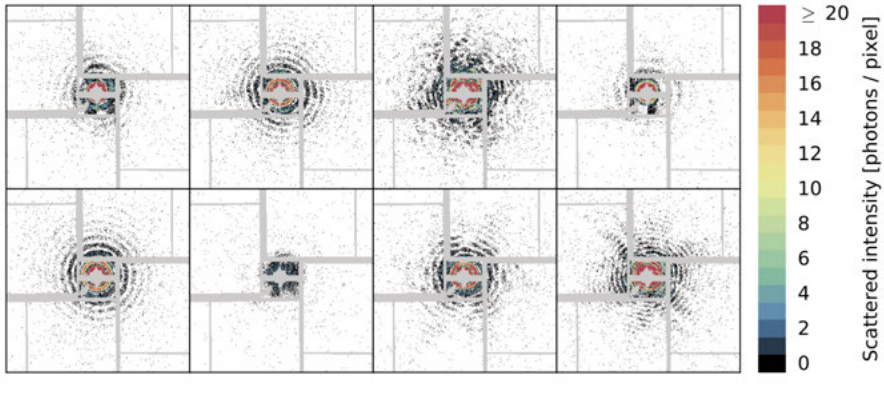

Figure 4.2. Selection of diffraction hits showing a large variety of features. Figure adopted from Fig. 4 in **Paper I**.

4.2 Photon-counting detector

Thanks to upgrades to the firmware of the CSPADs [48], we were able to discriminate the 1-photon signal (sample diffraction or background) from the 0-photon signal (which is due to detector noise) as shown in panels (a) and (b) of *Figure 4.3*. This was a major improvment compared to previous FXI ex-

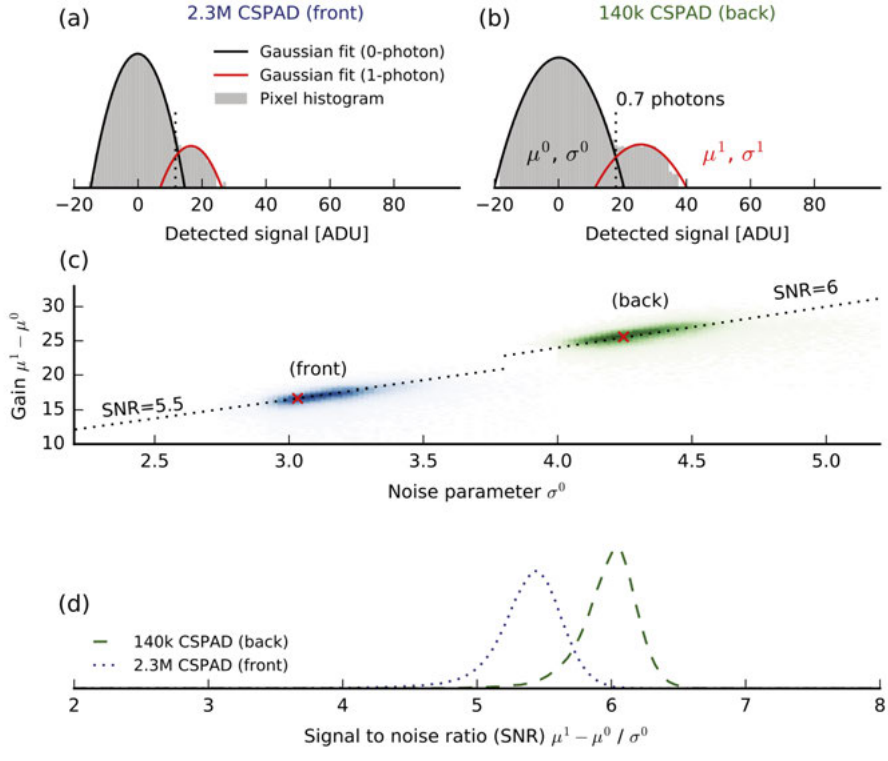

Figure 4.3. Photon counting on the CSPAD detectors. Single per-pixel histograms on back (b) and front (a) detector show a separation of the 0-photon and 1-photon peaks. Gain and noise parameters have been determined based on Gaussian fitting to many pixel histograms (c). Their linear relationship define average signal-to-noise ratios (SNRs) of 6 and 5.5 (d) for back and front detector, respectively. Figure adopted from Fig. 3 in **Paper I**.

periments with a similar setup, and it allowed us to define gain values for each pixel, based on fitting Gaussian functions to their individual histograms (see algorithm described in 6.1). Based on these gain values and per-pixel estimates for the detector noise, we obtained average signal-to-noise ratios (SNRs), as defined by gain over noise, of 6 and 5.5 for the back and front detector, respectively (see panels (c) and (d) in *Figure 4.3*). Using an empirically chosen threshold of 0.7 photons, we converted the corrected ADU signal to discrete per-pixel photon counts. This allowed us to accurately model the diffraction

signal using equations (3.39) and (3.40) together with (3.28) for the purpose of obtaining particle sizes and beam intensities.

4.3 Sample delivery

Over the past years, different sample delivery strategies for FXI have been suggested. Delivery methods which deposit the sample on a membrane [9, 43] or keep them in a thin water cell [42] suffer from significant background due to the sample substrate. Aeorosol injection methods [10--14, 49] have the advantage that potential background is reduced to the interaction with the carrier gas. This injection concept is outlined in *Figure 4.4*. The sample is

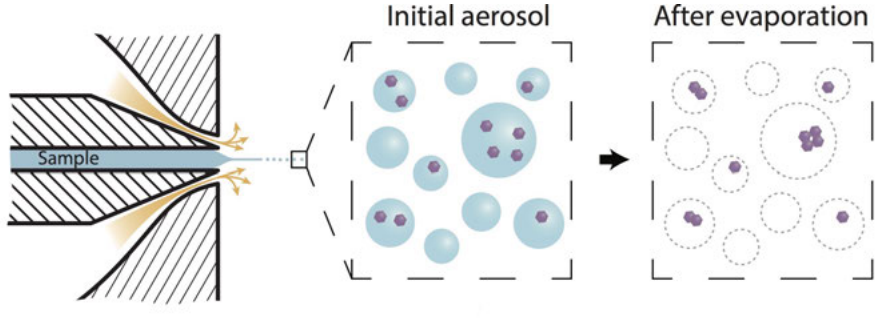

Figure 4.4. Schematic of droplet formation and evaporation for sample delivery with aeorosols using a GDVN. The sample in suspension is injected through the nozzle and guided by a carrier gas (e.g. Helium) to form a thin liquid jet which breaks up into droplets. As the droplets evaporate, the dry sample aerosol is left behind. Figure adopted from Fig. 1a in **Paper I**.

flowing through a GDVN forming a thin jet with the help a carrier gas. The jet is breaking into droplets which carry the sample as an aerosol to the interaction region and dry along the way due to the high pressure in the vacuum chamber, leaving the dry aerosol behind.

In **Paper I**, we detected particles with a large range of sizes as shown in (a) and (b) of *Figure 4.5*. We attributed this wide distribution to "caking" of non-volatile contaminants inside the large micron-sized droplets produced by the GDVN. In a more recent experiment, we used electrospray ionization (ESI) which is able to produce droplets much smaller in size ($\sim$ 100 nm). We injected virus particles of similar size (35 nm) and the data[1] results in a much cleaner particle size distribution, see panels (c) and (d) of *Figure 4.5*. This shows that producing sufficiently small droplets is important for FXI of small viruses and bio-molecules.

[1] unpublished, collected at AMO in June 2016 (experiment amol3416)

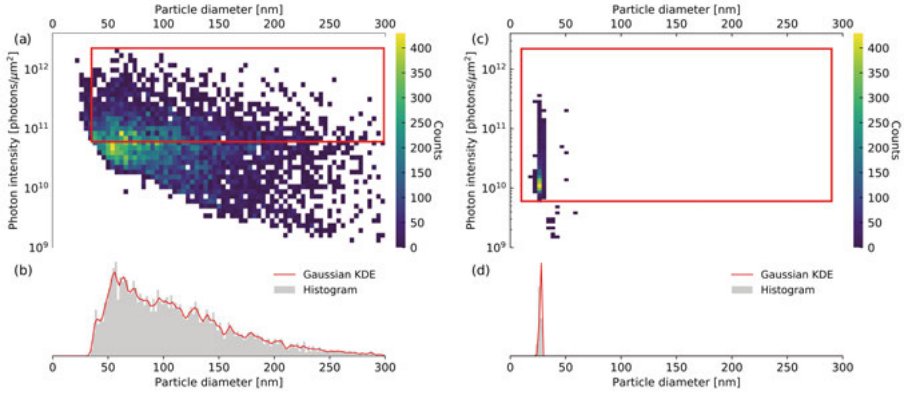

Figure 4.5. Comparison of particle size and intensity distributions with droplet formation via GDVN (a,b) and ESI (c,d). ESI produces smaller droplets and gives a much cleaner particle size distribution. Panels (a,b) are adapted from Fig. 5 in **Paper I**, while (c,d) are based on unpublished FXI data[1] on a 35 nm virus.

4.4 Signal and background

One of the core challenges for FXI at the moment is to maximize the ratio of signal to background. In **Paper I**, we estimated the maximum intensity on the sample to be 1.9×10^{12} photons/µm^2 and used the detected intensity distribution to argue that the intensity profile in the X-ray focus was larger than expected (> 100 nm). Furthermore, we found indications for local phase variation in the wavefront of the X-ray beam. All this numbers and observations are not as good as the current specifications of the CXI instrument and thus give reason to hope for an increase in signal levels with technical improvements done at present and future end stations and XFELs.

