Reconstruction Methods In Free Electron Laser X-Ray Diffraction

by

Xiaoyu Wang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2011 by the
Graduate Supervisory Committee:

John Spence, Chair
R Doak
Robert Ros
Uwe Weierstall
Kevin Schmidt

ARIZONA STATE UNIVERSITY

August 2011

ABSTRACT

One of the most important issues in femtosecond free electron laser X-ray diffraction is to reconstruct the 3D charge density of molecule from a mass of diffraction snapshots.

In order to determine the orientation of single molecule from diffraction patterns, we first determine the moments and products of inertia of this from 2D experiment data (diffraction patterns or EM images to obtain the elements of the inertia tensor. If diffraction patterns from uniformly random orientations or some preferred orientations are collected, the principal axes of the molecule can be extracted, together with the Euler angles which relate the principal axes of the molecule to the laboratory frame axes. This is achieved by finding the maximum and minimum values for the measured moments from many single-molecule patterns. Simulations for GroEL protein indicates that the calculation of the autocorrelation help eliminate the Poisson noise in Cryo-EM images and can make correct orientation determination.

The effect of water jacket surrounding the protein molecule is studied based on molecular dynamics simulation result. The intensities from water and interference is found to suppress those from protein itself. A method is proposed and applied to the simulation data to show the possibility for it to overcome the water background problem.

The scattering between Bragg reflections from nanocrystals is used to aid solution of the phase problem. We describe a method for reconstructing the charge density of a typical molecule within a single unit cell, if sufficiently finely-sampled diffraction data are available from many nanocrystals of different sizes lying in the same orientations without knowledge of the distribution of particle size or requiring atomic-resolution data.

Triple correlation of the diffraction patterns are made use of to recon-

structed the 3D diffraction intensities from 2D data set. An analytical, linear, and non-iterative algorithm is developed to tackle this problem with the assumption that the spherical harmonics expansion of intensities is band-limited. The algorithm is made feasible by decoupling the large nonlinear problem and numerical implementation shows it works with ideal data but error accumulation has to be overcome before applying to real world data.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Most of protein structures are solved by X-ray crystallography. X-ray crystallography is one of the most successful techniques ever developed for the study of structures with atomic resolution. Every year thousands of new structures are solved and posted to the Protein Data Bank. However, the success of the method depends on growing crystals of sufficient size and quality. The growth of high-quality crystals needs huge amount of investment, and even though, many proteins yield poorly grown crystals. In July 2007, there are more than 750,000 proteins sequenced, and the structures of less than 6% (44,700) of them are solved. However, only 460 of the proteins with determined structures are membrane proteins, which are extremely important as 70% of today's drugs aim at them, but are notoriously difficult to crystallize. Revolutions in both the experimental and theoretical approaches, such as new phasing and sorting algorithms, are expected to make breakthroughs in this fundamental field of nature science.

Serial Crystallography with Femtosecond X-ray Free Electron Laser (XFEL)

In conventional measurements, the necessary increase in X-ray dose to record data from crystals that are too small or lacking sufficient order leads to extensive damage before a diffraction signal can be recorded. Coherent X-ray Diffractive Imaging (CXDI), which uses a coherent, short and extremely bright pulse of X-rays to obtain a diffraction pattern which is then phased to reconstruct its charge distribution, has emerged as an promising alternative.

The idea of CXDI, as proposed by Sayre[3] in the early 1980s, is to determine the spatial distribution of electron density $\rho(\mathbf{r})$ in noncrystalline sample

from its far-field coherent diffraction pattern. There are two distinctions from conventional X-ray crystallography. Firstly, the reciprocal space data (diffraction pattern) is a continuous function for noncrystalline targets, as opposed to discrete Bragg peaks for a crystal. This allows the application of an iterative oversampling phasing algorithm[4, 5, 6] for phase retrieval and structure determination. Secondly, The method requires an intense fully coherent X-ray incident beam to preserve the phase information in the diffraction pattern and overcome the lack of periodicity.

This technique only becomes feasible recently with the availability of XFEL. XFELs are capable of producing intense X-ray pulse, which is a billion times brighter than the third generation synchrotron light source, and its pulses duration as short as a few femtoseconds (fs, 10^{-15}s). For example, the FLASH XFEL in Germany is the first soft XFEL, which was launched as the first hard XFEL int the world, generates photons with wavelength of 32nm and pulse length of 10fs. It has been updated to reach the wavelength of 6.5nm. In 2009, the Linac Coherent Light Source (LCLS) is launched as the first hard XFEL in the world, whose wavelength reached $1.5\mathring{A}$ in April 2009. This facility provides the ultimate instrument to solve the structure problem at atomic resolution.

However, the intense radiation may cause substantial damage to the biology sample. In a preliminary simulation, it has been indicated that the spatial resolution is limited to 10nm for organic samples an 1nm for inorganic samples (see Fig.(1.2)[8]. Further simulation has confirmed this result, finding that multiple single-file protein beams will be needed for sub-nanometer resolution on current third-generation synchrotrons where reconstruction of secondary protein structure at a resolution of $7\mathring{A}$ should be possible with relatively short exposure time[9, 10].

2
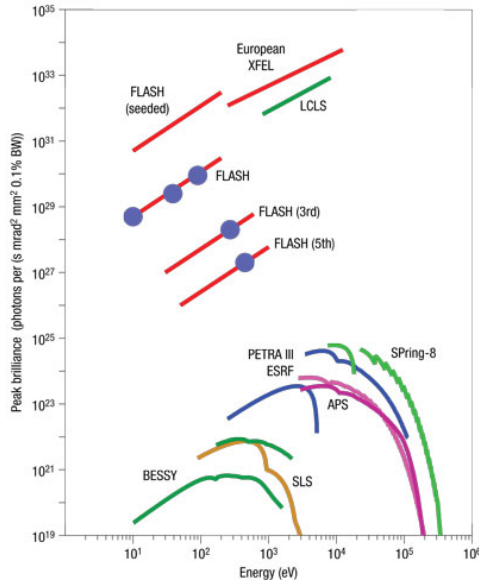
Figure 1.1: Peak brilliance of several X-ray light source[7]
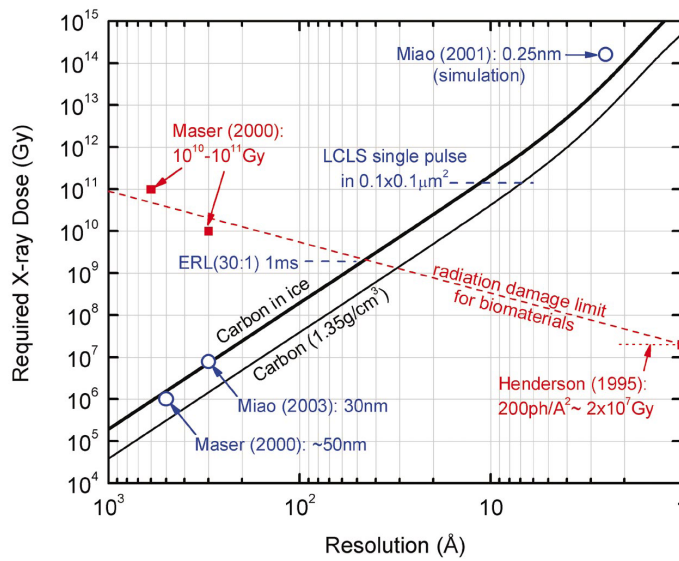
Figure 1.2: Required X-ray dose and the radiation-damage limit is shown as the dotted line connecting Henderson's limit at atomic resolution and microscopy studies against mass loss at low resolutions.[10]

Neutze[11] thus proposed a scheme to overcome the radiation-damage limit as described in Fig.(1.3)[12]. It makes use of the ultrashort pulse to record a high-resolution diffraction patter before the illuminated molecule explodes as a result of the exposure to the intense radiation. The intensity pattern formed from the intense X-ray pulse (incident from left) scattering off the object is recorded on a pixelated detector. The pulse also photoionizes the sample. This leads to plasma formation and Coulomb explosion of the highly ionized particle, so only one diffraction pattern (a single two-dimensional slice) can be recorded from the particle. Many of these individual diffraction patterns can be recorded from single particles in a jet (traveling from top to bottom). The particles travel fast enough to clear the beam by the time the next pulse (and particle) arrives. The data must be read out from the detector just as quickly. In the next stage, the full 3D diffraction data set is assembled from noisy diffraction patterns of identical particles in random and unknown orientations. Patterns are classified to group patterns of like orientation, averaged within the groups to increase signal to noise, oriented with respect to one another, and combined into a 3D map of reciprocal space. The image is then obtained by iterative phase retrieval. Although the high resolution signal in a single pattern collected is not strong enough for the future reconstruction of a protein molecule, millions of such patterns are available, and if some way can be found to sort and merge them properly, they can provide statistically significant signal at atomic resolution[13].

On the other hand, the application of the state-of-the-art facility and the novel experimental scheme also poses some key challenges to the community of diffraction physics. Firstly, the synchronized beam of protein are hydrated in order to improve the hit rate of incident pulse to an acceptable level. There-

4
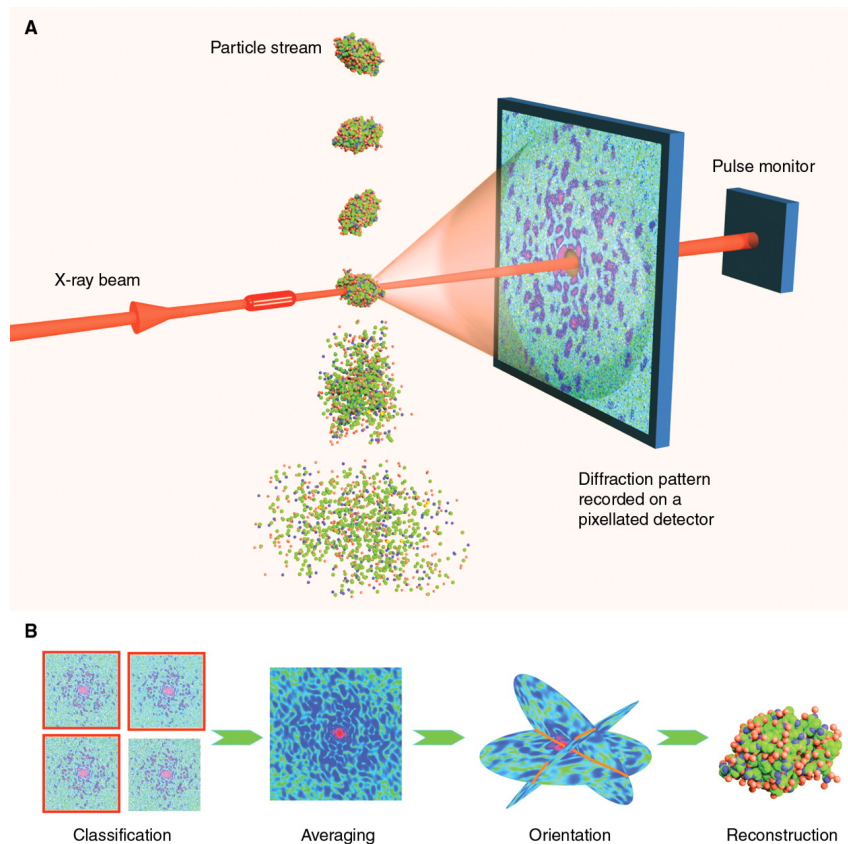
Figure 1.3: Schematic depiction of single-particle coherent diffractive imaging with an XFEL pulse.[11]

fore, the effect of water jacket surrounding the molecule remains to be studied to extract the signal from molecule alone. Secondly, even given the ultra high flux of XFELs, the photons scattered per shot is still very few ($\tilde{1}000$/shot), which fall onto about a million of pixels on the detector. Considering the task of next step to reconstitute the 3D intensity distribution from each 2D snapshot from unknown random orientation, such weak signals invalidate the most straightforward and widely used orientation determination method, common line method[14]. Moreover, the data had better be collected within 20fs after pulse arrival, at the moment after the molecule is blown up and before it has flown apart.

Although the serial crystallography has been accepted as the on of the mainstream applications of XFELs and real-world experiments has been conducted, which confirm the power of it[15], many of the problems emerging from it, especially at the stage of data process, still remain unclear up to now due to their difficulties. The aims of this thesis are to investigate several of the issues above via theoretical and computational methods to shed some light on them.

Coherent Diffractive Imaging

Diffractive (or lensless) imaging refers to the use of mathematical methods and computer algorithms to solve the phase problem for scattering by a non-periodic object. Additional information about the object, such as the sign of the scattering potential an the approximate boundary of the object, may be combined with the measured scattered intensity to solve for the phases of the scattered amplitudes. By avoiding the need for a lens, the aberrations and resolution limits introduced by lenses are thus avoided. Diffractive imaging promises a 3D resolution limited only by radiation damage, wavelength, the collected solid angle, and the number of photons or electrons collected.

Sayre is the first to consider the relationship between Shannon's sampling theorem and Bragg's law[16]. He found the fact that the Bragg diffraction undersamples the diffracted intensity was important and led to more specific proposals by the same author for X-ray diffractive microscopy of nonperiodic objects.

Since the proposal of the principle, a number of groups have been working to bring it from words to reality. Robinson and co-workers applied it to hard X-ray experiments on microcrystalline particles[17]. A 3D image at 40nm

resolution has been constructed tomographically.

A collaboration between Berkeley and Livermore laboratories and Arizona State University produced 3D imaging at $10 \times 10 \times 40$nm resolution of test samples[18]. In this work, a simple zone plate was used as a monochromator, following by a beam-defining aperture of about 10 μm in diameter, coherently filled. A nude soft X-ray CCD camera, employing 1024×1024 24-μm pixels was used. The sample is mounted in the center of a silicon nitride window fitted to a TEM single-tilt holder, which provides automated rotation about a single axis normal to the X-ray beam. The window is rectangular, with the long axis normal to both the beam and the holder axis. Diffraction patterns are recorded at 1° rotation increments, with a typical recording time of about 15 min per orientation. The maximum tilt angle is then limited by the thickness of the silicon frame around the window to perhaps 80°, resulting in a missing wedge of data. In addition, data may be missing around the axial beamstop. 3D interpolation of data points near the sphere is needed, and careful intensity scaling may be necessary if several exposures with different times are required to cover the full dynamic range of the data. It is often found that missing data points in the central region can be treated as adjustable parameters in the phasing. Once a roughly spherical volume has been filled in reciprocal space (perhaps with missing wedge and beam-stop region), the 3D iterations of the phasing algorithm may be applied. The computing demands are severe, as outlined below. The converged data will provide a 3D density map (see Fig.(1.4), proportional to the local charge density, if the single-scattering approximation of X-ray diffraction theory applies and if the spatial variation in attenuation of the beam due to the photoelectric effect can be neglected.

In the field of electron microscopy, lens aberrations limit electron imaging

7
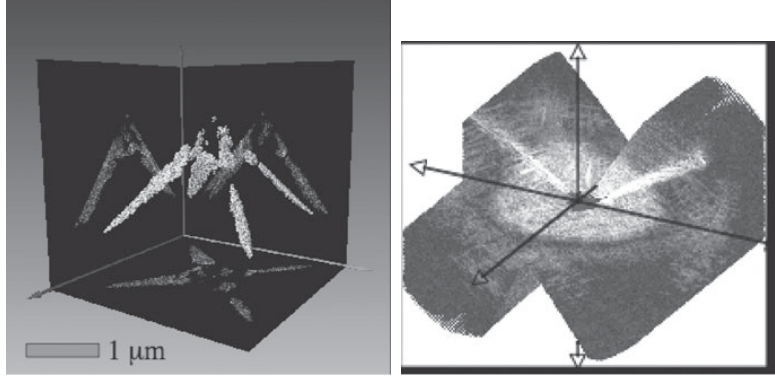
Figure 1.4: (A) Tomographic reconstruction from a soft X-ray diffraction pattern shown in (B). The object consists of gold balls (50 nm diameter) lying along the edges of a pyramidal-shaped silicon nitride structure. This is one image from a rotation series. From the complete series, three-dimensional surfaces of constant density can be constructed. (B) The volume of soft X-ray diffraction data collected to obtain the three-dimensional reconstruction in (A).[18]

resolution to about 1 angstrom. Resolution is reduced further by low contrast from weak scattering or from the limitations on electron dose for radiation-sensitive molecules. Zuo showedq that both high resolution and high contrast can be achieved by imaging from diffraction with a nanometer-sized coherent electron beam[19]. The coherent electron nanodiffraction is collected on a JEOL electron microscope with a field emission gun. The phase problem is solved by oversampling and iterative phase retrieval. Although the nominal point resolution of the machine is $2.2\mathring{A}$ for phase contrast imaging at the Scherzer focus condition, this technique is able to image a double-wall carbon nanotube at $1\mathring{A}$ resolution, revealing the structure of two tubes of different helicities (see Fig.(1.5, 1.6).

Iterative Phasing Algorithm

When we record the diffraction pattern intensity scattered by an object, the phase information is missing. Apart from normalization factors, an object of
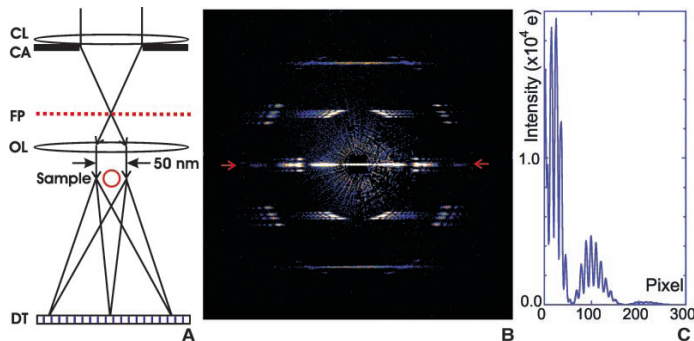
Figure 1.5: Coherent nanoarea electron diffraction: (A) a schematic ray diagram, (B) the recorded diffraction pattern from a DWNT, and (C) the intensity profile of (B) from the center along the line indicated by arrows. The nanometer-sized parallel electron beam is formed by illuminating the condenser aperture (CA) with a coherent electron beam from the field emission electron gun and focusing the beam with the use of the condenser lens (CL) onto the front focal plane (FP) of the objective lens (OL). The size of beam is 50 nm for a $10\mu m$ aperture. The far-field diffraction pattern of (B) is digitized with the use of imaging plates. The pixel resolution is $0.0025\AA^{-1}$, which defines a $400\AA$ field of view in real space. Intensities for the center 60 pixels by 60 pixels were obtained from the Fourier transform amplitude of a low-resolution electron image of the DWNT.[19]

density $\rho(\mathbf{r})$, $\mathbf{r}$ being the coordinates in the *object* (or *real*) space, generates a diffraction pattern equal to the modulus square of the Fourier transform (FT) $\tilde{\rho}(\mathbf{k})$:

$$
\begin{aligned}
I(\mathbf{k}) &= |\tilde{\rho}(\mathbf{k})|^2 \\
I(\mathbf{k}) &= \tilde{\rho}^{\dagger}(\mathbf{k})\tilde{\rho}(\mathbf{k}),
\end{aligned}
\tag{1.1}
$$

where $\mathbf{k}$ represent the coordinate in the Fourier (or Reciprocal) space. The inverse Fourier transform (IFT) of the measured intensity I provides the autocorrelation $\rho(-\mathbf{r}) * \rho(\mathbf{r})$ of the object:

$$
\text{IFT}[I(\mathbf{k})] = \rho(-\mathbf{r}) * \rho(\mathbf{r}).
\tag{1.2}
$$

The phase-retrieval problem consists of solving $\tilde{\rho}$ in Eq. (1.1) or ρ in Eq. (1.2), using some extra prior knowledge. In diffraction microscopy, solving such

9
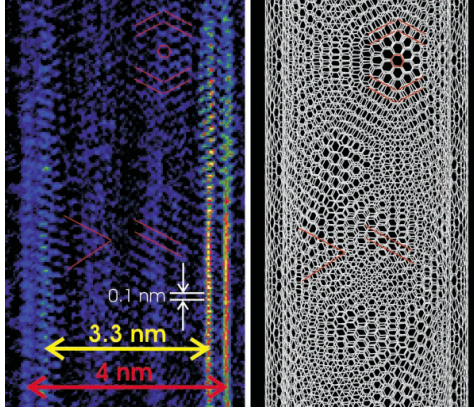
Figure 1.6: (left) A section of the reconstructed DWNT image at $1\mathring{A}$ resolution and (right) a structural model constructed with the use of the chiral vectors of (35, 25) and (26, 24) that were determined from the image and diffraction pattern. The DWNT imaged here is one of many in our catalytic chemical vapor deposition grown samples. Yellow and red lines mark the diameters of the inner and outer tubes, respectively. One side of walls is stronger than the other, which is because of the illumination. The DWNT is incommensurate. In projection, the structure has complex patterns showing both accidental co-incidences and Moire fringes, which are highlighted by hexagons and lines.[19]

problem is performed with giga-element large-scale optimization algorithms, described in the following section.

Since the intensity represents the FT of the autocorrelation function, and the autocorrelation is twice as large as the object, the diffraction pattern intensity should be sampled at least twice as finely as the amplitude to capture all possible information on the object. Finer sampling adds a 0-padding region around the recovered autocorrelation function

$$\rho(\mathbf{r}) = 0, \text{ if } \mathbf{r} \notin S. \tag{1.3}$$

which adds no further information (Shannon theorem). Less than critical sampling in the Fourier domain causes alias in the object space. Loosely, it can also be understood in the way that the missing half of the data (the phases) are compensated by requiring that half of the object values be known

10

(they are zero outside the support), so that the system of ellipsoid equations
(Eqn.(1.4)) are solvable in principle.

$$\left| \sum_{\boldsymbol{r} \in S} \rho(\boldsymbol{r}) \exp(i\boldsymbol{k} \cdot \boldsymbol{r}) \right|^2 =$$

$$\sum_{\boldsymbol{r}, \boldsymbol{r}' \in S} \exp(i\boldsymbol{k} \cdot (\boldsymbol{r} - \boldsymbol{r}')) \rho(\boldsymbol{r}) \rho^*(\boldsymbol{r}') = I(\boldsymbol{k}). \qquad (1.4)$$

Each value of $I(\boldsymbol{k})$ in reciprocal space defines an ellipsoid (Eq. (1.4)) in
the multidimensional space of the unknowns $\rho(\boldsymbol{r})$, $\{\boldsymbol{r} \in S\}$. The intersec-
tion of these ellipsoids forms our solution. Constant phase factors, inversion
with respect to the origin (enantiomorphs), and origin shifts $\rho(\pm\boldsymbol{r} + \boldsymbol{r_0})e^{i\phi_0}$
are undetermined and considered equivalent solutions. The presence of mul-
tiple non-equivalent solutions in two- and higher- dimensional phase retrieval
problems is rare[20]; it occurs when the density distribution of the object can
be described as the convolution of two or more non-centrosymmetric distri-
butions. Simple homometric structures for which the phase problem is not
unique [Buerger] exist in nature, but such non-uniqueness is less likely for
more complex structures.

