Ultrafast coherent diffractive imaging of nanoparticles using X-ray free-electron laser radiation

im Fachbereich Physik der Freien Universität Berlin
eingereichte Dissertation

vorgelegt von
Stephan Kassemeyer
geboren in Braunschweig

April 2014

Disputation: 20. Mai 2014

Gutachter: Prof. Dr. Robert Bittl (FU Berlin)
Hon.-Prof. Dr. Ilme Schlichting

Abstract

Coherent diffractive imaging with *X-ray free-electron lasers* (X-FEL) promises high-resolution structure determination of single microscopic particles without the need for crystallization. The diffraction signal of small samples can be very weak, a difficulty that can not be countered by merely increasing the number of photons because the sample would be damaged by a high absorbed radiation dose. Traditional X-ray crystallography avoids this problem by bringing many sample particles into a periodic arrangement, which amplifies the individual signals while distributing the absorbed dose. Depending on the sample, however, crystallization can be very difficult or even impossible. This thesis presents algorithms for a new imaging approach using X-FEL radiation that works with single, non-crystalline sample particles.

X-FELs can deliver X-rays with a peak brilliance many orders of magnitude higher than conventional X-ray sources, compensating for their weak interaction cross sections. At the same time, FELs can produce ultra-short pulses down to a few femtoseconds. In this way it is possible to perform ultra-fast imaging, essentially "freezing" the atomic positions in time and terminating the imaging process before the sample is destroyed by the absorbed radiation.

This thesis primarily focuses on the three-dimensional reconstruction of single (and not necessarily crystalline) particles using coherent diffractive imaging at X-FELs: in order to extract three-dimensional information from scattering data, two-dimensional diffraction patterns from many different viewing angles must be combined. Therefore, the diffraction signal of many identical sample copies in random orientations is measured. The main result of this work is a globally optimal algorithm that can recover the sample orientations solely based on the diffraction signal, enabling three-dimensional imaging for arbitrary samples. The problem of finding three-dimensional orientations is reduced to one-dimensional sub-problems by arranging diffraction patterns in geodesic similarity sequences. Relations between the one-dimensional sub-problems are established by identifying rotations about the X-ray axis and one-dimensional solutions are combined into a three-dimensional orientation recovery. The global optimization approach ensures that information is extracted from the whole diffraction dataset, not only individual diffraction patterns. Therefore this method can cope with diffraction data sets consisting of individual diffraction patterns with weak signals. The geodesic approach can handle datasets from inhomogeneous samples as well as samples with symmetries. A successful application to experimental X-FEL data is shown, resulting in the first three-dimensional reconstruction of a nanoparticle using X-FEL coherent diffractive imaging.

Zusammenfassung

Kohärente Abbildung mit Röntgenlasern (*X-ray free-electron lasers*, X-FEL) ermöglicht die Strukturbestimmung von einzelnen mikroskopischen Teilchen mit hoher Auflösung, ohne dass ihre Kristallisation notwendig ist. Das gestreute Signal von kleinen Proben kann jedoch sehr schwach sein. Diese Schwierigkeit kann nicht einfach durch mehr einfallende Photonen umgangen werden, da die Probe bei der Absorption einer hohen Strahlendosis Schaden nimmt. Herkömmliche Kristallographie vermeidet dieses Problem durch das periodische Anordnen vieler Probenteilchen, wodurch das Signal verstärkt und die Strahlendosis verteilt wird. Je nach Probe kann die Kristallisation jedoch sehr aufwändig oder gar unmöglich sein. Diese Arbeit behandelt Algorithmen für ein neues bildgebendes Verfahren mit X-FEL Strahlung, das ohne Kristallisation auskommt.

Mit X-FELs können Röntgenstrahlen mit sehr viel höherer Spitzenbrillanz erzeugt werden als mit herkömmlichen Röntgenquellen; somit können die schwachen Wechselwirkungsquerschnitte von Röntgenphotonen mit Materie kompensiert werden. Gleichzeitig können diese Röntgenstrahlen sehr kurz gepulst werden, bis hin zu wenigen Femtosekunden. Dadurch kann eine Bildgebung erreicht werden, die so schnell ist, dass die Atompositionen zeitlich „eingefroren“ werden und ein Abbild der Probe erzeugt wird, bevor diese durch die absorbierte Strahlung zerstört wird.

Das Hauptaugenmerk dieser Arbeit liegt auf der dreidimensionalen Rekonstruktion: Um dreidimensionale Information aus Streudaten zu gewinnen ist es erforderlich viele zweidimensionale Streubilder aus verschiedenen Blickwinkeln zusammenzufassen. Dazu werden Streubilder von vielen identischen Kopien der Probe sequentiell gesammelt, wobei jede Probenkopie eine zufällige Orientierung hat. Das wichtigste Ergebnis dieser Arbeit ist ein global optimaler Algorithmus, der die Orientierungen allein mit Hilfe der Streubilder rekonstruiert, wodurch eine dreidimensionale Bildgebung für beliebige Proben möglich wird. Dazu wird das Problem dreidimensionale Orientierungen zu rekonstruieren in eindimensionale Teilprobleme unterteilt, indem Streubilder aufgrund ihrer Ähnlichkeit in geodätische Bildfolgen angeordnet werden. Die eindimensionalen Teilprobleme werden dann miteinander in Bezug gebracht, indem gemeinsame Drehungen um die Röntgenachse identifiziert werden. Somit können eindimensionale Lösungen in eine dreidimensionale Rekonstruktion der Orientierungen kombiniert werden. Die globale Optimierung stellt dabei sicher, dass die Information des gesamten Datensatz genutzt wird, anstatt nur einzelne Streubilder zu berücksichtigen. Aus diesem Grund kann diese Methode auch bei Datensätzen eingesetzt werden, bei denen einzelne Streubilder nur ein schwaches Signal erhalten. Die auf Geodäten beruhende Methode kann sowohl Datensätze von inhomogenen Proben bewältigen, als auch mit Objektsymmetrien umgehen. In dieser Arbeit wird eine erfolgreiche Anwendung auf experimentelle X-FEL Daten gezeigt, die die erste dreidimensionalen Rekonstruktion eines Nanopartikels mit Hilfe von kohärenten Abbildungen mit X-FELs ermöglichte.

Table of Contents

Part I Background

Part II Results

Publications

Major parts of this thesis have been published in the following articles:

Femtosecond free-electron laser X-ray diffraction data sets for algorithm development

Kassemeyer S, Steinbrener J, Lomb L, Hartmann E, Aquila A, Barty A, Martin AV, Hampton CY, Bajt S, Barthelmess M, Barends TR, Bostedt C, Bott M, Bozek JD, Coppola N, Cryle M, DePonte DP, Doak RB, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hartmann A, Hartmann R, Hauser G, Hirsemann H, Hömke A, Holl P, Jönsson O, Kimmel N, Krasniqi F, Liang M, Maia FR, Marchesini S, Nass K, Reich C, Rolles D, Rudek B, Rudenko A, Schmidt C, Schulz J, Shoeman RL, Sierra RG, Soltau H, Spence JC, Starodub D, Stellato F, Stern S, Stier G, Svenda M, Weidenspointner G, Weierstall U, White TA, Wunderer C, Frank M, Chapman HN, Ullrich J, Strüder L, Bogan MJ, Schlichting I.
Optics Express **20** 4, pp 4149-4158 (February 2012)

Optimal mapping of x-ray laser diffraction patterns into three dimensions using routing algorithms

Kassemeyer S⋆, Jafarpour A, Lomb L, Steinbrener J, Martin AV, Schlichting I⋆.
Physical Review E **88** 4, p 042710 (October 2013)

⋆ corresponding author

Further publications coauthored by Stephan Kassemeyer are listed in appendix D.

Introduction

Visualizing the microscopic world at the length scale of atoms is a fascinating undertaking that gives exciting insights to many scientific fields. The development of high resolution imaging methods has boosted physics, material science, chemistry and in particular biology. In biological systems, function and structure are inseparably intertwined. This holds for whole organisms as well as for cells and for individual molecules. One of the best known examples of the impact of structural knowledge on functional understanding is the structure of the DNA double helix that was first proposed in 1953 by Watson and Crick. As Watson and Crick already recognized, the base pairing in their model "immediately suggests a possible copying mechanism for the genetic material" [1]. It is probably fair to say that the insight provided by this single structure has revolutionized biology. Additional examples of such high impact results are the structures of RNA polymerase (for which a Nobel prize was awarded in 2006), the ribosome (Nobel prize in 2009), enzymes such as the ATP synthase (Nobel prize in 1996) and several membrane proteins (Nobel prizes in 1988, 2003 and 2012). At the time this thesis was written, the protein data bank (PDB) held over 99,000 structures of biological macromolecules. This considerable body of information has, over the last couple of decades, changed the way in which we view the living world.

Currently used imaging techniques

The vast majority of these structures was determined by X-ray crystallography, a technique that was pioneered 100 years ago by Sir William Henry Bragg and his son William Lawrence Bragg. Since then it has been successfully in use to determine structures at or near atomic resolution. X-rays are an ideal probe for high-resolution three-dimensional imaging because of their short wavelengths and their long penetration depth. X-ray crystallography takes advantage of the amplification that is achieved due to the scattering from many periodically arranged units, the *unit cells*, that adds up constructively, causing sharp maxima, the so called *Bragg reflections*. In this way, N identical copies of the sample arranged in a three-dimensional lattice result in an amplified scattering signal with a Bragg reflection intensity I that grows quadratically with N: $I \propto N^2$. Thus, using crystals with many unit cells, the dose of ionizing X-ray radiation can be reduced, mitigating the effects of radiation damage.

Even though the experiments are performed on crystalline samples which has the potential to introduce artifacts, it has been shown that biological molecules generally retain the same structure as in solution and often even at least part of their activity. This enables studies on biochemical mechanisms, for instance by introducing reaction partners into protein crystals and observing chemical reactions using crystallography.

However, the method crucially relies on the ability to grow macroscopic, well-ordered crystals, a requirement that is not always easily met. Membrane proteins and macromolecular complexes are difficult to crystallize. Even if a successful crystallization can be obtained, it is not guaranteed that the native sample environment is preserved. Moreover, while crystallographic amplification reduces the dose required to obtain sufficient signal, a typical radiation dose used to determine a protein structure can be as much as 20 MGy [2], a dose which is certain to cause significant damage to the sample. Data collection at cryogenic temperatures normally enables the crystal to endure such harsh treatments but in several cases the radiation causes significant alterations to the chemistry of the molecule, rendering the resulting structure useless for biological interpretation. Examples include metalloproteins containing atoms with high atomic numbers such as the manganese cluster in photosystem II [3] and heme proteins such as chloroperoxidase [4].

Nuclear magnetic resonance spectroscopy (NMR) enables structure determination without ionizing radiation and without requiring crystals. NMR is a spectroscopic method that can be applied to samples in solution or solid state which relies on the spin properties of atomic nuclei.

However, this method has disadvantages, too. NMR signals are typically weak so that many spectra need to be averaged to obtain a sufficiently high signal to noise ratio. This means that experiments often lasts for several days, requiring the sample to be stable over the course of the experiment. Moreover, in NMR spectroscopy, peak widths are inversely related to the speed with which the molecules tumble in solution. In general, this puts an upper limit of about 50 kDa on the molecular weight of the molecules that can be investigated, as large molecules give rise to broad, overlapping peaks, resulting in uninterpretable spectra.

Electron microscopy (EM) is an imaging method like X-ray crystallography, but uses electrons rather than X-rays. As these have a much higher scattering cross section, EM is capable of imaging *single molecules* so that crystals are not required [5, 6]. In a typical EM experiment, the molecules under study are embedded in a thin film of vitreous ice and imaged using as low a radiation dose as possible to minimize damage. This results in low signal-to-noise ratios of the individual images, so that several thousands of images need to be collected, containing many two-dimensional projections of different copies of the molecule. Individual molecules are then identified in these images and grouped into classes, representing different orientations and conformations, carrying dynamic information. Special algorithms are then employed to reconstruct three-dimensional structures from these projections. Like NMR, this allows structure determination in a near native environment without requiring crystals. In addition, the identification of dynamic features is possible.

To a large extent, achievable resolution is limited by radiation damage; the large cross sections that allow single particle imaging in the first place are also responsible for rapid deterioration of the sample during imaging. However, as shown by recent reports of near atomic resolution structures [7], electron microscopy is currently being revolutionized by new detectors [8].

At the low doses typically employed to counter this and other effects, identification of the molecules for reconstruction becomes difficult, if not impossible for small molecules, currently putting a lower size limit of around 200-300 kDa on the molecules that can be studied with EM. Given the mean free path of an electron in biological materials, there is an upper limit to the thickness of the sample, too, precluding the investigation of e.g. the inside of intact large viruses.

Towards a new method

As shown above, the structural biologist's toolkit contains a variety of methods for structure determination, each with its own advantages and disadvantages, with X-ray crystallography being the principal technique. However, as also shown, there is a need for methods that allow the study of molecules that are difficult to crystallize and cannot be studied by EM or NMR, for instance because of size limitations. Also, the high-resolution information that X-ray crystallography can provide is sometimes rendered useless by radiation damage, for instance in the case of highly radiation-sensitive cofactors. This provides the motivation for the work described in this thesis, which presents important first steps towards a new imaging method, three-dimensional *ultrafast X-ray Coherent Diffractive Imaging* (CDI).

The small interaction cross sections of X-ray photons preclude the imaging of single biological macromolecules as in EM, unless a sufficiently high number of photons is used. At such high doses however, radiation damage would completely destroy the molecule under study.

To counter this, it was proposed by Solem [9] to combine this large number of photons in a single pulse of ultra short duration, passing the sample before the molecule is destroyed by radiation damage. This 'outrunning' of radiation damage is called the '*Diffraction-before-Destruction*' approach.

Neutze et al. [10] performed theoretical calculations that supported this idea. In molecular dynamics simulations of a single lysozyme molecule, the effect of sudden, massive ionization as caused by a femtosecond X-ray pulse was analyzed. It was predicted that the structure remains largely intact during an X-ray pulse of a few femtoseconds duration and that the molecule is destroyed in a "*Coulomb explosion*" only after the pulse has passed.

Several authors have characterized the damage caused by high intensity femtosecond X-ray pulses in more detail [11, 12, 13, 14].

The dominant primary process is photoionization, whereby a photon is absorbed and an electron is ejected. The lifetimes of the initially generated electron holes relevant for radiation damage are on the order of femtoseconds [10]. Relaxation occurs when an electron from a higher-lying shell fills the vacant orbital. The energy of the relaxing electron can either result in fluorescence or in the ejection of another electron in a process known as the *Auger effect*. In heavier elements, fluorescence is the dominant process whereas in lighter elements such as carbon, nitrogen, oxygen and sulfur, the Auger effect is more likely. The electrons that are ejected by primary processes such as photoelectric effect, re-absorption of the X-ray fluorescences as well Auger electrons lead to avalanches of free electrons which do even more damage to the sample. The final result is the Coulomb explosion: eventually, all atoms of the sample are highly ionized, causing strong mutual electrostatic repulsion which causes the sample to explode. However analysis of the time scales involved predicts that significant radiation damage will only set in after the probing X-ray pulse has left the sample. Moreover, imaging experiments indeed confirm the feasibility of 'Diffraction-before-Destruction' [15].

As in crystallography, CDI requires the retrieval of the phases of the scattered electromagnetic wave. In contrast to crystallography, where the diffraction signal can only be sampled in Bragg reflections, in CDI the sample is non-periodic, giving rise to a continuous diffraction pattern. This enables a sampling rate beyond the Shannon limit [16]. As will be explained later on (see chapter 1.1.1), this provides a set of constraints that can be exploited in iterative phase retrieval algorithms.

Free-electron lasers

Imaging with ultrashort pulses has now become possible with the advent of X-ray *free-electron lasers* (FEL). These novel light sources provide the high peak *peak brilliance* (spectral brightness) and the short pulses required for the Solem experiment. X-ray FELs are sometimes referred to as fourth generation X-ray sources which reflects their descent from synchrotron radiation sources. Contrary to synchrotrons, where electrons are kept in a circular storage ring, in FELs electrons are accelerated to relativistic speed in a linear accelerator. The relativistic electrons are then introduced into an undulator; a periodic arrangement of magnets with alternating polarity, forcing the electrons on a sinusoidal path with a fixed phase relative to the electromagnetic wave. The resulting alternating acceleration causes the charged particles to emit radiation. The use of a linear accelerator allows to decrease the beam *emittance*; the average spread of electrons in six-dimensional phase space. If the emittance is sufficiently low, interaction occurs between electrons emitted photons, resulting in a nonlinear coupling that groups the electrons into *microbunches* with the periodicity of the generated radiation. This introduces a temporal correlation with the motion of the electrons. In the limit of perfect microbunching with bunch sizes much smaller than the radiation wavelength, all N_e electrons of a bunch train radiate in phase. This leads to a significant amplification of the radiation power ($\propto N_e^2$ as opposed to $\propto N_e$ in case of uncorrelated emission). The light that is produced in this way is spatially and temporally coherent, like laser light. Unlike conventional lasers, which use electrons in bound states as the gain medium, FELs use free electrons as a lasing medium. Consequently, the resulting wavelength can be tuned in a continuous fashion within a wide spectral range, including the X-ray region. The operating principle of FELs is similar to that of synchrotrons, however, the resonance effect leads to a much higher peak brilliance. In current realizations, the peak brilliance exceeds that of synchrotrons by nine orders of magnitude. At present, two hard X-ray FELs are available worldwide, the *Linac Coherent Light Source* (*LCLS*) LCLS which went operational for users in 2009 [17] and the *Spring-8 Angstrom Compact free electron LAser* (SACLA) which followed in 2011. Since then, the construction and planning of many more facilities has progressed.

Serial coherent diffractive imaging

While diffraction using FEL pulses yields diffraction information of the undamaged sample in femtoseconds, the following Coulomb explosion inevitably results in the complete destruction of the sample [10].

Because, as in electron microscopy, a large number of snapshots is required for a three-dimensional reconstruction and to improve the signal to noise ratio by averaging, the sample must continuously be replenished during the experiment. Therefore, the imaging experiment is performed in a serial way, as shown in figure 1. Single objects have to be separated from the bulk and guided to the interaction region. An additional complication is that the imaging (including scattering and detection) is performed in high vacuum ($\sim 10^{-7}$ mbar) to avoid absorption of the incoming X-ray beam and absorption of the signal, and to suppress background scattering. Three different, vacuum compatible techniques for serial sample delivery are currently used: aerosol injection, liquid jet injection and pre-aligned objects suspended on thin membranes.

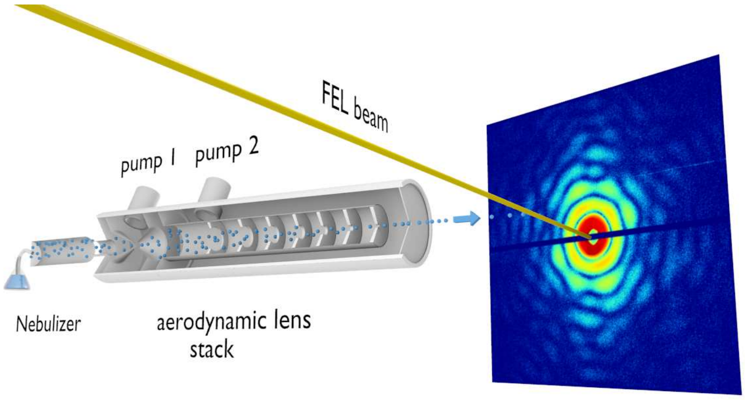


Fig. 1. Setup of coherent diffraction imaging experiments: a particle stream is intercepted by the pulsed X-ray beam. In this example, a nebulizer aerosolizes the sample and an aerodynamic lens stack is used to focus the particle stream. Diffraction patterns are measured by means of CCD pixel array sensors which are triggered with the X-ray pulses. The diffraction pattern depicted on the detector plane originates from a T4 bacteriophage virus.

Aerosol injection

The goal of aerosol injection is to introduce single sample particles into the X-ray beam without the use of substrates, conserving a liquid sample environment, reducing physical stress on the sample object and reducing background scattering by removing as much of the solvent as possible. This is achieved by aerosolizing the solvated sample and forming a molecular beam guided by a low density carrier gas. In order to increase the chance of hitting a single particle with every X-ray pulse, the particle beam is focused by means of gas-dynamic focusing [18]. The focusing principle is illustrated in figure 2, figure 1 shows the setup of an aerodynamic lens stack with differential pumping: the pressure is decreased along the path of the particle stream and excess gas is removed before particles enter the vacuum chamber.

This sample delivery method relies on evaporation of the solvent shell around sample particles during their flight into the interaction region. However, this evaporation is not necessarily uniform and sample particles may be left with different amounts of residual solvent resulting in differently sized droplets. Moreover, the drying process might leave a crust of dried buffer components on the surface of the particles. In addition, multiple particles might reach the interaction region at the same time, complicating the analysis of the diffraction pattern. Thus, all parameters affecting the injection process such as gas flow, initial aerosol droplet size, particle density etc. need to be optimized such that the chance of “hitting” single particles is maximized. The substrate-free aerosol injection offers the possibility to perform imaging on samples in their native state. However, only volatile buffers can be used.

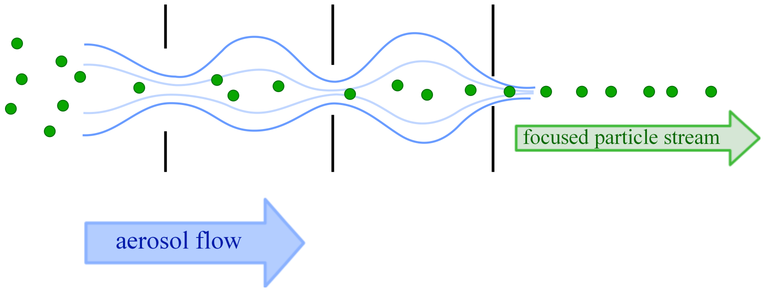


Fig. 2. Illustration of the focusing principle of an aerodynamic lens stack. At each aperture, the gas is compressed, focusing the embedded particles. After it passes the aperture, the gas can relax and expand, the particles however stay on the central line due to inertia. The diameters of the apertures decrease along the lens stack.

Liquid jet injection

In liquid jet injection the goal is to inject the sample into the FEL vacuum chamber while still in native aqueous solution, i.e. it bypasses the need for aerosolisation and drying required for aerosol injection and eliminates the need for volatile buffers. To generate a micro jet, the solution containing the sample particles is extruded while surrounded by a sheath of convergent, co-flowing gas. This gas sheath forms a virtual nozzle [19] (see figure 3), focusing the jet into a micron-sized liquid column. A virtual nozzle, i.e. one made of gas, rather than a “real”, physical nozzle is used since these are less prone to clogging, which is a considerable problem when producing microscopic jets of biological samples. In addition, the sheathing gas prevents freezing of the sample when the liquid enters the vacuum chamber. In order not to compromise the high vacuum environment of the experimental chamber, the co-flowing focusing gas as well as the liquid are encased within a differentially-pumped shroud.

Gas dynamic virtual nozzles (GDVNs) achieve much higher injection efficiencies than aerosol injectors, resulting in a much higher “hit rate”, the fraction of FEL pulses that results in a successful diffraction event. However, the solvent results in a high background signal, and a strong diffraction signature which complicates analysis is caused by the surface of the liquid column (see figure 3). Therefore, in order to achieve sufficient contrast between the solvent and the sample particles, the measurements have to be performed in the “water window” (at a wavelength between

2.3 nm and 4.4 nm) where the absorption contrast between carbon (sample) and oxygen (water) is high. Consequently, the achievable resolution is limited.

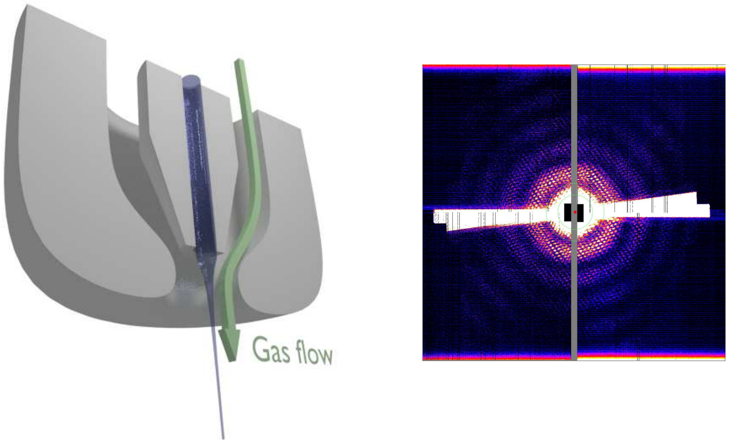


Fig. 3. Liquid jet. **Left**: cutaway drawing of a gas dynamic virtual nozzle. **Right**: diffraction pattern with horizontal (orthogonal to jet) streak signature from diffraction of the jet's surface (sample: Cro-virus). The vertical line is a gap between two detector halves.

Fixed target

Both aerosol and liquid jet injection employ a continuous stream of particles which is then intercepted by discrete X-ray pulses. The sample material injected between subsequent FEL pulses is wasted. This can be avoided by depositing single sample particles onto thin membranes analogous to the sample holders utilized in EM. A disadvantage of this method is high background scattering, also the orientations of the sample objects may be biased because of preferred alignment to the substrate in specific orientations. Fixed targets are nonetheless useful to perform two-dimensional imaging, providing a very stable mode of sample delivery and everything can be prepared and characterized beforehand.

Sorting diffraction patterns

The serial sample delivery needed for three-dimensional coherent diffractive imaging with FELs can introduce artifacts such as *sample aggregates*, *solvent droplets*, *blank shots*, diffraction from multiple objects (*multi-hit*) and alien objects from impure sample preparations (see figure 4). To a large extent, such artifacts can be recognized before a three-dimensional reconstruction is performed, easing the reconstruction process and reducing the chance of such artifacts impairing the result. Statistical sorting methods have been developed in this thesis to address this problem.

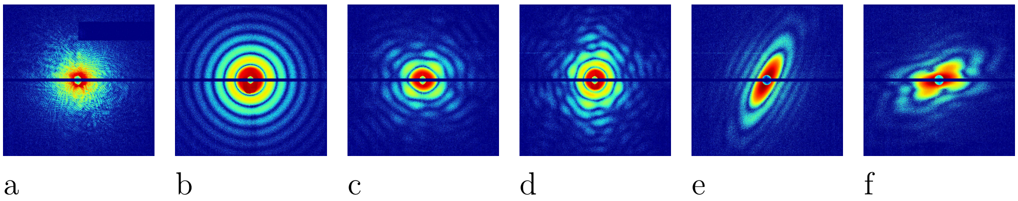


Fig. 4. Typical diffraction patterns of different samples and injection artifacts. In addition to the desired single sample particles, solvent droplets, sample aggregates and multiple particles may be present in the interaction region at the time the X-ray pulse passes. (a) diffraction pattern of a large aggregate, (b) solvent droplet, (c,d) single T4 virus phage particles, (f) single nanorice grain, (g) two grains of nanorice at the same time.

The orientation problem

A diffraction pattern provides only two-dimensional information. For a three-dimensional reconstruction, diffraction snapshots in different orientations need to be combined. In crystallography, this can be achieved by rotating the macroscopic crystal while measuring the diffraction signal. Discrete Bragg reflections allow an exact interpretation of the orientation of a given diffraction pattern.

However, the single particle delivery methods employed for CDI described above share one important limitation: the orientation of the sample object can neither be controlled nor directly measured. Instead, sufficient sampling of orientation space is achieved because the sample objects are injected in random orientations. If the orientations are known, two-dimensional snapshots can be arranged into a three-dimensional data set. The same problem occurs in electron microscopy with the important difference that in EM, orientations can be recovered from real space projections of the object under study, whereas in CDI the orientations need to be recovered from diffraction data. Another important difference to orientation recovery in EM is that the expected signal strength for CDI on individual molecules is significantly lower than in EM. This also necessitates the development of new approaches to sorting and orientation.

The *orientation problem*, the task of obtaining the orientations *a posteriori*, can be solved because a set of diffraction patterns originating from the same object, with diffraction snapshots that only differ in the object orientations will show correlations that can be identified and exploited to obtain relative orientation information.

Motivation and outline

The ultimate goal of CDI at X-ray FEL sources is the three-dimensional structure determination of *single molecules*. At present, many experimental and theoretical obstacles still remain and this goal is not within reach. However, during the time of this thesis, successful proof-of-principle experiments have been performed using larger particles, providing strong scattering signals. The demonstration of a three-dimensional reconstruction of such a nanoparticle by serial CDI marks an important milestone on the way towards single molecule imaging. Before this thesis, such a reconstruction using experimental data had never been shown.

The main goal of this work was therefore the development and successful application of data evaluation methods that combine the information of many weak two-dimensional diffraction signals into a reliable three-dimensional image of the sample object. This includes the development of a new algorithm for orientation recovery as well as statistical methods to sort diffraction patterns.

First, the concepts of coherent diffractive imaging are further developed in chapter 1. It is shown how a real-space reconstruction can be obtained from diffraction data using oversampling-based phase retrieval algorithms. Specifically, the challenges of three-dimensional reconstructions are considered.

Then, statistical classification methods for the sorting of diffraction snapshots are described in chapter 2. These methods are not only useful to automate the sorting of many thousands of diffraction patterns, but also to provide useful feedback during imaging experiments that aids in optimizing the injection parameters highlighted above.

After this, a novel orientation recovery algorithm, GIPRAL, is introduced and developed in chapter 3. Finally, the classification methods developed in chapter 2 and the GIPRAL method are successfully applied to real data collected at an X-ray FEL in chapter 5, resulting in the first demonstration of the three-dimensional reconstruction of a nanoparticle from serial CDI data.

Part I

Background

1

The theory of coherent X-ray diffractive imaging

Simple transmission lenses like those used in light microscopes are not available for X-rays, since the refractive index of all materials is close to 1 for such wavelengths. Grazing incidence reflective optics [20] are used instead to build X-ray focusing optics. However they can not withstand high intensities and introduce a considerable amount of absorption and aberration, limiting their usefulness in a high resolution imaging system.

Therefore, the work presented here focuses on "*lensless imaging*" where the image is reconstructed computationally, making use of the coherence of the incident light wave. Without lenses, the Fourier plane of the ray optics illustrated in figure 1.1 is placed at infinity and can approximately be detected by a detector placed at a long distance. The the inverse transformation back to real space has to be carried out computationally.

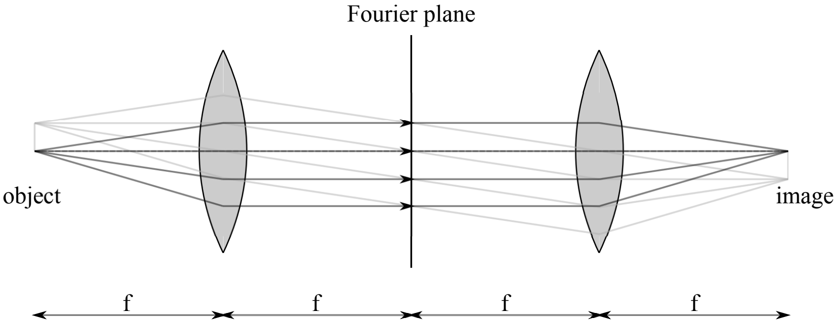


Fig. 1.1. Fourier optics in the $4f$ geometry. The first lens located at the focal length f refracts the cone of rays that originates from each object point into a parallel beam. Every point of the object is spread out into the Fourier plane where a superposition of the waves is formed that can be described by a Fourier transform (see eq. 1.1). The same parallel superposition happens without a lens on a detector that is placed at infinity. The second lens reverses the process which can be described by the inverse Fourier transform. Without a lens the rays keep diverging and thus the forward transform continues to infinity, but still the inverse process can be performed computationally.

When the light scatters forward from the sample object onto a detector, each point on the detector receives light from every point of the sample object, giving rise to a superposed detector signal which will be called *diffraction pattern*. Only elastic scattering will be considered in the following. In order to reconstruct the object image by reversing the formation of the diffraction pattern, the *far field approximation* can be used. The assumption is that the detector is practically placed at infinity when its distance is compared to the wavelength and the size of the sample object. Then, the rays that originate from arbitrary points of the object can be considered to be parallel. The intensity that is registered at the detector is the coherent sum over these rays. The addends of this sum are the electromagnetic waves that originate from single points $\boldsymbol{r}$ in the sample via *Thomson scattering* [21], the elastic scattering by the electrons of the sample. In the far field approximation, diffraction can be modeled by *Fraunhofer diffraction* which is the topic of many textbooks (see e.g. [22, 21]). In the following, this model is introduced following the notation of [23]. The amplitude $\mathrm{d}A$ of the scattered wave originating at a small volume element $\mathrm{d}V$ is proportional to the probability of encountering an electron inside $\mathrm{d}V$ which can be expressed in terms of the electron density $\rho(\boldsymbol{r})$ as $\mathrm{d}V \cdot \rho(\boldsymbol{r})$. The phase $\mathrm{d}\varphi$ of this wave relative to the wave originating from a reference point $\boldsymbol{r}_0 = (0,\ 0,\ 0)^T$ (which, for convenience defines the origin) and travels along the direction of its wave vector $\boldsymbol{k}$ can be calculated as $\mathrm{d}\varphi = (\boldsymbol{k} - \boldsymbol{k}_0) \cdot \boldsymbol{r}$ because the complete path difference relative to the incident beam is $\boldsymbol{k} \cdot \boldsymbol{r}/k - \boldsymbol{k}_0 \cdot \boldsymbol{r}/k$ (see fig 1.2) with $\boldsymbol{k}_0$ and $\boldsymbol{k}$ being the wave vector of the incident wave and the diffracted wave, respectively, and $|\boldsymbol{k}| = |\boldsymbol{k}_0| := k$ because the scattering is elastic. Assuming a planar incident wavefront $e^{ikx-\omega t}$, the wave that leaves the volume element $\mathrm{d}V(\boldsymbol{r})$ along the direction $\boldsymbol{k}$ can be expressed as the following plane wave at infinity:

$$\mathrm{d}A(\boldsymbol{r}, t) \;=\; \frac{A_0}{|\boldsymbol{r}_{\mathrm{det}}|}\, \rho(\boldsymbol{r})\, e^{i[k \cdot r_{\mathrm{det}} - (\boldsymbol{k}-\boldsymbol{k}_0)\cdot \boldsymbol{r} - \omega t]}\, \mathrm{d}V$$

where $\boldsymbol{r}_{\mathrm{det}}$ is the position of the detector relative to $\boldsymbol{r}_0$ and ω is the angular frequency of the incident radiation. In an exact formulation, $\boldsymbol{r}_{\mathrm{det}}$ would have to be replaced by the position of the detector relative to $\boldsymbol{r}$ instead of $\boldsymbol{r}_{\mathrm{det}}$. Not doing so affects both the phase and intensity of the wave front at the detector. However the phase shift is constant and so of no consequence in an intensity measurement, and the intensity difference is negligible given the mismatch in detector distance relative to object size. Moreover $\boldsymbol{r}_{\mathrm{det}} \,||\, \boldsymbol{k}$, therefore $\boldsymbol{k} \cdot \boldsymbol{r}_{\mathrm{det}} = |\boldsymbol{k}| \cdot |\boldsymbol{r}_{\mathrm{det}}| := k \cdot r_{\mathrm{det}}$.

The total amplitude of the wave that is detected at $\boldsymbol{r}_{\text{det}}$ can be calculated as the sum of all amplitudes over all sample volume elements, or the integral

$$A(t,\boldsymbol{q}) \;=\; \frac{A_0}{r_{\text{det}}}\, e^{-i(\omega t - k r_{\text{det}})} \int_{V_{\text{sample}}} \rho(\boldsymbol{r})\, e^{-i\boldsymbol{q}\boldsymbol{r}}\, \mathrm{d}V \tag{1.1}$$

with the *scattering vector* $\boldsymbol{q} := \boldsymbol{k} - \boldsymbol{k}_0$ defining *reciprocal space*.

The constant global phase factor of the incident beam that depends on the choice of $\boldsymbol{r}_0$ relative to the incoming wavefront can be neglected since it does not affect the intensity at the detector.