We also investaged present background levels, both from stray light scattering from the beamline and scattering from the gas surrounding the sample. We have observed strong background features close to the beam stop, which are attributed to scattering from X-ray optics or apertures and estimated the average background level to be on the order of 10^{-4} photons/pixel. The strong background features could be reduced by further optimizing the beam alignment with apertures and the existence of such features might be tolerable if the masked region does not become very large. The isotropic background scattering from the gas which is filling the entire detector, however, is a much bigger problem, especeially for high-resolution 3D imaging based on a large quantity of diffraction with very low signal at large diffraction angles. This issue can be adressed from two sides. Experimentally, it might be possible to reduce the amount of carrier gas used for the formation of droplets. Also, a post-sample shielding against background scattering from the gas should be beneficial [14, 50]. Another complimentary approach is to further develop and

integrate accurate background models into exisiting reconstruction algorithms [51].

4.5 2D Imaging

Due to the various challenges we faced in this experiment, as described in the previous sections, the collected data did not allow for a successful reconstruction of a three-dimensional structure based on many diffraction patterns as outline in section 3.5.4. However, we picked the sample-sized diffraction pattern with the strongest scattering signal and reconstuct a projected low-resolution 2D structure, as shown in *Figure 4.6*. We did 5000 phase searches with random

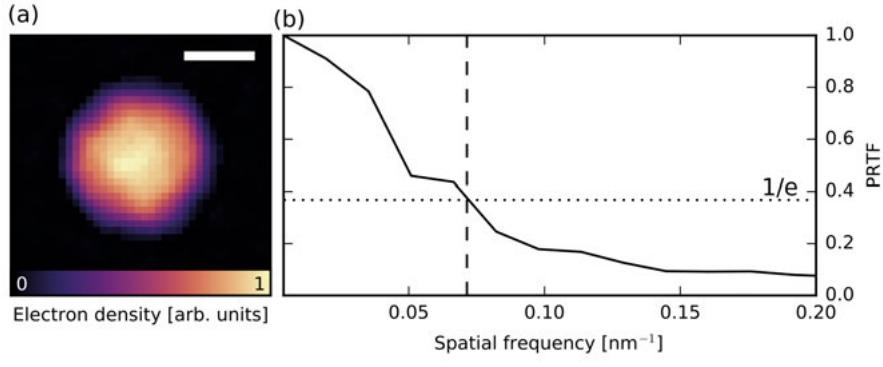

Figure 4.6. 2D reconstruction based on a strong diffraction pattern corresponding to a sample-sized particle. (a) Averaged magnitude image after 5000 independent reconstructions. The scale bar indicates 20 nm. (b) Radial average of the PRTF dropping below $1/e$ at a full-period resolution of 13.5 nm. Figure adopted from Fig. 11 in **Paper I**.

guesses for the initial image, each running 5000 iterations of the HIO algorithm followed by 1000 iterations of ER with a fixed support size of 48 nm. Based on the calcuated PRTF, we obtained a half-period resolution of 13.5 nm. Due to the limit in resolution, this reconstruction did not allow any further structural analysis. But this result shows that even at current signal and background levels, a large number of "pure" and homogeneous diffraction patterns might permit 3D imaging and could potentially improve the resolution down to about 5 nm, which would still be far from its potential, but nevertheless an important stepping stone towards high-resolution FXI.

5. Wavefront sensing of individual XFEL pulses

For many XFEL experiments, it is beneficial to have a good description of the wavefront in the X-ray focus. For FXI, the available tools for wavefront characterization are limited to ablative imprint studies [52] and attempts to recover properties of the X-ray beam based on the diffraction from aerosolized particles [53] (see also **Paper I**). In **Paper V**, we describe an experiment which uses mixed-state ptychgraphy [41] to image the wavefront of individual XFEL pulses and allows a comprehensive characterization of the X-ray beam in the focus.

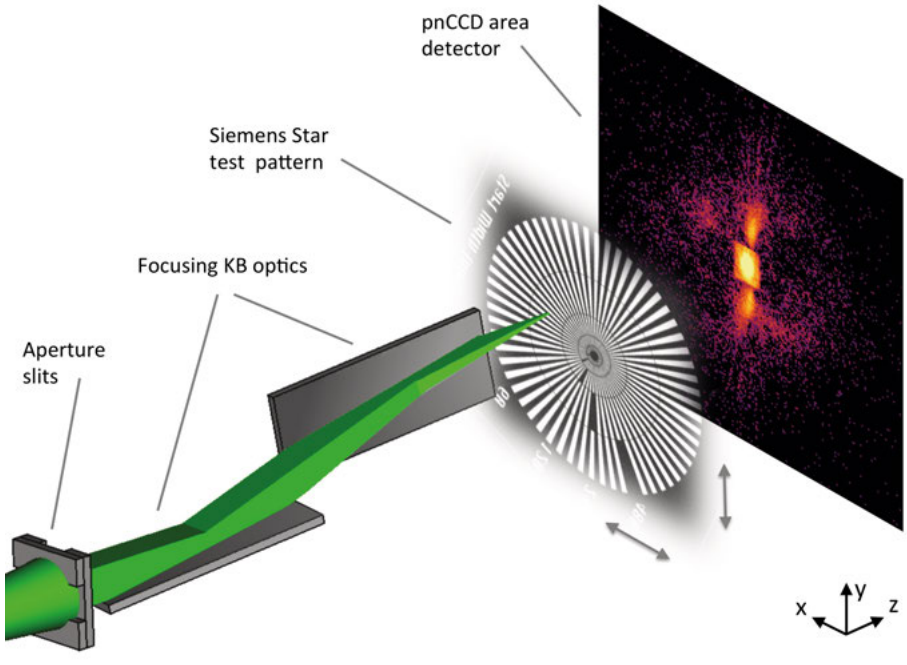

Figure 5.1. Experimental setup. An attenuated X-ray beam was focused onto a Siemens Star test pattern using KB optics. The test pattern was scanned in x and y and a total of 30 000 diffraction patterns were recorded, 300 per scan position. Figure adopted from Fig. 1 in **Paper V**.

The experiment was performed at the end station for Atomic, Molecular and Optical science (AMO) [54] at the LCLS. The experimental setup is shown

in *Figure 5.1*. The conditions were chosen such that they mimic typical FXI experiments, except that the beam was attenuated by a 4.5 m long nitrogen gas cell to avoid damaging the sample. The attenuated beam, with a photon energy of 1260 eV, was focused onto a Siemens Star test pattern using a pair of KB mirrors. While scanning the test pattern on a 10×10 grid with randomized offsets, we collected a total of 30 000 diffraction patterns, 300 for each scan position. After dark correction, signal conversion to photons and exclusion of erroneous events, we averaged 28 647 individual diffraction patterns $\hat{I}_{n\mathbf{q}}$ (with label n) into 98 diffraction patterns $\hat{I}_{j\mathbf{q}}$ (with label j) grouping them by scan position. We denote $\hat{I}_{n\mathbf{q}}$ the "full" dataset and $\hat{I}_{j\mathbf{q}}$ the "average" dataset.

5.1 Mixed-state reconstruction

Using the "average" dataset and the mixed-state model for ptychography, as defined in equation (3.50), we reconstructed the 20 most dominant probe modes alongside the object. We used 5000 iterations of the ML algorithm followed by 5000 iterations of the ML algorithm as implemented in the *ptypy* software package [55].

We projected the 20 probe modes $P_{\mathbf{x}}^{(m)}$ onto an orthogonal basis and interpreted them as orthogonal beam compontents $R_{\mathbf{x}}^{(\mu)}$. We scaled those components such that the object $O_{\mathbf{x}}$ inside (3.50), as defined by (3.48), matches the expected refractive index of the test pattern. This allowed us to perform a quantitative study of the wavefront based on the orthogonal beam components.

5.2 Pulse-to-pulse reconstruction

Assuming that most of the beam components obtained by the mixed-state reconstruction are describing coherent illumination of the object, we can model the diffracted intensities for an individual pulse with label n

$$I_{n\mathbf{q}} = \frac{k^2}{r^2} \left| \mathrm{DFT} \left[\sum_{\mu} c_{n\mu} R^{(\mu)}_{\mathbf{x}-\mathbf{x}_n} O_{\mathbf{x}} \right] \right|^2 \tag{5.1}$$

as a coherent sum, where $c_{n\mu}$ are unknown complex-valued coefficients. Using a conjugate gradient least squares method, we found these coefficients by solving

$$\arg\min_{c_{n\mu}} \left| I_{n\mathbf{q}} - \hat{I}_{n\mathbf{q}} \right|^2 \tag{5.2}$$

for each individual pulse. We performed this analysis for all 28 647 diffraction

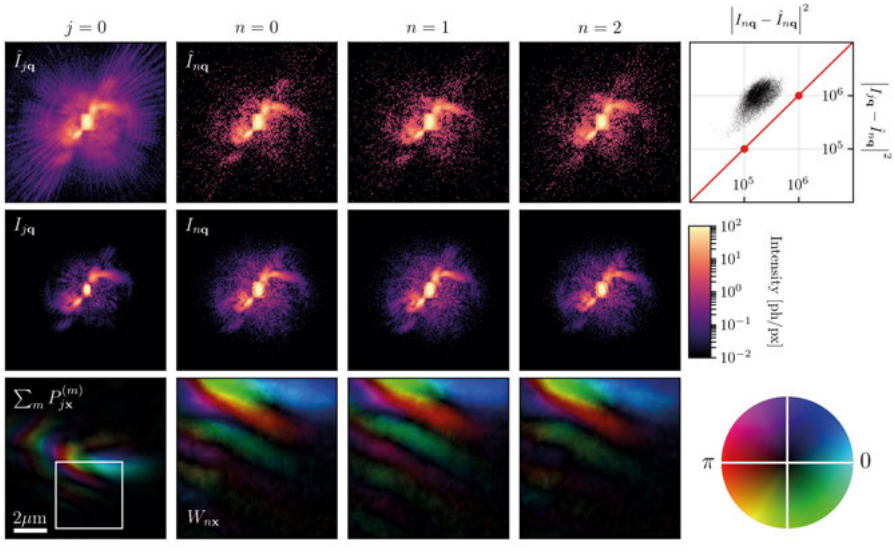

Figure 5.2. Overview of pulse-to-pulse analysis showing example diffraction patterns of the "average" and "full" dataset (top row), modeled diffraction intensities (middle row) and reconstructed wavefronts with shot-to-shot variations as visible inside for a zoomed in region (bottom row). The shot-to-shot residuals between modeled and measured diffraction intensities are reduced by an order of magnitude after performing the pulse-to-pulse analysis (top right panel). Figure adopted from Fig. 4 in **Paper V**.