Unfortunately this system of equations is difficult to solve, and has an
enormous number of local minima. In the early 1980s, the development of
iterative algorithms with feedback by Fienup[4], produced a remarkably suc-
cessful optimization method capable of extracting phase information. These
algorithms try to find the intersection between two sets, typically the set of all
the possible objects with a given diffraction pattern (modulus set), and the set
of all the objects that are constrained within a given area or support volume.
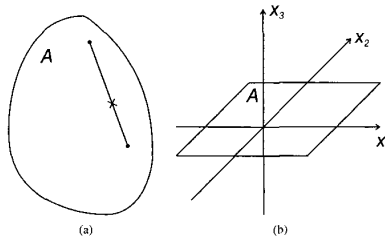
Figure 1.7: Convex sets A:(a) general geometry, (b) the convex set for a support constraint for which $x_3 = 0$[21]
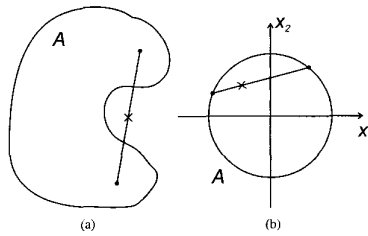
Figure 1.8: Nonconvex sets A:(a) general geometry, (b) the nonconvex set for a specified image energy. It also describes the Fourier magnitude constraint[21]

Constraints and Projections

A special role is played by constraints that are convex. Convex constraints are represented by convex constraint sets. A convex set is one for which the line joining any two points in the set is totally within the set (Fig. 1.7). An image with a given support is an example of a convex constraint. The constraints is characterized by certain pixels in the image being zero, and the constraints set is therefore a hyperplane in S, which has the property of convexity. However, the Fourier magnitude constraint set is nonconvex (Fig. 1.8)

A projector $\boldsymbol{P}$ is an operator that takes to the closest point of a set from the current point ρ. A repetition of the same projection is equal to one projection alone ($\boldsymbol{P^2} = \boldsymbol{P}$); its eigenvalues must therefore be $\lambda = 0, 1$. Another operator used here is the reflector $\boldsymbol{R} = \boldsymbol{I} + 2[\boldsymbol{P} - \boldsymbol{I}] = 2\boldsymbol{P} - \boldsymbol{I}$, which applies the same

12

step as the projector but moves twice as far.

Widely used in phase retrieval, projection onto the support $\mathbf{P}s$ involves setting to 0 the components outside the support, while leaving the rest of the values unchanged

$$
\boldsymbol{P}_s\rho(\boldsymbol{r}) =
\begin{cases}
\rho(\boldsymbol{r}) & \text{if } \boldsymbol{r} \in S \\
\\
0 & \text{otherwise,}
\end{cases}
\tag{1.5}
$$

And the projector onto the magnitude constraints in reciprocal space $\mathbf{P}_m$ is defined as

$$
\boldsymbol{P}_m = F^{-1}\tilde{\boldsymbol{P}}_m F \,,
\tag{1.6}
$$

where F and F^{-1} represent the forward and inverse Fourier transforms respectively and

$$
\tilde{\boldsymbol{P}}_m\tilde{\rho}(\boldsymbol{k}) = \tilde{\boldsymbol{P}}_m|\tilde{\rho}(\boldsymbol{k})|e^{i\varphi(\boldsymbol{k})} = \sqrt{I}(\boldsymbol{k})e^{i\varphi(\boldsymbol{k})} \,,
\tag{1.7}
$$

An important property of convex sets is that the projection onto it is unique for any point. For nonconvex sets, the projection is not necessarily unique, although in many cases it is unique for most points. The immediate result of that is it is always possible to reach the intersection of two convex set by iterative projection from anywhere. On the other hand, as a result of the nonconvexity of at least one constraint, the process may converge to a point in the intersection, converge to a point not in the intersection, converge to a limit cycle, or diverge, depending on the particular constraints and the starting point (see Fig. 1.9)[21]. Convergence to a limit cycle that is not close to the solution is often referred to as stagnation, which remains a important problem to be overcome for any practical iteration algorithm.

13
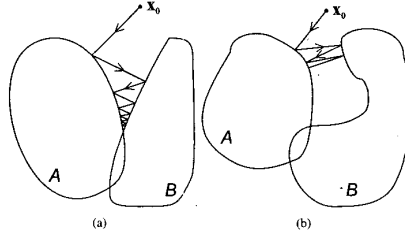
Figure 1.9: Geometric illustrations of (a) the projection onto convex sets, and (b) the projection onto a convex and a nonconvex set[21]

Moreover, there can be other constraints in real world such as positivity. Any number of constraint sets can be combined into a single constraint set, i.e.

$$\boldsymbol{P}_{s+}\rho\boldsymbol{r} = \begin{cases} \rho(\boldsymbol{r}) & \text{if } \boldsymbol{r} \in S \ \& \ \rho(\boldsymbol{r}) \geq 0 \,, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{1.8}$$

so that the number of constraint sets can be reduced to two (one in real space and another in reciprocal space). However, the sets become more complex, as do the projection operators, when sets are combined. Actually combining convex constraints generally gives a nonconvex constraint.

Iterative Projection Algorithm

Several algorithms based on these concepts have now been proposed and summarized here. The following algorithms require a starting point ρ^0, which is generated by assigning a random phase to the measured object amplitude (modulus) in the Fourier domain $|\tilde{\rho}(\boldsymbol{k})| = m(\boldsymbol{k}) = \sqrt{I(\boldsymbol{k})}$.

The first algorithm called *error reduction* (ER)[5] is the simplest implement of the projecting back and forth between two sets, it converges to the local minimum.

$$\rho^{(n+1)} = \boldsymbol{P}_s\boldsymbol{P}_m\rho^{(n)} \,, \tag{1.9}$$

14

Figure 1.10 shows that the step size is far from optimum, but that it guarantees linear convergence. A line search along this gradient direction would considerably speed up the convergence to a local minimum.

The *solvent flipping* (SF) algorithm[22] is obtained by replacing the support projector $\boldsymbol{P_s}$ with its reflector $\boldsymbol{R_s} = 2\boldsymbol{P_s} - \boldsymbol{I}$:

$$\rho^{(n+1)} = \boldsymbol{R_s P_m} \rho^{(n)}, \tag{1.10}$$

which multiplies the charge density ρ outside the support by -1.

The *hybrid input-output* (HIO) [4] (Fig. 1.10) is based on non-linear feedback control theory and can be expressed as:

$$\rho^{(n+1)}(x) = \begin{cases} \boldsymbol{P_m} \rho^{(n)}(x) & \text{if } x \in S, \\ (\boldsymbol{I} - \beta \boldsymbol{P_m}) \rho^{(n)}(x) & \text{otherwise.} \end{cases} \tag{1.11}$$

Since the output of HIO iteration is not necessarily a good estimate of the solution, the algorithm is always terminated with a few cycles of ER. The *difference map* (DM) is a general set of algorithms [23], which requires 4 projections (two time-consuming modulus constraint projections) (Fig. 1.10):

$$\rho^{(n+1)} = \{\boldsymbol{I} + \beta \boldsymbol{P_s} \left[(1 + \gamma_s) \boldsymbol{P_m} - \gamma_s \boldsymbol{I}\right] - \beta \boldsymbol{P_m} \left[(1 + \gamma_m) \boldsymbol{P_s} - \gamma_m \boldsymbol{I}\right]\} \rho^{(n)};$$

the solution corresponding to the fixed point is described in the same article. We will use in the upcoming tests what Elser suggested as the optimum, with $\gamma_s = -\beta^{-1}$ and $\gamma_m = \beta^{-1}$.

The *averaged successive reflections* (ASR)[24] algorithm is:

$$\rho^{(n+1)} = \tfrac{1}{2}[\boldsymbol{R_s R_m} + \boldsymbol{I}] \rho^{(n)}. \tag{1.12}$$

The *Hybrid Projection Reflection* (HPR)[25] algorithm is derived from a relaxation of the ASR:

$$\rho^{(n+1)} = \tfrac{1}{2}[\boldsymbol{R_s} \left(\boldsymbol{R_m} + (\beta - 1)\boldsymbol{P_m}\right) + \boldsymbol{I} + (1 - \beta)\boldsymbol{P_m}] \rho^{(n)}.$$

15
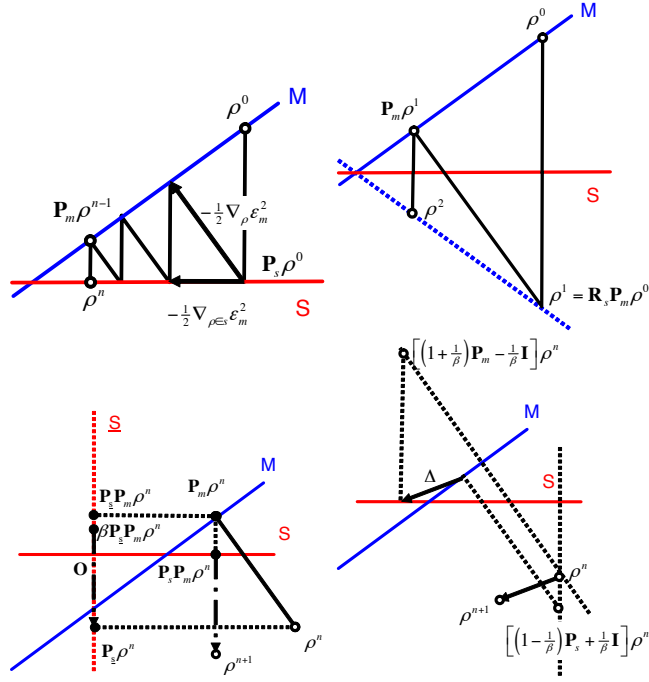
Figure 1.10: Geometric representation of various algorithms using a simplified version of the constraint: two lines intersecting. (a) Error reduction algorithm: we start from a point on the modulus constraint by assigning a random phase to the diffraction pattern. The projection onto the modulus constraint finds the point on the set which is nearest to the current one. The arrows indicate the gradients of the error metric. (b) The speed of convergence is increased by replacing the projector on the support with the reflector. The algorithm jumps between the modulus constraint (solid diagonal line) and its mirror image with respect to the support constraint (dotted line). (c) Hybrid input–output. The space perpendicular to the support set is represented by the vertical dotted line. (d) Difference map.[27]

It is equivalent to HIO if positivity is not enforced, but it is written in a recursive form, instead of a case-by-case form such as Eq. (1.11). It is also equivalent to the DM algorithm for $\gamma_s = -1$, $\gamma_m = \beta^{-1}$. Finally the *relaxed averaged alternating reflectors* (RAAR) algorithm[26]:

$$\rho^{(n+1)} = \left[\tfrac{1}{2}\beta \left(\boldsymbol{R_s R_m} + \boldsymbol{I} \right) + (1 - \beta)\boldsymbol{P_m} \right] \rho^{(n)} .$$ (1.13)

For $\beta = 1$, HIO, HPR, ASR and RAAR coincide.
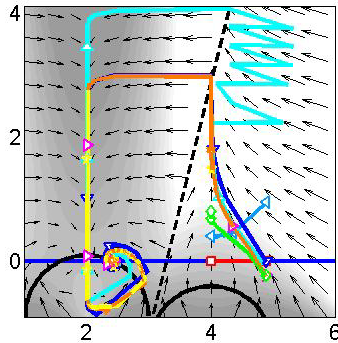
Figure 1.11: The horizontal line represents a support constraint, while the two circles represent a non-convex constraint, i.e. the modulus constraint.[27]

In conclusion, these algorithms can be grouped in two categories[27]: (1) local minimizers such as ER, SF, steepest descent and conjugate-gradient methods, with Solvent Flip having some moderate ability to escape local minima[22] (2) more global minimizers such as HIO, DM, ASR, HPR which use a feedback to reach the solution. RAAR and ER+HIO fall somewhere in between the two categories, depending on an adjustable parameter.

In order to show this important difference, a numerical example is shown in Fig. 1.11. Here the circumference of two circles represents a non-convex set (modulus constraint), while the support constraint is represented by a line. The convex set represents a simplified modulus constraint in a phase-retrieval problem. The gradient-type (ER and SF) algorithms converge to the local minimum, while HIO and its variants follow the descent-ascent direction indicated by the arrows.

A simple 2-D phase-retrieval problem, where only two variables (pixel values) are unknown. The solution is the atop minimum in the figures, is also used to test the most widely used algorithms, ER and HIO The primary advantages of iterative projection algorithms for solving inverse problems are

17
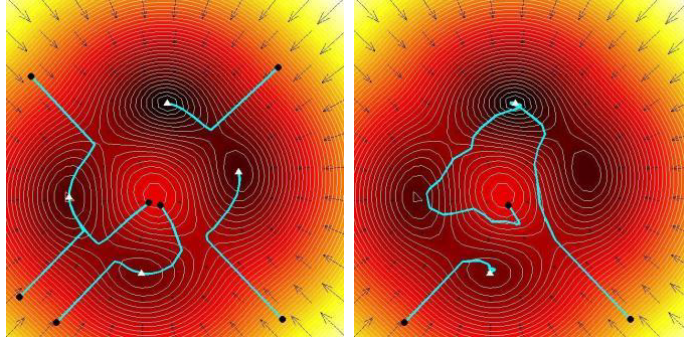
Figure 1.12: The error reduction algorithm proceeds toward the local minimum while the HIO method generally converges to the global minimum, however some rare starting points converge to a local minimum.[27]

1. they are computationally efficient compared to many other optimization methods (the projection operators are fast to compute)

2. the projection operators are relatively simple to implement

3. the more advanced algorithm are quite resistant to becoming trapped in local minima (stagnation)

Beam Stop Problem

There are several approaches made to solve the problem of data lost behind a synchrotron beam stop, which is essential to protect a sensitive area detector. In several HIO applications these missing values have simply been treated as free adjustable parameters, and the algorithm was found to converge.

Here the simplest initial choice of support is the boundary of the autocorrelation function (obtained by Fourier transform of the diffracted intensity). The part of the diffraction pattern covered by a central beam stop from the transform of the current estimate of the object. Low-frequency components are treated as free parameters. Every 20 iterations the reconstructed image is convolved with a Gaussian peak to find the new support mask. The mask
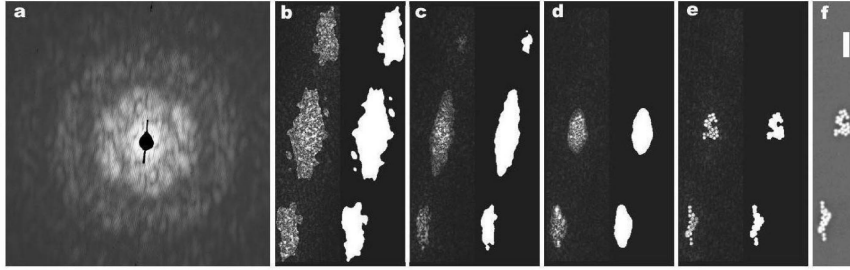
Figure 1.13: Image reconstruction from an experimental x-ray-diffraction pattern. (a) X-ray diffraction pattern of a sample of 50nm colloidal gold particles, recorded at a wavelength of 2 nm. (b - e) shows a sequence of images produced by the algorithm as it converges. Number of iterations: 1 (b), 20 (c), 100 (d), and 1000 (e). The reconstruction progresses from the autocorrelation function in (b) to an image in (e) with a steady improvement of the support boundary shown at the bottom of each frame. For comparison, a scanning electron micrograph of the object is shown in (f). The scale bar length is 300 nm and the resolution of our reconstructed image is about 20 nm.[28]

is then obtained by applying a threshold at 20% of its maximum. The width is set to 3 pixels in the first iteration, and reduced by 1% every 20 iterations down to a minimum of 1.5 pixels. This estimate is rapidly improved upon by the so called shrinkwrap algorithm[28]. This appears to be the most useful practical algorithm at present.

There is another solution using a sample consisting of an unknown object filling a small hole in an otherwise opaque mask[29, 30]. The use of very small silicon nitride windows greatly reduces the intensity of the direct beam and blooming effects. However, the detailed shape of the partially transparent silicon wedge round the window must then be estimated and used as a support for inversion. Finally, the diffuse X-ray scattering around Bragg peaks from a cryallite has been inverted to an image, thus avoiding the direct-beam scattering.

Two-wavelength inversion of multiply scattered soft X-ray intensities to charge density[1]

For 2D soft X-ray and electron diffraction in the projection approximation, the reconstruction of the charge density (or potential) may be understood to be a two step process. The phase problem is solved for the far-field diffracted intensities, and the 2D Fourier transform of the intensities and phases provides the complex exit-wave function. However, without the assumption of single scattering, the real-space map of charge density (or potential) cannot be obtained. The parameters that define the validity domain of single-scattering approximation include extinction distance, sample thickness, and inelastic effects which may provide an effective limit on thickness.

The presence of multiple scattering destroys the simple Fourier Transform relationship of the first Born approximation between scattered amplitudes and the sample charge density. It has always been considered a severe limitation to this diffractive imaging technique[31]. However it has frequently been pointed out that in a sense multiple scattering solves the phase problem, since it allows interference between different Bragg beams in crystals and makes them sensitive to structure factor phases. An approach to inversion with multiple scattering which takes advantage of this effect, based on projection between constrained sets in the manner of the HiO algorithm, was described by Spence et al[32]. A dynamical Ptychography approach has also been described in [33].

Firstly, it is important to make sure that multiple-scattering effects can be ignored under strong absorption, although it appears an intuitional conclusion. This is due to the lack of general form of the distribution of multiple elastic scattering from a 3D nonperiodic sample apart from statement about sym-

metry. Multislice calculation at different energy is carried out on a mixture of protein and water, whose refractive parameters are obtained from CXRO website and the assumption that coherent scattering from the mixture is not greatly different from that from separated regions of each is made. The result shows that both single and multiple scattering curves commence with the correct parabolic thickness dependence, the low-angle multiple scattering a premature roll off, not present at high angles, which will enhance high spatial frequencies in images. The final effect is that although absorption does greatly attenuate the effects of multiple scattering, it still must be taken into account in order to obtain quantitative agreement for samples thicker than about half the extinction distance of the low order scattering.

Therefore, it is necessary to study the multiple scattering effect under any circumstance. Fortunately, it is possible to show that in soft X-ray diffraction the multiple scattering is a function a function only of the product λt, where λ is the wavelength while t the thickness of sample.

In the Bloch-wave representation of scalar multiple-scattering theory, the Fourier coefficients of the dynamical wave field at depth z within the sample can be formed into a column vector $\mathbf{u}$, where

$$\frac{d\mathbf{u}}{dz} = -2\pi i \mathbf{A}(z)\mathbf{u}(\mathbf{z}) \tag{1.14}$$

Assuming that the second-order derivative of $\mathbf{u}(z)$ in direction z is negligible[34], for soft X-ray diffraction, the off-diagonal elements of $\mathbf{A}$ are proportional to the set of z-dependent Fourier coefficients $A_{\mathbf{g}-\mathbf{h}}(z)$ of the 2D charge density.

If $\mathbf{A}$ is independent of z for a charge density that depends only on the 2D vector normal to the beam, the solution to Eqn.(1.14) is

$$\mathbf{u}(t) = \mathbf{S}\mathbf{u}_0 = \exp(2\pi i \mathbf{A} t)\mathbf{u}_0 \tag{1.15}$$

21
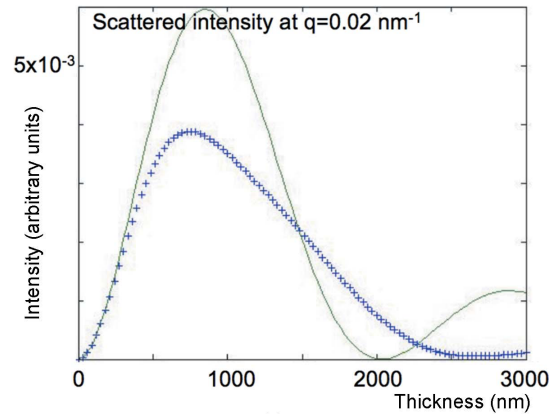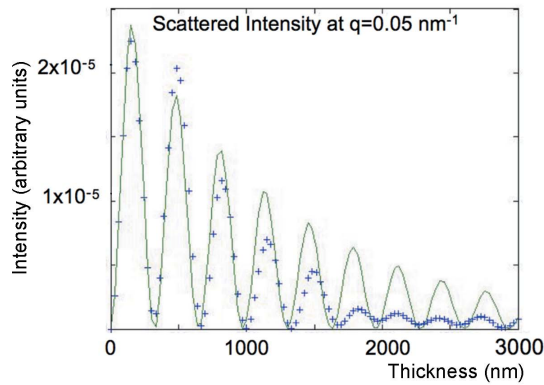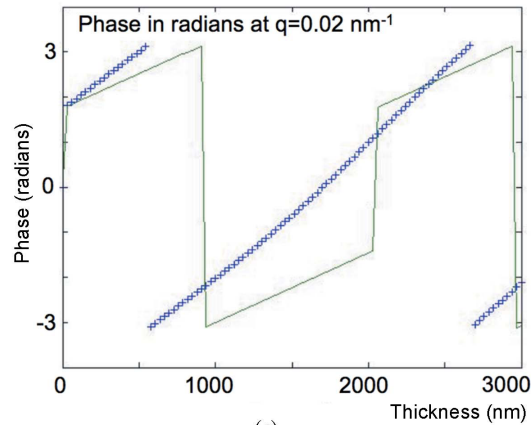
Figure 1.14: (a) The thickness dependence of the scattered intensity at 500 eV and $q = 0.02nm^{-1}$, both curves with absorption. The continuous curve shows the single scattering, the crosses show the multiple scattering. (b) Similar to (a) for $q = 0.05nm^{-1}$, showing more rapid oscillations. (c) Variation of phase (in radians) with thickness at $q = 0.02nm^{-1}$. Note the second jump (phase reversal) in the continuous curve at the minimum of the single-scattering curve in (a).[1]

22

where **u**0 is a column vector containing Fourier coefficients of the incident coherent beam and **S** is unitary in the absence of spatially varying absorption. Now the matrix **A** has diagonal elements which are the excitation errors of the computational superlattice[35]

$$S_q = |\mathbf{q}|^2 \lambda / 2 \tag{1.16}$$

And the off-diagonal elements in **S** are the positive quantities $A_{\mathbf{q}} = r_e \rho_{\mathbf{q}} \lambda / 2\pi$, where $\rho_{\mathbf{q}}$ is the complex Fourier coefficient of the effective charge density. Then there is

$$\mathbf{S} = \exp(2\pi i \mathbf{A}' \lambda t) \tag{1.17}$$

where **A**′ is independent of λ and thickness t over any range of beam energy for which the effective number of electrons. Taking the derivation on both size gives

$$\frac{d\mathbf{S}}{d\lambda} = (2\pi i \mathbf{A}' t) \mathbf{S} \tag{1.18}$$

As **S** can be measured from experiment, **A** is allowed to be found if t is known.