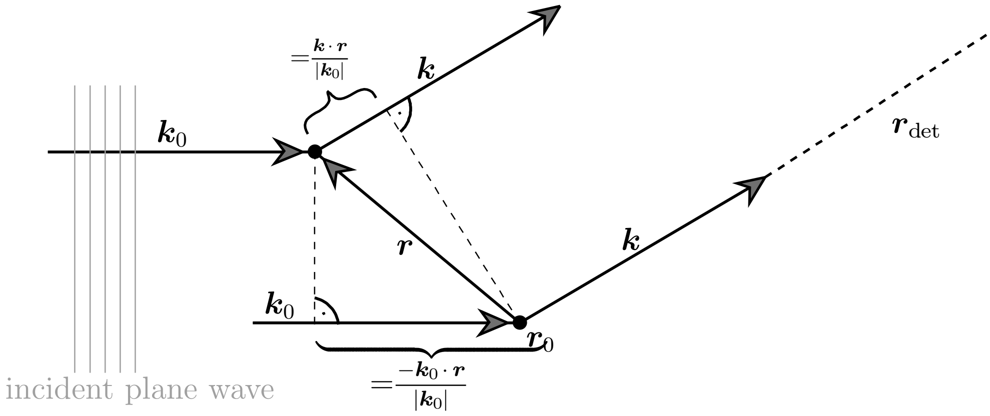


Fig. 1.2. path difference between incident wave $\boldsymbol{k_0}$ and diffracted wave $\boldsymbol{k}$. Figure adapted from [23]

In summary, the amplitude of the wave detected as a function of the scattering vector $\boldsymbol{q}$ is proportional to the Fourier transform of the electron density $\rho(\boldsymbol{r})$, as can be seen from eq. 1.1.

The amplitude factors that encode phase and intensity depend on the detector distance and on the incident intensity A_0. For imaging purposes, they are omitted in the following, as their time-average is constant and only results in a constant intensity factor. Only the time independent part of eq 1.1 that contains information about the sample $\rho(\boldsymbol{r})$ is of interest, the so-called *scattering amplitude* $\mathcal{A}$[1.1]:

$$\mathcal{A}(\boldsymbol{q}) \;=\; \int_{V_{\text{sample}}} \rho(\boldsymbol{r})\, e^{-i\boldsymbol{q}\boldsymbol{r}}\, \mathrm{d}V. \tag{1.2}$$

[1.1] $\mathcal{A}(\boldsymbol{q})$ and $\mathcal{I}$ are introduced for simplicity. The factors that are omitted compared to $A(t,\boldsymbol{q})$ and I do not change the concept, they are reflected in the actual measurements, though.

Note that $\mathcal{A} \propto A$. The inverse Fourier transform then takes the form

$$\rho(\boldsymbol{r}) \;=\; \frac{1}{(2\,\pi)^3} \int \mathcal{A}(\boldsymbol{q})\, e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\, \mathrm{d}^3\boldsymbol{q} \tag{1.3}$$

Once the scattering amplitudes $\mathcal{A}(\boldsymbol{q})$ are sampled in three-dimensional *reciprocal space*, the imaging process can be completed by mapping back the diffraction signals to the electron density $\rho(\boldsymbol{r})$ by means of the inverse Fourier transform 1.3, which reveals information about the structure of the sample.

The amplitude $A(t,\, \boldsymbol{q})$ cannot be measured by the pixel array detectors used for CDI which only detect the intensity averaged over time $I(\boldsymbol{q}) = \langle I(t, \boldsymbol{q})\rangle \propto |A(0, \boldsymbol{q})|^2 \propto |\mathcal{A}(\boldsymbol{q})|^2$. Again, the proportionality constant that accounts for the intensity of the incident wave and the decrease in signal with increasing detector distance will be omitted, and the intensity $\mathcal{I}$ that is invariant to these conditions will be defined as $\mathcal{I}(\boldsymbol{q}) := |\mathcal{A}(\boldsymbol{q})|^2$. The measured intensity I is proportional to $\mathcal{I}$ (see footnote 1.1). The scattering amplitude is complex-valued. The phase information is lost in the time-averaged detection and has to be determined by other means in order to infer $\mathcal{A}(\boldsymbol{q})$ from $|\mathcal{A}(\boldsymbol{q})|^2$.

Due to the relation of the elastic scattering $|\boldsymbol{k}| = |\boldsymbol{k}_0|$, sampling the scattering amplitudes from a single diffraction pattern constrains the scattering vector $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}_0$ to lie on the two-dimensional surface of a sphere with radius $r = |\boldsymbol{k}_0|$. This sphere is called *Ewald's sphere* (see fig. 1.3). Therefore in order to sample $\mathcal{A}$ completely in the three-dimensional reciprocal space, the imaging process needs to be repeated with different directions of the incident beam relative to the sample, e.g. by rotating the sample and measuring the intensities. Seen from the laboratory frame the Ewald sphere is fixed by the incident X-ray beam while the orientation of the scattering volume is given by the sample orientation. Thus different sample orientations sample the scattering amplitude at the resulting intersections between the Ewald sphere and the scattering volume, resulting in a three-dimensional sampling of $\mathcal{I}(\boldsymbol{q})$.

If the orientations of all diffraction patterns are known, it is straightforward to assemble the three-dimensional *diffraction volume* $\mathcal{I}(\boldsymbol{q})$. Averaging the diffraction patterns of similar orientations increases the signal to noise ratio. Next, the phase information needs to be recovered and combined with $\mathcal{I}(\boldsymbol{q})$ to obtain the amplitudes $\mathcal{A}(\boldsymbol{q})$ and eventually it is possible to use eq. 1.3 to obtain the structure of the sample object.

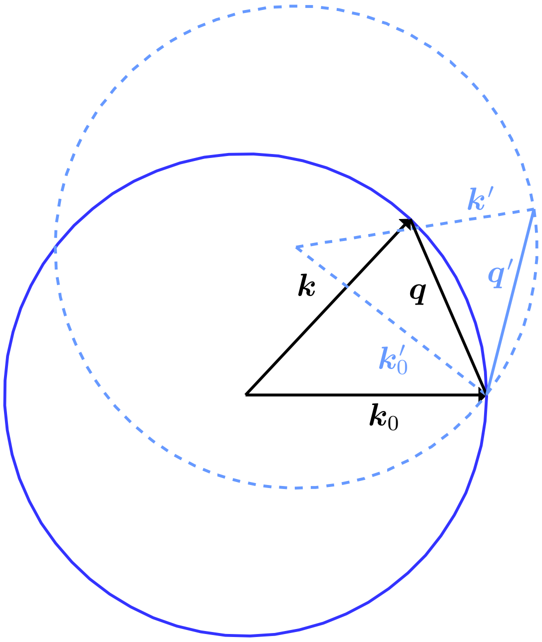


Fig. 1.3. Ewald sphere: Due to the elastic nature of the scattering, the scattering vector $\boldsymbol{q}$ is constrained to the *Ewald sphere*, as shown here in a coordinate frame that is rotated along with the sample object. The full three-dimensional scattering amplitude can be measured by sampling reciprocal space ($\boldsymbol{q}$) with different incident beam orientations $\boldsymbol{k}_0'$ relative to the object, leading to rotated Ewald spheres, as suggested by the dashed lines. All rotated Ewald spheres intersect at the origin of reciprocal space $\boldsymbol{q}=0$.

1.1 Lensless imaging: the phase problem

The phase problem arises because the time-averaging nature of the pixel array detectors like CCDs prevents the phase information from being measured. The missing phase information complicates the determination of the electron density distribution ρ because, without the phase information, eq. 1.3 can not be used to invert the diffraction pattern formation process. Essentially the information content is cut in half when the complex amplitudes $\mathcal{A}$ are transformed into real valued intensities $\mathcal{I}=\mathcal{A}^*\mathcal{A}$. In fact, the loss in information is reflected in a symmetry: The electron density ρ was introduced as a real function that describes the probability $\rho\,\mathrm{d}V$ to find an electron in the volume $\mathrm{d}V$. It can be generalized to a complex function with the imaginary part describing absorption. Without absorption, $\rho^*=\rho$ and it can be seen from eq. 1.2 that $\mathcal{A}^*(\boldsymbol{q})=\mathcal{A}(-\boldsymbol{q})$, therefore

$$\mathcal{I}(-\boldsymbol{q}) \;=\; \mathcal{A}^*(-\boldsymbol{q})\,\mathcal{A}(-q)$$

$$
\begin{aligned}
&= \mathcal{A}(\boldsymbol{q})\,\mathcal{A}(-\boldsymbol{q})\\
&= \mathcal{A}(-\boldsymbol{q})\,\mathcal{A}(\boldsymbol{q})\\
&= \mathcal{I}(\boldsymbol{q})
\end{aligned}
$$

This result is known as *Friedel's law* [24] which states that, in the absence of absorption, the intensity is centrosymmetric. For very short wavelengths, at which the Ewald sphere can be approximated by a plane, this symmetry is visible in individual diffraction patterns. The symmetry introduces an ambiguity that is directly related to the phase problem.

Given that the strength of the diffraction signal at a detector pixel is determined by the interference of all elementary waves originating in the sample and that the interference is determined by the relative path differences of the waves, it is not surprising that the diffraction pattern contains information about relative distances in the electron density. One way to eliminate the absolute position dependence of the electron density is to introduce relative distances x and integrate over the correlations at all absolute positions ξ, which leads to the *autocorrelation function* $\rho \star \rho$ of the electron density ρ. Here, the *cross-correlation function* $f \star g$ is defined as

$$
f(x) \star g(x) \;:=\; \int_{\mathbb{R}} f(\xi)^{*}\, g(\xi + x)\,\mathrm{d}\xi.
$$

For clarity the problem is restricted to one dimension which can easily be generalized to three dimensions. Consider the Fourier transform of the autocorrelation of the electron density:

$$
\begin{aligned}
\mathfrak{F}(\rho \star \rho) &= \int_{\mathbb{R}} e^{-iqx} \int_{\mathbb{R}} \rho^{*}(\xi) \star \rho(\xi + x)\,\mathrm{d}\xi\,\mathrm{d}x\\
&= \int_{\mathbb{R}} e^{(\xi - z)iq} \int_{\mathbb{R}} \rho^{*}(\xi) \star \rho(z)\,\mathrm{d}\xi\,\mathrm{d}z \quad \text{with } z := \xi + x\\
&= \int_{\mathbb{R}} \rho^{*}(\xi) e^{iq\xi} \int_{\mathbb{R}} \rho(z)\, e^{-iqz}\,\mathrm{d}z\,\mathrm{d}\xi\\
&= \int_{\mathbb{R}} \rho^{*}(\xi)\, e^{iq\xi}\,\mathrm{d}\xi \int_{\mathbb{R}} \rho(z)\, e^{-iqz}\,\mathrm{d}z\\
&= \mathcal{A}^{*} \cdot \mathcal{A}\\
&= |\mathcal{A}(q)|^{2}\\
&= \mathcal{I}(q)
\end{aligned} \tag{1.4}
$$

Here, *Fubini's theorem* has been applied, which allows the double integral to be separated. Generalized to three dimensions and with the Fourier transform inverted, this shows that the autocorrelation of the electron density is equal to the inverse

Fourier transform of the intensity (see [25, 26]):

$$\rho(\boldsymbol{x}) \star \rho(\boldsymbol{x}) \;=\; \mathcal{F}^{-1}(\mathcal{I}(\boldsymbol{q})). \tag{1.5}$$

This means that relative information about the structure of the sample object can be revealed, even without measuring the phases of the scattered waves at the detector. Figure 1.4 elucidates the geometric interpretation of the autocorrelation function and explains that the *support* $\mathrm{supp}(\rho \star \rho)$

$$\mathrm{supp}(f) := \{x \in \mathbb{R}^3 | f(x) \neq 0\} \qquad f\colon \mathbb{R}^3 \to \mathbb{C}$$

of the autocorrelation covers twice the volume of the sample, in other words its extent $\mathrm{extend}(\mathrm{supp}(\rho \star \rho))$ is twice the size of the sample. Here, the extent is defined as the supremum of distances within a set:

$$\mathrm{extent}(X) := \sup_{x,y \in X} \|x - y\| \qquad X \subseteq \mathbb{R}^3$$

Fig. 1.4. Two-dimensional autocorrelation (right) of the source image (left). Because the relative relations from one point to another point can also be found with reversed direction, the support of the autocorrelation covers twice the area of the support of the input signal.

1.1.1 Phase retrieval

A suitable way of looking at the phase problem is provided by the *Nyquist-Shannon sampling theorem* [27] which is usually formulated in terms of signal processing: if a signal is band limited in the frequency domain with B being the highest frequency, then it is completely determined by equispaced samples that are $\frac{1}{2B}$ apart in the time domain (when the sampling frequency is twice the maximum frequency). Sayre [16] realized that this has implications for the sampling of diffraction patterns: the domains of interest in coherent diffractive imaging are not time and frequency, but the real space positions of scatterers on one hand and the spatial frequencies of the object which give rise to the diffraction pattern on the other hand. Sample objects usually have a finite size d and therefore the electron density ρ is “band limited” in the real space domain, with $\mathrm{extent}(\mathrm{supp}(\rho)) = d$. Then the full diffraction information can be obtained by sampling reciprocal space at the points $0, \pm\frac{1}{d}, \pm\frac{2}{d}, \ldots$ which can be understood in the following way:

For a real-valued signal $f(x)$ and a real-valued sampling function $g(x)$, the convolution $f \circ g$

$$f \circ g := \int_{\mathbb{R}} f(\xi)\, g(\xi - x) \mathrm{d}\xi$$

is the same as the correlation $f \star g(-x)$ and, following equation 1.4, the *convolution theorem* $\mathcal{F}(f \circ g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$ can be found [28, 26]. With the intensity $\mathcal{I}$ as the signal and with a sampling function that consists of a grid of evenly spaced delta peaks with a distance h, resembling the pixel array of a detector, the measurement of the diffraction signal is a multiplication of the sampling function and the intensity (see figure 1.5). In Fourier space, this means a convolution of the intensity spectrum with the Fourier transform of the sampling grid, which is still a grid after the Fourier transform has been made. If the spacing $1/h$ of the convolution grid is larger than (or equal to) the size d of the support of the intensity spectrum, then the convolution does not lead to overlaps and the intensity spectrum can be separated (or deconvoluted) by multiplying with a step function (see figure 1.5). The Fourier transform of this step function gives rise to an interpolation function $\mathrm{sinc}(x)$ and it suffices to sample the diffraction pattern at points $0, \pm\frac{1}{d}, \pm\frac{2}{d}, \ldots$ to gain the maximum extractable information about the object when the phases ϕ are known. Without noise, oversampling of a factor of 2 provides enough additional constraints to recover the phase information [29]. This increases the measured “bandwidth” of ρ. One can think of that as including the empty space around the object into the measurement. Prior knowledge about the sample can be used to constrain the position and size of this empty space, and constraints like this along with oversampling can be used to solve the phase problem which is otherwise under-constrained.

In crystallography, oversampling would require measurements between Bragg reflections. For infinitely-extended crystals (and, in a first approximation, macroscopic crystals used in crystallography), however, the scattered intensity is rigorously zero between the Bragg reflections (which become delta-functions), and phase retrieval by over-sampling is precluded. The diffraction intensities of samples with finite size are not constrained to Bragg reflections but are distributed in a continuous fashion, facilitating oversampling (eq. 1.7 describes how object size and oversampling are related).

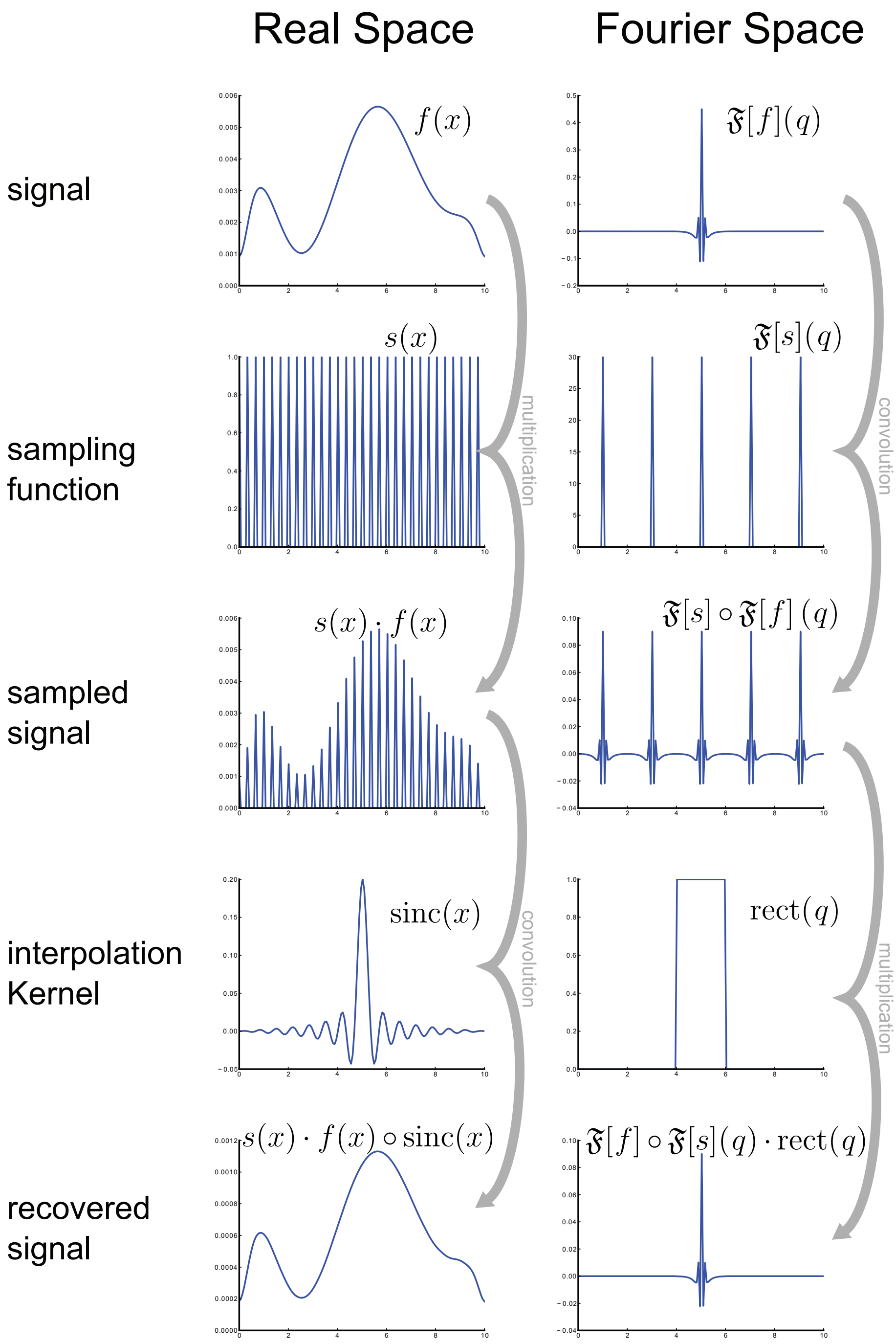


Fig. 1.5. Sampling theorem: a band limited signal $f(x)$ can be fully recovered if the sampling rate is higher than the Nyquist rate - in this case the replicated spectra in the Fourier transform of the sampled signal do not overlap and they can be separated to reconstruct the original signal.. Figure adapted from [30]. Only the real part (cosine waves) is plotted.

The complex scattering amplitude $\mathcal{A}$ can be decomposed into amplitude $|\mathcal{A}|$ and phase $\phi(\boldsymbol{q})$ in the following way:

$$\mathcal{A}(\boldsymbol{q}) \;=\; |\mathcal{A}(\boldsymbol{q})| \cdot e^{i\phi(\boldsymbol{q})} \tag{1.6}$$

(since $|e^{ix}| = 1$). Prior knowledge puts constraints on the phase ϕ and these constraints can be used in an optimization approach to find the function ϕ which agrees best with the measurements and constraints. The most popular constraint is the *support constraint*. It is a real space constraint which makes assumptions about the support $\mathrm{supp}(\rho(\boldsymbol{r}))$. As explained in fig. 1.4, the autocorrelation can be used to find a loose constraint of the support, but also a rough size estimate of the sample obtained by other means can be used. A unified evaluation of phase retrieval algorithms can be found in [31].

1.1.2 Iterative phase retrieval algorithms

Image reconstruction by phase retrieval can be described as the search for an image that satisfies both the constraints of the measurement and the constraints imposed by prior knowledge. According to eq. 1.6, the scattering amplitude can be decomposed into amplitude $|\mathcal{A}|$ and phase. The amplitude is constrained by the measurement of the intensities by $I \sim |\mathcal{A}|^2$. Due to eq. 1.3, the search for the image of ρ is equivalent to the search for the scattering amplitude $\mathcal{A}$. Since $|e^{i\phi}| = 1$, the intensity constraint is fulfilled by filling in the intensities into $\mathcal{A} = |\mathcal{A}|\, e^{i\phi} = \sqrt{I}\, e^{i\phi}$ and the search is limited to the phase ϕ which satisfies the additional constraint imposed by prior knowledge, such as the aforementioned support constraint. It has been shown [32] that the result of this search is unique given that the oversampling is sufficient. Since noise is always present in real diffraction patterns, as well as other impairments caused by the detection, this search comes down to an optimization: the search for the phases that fulfill the constraints best. The search space for this optimization is huge, too large for an exhaustive search. Therefore, iterative algorithms [33, 34] have been proposed to find a solution in acceptable time.

The following is a description of the error reduction algorithm [32] which makes use of the support constraint:

- Initial phases ϕ_0 are chosen and, after combination with the amplitudes $|\mathcal{A}|$ (and as such ensuring compliance with the intensity constraint), the resulting preliminary scattering amplitude $\mathcal{A}_0$ is inversely Fourier transformed into real space.

- Here, the real-space constraint can be enforced, in this case by forcing the values of the preliminary image ρ_0 to be zero outside of the assumed support.
- Taking the Fourier transform takes the image ρ_0 back to scattering amplitudes $\mathcal{A}_1$, which do not necessarily fulfill the intensity constraint anymore.
- The iterative enforcement of the intensity constraint and real space constraint eventually converges to an image ρ_N which fulfills both constraints and therefore constitutes a valid reconstruction, or the convergence ends in a local minimum which cannot be escaped in further iterations.

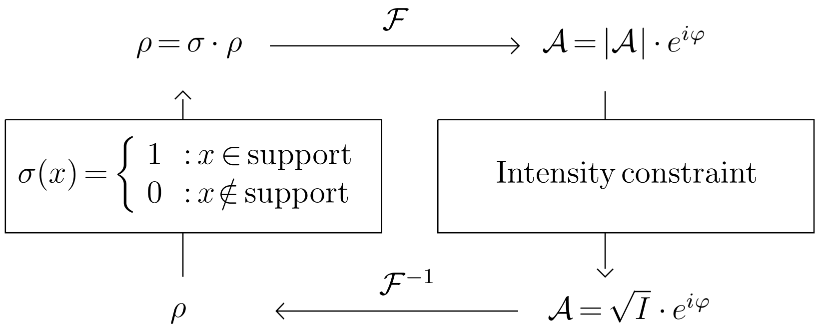


Fig. 1.6. iterative phase retrieval: error reduction algorithm with support and modulus constraint

The use of iterative phase retrieval algorithms requires the diffraction patterns to be sampled finer than the Nyquist frequency. The missing phase information cuts the information content in half and thus an two-fold oversampling suffices in theory [29], but in practice it can not be guaranteed that the phase retrieval algorithm finds the correct phases. Increasing the oversampling ratio increases the chances that the iterative phase retrieval converges to the correct solution. An oversampling ratio >10 per dimension has been used to obtain the diffraction data presented here.

The Nyquist frequency corresponds to the size of the Shannon pixel on the detector (see section 1.3) while the sampling frequency corresponds to the pixel size p on the detector. Thus the oversampling ratio σ is determined by the object size d, the detector distance L and the wavelength λ:

$$\sigma \;=\; \frac{\lambda\, L}{d \cdot p} \tag{1.7}$$

In real space, this ratio can be seen as the maximum sample volume that can be described by a Fourier synthesis of all discrete frequencies represented by detector pixels (this volume is governed by the lowest spatial frequency) divided by the volume of the sample object. Increasing the detector distance or decreasing the wavelength increases the oversampling ratio but decreases the maximum possible resolution.

Another factor that determines the chance to obtain the correct solution from iterative phase retrieval methods is the location of missing information within the detector region. A significant part in the center of the diffraction pattern can not be recorded because it is overlayed by the direct beam which has to be blocked. The affected pixels correspond to low spatial frequencies which are important in determining the coarse structure of the sample. Simulations in [35, 36] suggest that 10 missing *speckles* can be tolerated. In diffraction patterns, speckles are connected regions of high intensity surrounded by low intensity, the size s of speckles $s = \frac{2\pi}{d}$ is determined by the sample size d.

1.2 The orientation problem

As described earlier, the parts of the diffraction volume that can be measured are constrained to lie the two-dimensional surface of the Ewald sphere and thus the information content of a diffraction pattern is only two-dimensional. In the considerations that led to eq. 1.2, the incident beam direction $\boldsymbol{k}_0$ served as a reference orientation and a look at fig. 1.2 reveals that rotating the incident beam results in a equally rotated diffraction signal. The same happens when the sample object is rotated. Mathematically, since both a rotation $\boldsymbol{R}$ and the Fourier transform are linear operations and since the integration of the Fourier transform covers the whole space and as such is independent of rotations which preserve length and volume, rotations and the Fourier transform commute:

$$\begin{aligned}
\mathcal{F}[\rho(\boldsymbol{R}\boldsymbol{x})](\boldsymbol{q}) &= \int \rho(\boldsymbol{R}\boldsymbol{x})\, e^{-2\pi i \boldsymbol{q}\boldsymbol{x}} \mathrm{d}V \\
&= \int \rho(\boldsymbol{z})\, e^{-2\pi i \boldsymbol{q}\boldsymbol{R}^{-1}\boldsymbol{z}} \mathrm{d}V \quad \text{with } z := x = \boldsymbol{R}^{-1} x \\
&= \int \rho(\boldsymbol{z})\, e^{-2\pi i \boldsymbol{R}\boldsymbol{q}\boldsymbol{z}} \mathrm{d}V \quad \begin{array}{l}\text{because } \boldsymbol{R} \text{ is an isometry and} \\ \text{preserves the dot product}\end{array} \\
&= \mathcal{F}[f(\boldsymbol{x})](\boldsymbol{R}\boldsymbol{q}).
\end{aligned}$$

Generalizing the rotation $\boldsymbol{R}$ to an operator with $\boldsymbol{R}f := f(\boldsymbol{R}\,\boldsymbol{x})$, the commutation of $\boldsymbol{R}$ and $\mathcal{F}$ can be stated as

$$\boldsymbol{R}\mathcal{F}(\rho) = \mathcal{F}(\boldsymbol{R}\rho). \tag{1.8}$$

Therefore, rotations in reciprocal diffraction space and rotations in object space can be treated equally and a rotation of the sample object results in an equal rotation of the scattering amplitudes $\mathcal{A}$ in reciprocal space. A diffraction snapshot samples the scattering amplitudes on the surface of the Ewald sphere which is fixed in the laboratory frame. Rotating the sample object rotates the scattering amplitudes $\mathcal{A}$ through this sphere and hence $\mathcal{A}$ can be sampled three-dimensionally in reciprocal space. Consequently diffraction snapshots of different orientations can be combined into a three-dimensional diffraction volume. This merge can be carried out if the relative orientations of the sample objects between the different snapshots are known. Because of the equivalence of rotations in real space and diffraction space, it is sufficient to know the relative orientations between the diffraction patterns.

Intuitively, the formation of a single diffraction pattern can be recognized as a projection from the three-dimensional object to a two-dimensional diffraction pattern. And indeed, in the flat Ewald sphere approximation this projection turns out to be the orthographic projection of the sample object along the X-ray beam, as the following considerations show:

Provided the wavelength λ is small compared to the distance of the detector, the radius of the Ewald sphere $|\boldsymbol{k}_0| = \frac{2\pi}{\lambda}$ is large and the detected part of the Ewald sphere becomes almost flat. In this approximation, a diffraction pattern is a planar slice through the three-dimensional diffraction volume $\mathcal{I}(\boldsymbol{q})$ instead of a spherical slice. For clarity, consider only the case of a two-dimensional diffraction volume $\mathcal{I}(q_x, q_y)$ originating from the two-dimensional electron density $\rho(x, y)$ and a linear slice $s(\boldsymbol{q})$ thereof. Without loss of generality, the slice can be described in a coordinate system which is aligned to the slice:

$$\begin{aligned} s(\boldsymbol{q}) &= \mathcal{A}(q_x, q_y = 0) \\ &= \iint \rho(x, y)\, e^{-i\left(xq_x + \underbrace{yq_y}_{=0}\right)} \mathrm{d}x\, \mathrm{d}y \end{aligned}$$

$$
\begin{aligned}
&= \int \left[\underbrace{\int \rho(x,y)\mathrm{d}y}_{:=p(x)} \right] e^{-ixq_x}\,\mathrm{d}x \\
&= \mathcal{F}(p(x))
\end{aligned}
$$

This result is known as the *projection-slice theorem* [25]. It can be seen that the function $p(x)$ is a projection of the electron density along the y-direction and, after generalizing to three dimensions, we can note that a slice through the diffraction volume is the same as the diffraction of the projection of the electron density on the slicing plane. In other words, every diffraction pattern contains information about the projection of the sample along the X-ray beam. The three-dimensional image can be assembled from two-dimensional projections of different orientation if their relative orientations are known. In the general case where the Ewald sphere is not flat, the principle remains the same although the "curved" projections do not have a simple geometric interpretation.

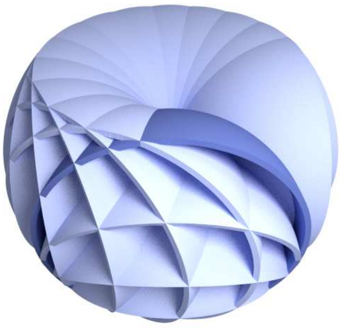


Fig. 1.7. 3D Merge: The diffraction Intensities sampled at multiple Ewald spheres are combined into the 3D diffraction volume

1.3 Resolution and the number of required diffraction snapshots

The number of diffraction snapshots that are required for a given resolution can be obtained from the Nyquist-Shannon sampling theorem (see 1.1), which states that for an object of size d the full diffraction information can be obtained by sampling reciprocal space at the points $0, \pm\frac{1}{d}, \pm\frac{2}{d}, \ldots$ This leads to the "Shannon pixel size" of $\mathrm{sp} = \frac{L\lambda}{d}$ of a detector placed at distance L when the diffracted photons have the wavelength λ.

To resolve sample object features down to the size r, the reciprocal space has to be measured in the range $[-1/r, 1/r]$ along each dimension. For the two dimensions that can be covered by a single diffraction pattern, the detector pixels have to be smaller than (or equal to) the Shannon pixel to obey the sampling theorem. For the third dimension, namely that out of the detector plane, the sampling theorem constrains the orientational sampling. At the reciprocal detector edge $1/r$, the closest pixel along the third dimension lies on another diffraction pattern. This pixel can not be further away than $1/d$ without disobeying the sampling theorem. Therefore, the angle θ_{min} between closest (in an angular sense, see figure 1.8) diffraction patterns has to obey $\theta_{\text{min}} \approx \tan\,(\theta_{\text{min}}) < \frac{1/d}{1/r} = \frac{r}{d}$ in order to fulfill the sampling condition. The following considerations are based on [37]. Assuming uniformly distributed orientations that can be represented as points on the hypersphere S^3, the average volume that a snapshot orientation takes on the three-dimensional "surface" of S^3 is $\left(\frac{\theta}{2}\right)^3$ because every point on S^3 represents the same orientation as its antipode (inverting the angle is the same as inverting the rotation axis). The full area of S^3 is $2\,\pi^2$ but again, because of the antipodes, the area of all orientations is only π^2. Consequently, the minimum number of diffraction patterns that are necessary to obtain the resolution r is (see [37])

$$N \;=\; \frac{8\,\pi^2\,d^3}{r^3\,S}.$$

If the sample object is symmetric, this number is reduced by the cardinality S of the symmetry group.

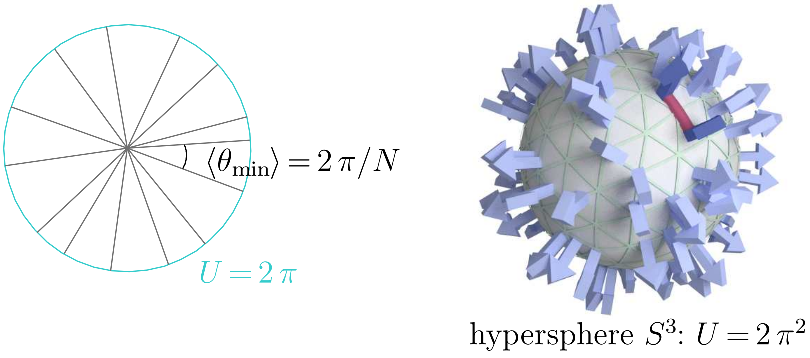


Fig. 1.8. Mean nearest neighbor angular distance between two diffraction snapshots used to estimate the number of required snapshots. Drawn are the orientations of the sample object (**left**: one degree of freedom, **right**: three degrees of freedom, a pair of closest neighbors is highlighted in dark blue, θ_{min} is visualized in red). Only the 2-sphere S^2 can be visualized. The third dimension is encoded in the orientation of the blue arrows. The curved tetrahedral grid is projected onto a curved triangular grid.

2

Digital image analysis and pattern recognition

Three-dimensional coherent diffractive imaging is based on the measurement of many thousands of discrete diffraction snapshots, namely two-dimensional diffraction patterns. This chapter borrows ideas from the fields of computer vision and pattern recognition to categorize and manage these patterns in an automated way.

2.1 Classification

In order to obtain enough diffraction snapshots of good quality to obtain a three-dimensional image of the sample object with the desired resolution, data acquisition is performed in a serial manner with subsequent X-ray pulses that intercept a stream of identical sample objects. It is inevitable to also record snapshots that are either undesired or have to be treated in a special way, see fig 4 for examples. These “special” snapshots fall into different classes:

i. **empty snapshots** where the X-ray pulse did not coincide with the presence of a sample object in the interaction region
ii. ”**multiple hits**” where an X-ray pulses coincides with the presence of more than one particle in the interaction zone
iii. **clusters** or **big aggregates** of the sample object
iv. **alien sample objects** from impure sample preparations
v. in case of aerosol injection: **empty droplets** of the solvent in which the sample objects are immersed before aerosolization

These different classes of diffraction snapshots make it desirable to find a way to automatically sort the data before continuing with the data analysis. Automating this task is possible by borrowing ideas from *statistical classification* which is a vivid field in statistics and machine learning.

2.2 Feature extraction

In principle, each pixel of a diffraction snapshot can be seen as an independent variable that can help to classify the snapshot. Thus, an obvious solution to the classification would be to feed all pixel values to a statistical classification algorithm. In practice however, performing the classification in such a high-dimensional space is problematic: With increasing dimensionality, individual observations get more and more isolated unless a vast amount of observations are added. This is due to the rapid increase of volume with increased dimensionality and is known as the *curse of dimensionality*. It is therefore worthwhile to base the classification on a small number of explanatory variables, called *features*, that carry enough decisive information. Such features can be extracted from the diffraction snapshots. The following sections describe different types of features that have turned out to be a good choice for the classification of diffraction patterns:

- section 2.2.1: Intensity variations (Viola-Jones Features [38])

  The spatial distribution of intensity variations in single-particle diffraction patterns contains information on the size of the sample specimen
- section 2.2.2: Maxima of the rotational auto-correlation function and maxima of the angular distribution function

  Extracts information about rotational symmetries
- section 2.2.3: "Eigenpatterns", the *score* (see below) of eigenvectors obtained from PCA (*Principal Components Analysis*) [39]

  Eigenpatterns are special directions in the high-dimensional vector space of diffraction snapshots, the directions that have the largest variance and thus carry a large amount of information. For example, Eigenpatterns disentangle signal from background signal. Also, variations in the signal strength are encoded in separate Eigenpatterns.