patterns $\hat{I}_{n\mathbf{q}}$, the results are shown in *Figure 5.2* with examples from the "average" and "full" datasets, $\hat{I}_{j\mathbf{q}}$ and $\hat{I}_{n\mathbf{q}}$ respectively. Modeling the individual diffraction patterns with $I_{n\mathbf{q}}$ and fitting the unknown coefficients reduces the residuals $|I_{n\mathbf{q}} - \hat{I}_{n\mathbf{q}}|^2$ by an order of magnitude compared to $|I_{j\mathbf{q}} - \hat{I}_{n\mathbf{q}}|^2$, a description of individial shots using the averaged model. This approach allowed us to estimate the complex wavefront of individual pulses as

$$W_{n\mathbf{x}} = \sum_{\mu=0}^{M-1} c_{n\mu} R_{\mathbf{x}}^{(\mu)} \tag{5.3}$$

which seems to fluctuate between the different examples given in *Figure 5.2*. The features inside a zoomed in region (indicated by the white box) are changing and the individual wavefronts appear to be slightly deformed.

5.3 Properties of a focused XFEL beam and its relation to FXI experiments

We numerically propagated the main orthogonal component, representing the major features of the reconstructed X-ray beam, 25 mm upstream and downstream along the beam axis. Vertical and horizontal sections through the propagated focus are shown in panel (a) and (b) of *Figure 5.3*. In the context of FXI, the relevant part of the X-ray focus is 10 µm upstream and downstream of the interaction region, since the injected stream of particles is typically 20 µm wide. In this range, the wavefront seems to be about the same in all propagated planes, as shown in panels (h) and (i) in *Figure 5.3*. This means that different

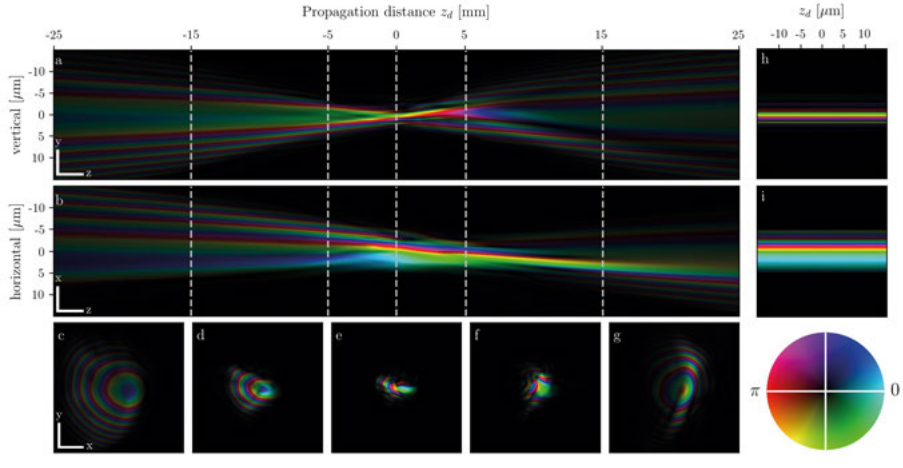

Figure 5.3. Main orthogonal probe mode of the mixed-state reconstruction $R_{\mathbf{x}}^{(0)}$, numerically propagated by 25 mm upstream and downstream of the focus with vertical (a) and horizontal (b) sections. A closer zoom into the focus is shown in (h) and (i). The dashed gray lines indicate cuts perpendicular to the beam axis which are shown in (c-g). Amplitude and phase are mapped to brightness and hue, respectively. Figure adapted from Fig. 5 in **Paper V**.

particles should see the same profile of the wavefront, no matter where they are along the beam axis. More relevant for the interaction with aerosolized particles, is the intensity profile in the focus. The most intense region has an estimated full-width at half the maximum (FWHM) of 2 µm in the horizontal and 1 µm in the vertical direction.

By accounting for the transmission through the nitrogen gas attenuator, we can estimate the fluence distribution withing the intensity profile $|W_{n\mathbf{x}}|^2$ of a given pulse. For particles smaller than 50 nm (the size of a pixel in the reconstructed wavefront), this distributions reflects the sampling of different fluences based on where the interaction of the particle with the X-ray beam occurs. The obtained distribution is shown in panel (a) of *Figure 5.4* in terms of mean and standard deviation of 200 logarithmic fluence bins across 5000 different

pulses. In panel (b) of *Figure 5.4*, we make a comparison to fluence estimates

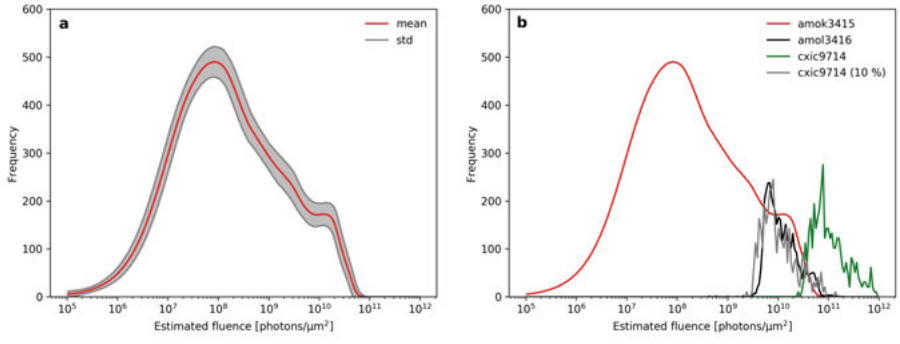

Figure 5.4. Estimated mean (red) and spread (gray) of 5000 fluence distributions (a) compared to similar distributions obtained in two FXI experiments by means of a sphere fitting analysis (b). Since the FXI distributions are truncated due to hit-finding limits, they have been scaled to match the counts inside the accessible fluence region. The black curve describes an AMO experiment with a 35 nm virus (same data as shown in panel (c) of *Figure 4.5*). The green curve is based on the data described in **Paper I** selected for particle sizes smaller than 50 nm, the gray curve is the same as the green curve but with 10 % of the fluence. All FXI based distributions seem to qualitatively match the fluences obtained by the wavefront analysis (red curve). Figure adapted from Fig. 7 in **Paper V**.

based on a sphere fitting analysis (see sections 3.5.1 and 6.4) for different FXI experiments. For the FXI experiment performed at the AMO end station on a 35 nm virus (amol3416) under similar conditions, the estimated peak fluence of about 10^{11} photons/μm^2 and the qualitative trend towards lower fluences seem to match. For the FXI experiment performed at the CXI end station on a 40 nm virus (cxic9714, **Paper I**), the peak fluence was an order of magnitude higher (tighter focus, different pulse energy and beamline transmission) and is therefore difficult to compare, but the qualitative trend towards smaller fluence appears to be similar. This comparison strenghtens the validity of fluence estimates for FXI experiments based on a sphere fitting analysis. Furthermore, the obtained fluence distribution shown in *Figure 5.4*a can be used in FXI simulations to prepare for future experiments or improve the behaviour of reconstruction algorithms under experimental conditions.

6. Algorithms and Software

The development of algorithms and the implementation of software is a core part of realizing experiments. Over the years, we have worked on many computational tools, which are elemental to the success of FXI and ptychography experiments. We have established robust algorithms for per-pixel gain correction, hit-finding, center finding and classification of FXI diffraction data. Furthermore, we have created a number of software packages: *Hummingbird* - a tool for real time analysis of FXI data, *Owl* - a visualization tool for diffraction data, *SHARP* - a fast GPU solver for ptychography and *Nanosurveyor* - a framework for real-time ptychography. Most of the code is written in Python which makes it easy to use, extend and share with others.

6.1 Per-pixel gain correction

To establish a per-pixel gain map for the CSPAD in **Paper I**, we developed an algorithm with involves the following steps for each detector pixel

1. Produce a histogram of corrected ADU values
2. Estimate initial parameters of the 0-photon peak, namely its amplitude A^0, expectation value μ^0 and noise σ^0
3. Establish a mask which blocks signal above the 0-photon peak (based on σ^0)
4. Fit a Gaussian to the 0-photon peak using initial parameters and mask from 2. and 3.
5. Find the bottom of the valley between 0- and 1-photon peak
6. Estimate the location μ^1 and amplitude A^1 of the 1-photon peak based on the maximal value in the histogram beyond the valley
7. Establish a mask which blocks signal below the 1-photon peak
8. Fit a Gaussian to the 1-photon peak using initial parameters and mask from 5. and 6.
9. Return the fitted values A^0, μ^0, σ^0, A^1, μ^1 and σ^1

which can be used for gain correction, photon counting and SNR estimation, as described in 4.2. This algorithm is implemented as part of the data processing description[1] for **Paper I**, in particular the Python module `src/fit.py`.