From Eqn.(1.14) and Eqn.(1.18), the exit wave can be written in form of

$$\Psi(\mathbf{r}, t, \lambda) = \Psi(\mathbf{r}, t\lambda) \tag{1.19}$$

as $\mathbf{u}(t) = \mathbf{u}(\mathbf{q}, t, \lambda)$ contains the Fourier coefficients of $\Psi(\mathbf{r}, t, \lambda)$. The dynamical solution will be unchanged if $\delta(t\lambda) = 0$ or $t\delta\lambda = -\lambda\delta t$, therefore the small amount of change in wavelength is equivalent to the change in thickness

$$
\begin{aligned}
\Delta\Psi_{\mathbf{r}} &= \Psi_{\mathbf{r}}(\lambda(t + \Delta t)) - \Psi_{\mathbf{r}}(\lambda t) \\
&= \Psi_{\mathbf{r}}((\lambda - \Delta\lambda)t) - \Psi_{\mathbf{r}}(\lambda, t), \Delta\lambda = -\frac{\lambda}{t}\Delta t
\end{aligned}
\tag{1.20}
$$

The equation can be made use of by both X-ray diffraction, which is relatively easier to change wavelength, and electron diffraction, which is easier to move

23

the sample to change the thickness. Now the multislice iteration is

$$\Psi(\mathbf{r}, t + \Delta t, \lambda) = \Psi(\mathbf{r}, t, \lambda) \exp(r_e|\rho(\mathbf{r})|\Delta t\lambda) \otimes P(\mathbf{r}, \lambda\Delta t) =$$

$$\Psi(\mathbf{r}, t, \lambda + \Delta\lambda) = \Psi(\mathbf{r}, t, \lambda) \exp(r_e|\rho(\mathbf{r})|t\Delta\lambda) \otimes P(\mathbf{r}, t\Delta\lambda) \qquad (1.21)$$

where the Fresnel propagator P is

$$P(r) = \exp(-i\pi|\mathbf{r}|^2/\lambda t) \qquad (1.22)$$

Therefore, in case of X-ray diffraction, it is possible to obtain exponential of the charge density from the difference or division of complex images recorded at two adjacent wavelengths by deconvoluting

$$\frac{\Psi(\mathbf{r}, t, \lambda + \Delta\lambda)}{\Psi(\mathbf{r}, t, \lambda)} = \exp(r_e|\rho(\mathbf{r})|t\Delta\lambda) \otimes P(\mathbf{r}, t\Delta\lambda) \qquad (1.23)$$

Especially when $\Delta\lambda$ is small, the exponential can be approximated by the first order term of expansion, which provides the direct evaluation of $\rho(\mathbf{r})$ (see Fig. 1.15).
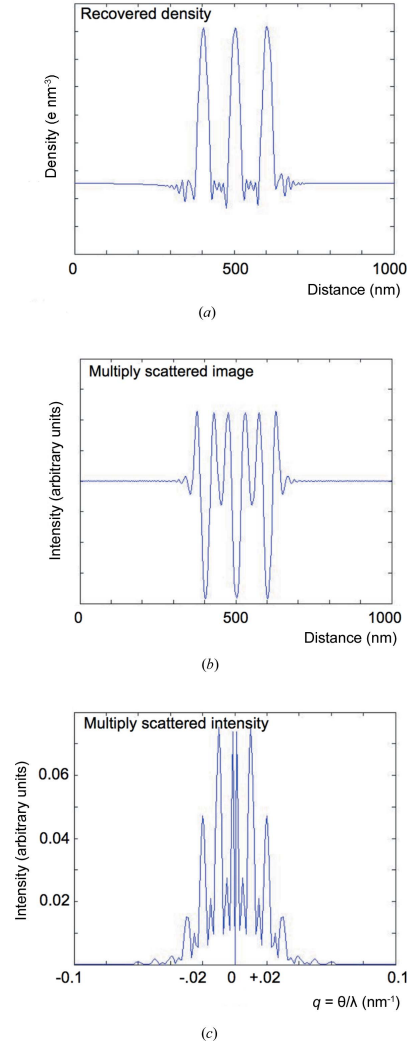
Figure 1.15: (a) The charge density recovered by this inversion algorithm using diffraction-pattern intensities simulated for 500 and 450 eV. The ordinal increment is $50nm^3$. The thickness is 0.5 mm, at the onset of multiple-scattering perturbations, particularly in phase at low angles. (b) The multiply scattered image intensity at 500 eV, showing severe distortion. This is the Fourier transform of the complex pattern shown in (c). (c) Diffraction pattern intensity at 500 eV used to obtain (a), showing strong multiples scattering perturbations. [1]

CHAPTER 2

THE MOLECULAR ALIGNMENT PROBLEM

2.1   Orientation Determination from Three Beam Diffraction Pattern

The use of single-shot X-ray laser pulses from individual molecules has been
suggested as a method for determining the structure of proteins that are dif-
ficult to crystallize[11, 12]. If this should prove technically possible, it will be
necessary to merge many single-molecule two-dimensional diffraction patterns
from molecules lying in random orientations into a single three-dimensional
data set. The subsequent destruction of the sample following the initial elastic
scattering event, however, has precluded the possibility of three-dimensional
(tomographic) imaging of unique structures. Several approaches to the re-
sulting problem of molecular orientation determination have been proposed
[14, 36, 37, 38, 9, 15, 39]. A-priori molecular alignment using experimen-
tal techniques such as liquid flow alignment, laser alignment and alignment
in electrical and magnetic fields have also been proposed and demonstrated
with varying degrees of success[40]. In this section, we suggest a means for
overcoming this limitation.

Consider the arrangement shown in Fig. 2.1. A beam-splitter and reflecting
crystals direct three orthogonal beams onto a non-periodic target particle (only
two beams are shown for clarity) producing three far-field diffraction patterns
prior to destruction of the target. (Thermal vibration prevents overlap of these
patterns.) We assume that all three two-dimensional patterns are read out
after each X-ray pulse, whereupon a new, identical target such as a biomolecule
is inserted in a new orientation. If the phase problem can be solved these
patterns can provide three orthogonal projections of the target charge density.

We now consider the problem of defining an internal ("body", or principal
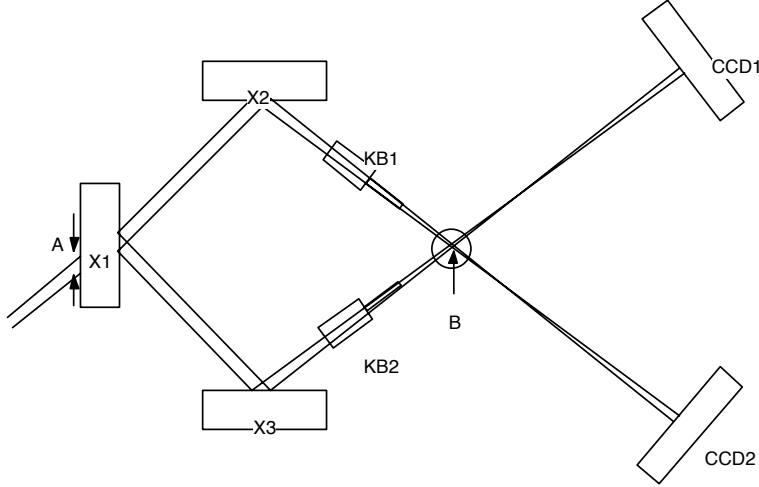
Figure 2.1: Scheme for tomographic femtosecond diffraction, drawn for only two beams for simplicity (Three orthogonal beams are proposed in the text). Beamsplitter X1 is set to the dynamical 3-beam diffraction condition. Crystals X2 and X3 operate at the 2-beam dynamical condition. KB1 and KB2 are Kirkpatrick-Baez focussing mirrors for target at B, with area detectors CCD1 and CCD2.[37]

axes) coordinate system for the target, and of finding the orientation of this with respect to the laboratory frame for the case where the structure of the target is unknown. If this process can be repeated for each particle, the relative orientations of successive particles will also be determined.

The use of principal axes in crystallography has been suggested recently [41]. We propose using the experimental data to determine directly the principal axes of the molecule which provide a natural means of specifying its orientation relative to the laboratory frame defined by the incident probe beams. We assume that multiple scattering can be neglected, i.e., the first Born approximation is valid, so that the patterns have inversion symmetry and the target density is a real function.

In order to expose the principle of the method, we first assume that the phase problem can be solved, by, for example, iterative methods . (We will
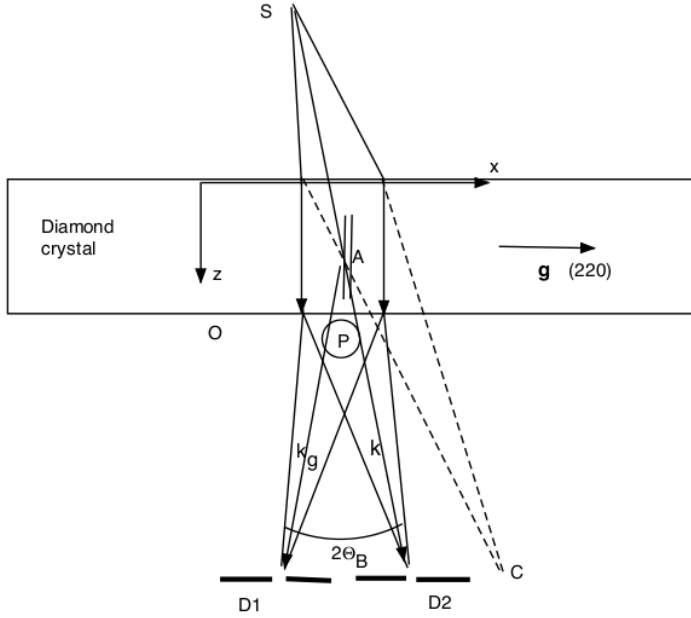
Figure 2.2: Two-beam beamsplitter with sample shown at P lying on the exit face of the beamsplitter. The source S is focused onto two area detectors D1 and D2 containing central beam-dump holes. The two vertical arrows show the direction of the Poynting vector. Three such orthogonal diffracted beams, rather than the two shown, are proposed in the text. [37]

relax this assumption later.) Then, at high energy, each beam delivers a projection in real space, along the direction of the corresponding beam, of the scattering strength per unit volume within the particle. The projections will be referred to different (randomly positioned) origins, and both enantiomorphs (related by inversion symmetry) will be present with equal likelihood. However once a particular enantiomorph is chosen for one projection, the resulting two-dimensional envelope will constrain the choice of enantiomorph for the other two projections.

Consider the moments of the mass density $\rho(\mathbf{r})$ for the target[42]. The zeroth moment delivers the total mass, the first moment delivers the center of mass vector, and the second moment delivers the moment of inertia tensor.

28

By diagonalizing this, the principal axes of the target may be found and hence its orientation relative to the lab frame. Taking the center of mass position as the origin, the inertia tensor is

$$\mathbf{I} = \iiint \rho(\mathbf{r}) \left(r^2 \mathbf{E} - \mathbf{rr} \right) d\mathbf{r}, \tag{2.1}$$

where $\mathbf{E}$ is the unit tensor and $\mathbf{rr}$ is the outer product of the position vector with itself. As with any symmetric tensor, $\mathbf{I}$ has only six independent elements, real eigenvalues, and orthogonal eigenvectors corresponding to different eigenvalues.

We now interpret $\rho(\mathbf{r})$ as the electronic density of the target, whose projections in three orthogonal directions are provided by the phased data, and which define the x, y, and z directions specified by unit vectors $\mathbf{e}_i$ in the lab reference frame. The six independent elements of the "inertia" tensor then have the form

$$\begin{aligned}
I_{zz} &= \iint \rho_z(x,y)(x^2 + y^2) \, dxdy \\
I_{xy} &= \iint \rho_z(x,y)xy \, dxdy
\end{aligned} \tag{2.2}$$

and similarly for I_{xx}, I_{yy}, I_{yz} and I_{xz}. Here ρ_α is the projected density along the α-direction. Two of these six tensor elements can be computed from each of the three projections, e.g. I_{zz} and I_{xy} from the projection along the z-axis. Hence the inertia tensor of the target is fully specified by computing moments and products of inertia from the three projections. While our "inertia" tensor (based solely on electron density) may differ from one based on true mass (including nuclear masses), it need only provide a consistent set of body axes fixed to the molecule to be useful for our purposes. Being symmetric, this tensor may be diagonalized in the usual fashion by constructing the eigenvalue

equations

$$\mathbf{I} \cdot \mathbf{B} = b\mathbf{B} \qquad (2.3)$$

for the three eigenvalues b and corresponding eigenvectors $\mathbf{B}$, solving the secular equation for the eigenvalues, inserting these into the eigenvalue equation, and solving that set of equations for the eigenvectors. These eigenvectors define a new orthogonal coordinate system $\mathbf{e}'_j$ in which the three unit vectors lie along the principal axes of the inertia tensor (the principal axis or "body axis" reference frame). Barring degeneracy among the eigenvalues, the three eigenvectors are unique to within a sign, and therefore offer a natural means of specifying the orientation of the target relative to the incident beam directions (lab frame). With the unit vectors $\mathbf{e}_i$ of the lab frame and $\mathbf{e}'_j$ both known, the angles between the principal axes of the target and the lab frame can immediately be computed. Thus the orientation of the target has not only been defined by introducing the principal axes of the inertia tensor, but also specified (within polarity) relative to the lab coordinates.

To summarize, the procedure to establish the orientation of the particle is as follows: (i) Record three diffraction patterns, one for each of the three incident beam directions. (ii) Invert the diffraction patterns using phase retrieval techniques to yield three real-space projections of the scattering strength. (iii) Compute the first moment of each projection to obtain the center of mass position for that projection. (iv) Compute the second order moments of each projection (products of inertia) about the center of mass to obtain one diagonal and one off-diagonal tensor element. (v) Diagonalize the resulting tensor to obtain the eigenvectors of the tensor. (vi) Compute the orientation of each beam relative to the eigenvectors of the target in order to determine the angles

between laboratory and body (principal axes) coordinates. (vii) If this process is repeated for many successive identical targets in random orientations, their relative orientations can be found, and hence a complete three-dimensional tomographic image can be assembled by standard tomographic techniques such as filtered backprojection.

We now show that the principal axes may be found without the need to solve the phase problem by working with the autocorrelation of the sample density

$$A(\mathbf{r}) = \int d\mathbf{r}' \rho(\mathbf{r} + \mathbf{r}')\rho(\mathbf{r}') . \tag{2.4}$$

A typical product of inertia is

$$I_{xy}^A = \int d\mathbf{r} \; xy A(\mathbf{r}) = \int d\mathbf{r}' \rho(\mathbf{r}') \left[\int d\mathbf{r} xy \rho(\mathbf{r} + \mathbf{r}') \right]$$
$$= \int d\mathbf{r}' \rho(\mathbf{r}') \left[I_{xy}^\rho + x'y'M \right] = 2M I_{xy}^\rho \tag{2.5}$$

where $M = \int d\mathbf{r}\rho(\mathbf{r})$, and we have used the parallel axis theorem to calculate the product of inertia for the shifted coordinates. The principal axes of the autocorrelation function are the same as the principal axes of the corresponding density.

In the high energy projection approximation, the Fourier Transform of each diffraction pattern (intensity) provides a projection of the three-dimensional autocorrelation function of the density, and the analysis simply requires changing $\rho(\mathbf{r})$ to $A(\mathbf{r})$ in Eqn. 2.2.

Alternatively, the moment of inertia can be calculated from the second derivative of the Fourier transform denoted by a tilde,

$$I_{xy}^A = \int d\mathbf{r}^3 A(\mathbf{r}) xy = \lim_{q \to 0} \int d\mathbf{r}^3 A(\mathbf{r}) xy e^{-i\mathbf{q}\cdot\mathbf{r}} = -\lim_{q \to 0} \frac{\partial}{\partial q_x} \frac{\partial}{\partial q_y} I(\mathbf{q}) \tag{2.6}$$

31

where $I(\mathbf{q}) = \tilde{\rho}(\mathbf{q})\tilde{\rho}(-\mathbf{q})$ is the scattering intensity in reciprocal space. So that

$$I_{xy}^A = -[\tilde{\rho}(\mathbf{q})\frac{\partial}{\partial q_x}\frac{\partial}{\partial q_y}\tilde{\rho}(-\mathbf{q}) - \frac{\partial}{\partial q_x}\tilde{\rho}(\mathbf{q})\frac{\partial}{\partial q_y}\tilde{\rho}(-\mathbf{q}) - \frac{\partial}{\partial q_x}\tilde{\rho}(-\mathbf{q})\frac{\partial}{\partial q_y}\tilde{\rho}(\mathbf{q}) + \tilde{\rho}(-\mathbf{q})\frac{\partial}{\partial q_x}\frac{\partial}{\partial q_y}\tilde{\rho}(\mathbf{q})]|_{q=0}$$

(2.7)

Due to the inversion symmetry of $\tilde{\rho}(\mathbf{q})$ at $q = 0$, the first derivative of it at any direction at $q = 0$ is zero, therefore the second and the third terms vanish. So that the equation above can be written as

$$I_{xy}^A = -2\tilde{\rho}(0)\frac{\partial}{\partial q_x}\frac{\partial}{\partial q_y}\tilde{\rho}(\mathbf{q})|_{q=0} = 2\tilde{\rho}(0)\lim_{r\to 0}\int d\mathbf{r}^3\rho(\mathbf{r})xye^{-i\mathbf{q}\cdot\mathbf{r}} = 2MI_{xy}^\rho \quad (2.8)$$

as before. Replacing $\tilde{A}(q_x, q_y, 0)$, corresponding to the high-energy limit, with the correct Ewald sphere diffraction pattern, for an incident wave vector $\mathbf{k}$ along z, shows

$$\partial_{q_x}\partial_{q_y}\tilde{A}\left(q_x, q_y, \sqrt{k^2 - q_x^2 - q_y^2} - k\right)\bigg|_{q_x=q_y=0} =$$
$$\partial_{q_x}\partial_{q_y}\tilde{A}(q_x, q_y, 0)\bigg|_{q_x=q_y=0} \quad (2.9)$$

so that the moments of the Fourier transform of the diffracted data give the same principal axes.

The orientation relationship between the two molecule can be determined with the principle axes of auto-correlation function:

$$\mathbf{R_1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (2.10)$$

However, the other 3 matrixes can also indicate the same principle axes while

not change the handedness of the coordinates:

$$\mathbf{R_2} = \begin{pmatrix} -a_{11} & -a_{12} & a_{13} \\ -a_{21} & -a_{22} & a_{23} \\ -a_{31} & -a_{32} & a_{33} \end{pmatrix}, \mathbf{R_3} = \begin{pmatrix} -a_{11} & a_{12} & -a_{13} \\ -a_{21} & a_{22} & -a_{23} \\ -a_{31} & a_{32} & -a_{33} \end{pmatrix}, \mathbf{R_4} = \begin{pmatrix} a_{11} & -a_{12} & -a_{13} \\ a_{21} & -a_{22} & -a_{23} \\ a_{31} & -a_{32} & -a_{33} \end{pmatrix}$$
(2.11)

The correct one is among the 4 matrixes. However, if we want to move a point on the diffraction pattern to a point in its proper position 3D reciprocal space, we do the operation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix}$$
(2.12)

But if using another matrix, i.e.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} -a_{11} & a_{12} & -a_{13} \\ -a_{21} & a_{22} & -a_{23} \\ -a_{31} & a_{32} & -a_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix}$$
(2.13)

Then obviously

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \neq - \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
(2.14)

hence the Fried's law cannot be applied. It really matters whether proper matrix is chosen when filling the reciprocal space if the object doesn't have the reflection symmetry.

Now consider the intensity of the diffraction $I(\mathbf{r})$ in reciprocal space. When the molecule is rotated according to the matrix $\mathbf{R}$, the operator P_R is applied on it so that $P_R I(\mathbf{r}) = I(\mathbf{R}^{-1}\mathbf{r}) = I(\mathbf{r}')$. The diffraction patterns, which are

33

the only experiment result available, can be regarded the slices of reciprocal space going through the origin if the wave length is not too long. The equation of the slices, namely the diffraction patterns, obtained from x,y,z detectors are I(x=0), I(y=0), I(z=0) for all the orientations.

Therefore, in order to find the common line from the diffraction patterns of different orientation, i.e. two different patterns on z detector, take the first one as reference, then $z = 0$, and the other rotated, $z' = 0$. Then we have the equation:

$$\left. \begin{array}{l} z = 0 \\ z' = (\mathbf{R_1}^{-1}\mathbf{r})_z = a_{13}x + a_{23}y + a_{33}z = 0 \end{array} \right\} a_{13}x + a_{23}y = 0 \qquad (2.15)$$

which is the common line equation in the first(reference) diffraction pattern. The common line equation in the second(rotated) diffraction pattern is just

$$\left. \begin{array}{l} z' = 0 \\ z = (\mathbf{R_1}\mathbf{r'})_z = a_{31}x' + a_{32}y' + a_{33}z' = 0 \end{array} \right\} a_{31}x' + a_{32}y' = 0 \qquad (2.16)$$

Substituting $\mathbf{R_1}$ with $\mathbf{R_2}$, $\mathbf{R_3}$ and $\mathbf{R_4}$ gives all the 4 possible pairs of common lines in the two diffraction patterns:

$$\begin{cases} a_{13}x + a_{23}y = 0 \\ a_{31}x' + a_{32}y' = 0 \end{cases} \begin{cases} a_{13}x + a_{23}y = 0 \\ -a_{31}x' - a_{32}y' = 0 \end{cases} \begin{cases} -a_{13}x - a_{23}y = 0 \\ -a_{31}x' + a_{32}y' = 0 \end{cases} \begin{cases} -a_{13}x - a_{23}y = 0 \\ a_{31}x' - a_{32}y' = 0 \end{cases}$$
$$(2.17)$$

Although the first two pairs and second two pairs indicate the same common lines, at least there will be 2-fold ambiguity instead of 4. The next step is do the same thing on the diffraction patterns obtained by y or z detector and it can be shown that only one of the rotation matrix is consistent, which is the correct one.
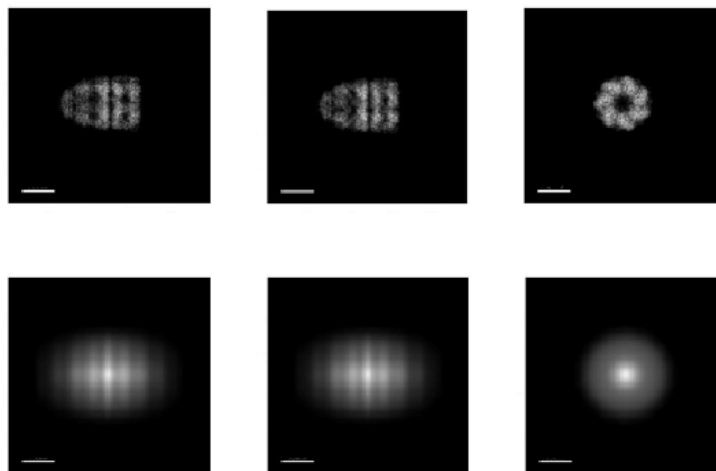
Figure 2.3: This figure shows the three orthogonal projections of the GroEL charge density (upper) and the corresponding projections of the autocorrelation function (lower). The bar indicates 10 nm.

Numerically, once we get the matrix(eigenvector of moment of inertia), make all the 4 rotation matrixes with the correct handedness. It is not necessary to find the common line of the 2 diffraction patterns. Since its possible position can be predicted by rotation matrix, just compare the similarity of the pairs of lines and choose the pair with higher similarity. The corresponding rotation matrix should be the correct one. This method has been verified with several sample problems and gives the correct one for all of them.

We have investigated this procedure using detailed numerical simulations based on data in the Protein Data base for GroEL protein (PDB entry 1SVT) in order to evaluate errors. The three-dimensional density was synthesized from tabulated atomic coordinates. Fig. 2.3 shows the projected densities and corresponding projected autocorrelation functions using the principal axes obtained from Eqn. 2.2. A second density was then generated in a random orientation 2 with respect to the first, as shown in Fig. 2.4. For each of these

35
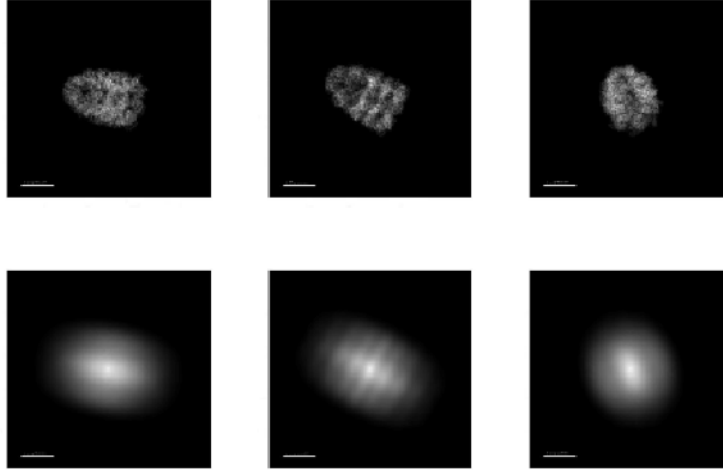
Figure 2.4: This figure shows the projections of GroEL density (upper) and autocorrelation function (lower) in a second random orientation.

orientations the principal axes were determined using both the densities and the autocorrelation functions, giving similar results. When autocorrelation functions were used for the two orientations to determine the principal axes, a rotation matrix needed to rotate the principal axes of the first orientation shown in Fig. 2.4 into the second orientation of Fig. 2.4 we call R.

However as a result of the inversion symmetry in the diffraction patterns and autocorrelation functions (not present in the density), there are three other distinct choices for the rotation matrices whose elements differ from R by alternative choices of the signs of the eigenvectors. Only one of these is correct. The correct rotation matrix may be obtained by applying all these rotation matrices to the diffraction patterns in an attempt to predict the locus along lines of intensity common to two different orientations. (Any two planes in reciprocal space passing through the origin must intersect along a common line). In this way only one rotation matrix will be found to give consistent results. Numerical trials have found this procedure to be reliable with several

different test objects. The use of common lines of intensity here differs from that described elsewhere [14], where it is shown that a complete orientation determination cannot be made from these alone.