In detail, these features are described in the following sections.

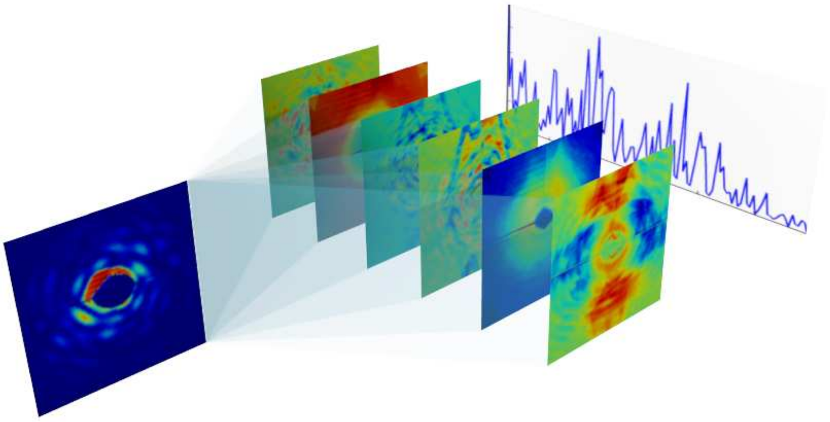


Fig. 2.1. Feature analysis of a diffraction pattern. This illustration shows a diffraction pattern decomposed into the first six principal components of a diffraction dataset. Their scores (graph in the background) are used to extract descriptors containing discriminative information from the diffraction pattern, reducing the dimensionality of the data that is fed into statistical classifiers.

2.2.1 Intensity variations

The idea is to capture the *speckle* sizes (see section 1.7) and their spatial distribution by analyzing intensity variations within the diffraction patterns. Large objects will lead to smaller speckles in the diffraction pattern than smaller objects, while the spatial distribution is determined by the internal structure of the sample objects. Extracted from the diffraction patterns, this information will provide features that can be used for classification.

The implementation is based on the successful face recognition scheme developed by Paul Viola and Michael Jones in 2001 [38]. In the original formulation, differences in brightness between facial regions like eyes, nose and mouth are estimated by calculating the difference of integrated intensities of rectangular templates. In the context of diffraction patterns the significant regions with different intensities correspond to diffraction speckles. These features have the advantage that their computation requires only a few look-ups in the *integral image* $I(x,y) = \sum_{x' \leqslant x, y' \leqslant y} p(x', y')$ which can be precomputed in a single pass over the diffraction pattern p and reused for several differently shaped or sized Viola-Jones features. Instead of integrating the rectangular patches, the computation can be performed by just adding/subtracting the corners of the rectangular regions in the corresponding integral image (see fig.

2.2):

$$\sum_{\substack{x_1 \leqslant x \leqslant x_2 \\ y_1 \leqslant y \leqslant y_2}} p(x,y) = I(x_2,y_2) - I(x_1,y_1)$$

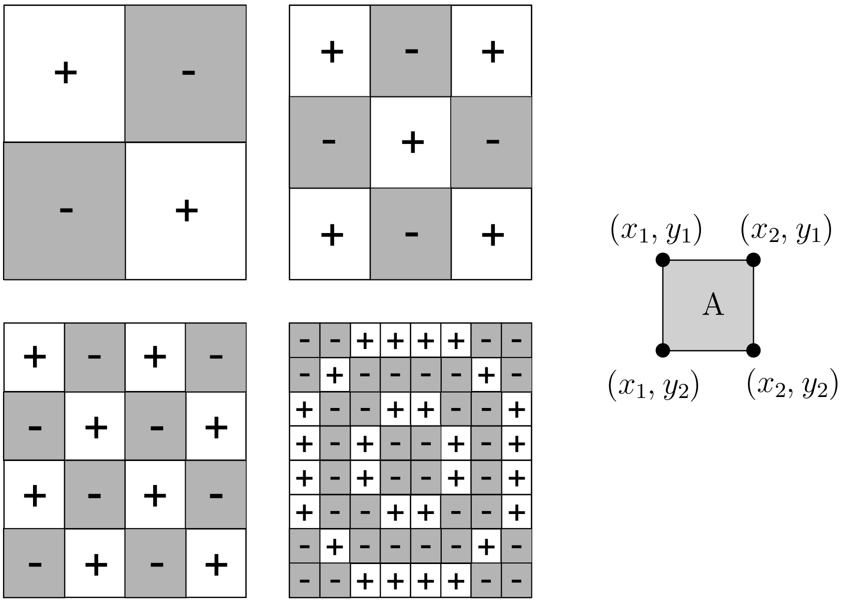


Fig. 2.2. Illustration of different Viola-Jones Features. The integrals over the intensities in the rectangular regions can be looked up in the integral image I: $A = I(x_2, y_2) - I(x_1, y_1)$.

2.2.2 Rotation symmetry

Since rotational symmetry can often be found in nature and some of the samples that were used for this thesis display icosahedral and dihedral symmetries which are reflected in symmetries of the diffraction patterns, an analysis of the two-dimensional rotational symmetries reveals discriminative information that can be used for classification purposes. As a fuzzy measure for symmetries, the positions of the first few maxima of the angular autocorrelation function are used.

2.2.3 “Eigenpatterns”

The goal of feature extraction is to combine the discriminative information of many pixels into a few numbers that build the basis for classification. This can be done using prior information about the structure of the information that is hidden in the pixels, such as the size and angle based features described above, or it can be done by statistically analyzing the relations between pixels and learning about their relations from examples. Here, this is performed by forming groups of correlated pixels and treating them as basis vectors. When a single diffraction pattern is represented in such a basis, the corresponding coefficients can be interpreted as features that carry discriminative information. It turns out that a linear analysis of inter-pixel correlations is sufficient for this task and therefore the well-established *Principal Component Analysis* (PCA) [39] can be used.

The following matrix representation of n snapshots consisting of m pixels will be assumed:

$$\boldsymbol{X} \in \mathbb{R}^{n\times m} \quad \text{with } \boldsymbol{X}_{i,j} = \begin{array}{l}\text{intensity measured}\\ \text{at the } j^{\text{th}} \text{ pixel in the}\\ i^{\text{th}} \text{ snapshot}\end{array}$$

The mean μ_j of the j^{th} pixel is defined as $\mu_j := \frac{1}{n}\sum_i \boldsymbol{X}_{i,j}$ and the *centered* data matrix is the data matrix with the mean μ_j subtracted from the j^{th} column: $\tilde{\boldsymbol{X}}_{i,j} := \boldsymbol{X}_{i,j} - \mu_j$.

In this representation, the covariance between the a^{th} and the b^{th} pixel is given by

$$\begin{aligned}\text{cov}(a,b) &= \frac{1}{n-1}\sum_{i=0}^{n}(\boldsymbol{X}_{i,a}-\mu_a)\cdot(\boldsymbol{X}_{i,b}-\mu_b)\\ &= \frac{1}{n-1}\sum_{i=0}^{n}\tilde{\boldsymbol{X}}_{i,a}\cdot\tilde{\boldsymbol{X}}_{i,b}\\ &= \frac{1}{n-1}\sum_{i=0}^{n}\tilde{\boldsymbol{X}}^T_{a,i}\cdot\tilde{\boldsymbol{X}}_{i,b}\end{aligned}$$

which means that the covariance matrix $\boldsymbol{C}$ with $\boldsymbol{C}_{a,b} := \text{cov}(a,b)$ can be written in terms of a matrix multiplication:

$$\boldsymbol{C} = \frac{1}{n-1}\tilde{\boldsymbol{X}}^T\cdot\tilde{\boldsymbol{X}}$$

The covariance matrix is symmetric and can thus be diagonalized. The coordinate system in which $\boldsymbol{C}$ is a diagonal matrix consists of the eigenvectors of $\boldsymbol{C}$ which have no crossed variances: $\boldsymbol{C}_{\text{diag}\,i,j} = 0$ for $i \neq j$. The eigenvalues $\boldsymbol{C}_{\text{diag}\,i,i}$ of the covariance matrix reflect the degree of "variance" that the corresponding i^{th} eigenvector encodes. A large variance along an eigenvector means that many different diffraction patterns influence this eigenvector. The variance is therefore a measure of the importance of that particular eigenvector and we can use it to reduce the whole training data set to just a few very discriminative coefficients: the coefficients of the eigenvectors with the highest eigenvalues.

In the case where the number of pixels $m > n$ is larger than the number of snapshots, the eigenvalue problem can be simplified since only a few eigenvectors with the highest eigenvalues are of interest:

let $\boldsymbol{v}$ be the eigenvector and λ be the eigenvalue that solves the eigenvalue problem

$$\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\boldsymbol{v} = \lambda\boldsymbol{v}$$

then a left multiplication with $\tilde{\boldsymbol{X}}^T$ yields

$$\tilde{\boldsymbol{X}}^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T = \lambda\tilde{\boldsymbol{X}}^T\boldsymbol{v}$$

which means that $\tilde{\boldsymbol{X}}^T\boldsymbol{v}$ is an eigenvector of the original eigenvalue problem for $\tilde{\boldsymbol{X}}^T\tilde{\boldsymbol{X}}$. Since $\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T$ is an $n \times n$ matrix, the eigenvalue problem is easier to solve than for the larger $m \times m$ matrix $\tilde{\boldsymbol{X}}^T\tilde{\boldsymbol{X}}$.

It turns out that for real data collected with pnCCD detectors at the LCLS (see chapter 5), the first 120 eigenvectors with the highest eigenvalues capture enough information for a successful classification. Fig. 2.3 shows a diffraction pattern in the severely compressed PCA representation. The original basis consists of 1048576 base vectors (each representing a single detector pixel). In the basis of just 120 *eigenpatterns*, the structure of the signal is preserved and the diffraction fringes are still visible. For classification, the *scores*, the components of the reduced basis of eigenpatterns, are used as features.

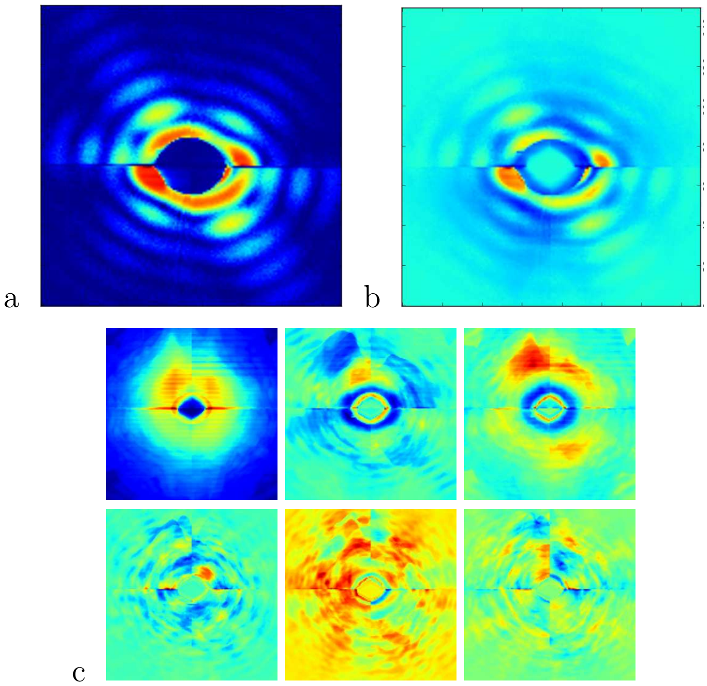


Fig. 2.3. **(a)** Detail of a diffraction pattern (sample: Cro-virus). **(b)** the same diffraction pattern expressed in a basis of only the first 120 principal components. **(c)** the six most important principal components or *eigenpatterns* for the entire training dataset of 1000 diffraction images. Note that these diffraction patterns consist of two halves separated by a horizontal gap which has been removed in this representation.

2.3 Supervised classification

The features described in section 2.2 provide enough discriminative information for sorting and classifying diffraction patterns. The classification process has to be automated in a reliable and flexible way. A supervised classification scheme was chosen in which the system learns the optimal classification based on a training data set that consists of diffraction images with known classification results. This eliminates the need to carefully optimize the parameters of the classification system every time the experimental conditions change. If parameters like the wavelength, detector settings or detector geometry change, all that needs to be done is to provide the system with a valid training data set. Of course this is also true when switching to a different sample object. The training data set can easily be created manually, defining the classes based on visual inspection of the diffraction patterns. It turns out that a training set with the surprisingly low number of ∼20 diffraction patterns per class is enough to create accurate classifiers that are able to sort thousands of diffraction patterns automatically.

2.3.1 Partitioning the *feature space*

The features introduced in section 2.2 constitute a *feature space*. A general approach to the classification problem is to find a partitioning of the feature space into disjoint regions. The class affiliation of a given feature vector can then be predicted based on what region it falls into. In the case of supervised classifications, the partitioning can be found by looking at the training data set where for each feature vector, a valid class label is known. A decision boundary can be represented by a binary space partitioning; the result is a *decision tree*. Note that the resulting decision tree can perfectly resemble the class affiliations of the training data set provided the depth of the tree is large enough. Perfect resemblance may seem like a valid goal, but in the presence of noise there is a risk of over-fitting. In case of over-fitting, the decision tree provides insufficient generalization to new data that is not part of the training data set because a particular instance of noise in the training set is included into the model and the added complexity can amplify small variations in the data.

Assuming a simple noise model for the training data $t_i = f(x_i) + \varepsilon$, $i = 1, \ldots, N$ with f being the true classification function and a statistical error ε with $E\{\varepsilon\} = 0$, the *mean squared error* (MSE) of the classification model $y(x)$ that approximates

$f(x)$ can be estimated by validating the predictions of the training data:

$$\text{MSE} \;=\; \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2$$

The error MSE states how well the classification model can explain a specific realization of the training data. Since the training data underlies randomness, so does the MSE and a better estimate of the classification error is the expectation value

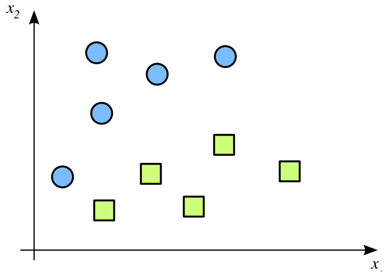
$$E\{\text{MSE}\} \;=\; \frac{1}{N}\sum_{i=1}^{N}E\{(t_i - y_i)^2\}$$

The analysis of the expectation value of just one test datum is enough to establish a decomposition of the MSE into different error classes:

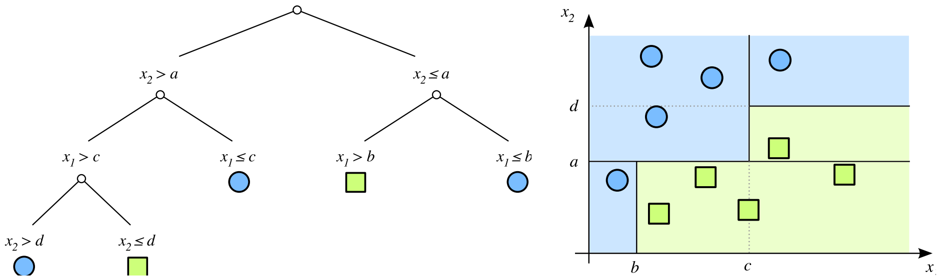
$$\begin{aligned}
E\{(t_i - y_i)^2\} &= E\{(t_i - f_i + f_i - y_i)^2\} \\
&= E\{(t_i - f_i)^2\} + E\{(f_i - y_i)^2\} + 2\,E\{(t_i - f_i)\,(f_i - y_i)\} \\
&= E\{\varepsilon^2\} + E\{(f_i - y_i)^2\} + 2\,(\underbrace{E\{t_i\,f_i\}}_{=f_i^2} + E\{f_i\,y_i\} - \underbrace{E\{f_i^2\}}_{=f_i^2} - \\
&\quad E\{t_i\,y_i\}) \\
&= E\{\varepsilon^2\} + E\{(f_i - y_i)^2\} + 2\,(E\{f_i\,y_i\} - E\{(f_i + \varepsilon)\,y_i\}) \\
&= E\{\varepsilon^2\} + E\{(f_i - y_i)^2\} + 2\,(E\{f_i\,y_i\} - E\{f_i\,y_i\} + \underbrace{E\{\varepsilon\,y_i\}}_{=0}) \\
&= E\{\varepsilon^2\} + E\{(f_i - y_i)^2\} \\
&= E\{\varepsilon^2\} + E\{(f_i - E\{y_i\} + E\{y_i\} - y_i)^2\} \\
&= E\{\varepsilon^2\} + \underbrace{E\{(f_i - E\{y_i\})^2\}}_{:=\text{bias}^2} + \underbrace{E\{(E\{y_i\} - y_i)^2\}}_{:=\text{Var}(y_i)} + 2\,E\{(E\{y_i\} - \\
&\quad y_i)\,(f_i - E\{y_i\})\} \\
&= E\{\varepsilon^2\} \quad + \quad \text{bias}^2 \quad + \quad \text{Var}(y_i) \quad + \\
&\quad \underbrace{2\,(f_i\,E\{y_i\} - E\{y_i\}^2 - f_i\,E\{y_i\} + E\{y_i\}^2)}_{=0} \\
&= E\{\varepsilon^2\} + \text{bias}^2 + \text{Var}(y_i)
\end{aligned}$$

Thus, the classification error can be decomposed into the variance of the noise, the model *bias* and the *variance*. The *bias* describes the error introduced by a systematic classification error due to an unfit model. The *variance* describes the dependence of the prediction on a given realization of the training data. A low bias means that on average the predictions are accurate, a low variance means that the predictions do not change much as the training set varies.

The variance of the error $E\{\varepsilon^2\}$ can not be reduced, but it is clear that in order to construct an optimal classifier, both bias and variance have to be minimized. However, they cannot be minimized independently as a simple example shows: a classifier that always gives the same prediction (ignoring the features) has perfect variance $\mathrm{Var} = 0$. On the other hand the bias of this classifier would be tremendously large. Another classifier that perfectly resembles the training data would lead to a vanishing bias term for the training data set; in contrast the variance will become equal to the variance of the noise, this classifier suffers from *over-fitting*.



(a) samples of different classes (blue and green) in two-dimensional feature space



(b) decision tree (c) partitioned feature space

Fig. 2.4. Illustration of the feature space partitioning by means of decision trees. This example shows a two-dimensional feature space and several data points belonging to two different classes (blue and green). The partitioning is performed by dividing the feature space recursively in a binary fashion along the feature dimensions x_1 and x_2, resulting in decision boundaries a, b, c, d. This method can be easily generalized to higher dimensions as well as more classes.

2.3.2 Random forest classifier

The simultaneous minimization of *bias* and *variance* can be accomplished by *ensemble learning*. Here, the classification is not based on a single model, but on an ensemble of classifiers whose results can be combined, e.g. by a weighted sum. This is an improvement over single classifiers like the decision tree described above, because the risk of learning a wrong model is distributed over several hypothesis and therefore alleviated. At the same time, combining different classifiers adds more flexibility to the model that is used to fit the data and therefore the bias can be reduced. See [40] for an empirical study.

The ensemble learning method "random forest" developed by Leo Breiman and Adele Cutler [41] was used as the main classifier in this work. This method uses an ensemble of *decision trees*, and injects two sources of randomness into each tree to introduce diversification in the ensemble of trees, hence the name random forest. The first source of randomness is that in each node of the tree, the feature-dimension that defines that node's decision hyperplane is chosen from a random subset of features instead of searching for the best split among all features. This limits the generalization ability of a single tree and therefore makes the average of all trees less prone to over-fitting. The second source of randomness is that the training data set for each tree is limited to a random subset of the whole available training data. This limits the variance error term described above because the variance of the individual trees is averaged at the end. Moreover, the available training data that is not used for a given tree can be used to estimate the generalization error of that tree ("*out of bag error*") and therefore the generalization can be optimized.

In this work the generic computer vision library VIGRA [42] was used, which provides an implementation of the random forest algorithm.

With this tools at hand, the enormous amount of diffraction patterns can be presorted. As will be shown in chapter 5.1.5, false positives can be tolerated by the following orientation recovery step, but the efficiency of the presorting is very important since the orientation recovery treats the diffraction dataset as a whole ensemble and diffraction patterns can not be treated independently, resulting in an unfavorable non-linear scaling of the computational complexity. The statistical classification however treats diffraction patterns independently from each other, resulting in a computational complexity that scales linearly with the number of diffraction patterns. Reducing the number of unusable diffraction patterns that are fed into the orientation recovery stage represents an efficient way to reduce the overall computing time. Only a split second is needed for the classification of each diffraction pattern and in addition the tasks can easily be parallelized and distributed among different computers.

Part II

Results

3

Geodesic orientation recovery

The assembly of many diffraction patterns with different orientation into a three-dimensional scattering volume is straightforward once the relative angles between the diffraction patterns are known. Finding a method that recovers the orientation solely based on the diffraction signals is beneficial for the versatility of this imaging scheme. This chapter describes a solution to this problem that has been developed as the main result of this thesis.

3.1 Establishing and interpreting similarities among diffraction patterns

The approach to the orientation problem that is proposed here is to assemble many local angular estimates into a robust global angular measure while reducing the three-dimensional orientation problem to many one-dimensional sub-problems. Therefore, pairs of measured diffraction patterns are compared and a global similarity measure is found by global optimization over all such pairs. The underlying principle is that objects of incrementally different orientation will yield diffraction patterns that also differ only incrementally. Given a large enough number of measured diffraction patterns, it becomes possible - via pairwise comparisons - to arrange the entire ensemble of measured diffraction patterns in what might be called *similarity sequences*. In orientation space, such sequences can be interpreted intuitively:

the most direct transition between two different orientations is a rotation about one single rotation axis. In other words, a *geodesic* on the special orthogonal group SO(3) with respect to the angular distance measure $d_\angle$ is a rotation about a single axis, generally called the *Euler axis*. Relating these geodesics to the aforementioned similarity sequences allows this specific axis to be identified and thereby the angular relationships between the 2D diffraction patterns can be established.

The formation of the diffraction pattern can be seen as a mapping from the space of object orientations to the space of diffraction patterns. According to eq. 1.2, this mapping is described by the Fourier transform

$$\mathcal{A}(\boldsymbol{q}, \boldsymbol{R}) \;=\; \int_{V_\text{sample}} \rho(\boldsymbol{R}\,\boldsymbol{r})\, e^{-i\boldsymbol{q}\boldsymbol{r}}\, \mathrm{d}V$$

where $\mathbf{R}$ is the rotation of the sample object. Using relation eq. 1.8 this becomes the rotation $\boldsymbol{R}\,\mathcal{A}(\boldsymbol{q})$ of $\mathcal{A}(\boldsymbol{q}, \mathbb{1}) = \mathcal{A}(\boldsymbol{q})$. This means that the mapping from orientations to diffraction patterns is continuous in the sense that a slight rotation of the object induces only a correspondingly slight change in the diffraction pattern. Consequently, a measure $d_\text{diss}(P_1,\, P_2)$ of the “dissimilarity” between two different diffraction patterns P_1 and P_2 provides a local measure of the angular separation between the two corresponding objects. One can think of many methods to capture the dissimilarity, the simplest would be the Euclidean distance between the vectors containing the pixel values as components (so each detector pixel is a basis vector). Here the *Pearson correlation coefficient* is used to estimate the dissimilarity d_diss, because this coefficient is invariant to offset (varying background signal) and scale (varying signal strength). In order to extend this local estimate of angular distances (slight changes in diffraction patterns corresponding to slight object rotations) to global quantities (arbitrary changes in diffraction patterns corresponding to possibly large rotations), the *geodesic dissimilarity* d_geo between two patterns P_1 and P_2 is defined to be the shortest accumulated dissimilarity of all possible sequences $\{\gamma_k\}$ of diffraction patterns starting from P_1 and ending at P_2:

$$d_\text{geo}(P_1, P_2) \;=\; \min_{\gamma \in \Gamma(P_1, P_2)} \sum_{i=0}^{|\gamma|-1} d_\text{diss}(\gamma_i,\, \gamma_{i+1}).$$

To cope with the local nature of d_{diss}, the optimization is limited to a subset Γ of sequences with similar consecutive elements based on a threshold ε:

$$\{\Gamma(P_1,P_2)=\{\gamma_k\}_{k=0}^{N}|(\gamma_0=P_1)\wedge(\gamma_N=P_2)\wedge(d_{\mathrm{diss}}(\gamma_i,\gamma_{i+1})<\varepsilon\forall i=0,N-1\}.$$

The optimal sequence $\gamma(P_1,\, P_2)$ is named the *geodesic sequence* or *shortest path* between P_1 and P_2. Because of the close connection between rotations in real space and rotations in diffraction space, it is useful to associate the shortest similarity path between two diffraction patterns with a geodesic trajectory that can be interpreted as a rotation about an Euler axis. This correspondence is exact if the true angular distance $d_\angle$ is used, which is unknown in general. Since d_{diss} locally (for small angles) correlates with $d_\angle$ it can be used to approximate $d_\angle$ locally, which is all that is needed in order to find geodesics. In this thesis, the validity of this approximation has been tested by numerical simulations (see figures 3.1 and 3.4). It turns out that the correlation between d_{diss} and $d_\angle$ is strong enough to find geodesics for a wide range of object shapes and quite generally the "shortest" similarity sequence that smoothly connects two diffraction patterns (i.e., the path through diffraction-pattern space having minimal cumulative dissimilarity) corresponds to the smallest real space rotation of the object.

This mapping of similarity geodesics onto real space geodesics must be considered very carefully. In general, the distance measure defined by the dissimilarity can be distorted by any anisotropy of the experimental geometry or of the object itself and thus deviates from the round metric of S^3. One obvious anisotropy, as pointed out in Ref. [43], is that due to the unidirectional nature of the x-ray beam which, as a first approximation, results in a projection. However, this distortion is easily treated by simply distinguishing rotations about the x-ray beam axis (which will be termed *in-plane rotations*) from rotations about axes orthogonal to the x-ray beam axis (*out-of-plane rotations*, see below). The consequences of anisotropies in object shape are difficult to characterize in a general fashion. However, in simulations with parameters typical for coherent diffraction imaging experiments and reasonable sample object shapes (finite size and thickness), the effect of distortions due to object shape is negligible. A result of such simulations is shown in figs. 3.1 and 3.4. It can be seen that even under severe distortion the geodesics do not deviate much from great circles on S^3 and thus they correspond to single axis rotations. Accordingly, this mapping of similarity geodesics (within diffraction patterns) onto real space geodesics (of object orientation) can be considered to be generally valid and in the following, either will be referred to as simply a “geodesic”.

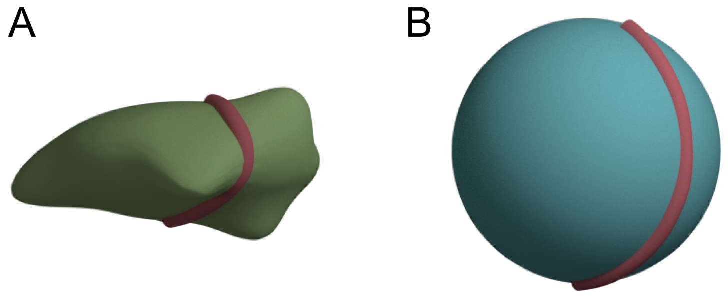


Fig. 3.1. Simulating the effect of distortions on geodesics. The similarity metric of diffraction snapshots can be distorted, deviating from the round metric induced by S^3. (**A**) shows a 3D slice through the heavily distorted three-sphere S^3 (only its projection S^2 can be shown here) and a geodesic line (red). (**B**) shows that even under heavy geometric distortion the geodesic connecting two poles is approximately preserved in a topological sense, i.e., it is a great circle when mapped back to S^3. Figure taken from [44].

The topological information on geodesic sequences can be translated into geometric information on the orientations by comparison with distinct angles such as the maximum geodesic angle. Provided that the number of diffraction patterns is sufficient to approximate a complete sampling of the orientation space, the longest geodesic sequence found in the data corresponds to the maximum geodesic angular separation, which depends on the sample symmetry (180° for asymmetric objects). The object symmetry can be assessed from the diffraction patterns, assisted by the observation that the geodesic sequences end on symmetry poles, since beyond these the diffraction patterns increasingly resemble the starting diffraction pattern.

3.2 Identifying in-plane and out-of-plane rotations and combining them to span the orientation space

Two steps are required to successfully recover the orientations of all collected diffraction patterns of given sample objects via geodesic analysis: First, each pattern has to be assigned to a geodesic sequence and then the relations between their respective Euler axes need to be established. The former involves an optimization that can be carried out efficiently by dynamic programming algorithms like Dijkstra's shortest path algorithm [45] while the latter can be realized by adding another source of angular information with the aid of *in-plane* rotations. Typical setups for diffraction experiments are symmetric with respect to the x-ray beam axis. Due to this symmetry, rotating the specimen about the x-ray beam axis corresponds to a rotation of the diffraction pattern in the detector plane by that same angle (see fig. 3.2). Starting from a diffraction pattern P we can therefore identify or generate a "synthetic" diffraction pattern $P(\alpha)$ that is rotated in-plane through an angle α. Given two diffraction patterns P_1 and P_2 whose orientations are related by the Euler axis $\vec{\boldsymbol{E}}_{0,0}$, introducing synthetic in-plane rotations $P_1 \rightarrow P_1(\alpha)$ and $P_2 \rightarrow P_2(\beta)$ will lead to the Euler axis $\vec{\boldsymbol{E}}_{\alpha,\beta}$ of a geodesic sequence $\gamma(P_1(\alpha), P_2(\beta))$. The fraction of diffraction patterns that can be assigned to such sequences depends on the angular separation θ and on the orientation of the Euler axis $\vec{\boldsymbol{E}}_{0,0}$ relative to the X-ray axis $\vec{\boldsymbol{c}}$ (the in-plane axis). The second dependence can be understood as follows: in-plane rotations are also geodesic rotations and in the extreme case where $\vec{\boldsymbol{E}}_{0,0} = \vec{\boldsymbol{c}}$, there is no difference between in-plane rotations and the geodesic rotation that rotates P_1 to P_2 and thus the in-plane rotations do not provide additional information. To maximize the information gain that can be obtained from in-plane rotations the overlap between in-plane rotations and the geodesic rotation has to be minimized by choosing P_2 such that the geodesic rotation that rotates P_1 to P_2 is orthogonal to the in-plane rotations. These orthogonal rotations will be named *out-of-plane* rotations. This suggests that the maximum number of diffraction patterns can be assigned to geodesic sequences $\gamma(P_1(\alpha), P_2(b))$ if P_1 and P_2 are separated by the maximum geodesic angle $\theta = 180°$ and if the corresponding Euler axis $\vec{\boldsymbol{E}}_{0,0}$ is orthogonal to the X-ray axis $\vec{\boldsymbol{c}}$. In fact, as shown below (also see formal proof in appendix C.1), and illustrated in Figs. 3.2(b) and 3.2(c), this covers all diffraction patterns and all orientations. In-plane and out-of-plane rotations can be orthogonalized by artificially setting $d_{\mathrm{diss}}(P_1, P_1(\alpha))$ to zero for all values of α. Then, initial in-plane rotations will be preferred in the search for the shortest path because they are cost-free. In this way, the in-plane component is only contained in the selection step between P_1 and the next diffraction pattern in the geodesic sequence and can easily be removed, leading to out-of-plane geodesic sequences which are orthogonal to the in-plane rotations.

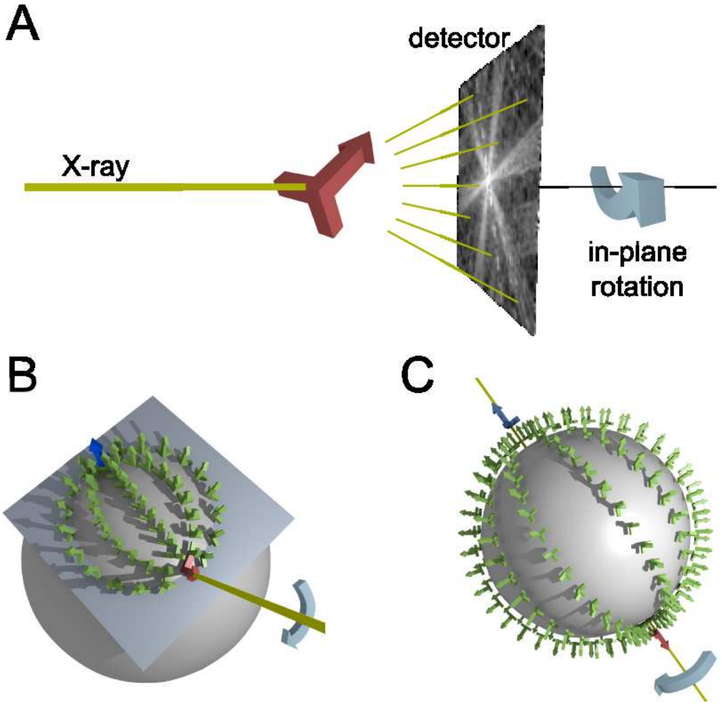


Fig. 3.2. Illustration of geodesic and in-plane rotations. (A) Geometry of diffraction experiment. The triad (red) denotes the object and its orientation. In-plane rotations correspond to a rotation of both object and diffraction pattern around the X-ray axis. (B+C) Illustration of combined geodesic and in-plane rotations. For clarity, only the red object $O(P_1)$ corresponding to diffraction pattern P_1 is rotated in-plane around the X-ray beam (yellow line). The geodesic sequences connecting the orientations of each in-plane rotation of $O(P_1)$ (red) to $O(P_2)$ (blue) are shown in light green. (B) In-plane rotations of $O(P_1)$ with arbitrary orientation (see www.gipral.org for an interactive illustration). (C) Maximum separation between the object orientations $O(P_1)$ and $O(P_2)$ leads to full coverage of SO(3). The orientations of the red and the blue arrows constitute the poles on S^3. Note that only its projection S^2 can be shown here. Figure taken from [44].

In order to calculate the effect of combining in-plane and out-of-plane rotations, consider a diffraction pattern $P_{\alpha,\beta,\varphi}$ which is part of the geodesic sequence $\gamma(P_1(\alpha), P_2(\beta))$ with the geodesic angle φ (see Fig. 3.3). The following is a derivation of the orientation of $P_{\alpha,\beta,\varphi}$ relative to the orientation of P_1 under the condition that P_2 is related to P_1 by a true out-of-plane rotation (without in-plane components) through the angle θ. The coordinate system is chosen such that the y axis coincides with the rotation axis $\vec{\boldsymbol{a}}$ of the out-of-plane rotations and the z axis coincides with the X-ray axis $\vec{\boldsymbol{c}}$ (= in-plane axis). Orientations are described as rotations of a reference orientation so they can be expressed in quaternions

$$\boldsymbol{q}_{\vec{\boldsymbol{e}},\vartheta} = \begin{pmatrix} \sin(\vartheta/2)\cdot\vec{\boldsymbol{e}} \\ \cos(\vartheta/2) \end{pmatrix},$$

where $\left(\vec{\boldsymbol{e}}, \vartheta\right)$ is the Euler axis/angle representation. In this orientation representation, the unit quaternion $\boldsymbol{q} = \mathbb{1}$ stands for the reference orientation which, for convenience, is defined as the orientation of the diffraction pattern P_1.

$$\begin{tikzcd}
P_2 \arrow[rr, "\boldsymbol{R}(\vec{\boldsymbol{c}},\beta)"] & & P_2(\beta) \\
& & P_{\alpha,\beta,\varphi} \arrow[u, dashed] \\
P_1 \arrow[uu, "\boldsymbol{R}(\vec{\boldsymbol{a}},\theta)"'] \arrow[rru, "\boldsymbol{q}_{P_{\alpha,\beta,\varphi}}"] \arrow[rr, "\boldsymbol{R}(\vec{\boldsymbol{c}},\alpha)"] & & P_1(\alpha) \arrow[u, dashed, "\boldsymbol{R}(\vec{\boldsymbol{e}}_{\boldsymbol{\Delta}},\varphi)"']
\end{tikzcd}$$

Fig. 3.3. Relations between diffraction patterns and rotation operators $\boldsymbol{R}$. Vertical operators describe geodesic out-of-plane operations (rotations about $\vec{\boldsymbol{a}}$), horizontal operators describe in-plane operations (rotations about $\vec{\boldsymbol{c}}$). Note that in the special case $\theta = 180°$ the dashed arrows are out-of-plane rotations, too.