[1] https://github.com/FXIhub/cxic9714-analysis

6.2 Hit-finding

To distinguish "hits" from "misses" in **Paper I**, we have used the well established lit pixel counter, which includes the following steps

1. Define the expected ADU value of a single photon event based on an accumulated histogram or tabulated gain value
2. For each event, count the number of pixels which are "lit", meaning that they have a value above the expected single photon signal
3. Monitor the lit pixel hit score for a few events and define a threshold based which separates strong outliers from the baseline
4. Label each event as either a "hit" or a "miss" based on the defined threshold

This is also described in **Paper II** as an example for real-time analysis with *Hummingbird*.

6.3 Center finding

If the local wavefront, as seen by different particles during an FXI experiment, is changing, the center position of each diffraction pattern fluctuates. Making use of the Friedel symmetry[2] which applies for most FXI data, we developed the following procedure to find the center of diffraction for each event

1. For a given range of expected center positions, extend a given bad pixel mask to include the centro-symmetric partner of each bad pixel
2. Calculate pair-wise cross correlations for each non-masked pixel and its centro-symmetric partner and
3. Find the center position which maximizes this correlation

and implemented this algorithm inside the Python module `src/_spimage_find_center.py` of the backend[3] to the software package *Hawk* [56]. Finding the center for individual events is an important pre-step when modeling diffraction data and has further been used in **Paper I** to build a map of local phase variations in the X-ray beam.

6.4 Classification based on sphere diffraction

For the classification of diffraction patterns in **Paper I** based on determining the corresponding particle size and incident intensity, we have implemented the following model-based sphere fitting method

[2] $\mathrm{DFT}[f](q_x, q_y, q_z) = \mathrm{DFT}[f](-q_x, -q_y, -q_z)$ for $f \in \mathbb{R}$
[3] https://github.com/FXIhub/libspimage

1. Find the diffraction center using the algorithm described in 6.3
2. Model the diffraction using equation (3.39) with given experimental parameters and assume particles to be homogeneous spheres with constant refractive index of a given material
3. Define a circular mask given a radius that indicates the maximal diffraction angle to be used for the fitting steps below
4. Find a rough estimate for the particle size by maximizing the Pearson correlation between modeled and measured intensities, since this correlation is insensitive to intensity fluctuations
5. Find a rough estimate for the incident intensity by minimizing the sum of squared differences between modeled and measured intensities at a fixed particle size
6. Refine the estimates for center position, particle size and incident intensity by minimizing the sum of squared differences between modeled and measured intensities

into the Python module `src/_spimage_sphere_model.py` of the backend[4] to the software package *Hawk* [56].

6.5 *Hummingbird*: Flash X-ray imaging in real time

With more than a 100 events per second at the LCLS, the data rate at current FXI experiments is fairly high, and it is expected to increase even more with the new light sources EuXFEL and LCLS-II coming online. Since access to those facilities is very limited, users are under pressure to collect high quality data within a short amount of time. Under those circumstances, it is crucial to have access to a software tool such as *Hummingbird*, which is capable to process data in real time and gives immediate feedback on data quality and current experimental conditions (**Paper II**).

Hummingbird is a Python-based software tool designed with a server-client architecture, as shown in *Figure 6.1*. The backend (server) is meant to operate on dedicated workers with fast access to the data, while the frontend (client) can be any desktop computer or laptop with a graphical display. The data flow from a file or a shared memory stream to a specific plot in the graphical user interface (GUI) can be defined by the user through a Python-based configuration file (conf.py). Based on the user facility, *Hummingbird* translates native events into a common data structure (the `evt` variable), which makes it possible to use the same user-specific analysis pipeline at different experiments and facilities. To sustain a high throughput of data, the backend is scalable to as many workers as necessary. Since most FXI analysis is event-based, these workers are mostly independent, but communication between workers is possible using a message passing interface (*MPI*) which can be very useful in

[4] https://github.com/FXIhub/libspimage

certain cases (e.g. calculating hit ratios). At the end of a configuration file, the user specifies how the analysis results should be displayed and *Hummingbird* sends this information across the network to the GUI.

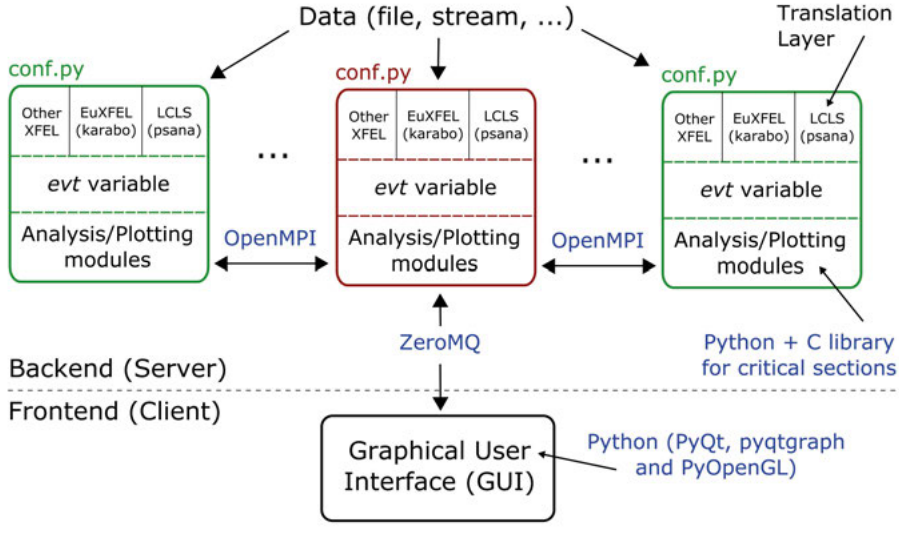

Figure 6.1. The client–server architecture of *Hummingbird*. Depending on the configuration (conf.py), the backend workers (master in red, slaves in green) read in data, translate native events based on the facility, perform user-specific data analysis and send plots off to the frontend. Figure adopted from Fig. 1 in **Paper II**.

Hummingbird has been created for the particular needs of FXI. However, since its structure is quite generic, it can be used for almost any experiment performed at modern light sources. It has already proven to be of great value in numerous experiments at the LCLS and FLASH. The software is open source[5] and documentation[6] with examples is available online.

6.6 *Owl*: Visualization of X-ray diffraction data

Before, during or after processing FXI data, one typically wants to browse through a large amount diffraction data. For that purpose, we developed the Python-based visualization tool *owl*. It allows to open very large files and still be able to skim through many images in a seamless fashion, since it keeps a buffer of nearby images and uses *OpenGL* for rendering. *Owl* is mainly designed for displaying images, but it can also visualize one-dimensional data such as spectral traces. It also supports a display of experimental parameters (e.g. motor positions or beam intensities). With the tagging tool, it is easy

[5] `https://github.com/FXIhub/hummingbird`
[6] `http://www.lmb.icm.uu.se/hummingbird/docs`

to manually identify and classify images. A built-in sizing tool can be used to display the sphere diffraction model next to the data and quickly assess particle size and intensity based on a diffraction pattern (see *Figure 6.2*).

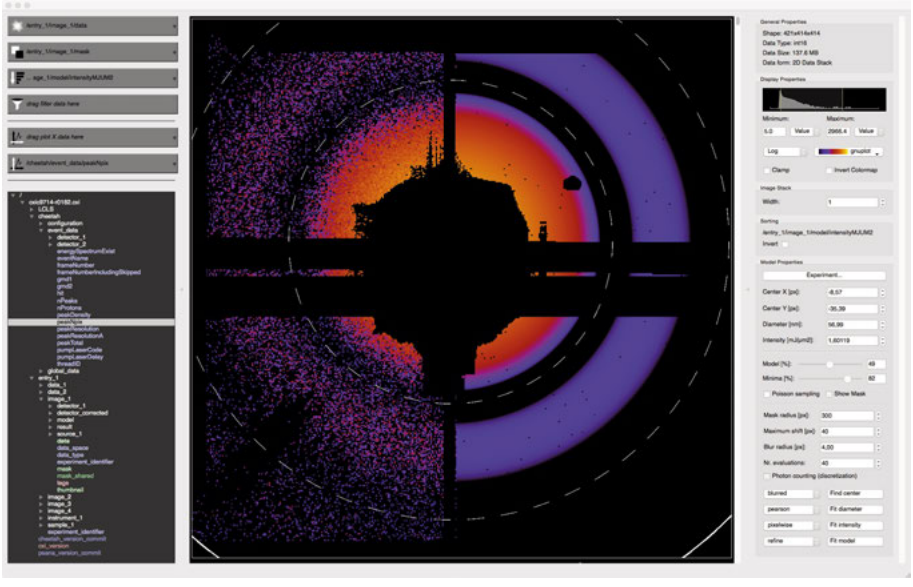

Figure 6.2. GUI of the visualization tool *owl*. The file tree is displayed on the left. The main stage in the middle shows a continuous stack of diffraction images. The panel on the right can be used to adjust visualization and other parameters. *owl* also includes analysis tools such as displaying a diffraction model next to the data.

Owl is open source[7] and can read from data stored in the *Hierarchical Data File Format Version 5* (HDF5) which adhere to the CXIDB file structure [57].

6.7 *SHARP*: A fast GPU solver for ptychography

X-ray ptychography has become a popular imaging technique, mostly since it provides access to high resolution. However, it comes with a fairly advanced and time-consuming reconstruction scheme. In **Paper III**, we describe the software package *SHARP* which uses graphics processing units (GPUs) to allow fast ptychographic reconstructions. The computational backend of *SHARP* is written in C/CUDA and implemented for NVIDIA GPU architectures. *SHARP* also provides a Python interface which makes it easy for the user to customize and configure the reconstruction to the given needs of a particular experiment.

For a dataset with 10 000 frames of size 128×128, it takes less than 2 s for *SHARP* to obtain a full ptychographic reconstruction using a GPU cluster with

[7] `https://github.com/FXIhub/owl`

16 Nvidia GTX Titan cards (see Fig. 3 in **Paper III** for more details). Having such a fast GPU-based solver implemented opens up new possibilities to do high-resolution imaging with ptychography in real time.