We have also considered the case where the beams from the beamsplitter are not orthogonal. Reciprocal vectors can be defined in the usual way, so that each pattern lies in the plane of two of these vectors. Taking the direction of the beam x' and the 2D Cartesian coordinates on the detector x', y' gives a new 3D Cartesian coordinates with respect to the laboratory one (x, y, z). They are related by the matrix R'_{ij} so that $\mathbf{r'} = \mathbf{R'r}$. Generally, if the direction of the three beams are z', z'', z''', the Fourier transformation of the diffraction patterns are actually:

$$
\begin{cases}
F_1 = \int A(x, y, z)dz' \\
F_2 = \int A(x, y, z)dz'' \\
F_3 = \int A(x, y, z)dz'''
\end{cases}
\tag{2.18}
$$

where $A(x, y, z)$ is the auto correlation function.

Calculating the matrix element of the moment of inertia as usual yields a linear combination of these matrix elements:

$$
\int A(x, y, z)x'y'd^3r' = \int A(x, y, z)(R'_{11}x + R'_{12}y + R'_{13}z)(R'_{21}x + R'_{22}y + R'_{32}z)d^3r
$$
$$
= R'_{11}R'_{21}I_{xx} + R'_{12}R'_{22}I_{yy} + R'_{13}R'_{32}I_{zz} +
$$
$$
(R'_{11}R'_{22} + R'_{21}R'_{12})I_{xy} + (R'_{11}R'_{32} + R'_{21}R'_{13})I_{xz} + (R'_{12}R'_{32} + R'_{13}R'_{32})I_{yz}
$$
$$
\tag{2.19}
$$

Do the same thing for $\int A(x, y, z)(x'^2 + y'^2)d^3r'$ and there will be 2 equations for this coordinator system. Thus we have 6 linear equations in total for 6 I_{ij} that we are interested in. Therefore, the products of inertia may be simply evaluated in terms of these reciprocal vectors, and finally transformed into the required lab frame moments.

2.2 Orientation Determination from the Single Beam Diffraction Pattern and its Application to Cryo-EM

In the previous section, we have suggested a method which makes use of diffraction patterns projected simultaneously from one molecule by a single pulse, divided by a beamsplitter, and arriving from three different directions.

In this section a more modification of this method, in which just one beam is needed, is proposed to meet the common experiment setup. In addition to X ray diffraction, it is found especially suitable when applied to Cryo-EM. The routine of Cryo-EM consists of averaging, classification, and then 3D reconstruction. The most common-used 3D reconstruction methods[43] are random-conical the common line method. However, the most important difficulty in molecular biology is specimen damage, which generally prevents sufficient data for a 3D reconstruction to be obtained from a single object. In order to deal with this kind of difficulty, Kam[44] and Provencher[45] designed their own sophisticated schemes using the data of statistical uniformity. Although they both overcome the dose problem, their strict requirement of uniformity can seldom be fulfilled. The method in this paper, originated from X-ray diffraction, is found robust under noise, while depends much less on the uniformity.

Method
Principle Moment of Inertia extracted from Diffraction Patterns

The method is based on a determination of the principle axes of the electronic "mass" distribution $\rho(\mathbf{r})$ of the molecule from its autocorrelation function $A(\mathbf{r})$. Although its general principle has something in common with the three beam scheme, which has been described in detail in the previous section, it is briefly

restated her to make this section self-contained.

The definition of the auto correlation is:

$$A(\mathbf{r}) = \int \rho(\mathbf{r} + \mathbf{r}')\rho(\mathbf{r}) \tag{2.20}$$

A typical product of inertia is

$$I_{xy}^A = \int d\mathbf{r}' x'y' \int \rho(\mathbf{r} + \mathbf{r}')\rho(\mathbf{r})d\mathbf{r} = \int d\mathbf{r}' x'y' \rho(\mathbf{r} + \mathbf{r}') \int d\mathbf{r}\rho(\mathbf{r})d\mathbf{r} \tag{2.21}$$

The second term can be evaluated as

$$\int d\mathbf{r}' x'y' \rho(\mathbf{r} + \mathbf{r}') = \int d\mathbf{t}(t_x - x)(t_y - y)\rho(\mathbf{t}) = \int d\mathbf{t} t_x t_y \rho(\mathbf{t}) + xy \int d\mathbf{t}\rho(\mathbf{t}) = I_{xy}^\rho + xyM \tag{2.22}$$

where M is the total density and the origin is at the center of mass. Then

$$I_{xy}^A = \int d\mathbf{r}\rho(\mathbf{r})[I_{xy}^\rho + xyM] = MI_{xy}^\rho + M\int d\mathbf{r}xy\rho(\mathbf{r}) = 2MI_{xy}^\rho \tag{2.23}$$

In addition, $A(\mathbf{r})$ provides the moments and products of inertia needed to define the inertia tensor of the molecular density, and the eigenvectors of the inertia tensor supply the euler angles which relate the principal axes of the molecule to the laboratory frame coordinates (x,y,z) in which the diffraction patterns are recorded. In the short wave length approximation, where curvature of the Ewald sphere is neglected, the diffraction pattern is a planar slice in reciprocal space, and the Fourier transformation of the diffraction pattern intensities gives us the two-dimensional auto-correlation function $A(x, y)$. This is a projection, taken normal to z, of the three-dimensional autocorrelation $A(x, y, z)$ of the charge density $\rho(r)$. Then the 2 x 2 inertia tensor $\mathbf{Q}$ of the

39

two-dimensional auto-correlation function $A(x, y)$ has elements:

$$Q_{xx} = \int A(x, y)y^2 dxdy = \int A(x, y, z)y^2 dxdydz \quad (2.24)$$

$$Q_{yy} = \int A(x, y)x^2 dxdy = \int A(x, y, z)x^2 dxdydz \quad (2.25)$$

$$Q_{xy} = Q_{yx} \int A(x, y)xydxdy = \int A(x, y, z)xydxdydz = I_{xy} = I_{yx} \quad (2.26)$$

The moments of inertia of the three-dimensional autocorrelation function A(x,y,z), provide the elements of the 3x3 inertia tensor $\mathbf{I}$, such as

$$I_{zz} = \int A(x, y, z)(x^2 + y^2)dxdydz = Q_{xx} + Q_{yy} \quad (2.27)$$

Hence three entries I_{zz}, I_{xy} and I_{yx} in the 3x3 symmetric tensor I (which contains only 6 distinct elements) can be obtained from quantities measurable from the Fourier transforms of the diffraction pattern intensity of each molecule. Because each molecule lies in a different orientation, these quantities will be different for each molecule. The matrix $\mathbf{I}$ can be diagonalized using $\mathbf{P} = \mathbf{R}^{-1}\mathbf{I}\mathbf{R}$, where $\mathbf{R}$ is a rotation matrix (different for each molecule) whose entries are direction cosines of Euler angles, which give the rotation between the laboratory frame (x,y,z) and the principle axes system of the molecule, in which $\mathbf{I}$ is diagonal. Here $\mathbf{P}$ is a diagonal matrix of eigenvalues containing the principle moments of inertia $p1 \geq p2 \geq p3$. This follows using $\mathbf{I} = \mathbf{R}\mathbf{P}\mathbf{R}^{-1}$, since $I_{zz} = R_{31}^2 p1 + R_{32}^2 p2 + R_{33}^2 p3$, where the coefficients are the components of the last row of matrix $\mathbf{R}$. Then, since $R_{31}^2 + R_{32}^2 + R_{33}^2 = 1$, it can be shown that $p1 \geq I_{zz} \geq p3$. Since an equivalent ellipsoid exists for any object, with the same principle moments of inertia, and since the maximum and minimum moments of an ellipsoid run along principle axes, we may find the principle axes of the molecule by searching for maximum and minimum val-

ues of $I_{zz} = Q_{xx} + Q_{yy}$ amongst the diffraction patterns from many randomly oriented molecules.

Firstly, suppose a very large number of diffraction patterns or Cryo-EM images from uniformly random orientations are collected, from each of these, we can calculate its autocorrelation and then the trace of the 2X2 inertia tensor $\mathbf{Q}$ for each molecule. From Equ.(3.18) it can be shown that I_{zz} is a maximum when the lab frame coordinate z runs along the shortest principle axis of the molecule, which will occur by chance. Then $I_{zz} = p1$. Similarly, I_{zz} is minimized when the beam runs along the longest axis. Therefore the maximum and minimum of these traces of $\mathbf{Q}$ are just p1 and p3 respectively

$$p_1 = \max\{Q_{xx} + Q_{yy}\} \tag{2.28}$$

$$p_3 = \min\{Q_{xx} + Q_{yy}\} \tag{2.29}$$

The next step is to find the value of p2. If all the diffraction patterns occur with equal frequency from all possible orientations, we can take the average of all the $Q_{xx} + Q_{yy}$ and this will be equal to the integral of I_{zz} with respect to the Euler angles.We therefore want to write down matrix $\mathbf{I}$ explicitly in terms of the 3 Euler angles. We choose an orientation in which the three principal axes of a molecule coincide with the x,y,z axes in the laboratory coordinator frame, and define these Euler angles as (0,0,0). Then the matrix $\mathbf{I}$ at (α, β, γ) is:

$$\mathbf{I}(\alpha, \beta, \gamma) = \mathbf{RPR}^{-1} \tag{2.30}$$

in which $\mathbf{R}$ is Euler rotation matrix

$$\begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma, & -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma, & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma, & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma, & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma, & \sin\beta\sin\gamma, & \cos\beta \end{pmatrix},$$

and $\mathbf{P} = \begin{pmatrix} p_1 & & \\ & p_2 & \\ & & p_3 \end{pmatrix}.$

This yields

$$I_{zz} = \sin^2\beta\cos^2\gamma \cdot p_1 + \sin^2\beta\sin^2\gamma \cdot p_2 + \cos^2\beta \cdot p_3 \qquad (2.31)$$

The average of I_{zz} is the integral with respect to Euler angle:

$$\bar{I}_{zz} = \frac{1}{8\pi^2} \int_{-\pi}^{\pi} d\alpha \int_0^{\pi} \sin\beta d\beta \int_{-\pi}^{\pi} d\gamma \cdot I_{zz} = \frac{1}{3}(p_1 + p_2 + p_3) \qquad (2.32)$$

Since p1, p3 and I_{zz} are known, p2 can be found from this expression. Note that this is the only place where sttistical uniformity is required. Such uniformity of single molecule diffraction can be replaced by the powder diffraction data without texture, if available.

Even without uniformly random orientation, the extraction of these principle values is still possible from a set of orientation-preferred images. Because large areas contact are energetically preferred[46], the most probable orientation of a molecule is its largest and second largest principle axes lying on the horizontal plane while the least one vertical. The eigenvalues of the corresponding Q are actually $\int A(x,y,z)x^2 dxdydz$, $\int A(x,y,z)y^2 dxdydz$ of the unrotated object. Once the orientation that the largest and smallest principle axes on the horizontal plane is found, either by chance, which is not impossible, or by tilting the sample, the last moment $\int A(x,y,z)z^2 dxdydz$ can also be determined. With this model, and giving that the moment of inertia is not

sensitive under noise (shown in the next section), the three principle values can also be calculated even without random distributed data.

Orientation determination of individual molecules

It remains to determine the orientation of the individual molecules, and hence of their diffraction patterns, from this determination of the eigenvalues of the inertia tensor. We have $\mathbf{I} = \mathbf{R}\mathbf{P}\mathbf{R}^{-1}$, in which $\mathbf{P}$ and some of I are now known, and we wish to find the rotation $\mathbf{R}$ for a particular molecule (given one of its diffraction patterns, which supplied the entries in $\mathbf{I}$). There are three independent elements in the inertia matrix $\mathbf{Q}$ of the two-dimensional auto-correlation Q_{xx}, Q_{yy}, Q_{xy}. Using the fact that $I_{xx} + I_{yy} + I_{zz} = p1 + p2 + p3 = p$, we then have:

$$
\begin{aligned}
Q_{xx} &= \tfrac{p}{2} - I_{xx} = \tfrac{p}{2} - R_{11}^2 p_1 + R_{12}^2 p_2 + R_{13}^2 p_3 \\
Q_{yy} &= \tfrac{p}{2} - I_{yy} = \tfrac{p}{2} - R_{21}^2 p_1 + R_{22}^2 p_2 + R_{23}^2 p_3 \quad\quad (2.33) \\
Q_{xy} &= I_{xy} = R_{11} R_{21} p_1 + R_{12} R_{22} p_2 + R_{13} R_{23} p_3
\end{aligned}
$$

in which R_{ij} is the matrix element of $\mathbf{R}$ and due to the orthogonality of $\mathbf{R}$, only 3 of 6 elements are independent, corresponding to the 3 Euler angles. The equations can be solved since we have 3 of them for 3 unknown variables.

Practical algorithm

The nonlinear equations above are not only difficult to solve numerically but also sensitive to errors due to noise or error during measurement. A method needs to be found to transform the equations to a linear problem. This can be done if the remaining two elements in I can be found.

Note that from one diffraction pattern we have 4 of the 6 independent moments of inertia matrix elements. The other two can be solved from the scalar equations with the known eigenvalues, so that only 2 of 3 are needed

since the third can be determined from the trace. The secular equation is of the form:

$$
\begin{vmatrix}
X & c & a \\
c & Y & b \\
a & b & Z
\end{vmatrix} = 0
\qquad (2.34)
$$

,in which a, b are unknown elements, $c = Q_{xy}$ and X, Y, Z are constants, dependent on the specific eigenvalues. After some algebra this equation can be written as:

$$
\begin{aligned}
Y_1 a^2 - 2cab + X_1 b^2 &= D_1 \\
Y_2 a^2 - 2cab + X_2 b^2 &= D_2
\end{aligned}
\qquad (2.35)
$$

where the subscripts indicate the different corresponding eigenvalues p_i. These lead to a quadratic equation about b^2 and can be solved analytically. Once the moment of inertia matrix is found, the rotation matrix can be easily obtained by calculating its eigenvectors.

It often occurs for some proteins (such as GroEL) that two of the principal values are almost identical. Then only one principal axis can be found and the other two lie in the plane normal to that axis. The current method is not able to deal with this difficulty, however, it is able to reduce the orientation problem from 3 parameters to just 1 parameter, which is accessible to the newly developed GTM method[36] and related fiber-diffraction techniques (for an application of iterative phasing to fiber diffraction, see[38]).

Result and Discussion

The method above suffers some ambiguities. If $R'_{1j} = -R_{1j}$ and $R'_{2j} = -R_{2j}$ the equations are still satisfied for R'. However, on looking into the form of $\mathbf{R}$, we find that this ambiguity occurs when $\alpha' = \alpha + \pi, \beta' = \beta, \gamma' = \gamma$. Considering

also the inversion symmetry of auto-correlation functions, we can choose any of them, and that would not affect the reconstruction of the 3D auto-correlation function. Another ambiguity is $\alpha' = \alpha, \beta' = \pi - \beta, \gamma' = \pi - \gamma$ and this leads to a different orientation which is not consistent with the Friedel's law. This ambiguity corresponds to the 2 possible solutions (a,b) and (-a,-b) for Eqn.(5.25).

In the previous section, we showed that the moments of inertia we calculated using a curved Ewald Sphere (in the the long wave length case) gives the same result as a flat Ewald sphere. This proof remains applicable here, so that we can ignore curvature of Ewald sphere.

In order to apply this method experimentally, the influence of noise must be taken into account. We should not expect to get diffraction patterns as good as in these simulations. The most important source of error is from insufficient photons at the detector in order to prevent the radiation damage, and this kind of error can be described as Poisson noise[14].

The other problem with X ray diffraction is the beam stop. All the detectors used in the synchrotrons have a hole in the center so that the central beam would not damage the detector. However, without the central data the autocorrelation calculated from the diffraction pattern would be zero. In order to fix it, the shrinkwrap[28] method can be applied to reconstruct this part of data, with an additional 10% percent of error introduced. Therefore, Cryo EM, with the central beam diverged to the whole image, is more favorable to this method.

Now first consider the case when noisy real space image is available, and the noise is mainly Poisson noise. Define the relative error of the noisy image, its corresponding autocorrelation and $\mathbf{Q}$ tensor to be:
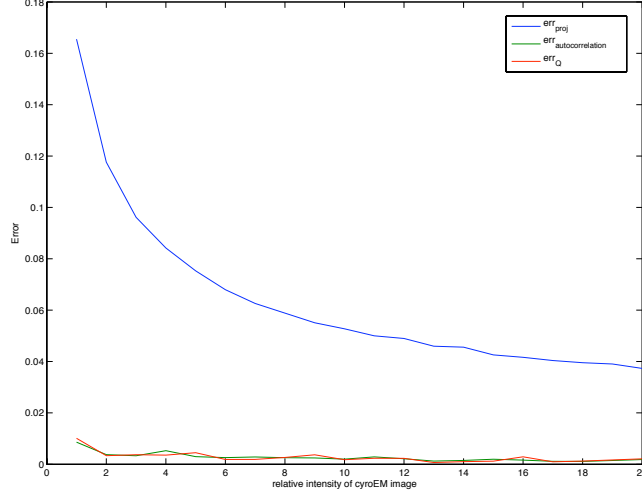
Figure 2.5: This figure shows the dependence of the relative error of Cryo-EM images with Poisson noise. The x-axis is the relative intensity of the original image and the y-axis is the error of the noisy image from Eqn.(2.36). The increase of dose help reduce the Poisson noise of the image, while the other indirect measurement of electrical potential project suffer much fewer noise.

$$
\begin{aligned}
err_{img} &= \frac{\sum\limits_{all\ pixels} |Img_{noisy} - Img_{noise-free}|}{\sum\limits_{all\ pixels} |Img_{noise-free}|} \\
err_A &= \frac{\sum\limits_{all\ pixels} |A_{noisy} - A_{noise-free}|}{\sum\limits_{all\ pixels} |A_{noise-free}|} \\
err_Q &= \frac{\sum\limits_{all\ elements} |Q_{noisy} - Q_{noise-free}|}{\sum\limits_{all\ elements} |Q_{noise-free}|}
\end{aligned}
\tag{2.36}
$$

The change of these errors with the intensity of initial image is plotted and it shows that during the calculation of autocorrelation error is greatly reduced. This result is not hard to predict since each pixel in autocorrelation depends on the pixels of the whole original image, and the error is cancelled during the accumulation. Actually this is the primary reason to transform potential projection to autocorrelation.

In the end the algorithm is tested for a specific case. As shown in Fig.(2.5), we have studied a GroEL protein molecule, lying in all possible orientations

46
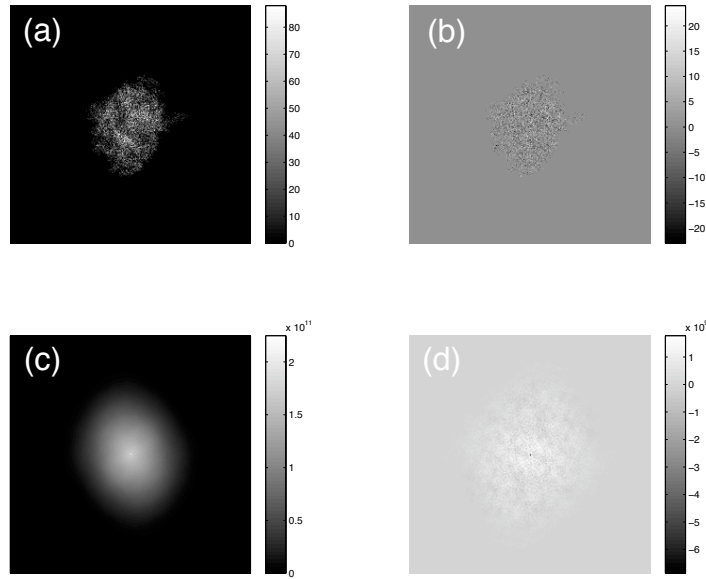
Figure 2.6: (a) The simulated noisy charge density projection of GroEL in grey scale (the size of the molecule $\simeq 50\text{Å}$ and orientation specified by Euler angles (45,60,80) degrees; (b) The difference between (a) and original value, the relative error is about 20%; (c) The autocorrelation calculated from (a); (d) The difference between (c) and autocorrelation calculated and true value, the relative error is of the order 1%

within a 70nm×70nm cell. For several different set of Euler angles, it is found that the first column of the pre-assigned rotation matrices and reconstructed rotation matrix agree quite well with an error of 1 degrees, which is predictable from the low error of the $\mathbf{Q}$ tensor. This indicates that the axis with a distinct eigenvalue is readily obtained. The agreement of the other two columns is poor, since the other two eigenvalues of GroEL are almost identical. However, there is only one undetermined Euler angle now, and such problem can be easily solved by those methods using the detail of the image, such as GTM[36].

2.3 Generative Topology Mapping method used for Orientation Determination

As the orientation of single molecule is hard to be pre-assigned when its snapshot is taken, it is essential to retrieve such information from a set of diffraction patterns (DP) taken from unknown orientations to enable the merging of intensities. Generative topographic mapping (GTM) has been used in data mining to reveal such intrinsic latent variables from the distribution of the observable data. In this case, the latent variables are the Euler angles that determine the orientation, and the data are experimental DP, which is a vector whose components come from every pixel of the DP. In the application here, compared with the straightforward method such as common line, the statistical method GTM is used to benefit the sorting of diffraction pattern from two of its properties:

1. making full use of flux

2. robustness under noise

GTM help figure out the mapping relationship from latent space to data space, as shown in Fig. 2.7, which make it possible to map the experimental data back to orientation.

GTM helps figure out the mapping relationship from latent space to data space, thus makes it possible to map the experimental data back to orientation. In numerical calculation, however, both of the two spaces are discrete, therefore we are dealing with the data points in them. Ideally, point in L space corresponds to point in the $n \times n$ dimensional D space. The GTM method works provided that the mapping function $\mathbf{y} = (\mathbf{x}; \mathbf{W})$, where $\mathbf{W}$ is a set of
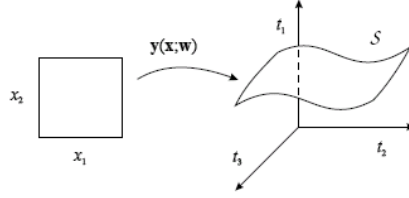
Figure 2.7: The non-linear function y(x;W) defines a manifold S embedded in data space given by the image of the latent-variable space under the mapping x to y.[47]

mapping parameters, is smooth and continuous, so that the projected points $\mathbf{y} = (\mathbf{x}_i; \mathbf{W})$ will necessarily have a topographic ordering in the sense that any two points $\mathbf{x}_A$ and $\mathbf{x}_B$ that are close in L space will map to points $\mathbf{y}_A$ and $\mathbf{y}_B$, which are close in data space.