Thus the orientation of $P_1(\alpha)$ is the in-plane rotated reference orientation and can be written as

$$\boldsymbol{q}_{P_1(\alpha)} = \begin{pmatrix} 0 \\ 0 \\ \sin(\alpha/2) \\ \cos(\alpha/2) \end{pmatrix}.$$

The orientation of P_2 is related to that of P_1 by an out-of-plane rotation through the angle θ and can be written as

$$\boldsymbol{q}_{P_2} = \begin{pmatrix} 0 \\ \sin(\theta/2) \\ 0 \\ \cos(\theta/2) \end{pmatrix}.$$

The orientation of $P_2(\beta)$ can be obtained by adding an in-plane rotation through the angle β to $\boldsymbol{q}_{P_2}$:

$$\begin{aligned} \boldsymbol{q}_{P_2(\beta)} &= \begin{pmatrix} 0 \\ 0 \\ \sin(\beta/2) \\ \cos(\beta/2) \end{pmatrix} \otimes \boldsymbol{q}_{P_2} \\ &= \begin{pmatrix} -\sin(\theta/2)\sin(\beta/2) \\ \sin(\theta/2)\cos(\beta/2) \\ \cos(\theta/2)\sin(\beta/2) \\ \cos(\theta/2)\cos(\beta/2) \end{pmatrix}. \end{aligned}$$

The geodesic rotation $\boldsymbol{q}_{\boldsymbol{\Delta}}$ from $P_1(\alpha)$ to $P_2(\beta)$ is given by

$$\boldsymbol{q}_{\boldsymbol{\Delta}} = \boldsymbol{q}_{P_2(\beta)} \otimes \boldsymbol{q}_{P_1(\alpha)}^{-1} = \begin{pmatrix} -\sin(\theta/2)\sin(\beta/2+\alpha/2) \\ \sin(\theta/2)\cos(\beta/2+\alpha/2) \\ \cos(\theta/2)\sin(\beta/2-\alpha/2) \\ \cos(\theta/2)\cos(\beta/2-\alpha/2) \end{pmatrix}.$$

The rotation axis $\vec{\boldsymbol{e}}_{\boldsymbol{\Delta}}$ of all rotations that are part of the geodesic connecting $P_1(\alpha)$ and $P_2(\beta)$ can be extracted from the vector part of $\boldsymbol{q}_{\boldsymbol{\Delta}}$:

$$\vec{\boldsymbol{e}}_{\boldsymbol{\Delta}} = \frac{1}{|\vec{\boldsymbol{q}}_{\boldsymbol{\Delta}}|} \cdot \begin{pmatrix} -\sin(\theta/2)\sin(\beta/2+\alpha/2) \\ \sin(\theta/2)\cos(\beta/2+\alpha/2) \\ \cos(\theta/2)\sin(\beta/2-\alpha/2) \end{pmatrix}$$

with

$$|\vec{\boldsymbol{q}}_{\boldsymbol{\Delta}}| = \sqrt{\sin^2\left(\frac{\theta}{2}\right)\sin^2\left(\frac{\beta}{2}+\frac{\alpha}{2}\right)+\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\beta}{2}+\frac{\alpha}{2}\right)+\cos^2\left(\frac{\theta}{2}\right)\sin^2\left(\frac{\beta}{2}-\frac{\alpha}{2}\right)}.$$

After introducing the angle φ to parametrize rotations along this geodesic the orientations of the geodesic sequence are given by

$$\boldsymbol{q}_{P_{\alpha,\beta,\varphi}} = \begin{pmatrix} \sin(\varphi/2)\cdot\vec{\boldsymbol{e}}_{\boldsymbol{\Delta}} \\ \cos(\varphi/2) \end{pmatrix} \otimes \boldsymbol{q}_{P_1(\alpha)}$$

and thus

$$\boldsymbol{q}_{P_{\alpha,\beta,\varphi}} = \begin{pmatrix} -\frac{1}{|\vec{\boldsymbol{q}_\Delta}|}\sin\left(\frac{\varphi}{2}\right)\sin\left(\frac{\theta}{2}\right)\sin\left(\frac{\beta}{2}\right) \\ \frac{1}{|\vec{\boldsymbol{q}_\Delta}|}\sin\left(\frac{\varphi}{2}\right)\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\beta}{2}\right) \\ \sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\varphi}{2}\right)+\frac{1}{|\vec{\boldsymbol{q}_\Delta}|}\cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\beta_j}{2}-\frac{\alpha_i}{2}\right) \\ \cos\left(\frac{\alpha}{2}\right)\cos\left(\frac{\varphi}{2}\right)-\frac{1}{|\overrightarrow{\boldsymbol{g}_\Delta}|}\sin\left(\frac{\alpha}{2}\right)\sin\left(\frac{\varphi}{2}\right)\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\beta_j}{2}-\frac{\alpha_i}{2}\right) \end{pmatrix}. \tag{3.1}$$

In the special case $\theta = 180°\quad \Rightarrow \quad |\vec{\boldsymbol{q}_\Delta}| = 1$, and

$$\boldsymbol{q}_{P_{\alpha,\beta,\varphi}}|_{\theta=180°} = \sin\left(\frac{\varphi}{2}\right) \cdot \begin{pmatrix} -\sin\left(\frac{\beta}{2}\right) \\ \cos\left(\frac{\beta}{2}\right) \\ \sin(\alpha/2)/\tan(\varphi/2) \\ \cos(\alpha/2)/\tan(\varphi/2) \end{pmatrix}. \tag{3.2}$$

As can be seen from equation 3.2, in this case the orientations $\boldsymbol{q}_{P_{\alpha,\beta,\varphi}}$ cover the whole orientation space and thus every possible diffraction pattern can be assigned to the sequences $P_{\alpha,\beta,\varphi}$. A proof using the Rodrigues formalism [46] can be found in appendix C.1.

3.3 GIPRAL - an orientation recovery algorithm in ten steps

The description of any orientation in terms of α, β and φ motivates an algorithm for orientation recovery which will be referred to as *geodesic in-plane rotation algorithm* (**GIPRAL**). It can be outlined as follows:

1. Calculate diffraction pattern cross correlations $\tilde{C}(P_1, P_2)$ between all pairs (P_i, P_j) of diffraction patterns. Normalize and invert in order to obtain the dissimilarity measure as follows:

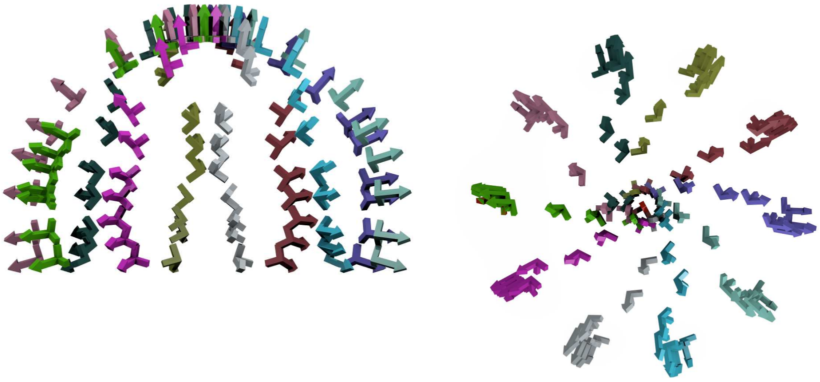
$$\tilde{d}_{\text{diss}}(P_i, P_j) \;=\; 1 - \tilde{C}(P_i, P_j)/\max_{k,l} \tilde{C}(P_k, P_l).$$

2. Threshold nearest neighbors to enforce the local range of validity of $\tilde{d}_{\text{diss}}$:

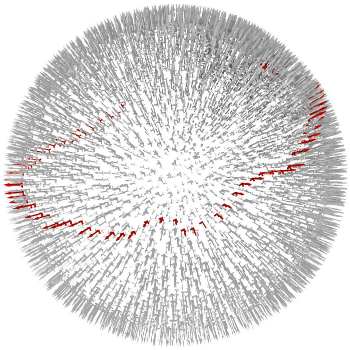
$$d_{\text{diss}}(P_i, P_j) \;=\; \begin{cases} \tilde{d}_{\text{diss}}(P_i, P_j) & : \tilde{d}_{\text{diss}}(P_i, P_j) < \varepsilon \\ \infty & : \text{otherwise} \end{cases}.$$

3. Select the initial diffraction pattern P_1 random or by visual inspection to guarantee that the desired object is chosen (as opposed to a blank shot or a shot containing artifacts as described in Refs. [47, 48] like solvent droplets or clusters of the specimen object, see figure 4).

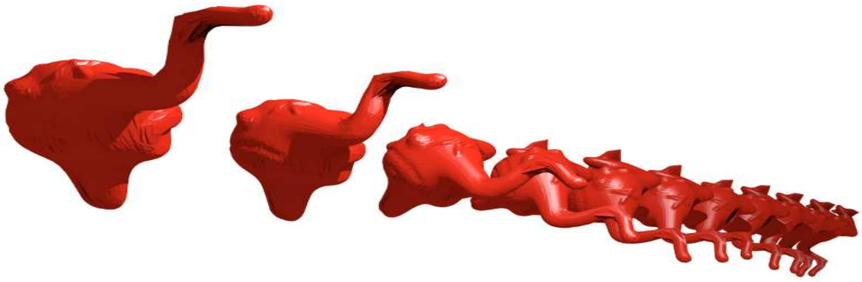
4. Use Dijkstra's algorithm [45] to find the sequences with minimum accumulated dissimilarity from P_1 to every other diffraction pattern that is connected to P_1 (directly or indirectly through a sequence of other diffraction patterns in between, without disobeying ε). The threshold ε should be chosen to be high enough so that all diffraction patterns are indirectly connected to P_1 and low enough so that only very similar diffraction patterns are directly connected to each other. In order to remove in-plane components from the sequence, add in-plane rotated copies $P_1(\alpha)$ of P_1 to the pool of diffraction patterns and set their dissimilarities $d_{\mathrm{diss}}(P_1, P_1(\alpha))$ to zero before running Dijkstra's algorithm.
5. Identify the end pattern P_2 as the one that maximizes $d_{\mathrm{geo}}(P_1,\ P_2)$. Provided that the dataset covers a complete sampling of the orientation space this is the antipode to P_1 on S^3, the three-sphere representing the orientations of SO(3). Because in-plane components have been suppressed in the previous step P_1 and P_2 are related by a true out-of-plane rotation.
6. In case of object symmetries, start again, this time choose the previous P_2 as the new initial image P_1'. This avoids premature termination of the geodesic sequence due to symmetry (see appendix C.4).
7. Generate in-plane rotated diffraction patterns $P_1(\alpha_i)$ of P_1 and $P_2(\beta_j)$ of P_2 and keep track of their in-plane angles α_i and β_j. Add them to the pool of diffraction patterns and repeat step 1 and 2. At this stage it is not necessary to calculate every correlation anew; $d_{\mathrm{diss}}(P_i, P_j)$ has to be updated only for the new diffraction patterns.
8. Similar to step 4 but without in-plane component suppression: Find geodesic sequences between all pairs $(P_1(\alpha_i), P_2(\beta_j))$.
9. Determine the angle $\varphi_{i,j,k}$ between $P_1(\alpha_i)$ and the k^{th} diffraction pattern $P_{i,j,k}$ in the geodesic sequence between $P_1(\alpha_i)$ and $P_2(\beta_j)$ by interpreting the dissimilarity value between P_1 and P_2 as a single axis rotation of 180° (adapt in case of object symmetries).
10. Relate the different rotation axes of different geodesics to each other using the known in-plane angles and $\varphi_{i,j,k}$. The quaternion corresponding to the orientation of diffraction pattern $P_{i,j,k}$ with respect to P_1 is given by Eq. 3.2.



different geodesic sequences corresponding to different in-plane rotations
represented in orientation space
(**left**: side view, **right**: top view along the X-ray beam)



all orientations (ground truth) of the simulated diffraction snapshots,
with a geodesic sequence (based on diffraction patterns) highlighted in red



geodesic sequence in real-space, showing a single-axis rotation of the sample object.

Fig. 3.4. Geodesic sequences of simulated diffraction data. Diffraction snapshots of an object were simulated in 10000 random orientations. Next, diffraction geodesics were found and related to each other by in-plane rotations as outlined above. In the top row, the resulting geodesics are represented based on the ground truth of orientations which are known for the simulated object. It can be seen that the diffraction geodesics (similarity sequences of diffraction patterns) clearly correspond to orientation geodesics (single axis rotations). This result was observed for different object shapes as well.

3.4 Computational complexity

The computational bottleneck of the proposed geodesic algorithm is the computation of pairwise dissimilarities between all pairs of diffraction patterns with a computational complexity of $\mathcal{O}(N^2)$ for N diffraction patterns. This can be sped up by parallelization, since the individual pairs are independent of each other. Moreover, the threshold ε sparsifies the dissimilarity matrix that is needed. Under the assumption that the triangle inequality holds for the estimate of d_{diss}, then a table of upper and lower bounds can be updated for d_{diss} iteratively while adding entries to the dissimilarity matrix. These bounds can guide the computation of the next matrix elements, since elements with a lower bound $>\varepsilon$ can be rejected while elements with a small upper bound will be preferred. This way, only a fraction of all pairs of diffraction patterns needs to be taken into account.

Parallelization and other low level optimization techniques (see appendix B) make the search for geodesics feasible for datasets consisting of ~100.000 diffraction patterns. The computing time for such datasets amounts to 2-3 days. For larger datasets, another idea for optimization - which has not been implemented in this work - is the use of hierarchical search trees in which large numbers of comparisons between pairs of diffraction patterns are avoided on the basis of a few inexpensive features which are only computed once per diffraction pattern as opposed to every possible pair of two diffraction patterns. For example, diffraction patterns with big speckles do not need to be compared to diffraction patterns with small speckles.

3.5 Generalization to symmetric objects

Object symmetries complicate matters since symmetry operators can be applied to any orientation without altering the diffraction pattern dissimilarity. This leads to shortcuts on geodesic paths that act as “wormholes” in orientation space. A portion of the orientation space that consists of symmetrically irreducible orientations and therefore does not contain any wormholes can be constructed by applying symmetry operators to map every orientation to a symmetrically equivalent orientation such that the angular distance to a given reference orientation is minimized. These fundamental zones take very convenient shapes when expressed in Rodrigues-Frank (RF) parametrization [46]: For finite symmetry groups they are polytopes with planar boundaries. Moreover, geodesic paths are straight lines in RF space [46] (see appendix C.1 for an illustration), which makes RF space a natural choice for the formal treatment of rotational geodesics. The maximum angular separation that is possible under a given symmetry can be calculated from the shapes of all finite symmetry group classes [49]. These angles (see Ref. [50] for a complete list) correspond to the longest Euclidean distances between corners of the fundamental zone and can be used to relate geodesic paths of maximum length to angles in the case of symmetric objects. In Appendix C.2, it is shown how this can be used to navigate in the orientation space of symmetric objects and which part of the fundamental zone can be covered. With this extension, the geodesic method can be used to recover the 3D diffraction volume using an ensemble of 2D diffraction snapshots irrespective of the underlying object symmetry.

4

On-line analysis

Current FEL X-ray light sources are used for many different experiments which leaves only limited beam-time for imaging experiments as carried out in the course of this thesis. Thus, most of the experiment has to be set up just for a few days and can not be carefully optimized over a longer period of time in a continuously running experiment. This short-lived nature of the experiments requires perfect preparation. But since almost every aspect of this new imaging method is being developed from scratch, many experimental and technical uncertainties exist which can only be resolved during the experiment, when the X-ray beam is available. Therefore, it is crucial to get instant feedback about the quality of the diffraction data to be able to optimize all accessible parameters.

4.1 On-line hit rate estimation

An example of such feedback is the *hit rate*: Maximal overlap between the injected particle stream and the X-ray beam is required in order to minimize the number of X-ray pulses that do not intercept any sample object which waste precious beam-time and sample material. The injected particle beam or liquid jet containing the sample is not always stable and has to be adjusted constantly. The feedback for the continuous optimization of the injection parameters can be obtained from the diffraction signal. It is helpful to automate the distinction between empty shots and hits in order to estimate the hit rate which can then be maximized. To keep up with the rate of the data acquisition, which currently clocks at 120 Hz, computational speed is crucial. Therefore, using the offline classification scheme described in section

2.1 is not an option. It turns out that a simple median based outlier detection can be used instead. A hit detection system is needed the most in situations where the hit rate is low. Then, empty shots occur frequently enough such that hits can be considered as statistical outliers. An obvious feature space for the hit detection is the total scattered intensity. However it is possible that the magnitude of background fluctuations is in the same range as the intensity of very weak diffraction signals. Thus, additional features need to be considered. As described before (see 2.2.1) the Viola-Jones features can be calculated very efficiently and add useful discriminative information.

After calculating n features for each snapshot, on-line hit detection is performed by updating a running average f_{avg} of the current feature vector for each snapshot. The feature vector $f_{\text{curr}} = \begin{pmatrix} f_{\text{curr}\,1} \\ \vdots \\ f_{\text{curr}\,n} \end{pmatrix}$ of the current snapshot is then compared to the average feature vector by means of the *Mahalanobis distance* [51]

$$d_{\text{Mahalanobis}} = \sqrt{(f_{\text{curr}} - f_{\text{avg}})^T\, \Sigma^{-1}\, (f_{\text{curr}} - f_{\text{avg}})}$$

where Σ is the covariance matrix which contains the covariance between different features $f_{\text{curr}\,i}$, $f_{\text{curr}\,j}$ within the last N snapshots:

$$\Sigma_{i,j} = \operatorname{cov}(f_{\text{curr}\,i}, f_{\text{curr}\,j}) = E[(f_{\text{curr}\,i} - f_{\text{avg}\,i})\,(f_{\text{curr}\,j} - f_{\text{avg}\,j})].$$

Here, E denotes the expected value estimated by calculating the sample mean over N snapshots. The Mahalanobis distance can be thought as a generalization of the Euclidean distance

$$d_{\text{euclidean}} = \sqrt{(f_{\text{curr}} - f_{\text{avg}})^T\, (f_{\text{curr}} - f_{\text{avg}})}$$

which takes into account the correlations of the data set. If a feature varies a lot in a specific direction in the n-dimensional feature space, then the weight of this direction is reduced in the Mahalanobis distance. If, on the other hand, a specific direction is very stable, then its weight is increased. That way, many different features with different statistics can be combined, while their statistical differences are normalized.

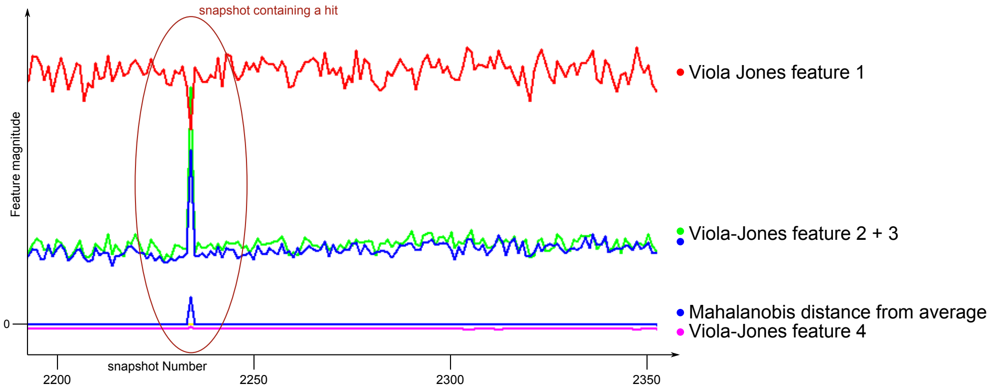


Fig. 4.1. Mahalanobis distance: instead of considering different features individually, the Mahalanobis distance can be used to determine a distance from the average that combines all features, normalizing their covariance. The resulting signal is stable enough to recognize hits with a simple threshold. The plot shows different four feature values as a function of snapshot number as well as the Mahalanobis distance from average in this four-dimensional feature space. A snapshot containing a hit is marked (red ellipse).

As can be seen in fig. 4.1, the peak indicating a hit is more pronounced when multiple features are combined using the Mahalanobis distance. This on-line hit-finding scheme was used successfully in combination with aerosol sample injection the hit-rate of which never exceeded 30%. During the experiment, a graph of the hit-rate was displayed on a monitor next to the person steering the injection system. This proved to be a very useful feedback to optimize the sample delivery.

4.2 On-line size estimation

As will be shown in chapter 5.2, the apparent sample size of biological objects as estimated from diffraction patterns varied strongly, in contrast of what was expected from the biological properties of the samples confirmed by pre-characterization using transmission electron microscopy. This apparent size distribution most likely is an artifact caused by the injection system and needs to be monitored. Continued injection of a sample with huge size instabilities is a waste of beam-time and has to be avoided. In order to estimate the sample size, diffraction fringes of simplified model-objects are fitted to the observed diffraction fringes. A spherical model was used to approximate the biological objects (viruses) with icosahedral shape that were used in this thesis. The fit was performed semi-automatically: per mouse-click, a few positions of a single diffraction fringe were submitted to a least-squares fitting algorithm [52] to obtain the parameters of the circular approximation of the fringe.

From the radius r_{diff} of a spherical diffraction fringe and the diffraction order n_{diff} of the fringe, the radius r_{ob} of the object is given by

$$r_{\mathrm{ob}} = \frac{\mathrm{Jz}_1(n_{\mathrm{diff}})\,\lambda}{4\,\pi\,\sin\left[\frac{1}{2}\mathrm{atan}\frac{r_{\mathrm{diff}}}{r_{\mathrm{det}}}\right]}$$

where r_{det} is the detector distance, $\mathrm{Jz}_1(n)$ is the n-th zero of the Bessel function of the first kind.

Another tool that has been used to analyze the sample size is the autocorrelation function obtained by eq. 1.5. The size can be estimated as half of the extent of the autocorrelation support, see fig. 1.4.

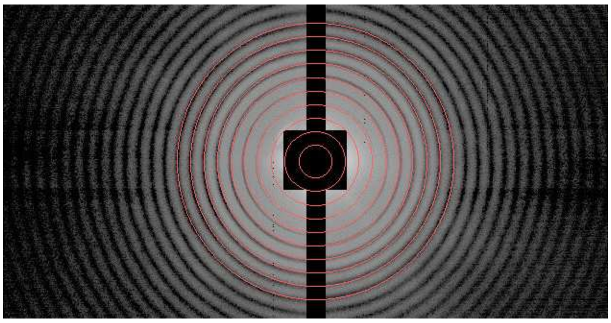
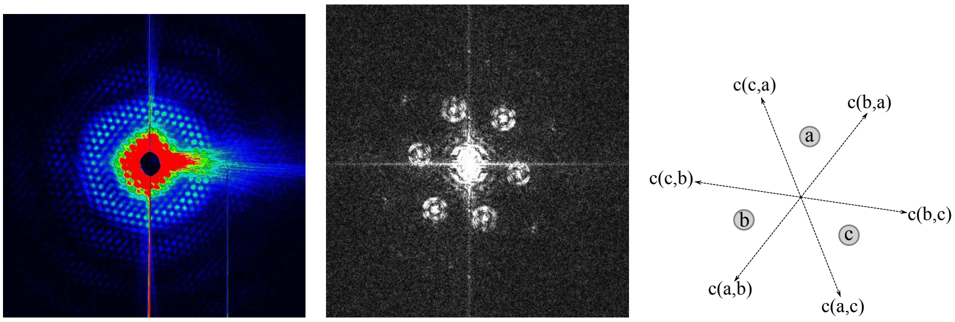


Fig. 4.2. Semi-automatic fit of **airy rings** (red) to the diffraction pattern of a spherical object.

4.3 On-line feedback on sample concentration

For the reconstruction methods described in this thesis, diffraction patterns of *single* sample particles are required. If the concentration of the injected sample is too high, it can happen that two or more particles reside in the interaction region during the arrival of an X-ray pulse. In principle, diffraction patterns of single objects can be extracted from diffraction patterns of multiple object copies by analyzing well-separated cross-correlation terms within the autocorrelation function [53] (see fig. 4.3), but a higher resolution can be achieved when single-particle diffraction patterns are used to begin with. Therefore, the occurrence of multiple-particle diffraction patterns needs to be monitored during the experiment and the sample concentration needs to be reduced should they occur. This was done based on visual inspection of the autocorrelation function, which shows cross-correlation terms once more than one particle is present (see fig. 4.3).



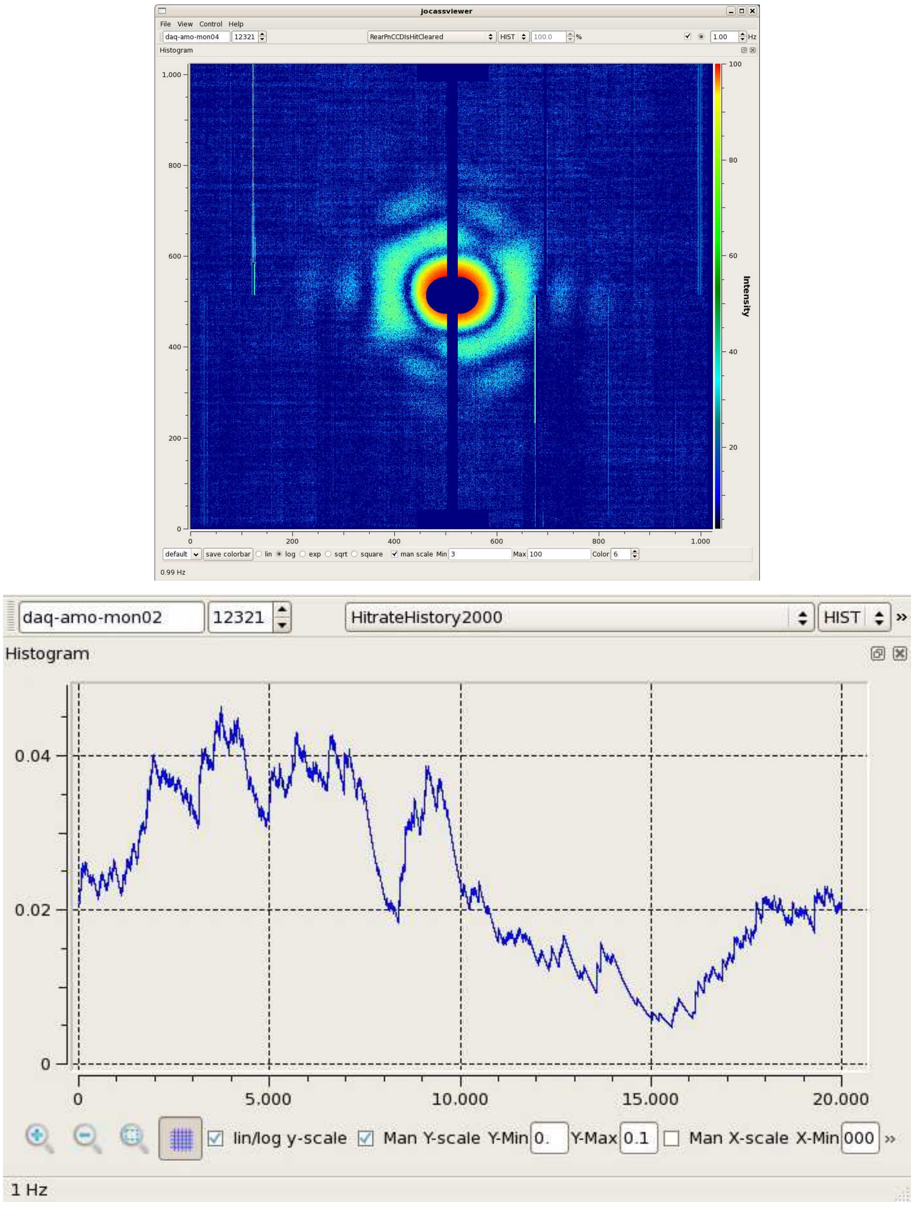
a **b** **c**

Fig. 4.3. Multi particle hit. **(a)**: diffraction pattern of multiple Cro-virus particles injected with liquid jet. **(b,c)**: the autocorrelation shows 6 cross-correlation parts. Since identical particles have been injected, this can be interpreted as cross-correlations of 3 particles, each pair (a,b) results in two cross-correlation terms c(a,b) and c(b,a).

4.4 CASS - a framework for on-line analysis

All on-line analysis tools have been implemented as modular packages for the CFEL ASG Software Suite (CASS) [54]. CASS can easily be customized to hook into the live data streams of FEL experiments carried out at LCLS, SACLA and FLASH. CASS can access the data of different detectors (e.g. CCD pixel array detectors, reaction microscopes) as well as beam-line data (e.g. photon Energy, pulse length). It is a very modular system that can easily be adapted to work with other data sources as well. Computation tasks are split into elementary units and a dependency graph is used to manage the order in which they are computed. This way, the elementary units can be sent to different worker threads, resulting in a huge speed-up by parallelization. Analysis tasks can be defined in a scripting language so that modifications to the on-line analysis chains can be performed very fast, without the need of recompiling software. In the course of this thesis CASS was extended to read and write different detector data-streams (frms6: pnCCD [55], cbf: PILATUS [56, 57], MarCCD [58], xtc: CSPAD [59]) and a client software was developed that can connect to CASS and request analysis results which can then be displayed as images or graphs in a graphical user interface. The software of this graphical user interface has been added to the main CASS software repository which can be obtained as open-source from [60].

Fig. 4.4 shows a screen-shot of the graphical user interface which has been set up to show a time-average of the estimated hit-rate as well as the diffraction pattern of the last found hit.



daq-amo-mon04 12321 RearPnCCDIsHitCleared HIST 100.0 % ☑ ● 1.00 Hz

○ lin ● log ○ exp ○ sqrt ○ square ☑ man scale Min 3 Max 100 Color 6

0.99 Hz

daq-amo-mon02 12321 HitrateHistory2000 HIST

☑ lin/log y-scale ☑ Man Y-scale Y-Min 0. Y-Max 0.1 ☐ Man X-scale X-Min 000

1 Hz

Fig. 4.4. graphical user interface used for tuning sample injection on-line. A time-average of the hit-rate is displayed (**bottom**) as well as the last diffraction pattern classified as hit (**top**). Sample: PBC virus. Photon energy: 1.8 keV. Injection: aerodynamic lens stack.

5

Application

Despite the attractive prospects of X-ray single particle imaging, so far, no complete three-dimensional reconstruction has been demonstrated. The main reason is that almost every aspect of this imaging method is currently at an early stage of development. This includes the development of the light source itself, which can be quite unstable at times, the injection and handling of the single particles in vacuum, the detection of the diffraction patterns and the development of algorithms that can analyze the resulting data. The work presented here focuses on the algorithmic side of the problems. Most of the published algorithms [61, 37, 62, 63] have been shown to work with simulated data, where all parameters that affect the formation of the diffraction signal can be controlled. Real data, however, still contain many artifacts that have yet to be understood. Therefore, it is inevitable to use a model system for development, for which most of the uncertainties can be eliminated. The model system that is presented here is an iron-oxide nanoparticle called *nanorice*, which is chemically stable, can be produced in huge quantities, is commercially available and scatters strongly, providing a very good diffraction signal. The nanorice particles were inhomogeneous in size and shape which reflects the situation expected for biological objects. The application of the algorithms described in chapter 3 to real data is presented in the following, including the first three-dimensional reconstruction of serial CDI data.

5.1 *nanorice* - an ellipsoidal iron oxide nanoparticle

Nanorice is a nanoparticle with a strongly scattering atomic composition. It consists of an ellipsoidal iron oxide core (Fe_2O_3) with principal axes of ∼200 nm and 50 nm and it is coated with silicon oxide (SiO_2). See figure 5.1 for transmission electron micrographs. It was purchased from Corpuscular, Inc. (Cold Spring Harbour, NY, USA).

Nanorice is an excellent model system for the development of orientation classification algorithms since its symmetry axis C_∞ reduces the rotational degrees of freedom from three to two and the orientations can, in principle, be recovered “manually” from individual diffraction patterns, making this sample an ideal test case for general methods of orientation recovery. In section 5.1.6 this nice property of nanorice will be used as a control.

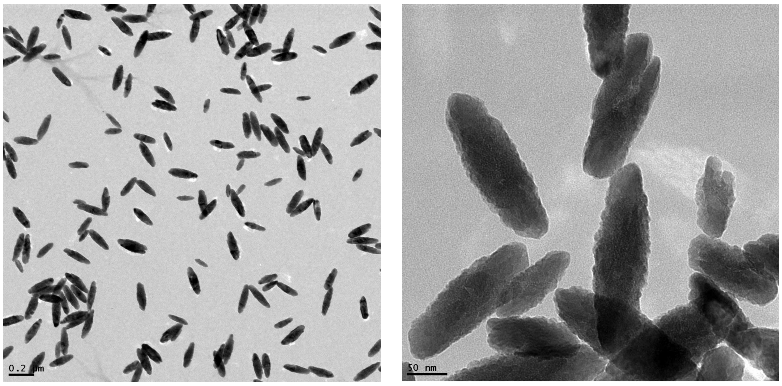


Fig. 5.1. TEM images of **nanorice particles**. All TEM images reproduced in this thesis are used with the kind permission of Thomas Barends.

5.1.1 Data acquisition

Diffraction patterns were obtained during imaging experiments at the AMO beamline at LCLS. These experiments were conducted in the vacuum chamber of the CAMP end station [55] which has been installed into the AMO hutch in 2009 and was available for three years. The CAMP instrument was designed by the Max-Planck Advanced Study Group (ASG) for various types of experiments, including coherent diffractive imaging experiments. It is equipped with ion and electron spectrometers as well as two large-area, one megapixel pnCCD detectors. The pnCCD detectors can detect single photons and they can also be operated in a spectroscopy mode where the deposited energy can be resolved to within 40 - 200 eV in the energy range between 100 eV and 25 keV. The X-ray beam is focused to a spot of $10\,\mu m^2$ using a pair of *Kirkpatrick-Baez* (KB) mirrors. The nominal X-ray pulse length (determined from the electron pulse length) was 150 fs and the photon energy was 1.2 keV with a pulse energy of 3.2 mJ. The FEL repetition rate was 60 Hz.

The sample was injected into the X-ray beam using aerosol injection (see page 6). The aerosol was generated from a liquid solution containing the sample objects by a gas dynamic virtual nozzle [19] and by a commercial nebulizer (Burgener Mira Mist CE nebulizer, AHF Analysetechnik, Tübingen, Germany). While the FEL focus was 10 μm^2, the diameter of the particle beam was approximately 440 μm at the X-ray interaction region (22 mm from the aerodynamic lens stack exit). The particle speed was estimated to be on the order of 100 m/s [64]. These parameters allow a maximal achievable hit rate of 2.6% [47].

Diffraction patterns were recorded using a pnCCD detector [55] placed at a distance of 738 mm from the interaction region. The detector consists of 1024×1024 pixels (each sized $75\mu m \times 75\mu m$) with detector-halves that can be positioned such that the direct X-ray beam passes through the gap between the two halves. In addition to the gap, a circular central region with a radius of 30 pixels is cut out for the direct beam. The detector area was shielded by a 3 μm thick polymide filter to prevent the contamination by sample particles. The detector readout was synchronized with the repetition rate of the X-ray pulses of 60 Hz. The pnCCD detector was operated in a gain mode that allows each pixel to store charges created by 10^3 photons per pixel (measured at 2.0 keV [65]). A postprocessing was performed to make use of calibration methods described in [66].

5.1.2 Classification results

The classification scheme described in chapter 2.1 was used without any specific adjustment to the nanorice diffraction data apart from providing a training set consisting of three different classes of nanorice diffraction patterns:

- Class 1: diffraction patterns of single nanorice particles (“single hits”)
- Class 2: diffraction patterns of two nanorice particles (“double hits”)
- Class 3: the remaining diffraction patterns, containing diffraction patterns of multiple nanorice particles, empty shots and saturated shots.

The training dataset consists of 20 diffraction patterns per class that were selected based on visual inspection. A random-forest classifier was built for this training dataset and the size-, symmetry- and PCA features described in chapter 2.1 proved powerful enough to classify the whole nanorice data set, as visual inspection of random images showed that 99% of the diffraction patterns were correctly classified. Figure 5.2 shows a random subset of automatically classified diffraction patterns.