6.8 *Nanosurveyor*: Ptychography in real time

The idea of streaming ptychography which enables imaging in real time has been already outlined in **Paper III**. With the implementation of *Nanosurveyor* (**Paper IV**), we provided a framework which facilitated that promise. The structure of this real-time processing framework is sketched in *Figure 6.3* using the ptychographic streaming pipeline implemented at the Advanced Light Source (ALS) as an example. Triggered by the start of a new scan, raw data is

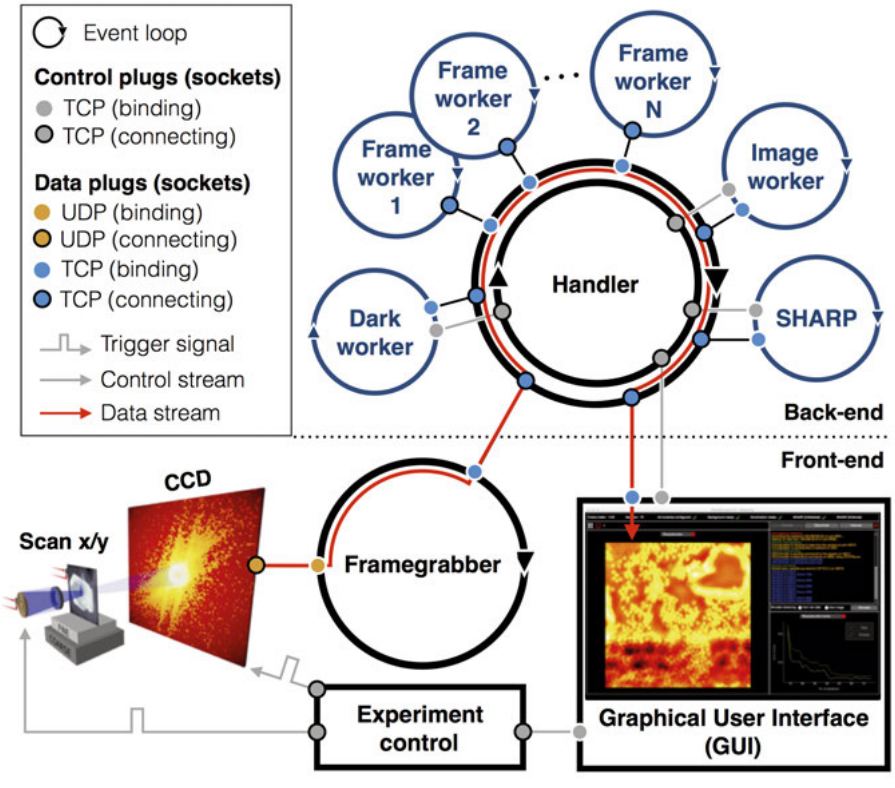

Figure 6.3. Ptychographic streaming pipeline implemented at the Advanced Light Source (ALS). Once a new scan has been triggered by the experimental control, a framegrabber continuously receives raw data packets from the camera, assembles them to a frame and sends raw frames to the back-end. Incoming frames are processed by different (and independent) workers of the back-end and reduced data are sent back to the front-end and visualized in a GUI. Figure adopted from Fig. 2 in **Paper IV**.

sent via a framegrabber to a powerful back-end (e.g. high-performance compute cluster) which immediately starts the ptychographic reconstruction using *SHARP*, while the scan is still running and new raw data is coming in. Once reconstructed images are available, they are sent directly to the GUI along with other data which includes information on the current scan. In the back-end, an event handler is coordinating the data and communication workflow in which different types of control and data plugs (sockets) are used. While most of the components communicate via the transmission control protocol (TCP), the raw data packets from the camera are sent via the user datagram protocol (UDP). All the individual components of this pipeline are running asynchronous event loops, which allows the information obtained by means of diffraction patterns to be transmitted to the user in forms of reconstructed images as fast as possible, and already during data collection. This makes ptychographic data collection more efficient and leaves more time for the user to find the interesting features in his sample, which leads to higher throughput and probably more scientific outcome.

Summary and Outlook

Achievements

In this work, we explored the possibilities of FXI in particular when approaching high-resolution structural imaging of small biological objects (below 50 nm) with high photon energies (above 5 keV). We described an experiment in which we used X-ray ptychography to image the wavefront of a focused XFEL under relevant FXI conditions. Furthermore, we presented a large number of algorithms, computational workflows and software tools for FXI and X-ray ptychography experiments.

We have seen in **Paper I** that sample delivery is a key challenge for FXI. The most common delivery method is aerosol injection using a GDVN to form micron-sized droplets. For biological targets much smaller than a droplet, a large fraction of the droplet is unoccupied. This increases the likelihood that non-volatile contaminants are accumulated in the drying process leading to "caking" effects. As a result, the diffraction data is corrupted which makes successful 3D imaging extremely difficult. A strategy for future experiments could be to work on alternative droplet formation methods which produce droplets similar to the size of the particle, and also to take additional care in terms of maintaining a clean buffer when preparing the sample. Recent experiments have shown that droplet formation using electrospray ionization (ESI) substantially reduces the "caking" problem since it can produce droplets as small as 100 nm.

Another important aspect of FXI is the ratio of signal to background. We have shown in **Paper I** that it is possible to count individual photons on a CSPAD detector. This is a great achievement, but it is equally important for the success of FXI that a large fraction of those photon events results from sample diffraction. Otherwise, erroneous background photons will compromise the data and might lead to artifacts in the image reconstruction. In future experiments, the level of photon background should therefore be further reduced. A complementary approach is the integration of background modeling into current reconstruction schemes. However, this requires a good and realistic description of the background, which is hard to model.

Despite the challenges stated above, we were able to reconstruct a 2D projection image of a 40 nm particle (matching the size of the virus being imaged) based on a single noisy diffraction pattern, at a full-period resolution of 13.5 nm.

In **Paper V**, we have used X-ray ptychography as a method to fully characterize the focused beam of an XFEL under conditions which are relevant for

FXI experiments. We were able to map out the wavefront of individual pulses and could show that it changes from shot to shot. Such a method could form the basis for a future routine wavefront characterization step in FXI experiments.

The continuous development of algorithms and software is an integral part of CDI. We have contributed to this development with an outline of a complete computational pipeline for FXI, from raw diffraction data to reconstructed images (**Paper I**), and provided fast software tools for FXI (**Paper II**) and ptychography (**Paper III** and **IV**) which enable real-time analysis.

Open science

Modern science should be transparent and must be reproducible. Following this principle, the collective brain of an entire community can be utilized to address and fix problems, which allows research fields to reach their potential in a fast and robust manner. It is particularly important that experiments, including computational reconstruction schemes, are reproducible. A crucial aspect in that regard is the validation of results, especially in the particular case of method development for imaging. With a constantly increasing number of new algorithms for image reconstruction, there is also a need for comprehensive and standardized validation techniques. In many cases, the data analysis and also its validation are targeted towards a particular experimental setup or sample. This makes it challenging to define a common validation method or criterion which applies for all (or at least most) data sets. However, with access to data and code repositories, the development of unified and widely applicable validation tools should be possible. Therefore, in the spirit of open science, it should be a given that any scientific publication is accompanied with a deposition of all relevant data as well as the code, which translated the data into the final scientific outcome.

For **Paper I**, we deposited all raw and processed data in the CXIDB [57] and created a public code repository[8] which includes a comprehensive description of the data analysis pipeline and a collection of scripts and Python modules used to create the published figures. For **Paper II** and **IV**, the source code is available through public repositories. For **Paper III**, the source code is currently available on request only, due to legal reasons. For **Paper V**, data and code will be deposited together with its publication. Furthermore, all of the FXI related algorithms and software tools described in this thesis are available through public code repositories[9].

[8] `https://github.com/FXIhub/cxic9714-analysis`
[9] `https://github.com/FXIhub`

Outlook

CDI is still a young and fast growing field. While X-ray ptychography has already advanced into a frequently used tool at synchrotrons and will probably develop further into a more and more automated imaging method, FXI has not yet reached its full potential. But several new experimental end stations at the EuXFEL and LCLS-II are targeting FXI, like the SPB beamline at the EuXFEL, which started operation in September 2017. They offer new opportunities with a tremendous increase in data rates up to kHz in the near future and maybe even MHz in the more distant future. FXI still has to improve on many levels and there are remaining challenges to tackle. But with the continuous development of sample delivery methods (e.g. smaller droplets), an increase in beam intensity and a reduction of background signal, the promised high data rates might lead the way towards three-dimensional studies of homogeneous biological structures at atomic resolution. In combination with further algorithmic development, we might even be able to study heterogeneous structures by mapping a large number of conformations. With cryo-electron microscopy (cryoEM) there is another imaging technique targeting similar applications, which has shown tremendous progress over the past few years [58]. However, cryoEM requires cryogenic fixation of the sample, demands relatively long exposure times and only permits a limited amount of exposures. For time-resolved studies, the resolution might be limited to the time scales of the freezing process (milliseconds). FXI has the potential to surpass those limitations. Furthermore, the success of cryoEM in resolving different conformations [59] suggests that a similar approach should also work for FXI. The short duration of XFEL pulses (femtoseconds) could allow for studies of fast structural dynamics via pump-probe FXI experiments. This approach, together with the capability of mixing components prior to injection [60], might permit studies of e.g. virus assembly, protein folding or the conformational landscape of macromolecules.

Author's contribution

Paper I
I performed the experiment together with others at the LCLS and contributed to onsite data analysis. I pre-processed the raw data, made the classification based on particle size and intensity, and extracted all results reported in the paper. I made simulations to validate the classification approach. I wrote most of the manuscript in collaboration with others and deposited an extensive description of the analysis in a public repository.