In order to take account of the noise, the point in D space can expand to a sphere of continuous density, as shown in Fig.2.8. We choose the distribution of $\mathbf{t}$, for given $\mathbf{x}$ and $\mathbf{W}$, to be a radially symmetric gaussian centered on $\mathbf{y}=(\mathbf{x};\mathbf{W})$ having variance β^{-1} so that

$$p(\mathbf{t}|\mathbf{W}, \beta) = (\frac{\beta}{2\pi})^{D/2} \exp\{-\frac{\beta}{2}\|\mathbf{y}(\mathbf{x}; \mathbf{W}) - \mathbf{t}\|^2\} \qquad (2.37)$$

This function describes the probability of certain DP is related with some orientation if the correct mapping (W,β) is known. Fig.2.9 is the schematic diagram for such mapping where there is just one latent variable, the angle of an object that can rotate along one axis (not necessarily the same direction of incident beam), and the data space is reduce to the intensities on three pixels for the sake of visualization, which can be regarded as the projection of the manifold to the 3D subspace. The distribution of sampling in latent space is known as uniform, and the measurement distributes inside the density sphere.
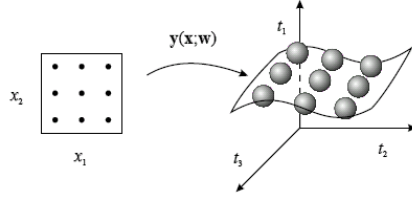
Figure 2.8: We consider a prior distribution p(x) consisting of a superposition of delta functions, located at the nodes of a regular grid in latent space. Each node x_i is mapped to a corresponding point $y(x_i; W)$ in data space, and forms the center of a corresponding Gaussian distribution.[47]
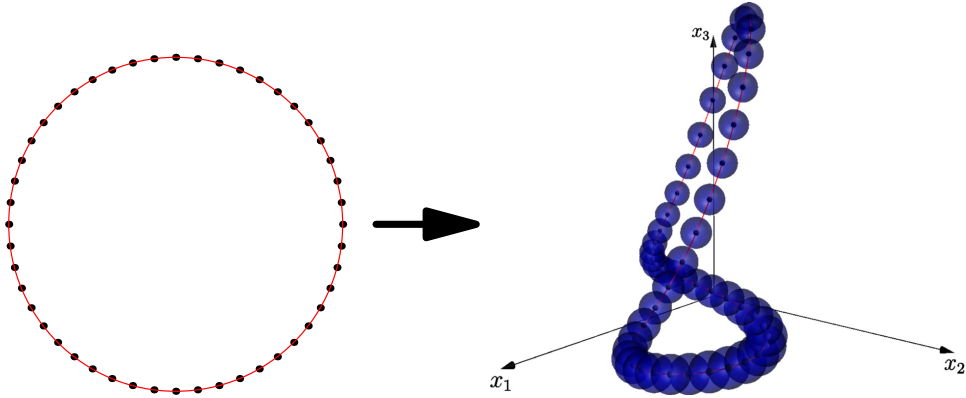
Figure 2.9: The mapping from latent variable orientation to the scattering intensities on 3 pixels. Note that the point diffuses into density sphere, and the compactness of latent space is preserved in the data space

The GTM method works only when the prior distribution $p(\mathbf{x})$ of $\mathbf{x}$ in L space is known. But in our experiment the DP is obtained with pulse of XFEL, $p(\mathbf{x})$ can be regarded as a sum of delta function, , then the distribution in D space is then obtained analytically by

$$p(\mathbf{x}) = \frac{1}{K} \sum_{i=1}^{K} \delta(\mathbf{x} - \mathbf{x}_i), \tag{2.38}$$

then the distribution in D space is then obtained by

$$p(\mathbf{t}|\mathbf{W}, \beta) = \int p(\mathbf{t}|\mathbf{x}, \mathbf{W}, \beta) p(\mathbf{x}) d\mathbf{x} = \frac{1}{K} \sum_{i=1}^{K} p(\mathbf{t}|\mathbf{x}_i, \mathbf{W}, \beta) \tag{2.39}$$

50

Apparently, only $\mathbf{y}(\mathbf{x}_i)$ close to $\mathbf{t}$ contributes to the probability. From that we can define a log likehood function

$$L(\mathbf{W}, \beta) = \ln \prod_{n=1}^{N} p(\mathbf{t}_n | \mathbf{W}, \beta) = \sum_{n=1}^{N} \ln \frac{1}{K} \sum_{i=1}^{K} p(\mathbf{t}_n | \mathbf{x}_i, \mathbf{W}, \beta). \qquad (2.40)$$

If the mapping parameters $\mathbf{W}$ are correct so that the manifold in D space go through the center of the swarm of data points, we can image that every term in the above summation reaches its maximum and the likehood function reaches its maximum. Therefore, we can determine $\mathbf{W}$ and β by maximizing L.

The way we do that is EM algorithm. We can choose $\mathbf{y}=(\mathbf{x};\mathbf{W})$ to be given by a linear model so that the component of $\mathbf{y}$ can be written as

$$y_i = W_{i1}\phi_1(\mathbf{x}) + W_{i2}\phi_2(\mathbf{x}) + \ldots + W_{iM}\phi_M(\mathbf{x}), \qquad (2.41)$$

in matrix form

$$\mathbf{y} = \mathbf{W}\phi(\mathbf{x}) \qquad (2.42)$$

The base function ϕ_i is generally choose to be radially symmetric gaussian whose centers are distributed on a uniform grid in L space with a common width parameter assigned at the beginning. Such parameters and the number of sample points in L space should be able to make the manifold smooth enough while keep the precision.

Now we can start the training iterations from a initial set of parameters $\mathbf{W}_{old}, \beta_{old}$. Apparently, there is no direct mapping relationship between the sample point in L space and experiment point in D space. But we can obtain

the contribution, or responsibility of every sample point to a single data point

$$
\begin{aligned}
R_{in}(\mathbf{W}_{old}, \beta_{old}) &= p(\mathbf{x}_i|\mathbf{t}_n, \mathbf{W}_{old}, \beta_{old}) \\
&= \frac{p(\mathbf{t}_n|\mathbf{x}_i, \mathbf{W}_{old}, \beta_{old})}{\sum\limits_{i'=1}^{K} p(\mathbf{t}_n|\mathbf{x}_i, \mathbf{W}_{old}, \beta_{old})}.
\end{aligned}
\tag{2.43}
$$

Also only those can be mapped very close to $\mathbf{t}_n$ can make primary contribution to it.

Now the expectation of the log likelihood is

$$
\langle L(\mathbf{W}, \beta) \rangle = \sum_{n=1}^{N} \sum_{i=1}^{K} R_{in}(\mathbf{W}_{old}, \beta_{old}) \ln[p(\mathbf{t}_n|\mathbf{x}_i, \mathbf{W}, \beta)].
\tag{2.44}
$$

Maximizing Eq.2.44 with respect to $\mathbf{W}$, we obtain

$$
\sum_{n=1}^{N} \sum_{i=1}^{K} R_{in}(\mathbf{W}_{old}, \beta_{old})[\mathbf{W}_{new}\phi(\mathbf{x}_i) - \mathbf{t}_n]\phi^T(\mathbf{x}_i) = 0.
\tag{2.45}
$$

This can conveniently be written in matrix notation in the form

$$
\Phi^T \mathbf{G}_{old} \Phi \mathbf{W}_{new}^T = \Phi^T \mathbf{R}_{old} \mathbf{T}
\tag{2.46}
$$

where Φ is a $K \times M$ matrix with elements $\Phi_{ij} = \phi_j(\mathbf{x}_i)$, $\mathbf{T}$ is a $N \times D$ matrix with elements t_{nk}, $\mathbf{R}$ is a $K \times M$ matrix with elements R_{in}, and $\mathbf{G}$ is a $K \times K$ diagonal matrix with elements

$$
G_{ii} = \sum_{n=1}^{N} R_{in}(\mathbf{W}, \beta)
\tag{2.47}
$$

Here the M, N, K, and D are the number of basic functions, the number of data points $(t_1, ..., t_N)$, the number of latent space sampling points, and the number of dimensions of data space, respectively.

Now the $\mathbf{W}_{new}$ can be solved with standard matrix techniques, based on singular value decomposition to allow for possible ill conditioning. Note that Φ is constant and only need to be evaluated once at the start.

This training process of the data set should be repeated until the likelihood function stop increasing. Meanwhile the $\mathbf{W}_{new}$ will also be changed a very little bit. Then we can regard the mapping with this $\mathbf{W}$ are most close to the true one.

With the above mapping relationship, taking any DP, its orientation, which is represented as $\mathbf{x}$, can be estimated by summarizing the posterior by its mean, given by

$$\langle \mathbf{x}|\mathbf{t}_n, \mathbf{W}, \beta \rangle = \sum_{i=1}^{K} R_{in}\mathbf{x}_i \qquad (2.48)$$

A more straightforward way is to calculate the distance of this DP in D space with those generated from sample points in L space had we know the mapping relationship. Finding out which sample point can give the DP closest to the actual one. If the sample points are dense enough, the error between the two methods is actually equivalent.

The GTM algorithm is purposed by Bishop et al [47] and successfully applied to diffraction pattern classification by Fung et al [36].

2.4 Expansion-Maximization-Compression method

Loh et al[48] proposed the Expansion-Maximization-Compression (EMC) algorithm for reconstructing a particle's 3D diffraction intensity from very many diffraction patterns, when the orientation in each pattern is not determined. The algorithm consists of a maximization step (M) of a logarithm likelihood function, a expansion step (E) and a compression (C) steps that map the 3D intensity model to a redundant tomographic representation and back again. The working process of the method is concluded as follow.

Firstly, define $W_{ij} = W(\mathbf{R}_j \cdot \mathbf{q}_i)$ as the average photon number of model at detector pixel i when the particle has orientation j, $W(\mathbf{q})$ the time-integrated

scattered intensity at spatial frequency $\mathbf{q}$ when the particle is in some reference orientation, and K_{ij} the measurement of photon number at detector pixel i when the particle has orientation j.

The algorithm is based on expectation maximization (EM), which reconstruct a model from statistical data that is incomplete by maximizing a log-likelihood function $Q(W')$, then updating the model from $W \to W'$. It is necessary to determine the form of the likelihood function at first. The logarithm likelihood function for the photon number W'_{ij} is the logarithm of the Poisson distribution:

$$Q_{ijk}(W') = K_{ik} \log W'_{ij} - W'_{ij}. \tag{2.49}$$

where K_{ik} is the photon count at pixel i in measurement k. Therefore the log-likelihood function for a single diffraction pattern with independent Poisson distribution on each pixel can be written as the summation of $Q_{ijk}(W')$

$$Q_{jk}(W') = \sum_{i=1}^{M_{\text{pix}}} Q_{ijk}(W'). \tag{2.50}$$

Then define the the conditional probability of each diffraction pattern

$$R_{jk}(W) = \prod_{i=1}^{M_{\text{pix}}} W_{ij}^{K_{ik}} \exp\left(-W_{ij}\right). \tag{2.51}$$

Now any prior distribution of the orientation j can be written as

$$P_{jk}(W) = \frac{w_j R_{jk}(W)}{\sum_j w_j R_{jk}(W)}. \tag{2.52}$$

given the normalized weights of orientations w_j. Then it is possible to write down the total log-likelihood function of the whole set of diffraction patterns

$$
\begin{aligned}
Q(W') &= \sum_{k=1}^{M_{\text{data}}} \sum_{j=1}^{M_{\text{rot}}} P_{jk}(W) Q_{jk}(W') \\
&= \sum_{i=1}^{M_{\text{pix}}} \sum_{j=1}^{M_{\text{rot}}} (A_{ij} \log W'_{ij} - B_j W'_{ij})
\end{aligned}
\tag{2.53}
$$

54

where

$$A_{ij} = \sum_{k=1}^{M_{\text{data}}} P_{jk}(W)K_{ik}$$

$$B_j = \sum_{k=1}^{M_{\text{data}}} P_{jk}(W).$$

The EMC iterations can be started from random 3D intensities. In the E-step, the initial or previously calculated 3D intensities W on the grid are expanded into a tomographic model W_{ij} of diffraction patterns for the calculation of log-likehood function by rotation and interpolation.

$$W'_{ij} = \sum_{\mathbf{p}} f(\mathbf{p} - \mathbf{R}_j \cdot \mathbf{q}_i)W(\mathbf{p}). \tag{2.54}$$

Although the W_{ij}s are redundant, they are treated as independent variables by the next step.

In the M-step, the data is classified and their aggregation into tomographic model is improved by maximizing the log-likehood function. Such function is very easy to maximize noting that each term of the summation (2.53) is of the form $a \log W - bW$ where a and b are positive constants and all these terms are independent. The global maximum can be obtained when each term is maximized, which updates the W_{ij} into W'_{ij}

$$W'_{ij} = A_{ij}/B_j \tag{2.55}$$

In the C-step, the redundant W_{ij}s are "condensed/compressed" back into the intensities on the regular 3D grid. In this process, it is necessary to define the interpolation weights $f(\mathbf{q})$ that vanish for large $|\mathbf{q}|$ and is normalized

$$1 = \sum_{\mathbf{p}} f(\mathbf{p} - \mathbf{q}) \tag{2.56}$$

Then the 3D intensities are given by

$$W(\mathbf{p}) = \frac{\sum_{i=1}^{M_{\text{pix}}} \sum_{j=1}^{M_{\text{rot}}} f(\mathbf{p} - \mathbf{R}_j \cdot \mathbf{q}_i) W_{ij}}{\sum_{i=1}^{M_{\text{pix}}} \sum_{j=1}^{M_{\text{rot}}} f(\mathbf{p} - \mathbf{R}_j \cdot \mathbf{q}_i)}. \tag{2.57}$$

The progress of iterations are monitored by the update magnitudes:

$$\Delta W^2 = \langle |W'(\mathbf{p}) - W(\mathbf{p})|^2 \rangle_{\mathbf{p}}. \tag{2.58}$$

The vanishing of ΔW is used as the stopping criterion for the merging stages, and the reconstructed intensities can b used in further phasing stage.

CHAPTER 3

HYDRATION EFFECT ON CXDI IMAGE RETRIEVAL
3.1   Introduction

In order to solve the structure of proteins that are hard to crystallize, new experimental methods called femtosecond X-ray single molecule diffraction have been proposed to do the crystallography of single molecules[11, 15, 40]. Since most of the conventional methods in X-ray crystallography dealing with crystal samples are not applicable here. Therefore, new algorithms concerning the sorting[14, 39, 38] and phasing[27] of single molecule diffraction patterns (DPs) have been proposed and tested to ensure their effectiveness and robustness under all kinds of error and noise. Some of them have been proved to be feasible when applied to simplified sample objects[18].

The next step toward the application of these methods on real world protein is to add the water background, since the experiment requires that the single molecule sample should be dissolved and wrapped in the water so that aggregation can be avoided. The single molecule sample can be ejected one by one to scatter the incident x-ray and increase the hit rate. The nozzles made at ASU[49, 50], which have been used at the LCLS (Linac Coherent Light Source), are able to generate droplets or liquid stream with the dimensions as small as 1 micron.

In recent work[51], simulation is carried out on a satellite tobacco necrosis virus (STNV), whose capsid structure has been solved by x-ray crystallography (Protein Data Bank ID: 2BUK): object size 17 nm, icosahedral symmetry. Realistic water shells around the virus using the Tip3P model of liquid water with average thicknesses of 0.5, 1.5, and 2.5 nm is added. It is found that the limiting orientational (angular) resolution is weakly inuenced by water layer

or Poisson noise. The effects of the random water layer or Poisson noise on the relative error are of comparable magnitude for q below the water peak value. For q above the water peak value the effect of WL is dominant and leads to a large relative error. However, the water layer here is too thin to be realistic.

Compared with the radius of a typical protein molecule (around 10nm), the amount of water is so significant that it can be easily expected that the DP of the current experimental setup cannot be simply that of the single molecule alone. It is calculated in the next section that generally it is impossible to separate the DP into the water part and protein part due to the coupling term. A method to overcome the difficulty is also suggested and tested against all kinds of possible experimental errors.

3.2   Water window

The water window is the soft x-ray energy range between the carbon and the oxygen K edges, where water has a much lower x-ray attenuation coefficient than carbon-containing cells. The contrast between water and protein is therefore magnified, and the technique has been widely used in the microscopy of cells.

Theory

Firstly, it should be noted that the effect of exponential decay is a dynamics effect. The macroscopic refractive index is related with the microscopic structure factor by

$$n = 1 - \frac{n_a r_e \lambda^2}{2\pi}(f_1 + if_2) \tag{3.1}$$

where f_1 and f_2 are the real and imaginary part of the structure factor respectively. f_1 and f_2 are energy-dependent and have been tabulated by Henke et al[52] for all the elements. When the atom is much smaller than the X-ray

58

wavelength, as is the case in the soft X-ray region, the scattering amplitude of each electron can be added coherently for all the ks , where the f_1 and f_2 table is applicable and independent of k.

Start from Schrodinger equation, neglecting the spin and doing the scalar scattering theory

$$\nabla^2\psi + i4\pi\mathbf{k}_0 \cdot \nabla\psi = -4\pi^2(f_1 + if_2)\psi \tag{3.2}$$

and separate the z-component from the xy-component

$$\frac{\partial^2\psi}{\partial z^2} + i4\pi k_{0z}\frac{\partial\psi}{\partial z} + \nabla^2_{xy}\psi + i4\pi\mathbf{k}_0 xy \cdot \nabla_{xy}\psi = -4\pi^2(f_1 + if_2)\psi \tag{3.3}$$

If the high-energy approximation that the Laplacian derivative along the bean direction is negligible can be used here, the equation becomes

$$\frac{\partial\psi}{\partial z} = (\bar{\nabla}^2 + \bar{V})\psi \tag{3.4}$$

where

$$\bar{\nabla}^2 \equiv \frac{i}{4\pi k_{0z}}(\nabla^2_{xy} + i4\pi\mathbf{k}_{xy} \cdot \nabla_{xy}) \tag{3.5}$$

$$\bar{V} \equiv \frac{i\pi}{k_{0z}}(f_1 + if_2)$$

Therefore, the geometry of the problem $(\bar{\nabla}^2)$ is separated from the electron interaction $(\bar{V})$ and can be treated respectively.

For a very thin slice of material, it is reasonable to set $\bar{\nabla}^2 = 0$, which implies no lateral scattering, and the equation is solved by direct integral along the slice of thickness Δz

$$\psi(x, y) = e^{\int_0^{\Delta z} \bar{V}dz} = e^{\frac{i\pi}{k_{0z}}f_1\Delta z}e^{-\frac{\pi}{k_{0z}}f_1\Delta z} \tag{3.6}$$

which is of the same form as the exponential decay in classical electrodynamics, where the media is assumed uniform so that the assumption $\bar{\nabla}^2 = 0$ also

establishes due to the translational symmetry on the transverse xy-plane. On the other hand, it is not guaranteed that the index of the exponential is small enough to be approximated by its first order

Then, for the vacuum, we solve

$$\frac{\partial \psi}{\partial z} = \bar{\nabla}^2 \psi \tag{3.7}$$

Note that the second term in 3.5 vanishes if the beam is incident along the z-axis. Doing Fourier transformation on x, y components and transforming back leads to the solution that

$$\psi(x, y, \Delta z) = e^{\Delta z \bar{\nabla}^2} \psi(x, y, 0) = \mathscr{F}^{-1}[e^{-i\pi\lambda\Delta z q^2}] \otimes \psi(x, y, 0) \tag{3.8}$$

If the scattering object is divided into N slices of thickness Δz, then we have

$$\psi(x, y, \Delta z) = e^{\Delta z \bar{\nabla}^2} e^{\bar{V}^N \Delta z} \cdots e^{\Delta z \bar{\nabla}^2} e^{\bar{V}^2 \Delta z} e^{\Delta z \bar{\nabla}^2} e^{\bar{V}^1 \Delta z} \psi(x, y, 0) \tag{3.9}$$

which is the form of the Cowley-Moodie multislice method[53] in real space. Thibault et al[31] claimed that this method uses the small angle approximation and developed an actually equivalent propagation method. However, it is reasonable to apply this assumption for even soft X-ray scattering if the slice is thin enough. The transverse scattering is already included in 3.8.

Additionally, it is worth mentioning that for materials with a real potential, for its first term in the Born expansion, the optical theorem give the total cross section as zero, which surely cannot be true. This also indicates that the first Born approximation is not suitable to the absorption problem. If we consider up to the second Born approximation

$$\begin{aligned}
\psi^{(2)}(\mathbf{r}) &= \frac{\mu}{4\pi} \int \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} V(\mathbf{r}')\psi^{(1)}(\mathbf{r}') d\mathbf{r}' \tag{3.10} \\
&= \frac{\mu^2}{16\pi^2} \int d\mathbf{r}' \int d\mathbf{r}'' \frac{e^{-ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \frac{e^{-ik|\mathbf{r}'-\mathbf{r}''|}}{|\mathbf{r}'-\mathbf{r}''|} V(\mathbf{r}')V(\mathbf{r}'')
\end{aligned}$$

60

Here $\psi^{(1)}$ cannot be written as

$$\frac{\mu}{4\pi}\frac{e^{ik_0R}}{R}\int V(\mathbf{r}')e^{i\mathbf{q}\cdot\mathbf{r}'}d\mathbf{r}' = \frac{\mu}{4\pi}\frac{e^{ik_0R}}{R}\tilde{V}(\mathbf{q}) \qquad (3.11)$$

which is the Fourier transformation of the potential and quite widely used when single scattering approximation is used. However, this form implies that $R = |\mathbf{r}'|$ is very large compared with the dimensions of the scattering field $|\mathbf{r}''|$, and this approximation fails in the integral of Eqn.(3.10) as it has to be evaluated everywhere in $|\mathbf{r}'|$. The result is that the second term of Born series Eqn.(3.10) is no longer real given the real potential V and the total cross section is nonzero in the optical theorem.

Simulation

The multislice method is implemented to see the effect of multiple scattering and difference in attenuation coefficients of water and protein on the exit wave. The PSI molecule (1JB0) is chosen as the sample and it is immersed in a water droplet with varying size. The refractive indices were obtained from CXRO library ($\delta_{water} = 0.0022, \delta_{protein} = 0.0019, \beta_{water} = 0.000157, \beta_{protein} = 0.00128$);

The simulation shows that the exit wave is quite different from the projection of charge density. Although the contrast is enhanced due to the difference in attenuation factors of water and protein.

Some kind of inversion algorithm needs to be developed to convert the exit wave back to the original structure.
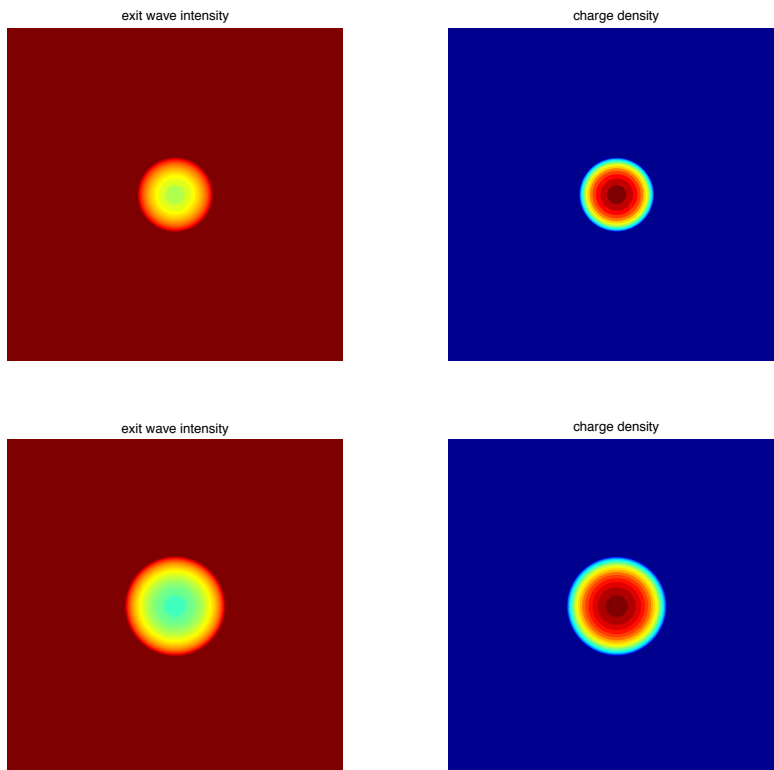
Figure 3.1: This figure shows the exit wave intensity calculated from multiple slice method (slice thickness is 5Å) and charge density projection of protein PSI in water droplet of radius 56Å(a) and 94Å(b).