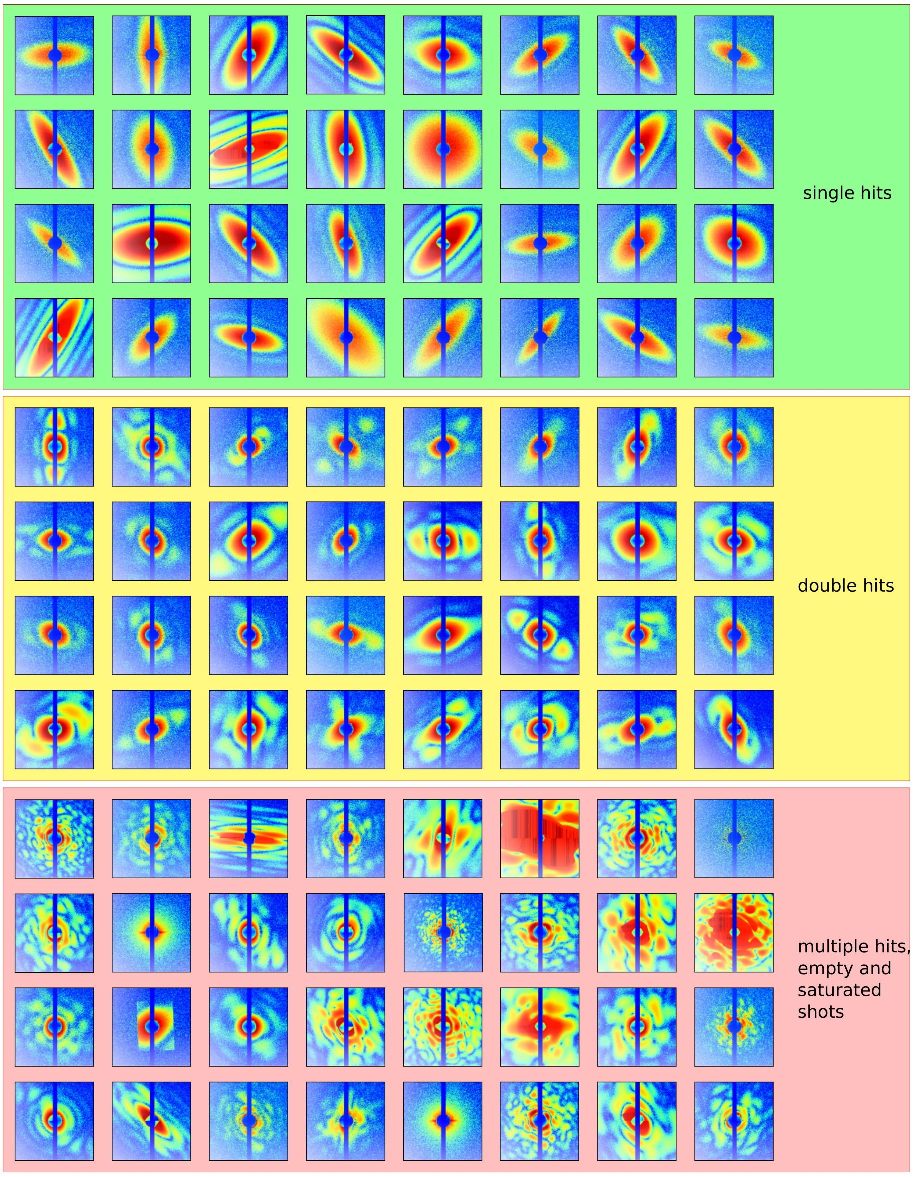
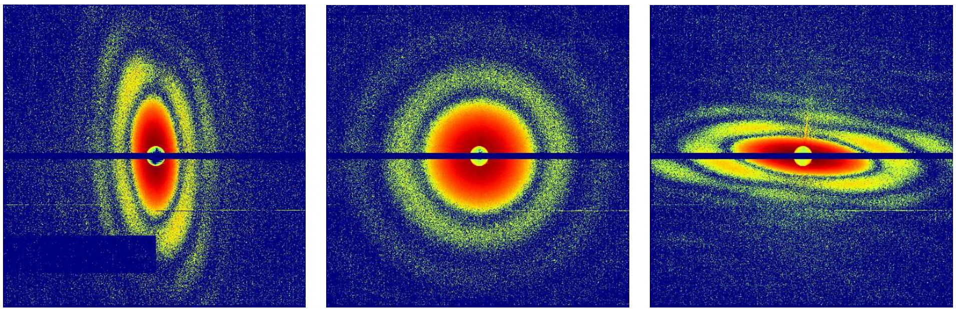


Fig. 5.2. Automatic classification results. Only a random subset of all 4248 diffraction patterns is shown and only a central detail is shown for each diffraction pattern, the diffraction extends beyond the borders shown here.

The nanorice diffraction dataset consists of a total of half a million snapshots that were collected with 1.2-keV photons at the LCLS. After removing most of the empty shots by applying a pre-processing hit-finding filter (see chapter 4) 4248 patterns remained of which 1389 were recognized as single shots by the automated classification employing the random-forest classifier.

5.1.3 Orientation recovery results

The single nanorice patterns that were identified by the statistical classification method (see chapter 2.3) were used as the input for the geodesic orientation recovery algorithm described in chapter 3. Figure 5.3 shows how P_1 and P_2, the two diffraction patterns between which all geodesic sequences will be spanned to find orientation relationships, were found following steps 1-6 outlined in section 3.3.



$\tilde{P}_1$: randomly picked diffraction pattern

P_2: out-of-plane antipode of P_1

P_1: out-of-plane antipode of P_2

Fig. 5.3. Finding P_1 and P_2, the two diffraction patterns between which all geodesic sequences will be spanned to find orientation relationships. $\tilde{P}_1$ is picked randomly, then P_2 is found as the diffraction pattern with maximum geodesic out-of-plane distance to $\tilde{P}_1$. In order to overcome stops in the geodesic sequences that are caused by object symmetry, the initial pattern $\tilde{P}_1$ is then replaced by P_1, the diffraction pattern with maximum geodesic out-of-plane distance to P_2.

After the diffraction patterns P_1 and P_2 are chosen, geodesic sequences between the in-plane rotated diffraction patterns $P_1(\alpha)$ and $P_2(\beta)$ are searched for. The rotational symmetry of the nanorice particle is reflected in the diffraction pattern P_2, therefore all resulting geodesic sequences can be visualized in two dimensions as shown in figure 5.4. Not all of the 1389 single-particle diffraction patterns could be assigned to geodesic sequences in between $P_1(\alpha)$ and $P_2(\beta)$. Importantly, the resulting orientation map covers only a subset of 128 diffraction patterns. As will be shown in section 5.1.5, the reason for this is a drastic inhomogeneity of the dataset, as the remainder of the diffraction patterns belong to nanoparticles with different sizes and shapes.

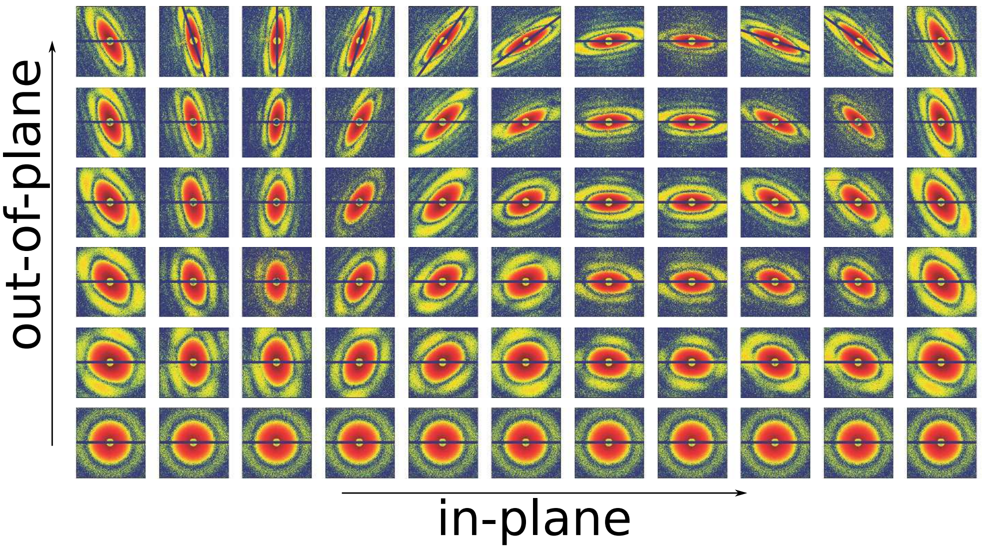


Fig. 5.4. orientational alignment of XFEL nanorice diffraction snapshots using the geodesic orientation recovery algorithm. This orientation map shows only a subset of the 128 aligned diffraction patterns that were collected at LCLS. Figure taken from [44].

The angular orientation map shown in figure 5.4 can be used to interpret the orientations of the individual diffraction patterns. The bottom row shows diffraction patterns of the rice grain hit head-on, while the top row shows diffraction patterns of the particles hit on their broadside. Thus the geodesic sequences in between describe a rotation through an angle of 90 degrees. This is in line with the symmetry of the particle: without symmetry the angle would be 180 degrees but it is reduced by a factor of two by the two-fold symmetry axis of the nanoparticle. The geodesic distance of these sequences is linearly mapped to angles between 0 and 90 degrees and eq. 3.1 can be used with $\varphi = 90°$ to calculate the orientations of all diffraction patterns.

Now that the orientations are known, the diffraction patterns can be used to construct the three-dimensional diffraction volume. Therefore the i-th detector pixel position $\overrightarrow{p_{i,j}}$ of the j-th snapshot is represented in a laboratory frame the origin of which is located in the interaction region, with the z-axis aligned to the X-ray beam. Let $\boldsymbol{R}_j$ be the rotation matrix of the j-th snapshot, λ the wavelength and $\boldsymbol{k}_0 = \vec{z}\,\frac{2\pi}{\lambda}$

the wave-vector of the incident beam, then $\overrightarrow{p_{i,j}}$ can be mapped to reciprocal space $\boldsymbol{q}_{i,j}$ in the following way:

$$\begin{aligned}
\tilde{\boldsymbol{k}}_{i,j} &= \boldsymbol{R}_j \overrightarrow{p_{i,j}} \\
\boldsymbol{k}_{i,j} &= \frac{2\pi}{\lambda} \cdot \frac{\tilde{\boldsymbol{k}}_{i,j}}{\left|\tilde{\boldsymbol{k}}_{i,j}\right|} \\
\boldsymbol{q}_{i,j} &= \boldsymbol{k}_{i,j} - \boldsymbol{k}_0
\end{aligned}$$

Figure 5.5 shows slices of the resulting diffraction volume and it can be seen that the averaging of diffraction patterns with similar orientation enhances the intensity of the diffraction fringes.

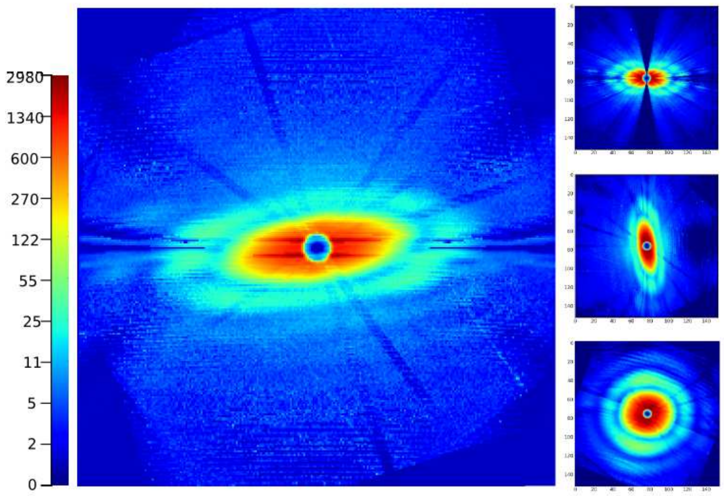
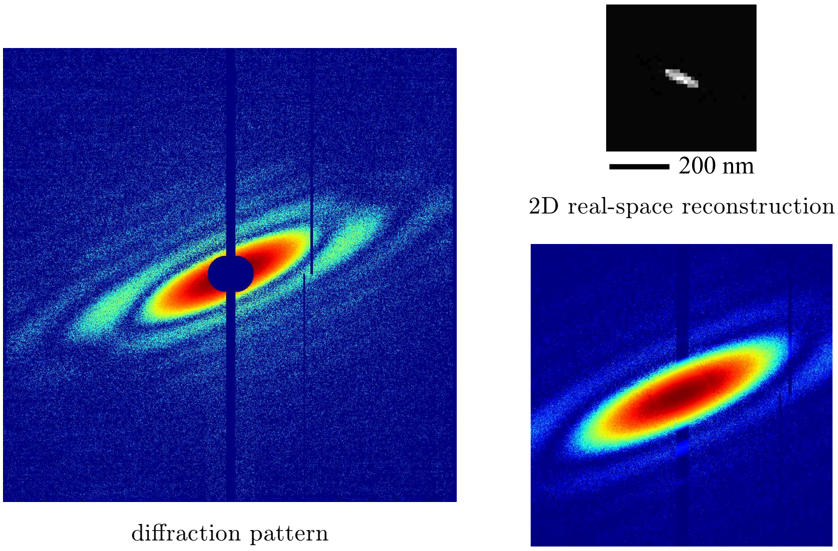


Fig. 5.5. Slices through the 3D diffraction volume assembled by the geodesic method. Fig. 5.4 shows a subset of the 128 diffraction patterns that contributed to this diffraction volume. The color-scale shows the intensity values in the arbitrary digital units of the pnCCD detectors. **Left**: slice with arbitrary rotation. **Right**: orthogonal slices aligned to the major axes of the particle.

5.1.4 Phase retrieval

Two algorithms (HIO [33] and RAAR [67]) for two-dimensional phase retrieval were implemented and extended with a nonlinear modulus constraint which amplifies strong intensities in the first few iterations in order to prefer the reconstruction of intense low angle signals in the region where data is missing due to the beamstop. Also, HAWK [68] was used, a phase retrieval software suite and software library developed by Filipe Maia which is unsurpassed in speed because it uses hardware acceleration (GPU computing) for the discrete Fourier transformation. For the three-dimensional phase retrieval, the software library `dm_recon` [69] implemented by Jan Steinbrener was used which is documented in [70]. This implementation uses a distributed memory model that can connect the main memory of many computers, forming a giant working memory that can hold the large three-dimensional diffraction- and reconstruction data. At the same time, it uses message passing to distribute work over the network of connected computers resulting in a huge speedup by parallelization. The library `dm_recon` implements the algorithms HIO [33] and difference map [71] in three dimensions.



2D real-space reconstruction

diffraction pattern

reconstructed missing information

Fig. 5.6. Two-dimensional phase retrieval of a single nanorice diffraction pattern. The real-space projection of the nanorice grain was reconstructed using the HIO algorithm [33]. The lower right image shows the reconstruction of the low angle diffraction intensities that were missing in the experimental data (left) because of a gap int the detector and a central hole for the direct X-ray beam.

For three-dimensional phase retrieval a box 10 times larger than the expected volume of the nanorice particle was used as a rough initial support estimate and the difference map algorithm [71] was used to recover the phase information and the electron density map which is shown in fig. 5.7. The algorithm converged after 1000 iterations. No sophisticated update of the support constraint or tweaking of parameters was necessary. The width and length of the reconstruction correspond to 39 nm and 150 nm, respectively. The result is consistent with the size and shape that has been determined by transmission electron microscopy and is also consistent with the individual two-dimensional diffraction patterns (see fig. 5.6). The theoretical optical resolution of the imaging setup was 9.9 nm. However, the diffraction pattern did not extend to the detector edge, and only 7 diffraction orders were visible, resulting in a resolution of 21 nm.

Figure 5.7 shows, the latter was used for the reconstruction of the nanorice shown in fig. 5.7.

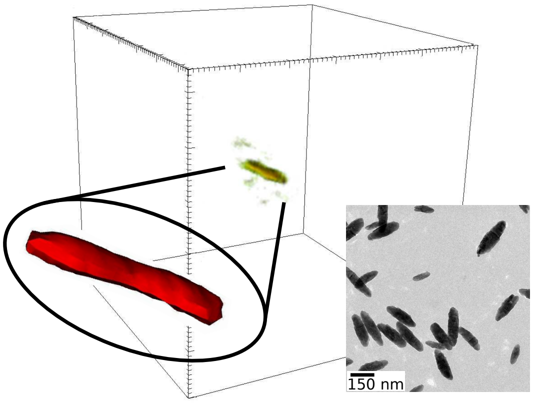


Fig. 5.7. 3D reconstruction of the electron density of a medium-sized 150 nm long nanorice particle. The size distribution of the sample can be seen in the TEM micrograph (inset). The bounding box shows the oversampling volume, the red magnified object shows an isosurface representation.

5.1.5 Data inhomogeneity

A closer look at the geodesic three-dimensional reconstruction showed the inhomogeneity of the diffraction snapshots as a fundamental difficulty of experimental CDI data that could be successfully addressed by the geodesic approach. The size and shape inhomogeneity (see fig. 5.7 for a TEM micrograph) of the nanorice particles splits the diffraction data in different groups of diffraction patterns that match into common three-dimensional diffraction volumes. Diffraction snapshots of particles that are indistinguishable based on the optical resolution end up in the same group. The geodesic approach singles out one such group by only considering diffraction snapshots that are interconnected by similar pairs of diffraction patterns, starting from a (randomly) user-chosen diffraction pattern. The degree of similarity is governed by the threshold ε introduced in chapter 3.1.

As mentioned before, a subset of 128 diffraction patterns out of 1389 was assigned to geodesic sequences starting from the initial diffraction pattern $\tilde{P}_1$ shown in fig. 5.3. Interestingly, picking a diffraction image from the remaining unassigned diffraction patterns and starting the algorithm again with this pattern as $\tilde{P}_1$ revealed another subset of diffraction patterns which corresponds to an object smaller than the one reconstructed in section 5.7. The orientation map can be seen in fig. 5.8. These results confirm that the geodesic algorithm can robustly handle inhomogeneous data sets by choosing the biggest subset of diffraction patterns that can be matched into a three-dimensional diffraction volume. This is not possible without modification for the common arc method, the Bayesian approaches and the diffusion map described in appendix A, since these methods include all diffraction patterns into the three-dimensional scattering volume whereas the diffraction patterns that are combined by the geodesic approach are contained within the border that is defined by the diffraction patterns P_1, P_2 and all their in-plane rotations.

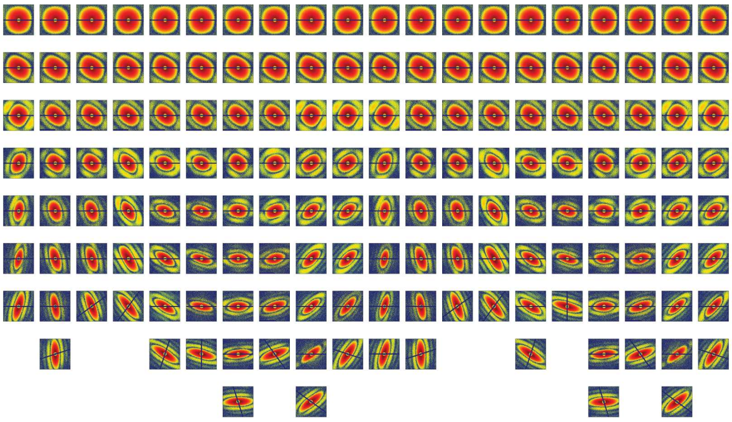


Fig. 5.8. orientation map of a second subset of matching diffraction patterns, corresponding to a smaller nanorice grain than the one resulting in fig. 5.4.

5.1.6 Using a simple geometric consideration as a control

In order to test the recovered orientations and to obtain more evidence that the different subsets found in the diffraction data are caused by sample inhomogeneities, the high symmetry of nanorice was exploited to verify the previous results.

The nanorice nanoparticle looks like a grain of rice (see fig. 5.7). Its minor axis is invariant under projections and the Fourier transform transforms it into a major axis in reciprocal space. The length of this major axis is constant in all diffraction patterns. The real-space major axis of the rice grain however is affected by the projection and the Fourier transform leads to a minor axis in reciprocal space whose size depends on the orientation. The ratio of the two axes in the diffraction pattern varies between 1 (the head-on orientation that aligns the rice grain with the X-ray beam, see bottom row of fig. 5.4) and an extremum r_{min} (broad-side orientation, see top row of fig. 5.4). The ratio r can be read out automatically by fitting ellipses to the diffraction fringes. When the rice grain is approximated by a cylinder of length d_{max} with spherical caps with diameter d_{min} (see fig. 5.9), the relation between angle and projection d is given by

$$\begin{aligned}\sin(\beta) &= \frac{d - d_{\mathrm{min}}}{d_{\mathrm{max}} - d_{\mathrm{min}}} \\ &= \frac{d/d_{\mathrm{min}} - 1}{d_{\mathrm{max}}/d_{\mathrm{min}} - 1}.\end{aligned}$$

The ratios d/d_{min} and $d_{\mathrm{max}}/d_{\mathrm{min}}$ translate to $1/r$ and $1/r_{\mathrm{min}}$ in reciprocal space. The angle β can be identified as the second Euler angle of the rice grain's orientation and can be determined from the ratios of major and minor axes in the diffraction pattern:

$$\beta = \arcsin \frac{1/r - 1}{1/r_{\mathrm{min}} - 1} \tag{5.1}$$

The first Euler angle α of the rice grain's orientation is given by the in-plane orientation of the fitted ellipse. The third Euler angle γ is completely degenerate because of the S^{∞} symmetry along the axis of the rice grain.

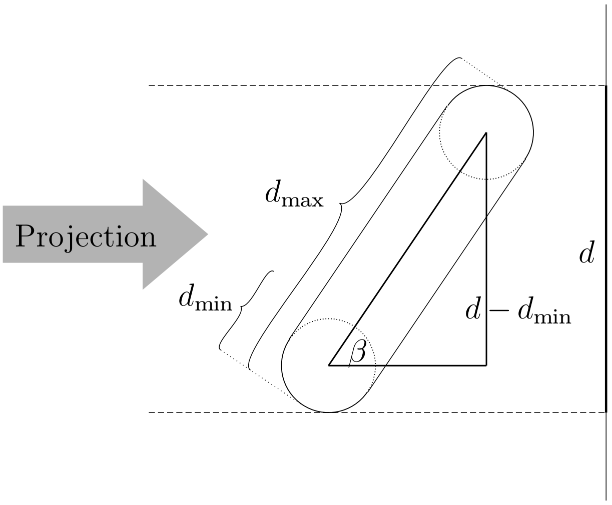


Fig. 5.9. Simplified projection model of a nanorice grain.

Assembling the diffraction patterns to the three-dimensional diffraction volume with the orientations obtained by the simple and robust "unprojection" (eq. 5.1) resulted in a surprise: the diffraction volume was not usable for three-dimensional reconstruction; the diffraction fringes were washed out. As it turned out, the diffraction data could not be merged into a single dataset because of inhomogeneities of the sample objects which was confirmed by transmission electron microscopic images of the sample: the rice grains had slightly different sizes and shapes (see TEM micrograph in fig. 5.7. The more general geodesic approach, however, without being equipped with any prior information about the sample, found subsets of diffraction patterns with matching fringes (see fig. 5.4). The averaging of several diffraction patterns enhanced the diffraction fringes instead of washing them out (see fig. 5.5). The result was consistent with the simple geometric check described in this section.

It turns out that such an inhomogeneity is also present in the diffraction data collected from biological samples such as viruses. Here, the problem is so severe that a three-dimensional reconstruction was not possible (see section 5.2).

5.2 Preliminary application to virus diffraction data

The successful 3D-reconstruction of the nanorice particle motivated the collection of diffraction datasets of biological samples. Viruses with different sizes and shapes were used: T4 bacteriophage, PBCV-1, CroV and Mimivirus.

5.2.1 Samples

The T4 virus is a large bacteriophage with an elongated icosahedral head (~80nm diameter) that contains a 172 kilobase pair DNA genome. A helical, hollow, contractile tail (~20nm × 100nm) is attached to the capsid.

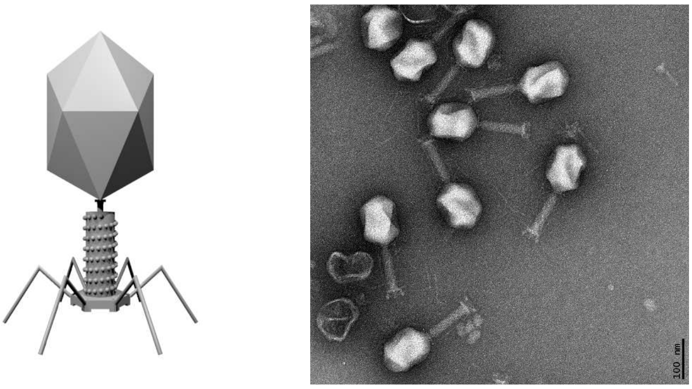


Fig. 5.10. schematic drawing of a T4 bacteriophage (left) and TEM image (right, already published in [47]).

The *Paramecium bursaria chlorella virus* PBCV-1 and the *Cafeteria roenbergensis virus* CroV are both giant icosahedral double-stranded DNA viruses. PBCV-1 infects fresh water algae (*Chlorella variablilis*) [72], CroV infects a marine heterotrophic flagellate (*Cafeteria roenbergensis*) [73]. CroV is of particular biological interest because it is a virus that itself can get parasitized by another virus named Mavirus and no three-dimensional structural information is available. Mimivirus is one of the larges known viruses, its icosahedral capsid has a diameter of 400 nm and fibrils attached to the capsid add to the total diameter which is about 750 nm [74]. A two-dimensional reconstruction of single Mimiviruses was achieved using X-FEL diffractive imaging [75].

5.2.2 Results - aerosol injection

The viruses, like all biological samples, need a hydration shell provided by residual buffer solution which remains bound to the particle when an aerosol injection system is used. The diffraction patterns were collected using the CAMP instrument at the AMO beam-line at the LCLS. The experimental parameters correspond to the parameters used for the nanorice dataset as described in 5.1.1.

The classification scheme described in 2.1 that was also used for automatic classification of the nanorice dataset proved general enough to handle virus diffraction data without modifications, apart from specifying a different training dataset for each sample. Again, a training data set of three classes (single hits, multiple hits and a class containing the rest) consisting of 20 diffraction patterns per class was selected and classification was performed using a random-forest classifier. Figure 5.11 shows a subset of automatically classified virus diffraction patterns.

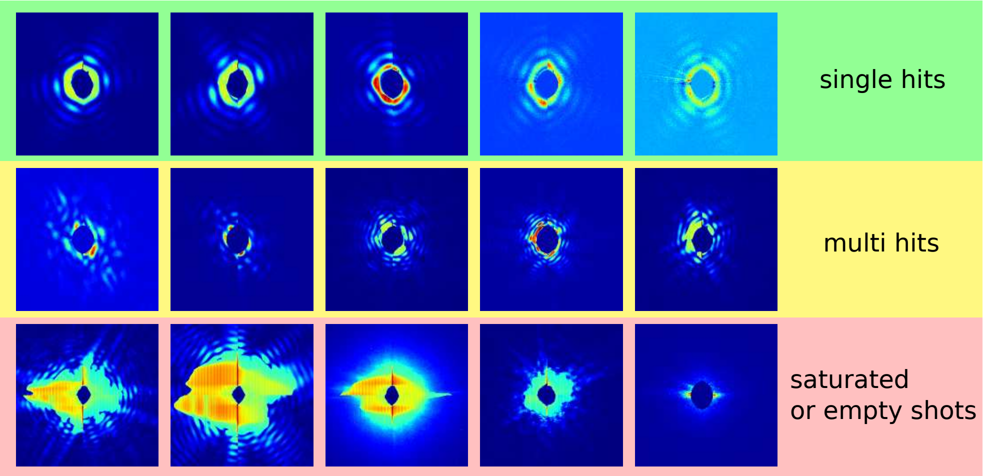


Fig. 5.11. Classification: Randomly selected diffraction snapshots of Mimivirus and their automatically obtained classes.

Unexpectedly, the aerosol dataset displayed huge size inhomogeneities. The object size was estimated by fitting diffraction rings of a spherical object to the diffraction patterns. Fig. 5.12 shows that the apparent size of the T4 virus varies dramatically from snapshot to snapshot and the average size of ∼330 nm is much larger than the actual particle size (∼100 nm T4 head). This effect only appears in combination with aerosol injection, the size distribution obtained from TEM as well as from liquid jet diffraction data (see 5.2.3) is homogeneous. A possible explanation for the origin of this size inhomogeneity are varying particle to detector distances which will only change the apparent size without really altering the particle. It is also possible that structural changes occur or that a layer of solvent remains around the particle, leaving a salt crust when it dries in the vacuum chamber. The effect is less severe for larger particles (see fig. 5.12), suggesting that either the focusing of the aerodynamic lens works better with bigger particles or that the size of the bigger particles is similar to the preferred size of solvent droplets, reducing solvent-shell artifacts as illustrated in fig. 5.12.

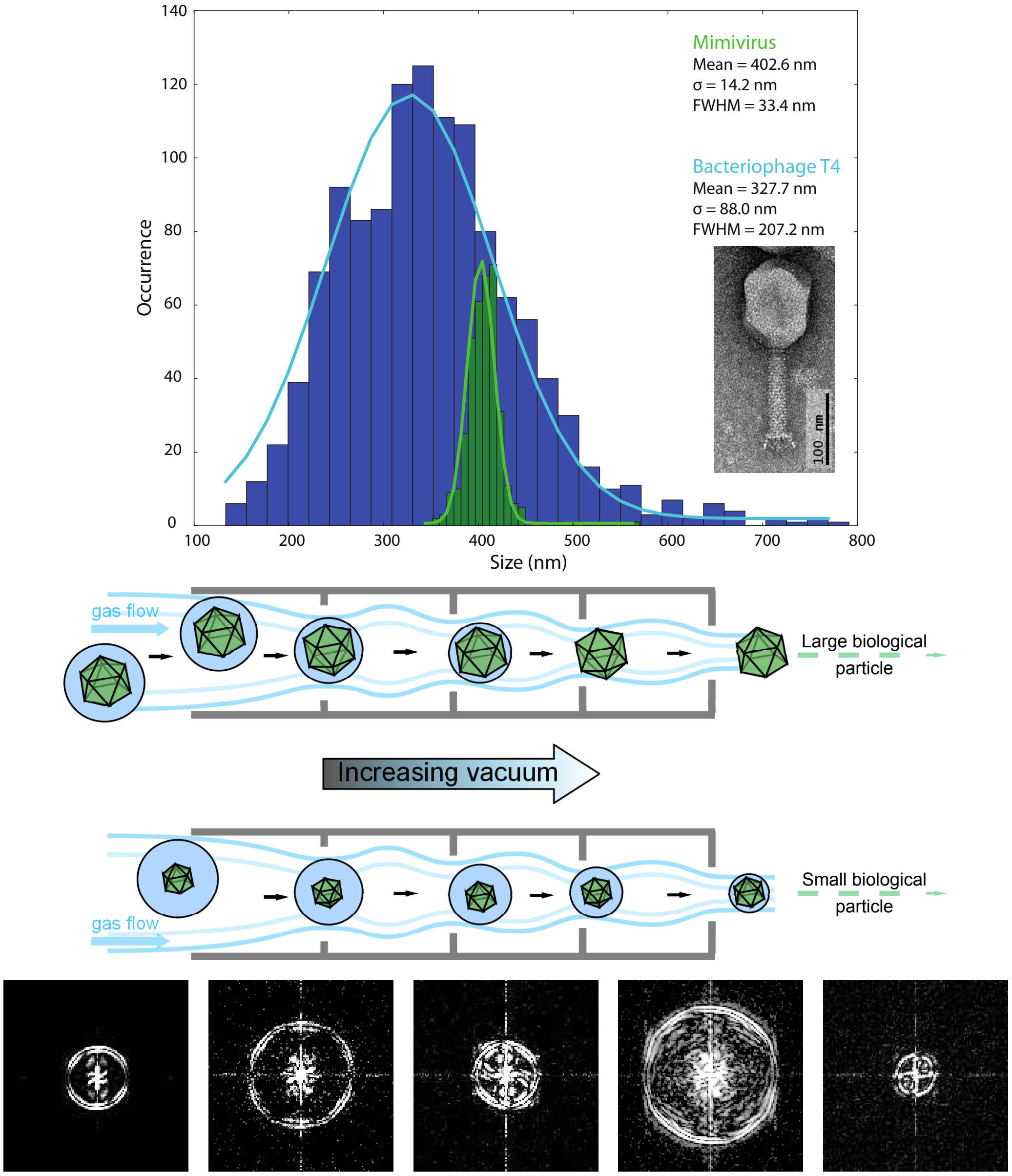


Fig. 5.12. Top: Size histogram of the small T4 virus (blue) and the Mimivirus (green). The size was estimated by fitting diffraction rings of a spherical object to the diffraction patterns. The apparent size of the larger Mimivirus matches the nominal size obtained from electron microscopy, but the obtained size of the smaller T4 virus (328 nm) differs from the nominal size (100 nm icosahedral head). TEM image: T4 virus, courtesy of Thomas Barends. **Middle:** aerosol injection of particles with different sizes. The solvent layer around the sample evaporates as the particle moves along the injector with increasing vacuum. Equal initial droplet sizes may leave smaller samples with a thicker remaining solvent shell, while a larger sample may dry off completely. **Bottom:** autocorrelation functions of different T4 snapshots, displaying different apparent object sizes.

The severe size inhomogeneities render the current aerosol data of mid-sized viruses unusable for combination into a single three-dimensional dataset. Also selecting a subset of matching particle sizes was not possible because too few diffraction patterns matched in apparent size.

5.2.3 Results - liquid jet injection

To avoid inhomogeneities in the apparent virus size, another injection system was used, the liquid jet system introduced on page 6. In a liquid jet, the amount of solvent in the vicinity of a sample particle is constant in contrast to aerosolized particles with varying solvent shells; thus injection artifacts like varying salt crusts are avoided.

In order to reduce the large background signal from the liquid jet which is much higher than the background of a small aerosol solvent shell, measurements were performed in the “water window” with a wavelength that had to be increased to 2.4 nm / 520 eV photon energy (as compared to 1 nm used for aerosol injection). Water window microscopy was first proposed by Wolter [76] in 1952, it is the wavelength region between the K-absorption edge of oxygen at a wavelength of 2.3 nm and the K-absorption edge of carbon at 4.4 nm where the contrast between biological samples (mostly carbon) and water (mostly oxygen) is maximized.

The virus diffraction dataset obtained by liquid jet injection does not display the size inhomogeneities observed in the aerosol dataset, but the resolution of this dataset is reduced because the reduced wavelength necessary for water window imaging. Still, the background scattering from the liquid jet severely hampered a three-dimensional reconstruction. Figure 5.13 shows that finding geodesic sequences is possible after masking out the intense background signal of the liquid jet, but since the signature of the jet changes rapidly, not enough matching sequences could be found to be combined into a three-dimensional diffraction volume.