Paper II
I wrote most code for the initial online monitoring project, which later became the software *Hummingbird*, and also had major contributions to the software in its current form. I did the performance benchmarking, created all figures and wrote the manuscript.

Paper III
I contributed to the *SHARP* code with the implementation of a Python interface, helped develop the idea of real-time ptychography and contributed to the writing of the manuscript.

Paper IV
I designed and implemented the code for the real-time streaming framework *Nanosurveyor* and wrote major parts of the manuscript.

Paper V
I helped in writing the proposal and planning for the wavefront sensing experiment. I performed the experiment together with others at the LCLS and contributed to onsite data analysis. I pre-processed the raw data and worked on the ptychographic reconstruction. I performed the pulse-to-pulse analysis and wrote major parts of the manuscript.

Sammanfattning på svenska

Vill man förstå livscykeln för en cell är det rimligt att först ta en bild av hela systemet och sedan titta närmare på de byggstenar som det är uppbyggt av. Det är därför ingen överraskning att avbildning har utvecklats till ett stort och viktigt forskningsområde.

Det första optiska mikroskopet kunde redan i slutet av 1600-talet avslöja information om världen som normalt är bortom det våra ögon kan se och uppfattas. Med ett sådant mikroskop kunde man plötsligt undersöka material på ett helt nytt sätt. Så länge materialet var genomskinligt kunde man förstora strukturer och titta närmare än vad som var möjligt med bara ögonen. I och med Wilhelm Röntgens upptäckt av röntgenstrålning 1895 öppnades en ännu större värld. Nu blev det även möjligt att undersöka ogenomskinliga material.

Både traditionell ljusmikroskopi och röntgenmikroskopi använder ett linssystem för att först fokusera ljus- eller röntgenstrålarna på provet och sedan använda en så kallad objektivlins för att skapa en bild. I princip är ett mikroskops upplösning begränsad av den våglängd som används, men för röntgenstrålning ligger begränsningen snarare i hur noggranna linser vi kan tillverka. Det är både svårt och dyrt att tillverka linser som kan fokusera röntgenstrålning tillräckligt bra. Det finns dock en annan metod för avbildning som inte behöver någon lins och därför ofta kallas för "linsfri" avbildning. Linsen är här utbytt mot en dator som återuppbyggar bilden. Den här metoden är känd som "Coherent diffractive imaging" eller CDI och förlitar sig istället på en tillräckligt tät sampling (urval av punkter) av diffraktionsdatan. Vid en CDI mätning så ser vi normalt sett bara amplituden hos den diffrakterade vågen medan fasinformationen går förlorad. Med hjälp av smarta upprepningsmetoder kan fasen återskapas förutsatt att amplituddatan är tillräckligt tätt samplad.

Upplösningen som kan nås med CDI metoder ligger nära våglängden av röntgenstrålning vilket innebär att man få tillgång till "nano" världern. CDI är därför väletablerad och används inom många vetenskapsområden. Metoden har blivit mer populär i och med utvecklingen av röntgenlasrar som producerar röntgenpulser med enastående egenskaper: de är extremt korta och har även extremt starka intensitet. En metod som utnyttjar dessa egenskaper är "Flash X-ray imaging" (FXI) som används för att avbilda biomolekyler och virus. I ett FXI experiment sprutas en aerosol av många små partiklar in i vakuum som träffas av röntgenpulsen, vilken är tillräckligt kort för att undvika att partikeln skadas innan bilden är tagen. Denna princip, som kallas för "diffraktion före destruktion", möjliggör nya sätt att studera biologiska strukturer. Redan år 2006 bevisades det att metoden fungerar genom ett lyckat experiment vid röntgenlasern FLASH i Hamburg och sedan dess har den metoden även använts vid

röntgenlasern LCLS vid Stanford Universitet. Än så länge har dock inte FXI uppnått sin fulla potential.

En närliggande metod heter röntgenptychografi. Stora prover skannas systematisk då med en fokuserad röntgenstrålning och på ett sådant sätt att närliggande scannpunkter överlappar med varandra. Denna metod ger högupplöst och kvantitativ avbildning och rekonstruerar även ljusstrålens form. Tack vare detta går det även att använda ptychografi som ett verktyg för att beskriva ljusstrålar.

I denna avhandling sammanfattar jag det nuvarande läget inom FXI-metoden och utvecklar nya strategier för att förbättra tekniken, för att uppnå avbildning av biomolekyler och virus med atomär upplösning. Jag visar de viktigaste delarna av ett sådant experiment nämligen partikelinjektion, röntgendetektor och ljuspulsens egenskaper. Dessutom beskriver jag ett experiment där ptychografi används för att studera ljuspulserna i ett FXI experiment.

Slutligen beskriver jag algoritmer och programvara som jag har utvecklat för FXI och ptychografi och som kan användas för att analysera data både i realtid och efter att experimentet är utfört. Realtidsanalysen är mycket viktig för det ger återkoppling redan medan datan samlas in.

Jag hoppas att detta kan vara ett värdefullt bidrag till utvecklingen av CDI och speciellt FXI och ptychografi. Jag är övertygad om att dessa metoder kommer att vara viktiga verktyg för högupplöst avbildning med tillämpningar inom många vetenskapsområde såsom strukturbiologi och att vi kommer så småningom förstå cellens liv.

Acknowledgements

This work is based on the collaborative efforts of an interdisciplinary and international team. It is an honor and a privilege that I was given the chance to work with so many great people, and I am grateful to all the help and support I got from my colleagues and friends all around the globe.

In particular, I would like to thank **Filipe Maia** who has been such an awesome and patient supervisor. I am deeply thankful for the many things that I learned from you, and for the opportunity to participate in so many exciting projects. I would also like to give special thanks to **Janos Hajdu** for his support and encouragement throughout the years, and for sharing his experience and knowledge with me. **Tomas Ekeberg**, you have helped me through my first days in Uppsala, you were always available if I needed advice (both work and non-work related), coffee and lunch break discussion with you were always fun, and you became a regular climbing companion: Thank you! **Max Hantke**, I would like to thank you for being my friend, great colleague, sailing mate and travel companion. Also, I thank **Daniel Larsson** and **Kenta Okamoto** for their amazing support, especially when I was writing my first paper. Thanks to **Nicuşor Tîmneanu** for always having an open door and for so many interesting discussions. Thanks to all the computer nerds, **Alberto Pietrini**, **Gijs van der Schot**, **Ida Lundholm-Benzi**, **Carl Nettelblad**, **Jessica Nettelblad**, **Jing Liu**, **Christofer Östlin**, **Olof Jönsson** and **Jonas Sellberg** for productive discussions and the sleepless and fun times during experiments. Thanks to **Johan Bielecki**, **Daniel Westphal**, **Federico Benzi**, **Kerstin Mühlig**, **Marvin Seibert**, **Alessandro Zani**, **Olena Kulyk** and **Jakob Andreasson** for the many hours spent on developing, building, setting up and operating the instrumental equipment for the XFEL experiments. Also many thanks to my colleagues **Hemanth Kumar**, **Laura Gunn**, **Anna Larsson**, **Anna Munke**, **Gunilla Carlsson**, **Dirk Hasse**, **Margareta Ingelman**, **Karin Valegård**, **Martin Svenda**, **Peter Kasson** and **Inger Andersson** for providing us with biological samples and for teaching me about biology. I would like to thank my first student, **Louis Doctor**, it was a pleasure working with you! Also, I am very grateful to **Duane Loh** for the many fruitful discussion over Skype and during experiments, for sharing so many ideas and for hosting me in Singapore (while these lines are being written). Thanks to **Simone Sala** and **Pierre Thibault** for your support and work on the wavefront project. I am indebted to **Stefano Marchesini** for hosting me during my stay in Berkeley and for many inspiring conversations. I would also like to thank the rest of the CAMERA team, **Hari Krishnan**,

Talita Perciano, **David Shapiro** and **James Sethian** for investing time and efforts into software development for imaging. Furthermore, I thank **Atsushi Nakagawa** and **Akifumi Higashiura** for being my hosts in Osaka, it was such a great and insightful experience. I also thank the **ICM administration** for their help.

For their indirect but essential contribution to this work, I would like to thank my parents **Gabriele** and **Hans**, my brother **Nikolaus**, and my friends **Sarina**, **Danna**, **Mayank**, **Preeti**, **Daniel**, **Mattias**, **Arvid**, **Rikard**, **Agnė**, **Tobias**, **Linnea**, **Disa**, **Yolanda**, **Christoffer**, **Claudia**, **Johan**, **Sebastian**, **Giwrgos**, **Fredrik**, **Prune**, **Pernilla**, **Isaak**, **Katrin**, **Hugo**, **Caro**, **Venkat**, **Stephan**, **Lorenz**, **Simone**, **Flo** and **Paola**, **Anna** and **Andreas**, **Carina**, **Barbara**, **Isabell**, **Phillipp**, **Thomas**, **Raphael**, **Matthias**, **Leo**, **Tom**, **Hannah**, and the members of **Salongsorkestern**.

References

[1] C. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10--21, 1949.

[2] D. Sayre. Some implications of a theorem due to Shannon. *Acta Crystallographica*, 5(6):843--843, 1952.

[3] R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik*, 35:237--246, 1972.

[4] J. R. Fienup. Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters*, 3(1):27, 1978.

[5] S. Marchesini. Invited Article: A unified evaluation of iterative projection algorithms for phase retrieval. *Review of Scientific Instruments*, 78(1):011301, 2007.

[6] J. Miao, P. Charalambous, J. Kirz, and D. Sayre. No Title. *Nature*, 400(6742):342--344, 1999.