3.3 Calculating the diffraction pattern of the protein molecule within water jacket

When the protein molecule is immersed in water, it is well known that its structure changes. In simulation, this can be done by putting such a molecule inside a spherical container and the molecular dynamics software (in this paper VMD/NAMD[54]) is used to fill the sphere with water molecules and find the position of every atom in the sphere when the system is stable (see Fig. 3.2 for example). However, due to the limitation of computational ability, it is impossible to get the position of each atom in a droplet with practical size,

i.e. 1000Åin radius. Therefore, some structural properties of water are used to make the calculation feasible while realistic.
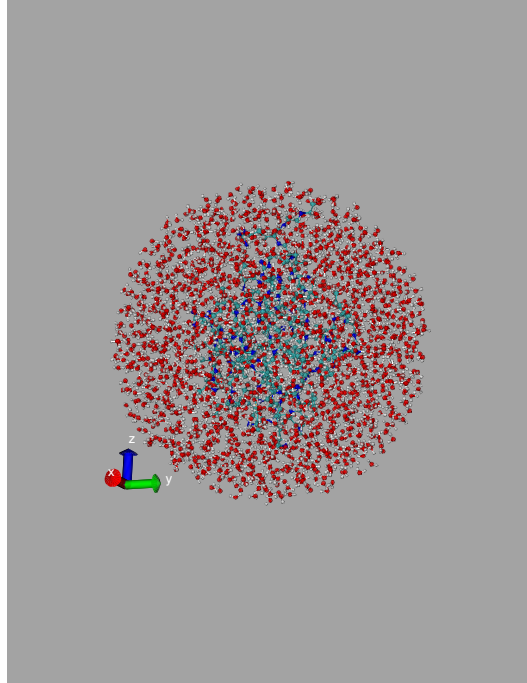
Figure 3.2: This figure shows the atom model of protein 1UBQ inside water droplet as the result of molecular dynamics simulation. The inner blue atoms are protein and the outer red ones are water.

In the calculation of the spherical droplet, it is divided into 3 regions.

- The core of the droplet is as small as possible to contain the protein molecule. The rest of the volume is filled with water by molecule simulation.

- A thin inner water shell surrounding the core with the thickness 8Å. The position water atoms in this shell is also found by molecule dynamics simulation

- The outer water jacket surrounding the thin water shell, the thickness of which goes up to the radius of the droplet.

The scattering intensity of hydrated protein (Region I) can easily be calculated since all the positions of atoms in this region are known. That of the water jacket (Region II & III) can be evaluated directly from the experimental results, which gives the first peak (at around $\frac{2\pi}{2.8}\mathring{A}^{-1}$) [55]corresponding to the nearest neighborhood distance between oxygen atoms in addition to the central peak. The radial distribution function derived from the intensity spectrum also indicates that the water can be regarded as homogeneous from the position 8Åaway, which is the reason that the thickness of Region II is chosen.

Therefore, the interference term within this range of resolution between region I and the outer water is

$$I_{inter}(\mathbf{k}) = Re[\sum_{I}\sum_{II,III} f_I f_{II,III} e^{i\mathbf{k}(\mathbf{r_I}-\mathbf{r_{II,III}})}] \quad (3.12)$$

Since the water in region III is uniform when interacting with protein atoms, the summation in region III can be reduced to an integral and can be evaluated analytically. For the spherical droplet, this term is

$$\sum_{I}\sum_{III} f_I f_{III} e^{i\mathbf{k}(\mathbf{r_I}-\mathbf{r_{III}})} = \sum_{I} f_I e^{i\mathbf{kr_I}} \sum_{III} f_{III} e^{-i\mathbf{kr_{III}}} \quad (3.13)$$

and in which

$$F_{III} = \sum_{III} f_{III} e^{-i\mathbf{kr_{III}}} = \rho f_O \int_{III} e^{-i\mathbf{kr_{III}}} d\mathbf{r_{III}} = \rho f_O S \quad (3.14)$$

where S is the shape factor, and for a sphere droplet

$$
\begin{aligned}
S &= \int_r^R \int_\Omega \sum_{l=0}^\infty r^2 \cdot i^l (2l+1) j_l(kr) P_l(\cos\gamma) dr d\Omega \\
&= 4\pi \int_r^R r^2 \frac{\sin kr}{kr} dr = \frac{4\pi}{k^3}[(\sin kR - kR\cos kR) - (\sin kr - kr\cos kr)]
\end{aligned}
$$
$$(3.15)$$

in which r is the inner radius of region III and R is the outer one. ρ is the density of water (calculated to be $0.033\frac{atoms}{A^3}$ from its density). F_{III} oscillates and goes to zero with the envelop of $\frac{2\pi\rho(R-r)}{k^2}$ $\frac{2\pi}{100}A^{-1}$.

64

In the end, in order to calculate the intensity from pure water, the only calculation needed is to convert the reduced intensity $H(\mathbf{k})$ from literature[56] to the real intensity of a specific droplet. Note that for the bulk sample used in a conventional experiment, at the low resolution where the short range structure can be ignored, the water background can be considered to be of inverse symmetry. Apparently the scattering amplitude is real, and the following equation can be used:

$$I_{water} = |\sqrt{N(H(\mathbf{k}) + 1)f_o(\mathbf{k})} \otimes S(\mathbf{k})|^2 \qquad (3.16)$$

In the equation above N is the number of water molecules in the droplet, while $S(\mathbf{k})$ is the shape factor determined by the size of a droplet and can be evaluated from Eq.(3.15) since the experimental spectrum is measured from bulk water where $S(\mathbf{k}) = \delta(k)$.

In the sample calculation, 4PFK, which is a medium molecule containing ~2300 atom for the sake of molecule dynamics simulation speed, the intensity of DP in region I and the interference term between I-II and I-III are plotted in Fig.(3.3) along a line going through the center of the pattern and the range of resolution that we are interested in (from 20Åto 6Å).

It is found that the intensity of the cross term is significant compared with that of protein diffraction. Consequently, the DP obtained from experiment cannot be simply separated into water intensity and protein intensity at any resolution. The thinner water jacket will help reduce the interference, however, even when the thickness is 10Å, the effect of water is still not negligible.

Compared with the spherical droplet, the only difference of the water stream is the change of shape factor S. According to experimental setup, the water stream illuminated by the incident beam can be regarded as a cylin-
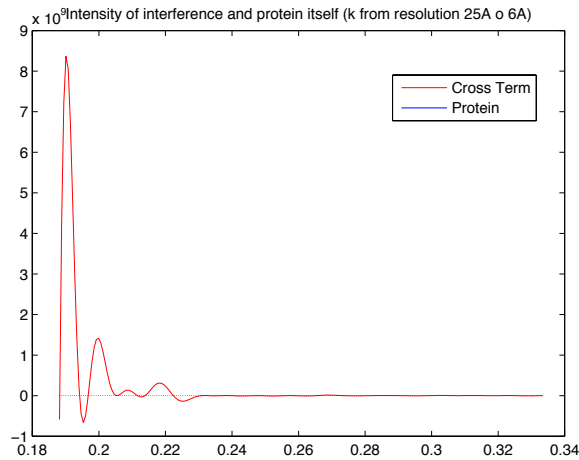
Figure 3.3: This figure shows the comparison of intensity between diffraction intensity of protein alone and interference term between protein and outer water jacket (droplet radius 10000Å). Note that the first term is negligible compared with the interference term. The x-axis is the radius in reciprocal space, corresponding to resolution from 25Åto 6Å

der, the length L of which is the segment in the beam spot. Therefore S can also be evaluated analytically

$$S = \int_{-L/2}^{L/2} e^{ik_z z} dz \int_D e^{i\mathbf{k}\cdot\mathbf{r}} d^2\mathbf{r} = \frac{\sin \pi \mathcal{L} k_z}{\pi k_z} \cdot \frac{R}{2} \frac{J_1(\pi R k_r)}{k_r} \qquad (3.17)$$

.

3.4   Making use of the cross term

Since the cross term is so strong that the protein term is suppressed, the conventional method of solving protein structure by phasing the intensity scattered by protein is no longer valid with the presence of the water jacket and the cross term it introduces. However, the cross term of the water jacket itself can serve as an important source of structural information.

The DP of protein inside the water jacket can be written as

$$I = \mathscr{F}[\rho_{water} + \delta\rho] \cdot \mathscr{F}[\rho_{water} + \delta\rho]^* \qquad (3.18)$$

and Fourier transformation of both side yields

$$
\begin{aligned}
A &= (\rho_{water}(\mathbf{r}) + \delta\rho(\mathbf{r})) \otimes (\rho_{water}(-\mathbf{r}) + \delta\rho(-\mathbf{r})) \\
&= A_{water} + \rho_{water} \otimes [\delta\rho(\mathbf{r}) + \delta\rho(-\mathbf{r})] + \delta\rho(\mathbf{r}) \otimes \delta\rho(-\mathbf{r})
\end{aligned}
\tag{3.19}
$$

,where we assume again that the water is inversely symmetrical so that $\rho_{water}(\mathbf{r}) = \rho_{water}(-\mathbf{r})$ can be used. Note that here $\delta\rho$ is not the charge density of protein, instead, $\delta\rho = \rho_{mol} - \rho_{water}$ is the fluctuation of the protein charge density with respect to the water background. Considering the protein molecule itself is composed of light atoms whose Z number is close to that of the water, this term should be much smaller than ρ_{mol}.

The left hand side of Eqn.(3.19) can be obtained by experiments, the first term on the right hand side is from the preliminary knowledge of the shape and size of water, while the last term can be neglected if $\delta\rho$ is much less than $\rho_{water}(\mathbf{r})$. Then $\delta\rho(\mathbf{r}) + \delta\rho(-\mathbf{r})$ can be calculated from experimental results by deconvolution. And if the molecule is not centered and small enough it can be separated from its space inverse copy and thus the water serves as the reference to solve the phase problem. Fortunately, both the above requirements of the $\delta\rho$ are consistent to the real experiment. The intermediate results of each step is shown in Fig.(3.4).

Numerical calculation shows that if only the charge density projection is less than 1/20 of the water background it can be successfully reconstructed and the smaller the ratio the better. This requirement should not be difficult to satisfy considering the fact that the size of the water jacket is of the order of micron, while that of the molecule is less than 100nm, and that $\delta\rho$ itself is a small term.
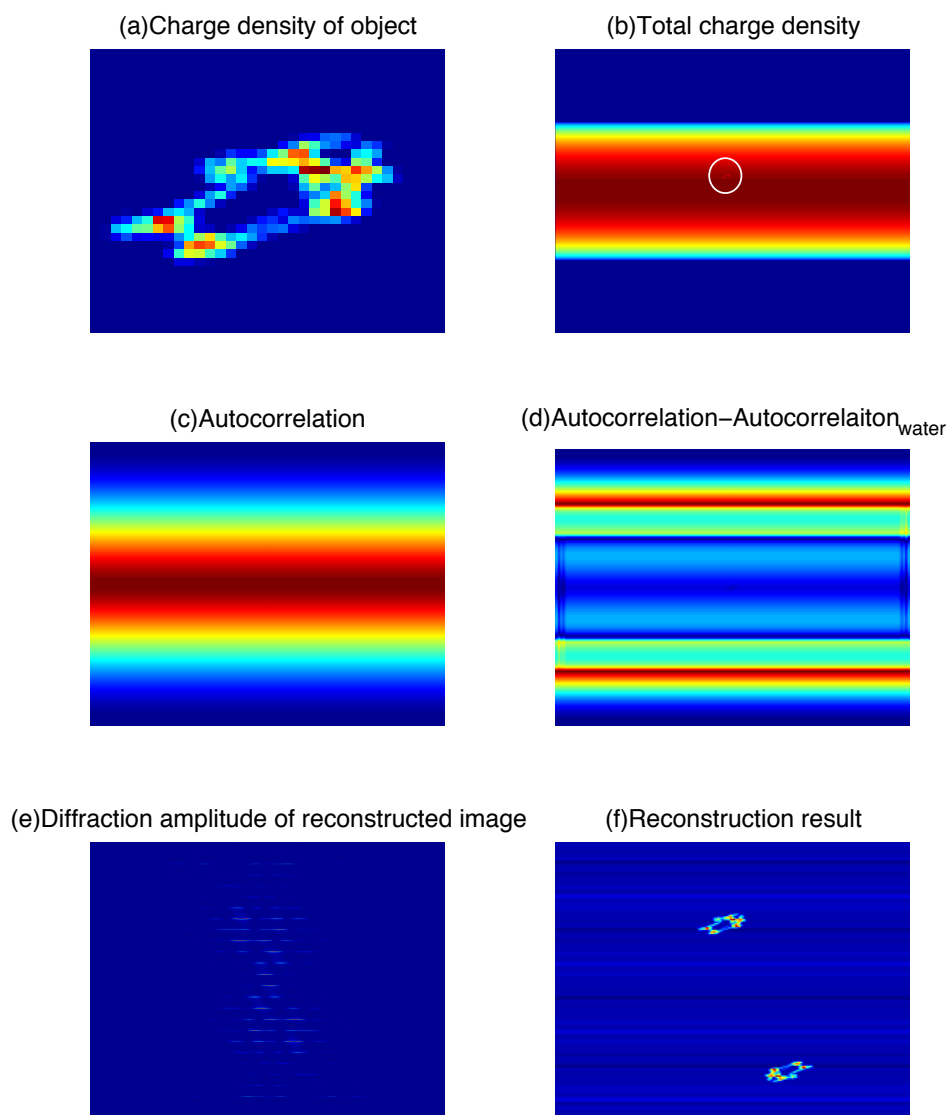
Figure 3.4: This figures shows the process of water background deconvolution: (a) is the charge density projection of the original object (b) is the charge density projection of the object in the water stream. The width of the streamm is 300nm and the object is circled as it can barely be seen. (c) The autocorrelation of the protein-stream mixture. (d) The difference between the autocorrelation of the protein-stream mixture and that of pure water stream $A - A_{water}$. (e) The intermediate step of deconvolution, which is actually the Fourier transformation of object with its inversion.(f) The enlarged area of reconstruction result, showing two inverse symmetrical object.

3.5 Effect of errors and noise

The validity of the algorithm under ideal circumstance is far from enough, it is largely determined by its performance on the experimental data which contain all possible errors and noise. In this section several major source of errors are simulated for the upcoming experiment of virus MSII in LCLS to test the robustness of the deconvolution method above.

For the CCD cameras used at present, most of them have a beam stop to block the strong transmission beam, thereby avoid its damage to the detecto. Thus the low resolution data is missing. The other important error source is the poisson noise of the detector pixels which count the number of x-ray photons scattered to them, which is related to the number of photons of each incident pulse. Finally, the success of the method heavily depends on the *a prior* knowledge of the shape of the water background, which is not always available.

The evaluation of these effects to the final result of reconstruction is conducted under the combination of all the above factors. If not specified, the parameters are chosen to match the LCLS experiment: The photon energy is 8keV, the resolution at the edge of DP is set to 6Å, the CCD camera is a 1024×1024 array with the beam stop 8 pixels in radius. The flux of photons is $10^{12}/pulse$ and Poisson noise is applied to the photon amount collected by each pixel of the detector. The water background is a cylinder whose diameter is 300nm. The sample used is the virus whose PDB entry is 1Z8Y.

Firstly, it is found that given the correct estimate of the water background, the quality of the reconstructed image is relatively insensitive to the size of the beam stop (see Fig.(3.5)). Thus satisfactory images are expected to be

69

obtained with the CCD detector used in LCLS. The missing of low resolution data is once an important issue in diffractive imaging. Some special algorithms used to interpolate[18] or reconstruct such information iteratively[28]. Sometimes low resolution optical image is used to complement this part. With the water background deconvolution method, all the above trouble can be avoided and the result follows directly from the original data. (I think some equations would make the argument above more solid, but still don't have any idea)
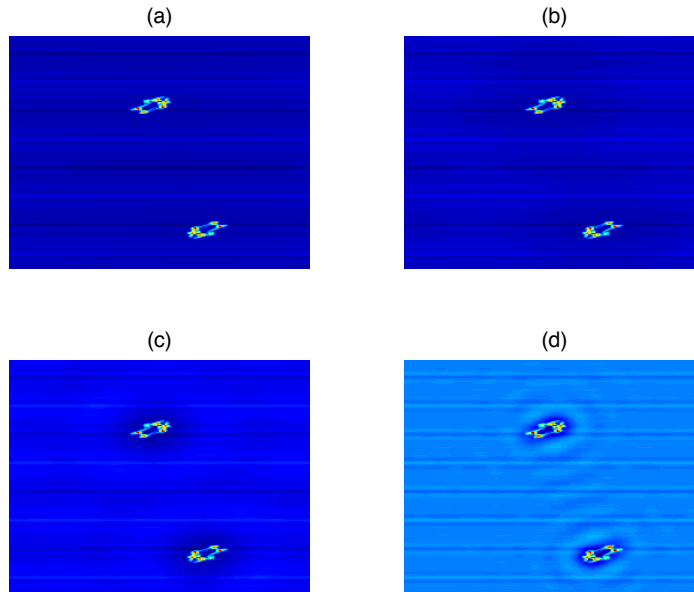
Figure 3.5: This figure shows the effect of beam stop: result with beam stop radius 5(a), 10(b), 20(c) and 50(d) pixels respectively for a. The quality of reconstruction is not sensitive to the beam stop size.

The success of the method is closely related to the estimate of the water background. In Fig.(3.6), the estimated value is changed a little bit from the correct size of the water background and the same process of deconvolution is conducted. Apparently the method collapses when the error is increased to some extent. The numerical experiment shows that the threshold value is

0.15%, which requires very accurate method of the size of the water stream. Although practically impossible currently, it is still possible to search for that value by going through the small portion of parameter space, which is not difficult to estimate, including the true value, with fine enough intervals. When the feature of the isolated object appears in the reconstructed image, the reconstruction can be regarded successful and the background parameter used is therefore the true value, and therefore the estimate problem is solved in expense of computational time.
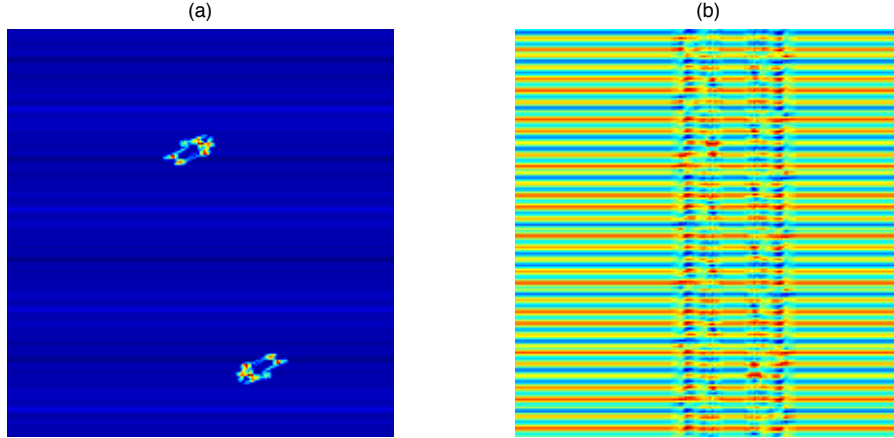
Figure 3.6: This figure shows the reconstructed image with 0.1% error of the estimated water background width (a) and 0.3% error (b). The quality of reconstruction is sensitive to the estimated water background, which should be less than 0.2% within the real size.

Additionally, there is also restriction on the size of the water stream. As stated above, If the stream size too small, the protein term cannot be regarded as perturbation and algorithm fails. On the other hand, if the stream size is too big, which might be more often in the experiment, the performance of the method is also harmed. This is because in that case the increase of water noise, with the increase of water signal, overwhelms the subtle difference between

protein-water and pure water. Fortunately, numerical calculation shows that for the current simulation configuration, the stream diameter can be safely increased to 1 micron, which can be practically produced with ASU nozzle.

The incident photon flux is also an important factor that determines the quality of the result. For Poisson noise, more photons lead to higher signal-noise ratio. Without controlling the noise under some certain level, as mentioned above, the signal from the perturbation cannot be distinguished. The simulation result that the image quality reconstructed from a $10^{11}photons/pulse$ DP badly degenerates indicates the necessity of the ultra high brightness of LCLS.

In the successful reconstruction, it is found that the noise and the perturbation signal are of the same order, which indicate the method is noise-robust to some extent. It is a natural result that the noise distributes all over the image, while the perturbation concentrates on a small portion of the image.

3.6   Conclusion

Immersing the protein single molecule in water and injecting them for diffraction has been proved to be one of the most feasible plan to do single molecule diffraction, and the experiment preparation is almost ready for the LCLS in the coming year. However, processing the DP of water-sheltered protein requires new method since the scattering intensity of protein alone cannot be directly separated from the whole pattern.

The following process is proposed to overcome the difficulty of water:

1. Do the Fourier transformation of the experimental(simulated) DP to get the autocorrelation of the whole system

2. Make a estimation of the water background, and calculate the autocorrelation of such background with the same experimental parameters

3. Deconvolute the difference of the above result with the assumed water background. If two isolated objects with inverse symmetry appear, the reconstruction is successful, or change the assumed background a little bit and repeat the process.

This method is tested numerically and found to work properly under the combination of all kinds of simulated errors and noise during real world experiment. Therefore, it is a suitable candidate algorithm to process the experimental data in the near future.

CHAPTER 4

PHASING NANO-CRYSTALS DIFFRACTION PATTERNS

4.1   Introduction

It has been appreciated for many years that if scattering could be obtained at points between Bragg reflections, the additional information so provided would greatly assist solution of the phase problem. For protein crystals, variations in water content can be used to vary the cell dimensions, offering the possibility of obtaining this information. Since they contain acentric alpha-helices, protein crystals are invariably non- centrosymmetric, however they may contain projections which are centric, such as those along 2, 4 or 6-fold axes, in which case the structure factors on the corresponding plane in reciprocal space will be real, with an ambiguity of sign only. Thus it was pointed out at an early stage[16] that on these planes, for adjacent structure factors to have opposite signs, the continuous Fourier Transform of the unit cell density (which modulates all scattering) must have a zero crossing between Bragg reflections. This could be identified if intensity were detectable at the half-order positions, thus solving the phase problem for centrosymmetric crystals and projections. Bragg "sampling" of intensities is consistent with the Shannon sampling theorem requirement for reconstruction of the autocorrelation of the molecular density; half-orders with reconstruction of the molecular density. The modern field of diffractive imaging, whose methods we will use, has developed, from these and other considerations, iterative algorithms which are capable of solving the phase problem for non-periodic objects under certain conditions[27]. For X-ray diffraction at higher energies well away from absorption edges, the occurance of a real charge density of known sign greatly facilitates these iterative methods. It is immediately clear then that one ap-

proach to phasing nanocrystal data would be to treat the entire nanocrystal as a single non-periodic object, and phase it as such. For a large protein, however, the computational demands of this approach be very great and, if the surfaces were not of interest, the resulting identical molecular densities would need to be summed to reduce noise. We describe a more efficient approach below.

In both electron and X-ray diffraction, nano-crystallography has developed rapidly in recent years. For a crystallite consisting of a few unit cells, the sharp Bragg reflections are convoluted with a shape-transform function, which therefore provides the required inter-Bragg scattering needed to assist in solving the phase problem. This paper develops that idea, in the light of recent results from the Linac Coherent Light Source, in which femtosecond X-ray snapshots where obtained from protein nanocrystals of Photosystem 1 containing a dozen or more unit cells on a side[57]. These show strong inter-Bragg interference fringes. Elementary theory [58] shows that, as for a finite grating, the number of unit cells N between crystal facets along direction g is just equal to (n+2) where n is the number of subsidiary maxima between Braggs along direction g. Thus the dimensions of the nanocrystal can be found by counting fringes along various directions in reciprocal space. In this paper we present simulations for X-ray diffraction from a protein nanocrystal, and demonstrate solution of the phase problem for the general non-centrosymmetric (acentric) case by using the inter-Bragg scattering. In simple terms we show that a finite grating solves the phase problem.