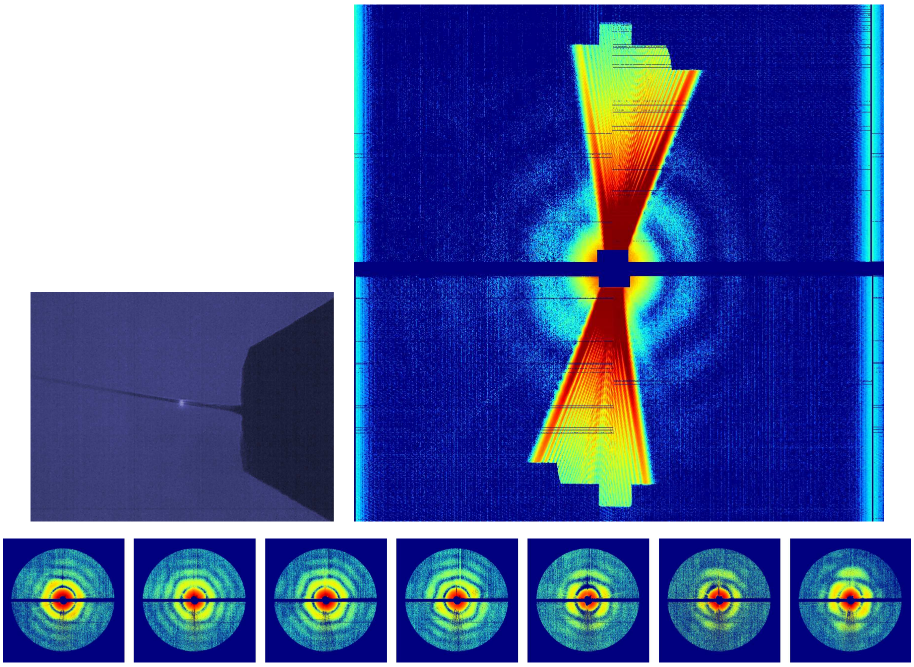


Fig. 5.13. CroV samples injected with a liquid jet. **Top left**: microscopic image of the jet, the bright point shows fluorescence of the X-ray beam. **Top right**: strong diffraction signature of the liquid jet (vertical streak). **Bottom**: geodesic sequence of diffraction patterns. Only diffraction patterns with low jet signature were used to find the sequence and the jet streak was masked out heavily.

6

Discussion

The result of this thesis is a set of data evaluation methods for three-dimensional serial coherent diffraction imaging of microscopic samples. A major achievement is a new algorithm for orientation recovery of diffraction patterns - the Geodesic In-Plane Rotation ALgorithm **GIPRAL**. This algorithm is able to cope with the practical difficulties posed by experimental conditions like noise, weak signals and the possibility of inhomogeneous specimen which can be contaminated by artifacts or the presence of multiple conformations. For the first time the successful application to experimental data was demonstrated with XFEL diffraction patterns of inhomogeneous nanoparticles which where automatically sorted and orientationally aligned, resulting in the first three-dimensional reconstruction[6.1].

6.1 Comparison to other orientation recovery approaches

During the last six years, several orientation classification algorithms [63, 77, 62, 43] have been published and shown to work with simulated data. Although some are reported to cope very well with simulated noise and very low photon counts [62, 78], none have been used to successfully reconstruct an object from serial single particle CDI data. This is not due to lack of experimental data, since we made the data used in this work publicly available at cxidb.org [47]. A detailed overview over existing approaches is presented in appendix A.

The essence of the method developed in the framework of this thesis is the finding that geodesic sequences of diffraction snapshots in reciprocal space can be assigned to geodesic sequences of object rotations in real-space. A distinguished property of

---

[6.1] Loh et al. [77] demonstrated a successful application of the EMC algorithm to diffraction snapshots of nanorice, but no real-space reconstruction could be obtained.

such sequences is their globally optimal nature: the geodesic distance is a combination of local distances between many data points into a conformable global distance, a principle which makes the isomap [79] algorithm so powerful and robust against outliers and noise. Fast dynamic programming approaches like Dijkstra's shortest path algorithm [45] are very efficient and guarantee globally optimal solutions.

GIPRAL combines two reliable sources of angular information (in-plane, out-of-plane) without propagating the error exponentially by nesting steps, a problem which is treated by additional averaging in common-line or -arc methods [61, 80]. In-plane angles can be obtained with an accuracy that is only limited by the discrete nature of the pixel based diffraction detection. The introduction of artificially in-plane rotated diffraction patterns increases the effective density of the orientation sampling, a technique which could also benefit other approaches.

Moreover, every pixel of the whole diffraction pattern contributes to the angular information and is considered in GIPRAL, whereas in common-line methods used frequently in EM, only lines within diffraction patterns are compared. This makes GIPRAL applicable to data sets with much fewer scattered photons. The method of moments [81] used in EM which has also been described for X-ray diffraction snapshots [82], analyzes an ellipsoid which has the same moments of inertia as the sample in order to recover its orientation. This method, however, suffers from ambiguities, especially when no phase information is available, as is the case in X-ray CDI. GIPRAL is also applicable to objects that have identical or very similar principal axes, such as icosahedral viruses. While these limitations prevent the methods established in EM from generally being used in CDI, GIPRAL can only benefit from the additional phase information of real space projections obtained from EM. An application of GIPRAL to EM data has not been tested yet, however, there are no theoretical obstacles that would prevent this.

Orientation classification schemes proposed in Refs. [83, 80] make use of the Pearson correlation coefficient to estimate the diffraction pattern similarity while [43] uses the Euclidean distance measure. In the described geodesic approach, the Pearson correlation coefficient is favored because of its invariance under linear transformations of the diffraction intensities. As with Bayesian methods [62, 63, 37] the ensemble information of all diffraction patterns combined is used to infer object orientations, making GIPRAL also useful for data with very low photon counts [62, 84]. Unlike the expectation maximization algorithms used in the Bayesian methods, the dynamic programming algorithms applied in the geodesic search ensure that the global optimum is found. The Bayesian methods make use of a statistical noise model to estimate the likelihood that a diffraction snapshot fits into the diffraction volume with a given orientation. Such noise models are beneficial for low photon counts. While this was not necessary for the data shown in this thesis, GIPRAL can easily be extended with such noise models. Then the notion of distance has to be replaced with the Bayesian interpretation of likelihood and the noise models can be used to estimate the likelihood that two diffraction snapshots are "nearest neighbors".

Compared to the graph-theoretic analysis of scattering data [43], GIPRAL refines simple pairwise local distances into an accurate integral distance measure and uses in-plane angles as an additional source of information. The geodesic analysis reduces the orientation recovery to one-dimensional sub problems, making it fast and physically intuitive.

GIPRAL is the first single-particle CDI orientation recovery method that has been used for a successful three-dimensional reconstruction of X-ray diffraction snapshots with random orientations. Its success is based on the ability to handle inhomogeneous data as well as its ability to find globally optimal solutions. Since the intermediate one-dimensional snapshot-sequences are easy to visualize, the whole process can be supervised and the parameters (nearest neighbor threshold ε) can be tweaked. In conclusion, GIPRAL is a method that can be used for single particle CDI reconstructions in a very general sense: it requires no assumptions about the sample object other than its symmetry group (which can be obtained from diffraction patterns), it can cope with weak diffraction signals and inhomogeneous data and it can handle sample object symmetries.

6.2 Towards the imaging of biological samples

The ultimate goal of X-ray CDI, the imaging of single biological molecules, cannot not be realized yet, because the peak brilliance of FEL X-ray pulses is currently not high enough. While improvements in the X-ray sources making the necessary intensities available eventually may be expected, the development of algorithms for sorting and orientation recovery of such samples is required, too. The methods presented in this thesis constitute a major step towards this goal by demonstrating their applicability to experimental FEL data of inorganic nanoparticles.

The next logical step in this development is the application to biological model systems. Large viruses are ideal in this respect, since they have similar atomic constituents as single biological molecules, but provide stronger scattering signals due to their large size. In addition, many identical copies can easily be produced - a prerequisite for serial CDI. Since the imaging of viruses is of high biological interest in itself, too, a large amount of structural information obtained from cryo-electron microscopy is already available to benchmark results obtained with CDI.

In terms of data analysis, biological viruses are no different from the nanorice particle for which a three-dimensional reconstruction could be obtained. Experimentally, however, they have to be treated very carefully in order not to affect their integrity which makes sample injection difficult. In particular, the requirement for defined buffer compositions can be at odds with the need for volatile buffers that leave only minimal background for aerosol injection. Existing sample delivery methods are not yet optimized to cope with these difficulties.

Indeed, as shown in chapter 5.2, a three-dimensional reconstruction of the virus diffraction data was not possible because because of various reasons. In the case of liquid jet injection, the background signature of the liquid jet, which was not constant over time, combined with a high background scattering interfered with orientation recovery. In the case of aerosol injection, huge inhomogeneities of the injected sample objects were observed, as can be seen in Fig 5.12. Possible causes of these inhomogeneities are poor aerodynamic focusing resulting in different particle to detector distances or different amounts of residual solvent around the particles.

The data of CDI experiments with giant Mimiviruses (700 nm) does not display the injection artifacts, but these giant viruses are not perfectly reproducible, which was one of the factors preventing a high-resolution three-dimensional reconstruction.

If the injection problems encountered with T4 and CroV cannot be alleviated, other possibilities for sample introduction include a mechanical sample delivery with sample particles suspended on thin membranes analogous to the sample holders used in cryo-EM. Alternatively, electrospray can be used to disperse liquids containing the sample objects into fine aerosols using Coulomb repulsion.

It can be expected though that these problems will be overcome in the near future, making the algorithms proposed in this thesis applicable to single particle coherent diffractive imaging of nano-scale biological samples.

6.3 Room for improvements / outlook

6.3.1 Technical improvements

During the course of this thesis, much of the technology has not only been developed but also improved already. New imaging modes have been established for the pnCCD detectors which reduce the effects of charge spill, greatly improving the quality of the recorded diffraction patterns. A successor to the CAMP instrument has been built, LAMP [85] (LCLS ASG Michigan Project) including large gate-valves in order to separate the detectors from the interaction region, so that the vacuum chamber may be vented without having to warm up and cool down the detectors. This allows for rapid changes to and repair of components inside the chamber without wasting hours of valuable beam-time. These are just two examples of improvements that do away with factors that made life difficult for CDI pioneers. In addition, many more details will be solved and labor-some tasks will be automated, turning the exciting but sometimes "touchy" CDI experiments more and more into a standard imaging method.

On a more fundamental level, there is still room for improvement in the construction of detectors with higher dynamic range, which will greatly benefit data evaluation. As an example, samples of the giant Mimiviruses (700 nm including fibrils) lead to very high intensities at small scattering angles, exceeding the dynamic range of detectors available so far.

In addition, the repetition rate of future FEL facilities will be increased (several kHz up to MHz at supra-conducting European XFEL and LCLS-2) which will help tremendously as the number of diffraction snapshots is a limiting factor right now. More diffraction snapshots will improve the signal to noise ratio of the data ensemble, facilitate orientation recovery by improving the orientational sampling and will also soften the resolution limit outlined in section 1.3.

6.3.2 direct measurement and manipulation of orientations

At least for some types of sample, the necessity for *a posteriori* orientation recovery can be eliminated by directly measuring or even manipulating the orientation of the sample particles. While optical tweezers can be used to manipulate electrically neutral particles, these use very strong fields which might compromise the integrity of the sample object. However, weaker electric or electromagnetic fields can be used to control the orientation of sample particles with a permanent electric dipole moment [86]. Also, the orientation of such sample objects can be measured by analyzing the flight direction of fragments that are released after the destruction of the sample by the intense FEL pulse [87].

6.4 Conclusions

In this thesis, the first successful three-dimensional reconstruction from serial coherent diffraction data is demonstrated. To this end, novel algorithms such as feature-based classification and the Geodesic In-Plane Recovery ALgorithm GIPRAL for orientation recovery were conceived and developed, and subsequently applied to real data.

It may be expected that coherent diffractive imaging with X-ray FELs will evolve from its current, budding stage to a stage that provides new insights into biomolecular structures that cannot be obtained by other methods. X-rays stand out because of their long penetration depth, the high resolution that can be achieved. Moreover, X-ray FELs not only promise to "outrun" radiation damage but also provide the possibility to perform time resolved measurements with ultra-short pulses. However, a lot of method development is still required to reach this stage. Many aspects have to be improved in concert, ranging from the light source itself to sample-handling and data analysis. A better understanding of injection and detection artifacts will help to improve reconstruction algorithms.

Two-dimensional diffraction snapshots have already helped to analyze nanoparticles in their natural environment, an example is the investigation of the morphology of airborne particulate matter without introducing substrates [88] . The generalization to three-dimensional imaging opens up the door to a new exciting chamber of the nano-world.

Appendix A

Existing approaches to the orientation problem

While GIPRAL was the first algorithm that could demonstrate the feasibility of serial coherent diffractive imaging, several alternative orientation recovery methods have been published. This appendix provides a brief summary of the approaches relevant for CDI.

A.1 Correlation

Section 1.1 showed that the spatial correlations that exist within the electron density of a sample object can be obtained by calculating the auto-correlation function of the electron density based on diffraction intensities. Using phase retrieval algorithms, the spatial correlations can be turned into a spatial reconstruction of the sample.

Similarly, correlations within the diffraction signal can be used to reconstruct the diffraction volume. It can be thought of as an angular extension to small-angle X-ray scattering (SAXS), where radial information of the rotational average of many particle diffraction signals is used to extract information about the shape of the sample object. If the degree of this rotational averaging is reduced to a small number of different orientations (as opposed to a continuous distribution of orientations), small angular fluctuations remain in the averaged intensities, carrying information about the sample structure. These angular SAXS fluctuations can be obtained by freezing the sample particles in space or time, removing the influence of diffusion on the rotational average.

Kam [89] showed that angular correlations of a single-particle diffraction volume can be obtained by forming ensembles of discrete orientations (without rotational averaging caused by diffusion), summing up the scattering within one ensemble, correlating the summed up intensities between pairs of different scattering directions and then averaging over many ensembles. Uncorrelated pairs will cancel out during the averaging. The correlations can thus be determined from many single-particle diffraction patterns or from diffraction patterns of particle ensembles. Converting these correlations into structural information is an under-constrained problem. Just as in phase retrieval, additional constraints like information on the sample symmetry can help to overcome the information deficiency [90]. This method was successfully applied to experimental soft X-ray scattering of nanoparticles that were lying on a membrane only allowing rotations about a single axis, constraining their orientations to one dimension [91]. Theoretical considerations for application to three-dimensional XFEL data can be found in [92].

A.2 Common arc

The common arc method is borrowed from the common-lines [93] methods used for three-dimensional reconstructions from micrographs in *cryo-electron microscopy* (cryo-EM). Because of the large depth of focus of conventional electron microscopes, electron micrographs can be seen as a projection of the sample object onto a plane. Any two such projections of a three-dimensional object share a common one-dimensional projection [93]. It is possible to identify those common projections which can then be used to find relative orientations of three projections, which are fixed by three common-lines [94].

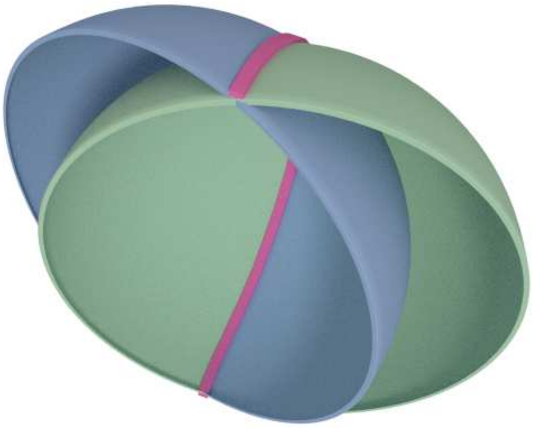


Fig. A.1. The common arc (red line) of two slices of the diffraction volume sampled by the Ewald sphere in two different orientations (blue and green). The two spheres intersect in the common arc.

The *projection-slice theorem* (see chapter 1.8) relates reciprocal slices to the Fourier transform of projections: the Fourier transform of a two-dimensional projection of the sample object is the same as a slice through the origin of the three-dimensional Fourier transform of the object parallel to the projection plane. These reciprocal slices share a common intersection.

This concept can be generalized to diffraction patterns which, depending on the wavelength, can not be considered to be orthographic projections corresponding to planar reciprocal slices (see chapter 1.8): the three-dimensional scattering volume (reciprocal space) is sampled by the Ewald sphere. When the sample object is rotated, so is the intersection of the scattering volume with this sphere. This spherical intersection always contains the origin and thus the sampling sphere of two diffraction snapshots always intersect in a *common-arc*, where the two diffraction snapshots share the same diffraction signal, allowing the common-arcs to be identified. Three common arcs of three diffraction snapshots uniquely determine their relative orientations (up to chirality) [83]. If the Ewald sphere is almost flat the common-arcs become common-lines, but unlike in electron microscopy, where the complex phase information can be measured, in diffractive imaging ambiguities are introduced because of Friedel symmetry. In [61] it is shown that a unique solution can still be obtained.

Since only local relationships are used in common-arc methods to establish the global orientations of all diffraction snapshots, local errors can accumulate and a global refinement is required, optimizing the self-consistency of all pairs of diffraction patterns [80]. In contrast to other methods (including GIPRAL), common-arc methods only consider correlations between small subsets of the diffraction patterns (the common-arc) instead of all the available scattering information.

A.3 Bayesian methods

The orientation recovery problem can be seen as an optimization problem in which a model of the diffraction volume and the orientations of the underlying diffraction patterns is sought that maximizes the likelihood of the model explaining the data. As shown in [95], both the Expansion-Maximization-Compression algorithm (EMC [63, 77]) and the Generative-Topographic-Mapping (GTM [62]) use Bayesian inference methods to find optimal models.

The EMC algorithm starts with a random initial guess of the diffraction volume. Tomograms (curved Ewald sphere sections) are extracted from this model volume and the likelihood is calculated of how well the measured diffraction patterns are explained by these tomograms, assuming a Poisson noise model. This likelihood is maximized by altering the model tomograms. Therefore, a tomogram is updated by averaging measured diffraction patterns weighted by the probability that the model tomogram explains the diffraction pattern. In a compression step, the consistency of the updated tomograms is enforced by interpolating all tomograms into a three-dimensional regular grid representing the model diffraction volume. These steps are repeated iteratively and can be thought of as an expectation maximization ([96]) technique.

The GTM method uses Gaussian functions as basis functions for a mapping from diffraction patterns to the underlying orientations. Similarly to the EMC method, the expectation maximization algorithm is used to maximize the likelihood. The central principle behind this method is the finding that while changing orientations of the sample are reflected in changes in high-dimensional 'manifest space' of measured pixel intensities, the true nature of this changes is only three-dimensional as different orientations only introduce three degrees of freedom. Thus, a latent three-dimensional manifold exists within manifest space, representing the orientations. GTM provides a way to embed this manifold into manifest space, mapping diffraction snapshots to their orientation.

A.4 Diffusion map / graph theory

The orientational relationships between diffraction snapshots that correspond to a finite number of different sample orientations can be encoded in a graph (see fig. A.2) with each vertex being a diffraction snapshot while the edge weights of the graph reflect the orientational distances, or angles. Just like in the geodesic orientation recovery developed in chapter 3.3, the orientational distances can be approximated locally by dissimilarities of diffraction patterns. Establishing the global structure of the graph also reveals information about the orientations. In the following, a geometrical interpretation of the graph theoretical take on the orientation problem is presented, which was described in a more formal way in [43].

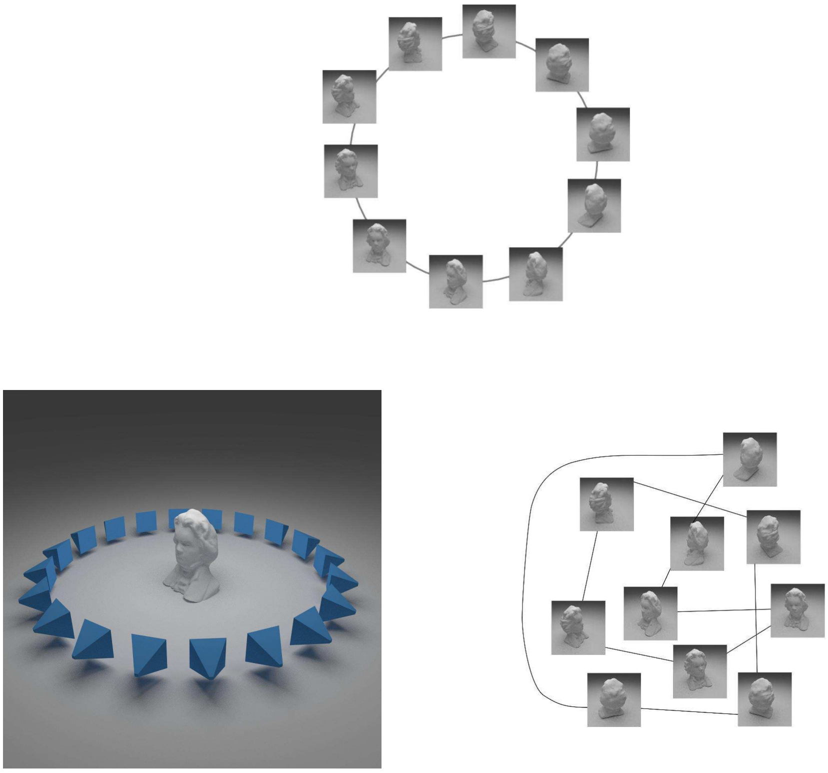


Fig. A.2. Snapshots of an object from different orientations (as viewed from the blue cameras around Beethoven's bust). The graph that reveals the orientations can be constructed by linking the two closest neighbors. Although the two graphs shown in this figure are identical in a topological sense, the orientations only become apparent in an adequate geometrical layout, as shown in the circular arrangement in the top image.

For simplicity, in this paragraph the dimensionality of the orientation problem is reduced to just one rotational degree of freedom, corresponding to a sample object which can only rotate about a fixed single axis. If only the two nearest neighbors of each vertex are taken into account, the graphical structure of this problem is that of a ring while the position on the ring encodes the orientation (see fig. A.2). Consider the one-dimensional *ring graph* R_n with n vertices. In a geometrical representation where the vertices are distributed equidistantly on a circle, the action of the *graph Laplacian* has a nice geometrical interpretation: The graph Laplacian is the discretized Laplacian differential operator and as such is symmetric in the sense that both the left and the right neighbor of a given vertex contribute equally to the effect of the Laplacian on said vertex. Therefore the Laplacian preserves the symmetry of the equispaced circular representation and it can only affect the radial position of the vertices, changing the scale of the circle. This shows that the circular representation behaves similarly to an eigenvector of the graph Laplacian, with the radial scaling factor (change in radius) being the eigenvalue. This scaling behavior still holds after decomposing the circle into the two components $\boldsymbol{x}(u) = \sin(u\,2\,\pi/n)$ and $\boldsymbol{y}(u) = \cos(u\,2\,\pi/n)$, two continuous eigenfunctions which, after evaluating at the vertex positions lead to two discrete eigenvectors $\boldsymbol{x}$ and $\mathbf{y}$ of the graph Laplacian of R_n. In this representation, the parameter u can be interpreted as the position along the circular ring graph and as such it encodes the angle of the underlying orientation. Thus, calculating the value of the Laplacian eigenfunctions evaluated at a given vertex (or snapshot) reveals the orientation of this snapshot.

The geometrical interpretation of the eigenvectors of the graph Laplacian was built on the requirement of equidistant vertices, or a uniform density of the orientation distribution. This requirement can be relaxed when the eigenfunctions of the continuous Laplace Beltrami operator are considered instead of the discrete eigenvectors of the graph Laplacian which might be sampled non uniformly [97]. A way to find these continuous eigenfunctions can be adopted from the analysis of physical diffusion processes where the Laplace operator plays a fundamental role. In a diffusion process, the spatial flux $\boldsymbol{j}$ is proportional to the spatial concentration gradient $\boldsymbol{j} = -D\ \nabla c$, while the temporal change of the concentration is directly linked to the flux and has to obey the continuity equation $\frac{\partial c}{\partial t} + \nabla \boldsymbol{j} = 0$. These two requirements can be combined into the diffusion equation

$$\frac{\partial c}{\partial t} = D\,\nabla^2 c$$

which displays the continuous Laplace operator. Simulating a diffusion process by means of a random walk among diffraction snapshots where the probability of a step is governed by the local angular similarity of the snapshots (the edge weights) can be used to find approximations of the wanted continuous Laplace Beltrami operator and its eigenfunctions even when the discrete sampling of the underlying continuous orientations are non uniform [43, 97]. In higher dimensions with three rotational degrees of freedom, the eigenvectors of the Laplace Beltrami operator are the Wigner D-functions [43, 98] which map the snapshot's orientation to the value of the Laplacian eigenfunction evaluated at its corresponding graph vertex, just like $\boldsymbol{x}$ and $\boldsymbol{y}$ do in the one-dimensional case.

Appendix B

Implementation

The computations carried out for this thesis were organized into a high-level controlling stage written in Python [99] and a low-level stage for time-critical computations written in C/C++. Therefore, numerical data was represented as continuous memory blocks that could be accessed from C/C++ code while NUMPY [100] wrappers were used for high-level access from Python. The following chapter focuses on performance optimizations that were performed to make GIPRAL applicable in acceptable time to big diffraction datasets (on the order of 100000 diffraction snapshots with a million pixels each). Starting from a naive implementation without optimizations, the computation times went down from months to days.

B.1 Hardware optimization

Several parts of the implementation required fast numeric linear algebra routines. SIMD (Single instruction, multiple data) hardware optimization was provided by the linear algebra library EIGEN [101] which uses SSE (*Streaming SIMD Extensions*) provided by the Intel x86 processors that were used for calculations. In addition, the library LAPACK [102] was used along with ATLAS [103] (*Automatically Tuned Linear Algebra Software*) that offered automatically performed extensive optimizations of the provided linear algebra implementations, tailored to the exact machines that were used for computations. NUMPY was used as a high-level wrapper for LAPACK and ATLAS. With those hardware optimizations, on the same machine a matrix multiplication of 10000x10000 random numbers with double floating point precision was two times faster than a commercial implementation (Matlab) that was not optimized for this specific machine.

B.2 Parallelization

The computational bottleneck for geodesic orientation recovery is the computation of a pairwise distance matrix of all pairs of diffraction patterns. The calculation of the distance between different pairs (resulting in single matrix elements) can be done independently, which means that this task can be parallelized easily. The distance matrix is divided into several parts and each part is sent out for calculation and distributed among several computing units ("*workers*"). There are two different types of workers: workers which share memory and thus data with the master process and workers that have isolated memory and need to send and receive data in a point to point type communication like a computer network. In the former, the *shared memory* type, communication between workers is only necessary if they need to write to a memory address which is concurrently read or written by a different worker. The latter, *distributed memory* type needs direct communication for every exchange of data or instructions. The computations carried out for this thesis were performed on a cluster of 164 machines that did not share main memory. The machines in turn consisted of 4 to 16 computing cores with shared memory, totaling in 1000 CPUs, of which approximately 300 CPUs could be reserved on average for this work. Therefore, a mixture of distributed and shared memory parallelization was used in a two-stage approach (see figure B.1).

B.2.1 Shared memory parallelization

Shared memory parallelization was implemented using Python threads and OpenMP [104]. Python threads were managed using a queue data structure to store tasks and several instances of processing classes that worked through the queue carrying out the tasks. While Python has a global interpreter lock (gil) that blocks concurrent access of threads to python data structures, low-level routines written in C/C++ can release the gil (using the Python C-API macros `Py_BEGIN_ALLOW_THREADS` and `Py_END_ALLOW_THREADS`) in parts of the code that does not use Python data structures. Computation heavy parts were implemented in C/C++ such that no Python data structures were necessary and the gil could be released, achieving true parallelism. OpenMP was used for C/C++ code to parallelize for loops using the compiler directive `#pragma omp for` (see listing B.2).

```
import Queue
inputQueue = Queue.Queue( nr_of_threads + 2)
import threading
class correlateThread(threading.Thread):
  def __init__(self, correlationclass):
    threading.Thread.__init__(self)
    self.correlationclass = correlationclass
  def run(self):
    self.finished = False
    while not self.finished:
      try:
        data_in = inputQueue.get()
      if data_in is None:
        # finished with all tasks
        inputQueue.task_done()
      break
    except Queue.Empty:
      break
    img1, mask1 = data_in[0]
    ii = data_in[1]
    jj = data_in[2]
    img2 = self.correlationclass.getimage(jj)
    mask2 = self.correlationclass.getmask(jj)
    cormat = self.correlationclass.cormat
    cormat[ii,jj] = self.correlationclass.image_
                         crossCorrelation(img1, img2, mask1, mask2)
    cormat[jj,ii] = cormat[ii,jj]
    inputQueue.task_done()
```

Listing B.1. A Python thread class that carries out tasks that are deposited in a queue (`inputQueue`)

```
#import <math.h>
#pragma omp for schedule (dynamic)
for (int aa=0; aa<npoints; ++aa)
  for (int bb=0; bb<npoints; ++bb)
    RESULT2(aa,bb) = sqrt( (VECTORS2(aa,0)-VECTORS2(
    bb,0))*(VECTORS2(aa,0)-VECTORS2(bb,0)) ) }
```

Listing B.2. Example code showing the simple parallelizations of loops with OpenMP [104].

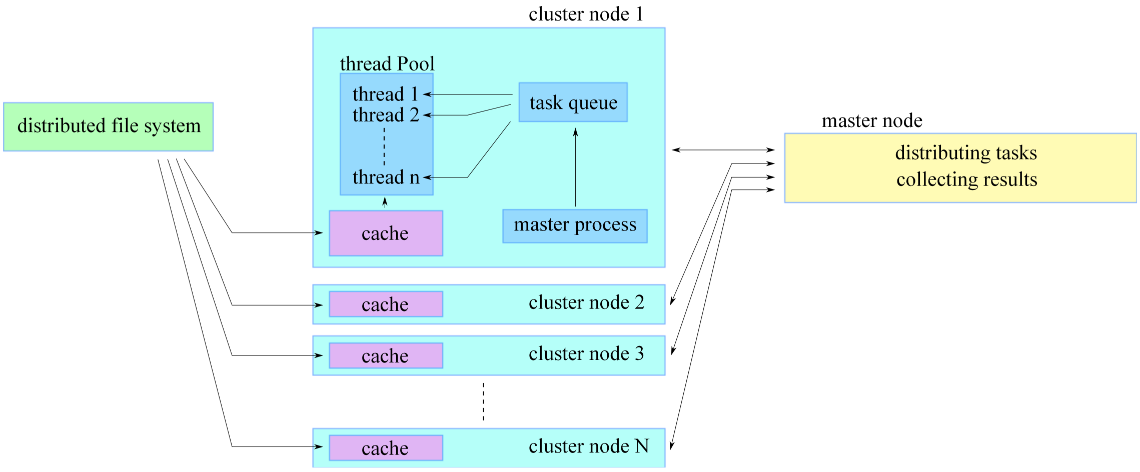


Fig. B.1. Two level parallelization architecture for computer clusters that consist of nodes with multiple processor cores.

B.2.2 Distributed memory parallellization

In order to harvest the big computational power of large computer clusters that do not use a single shared main memory, parallelization has to be managed by sending messages between the machines. Commands, input data and results are communicated with such messages. The implementation developed for this thesis uses *Open MPI*, a high performance message passing library. Open MPI provides an implementation of the standardized message passing interface, providing synchronization and communication between different processes on different machines. One of the cluster machines was designated as a master node that distributes tasks and collects results. The computational bottleneck of GIPRAL is the calculation of the dissimilarity distance matrix of pairs of diffraction patterns. A big contribution to the computation cost is slow loading of diffraction patterns from disk into memory. Smart caching was used to speed this up. Each node caches diffraction patterns in memory and computes all possible distance pairs for the cached diffraction patterns. In order to avoid unnecessary reloading, the parallel tasks are scheduled such that the cache of one machine can be reused in big parts for the next task (see figure B.2). Orchestrating the machines accordingly minimizes the amount of reloading to a large extend and the computational time is reduced tremendously. The computational time needed for the computation of the distance matrix of 50000 diffraction patterns went down from a half a year (naive implementation without parallelization and optimizations) to just a couple of days when a parallel implementation with smart caching was used.

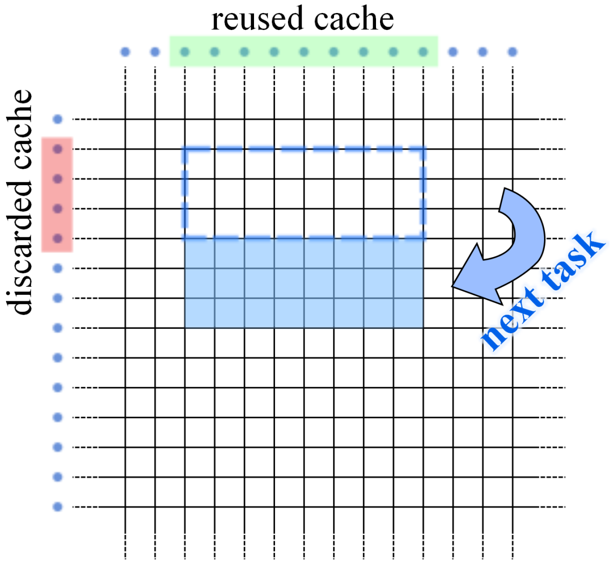


Fig. B.2. Parallel calculation of a distance matrix. A machine can re-use parts of the diffraction pattern cache if the tasks of computing sub-matrices are distributed accordingly. Carefully tuning the shape of the sub-matrices can lead to huge speed-ups.

The following listing shows the definition of a virtual base class that was used to manage cluster nodes. Classes can be derived from this base class, overwriting the worker functions to implement specific calculations.

```
from mpi4py import MPI
mpi_comm = MPI.COMM_WORLD
```

```
mpi_rank = mpi_comm.Get_Rank()
mpi_size = mpi_comm.Get_size()

class MPI_Pool:
  '''general class to distribute tasks
     to an Open MPI worker pool.'''
  tag_work = 0
  tag_stop = 1
  # the following worker functions have to be specialized
  # by derived classes
  def work_init(self, *initargs):
    # gets called before the nodes start working
    self.work_size = 0
  def work_finalize(self):
    # gets called after every node is finished
  def work_doWork(self, index, work_data):
    # does the actual work
  def work_checkInWorkResult(self, index, result):
    # gets called after a work chunk is finished
  def work_getWorkChunks(self, index):
    # splits work into chunks, returns chunk
    # number ‘index’
  # end of worker functions

  def process(self, *initargs, **keyargs):
    # main function of a computer node
    self.work_init(*initargs, **keyargs)
    if self.mpi_rank == 0:  # use node number 0 as master node
      self._master()
```

```
      result = self.work_getResult()
    else:
      self._slave()
      result = None
    self.work_finalize()
    return result

  def _master(self):
    work_index = 0
    num_completed = 0
    # fill all nodes with work
    for ii in range(1, min(self.mpi_size, self.work_size+1)):
      work_data = self.work_getWorkChunk(work_index)
      self.mpi_comm.send(work_index, ii, tag=self.tag_work)
      self.mpi_comm.send(work_data, ii)
      work_index += 1
    # receive results and distribute remaining work:
    for ii in range(self.mpi_size, self.work_size+1):
      result = self.mpi_comm.recv(source=MPI.ANY_SOURCE,
                                     tag=self.tag_work)
      result_index = status.tag
      self.work_checkinWorkResult(result_index, result)
      num_completed += 1
      proc = status.source
      work_data = self.work_getWorkChunk(work_index)
      self.mpi_comm.send(work_index, proce, tag=self.tag_work)
      self.mpi_comm.send(work_data, proc)
      work_index += 1
    # collect remaining results
    while num_completed < self.work_size:
```

```
      result = self.mpi_comm.recv(source=MPI.ANY_SOURCE,
                                     tag=MPI.ANY_TAG,
                                     status=status)
      result_index = status.tag
      self.work_checkinWorkResult(result_index, result)
      num_completed += 1
    # shut down workers:
    for proc in range(1, mpi_size):
      self.mpi_comm.send(-1, proc, tag=self.tag_stop)

  def _slave(self):
    status = MPI.Status()
    continue_working = True
    while continue_working:
      work_index = self.mpi_comm.recv(source=0, tag=MPI.ANY_TAG,
                                          status=status)
      if status.tag == self.tag_stop:
        continue_working = False
      else:
        work_data = self.mpi_comm.recv(source=0, tag=MPI.ANY_TAG,
                                          status=status)
        result = self.work_doWork(work_index, work_data)
        self.mpi_comm.send(result, dest=0, tag=work_index)
```

B.3 Class hierarchies

In order to make the implementation re-usable for many different data sources such as different detectors as well as simulated data encoded in different data formats, a flexible class hierarchy was used. This way, the algorithms can access diffraction snapshots by means of a unified interface.