[7] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature*, 406(6797):752--757, 2000.

[8] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro. High-resolution ab initio three-dimensional x-ray diffraction microscopy. *Journal of the Optical Society of America A*, 23(5):1179, 2006.

[9] M. Seibert, M, S. Boutet, M. Svenda, T. Ekeberg, F. R. N. C. Maia, M. J. Bogan, N. Tîmneanu, A. Barty, S. Hau-Riege, C. Caleman, M. Frank, H. Benner, J. Y. Lee, S. Marchesini, J. W. Shaevitz, D. A. Fletcher, S. Bajt, I. Andersson, H. N. Chapman, and J. Hajdu. Femtosecond diffractive imaging of biological cells. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 43(19):194015, 2010.

[10] M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, O. Jönsson, D. Odić, B. Iwan, A. Rocker, D. Westphal, M. Hantke, D. P. DePonte, A. Barty, J. Schulz, L. Gumprecht, N. Coppola, A. Aquila, M. Liang, T. A. White, A. Martin, C. Caleman, S. Stern, C. Abergel, V. Seltzer, J.-M. Claverie, C. Bostedt, J. D. Bozek, S. Boutet, a. A. Miahnahri, M. Messerschmidt, J. Krzywinski, G. Williams, K. O. Hodgson, M. J. Bogan, C. Y. Hampton, R. G. Sierra, D. Starodub, I. Andersson, S. Bajt, M. Barthelmess, J. C. H. Spence, P. Fromme, U. Weierstall, R. Kirian, M. Hunter, R. B. Doak, S. Marchesini, S. P. Hau-Riege, M. Frank, R. L. Shoeman, L. Lomb, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, C. Schmidt, L. Foucar, N. Kimmel, P. Holl, B. Rudek, B. Erk, A. Hömke, C. Reich, D. Pietschner, G. Weidenspointner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, I. Schlichting, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, S. Schorb,

D. Rupp, M. Adolph, T. Gorkhover, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, H. N. Chapman, and J. Hajdu. Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature*, 470(7332):78--81, 2011.

[11] M. F. Hantke, D. Hasse, F. R. N. C. Maia, T. Ekeberg, K. John, M. Svenda, N. D. Loh, A. V. Martin, N. Timneanu, D. S. D. Larsson, G. van der Schot, G. H. Carlsson, M. Ingelman, J. Andreasson, D. Westphal, M. Liang, F. Stellato, D. P. DePonte, R. Hartmann, N. Kimmel, R. a. Kirian, M. M. Seibert, K. Mühlig, S. Schorb, K. Ferguson, C. Bostedt, S. Carron, J. D. Bozek, D. Rolles, A. Rudenko, S. Epp, H. N. Chapman, A. Barty, J. Hajdu, and I. Andersson. High-throughput imaging of heterogeneous cell organelles with an X-ray laser. *Nature Photonics*, 8(12):943--949, 2014.

[12] G. van der Schot, M. Svenda, F. R. N. C. Maia, M. Hantke, D. P. DePonte, M. M. Seibert, A. Aquila, J. Schulz, R. Kirian, M. Liang, F. Stellato, B. Iwan, J. Andreasson, N. Timneanu, D. Westphal, F. N. Almeida, D. Odic, D. Hasse, G. H. Carlsson, D. S. D. Larsson, A. Barty, A. V. Martin, S. Schorb, C. Bostedt, J. D. Bozek, D. Rolles, A. Rudenko, S. Epp, L. Foucar, B. Rudek, R. Hartmann, N. Kimmel, P. Holl, L. Englert, N.-T. Duane Loh, H. N. Chapman, I. Andersson, J. Hajdu, and T. Ekeberg. Imaging single cells in a beam of live cyanobacteria with an X-ray laser. *Nature Communications*, 6:5704, 2015.

[13] T. Ekeberg, M. Svenda, C. Abergel, F. R. N. C. Maia, V. Seltzer, J.-M. Claverie, M. Hantke, O. Jönsson, C. Nettelblad, G. van der Schot, M. Liang, D. P. DePonte, A. Barty, M. M. Seibert, B. Iwan, I. Andersson, N. D. Loh, A. V. Martin, H. Chapman, C. Bostedt, J. D. Bozek, K. R. Ferguson, J. Krzywinski, S. W. Epp, D. Rolles, A. Rudenko, R. Hartmann, N. Kimmel, and J. Hajdu. Three-Dimensional Reconstruction of the Giant Mimivirus Particle with an X-Ray Free-Electron Laser. *Physical Review Letters*, 114(9):098102, 2015.

[14] A. Munke, J. Andreasson, A. Aquila, S. Awel, K. Ayyer, A. Barty, R. J. Bean, P. Berntsen, J. Bielecki, S. Boutet, M. Bucher, H. N. Chapman, B. J. Daurer, H. DeMirci, V. Elser, P. Fromme, J. Hajdu, M. F. Hantke, A. Higashiura, B. G. Hogue, A. Hosseinizadeh, Y. Kim, R. A. Kirian, H. K. Reddy, T.-Y. Lan, D. S. Larsson, H. Liu, N. D. Loh, F. R. Maia, A. P. Mancuso, K. Mühlig, A. Nakagawa, D. Nam, G. Nelson, C. Nettelblad, K. Okamoto, A. Ourmazd, M. Rose, G. van der Schot, P. Schwander, M. M. Seibert, J. A. Sellberg, R. G. Sierra, C. Song, M. Svenda, N. Timneanu, I. A. Vartanyants, D. Westphal, M. O. Wiedorn, G. J. Williams, P. L. Xavier, C. H. Yoon, and J. Zook. Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source. *Scientific Data*, 3:160064, 2016.

[15] H. K. Reddy, C. H. Yoon, A. Aquila, S. Awel, K. Ayyer, A. Barty, P. Berntsen, J. Bielecki, S. Bobkov, M. Bucher, G. A. Carini, S. Carron, H. Chapman, B. Daurer, H. DeMirci, T. Ekeberg, P. Fromme, J. Hajdu, M. F. Hanke, P. Hart, B. G. Hogue, A. Hosseinizadeh, Y. Kim, R. A. Kirian, R. P. Kurta, D. S. Larsson, N. Duane Loh, F. R. Maia, A. P. Mancuso, K. Mühlig, A. Munke, D. Nam, C. Nettelblad, A. Ourmazd, M. Rose, P. Schwander, M. Seibert, J. A. Sellberg, C. Song, J. C. Spence, M. Svenda, G. Van der Schot, I. A. Vartanyants, G. J. Williams, and P. L. Xavier. Coherent soft X-ray diffraction imaging of coliphage PR772 at the Linac coherent light source. *Scientific Data*, 4:170079, 2017.

[16] A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. DePonte, P. Drell, S. Doniach, K. H. Downing, T. Earnest, H. Elmlund, V. Elser, M. Gühr, J. Hajdu, J. Hastings, S. P. Hau-Riege, Z. Huang, E. E. Lattman, F. R. Maia, S. Marchesini, A. Ourmazd, C. Pellegrini, R. Santra, I. Schlichting, C. Schroer, J. C. Spence, I. A. Vartanyants, S. Wakatsuki, W. I. Weis, and G. J. Williams. The linac coherent light source single particle imaging road map. *Structural Dynamics*, 2(4), 2015.

[17] W. Hoppe. Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten. *Acta Crystallographica Section A*, 25(4):508--514, 1969.

[18] J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson. Hard-X-Ray Lensless Imaging of Extended Objects. *Physical Review Letters*, 98(3):034801, 2007.

[19] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer. High-Resolution Scanning X-ray Diffraction Microscopy. *Science*, 321(5887):379--382, 2008.

[20] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, P. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda. First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photonics*, 4(9):641--647, 2010.

[21] D. Paganin. *Coherent X-Ray Optics*. Oxford University Press, Oxford, 2006.

[22] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3 edition, 2007.

[23] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 19:297--301, 1965.

[24] H. N. Chapman, C. Caleman, and N. Timneanu. Diffraction before destruction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647):20130313--20130313, 2014.

[25] B. Henke, E. Gullikson, and J. Davis. X-Ray Interactions: Photoabsorption, Scattering, Transmission, and Reflection at E = 50-30,000 eV, Z = 1-92. *Atomic Data and Nuclear Data Tables*, 54(2):181--342, 1993.

[26] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758--69, 1982.

[27] D. R. Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1):37--50, 2005.

[28] V. Elser. Phase retrieval by iterated projections. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 20(1):40--55, 2003.

[29] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence. X-ray image reconstruction from a diffraction pattern alone. *Physical Review B*, 68(14):140101, 2003.

[30] D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A. M. Neiman, and D. Sayre. Biological imaging by soft

x-ray diffraction microscopy. *Proceedings of the National Academy of Sciences*, 102(43):15343--15346, 2005.

[31] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M. J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Hömke, C. Reich, D. Pietschner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, M. Messerschmidt, J. D. Bozek, S. P. Hau-Riege, M. Frank, C. Y. Hampton, R. G. Sierra, D. Starodub, G. J. Williams, J. Hajdu, N. Timneanu, M. M. Seibert, J. Andreasson, A. Rocker, O. Jönsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, K. E. Schmidt, X. Wang, I. Grotjohann, J. M. Holton, T. R. M. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, and J. C. H. Spence. Femtosecond X-ray protein nanocrystallography. *Nature*, 470(7332):73--77, 2011.

[32] J. Steinbrener, J. Nelson, X. Huang, S. Marchesini, D. Shapiro, J. J. Turner, and C. Jacobsen. Data preparation and evaluation techniques for x-ray diffraction microscopy. *Optics Express*, 18(18):18598, 2010.