4.2   X-ray scattering from a nanocrystal

In an earlier paper, we give the exact expression for diffraction from a finite crystal of parallelepiped form, illuminated by plane-polarized monochromatic incident radiation, with wavevector $\mathbf{k}_i$ ($|\mathbf{k}_i| = 1/\lambda$) and negligible beam divergence. Then the diffracted photon flux I (counts/time) at $\Delta\mathbf{k} = \mathbf{k}_i - \mathbf{k}_o$ produced by the n-th parallelepiped crystallite, consisting of N(n) = N1 x N2 x N3 unit cells, is given in the kinematic theory as

$$I_n(\Delta\mathbf{k}, \alpha, \beta, \gamma, N_i) = I_0|F(\Delta\mathbf{k})|^2 r_e^2 P(\Delta\mathbf{k}) \frac{\sin^2)(N_1\Psi_1}{\sin^2\Psi_1} \frac{\sin^2(N_2\Psi_2)}{\sin^2\Psi_2} \frac{\sin^2(N_3\Psi_3)}{\sin^2\Psi_3}\Delta\Omega$$

(4.1)

where $F(\Delta\mathbf{k})$ is continuous scattering from one unit cell, which we take here to contain one molecule. I_0 is the incident photon flux density (counts/time/area), r_e^2 the electron cross section, and $\Delta\Omega$ is the solid angle subtended by a detector pixel at the sample. Here

$$\Psi_1 = 2\pi a \sin\theta \cos\alpha/\lambda$$
$$\Psi_2 = 2\pi b \sin\theta \cos\beta/\lambda \qquad (4.2)$$
$$\Psi_3 = 2\pi c \sin\theta \cos\gamma/\lambda$$

where θ is half the scattering angle, and α, β, and γ define the crystal orientation as the angles which the scattering vector $\Delta\mathbf{k}$ makes with the directions of the real-space unit cell vectors $\mathbf{a}$, $\mathbf{a}$ and $\mathbf{c}$. $\Delta\mathbf{k}$ is defined by the position of the detector pixel and X-ray wavelength, and defines a point in reciprocal space where the Ewald sphere intersects the shape transform.

To simplify simulations, we take the incident X-ray wavevector $\mathbf{k}_i$ to lie parallel to the c axis of a cubic crystallite, and use a high X-ray energy ap-

proximation, in which $k_z = 0$, so that the Ewald sphere is approximately planar, resulting in two-dimensional diffraction. Then the third term in 4.1 is unity, and, if polarization effects are neglected, we may write 4.1 as

$$I_n(\Delta \mathbf{k}, N^{(n)}) = c|F(\Delta \mathbf{k})|^2 S_n(\Delta \mathbf{k}, N^{(n)}) \Delta \Omega \tag{4.3}$$

where $S(\Delta \mathbf{k}, N^{(n)})$ is the interference term in Eqn. 4.1 and c is a constant.

4.3   Phasing from shape-transform

We consider the case of femtosecond diffraction from a stream of nanocrystals, in which one diffraction pattern is read out each time one X-ray pulse hits one nanocrystal, and the process is repeated at the repetition rate of the X-ray laser (perhaps at 60 Hz). Our essential strategy is to divide the recorded intensity $I_n(\Delta \mathbf{k}, N_i)$ by the known interference function $S_n(\Delta \mathbf{k}, N_i)$ (which depends only on the number of unit cells), to obtain the modulus squared $|F(\Delta \mathbf{k})|^2$ of the Fourier Transform of the molecular density $\rho(r)$. The phase problem for the continuous function $|F(\Delta \mathbf{k})|^2$ may then be solved by iterative methods[28], to yield a density map for one unit cell.

The femtosecond diffraction method produces millions of diffraction patterns from randomly oriented crystals covering a range of sizes, so that the results may be sorted into various size classes. We consider a set of 2D diffraction patterns from nanocrystals differing only in size $N = N_1 \times N_2$. Then the problem of division by zero when inverting Eqn. 4.1 may be addressed by summing the required ratio I_n/S_n over n, so that the denominator passes through zero at different values of $\Delta \mathbf{k}$ in each term. We form

$$c|F(\Delta \mathbf{k})|^2 = \sum_n \frac{I_n(\Delta \mathbf{k}, N^{(n)})}{S_n(\Delta \mathbf{k}, N^{(n)}) + \epsilon} \tag{4.4}$$

77

Alternatively we may obtain the same result by summing Eqn. 4.2 over crystal sizes, giving

$$\sum_n I_n(\Delta\mathbf{k}, N^{(n)}) = \sum_n c|F(\Delta\mathbf{k})|^2 S_n(\Delta\mathbf{k}, N^{(n)})\Delta\Omega \qquad (4.5)$$

$$= c|F(\Delta\mathbf{k})|^2 \sum_n S_n(\Delta\mathbf{k}, N^{(n)})\Delta\Omega$$

so that

$$c|F(\Delta\mathbf{k})|^2 = \frac{\sum_n I_n(\Delta\mathbf{k}, N^{(n)})}{\sum_n S_n(\Delta\mathbf{k}, N^{(n)}) + \epsilon} \qquad (4.6)$$

If information is available from , for example, dynamic light scattering, on the distribution of nanocrystal sizes, we may consider the approximation

$$c|F(\Delta\mathbf{k})|^2 = \frac{\sum_n I_n(\Delta\mathbf{k}, N^{(n)})}{\sum_n F_n(N^{(n)})S_n(\Delta\mathbf{k}, N^{(n)}) + \epsilon} \qquad (4.7)$$

where $F_n(N^{(n)})$ is the particle size distribution. This avoids the need to determine individual values of N_1, N_2, and N_3.

4.4   Simulations and phasing
Two dimensions

We commence with simple non-statistical two-dimensional examples to clarify the method. Fig. 4.1 shows the density of a simple protein, Alpha-Conotoxin PNIB from Conos Pennaceus (1AKG in the PDB, orthorhombic, $a = 14.60\mathring{A}, b = 26.10\mathring{A}, c = 29.20\mathring{A}, P2_12_12_1$) , projected along the c axis. We have taken one molecules per unit cell, giving symmetry P1 for generality. Fig. 4.2 shows $|F(\Delta\mathbf{k})|^2$, the Fourier modulus squared for the molecule. (This function is equal to the structure factors at the lattice sites). Fig. 4.3 shows the interference function $S_n(\Delta\mathbf{k})$ for a nanocrystal of 6 unit cells on a side (N(1)=216 molecules). The calculation samples the shape transform at 7
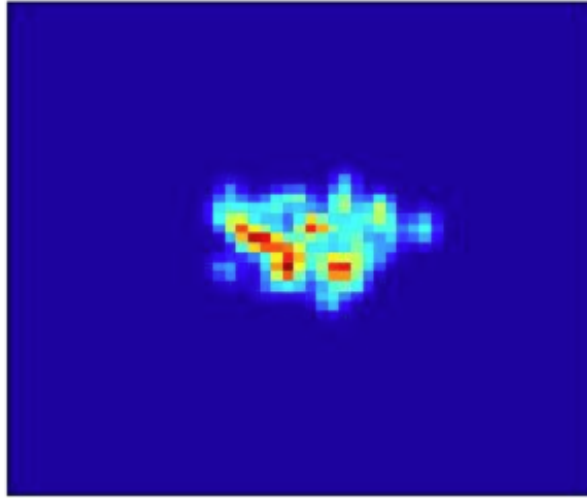
Figure 4.1: This figure shows the density projection of a simple protein, Alpha-Conotoxin PNIB from Conos Pennaceus 1AKG in the PDB, orthorhombic, a = 14.60Å, b= 26.10Å, c = 29.20Å, $P2_12_12_1$, projected along the c axis. The resolution is 1Å.
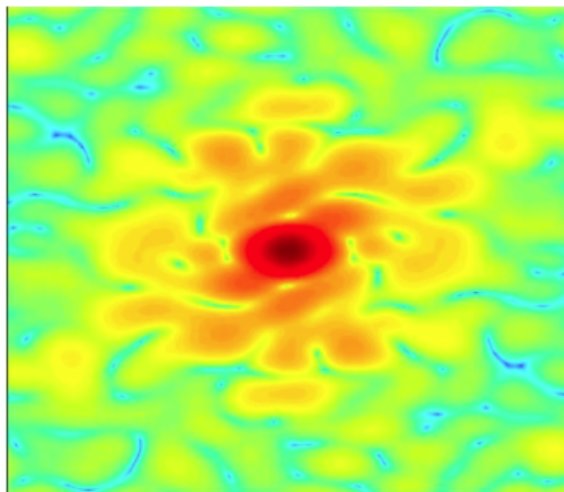
Figure 4.2: This figure shows $|F(\Delta\mathbf{k})|^2$, the Fourier modulus squared for the molecule in logarithm scale. (This function is equal to the structure factors at the lattice sites).

points within each sinc function subsidiary maximum. Fig. 4.4 shows the simulated diffracted intensity $|I_n(\Delta\mathbf{k}, N)$, extending to 2 Angstroms resolution, with the effects of Poisson noise added. Fig. 4.5 shows the recovered molecular transform $|F(\Delta\mathbf{k})|^2$ based on Eqn. 4.7, using just two nanocrystals in the sum for N=6 and N=8.
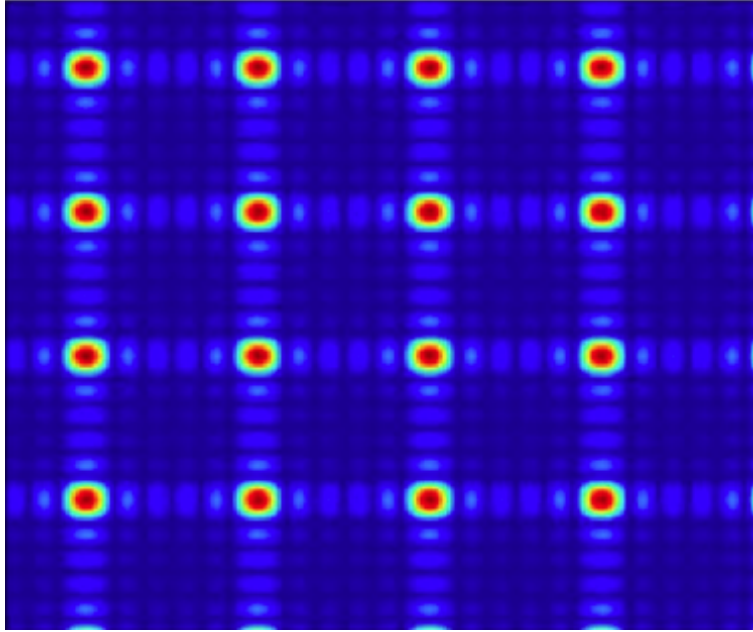
Figure 4.3: This figure shows the interference function $S_n(\Delta\mathbf{k})$ for a nanocrystal of 6 unit cells on a side ($N^{(1)}$=216 molecules). The calculation samples the shape transform at 7 points within each sinc function subsidiary maximum.

Three dimensions[2]

The restricting assumption of a flat Ewald sphere (and consequent resolution limit) may be eliminated by working directly with Eqn. 4.1 and Eqn. 4.7 in a statistical approximation, or with Eqn. 4.6, in which a particular interference function is modelled for each particle. The second approach requires a value of N_3, which might be extracted from high-order reflections, where the sphere cuts through the shape transform at a high angle. The results will

80
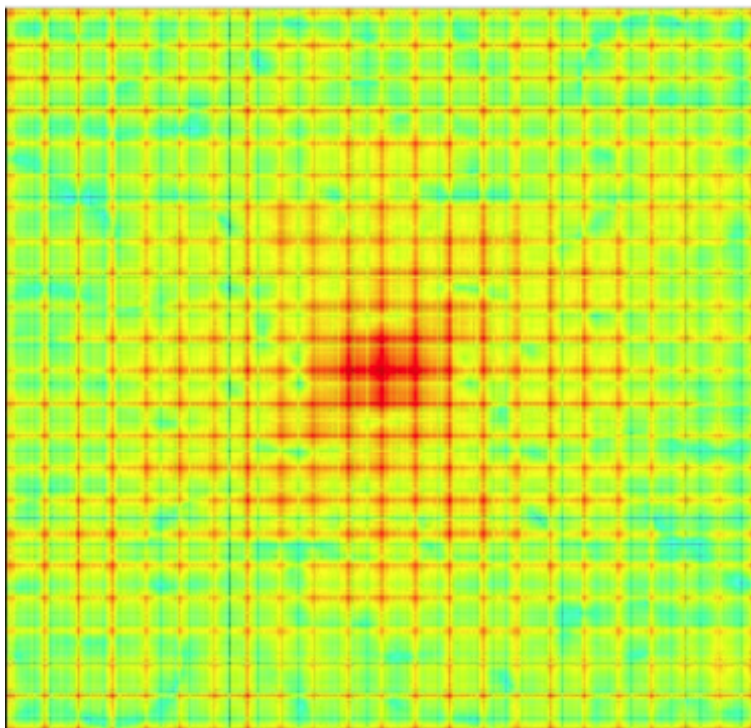
Figure 4.4: This figure shows the summation of diffraction patterns from different size of nanocrystals, as in the denumerator of Eqn. 4.7
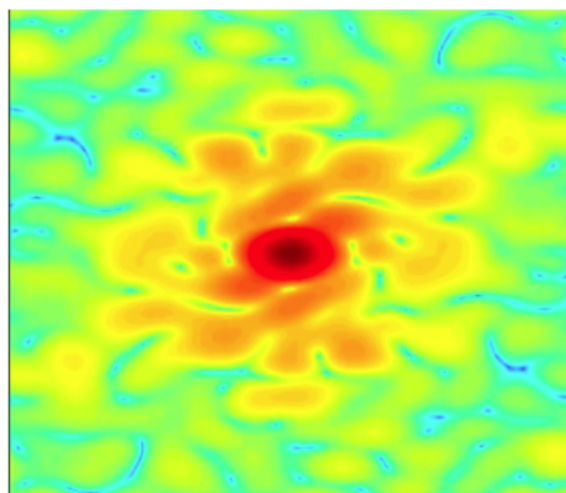
Figure 4.5: This figure shows the reconstructed single molecule diffracted intensity, extending to 2A, resolution, in logarithmic display. It is almost identical to the original one

duplicate those obtained in two dimensions. A more practical approach can be based on Eqn. 4.7, which uses independent measurements of the particle size distribution $F(N^{(n)})$. Then the sum in Eqn. 4.7 must be interpreted as follows. The diffraction patterns are first indexed[58] in order to determine their relative orientation. We then add shape transforms from different crystals together with the same Miller index, at each of the pixels which finely sample the shape transform. The result produces a three-dimensional diffraction volume including diffuse scattering between Bragg reflections. This is divided by the denominator in Eqn. 4.7. Fig. 4.6 shows 3D simulations. The accuracy of these is measured using a real-space R-factor, which measures the difference between the original model charge density, and that recovered by the algorithm.
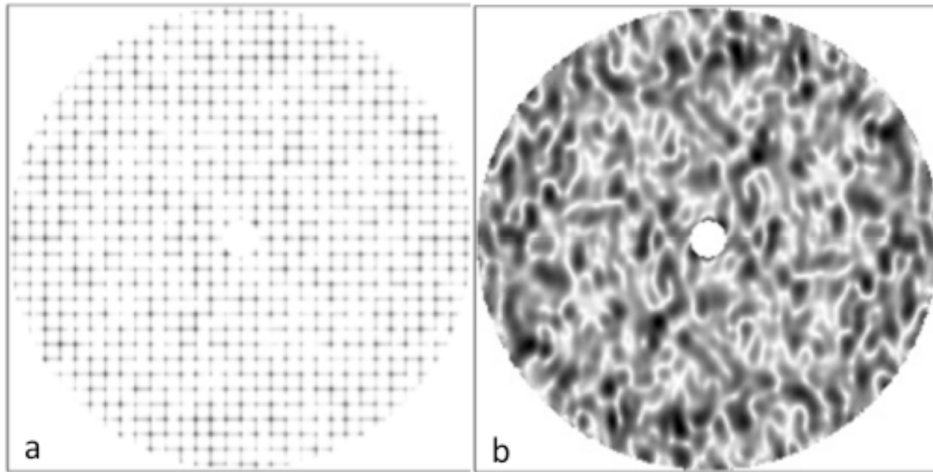
Figure 4.6: (a) Sum of many patterns (Eqn. 4.7) ; crystal size N=20, $\delta = 2$, on zone axis [001]. (b) The recovered molecular transform.

CHAPTER 5

EXTRACTION OF SPHERICAL HARMONICS EXPANSION

COEFFICIENTS FROM TRIPLE CORRELATION
5.1   Introduction

This chapter deals with the following problem. Given the hard X-ray snapshot
scattering from many randomly oriented, identical molecules in solution, re-
construct a charge-density map for one of them. This problem differs from the
problem addressed by the conventional small-angle X-ray scattering (SAXS)
method in that , for brief X-ray pulses, the molecules do not rotate during
the exposure, and are thus frozen in time. (The method we will discuss would
work equally well for molecules frozen in space, such at those in vitreous ice
irradiated with continuous X-rays). We assume that there is no interference
between X-rays scattered by different molecules. The resulting diffraction pat-
tern will show two-dimensional fluctuations in intensity, unlike the isotropic ,
one-dimensional patterns produced by conventional SAXS data. This indicates
that they contain more information, and it can thus be expected that the inver-
sion to a real-space image from these two-dimensional snapshot patterns will
be easier than inversion from one-dimensional SAXS patterns. Conventional
SAXS analysis relies on modelling and use of a-priori information - by contrast
we are concerned here with ab-initio methods. Iterative phasing may therefore
be used in the steps in which a density map is obtained from the angular au-
tocorrelation functions. In 1978 [59] , Z. Kam derived a remarkable result. He
found that the sum of the angular autocorrelation functions from many such
snapshot SAXS patterns (each containing contribuitons from many identical
molecules in solution) would converge to that of a single molecule, added to a
conventional SAXS background. In recent work [60] we have demonstrated the

practical application of this result for experimental soft X-ray scattering from a set of gold nanorods lying on their side on a transparent membrane. In this chapter, we discuss the extension of these methods to the more important case of proteins in solution, with the aim of obtaining a three-dimensional density map

The idea of using single molecule diffraction has been proposed as a method to solve the structure of macromolecule that are difficult to crystallize with has been proposed by[15, 40]. Essential parts of the whole experiment setup, such as the nozzle to feed the sample[50, 49] and the free electron laser X-ray light source (LCLS), of the whole experiment setup have already been launched. Experiments on nanocrystals at the LCLS has been considered successful and promising[57], and the experiments on single particles at atomic resolution will soon be carried out at this state-of-the-art facility. Although the new generation light source provides quite an effective weapon to attack structure problems which are formidable with conventional crystallography methods, the need for new data analysis algorithms arises at the same time.

Two of the most imminent challenges in data analysis are Poisson statistical noise and missing information about the orientation for the diffraction pattern from each X-ray pulse. In the analysis of nanocrystal data[61], the intensity is quite concentrated so that the noise is not significant and the crystallinity helps align the molecules and makes automatic indexing, and hence orientation determination possible.

However, even with the enormous flux of photons, the ability of single molecules to scatter photon is still so weak that the number of photons detected at high angles are quite limited. Consequently, the low signal-noise ratio makes it hard to classify the diffraction patterns of the same orientation pre-

cisely. And without the noise-free diffraction pattern, the conventional sorting method such as common line cannot work correctly, which make the whole 3D reconstruction fail.

In order to overcome the difficulties discussed above and make full use of the information contained in the millions of diffraction patterns collected, several sophisticated statistical algorithms[36, 48] have been proposed and tested on simulated data. The fact that all these algorithms make use of statistics to approach the most probable scattering intensity, which is regarded as the true intensity, indicates that the non-iterative analytical algorithm described in this chapter should also be based on the statistics of the noisy data to ensure the precision when exploiting the data.

Although only SAXS is left after simply averaging the diffraction over the spatial correlation, which is the product of intensity at different positions with the same relative coordinates averaged over many randomly oriented patterns, includes much more information,which is needed for reconstruction. Kam[44] used double correlation to process the electron micrograph of a macromolecule but cannot apply it to diffraction patterns. Kakarala[62] developed a systematic algorithm to recover a spherical function from its triple correlation. However, due to the fundamentally different nature of the practice, in X-ray diffraction the measurement is conducted in a spherical cap passing through the origin, instead of a set of homocentric spherical surfaces.

The most recent progress is that of Saldin et al[38]., who has successfully used cross correlation and triple correlation to reconstruct the diffracted intensity from a single particle from the diffraction patterns of many random particles that were randomly spatially distributed and in random orientation, and phase it. However, this gives a 2D projection instead of a 3D structure,

which people are most interested in. In this chapter, an entirely new algorithm is introduced to deal with the simplest single particle 3D problem.

5.2 General theory

Firstly, it would be convenient to expand the scattering intensity in reciprocal space in term of spherical harmonics

$$I^{\omega=0}(\mathbf{q}) = \sum_{\ell m} I_{\ell m}(q) Y_{\ell m}(\hat{\mathbf{q}}) \tag{5.1}$$

so that it is convenient to rotate it to any orientation explicitly with Wigner D-function

$$I^{\omega}(\mathbf{q}) = \sum_{\ell m m'} I_{\ell m}(q) D^{\ell}_{mm'}(\omega) Y_{\ell m'}(\hat{\mathbf{q}}) \tag{5.2}$$

It is natural to take the incident beam as z in spherical coordinates, then what we measure is a slice with ϕ from 0 to 2π, and θ depends on the radius q

$$\theta(q) = \frac{\pi}{2} - \sin^{-1}(\frac{q\lambda}{2}) \tag{5.3}$$

Now the diffraction pattern recorded at pixel i is

$$I^{\omega}(q_i, \theta_i, \phi_i) = \sum_{\ell m m'} I_{\ell m}(q_i) D^{\ell}_{mm'}(\omega) Y_{\ell m'}(\theta(q_i), \phi_i) \tag{5.4}$$

If the radial coefficients I_{lm} for each q_i can be obtained, it is expected that the intensity distribution is available and good enough for further phasing. Note that in theory the ℓ in the summation can be taken to inf, however, in the detector array used in LCLS experiments, the upper limit of ℓ for non-vanishing I_{lm} is determined by the granularity of pixels and can be estimated to be

$$\ell_{max}(q_i) = 2\pi i \tag{5.5}$$

The other property of the problem is that according to Frediel's law, the intensity is inversely symmetrical. This fact indicates that $I_{lm} = 0$ when ℓ is odd and provides the redundancy to deal with real world data.

86

Cross correlation and triple correlation

The spatial correlation can also be evaluated with the spherical harmonics expansion. The simplest case is the cross correlation

$$\langle I^\omega(q, \theta, \phi) I^\omega(q', \theta', \phi') \rangle_\omega = \tag{5.6}$$

$$\langle \sum_{\ell m m''} I_{\ell m}(q) D^\ell_{m m''}(\omega) Y_{\ell m''}(\theta(q), \phi) \sum_{\ell' m' m'''} I_{\ell' m'}(q') D^{\ell'}_{m' m'''}(\omega) Y_{\ell' m'}(\theta(q'), \phi') \rangle$$

where the $< ... >$ stands for $\int_{SO(3)} \cdots d\omega$, indicating the average over the ω of *uniformly random* orientation, which is the most important prerequisite for this method. As only the D-functions are included in the averaging, the orthogonality

$$\langle D^{\ell_1}_{m'_1 m_1}(\omega) D^{\ell_2}_{m'_2 m_2}(\omega) \rangle = C^{00}_{\ell_1 m_1 \ell_2 m_2} C^{00}_{\ell_1 m'_1 \ell_2 m_2'} \tag{5.7}$$

helps remove most of the cross terms in the multiplication and in the end we have

$$\langle I^\omega(q, \theta, \phi) I^\omega(q', \theta' \phi') \rangle_\omega$$

$$= \sum_\ell \sum_{m m'} C^{00}_{\ell m \ell' m'} I_{\ell m}(q) I_{\ell' m'}(q') \sum_{m m'} C^{00}_{\ell m \ell' m'} Y_{\ell m}(\theta(q), \phi) Y_{\ell' m'}(\theta(q'), \phi')$$

$$= \sum_\ell \{I_\ell(q) \otimes I_\ell(q')\}_{00} \{Y_\ell(i) \otimes Y_\ell(i')\}_{00} \tag{5.8}$$

where the direct product is a concise way to write the coupling of two spinors by Clebsch-Gordan coefficients, and $Y_\ell(i)$ is used to simplify the writing of $Y_\ell(\theta(q_i), \phi_i)$ from here on.

The quantities $\{Y_\ell(i_1) \otimes Y_\ell(i_2)\}_{\ell m}$ are the bipolar spherical harmonics. The $\ell = m = 0$ term are the bipolar scalar harmonics which are proportional to the Legendre polynomials

$$\{Y_\ell(i_1) \otimes Y_\ell(i_2)\}_{00} = \sum_{\ell m} C^{00}_{\ell m \ell - m} Y_{\ell m}(i_1) Y_{\ell - m}(i_2) = (-1)^\ell \frac{\sqrt{2\ell + 1}}{4\pi} P_\ell(i_1 \cdot i_2) \tag{5.9}$$

where $i_1 \cdot i_2$ indicates the angle between the directions given by $\theta(q_{i_1}), \phi_{i_1}$ and $\theta(q_{i_2}), \phi_{i_2}$.