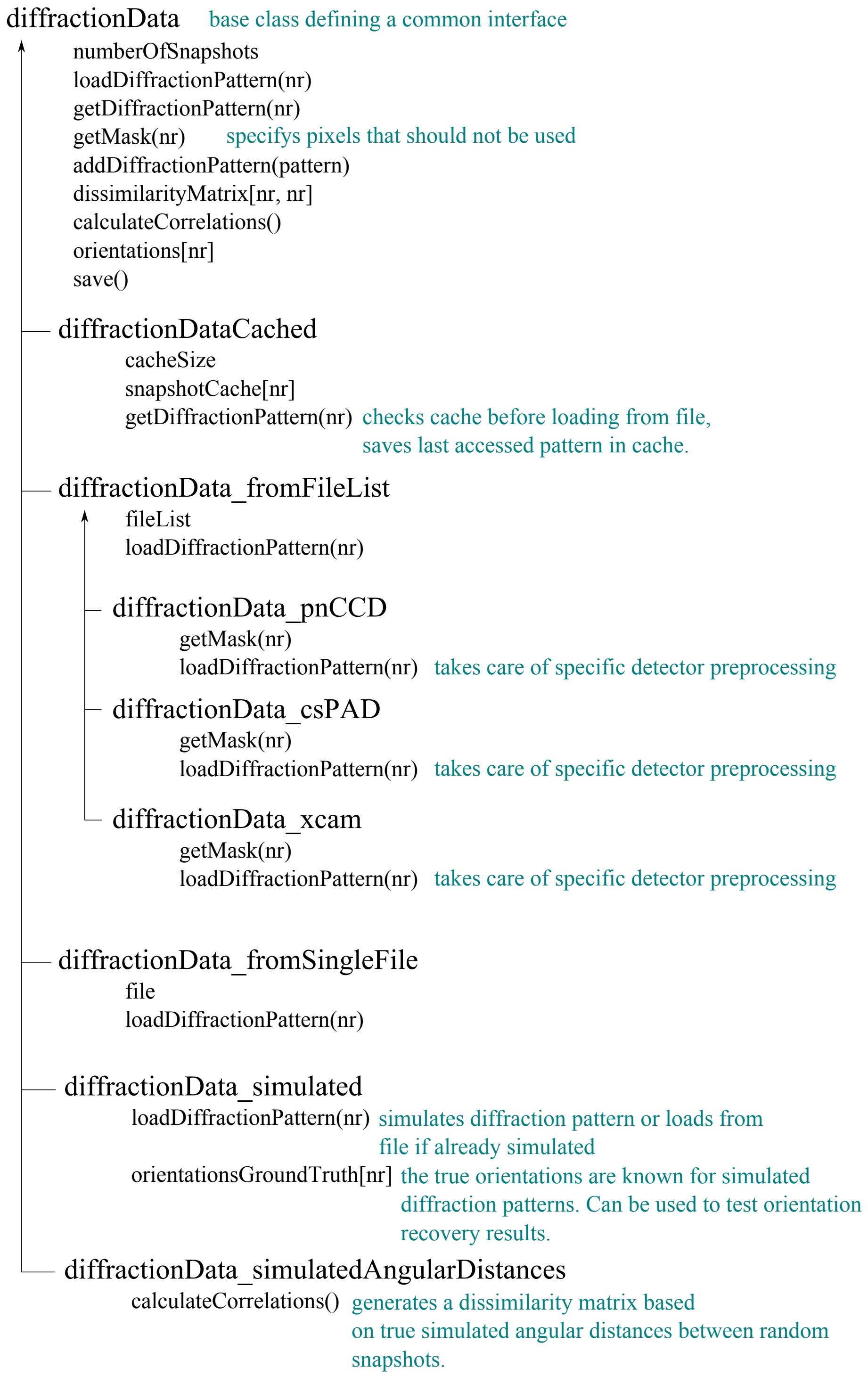


Fig. B.3. Class diagram showing inheritance relationships and specialized class members. This hierarchy provides a unified interface to diffraction data.

Appendix C

Mathematical Tools

C.1 Rodrigues Frank parametrization

Geodesics of objects with rotational symmetries can be treated elegantly in the Rodrigues/Frank (RF) parametrization. RF parametrization is a mapping from SO(3) to $\mathbb{R}^3$. $\mathbb{R}^3$ is not a natural space for rotations, because it does not reflect the curvature of SO(3). In RF space, this problem is addressed by "flattening" out the round structure that is inherent to rotations by mapping rotations of 180° to infinity via a factor $\tan\left(\frac{\alpha}{2}\right)$. The effect is that each circle representing a rotation has infinite curvature radius and is hereby flattened. RF space can be seen as a *gnomonic* projection of quaternions to Euclidean space. At the price of the nonlinear mapping $\tan\left(\frac{\alpha}{2}\right)$ comes a very nice property of RF space: the aforementioned flattening transforms geodesic lines into straight lines and the boundaries of Voronoi cells into planes. A rotation defined by a Euler axis $\hat{\boldsymbol{e}}$ and angle ϑ can be expressed as a RF vector

$$\boldsymbol{v} \; = \; \hat{\boldsymbol{e}} \tan\left(\frac{\vartheta}{2}\right). \tag{C.1}$$

As shown in [46] a rotation $\boldsymbol{r}_1$ followed by a rotation $\boldsymbol{r}_2$ then takes the form

$$\boldsymbol{r}_1 \circ \boldsymbol{r}_2 \; = \; \frac{\boldsymbol{r}_1 + \boldsymbol{r}_2 - \boldsymbol{r}_1 \times \boldsymbol{r}_2}{1 - \boldsymbol{r}_1 \cdot \boldsymbol{r}_2}. \tag{C.2}$$

From the definition C.1 it is clear that geodesic movements that start from the reference orientation (the origin in RF space) are straight lines in RF space, because they are rotations about a single, fixed axis. This fixed axis defines the direction of the RF vector and the angle modulates the length. From C.1 it can be seen that geodesics are straight lines, even if the reference orientation is changed by applying a rotation to a new reference orientation first; see [49]. This means that all geodesic curves are straight lines in RF space.

The proof of full orientation coverage of Eq. 3.2 can now be performed using the RF parametrization. Since the orientation of P_1 is used as the reference orientation, its RF vector can be found at the origin $\mathrm{RF}(P_1) = (0,\,0,\,0)^T$ (this means that no rotation is necessary to reach the orientation of P_1 from the reference orientation). The in-plane rotations of P_1 are single-axis rotations and therefore geodesics, thus the points $\mathrm{RF}(P_1(\alpha))$ lie on a straight line $\boldsymbol{l}_1$ and the points $\mathrm{RF}(P_2(\beta))$ describe a straight line $\boldsymbol{l}_2$. The geodesics $P_{\alpha,\beta,\varphi}$ between $P_1(\alpha)$ and $P_2(\beta)$ are then the straight-line segments $\boldsymbol{l}_{\mathrm{geo}\,\alpha,\beta}$ that start at $\boldsymbol{l}_1$ and end at $\boldsymbol{l}_2$. Since we are free to choose any combination of α and β, every combination of start and end points on $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ is possible, and the possible geodesics $\boldsymbol{l}_{\mathrm{geo}}$ fill the convex hull of $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ (see fig. C.1 for an illustration).

The out-of-plane rotation axis $\vec{\boldsymbol{a}}$ is orthogonal to the X-ray axis $\vec{\boldsymbol{c}}$, therefore a parametric representation of $\boldsymbol{l}_2$ is $\boldsymbol{l}_2(t) = \frac{\tan(\theta/2)\,\vec{\boldsymbol{a}} + \tan(t\,\alpha/2)\,\vec{\boldsymbol{c}} - \vec{\boldsymbol{d}}(t)}{1}$, where $\vec{\boldsymbol{d}}(t) \sim \vec{\boldsymbol{a}} \times \vec{\boldsymbol{c}}$ is perpendicular to both $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{c}}$. $\mathrm{RF}(P_2) = \tan(\theta/2) \cdot \vec{\boldsymbol{a}}$, so $\boldsymbol{l}_2(t) = \mathrm{RF}(P_2) + \vec{\boldsymbol{g}}(t)$, where $\vec{\boldsymbol{g}}(t)$ is orthogonal to $\mathrm{RF}(P_2)$, so the distance between the origin and $\boldsymbol{l}_2$ is $d = |\mathrm{RF}(P_2)| = \tan(\theta/2)$. $\boldsymbol{l}_1$ is parallel to $\vec{\boldsymbol{c}}$, so $\mathrm{RF}(P_2)$ is also orthogonal to $\boldsymbol{l}_1$ and since $\boldsymbol{l}_1$ contains the origin and $\boldsymbol{l}_2$ contains $\mathrm{RF}(P_2)$, d is the distance between $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ with $d \rightarrow \infty$ for $\theta \rightarrow 180°$.

As stated earlier, the convex hull of $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ contains all RF vectors that can be reached by a combination of out-of-plane geodesics and in-plane rotations. The boundaries of the convex hull of two infinite lines is given by two planes whose normals are orthogonal to both lines. The distance of these planes is the distance of the lines, and since $d \rightarrow \infty$, the half-space that is cut out of RF space by a plane that includes the origin is the space of all rotations that can be reached by patterns $P_{\alpha,\beta,\varphi}$. This half-space is sufficient to cover the full orientations space, since the other half represents equivalent rotations with inverted axes and angles.

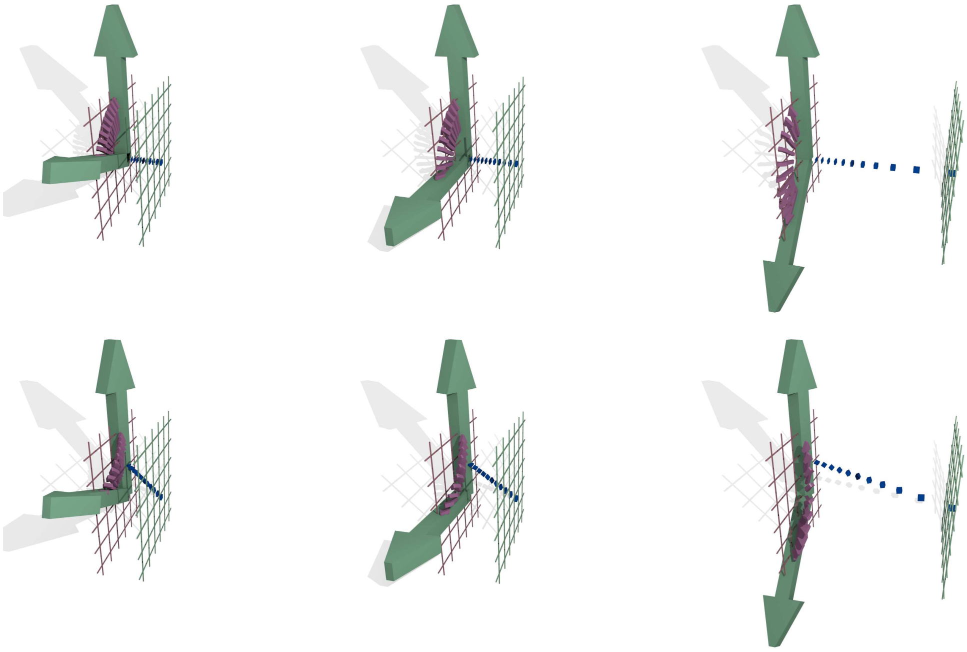


Fig. C.1. Illustration of geodesics in Rodrigues/Frank space. Purple arrows: geodesic sequence in real space which connect the start and end orientation (green arrows), blue squares: corresponding points in Rodrigues space. The grids indicate the planar borders of the convex hull of possible geodesics after adding in-plane rotations to the start and end orientations. For a separation $\theta = 180°$ between start and end orientation, the planes are infinitely far apart. The bottom row is the same as the top row except an in-plane rotation of the start orientation has been added. The green arrow pointing upwards depicts the start orientation. It points along the X-ray direction. For an interactive graphic, see www.gipral.org.

C.2 Object symmetries in Rodrigues-Frank space

The geodesic analysis of GIPRAL is based on a diffraction pattern distance which is subject to the rotational specimen symmetry. Only the asymmetric unit can be explored, like wave vectors in a crystal that always reside in the first Brillouin zone. The analysis of in-plane rotations does not underlay this restriction, since the true angular distance measure can be used here. This has implications for the geodesic paths which GIPRAL identifies as shortest paths. The longest of these can only span half of the maximum angle which is irreducible under the object's symmetry. If additional diffraction patterns are added to the longest geodesic pattern sequence, there will be a different sequence that acts as a shortcut to the additional diffraction patterns to which they will then be attributed instead. In principle, there are ways to find longer geodesic sequences, but the notion of shortest paths is simple, robust and efficient.

The fact that the "longest of all shortest" paths corresponds to a rotation of half the maximum irreducible angle can be used to calibrate the diffraction pattern distance to an angular distance. Therefore, the maximum possible object rotation within the fundamental zone of the symmetry has to be known. This is similar to the maximum misorientation angle used in crystallographic texture analysis, for which RF space has proved to be an elegant tool ([46],[50]).

Finite rotation groups can be classified into just a handfull of classes: a finite subgroup of SO(3) is isomorphic to either a cycle group, a dihedral group or one of the groups of regular (platonic) solids [105]. This result allows to categorize the fundamental RF cells of all possible symmetry classes. A list can be found in [50]. The maximum angle can be found as the longest RF vector within these cells, in the case of finite symmetry groups this is the RF vector of the cell vertices. Geometrically, it can be seen that the maximum angle is unique in the sense that it corresponds to rotation axes which are equivalent in terms of the symmetry operations. Thus, by identifying the "longest shortest paths", not only angles, but also the orientation of the rotation axis with respect to the symmetry axes of the specimen are determined. Due to the symmetry, there is a degeneracy of the maximum angle, since a rotation by an angle ω of the specimen around the maximal-angle-axis $\vec{\boldsymbol{d}}$ does not change the maximum geodesic distance. The axis $\vec{\boldsymbol{d}}$ restricts $\vec{\boldsymbol{c}}$ to the intersection $\boldsymbol{I}$ of the fundamental cell with a plane perpendicular to $\vec{\boldsymbol{d}}$ (because $\vec{\boldsymbol{c}} \perp \vec{\boldsymbol{d}}$). Within this plane the angle ω can be inferred with additional constraints: in-plane rotations are not restricted by the symmetry and thus the size of the fundamental cell in the direction of the in-plane rotations can be determined. The in-plane axis $\vec{\boldsymbol{c}}$ lies within the planar region $\boldsymbol{I}$ and the point p_{border} where it touches the border of the fundamental cell reveals the orientation of $\vec{\boldsymbol{c}}$ within $\boldsymbol{I}$ and can be used to obtain ω. The geodesic distance will be modulated by in-plane rotations such that jumps occur when in-plane rotations push $\boldsymbol{q}_{P_1(\alpha)}$ (or $\boldsymbol{q}_{P_2(\beta)}$) over the boundaries of the fundamental cell. These jumps can be used to identify p_{border}. Therefore, the orientation of both $\vec{\boldsymbol{d}}$ and $\vec{\boldsymbol{c}}$ with respect to the fundamental cell can be obtained and the diffraction snapshot orientations can be related to the object's symmetry axes.

When geodesics are identified as shortest paths the corresponding out-of-plane angle is restricted to only half of the maximum possible object rotation. Therefore the completeness of orientation coverage depends on the symmetry. However, because in-plane angles are not affected by the symmetry, the effect is not that severe. In the case of the dihedral symmetry of nanorice (see chapter 5.1.3), complete coverage can be achieved. As an example of higher symmetries, I numerically identified the possible coverage to be 92% for icosahedral symmetry (see figure S C.2). This coverage can be increased to 100% when results with different spanning patterns P_1 and P_2 are merged. Icosahedral symmetry is very common in nature and is of high importance for biological samples such as viruses.

Fig. C.2. Fundamental cell in Rodrigues space for icosahedral symmetry (dodecahedron). The blue part can be reached by a combination of out-of-plane geodesics and in-plane rotations of P_1 and P_2 in one go. Further iterations can then fill the whole fundamental cell. Due to the nonlinear deformation of Rodrigues space the gaps at the corner of the fundamental cell appear large but correspond to very small angular regions. In fact the blue region corresponds to 92% of all possible orientations.

C.3 extending geodesics

The orientation recovery of symmetric objects using the geodesic approach can be improved by eliminating the effect of the symmetry on the geodesic sequences. In the algorithm proposed in chapter 3.3, geodesic sequences are found by looking for shortest paths. Without any object symmetry, the maximum length of such a path corresponds to a rotation of 180°, which is reduced by symmetries. This reduction has to be taken into account when interpreting the geodesic sequences, which is somewhat alleviated by the RF formalism but still limits the number of orientations which are reachable by the geodesic search in one go. The in-plane rotations do not suffer from this limitation, they can be interpreted regardless of the object symmetry. In principle, this can also be achieved for the out-of-plane rotations if full geodesic paths can be found instead of just the sub-path that corresponds to the shortest path between two snapshots. Figure C.2 shows that this limitation is not severe, and can be worked around by repeated execution of the algorithm, but it is also possible to find full geodesic sequences. One solution would be to find geodesics based on local properties, but this works against a main advantage of the proposed method: finding shortest paths can be achieved efficiently in a globally optimal way. Mixing the global search with a more local search is possible with the following idea: the geodesic sub-paths found by a global search for shortest paths can be extended by other sub-paths as illustrated in fig. C.3: sub-paths can be patched together if they display a big intersecting part.

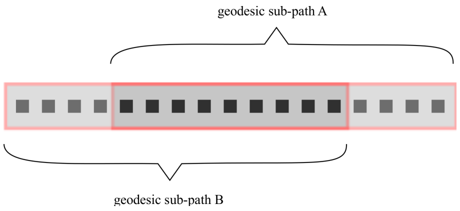


Fig. C.3. patching geodesic sub-paths A and B

The resulting full geodesic rings still reflect the objects symmetry, but they fill the gaps in the space of orientations that can be reached in one go.

Another method to fill these gaps is the repeated execution of the geodesic algorithm, picking new start end endpoints $\hat{P}_1$ and $\hat{P}_2$ from already reached orientations (blue part in fig. C.2) will result in geodesics between $\hat{P}_1(\alpha)$ and $\hat{P}_2(\beta)$ that protrude into previously unreached orientations.

C.4 Projections and mirror symmetry

Suppose the object under consideration is symmetric under a mirror operation $\boldsymbol{M}$ and the object orientation $\boldsymbol{O}$ is such that the X-ray axis coincides with the normal of the mirror plane. If the image formation process is approximated by a parallel projection $\boldsymbol{P}$ along the X-ray axis, it follows $\boldsymbol{P}p = \boldsymbol{P}\boldsymbol{M}p$ for every point p. Suppose

$\boldsymbol{R}^+$ is a rotation whose axis of rotation lies within the mirror plane of $\boldsymbol{M}$ and $\boldsymbol{R}^- := \boldsymbol{R}^{+^{-1}}$ is the inverse rotation (see figure S C.4). Then $\boldsymbol{R}^+ p = \boldsymbol{R}^+ \boldsymbol{M} \boldsymbol{M} p = \boldsymbol{M} \boldsymbol{R}^- \boldsymbol{M} p$ because mirroring inverts the rotation direction. It follows that $\boldsymbol{P}\boldsymbol{R}^+ p = \boldsymbol{P}\boldsymbol{M}\boldsymbol{R}^- \boldsymbol{M} p = \boldsymbol{P}\boldsymbol{R}^- \boldsymbol{M} p$. Applied to the set of object points, we can neglect the mirroring operation due to the symmetry $\Rightarrow \boldsymbol{P}\boldsymbol{R}^+ p = \boldsymbol{P}\boldsymbol{R}^- p$. Thus, starting from the orientation $\boldsymbol{O}$, the sequence of diffraction patterns obtained by intermediates of the rotation $\boldsymbol{R}^+$ will be the same as for intermediates of its inverse, $\boldsymbol{R}^-$. Inversely, the sequence starting from $[\boldsymbol{O}, \boldsymbol{R}^+]$ going to $[\boldsymbol{O}, \mathrm{Id}]$ gives a sequence of inverted element order compared to the sequence going from $[\boldsymbol{O}, \mathrm{Id}]$ to $[\boldsymbol{O}, \boldsymbol{R}^-]$. Only a single rotation axis is involved and thus the full sequence from $[\boldsymbol{O}, \boldsymbol{R}^+]$ to $[\boldsymbol{O}, \boldsymbol{R}^-]$ is of geodesic nature. However, the diffraction pattern based distance will increase until $[\boldsymbol{O}, \mathrm{Id}]$ is reached and then decrease until it falls to zero when $[\boldsymbol{O}, \boldsymbol{R}^-]$ is reached. When we search for the maximum geodesic sequence starting from $[\boldsymbol{O}, \boldsymbol{R}^+]$, then $[\boldsymbol{O}, \mathrm{Id}]$ will be the end of the found sequence when only diffraction pattern based similarities are taken into account. This means that the geodesic sequences tend to end at mirror axes as shown in figure C.4. This can be used to identify the symmetry of the object, as stated in the main text.

A stop at mirror axes means that geodesic sequences might become very short, depending on the the proximity of the start orientation $\boldsymbol{S}_0$ to mirror axes. But the sequence can be extended afterwards by making the orientation of the stop the new start orientation. The search for long geodesic sequences then yields sequences of maximal length.

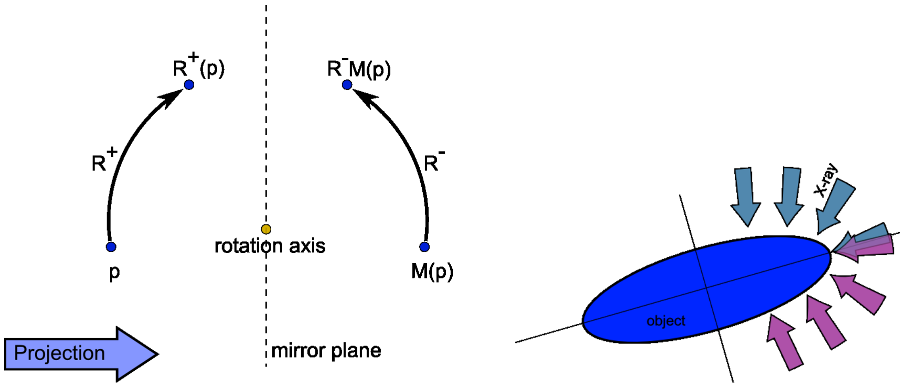


Fig. C.4. Mirror symmetry together with a projection operation leads to symmetry in rotation such that rotations in positive and negative direction yield the same projection. The diffraction pattern based geodesic sequence depicted by red arrows on the right side is equivalent to the blue sequence and it therefore stops at the mirror plane. It does not continue to the blue arrows as it would without symmetry.

C.5 Orthogonalizing in-plane and out-of-plane rotations

In the following considerations we will use a coordinate system that is fixed to the sample object. So instead of considering orientations of the sample we consider orientations of the x-ray beam and the detector. This implies that the in-plane axis changes from shot to shot. Three non collinear points are sufficient to represent orientations. Since we don't consider translations, all orientations are related by rotation axes that have one point in common: the origin $\mathcal{O}$ which is used as the first reference point r_0. r_0 is invariant for all shots. We choose the second reference point r_1 as the unit vector pointing along the x-ray beam and the third reference point r_2 is a point on the detector that does not coincide with the x-ray beam. The shortest (in an angular sense) rotation $\boldsymbol{R}_1$ that rotates r_1 of one shot to r_1' of a different shot is a rotation about an axis $\vec{\boldsymbol{e}}$ that is perpendicular to the x-ray beams of both shots: $\vec{\boldsymbol{e}} \perp \overline{r_0 r_1}$, $\vec{\boldsymbol{e}} \perp \overline{r_0 r_1'}$. After applying $\boldsymbol{R}_1$, the two reference points $\boldsymbol{R}_1 \cdot r_0 = r_0 = r_0' = \mathcal{O}$ and $\boldsymbol{R}_1 \cdot r_1 = r_1'$ are aligned to the new orientation. To complete the rotation to the new orientation, r_2 also has to be rotated to r_2' by a rotation $\boldsymbol{R}_2$. $\boldsymbol{R}_2$ needs to leave r_0 and r_1' invariant, so its rotation axis is $\overline{r_0 r_1'}$ which is the new x-ray axis of the second shot, meaning $\boldsymbol{R}_2$ describes an in-plane rotation. Thus the complete relative rotation between the two shots is $\boldsymbol{R} = \boldsymbol{R}_2 \circ \boldsymbol{R}_1$. By construction no rotation that rotates r_1 to r_1' can be shorter than $\boldsymbol{R}_1$, meaning that the angle of the composition $\boldsymbol{R}$ is minimized when $\boldsymbol{R}_2 = \mathbb{1}$. Since $\boldsymbol{R}_2$ is a in-plane rotation, an in-plane-rotation of the second shot can be found such that upon replacing the shot with its in-plane rotation, $\boldsymbol{R} = \mathbb{1} \cdot \boldsymbol{R}_1$. Thus finding the shortest rotation between the two shots while allowing cost-free in-plane-rotations of the second shot will yield a true out-of-plane rotation $\boldsymbol{R}_1$ with an axis that is orthogonal to the x-ray beam.

C.6 Discontinued: neighborhood preserving embedding

The following section describes an orientation recovery method that was developed initially for this thesis but was then abandoned in favor of the geodesic method described in chapter 3.

The orientation recovery problem can be formulated in terms of dimensionality reduction: the nominal dimensionality in detector space for a diffraction pattern is in the millions (one dimension for the intensity values of every pixel), but the intrinsic dimensionality is just three, because the only thing that differs between different diffraction snapshots is the orientation of the sample object which can be described by three Euler-angles. Embedding the diffraction snapshots in this three-dimensional orientation space would solve the orientation problem. Two things are known about this embedding:

1. the orientation space is "round" since orientations can be wrapped around by rotations through an angle of 2π. A natural choice for the target space of the embedding is therefore the three-dimensional surface of the hypersphere S^3.
2. The similarities should be preserved. If two diffraction snapshots are similar in the high dimensional detector space, they are also close by in orientation space.

This leads to the following embedding procedure: given the dissimilarity measure $\delta_{i,j} = d_{\text{diss}}(P_i, P_j)$ in detector space, N diffraction snapshots are embedded into a four dimensional Euclidean space while being constrained to the surface of a three-sphere C^3. Their positions $(x_1, ..., x_N)$ are thereby chosen to minimize the deviation of pairwise distances $\|x_i - x_j\|$ from the dissimilarity measure $\delta_{i,j}$:

$$\min_{x_1,...,x_N} \sum_{i<j} (\|x_i - x_j\| - \delta_{i,j})^2.$$

To account for geometric differences between the two metric spaces, the radius r of the three-sphere C^3 is chosen such that its hyper-area $2\,\pi^2\,r^3$ can be covered uniformly with N points such that the average distance between closest neighbors corresponds to the average minimum distance δ_{min} between all pairs of diffraction snapshots:

$$r \;=\; \sqrt[3]{\frac{N\,\delta_{\text{min}}}{4\,\pi^2}}.$$

The three-sphere is a double-cover of the orientation space (see section 1.3) To solve the optimization problem, a classic multidimensional scaling algorithm [106] as well as a gradient descent similar to simulated annealing [107] were used. Another approach was the use of a simulated physical model in which the data points where connected pairwise by springs the length of which represent the dissimilarity distance of the given pair. With carefully adjusted damping, a random initial state relaxes to the solution while being constrained to the surface of the three-sphere.

Only local similarity distances are correlated to angular distances. Figure C.5 shows the effect of missing global distance information. GIPRAL solves this problem by introducing a global geodesic distance measure.

Regardless of the optimization algorithm, the orientation recovery of simulated data was only successful when no noise was added to the diffraction snapshots. Even moderate noise lead to a solution that was stuck in a local minimum of the objective function. Also it is not easy to visualize and control the intermediate steps of the algorithm. Both problems can be overcome with the geodesic method as it guarantees global optima and the one dimensional geodesic snapshot sequences are easy to visualize.

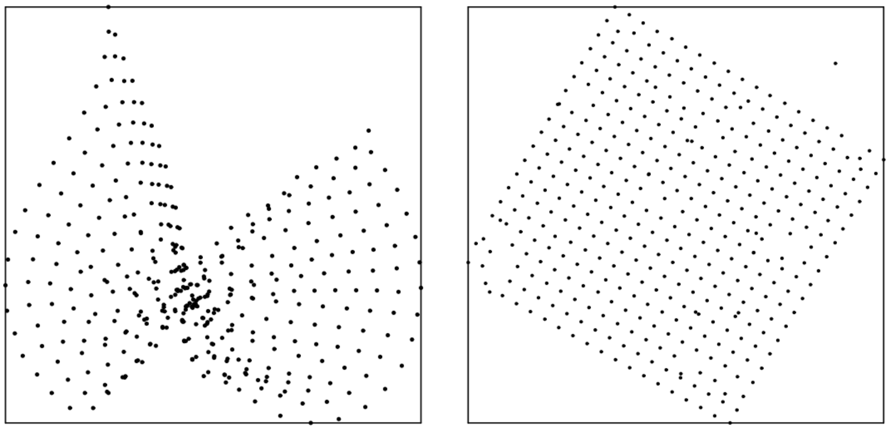


Fig. C.5. Global structure from local distances. Distances of points on a regular grid are fed into a multidimensional scaling algorithm. The left shows the resulting embedding into two-dimensional Euclidean space for the case where only distances of nearest neighbors were included, while the right images shows the result for distances in the 4-neighborhood with a result closer to the original regular grid.

Appendix D

Publications

My participation in several FEL projects resulted in publications listed in the following. My contributions include software development for data-acquisition, on-line supervision of the data quality and data analysis.

FEMTOSECOND FREE-ELECTRON LASER X-RAY DIFFRACTION DATA SETS FOR ALGORITHM DEVELOPMENT
Kassemeyer S, Steinbrener J, Lomb L, Hartmann E, Aquila A, Barty A, Martin AV, Hampton CY, Bajt S, Barthelmess M, Barends TR, Bostedt C, Bott M, Bozek JD, Coppola N, Cryle M, DePonte DP, Doak RB, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hartmann A, Hartmann R, Hauser G, Hirsemann H, Hömke A, Holl P, Jönsson O, Kimmel N, Krasniqi F, Liang M, Maia FR, Marchesini S, Nass K, Reich C, Rolles D, Rudek B, Rudenko A, Schmidt C, Schulz J, Shoeman RL, Sierra RG, Soltau H, Spence JC, Starodub D, Stellato F, Stern S, Stier G, Svenda M, Weidenspointner G, Weierstall U, White TA, Wunderer C, Frank M, Chapman HN, Ullrich J, Strüder L, Bogan MJ, Schlichting I.
Optics Express **20** 4, pp 4149-4158 (February 2012)

OPTIMAL MAPPING OF X-RAY LASER DIFFRACTION PATTERNS INTO THREE DIMENSIONS USING ROUTING ALGORITHMS
Kassemeyer S*, Jafarpour A, Lomb L, Steinbrener J, Martin AV, Schlichting I*.
Physical Review E **88** 4, p 042710 (October 2013)

* corresponding author

UNSUPERVISED CLASSIFICATION OF SINGLE-PARTICLE X-RAY DIFFRACTION SNAPSHOTS BY SPECTRAL CLUSTERING

Yoon CH, Schwander P, Abergel C, Andersson I, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Bozek J, Chapman HN, Claverie J, Coppola N, DePonte DP, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann E, Hartmann R, Hauser G, Hirsemann H, Holl P, **Kassemeyer S**, Kimmel N, Kiskinova M, Liang M, Loh ND, Lomb L, Maia FRNC, Martin AV, Nass K, Pedersoli E, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert M, Seltzer V, Shoeman RL, Sierra RG, Soltau H, Starodub D, Steinbrener J, Stier G, Strüder L, Svenda M, Ullrich J, Weidenspointner G, White TA, Wunderer C, Ourmazd A.

Optics Express **19** 17, pp 16542-16549 (August 2011)

CASS - CFEL-ASG SOFTWARE SUITE

Foucar L, Barty A, Coppola N, Hartmann R, Holl P, Hoppe U, **Kassemeyer S**, Kimmel N, Küpper J, Scholz M, Techert S, White TA, Strüder L, Ullrich J.

Computer Physics Communications **183** 10, pp. 2207-2213 (October 2012)

SINGLE-PARTICLE STRUCTURE DETERMINATION BY CORRELATIONS OF SNAPSHOT X-RAY DIFFRACTION PATTERNS

Starodub D, Aquila A, Bajt S, Barthelmess M, Barty A, Bostedt C, Bozek JD, Coppola N, Doak RB, Epp SW, Erk B, Foucar L, Gumprecht L, Hampton CY, Hartmann A, Hartmann R, Holl P, **Kassemeyer S**, Kimmel N, Laksmono H, Liang M, Loh ND, Lomb L, Martin AV, Nass K, Reich C, Rolles D, Rudek B, Rudenko A, Schulz J, Shoeman RL, Sierra RG, Soltau H, Steinbrener J, Stellato F, Stern S, Weidenspointner G, Frank M, Ullrich J, Strüder L, Schlichting I, Chapman HN, Spence JCH, Bogan MJ.