[33] N. D. Loh, M. J. Bogan, V. Elser, A. Barty, S. Boutet, S. Bajt, J. Hajdu, T. Ekeberg, F. R. N. C. Maia, J. Schulz, M. M. Seibert, B. Iwan, N. Timneanu, S. Marchesini, I. Schlichting, R. L. Shoeman, L. Lomb, M. Frank, M. Liang, and H. N. Chapman. Cryptotomography: Reconstructing 3D Fourier Intensities from Randomly Oriented Single-Shot Diffraction Patterns. *Physical Review Letters*, 104(22):225501, 2010.

[34] T. B. Edo, D. J. Batey, A. M. Maiden, C. Rau, U. Wagner, Z. D. Pešić, T. A. Waigh, and J. M. Rodenburg. Sampling in x-ray ptychography. *Physical Review A*, 87(5):053850, 2013.

[35] J. M. Rodenburg and H. M. L. Faulkner. A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20):4795--4797, 2004.

[36] A. M. Maiden and J. M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10):1256--1262, 2009.

[37] P. Thibault and M. Guizar-Sicairos. Maximum-likelihood refinement for coherent diffractive imaging. *New Journal of Physics*, 14(6):063004, 2012.

[38] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer. Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*, 109(4):338--343, 2009.

[39] C. M. Kewish, P. Thibault, M. Dierolf, O. Bunk, A. Menzel, J. Vila-Comamala, K. Jefimovs, and F. Pfeiffer. Ptychographic characterization of the wavefield in the focus of reflective hard X-ray optics. *Ultramicroscopy*, 110(4):325--329, 2010.

[40] A. Schropp, R. Hoppe, V. Meier, J. Patommel, F. Seiboth, H. J. Lee, B. Nagler, E. C. Galtier, B. Arnold, U. Zastrau, J. B. Hastings, D. Nilsson, F. Uhlén,

U. Vogt, H. M. Hertz, and C. G. Schroer. Full spatial characterization of a nanofocused x-ray free-electron laser beam by ptychographic imaging. *Scientific Reports*, 3(1):1633, 2013.

[41] P. Thibault and A. Menzel. Reconstructing state mixtures from diffraction measurements. *Nature*, 494(7435):68--71, 2013.

[42] T. Kimura, Y. Joti, A. Shibuya, C. Song, S. Kim, K. Tono, M. Yabashi, M. Tamakoshi, T. Moriya, T. Oshima, T. Ishikawa, Y. Bessho, and Y. Nishino. Imaging live cell in micro-liquid enclosure by X-ray laser diffraction. *Nature Communications*, 5:1--7, 2014.

[43] J. Fan, Z. Sun, Y. Wang, J. Park, S. Kim, M. Gallagher-Jones, Y. Kim, C. Song, S. Yao, J. Zhang, J. Zhang, X. Duan, K. Tono, M. Yabashi, T. Ishikawa, C. Fan, Y. Zhao, Z. Chai, X. Gao, T. Earnest, and H. Jiang. Single-pulse enhanced coherent diffraction imaging of bacteria with an X-ray free-electron laser. *Scientific Reports*, 6(1):34008, 2016.

[44] K. Okamoto, N. Miyazaki, D. S. D. Larsson, D. Kobayashi, M. Svenda, K. Mühlig, F. R. N. C. Maia, L. H. Gunn, H. Isawa, M. Kobayashi, K. Sawabe, K. Murata, and J. Hajdu. The infectious particle of insect-borne totivirus-like Omono River virus has raised ridges and lacks fibre complexes. *Scientific Reports*, 6(1):33170, 2016.

[45] D. P. DePonte, U. Weierstall, K. Schmidt, J. Warner, D. Starodub, J. C. H. Spence, and R. B. Doak. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics*, 41(19):195505, 2008.

[46] A. M. Gañán-Calvo, D. P. DePonte, M. A. Herrada, J. C. H. Spence, U. Weierstall, and R. B. Doak. Liquid Capillary Micro/Nanojets in Free-Jet Expansion. *Small*, 6(7):822--824, 2010.

[47] P. Hart, S. Boutet, G. Carini, A. Dragone, B. Duda, D. Freytag, G. Haller, R. Herbst, S. Herrmann, C. Kenney, J. Morse, M. Nordby, J. Pines, N. van Bakel, M. Weaver, and G. Williams. The Cornell-SLAC pixel array detector at LCLS. In *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, pages 538--541. IEEE, 2012.

[48] G. A. Carini, S. Boutet, M. Chollet, A. Dragone, G. Haller, P. A. Hart, S. C. Herrmann, C. J. Kenney, J. Koglin, M. Messerschmidt, S. Nelson, J. Pines, A. Robert, S. Song, J. B. Thayer, G. J. Williams, and D. Zhu. Experience with the CSPAD during dedicated detector runs at LCLS. *Journal of Physics: Conference Series*, 493(1):012011, 2014.

[49] M. J. Bogan, W. H. Benner, S. Boutet, U. Rohner, M. Frank, A. Barty, M. M. Seibert, F. Maia, S. Marchesini, S. Bajt, B. Woods, V. Riot, S. P. Hau-Riege, M. Svenda, E. Marklund, E. Spiller, J. Hajdu, and H. N. Chapman. Single Particle X-ray Diffractive Imaging. *Nano Letters*, 8(1):310--316, 2008.

[50] M. O. Wiedorn, S. Awel, A. J. Morgan, M. Barthelmess, R. Bean, K. R. Beyerlein, L. M. G. Chavas, N. Eckerskorn, H. Fleckenstein, M. Heymann, D. A. Horke, J. Knoška, V. Mariani, D. Oberthür, N. Roth, O. Yefanov, A. Barty, S. Bajt, J. Küpper, A. V. Rode, R. A. Kirian, and H. N. Chapman. Post-sample aperture for low background diffraction experiments at X-ray free-electron lasers. *Journal of Synchrotron Radiation*, 24(6):1--3, 2017.

[51] N. D. Loh. A minimal view of single-particle imaging with X-ray lasers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1647):20130328--20130328, 2014.

[52] J. Chalupsky, P. Bohacek, V. Hajkova, S. Hau-Riege, P. Heimann, L. Juha, J. Krzywinski, M. Messerschmidt, S. Moeller, B. Nagler, M. Rowen, W. Schlotter, M. Swiggers, and J. Turner. Comparing different approaches to characterization of focused X-ray laser beams. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 631(1):130--133, 2011.

[53] N. D. Loh, D. Starodub, L. Lomb, C. Y. Hampton, A. V. Martin, R. G. Sierra, A. Barty, A. Aquila, J. Schulz, J. Steinbrener, R. L. Shoeman, S. Kassemeyer, C. Bostedt, J. Bozek, S. W. Epp, B. Erk, R. Hartmann, D. Rolles, A. Rudenko, B. Rudek, L. Foucar, N. Kimmel, G. Weidenspointner, G. Hauser, P. Holl, E. Pedersoli, M. Liang, M. S. Hunter, L. Gumprecht, N. Coppola, C. Wunderer, H. Graafsma, F. R. Maia, T. Ekeberg, M. Hantke, H. Fleckenstein, H. Hirsemann, K. Nass, T. A. White, H. J. Tobias, G. R. Farquar, W. H. Benner, S. Hau-Riege, C. Reich, A. Hartmann, H. Soltau, S. Marchesini, S. Bajt, M. Barthelmess, L. Strueder, J. Ullrich, P. Bucksbaum, M. Frank, I. Schlichting, H. N. Chapman, and M. J. Bogan. Sensing the wavefront of x-ray free-electron lasers using aerosol spheres. *Optics Express*, 21(10):12385, 2013.

[54] K. R. Ferguson, M. Bucher, J. D. Bozek, S. Carron, J.-C. Castagna, R. Coffee, G. I. Curiel, M. Holmes, J. Krzywinski, M. Messerschmidt, M. Minitti, A. Mitra, S. Moeller, P. Noonan, T. Osipov, S. Schorb, M. Swiggers, A. Wallace, J. Yin, and C. Bostedt. The Atomic, Molecular and Optical Science instrument at the Linac Coherent Light Source. *Journal of Synchrotron Radiation*, 22(3):492--497, 2015.

[55] B. Enders and P. Thibault. A computational framework for ptychographic reconstructions. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 472(2196):20160640, 2016.

[56] F. R. N. C. Maia, T. Ekeberg, D. van der Spoel, and J. Hajdu. Hawk : the image reconstruction package for coherent X-ray diffractive imaging. *Journal of Applied Crystallography*, 43(6):1535--1539, 2010.

[57] F. R. N. C. Maia. The Coherent X-ray Imaging Data Bank. *Nature Methods*, 9(9):854--855, 2012.

[58] X. C. Bai, G. McMullan, and S. H. Scheres. How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences*, 40(1):49--57, 2015.

[59] S. Jonić. Cryo-electron Microscopy Analysis of Structurally Heterogeneous Macromolecular Complexes. *Computational and Structural Biotechnology Journal*, 14:385--390, 2016.

[60] K. R. Beyerlein, D. Dierksmeyer, V. Mariani, M. Kuhn, I. Sarrou, A. Ottaviano, S. Awel, J. Knoska, S. Fuglerud, O. Jönsson, S. Stern, M. O. Wiedorn, O. Yefanov, L. Adriano, R. Bean, A. Burkhardt, P. Fischer, M. Heymann, D. A. Horke, K. E. J. Jungnickel, E. Kovaleva, O. Lorbeer, M. Metz, J. Meyer, A. Morgan, K. Pande, S. Panneerselvam, C. Seuring, A. Tolstikova, J. Lieske, S. Aplin, M. Roessle, T. A. White, H. N. Chapman, A. Meents, and D. Oberthuer. Mix-and-diffuse serial synchrotron crystallography. *IUCrJ*, 4(6):1--9, 2017.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1589

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-329012

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017