Nature Communications **3** article no 1276 (December 2012)

FRACTAL MORPHOLOGY, IMAGING AND MASS SPECTROMETRY OF SINGLE AEROSOL PARTICLES IN FLIGHT

Loh ND, Hampton Y, Martin V, Starodub D, Sierra G, Barty A, Aquila A, Schulz J, Lomb L, Steinbrener J, Shoeman RL, **Kassemeyer S**, Bostedt C, Bozek J, Epp

SW, Erk B, Hartmann R, Rolles D, Rudenko A, Rudek B, Foucar L, Kimmel N, Weidenspointner G, Hauser G, Holl P, Pedersoli E, Liang M, Hunter M, Gumprecht L, Coppola N, Wunderer C, Graafsma H, Maia FRNC, Ekeberg T, Hantke M, Fleckenstein H, Hirsemann H, Nass K, White TA, Tobias HJ, Farquar GR, Benner WH, Hau-Riege SP, Reich C, Hartmann A, Soltau H, Marchesini S, Bajt S, Barthelmess M, Bucksbaum P, Hodgson KO, Strüder L, Ullrich J, Frank M, Schlichting I, Chapman HN, Bogan MJ.
Nature **486** pp 513-517 (June 2012)

FEMTOSECOND DARK-FIELD IMAGING WITH AN X-RAY FREE ELECTRON LASER
Martin AV, Loh ND, Hampton CY, Sierra RG, Wang F, Aquila A, Bajt S, Barthelmess M, Bostedt C, Bozek JD, Coppola N, Epp SW, Erk B, Fleckenstein H, Foucar L, Frank M, Graafsma H, Gumprecht L, Hartmann A, Hartmann R, Hauser G, Hirsemann H, Holl P, **Kassemeyer S**, Kimmel N, Liang M, Lomb L, Maia FRNC, Marchesini S, Nass K, Pedersoli E, Reich C, Rolles D, Rudek B, Rudenko A, Schulz J, Shoeman RL, Soltau H, Starodub D, Steinbrener J, Stellato F, Strüder L, Ullrich J, Weidenspointner G, White TA, Wunderer CB, Barty A, Schlichting I, Bogan MJ, Chapman HN.
Optics Express **20** 12 pp 13501-13512 (May 2012)

SENSING THE WAVEFRONT OF X-RAY FREE-ELECTRON LASERS USING AEROSOL SPHERES
Loh ND, Starodub D, Lomb L, Hampton CY, Andrew V Martin AV, Sierra RG, Barty A, Aquila A, Schulz J, Steinbrener J, Shoeman RL, **Kassemeyer S**, Bostedt C, Bozek J, Epp SW, Erk B, Hartmann R, Rolles D, Rudenko A, Rudek B, Foucar L, Kimmel N, Weidenspointner G, Hauser G, Holl P, Pedersoli E, Liang M, Hunter MS, Gumprecht L, Coppola N, Wunderer C, Graafsma H, Maia FRNC, Ekeberg T, Hantke M, Fleckenstein H, Hirsemann H, Nass K, White TA, Tobias HJ, Farquar GR, Benner WH, Hau-Riege S, Reich C, Hartmann A, Soltau H, Marchesini S, Bajt S, Barthelmess M, Strüder L, Ullrich J, Bucksbaum P, Frank M, Schlichting I, Chapman HN, Bogan MJ.
Optics Express **21** 10, pp 12385-12394 (May 2013)

HIGH-RESOLUTION PROTEIN STRUCTURE DETERMINATION BY SERIAL FEMTOSECOND CRYSTALLOGRAPHY

Boutet S, Lomb L, Williams GJ, Barends TRM, Aquila A, Doak RB, Weierstall U, DePonte DP, Steinbrener J, Shoeman RL, Messerschmidt M, Barty A, White TA, **Kassemeyer S**, Kirian RA, Seibert MM, Montanez PA, Kenney C, Herbst R, Hart P, Pines J, Haller G, Gruner SM, Philipp HT, Tate MW, Hromalik M, Koerner LJ, Bakel Nv, Morse J, Ghonsalves W, Arnlund D, Bogan MJ, Caleman C, Fromme R, Hampton CY, Hunter MS, Johansson LC, Katona G, Kupitz C, Liang M, Martin AV, Nass K, Redecke L, Stellato F, Timneanu N, Wang D, Zatsepin NA, Schafer D, Defever J, Neutze R, Fromme P, Spence JCH, Chapman HN, Schlichting I.
Science **337** 6092, pp 362-364 (July 2012)

RADIATION DAMAGE IN PROTEIN SERIAL FEMTOSECOND CRYSTALLOGRAPHY USING AN X-RAY FREE-ELECTRON LASER

Lomb L, Barends TRM, **Kassemeyer S**, Aquila A, Epp SW, Erk B, Foucar L, Hartmann R, Rudek B, Rolles D, Rudenko A, Shoeman RL, Andreasson J, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Bozek JD, Caleman C, Coffee R, Coppola N, Deponte DP, Doak RB, Ekeberg T, Fleckenstein H, Fromme P, Gebhardt M, Graafsma H, Gumprecht L, Hampton CY, Hartmann A, Hauser G, Hirsemann H, Holl P, Holton JM, Hunter MS, Kabsch W, Kimmel N, Kirian RA, Liang M, Maia FRNC, Meinhart A, Marchesini S, Martin AV, Nass K, Reich C, Schulz J, Seibert MM, Sierra R, Soltau H, Spence JCH, Steinbrener J, Stellato F, Stern S, Timneanu N, Wang X, Weidenspointner G, Weierstall U, White TA, Wunderer C, Chapman HN, Ullrich J, Strüder L, Schlichting I.
Physical Review B **84** 21, p 214111 (December 2011)

SELF-TERMINATING DIFFRACTION GATES FEMTOSECOND X-RAY NANOCRYSTALLOGRAPHY MEASUREMENTS

Barty A, Caleman C, Aquila A, Timneanu N, Lomb L, White TA, Andreasson J, Arnlund D, Bajt S, Barends TRM, Barthelmess M, Bogan MJ, Bostedt C, Bozek JD, Coffee R, Coppola N, Davidsson J, Deponte DP, Doak RB, Ekeberg T, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Fromme P, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann R, Hartmann A, Hauser G, Hirsemann H, Holl

P, Hunter MS, Johansson L, **Kassemeyer S**, Kimmel N, Kirian RA, Liang M, Maia FRNC, Malmerberg E, Marchesini S, Martin AV, Nass K, Neutze R, Reich C, Rolles D, Rudek B, Rudenko A, Scott H, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Spence JCH, Stellato F, Stern S, Strüder L, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Wunderer CB, Chapman HN.
Nature Photonics **6** 1, pp 35-40 (December 2011)

Toward unsupervised single-shot diffractive imaging of heterogeneous particles using X-ray free-electron lasers
Park HJ, Loh ND, Sierra RG, Hampton CY, Starodub D, Martin AV, Barty A, Aquila A, Schulz J, Steinbrener J, Shoeman RL, Lomb L, **Kassemeyer S**, Bostedt C, Bozek J, Epp SW, Erk B, Hartmann R, Rolles D, Rudenko A, Rudek B, Foucar L, Kimmel N, Weidenspointner G, Hauser G, Holl P, Pedersoli E, Liang M, Hunter MS, Gumprecht L, Coppola N, Wunderer C, Graafsma H, Maia FRNC, Ekeberg T, Hantke M, Fleckenstein H, Hirsemann H, Nass K, Tobias HJ, Farquar GR, Benner WH, Hau-Riege S, Reich C, Hartmann A, Soltau H, Marchesini S, Bajt S, Barthelmess M, Strueder L, Ullrich J, Bucksbaum P, Frank M, Schlichting I, Chapman HN, Bogan MJ, Elser V.
Optics Express **21** 23, pp 28729-28742 (November 2013)

Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography
Johansson LC, Arnlund D, Katona G, White TA, Barty A, DePonte DP, Shoeman RL, Wickstrand C, Sharma A, Williams GJ, Aquila A, Bogan MJ, Caleman C, Davidsson J, Doak RB, Frank M, Fromme R, Galli L, Grotjohann I, Hunter MS, **Kassemeyer S**, Kirian RA, Kupitz C, Liang M, Lomb L, Malmerberg E, Martin AV, Messerschmidt M, Nass K, Redecke L, Seibert MM, Sjöhamn J, Steinbrener J, Stellato F, Wang D, Wahlgren WY, Weierstall U, Westenhoff S, Zatsepin NA, Boutet S, Spence JCH, Schlichting I, Chapman HN, Fromme P, Neutze R.
Nature Communications **4** article no 2911 (December 2013)

Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser
Barends TRM, Foucar L, Shoeman RL, Bari S, Epp SW, Hartmann R, Hauser

G, Huth M, Kieser C, Lomb L, Motomura K, Nagaya K, Schmidt C, Strecker R, Anielski D, Boll R, Erk B, Fukuzawa H, Hartmann E, Hatsui T, Holl P, Inubushi Y, Ishikawa T, **Kassemeyer S**, Kaiser C, Koeck F, Kunishima N, Kurka M, Rolles D, Rudek B, Rudenko A, Sato T, Schroeter CD, Soltau H, Strüder L, Tanaka T, Togashi T, Tono K, Ullrich J, Yase S, Wada SI, Yao M, Yabashi M, Ueda K, Schlichting I.
Acta Crystallographica Section D **69** 5 pp 838-842 (May 2013)

NATIVELY INHIBITED TRYPANOSOMA BRUCEI CATHEPSIN B STRUCTURE DETERMINED BY USING AN X-RAY LASER
Redecke L, Nass K, DePonte DP, White TA, Rehders D, Barty A, Stellato F, Liang M, Barends TRM, Boutet S, Williams GJ, Messerschmidt M, Seibert MM, Aquila A, Arnlund D, Bajt S, Barth T, Bogan MJ, Caleman C, Chao T, Doak RB, Fleckenstein H, Frank M, Fromme R, Galli L, Grotjohann I, Hunter MS, Johansson LC, **Kassemeyer S**, Katona G, Kirian RA, Koopmann R, Kupitz C, Lomb L, Martin AV, Mogk S, Neutze R, Shoeman RL, Steinbrener J, Timneanu N, Wang D, Weierstall U, Zatsepin NA, Spence JCH, Fromme P, Schlichting I, Duszenko M, Betzel C, Chapman HN.
Science **339** 6116 pp 227-230 (January 2013)

LIPIDIC PHASE MEMBRANE PROTEIN SERIAL FEMTOSECOND CRYSTALLOGRAPHY
Johansson LC, Arnlund D, White TA, Katona G, Deponte DP, Weierstall U, Doak RB, Shoeman RL, Lomb L, Malmerberg E, Davidsson J, Nass K, Liang M, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Bozek JD, Caleman C, Coffee R, Coppola N, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann R, Hartmann A, Hauser G, Hirsemann H, Holl P, Hunter MS, **Kassemeyer S**, Kimmel N, Kirian RA, Maia FRNC, Marchesini S, Martin AV, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert MM, Sierra RG, Soltau H, Starodub D, Stellato F, Stern S, Strüder L, Timneanu N, Ullrich J, Wahlgren WY, Wang X, Weidenspointner G, Wunderer C, Fromme P, Chapman HN, Spence JCH, Neutze R.
Nature Methods **9** 3 pp 264-265 (January 2012)

IN VIVO PROTEIN CRYSTALLIZATION OPENS NEW ROUTES IN STRUCTURAL BIOLOGY

Koopmann R, Cupelli K, Redecke L, Nass K, Deponte DP, White TA, Stellato F, Rehders D, Liang M, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Boutet S, Bozek JD, Caleman C, Coppola N, Davidsson J, Doak RB, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann R, Hauser G, Hirsemann H, Holl P, Hunter MS, **Kassemeyer S**, Kirian RA, Lomb L, Maia FRNC, Kimmel N, Martin AV, Messerschmidt M, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Stern S, Strüder L, Timneanu N, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Williams GJ, Wunderer CB, Fromme P, Spence JCH, Stehle T, Chapman HN, Betzel C, Duszenko M.

Nature Methods **9** 3 pp 259-262 (January 2012)

TIME-RESOLVED PROTEIN NANOCRYSTALLOGRAPHY USING AN X-RAY FREE-ELECTRON LASER

Aquila A, Hunter MS, Doak RB, Kirian RA, Fromme P, White TA, Andreasson J, Arnlund D, Bajt S, Barends TRM, Barthelmess M, Bogan MJ, Bostedt C, Bottin H, Bozek JD, Caleman C, Coppola N, Davidsson J, DePonte DP, Elser V, Epp SW, Erk B, Fleckenstein H, Foucar L, Frank M, Fromme R, Graafsma H, Grotjohann I, Gumprecht L, Hajdu J, Hampton CY, Hartmann A, Hartmann R, Hau-Riege S, Hauser G, Hirsemann H, Holl P, Holton JM, Hömke A, Johansson L, Kimmel N, **Kassemeyer S**, Krasniqi F, Kühnel K, Liang M, Lomb L, Malmerberg E, Marchesini S, Martin AV, Maia FRNC, Messerschmidt M, Nass K, Reich C, Neutze R, Rolles D, Rudek B, Rudenko A, Schlichting I, Schmidt C, Schmidt KE, Schulz J, Seibert MM, Shoeman RL, Sierra R, Soltau H, Starodub D, Stellato F, Stern S, Strüder L, Timneanu N, Ullrich J, Wang X, Williams GJ, Weidenspointner G, Weierstall U, Wunderer C, Barty A, Spence JCH, Chapman HN.

Optics Express **20** 3 pp 2706-2716 (January 2012)

Acknowledgments

When I look back on the last four years that I spent on the work resulting in this thesis, I can summarize the positive experience with the many people that accompanied me along the way in a simple sentence: I would gladly work with all of you again!

Particular thanks are due to my supervisors Dr. Robert Bittl and Dr. Ilme Schlichting. The topic of my thesis and the fact that the experiments were carried out abroad separated me from the rest of Dr. Bittl's group, but I did not feel like an outsider as I was always welcome in Berlin where I had many helpful discussions.

Ilme Schlichting has given me the opportunity to work on ambitious, large-scale FEL projects. These projects can be very chaotic at times, especially since so many people are involved. Ilme always manages to master even the most chaotic beam times by keeping a good overview and guiding people. However, she did not restrict herself to a bare managing role but put her heart and soul in all aspects of this research, which I find exceptional and inspiring. Thank you, Ilme, for your trust in me and for showing me how to stand up to such big projects!

Good ideas usually emerge from lots of crazy ideas before they mature and become useful. Heinrich von Kleist [108] wrote a brilliant essay in which he pointed out how important it is to talk to friends before a solution to a difficult problem can be found. The person who had to listen to all of these crazy and sometimes very half-baked ideas was my dear friend Andreas Lindenblatt. Thank you, Andi, for listening and for helping me shape my ideas, but also for being a trustworthy “Seilpartner” in all kinds of alpine expeditions and for fighting my “barbarian” side by strengthening the “gourmet” side.

I would like to thank my colleagues Lukas Lomb, Jan Steinbrener, Aliakbar Jafarpour, Karol Nass, Wolfgang Kabsch, Thomas Barends, Lutz Foucar and Sabine Botha.

Lukas was half-way through his PhD project when I started mine - he introduced me to coherent diffractive imaging with FELs and related practical and theoretical approaches, getting me hooked on the subject. Lukas, Jan and I also had a good time meeting for activities outside our scientific lives, strengthening a bond that was helpful during harsh sleepless beam times. Sabine, thank you for injecting samples and still finding time to jump into the freezing cold Pacific Ocean with me during busy beam times. Lutz, thank you for your support in processing huge amounts of data and thank you for generously tolerating some messy last-minute hacks that I introduced into your code during some beam-times. Ali and Wolfgang, you are really smart guys and you helped a lot by discussing theoretical aspects of diffractive imaging and data processing. Thomas, thanks for being someone who shows polymaths still exist, thank you for sharing your all-round expertise also in hobby projects and for encouraging me to etch my own printed circuit boards.

I would like to thank Robert Shoeman and Bruce Doak for a lot of work on sample handling and sample injection and for proof reading and improving my publications. Elisabeth Hartmann, Matthias Fischer and Kerstin-Anikó Seifert put in weeks of work in order to contribute biological samples which were then eaten up in just a few hours of beam-time, thank you!

Chris, thank you for your excellent IT support that was always incredibly fast. The same goes to Frank Köck, the system administrator of the huge computer cluster at the Max-Planck institute for nuclear physics who is always able to deal with the most bizarre error messages.

Mario Bott helped by looking through hundreds of diffraction patterns to check the quality of the automated classification and to supervise the automated fit of diffraction fringes used to obtain the size distribution of the sample objects. Thank you for helping with this tiring but important task.

Big thanks go to my fellow PhD students, Cathleen Zeymer, Anikó Udvarhelyi, Juliane Benz, Aytac Dikfidan, Kristina Haslinger and many others for a lot of fun activities outside the lab, playing sports and enjoying fruit smoothies and baking chocolate cakes in the institute's kitchen. Cathleen started writing up her thesis just shortly before I did, confidently demonstrating that this can be done without panicking. Thanks a lot for leading the way.

My parents somehow sparked my tremendous, never resting curiosity for the inner workings of our world and could not stop me from approaching even the most banal problems in unconventional ways. I love you!

I sincerely thank all tax payers who understand that fundamental research is not so much a luxury but rather a necessity for our society.

Index

References

1 JD Watson and FHC Crick. A structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

2 Elspeth F Garman and Robin Leslie Owen. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 62(1):32–47, 2005.

3 Jan Kern, Roberto Alonso-Mori, Rosalie Tran, Johan Hattne, Richard J Gildea, Nathaniel Echols, Carina Glöckner, Julia Hellmich, Hartawan Laksmono, Raymond G Sierra, et al. Simultaneous femtosecond x-ray spectroscopy and diffraction of photosystem ii at room temperature. *Science*, 340(6131):491–495, 2013.

4 Thorsten Beitlich, Karin Kuhnel, Clemens Schulze-Briese, Robert L Shoeman, and Ilme Schlichting. Cryoradiolytic reduction of crystalline heme proteins: analysis by uv-vis spectroscopy and x-ray crystallography. *Journal of synchrotron radiation*, 14(1):11–23, 2006.

5 Richard Henderson. The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly reviews of biophysics*, 28(02):171–193, 1995.

6 Richard Henderson. Realizing the potential of electron cryo-microscopy. *Quarterly reviews of biophysics*, 37(01):3–13, 2004.

7 Alexey Amunts, Alan Brown, Xiao-chen Bai, Jose L Llácer, Tanweer Hussain, Paul Emsley, Fei Long, Garib Murshudov, Sjors HW Scheres, and V Ramakrishnan. Structure of the yeast mitochondrial large ribosomal subunit. *Science*, 343(6178):1485–1489, 2014.

8 Werner Kühlbrandt. The resolution revolution. *Science*, 343(6178):1443–1444, 2014.

9 Johndale C. Solem. Imaging biological specimens with high-intensity soft x rays. *J. Opt. Soc. Am. B*, 3(11):1551–1565, Nov 1986.

10 R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu. Potential for biomolecular imaging with femtosecond x-ray pulses. *Nature*, 406:752–757, 2000.

11 Lukas Lomb, Thomas RM Barends, Stephan Kassemeyer, Andrew Aquila, Sascha W Epp, Benjamin Erk, Lutz Foucar, Robert Hartmann, Benedikt Rudek, Daniel Rolles, et al. Radiation damage in protein serial femtosecond crystallography using an x-ray free-electron laser. *Physical Review B*, 84(21):214111, 2011.

12 A. Barty, C. Caleman, A. Aquila, N. Timneanu, L. Lomb, T.A. White, J. Andreasson, D. Arnlund, S. Bajt, T.R.M. Barends, et al. Self-terminating diffraction gates femtosecond x-ray nanocrystallography measurements. *Nature Photonics*, 6(1):35–40, 2011.

13 S.P. Hau-Riege, H.N. Chapman, J. Krzywinski, R. Sobierajski, S. Bajt, R.A. London, M. Bergh, C. Caleman, R. Nietubyc, L. Juha, et al. Subnanometer-scale measurements of the interaction of ultrafast soft x-ray free-electron-laser pulses with matter. *Physical review letters*, 98(14):145502, 2007.

14 Carl Caleman, Gösta Huldt, Carlos Ortiz, Filipe R. N. C. Maia, Erik G. Marklund, Fritz G. Parak, David van der Spool, and Nicusor Timneanu. Feasibility of nanocrystal imaging using intense and ultrashort x-ray pulses, 2010.

15 H.N. Chapman, A. Barty, M.J. Bogan, S. Boutet, M. Frank, S.P. Hau-Riege, S. Marchesini, B.W. Woods, S. Bajt, W.H. Benner, et al. Femtosecond diffractive imaging with a soft-x-ray free-electron laser. *Nature Physics*, 2(12):839–843, 2006.

16 David Sayre. Some implications of a theorem due to shannon. *Acta Crystallographica*, 5(6):843–843, 1952.

17 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, Hays G., Ph. 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 Galayda J. First lasing and operation of an angstrom-wavelength free-electron laser. *Nature Photonics*, 4(9):641–647, 2010.

18 Michael J. Bogan, W. Henry Benner, Sébastien Boutet, Urs Rohner, Matthias Frank, Anton Barty, M. Marvin Seibert, Filipe Maia, Stefano Marchesini, SaÂŽa Bajt, Bruce Woods, Vincent Riot, Stefan P. Hau-Riege, Martin Svenda, Erik Marklund, Eberhard Spiller, Janos Hajdu, and Henry N. Chapman. Single particle x-ray diffractive imaging. *Nano Letters*, 8(1):310–316, 2008.

19 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), 2008.

20 Paul Kirkpatrick and Albert Vincio Baez. Formation of optical images by x-rays. *JOSA*, 38(9):766–773, 1948.

21 D. Paganin. *Coherent X-Ray Optics*. Oxford Series on Synchrotron Radiation. OUP Oxford, 2013.

22 Wolfgang Demtröder. *Experimentalphysik 2*. Experimentalphysik / Wolfgang Demtröder. Springer, 2008.

23 S. Hunklinger. *Festkörperphysik*. Oldenbourg, 2007.

24 Georges Friedel. Sur les symétries cristallines que peut révéler la diffraction des rayons röntgen. *CR Acad Sci*, 157:1533–1536, 1913.

25 Ronald N Bracewell. Fourier transform and its applications. 1980.

26 A. Papoulis. *The Fourier integral and its applications*. McGraw-Hill electronic sciences series. McGraw-Hill, 1962.

27 Claude Elwood Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

28 Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge Mathematical Library. Cambridge University Press, 2004.

29 J. Miao, D. Sayre, and H.N. Chapman. Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 15(6):1662–1669, 1998.

30 Filipe R. N. C. Maia. *Ultrafast Coherent X-ray Diffractive Nanoimaging*. PhD thesis, Uppsala University, Molecular biophysics, 2010.

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

32 James R Fienup. Reconstruction of an object from the modulus of its fourier transform. *Optics letters*, 3(1):27–29, 1978.

33 James R Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.

34 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.

35 X. Huang, J. Nelson, J. Steinbrener, J. Kirz, J.J. Turner, and C. Jacobsen. Incorrect support and missing center tolerances of phasing algorithms. *Optics express*, 18(25):26441, 2010.

36 J. Miao, Y. Nishino, Y. Kohmura, B. Johnson, C. Song, S.H. Risbud, and T. Ishikawa. Quantitative image reconstruction of gan quantum dots from oversampled diffraction intensities alone. *Physical review letters*, 95(8):85503, 2005.

37 Brian Moths and Abbas Ourmazd. Bayesian algorithms for recovering structure from single-particle diffraction snapshots of unknown orientation: a comparison. *Acta Cryst. A*, 67(5):481–486, 2011.

38 Paul Viola and Michael Jones. Robust real-time object detection. *International Journal of Computer Vision*, 4, 2001.

39 I.T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer, 2002.

40 Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. *Proc. 13th International Conference on Machine Learning*, pages 148–146, 1996.

41 Leo Breiman. Random forests. In *Machine Learning*, pages 5–32, 2001.

42 Ullrich Koethe. Generic programming for computer vision.

43 D. Giannakis, P. Schwander, and A. Ourmazd. The symmetries of image formation by scattering. i. theoretical framework. *Opt. Express*, 20(12):12799–12826, 2012.

44 Stephan Kassemeyer, Aliakbar Jafarpour, Lukas Lomb, Jan Steinbrener, Andrew V. Martin, and Ilme Schlichting. Optimal mapping of x-ray laser diffraction patterns into three dimensions using routing algorithms. *Phys. Rev. E*, 88:042710, Oct 2013.

45 E.W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.

46 F.C. Frank. Orientation mapping. *Metall. Trans. A*, 19(3):403–408, 1988.

47 Kassemeyer S., Steinbrener J., Lomb L., Hartmann E., Aquila A., Barty A., Martin A.V., Hampton C.Y., Bajt S., Barthelmess M., Barends T.R., Bostedt C., Bott M., Bozek J.D., Coppola N., Cryle M., DePonte D.P., Doak R.B., Epp S.W., Erk B., Fleckenstein H., Foucar L., Graafsma H., Gumprecht L., Hartmann A., Hartmann R., Hauser G., Hirsemann H., H"omke A., Holl P., J"onsson O., Kimmel N., Krasniqi F., Liang M., Maia F.R., Marchesini S., Nass K., Reich C., Rolles D., Rudek B., Rudenko A., Schmidt C., Schulz J., Shoeman R.L., Sierra R.G., Soltau H., Spence J.C.H., Starodub D., Stellato F., Stern S., Stier G., Svenda M., Weidenspointner G., Weierstall U., White T.A., Wunderer C., Frank M., Chapman H.N., Ullrich J., Str"uder L., Bogan M.J., and Schlichting I. Femtosecond free-electron laser x-ray diffraction data sets for algorithm development. *Opt. Express*, 20(4):4149–4158, Feb 2012.

48 Chun Hong Yoon, Peter Schwander, Chantal Abergel, Inger Andersson, Jakob Andreasson, Andrew Aquila, Saša Bajt, Miriam Barthelmess, Anton Barty, Michael J Bogan, et al. Unsupervised classification of single-particle x-ray diffraction snapshots by spectral clustering. *Optics Express*, 19(17):16542–16549, 2011.

49 A. Morawiec and D.P. Field. Rodrigues parameterization for orientation and misorientation distributions. *Philosophical Magazine A*, 73(4):1113–1130, 1996.

50 A. Morawiec. Misorientation-angle distribution of randomly oriented symmetric objects. *J. Appl. Cryst.*, 28(3):289–293, 1995.

51 Prasanta Chandra Mahalanobis. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2, 1936.

52 V. Pratt. Direct least-squares fitting of algebraic surfaces. *Computer Graphics*, 21:145–152, 1987.

53 A.V. Martin, A.J. Morgan, T. Ekeberg, N.D. Loh, F.R.N.C. Maia, F. Wang, J.C.H. Spence, and H.N. Chapman. The extraction of single-particle diffraction patterns from a multiple-particle diffraction pattern. *Opt. Express*, 21(13):15102–15112, Jul 2013.

54 Foucar L., Barty A., Coppola N., Hartmann R., Holl P., Hoppe U., Kassemeyer S., Kimmel N., Küpper J., Scholz M., Techert S., White T.A., Strüder L., and Ullrich J. Cass – cfel-asg software suite. *Computer Physics Communications*, 183(10):2207–2213, 2012.

55 L. Strüder, S. Epp, D. Rolles, R. Hartmann, P. Holl, G. Lutz, H. Soltau, R. Eckart, C. Reich, K. Heinzinger, C. Thamm, A. Rudenko, F. Krasniqi, K.-U. Kühnel, C. Bauer, C.-D. Schröter, R. Moshammer, S. Techert, D. Miessner, M. Porro, O. Hälker, N. Meidinger, N. Kimmel, R. Andritschke, F. Schopper, G. Weidenspointner, A. Ziegler, D. Pietschner, S. Herrmann, U. Pietsch, A. Walenta, W. Leitenberger, C. Bostedt, T. Möller, D. Rupp, H. Adolph, H. Graafsma, H. Hirsemann, K. Gärtner, R. Richter, L. Foucar, R.L. Shoeman, I. Schlichting, and J. Ullrich. Large-format, high-speed, x-ray pnccds combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 614(3):483–496, 2010.

56 E.F Eikenberry, Ch Brönnimann, G Hüulsen, H Toyokawa, R Horisberger, B Schmitt, C Schulze-Briese, and T Tomizaki. Pilatus: a two-dimensional x-ray detector for macromolecular crystallography. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 501(1):260 – 266, 2003. Proceedings of the 10th International Workshop on Vertex Detectors.

57 Dectris AG, Baden, Swizerland.

58 MarUSA/Rayonix LLC (Evanston, IL 60201 USA).

59 S. Herrmann, S. Boutet, B. Duda, D. Fritz, G. Haller, P. Hart, R. Herbst, C. Kenney, H. Lemke, M. Messerschmidt, J. Pines, A. Robert, M. Sikorski, and G. Williams. Cspad-140k: A versatile detector for lcls experiments. *Nuclear Instruments and Methods in Physics Research Section A*, 718(0):550 – 553, 2013. Proceedings of the 12th Pisa Meeting on Advanced Detectors La Biodola, Isola d'Elba, Italy, May 20 ÂŰ 26, 2012.

60 Lutz Foucar. Cass source code. http://www.mpi-hd.mpg.de/personalhomes/gitasg/Downloads/cass.latest.tar.gz, 2014.

61 V.L. Shneerson, A. Ourmazd, and D.K. Saldin. Crystallography without crystals. i. the common-line method for assembling a three-dimensional diffraction volume from single-particle scattering. *Acta Cryst. A*, 64(2):303–315, 2008.

62 R. Fung, V. Shneerson, D.K. Saldin, and A. Ourmazd. Structure from fleeting illumination of faint spinning objects in flight. *Nature Physics*, 5(1):64–67, 2008.

63 D. Ne-Te Loh et al. A reconstruction algorithm for single-particle diffraction imaging experiments. *Phys. Rev. E*, 80(2):026705, 2009.

64 W.H. Benner, M.J. Bogan, U. Rohner, S. Boutet, B. Woods, and M. Frank. Non-destructive characterization and alignment of aerodynamically focused particle beams using single particle charge detection. *Journal of Aerosol Science*, 39(11):917–928, 2008.

65 R. Hartmann, S. Epp, H. Gorke, A. Hartmann, G. Hauser, S. Herrmann, P. Holl, N. Kimmel, N. Meidinger, C. Reich, et al. Large format imaging detectors for x-ray free-electron-lasers. In *SPIE Optics+ Optoelectronics*, pages 80780W–80780W. International Society for Optics and Photonics, 2011.

66 N. Kimmel, R. Andritschke, L. Englert, S. Epp, A. Hartmann, R. Hartmann, G. Hauser, P. Holl, I. Ordavo, R. Richter, et al. Calibration methods and performance evaluation for pnccds in experiments with fel radiation. In *SPIE Optics+ Optoelectronics*, pages 80780V–80780V. International Society for Optics and Photonics, 2011.

67 D.R. Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1):37–50, 2005.

68 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), 2010.

69 Jan Steinbrener. https://github.com/JanSteinbrener/dm_recon, 2010.

70 J.F. Steinbrener. X-ray diffraction microscopy: Computational methods and scanning-type experiments. *PhD Thesis*, 2010.

71 V. Elser. Phase retrieval by iterated projections. *Journal of the Optical Society of America A*, 20(1):40–55, 2003.

72 James L Van Etten and David D Dunigan. Chloroviruses: not your everyday plant virus. *Trends in plant science*, 17(1):1–8, 2012.

73 Matthias G Fischer, Michael J Allen, William H Wilson, and Curtis A Suttle. Giant virus with a remarkable complement of genes infects marine zooplankton. *Proceedings of the National Academy of Sciences*, 107(45):19508–19513, 2010.

74 Didier Raoult, Stéphane Audic, Catherine Robert, Chantal Abergel, Patricia Renesto, Hiroyuki Ogata, Bernard La Scola, Marie Suzan, and Jean-Michel Claverie. The 1.2-megabase genome sequence of mimivirus. *Science*, 306(5700):1344–1350, 2004.

75 M.M. Seibert, T. Ekeberg, F.R.N.C. Maia, M. Svenda, J. Andreasson, O. Joensson, D. Odic, 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. Homke, C. Reich, D. Pietschner, G. Weidenspointner, L. Struder, G. Hauser, H. Gorke, J. Ullrich, I. Schlichting, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kuhnel, R. Andritschke, C.-D. Schroter, 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.

76 Hans Wolter. Spiegelsysteme streifenden einfalls als abbildende optiken für röntgenstrahlen. *Annalen der Physik*, 445(1-2):94–114, 1952.

77 N.D. Loh, M.J. Bogan, V. Elser, A. Barty, S. Boutet, S. Bajt, J. Hajdu, T. Ekeberg, F. Maia, J. Schulz, et al. Cryptotomography: reconstructing 3d fourier intensities from randomly oriented single-shot diffraction patterns. *Physical review letters*, 104(22):225501, 2010.

78 H.T. Philipp, K. Ayyer, M.W. Tate, V. Elser, and S.M. Gruner. Solving structure with sparse, randomly-oriented x-ray data. *Arxiv preprint arXiv:1203.3353*, 2012.

79 J.B. Tenenbaum, V. De Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

80 Gábor Bortel and Miklós Tegze. Common arc method for diffraction pattern orientation. *Acta Crystallographica Section A*, 67(6):533–543, Nov 2011.

81 David B Salzman. A method of general moments for orienting 2d projections of unknown 3d objects. *Computer Vision, Graphics, and Image Processing*, 50(2):129 – 156, 1990.

82 X. Wang, K.E. Schmidt, and J.C.H. Spence. Molecular orientation determination from moments of inertia of their diffraction patterns. *Acta A*, 2009.

83 G Huldt, A Szőke, and Janos Hajdu. Diffraction imaging of single particles and biomolecules. *Journal of structural biology*, 144(1):219–227, 2003.

84 H.T. Philipp, K. Ayyer, M.W. Tate, V. Elser, and S.M. Gruner. Solving structure with sparse, randomly-oriented x-ray data. *Opt. Express*, 20(12):13129–13137, 2012.

85 Glenn Roberts Jr. LAMP: A new tool turns on at SLAC's x-ray laser. https://www6.slac.stanford.edu/news/2014-01-15-lamp-new-tool-turns-slacs-x-ray-laser.aspx, 2014.

86 J. Küpper, S. Stern, L. Holmegaard, F. Filsinger, A. Rouzée, A. Rudenko, P. Johnsson, A. V. Martin, M. Adolph, A. Aquila, S. Bajt, A. Barty, C. Bostedt, J. Bozek, C. Caleman, R. Coffee, N. Coppola, T. Delmas, S. Epp, B. Erk, L. Foucar, T. Gorkhover, L. Gumprecht, A. Hartmann, R. Hartmann, G. unter Hauser, P. Holl, A. Hömke, N. Kimmel, F. Krasniqi, K. Uwe Kühnel, J. Maurer, M. Messerschmidt, R. Moshammer, C. Reich, B. Rudek, R. Santra, I. Schlichting, C. Schmidt, S. Schorb, J. Schulz, H. Soltau, H.. Spence, J. C. D. Starodub, L. Strüder, J. Thøgersen, J.. Vrakking, M. J. G. Weidenspointner, T. A. White, C. Wunderer, G. Meijer, J. Ullrich, H. Stapelfeldt, D. Rolles, and H. N. Chapman. X-ray diffraction from isolated and strongly aligned gas-phase molecules with a free-electron laser. *Phys. Rev. Lett.*, 112:083002, Feb 2014.

87 B Erk, D Rolles, L Foucar, B Rudek, S W Epp, M Cryle, C Bostedt, S Schorb, J Bozek, A Rouzee, A Hundertmark, T Marchenko, M Simon, F Filsinger, L Christensen, S De, S Trippel, J Kpper, H Stapelfeldt, S Wada, K Ueda, M Swiggers, M Messerschmidt, C D Schrter, R Moshammer, I Schlichting, J Ullrich, and A Rudenko. Inner-shell multiple ionization of polyatomic molecules with an intense x-ray free-electron laser studied by coincident ion momentum imaging. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 46(16):164031, 2013.

88 Michael J Bogan, Dmitri Starodub, Christina Y Hampton, and Raymond G Sierra. Single-particle coherent diffractive imaging with a soft x-ray free electron laser: towards soot aerosol morphology. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 43(19):194013, 2010.

89 Z. Kam. Determination of macromolecular structure in solution by spatial correlation of scattering fluctuations. *Macromolecules*, 10(5):927–934, 1977.

90 D.K. Saldin, H.-C. Poon, P. Schwander, M. Uddin, and M. Schmidt. Reconstructing an icosahedral virus from single-particle diffraction experiments. *Opt. Express*, 19(18):17318–17335, 2011.

91 D.K. Saldin, H.C. Poon, M.J. Bogan, S. Marchesini, D.A. Shapiro, R.A. Kirian, U. Weierstall, and J.C.H. Spence. New light on disordered ensembles: Ab initio structure determination of one particle from scattering fluctuations of many copies. *Physical Review Letters*, 106(11):115501, 2011.

92 R.A. Kirian, K.E. Schmidt, X. Wang, R.B. Doak, and J.C.H. Spence. Signal, noise, and resolution in correlated fluctuations from snapshot small-angle x-ray scattering. *Physical Review E*, 84(1):011921, 2011.

93 R. A. Crowther, D. J. DeRosier, and A. Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 317(1530):319–340, 1970.

94 Marin Van Heel. Angular reconstitution: A posteriori assignment of projection directions for 3d reconstruction. *Ultramicroscopy*, 21(2):111 – 123, 1987.

95 Brian Moths and Abbas Ourmazd. Bayesian algorithms for recovering structure from single-particle diffraction snapshots of unknown orientation: a comparison. *Acta Crystallographica Section A: Foundations of Crystallography*, 67(5):481–486, 2011.

96 Arthur P Dempster, Nan M Laird, Donald B Rubin, et al. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal statistical Society*, 39(1):1–38, 1977.

97 R.R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

98 Gregory S Chirikjian and Alexander B Kyatkin. *Engineering applications of noncommutative harmonic analysis: with emphasis on rotation and motion groups*. CRC press, 2000.

99 G. van Rossum and F.L. Drake (eds). Python reference manual. http://www.python.org, 2001.

100 Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python. http://www.scipy.org/, 2001.

101 Gaël Guennebaud, Benoît Jacob, et al. Eigen v3. http://eigen.tuxfamily.org, 2010.

102 E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

103 R. Clint Whaley and Antoine Petitet. Minimizing development and maintenance costs in supporting persistently optimized BLAS. *Software: Practice and Experience*, 35(2):101–121, February 2005.

104 OpenMP Architecture Review Board. http://openmp.org/wp/openmp-specifications/, 2010.

105 Nigel Hitchin. A lecture on the octahedron. *Bulletin of the London Mathematical Society*, 35(05):577–600, 2003.

106 George AF Seber. *Multivariate observations*, volume 252. John Wiley & Sons, 2009.

107 Scott Kirkpatrick. Optimization by simulated annealing: Quantitative studies. *Journal of statistical physics*, 34(5-6):975–986, 1984.

108 Heinrich von Kleist. Über die allmähliche Verfertigung der Gedanken beim Reden. *Nord und Süd*, 4:3–7, 1878.