For triple correlation, the orthogonality becomes

$$\langle D^{\ell_1}_{m'_1 m_1}(\omega) D^{\ell_2}_{m'_2 m_2}(\omega) D^{\ell_3}_{m'_3 m_3}(\omega) \rangle = C^{00}_{\ell_3 m_3 \ell_3 - m_3} C^{00}_{\ell_3 m'_3 \ell_3 - m'_3} C^{\ell_3 - m_3}_{\ell_1 m_1 \ell_2 m_2} C^{\ell_3 - m'_3}_{\ell_1 m'_1 \ell_2 m'_2} \quad (5.10)$$

so that we also have the triple correlation in terms of tripolar harmonics expansion

$$\langle I^\omega(\mathbf{q}_1) I^\omega(\mathbf{q}_2) I^\omega(\mathbf{q}_3) \rangle \tag{5.11}$$

$$= \sum_{\ell_1 \ell_2 \ell_3} \{ \{ I_{\ell_1}(q_1) \otimes I_{\ell_2}(q_2) \}_{\ell_3} \otimes I_{\ell_3}(q_3) \}_{00} \{ \{ Y_{\ell_1}(i_1) \otimes Y_{\ell_2}(i_2) \}_{\ell_3} \otimes Y_{\ell_3}(i_3) \}_{00}$$

$$\tag{5.12}$$

The expansions of both cross correlation and triple correlation make sense as any function depending on two or three vector direction can be expanded in terms of bipolar harmonics or tripolar harmonics, and after angular averaging only the terms with $\ell = 0$ survive. In the following section it will be shown that these two kinds of correlations are sufficient for reconstruction.

Correlation under noise

Before moving on to the reconstruction algorithm, it would be necessary to evaluate whether triple correlation can be regarded as reliable data under noise, which is extremely important as all the following steps, from intensity reconstruction to phasing, depend on the precise measurement of this value.

If the real intensity I' is written as the sum of the true value I and the random noise δI (which can be much higher than I), the triple correlation

becomes

$$\langle [I^\omega(\mathbf{q}_1) + \delta I^\omega(\mathbf{q}_1)][I^\omega(\mathbf{q}_2) + \delta I^\omega(\mathbf{q}_2)][I^\omega(\mathbf{q}_3) + \delta I^\omega(\mathbf{q}_3)]\rangle$$

$$= \quad T(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) + \langle I^\omega(\mathbf{q}_1) I^\omega(\mathbf{q}_2) \delta I^\omega(\mathbf{q}_3)\rangle + \langle I^\omega(\mathbf{q}_1) \delta I^\omega(\mathbf{q}_2) \delta I^\omega(\mathbf{q}_3)\rangle$$

$$+ \quad \langle I^\omega(\mathbf{q}_1) \delta I^\omega(\mathbf{q}_2) I^\omega(\mathbf{q}_3)\rangle + \langle \delta I^\omega(\mathbf{q}_1) I^\omega(\mathbf{q}_2) I^\omega(\mathbf{q}_3)\rangle + \langle \delta I^\omega(\mathbf{q}_1) I^\omega(\mathbf{q}_2) \delta I^\omega(\mathbf{q}_3)\rangle$$

$$+ \quad \langle \delta I^\omega(\mathbf{q}_1) \delta I^\omega(\mathbf{q}_2) I^\omega(\mathbf{q}_3)\rangle + \langle \delta I^\omega(\mathbf{q}_1) \delta I^\omega(\mathbf{q}_2) \delta I^\omega(\mathbf{q}_3)\rangle \qquad (5.13)$$

If the expectation of the noise $E[\delta I] = 0$, the expected triple correlation can be firstly simplified by dropping terms that contain first order noise δI. Furthermore, note that $\delta I^\omega(\mathbf{q}_i)$ and $\delta I^\omega(\mathbf{q}_j)$ are *independent* if $i \neq j$, which leads to

$$E[\delta I^\omega(\mathbf{q}_i) \delta I^\omega(\mathbf{q}_j)] = E[\delta I^\omega(\mathbf{q}_i)] E[\delta I^\omega(\mathbf{q}_j)] = 0 \qquad (5.14)$$

Similarly,

$$E[\delta I^\omega(\mathbf{q}_i) \delta I^\omega(\mathbf{q}_j) \delta I^\omega(\mathbf{q}_k)] = 0 \qquad (5.15)$$

Substitute Equ.(5.14) and Equ.(5.15) into the expectation of Equ.(5.13), to obtain the relationship between the triple correlation of noisy data and that of ideal data:

$$E[\langle I'^\omega(\mathbf{q}_1) I'^\omega(\mathbf{q}_2) I'^\omega(\mathbf{q}_3)\rangle] = E[\langle I^\omega(\mathbf{q}_1) I^\omega(\mathbf{q}_2) I^\omega(\mathbf{q}_3)\rangle] \qquad (5.16)$$

This conclusion should be applicable to all kinds of error distribution whose expectation is zero, like Gaussian or Poisson distribution, the most possible random error during the measurement of number of photons. Consequently, the triple correlation can be used as ideal data in the following process.

Moreover, now in this sense the meaning of the average of "many" diffraction pattern is changed. Not only do enough patterns from different orientations need to be collected to cover every possible orientation, but that kind of covering also needs to be done many times to approach the expectation.

89

Uniqueness of spatial correlation

In would be useful to check the uniqueness before attempting to do the reconstruction, or the effort might be pointless. Obviously the cross correlation of intensity is unique, as it is the 3D Fourier transformation of the "intensity" of intensity, which lacks the phase information to correctly determine the original function.

The ambiguity of cross correlation will emerge in the next section when trying to get $I_{\ell m}$ from it, although it is still useful. This fact necessitates the use of triple correlation. The uniqueness of triple correlation up to a 3D rotation has been proved in [62]

5.3   Reconstruction

The triple correlation is a function of a 3-tuple of vectors. In the general case, for a set of $N \times N$ patterns without inverse symmetry, the independent triple correlation is $\frac{N^6}{6}$ due to the invariance on the interchange of the order of $\mathbf{q}$s. Thus it seems that triple correlation provides a sufficient number of equations to obtain the value of $I(\mathbf{q})$ on each pixel ($\frac{N^3}{2}$ of them). However, solving or doing optimization on so many nonlinear equations is impractical, the problem must be reduced and linearized to make it analytically possible.

In Equ.(5.18) the $\ell = 0$ component of 3-folded direct product can be simplified
as

$$\{\{I_{\ell_1}(q_1) \otimes I_{\ell_2}(q_2)\}_{\ell_3} \otimes I_{\ell_3}(q_3)\}_{00}$$

$$= (-1)^{\ell_1+\ell_2+\ell_3} \sum_{m_1m_2m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} I_{\ell_1 m_1}(q_1) I_{\ell_2 m_2}(q_2) I_{\ell_3 m_3}(q_3)$$

$$(5.17)$$

and the tripolar harmonics component bears the same form. Therefore, Equ.(5.18)
can be regarded as a linear equation with $(\ell_1^{max} + 1)(\ell_2^{max} + 1)(\ell_3^{max} + 1)$ vari-
ables, where ℓ_i^{max} depends on q_i. Larger q_i requires larger ℓ to achieve the
desired resolution. For any set of q_1, q_2, q_3, it is possible to make all kinds
of combination of $\omega_1, \omega_2, \omega_3$ to produce enough linear equations to solve each
$\{\{I_{\ell_1}(q_1) \otimes I_{\ell_2}(q_2)\}_{\ell_3} \otimes I_{\ell_3}(q_3)\}_{00}$. However, there are still too many equations
considering that ℓ_{max} may reach more than 1000.

The orthogonality of tripolar harmonics[63] can be made use of to readily
obtain the corresponding coefficients if all the triple correlations among the 3
spherical surface q_1, q_2 and q_3 are known. Unfortunately, this is impossible as
we can only measure a spherical cap in the experiment. On the other hand,
the fact that on the spherical cap we have the data of all the ϕ from 0 to 2π
for certain q and $\theta(q)$ indicates that the orthogonality of an the exponential
function can be exploited.

Fourier transformation of triple correlation on $\phi_i s$

Writing out the explicit form of the tripolar harmonics, and setting $\mathcal{I}_{\ell_1 \ell_2 \ell_3}(q_1, q_2, q_3) =$
$\{\{I_{\ell_1}(q_{i_1}) \otimes I_{\ell_2}(q_{i_2})\}_{\ell_3} \otimes I_{\ell_3}(q_{i_3})\}_{00}$, which is of interest here, we have the triple

correlation

$$T(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \sum_{\ell_1 \ell_2 \ell_3} \mathcal{I}_{\ell_1 \ell_2 \ell_3}(q_1, q_2, q_3)(-1)^{l_1+l_2+l_3} \sqrt{\frac{\ell_1+1}{4\pi} \frac{\ell_2+1}{4\pi} \frac{\ell_3+1}{4\pi}} \cdot$$

$$\sum_{m_1 m_2 m_3} \sqrt{\frac{(\ell_1-m_1)!}{(\ell_1+m_1)!} \frac{(\ell_2-m_2)!}{(\ell_2+m_2)!} \frac{(\ell_3-m_3)!}{(\ell_3+m_3)!}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \cdot$$

$$P_{\ell_1}^{m_1}(\cos\theta_{q_1}) P_{\ell_2}^{m_2}(\cos\theta_{q_2}) P_{\ell_3}^{m_3}(\cos\theta_{q_3}) e^{i(m_1\phi_1+m_2\phi_2+m_3\phi_3)} \quad (5.18)$$

The exponential function indicates the triple correlation is also a Fourier expansion, and it is fairly easy and fast to do the discrete Fourier transformation on ϕ_i. The result is equivalent to set all the ϕs in spherical harmonics to zero:

$$\widetilde{T}(m_1, m_2, m_3) = \sum_{\ell_1 \geq |m_1|, \ell_2 \geq |m_2|, \ell_3 \geq |m_3|} \mathcal{I}_{\ell_1 \ell_2 \ell_3}(q_1, q_2, q_3) \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \cdot$$

$$Y_{\ell_1 m_1}(\cos\theta_{q_1}, 0) Y_{\ell_2 m_2}(\cos\theta_{q_2}, 0) Y_{\ell_3 m_3}(\cos\theta_{q_3}, 0) \quad (5.19)$$

where the sign $(-1)^{\ell_1+\ell_2+\ell_3}$ disappears as all the ℓs are even. In the linear equation above, the $\widetilde{T}$ is from experiment, the coefficients of $\mathcal{I}_{\ell_1 \ell_2 \ell_3}$ are only dependent on ℓ, m and experimental conditions such as wave length, therefore, if there are enough number of $\widetilde{T}(m_1, m_2, m_3)$s, the tripolar coefficients can be determined from experiment.

However, it should be noted that only a few of the $\widetilde{T}(m_1, m_2, m_3)$s are independent considering that the Wigner $3j$ coefficients are non-vanishing only when $m_1+m_2+m_3 = 0$. This corresponds to the fact that the triple correlation itself only depends on the two angles between the vectors $\mathbf{q}_1, \mathbf{q}_2$, and $\mathbf{q}_3$ as it is invariant to rotation. Moreover, the symmetry of Wigner coefficients and

spherical harmonics

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{\ell_1 + \ell_2 + \ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix}$$

$$Y_{\ell m}(\cos\theta, 0) = (-1)^m Y_{\ell - m}(\cos\theta, 0) \qquad (5.20)$$

imply that $\widetilde{T}(m_1, m_2, m_3) = \widetilde{T}(-m_1, -m_2, -m_3)$, which halves the available equations again. The number is further reduced by the constraint that those $\widetilde{T}(m_1, m_2, m_3)$s whose $|m_1 + m_2| = |m_3| \geq l_3^{max}$ must be zero. Finally, the number of independent $\widetilde{T}(m_1, m_2, m_3)$s increases as $(2l_1^{max} + 1)(2l_2^{max} + 1)/2$ while that of tripolar coefficient increase as $(l_1^{max} + 1)(l_2^{max} + 1)(l_3^{max} + 1)$, which is an order higher and makes the direct solving of all of them impossible.

Fortunately, not all the $\mathcal{I}_{\ell_1 \ell_2 \ell_3}$s are independent. Firstly, there is the triangle constraint

$$|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2 \qquad (5.21)$$

Those coefficients whose subscripts fail to satisfy it are simply zeros. In addition, if some of the qs in the triple correlation are chosen to be the same, from Eqn(2.31) and symmetry property of $3jm$ symbols, the permutation of their corresponding ℓs leads to identical value up to a factor of -1:

$$\begin{aligned} \mathcal{I}_{\ell_1 \ell_2 \ell_3} &= \mathcal{I}_{\ell_2 \ell_1 \ell_3} = \mathcal{I}_{\ell_1 \ell_3 \ell_2} = \mathcal{I}_{\ell_3 \ell_2 \ell_1} \\ &= (-1)^{\ell_1 + \ell_2 + \ell_3} \mathcal{I}_{\ell_2 \ell_3 \ell_1} = (-1)^{\ell_1 + \ell_2 + \ell_3} \mathcal{I}_{\ell_3 \ell_1 \ell_2} \qquad (5.22) \end{aligned}$$

The actual number of independent, unknown $\mathcal{I}_{\ell_1 \ell_2 \ell_3}$s depends on the specific number of both $\ell_1^{max}, \ell_2^{max}, \ell_3^{max}$ and m_1, m_2, m_3 and has to be calculated at runtime. These two properties indicate that there should also be much less unknown than $(l_1^{max} + 1)(l_2^{max} + 1)(l_3^{max} + 1)$. Consequently, it is still possible to build more equations than unknown when ℓ^{max} is small ($\simeq 10$), and a least

93

square optimization routine can be applied to evaluate the $\mathcal{I}_{\ell_1 \ell_2 \ell_3}$s of interest at proper precision. This numerical calculation has been implemented and confirmed the expected result.

There is another advantage for using triple correlation worth mentioning. The rotational invariance of the triple correlation can be used to deal with fragmented data. $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ used in averaging in Eqn(5.18) can be rotated together along the z axis to avoid bad pixels without affecting the final result. It is not difficult to conclude that the omission of up to $\frac{2\pi}{3}$ of the circle can be tolerated.

From tripolar coefficients to spherical harmonics coefficients

In the previous subsection the radial components of the diffraction intensity were separated from the angular components in the triple correlation. However, the I of different ℓ, m and q are still coupled by the Clebsch-Gordan coefficients.

To solve this problem, note that firstly, all the $I_{00}(q)$ come directly from

$$\{\{I_0(q_i) \otimes I_0(q_i)\}_0 \otimes I_0(q_i)\}_{00} = I_{00}^3(q_i) \tag{5.23}$$

Then the cross correlation of $\ell = 2$, $B_2(q, q')$ can be calculated from the following relationships

$$
\begin{aligned}
\{\{I_2(q) \otimes I_2(q)\}_0 \otimes I_0(q')\}_{00} &= B_2(q, q) I_{00}(q') \\
\{\{I_2(q) \otimes I_0(q)\}_2 \otimes I_2(q')\}_{00} &= I_{00}(q) B_2(q, q') \\
\{\{I_0(q) \otimes I_2(q)\}_2 \otimes I_2(q')\}_{00} &= I_{00}(q) B_2(q, q') \tag{5.24}
\end{aligned}
$$

If the correct I_2s of some of the q_1s and q_2s are already solved (see the next subsection), the rest can only be calculated by solving the following linear

equations for different q_3s

$$\{\{I_2(q_1) \otimes I_2(q_2)\}_2 \otimes I_2(q_3)\}_{00} \tag{5.25}$$

Similarly, I_4 comes from

$$\{\{I_2(q_1) \otimes I_2(q_2)\}_4 \otimes I_4(q_3)\}_{00} \tag{5.26}$$

and I_6 from

$$\{\{I_2(q_1) \otimes I_4(q_2)\}_6 \otimes I_6(q_3)\}_{00}$$

$$\{\{I_4(q_1) \otimes I_4(q_2)\}_6 \otimes I_6(q_3)\}_{00} \tag{5.27}$$

I_8 from

$$\{\{I_2(q_1) \otimes I_6(q_2)\}_8 \otimes I_8(q_3)\}_{00}$$

$$\{\{I_4(q_1) \otimes I_4(q_2)\}_8 \otimes I_6(q_3)\}_{00}$$

$$\{\{I_4(q_1) \otimes I_6(q_2)\}_8 \otimes I_8(q_3)\}_{00}$$

$$\{\{I_6(q_1) \otimes I_6(q_2)\}_8 \otimes I_8(q_3)\}_{00} \tag{5.28}$$

Generally, there are $2\ell+1$ unknown variables for $I_\ell(q_3)$. On the other hand, the number of equations formed by different combinations of the triple correlation, in which two of I_ℓ are know with lower ℓ is much higher than this number. Firstly, to solve for $I_{\ell+2}$, there are $\ell^2/4$ pairs of $I_{\ell_1} \otimes I_{\ell_2}$ to be utilized, however, in order that $\ell_1 + \ell_2 \geq \ell_3$, only $\frac{\ell}{4}(\frac{\ell}{2}+1)$ pair of them are nonzero. And there are also another n^2 pairs, where n is the number of sampling along radial direction, of different q_1 and q_2 which increase the number to $\frac{\ell}{4}(\frac{\ell}{2}+1)n^2$. All the spherical harmonics coefficients at the qs where the tripolar coefficients are solvable from triple correlation can be calculated in this way.

95

The I_2 as starting point from cross correlation

For certain ℓ, choosing $2\ell + 1$ different qs, a $2\ell + 1$ dimension square matrix $\mathbf{I}$ can be composed with each row the I_ℓ of certain q

$$\mathbf{I}_{i,j} = I_{\ell j}(q_i) \tag{5.29}$$

so that

$$\mathbf{B} = \mathbf{I}\,\mathbf{I}^{\mathbf{T}} \tag{5.30}$$

The elements $\mathbf{B}_{qq'}$ is just $\{\{I_\ell(q) \otimes I_\ell(q')\}_{00}$. As $\mathbf{B}$ is a real symmetrical matrix, $\mathbf{I}$ can be solved analytically as a eigenvalue problem up to an unitary matrix $\mathbf{R}$.

$$\mathbf{B} = \mathbf{I}\,\mathbf{I}^{\mathbf{T}} = \mathbf{I}\mathbf{R}(\mathbf{I}\mathbf{R})^{\mathbf{T}} \tag{5.31}$$

The origin of this kind of ambiguity is that the initial orientation of the object when $\omega = 0$ is not specified. The $\mathbf{R}$ matrix is the $2\ell + 1$ dimensional representation of this arbitrary rotation. From the definition of triple correlation, $\{\{I_{\ell_1}(q_1) \otimes I_{\ell_2}(q_2)\}_{\ell_3} \otimes I_{\ell_3}(q_3)\}_{00}$ is invariant if the $\mathbf{R}$s of $I_{\ell_1}(q_1)$, $I_{\ell_2}(q_2)$ and $I_{\ell_3}(q_3)$ are different representations for the same SO(3) group element.

It is now clear that for each $\mathbf{I}$, $\mathbf{R}$ can be chosen freely when solving Equ.(5.30). However, these $\mathbf{I}$ values cannot be used together in a single spherical harmonics expansion as they come from the different orientation of the same method, and to which orientation it belongs to is unpredictable unless the whole structure is known. Actually, this fact corresponds to the ambiguity of cross correlation that has already been discussed in Sec.(5.2). However, it is still very useful as it provides a starting point for the recursive method using triple correlation to reduce the third order equations to linear equations. In the recursive process the orientation assigned by cross correlation is fixed as

triple correlation is unique up to a 3D rotation and this degree of freedom has been used up when choosing the eigenvalues as the solution of **I**.

5.4 Bootstrapping

Although the previous method can only be applied to the low frequency band limited qs, it provides the data to push ahead the calculation on those qs where tripolar coefficients are not solvable.

The general idea is to make use of Eqn(5.19), separating the summation into two parts with respect to the ℓ of interest, and substituting $m_3 = -m_1 - m_2$ make it an linear equations in $I_{\ell_3 m}$

$$
\widetilde{T}(m_1, m_2) = \sum_{\ell_1 \geq |m_1|, \ell_2 \geq |m_2|, \ell_3^{max} \geq \ell_3 > \ell} \mathcal{I}_{\ell_1 \ell_2 \ell_3}(q_1, q_2, q_3) \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \cdot
$$
$$
Y_{\ell_1 m_1}(\cos\theta_{q_1}, 0) Y_{\ell_2 m_2}(\cos\theta_{q_2}, 0) Y_{\ell_3 m_3}(\cos\theta_{q_3}, 0) +
$$
$$
\sum_{\ell_1 \geq |m_1|, \ell_2 \geq |m_2|, \ell \geq \ell_3 \geq |m_1+m_2|} \mathcal{I}_{\ell_1 \ell_2 \ell_3}(q_1, q_2, q_3) \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \cdot
$$
$$
Y_{\ell_1 m_1}(\cos\theta_{q_1}, 0) Y_{\ell_2 m_2}(\cos\theta_{q_2}, 0) Y_{\ell_3 m_3}(\cos\theta_{q_3}, 0) \tag{5.32}
$$

If the first term on the right hand side is available and we call $\widetilde{T}(m_1, m_2)_{partial}$, and choose the proper m_1, m_2 that $|m_1 + m_2| = \ell$, so that only the $\ell_3 = \ell$ is left in the summation, then expanding $\mathcal{I}_{\ell_1 \ell_2 \ell_3}$ gives

$$
\widetilde{T}(m_1, m_2) - \widetilde{T}_{partial} =
$$
$$
\sum_{\ell \geq m \geq -\ell} \sum_{\ell_1 \geq |m_1|, \ell_2 \geq |m_2|} \left(\sum_{\ell_1 \geq m_1' \geq -\ell_1, \ell_2 \geq m_2' \geq -\ell_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1' & m_2' & m \end{pmatrix} I_{\ell_1 m_1'}(q_1) I_{\ell_2 m_2'}(q_2) \right)
$$
$$
\begin{pmatrix} \ell_1 & \ell_2 & \ell \\ m_1 & m_2 & m_3 \end{pmatrix} Y_{\ell_1 m_1}(\cos\theta_{q_1}, 0) Y_{\ell_2 m_2}(\cos\theta_{q_2}, 0) Y_{\ell_3, m_3}(\cos\theta_{q_3}, 0) I_{\ell m}(q_3)
$$
$$
\tag{5.33}
$$

which is a linear equation in $I_{\ell m}(q_3)$. Had all the $I_{\ell m}(q_1)$ and $I_{\ell m}(q_2)$ been solved, $I_{\ell_3^{max} m}(q_3)$ could be solved at first as there is no $\widetilde{T}_{partial}$ at this moment. If ℓ_3^{max} is not much larger than ℓ_1^{max} or ℓ_2^{max}, enough equations can be collected from all the $\widetilde{T}(m_1, m_2), (|m_1 + m_2| = \ell_3^{max}, \ell_3^{max} - 1)$ from all the possible q_1, q_2 pairs. The exact number of equation also depends on these parameters and is up to runtime determination, but there is almost always enough of them compared with the $2 \times \ell + 1$ unknowns in practical simulation. Once all the $I_{\ell_3^{max} m}(q_3)$s are available, we can move on to do the same thing for $I_{\ell_3^{max}-2,m}(q_3)$ until we reach $I_{2m}(q_3)$, and in this process, all the higher order $I_{\ell m}(q_3)$ than the current one are used to calculate $\widetilde{T}_{partial}$. After all the spherical harmonics coefficients are solved for q_3, they can be used as q_1 or q_2 when taking the step up to the next q_i.

The whole algorithm can be concluded as a bootstrap process. It starts from some solved coefficients to solve the highest order ℓ for the next point along the radius, then uses the result for the next lower ℓ until all of them are calculated. That point is then marked as solved and used to repeat the same process for next point.

Obviously, in practical computation, when error is taken into account, the problem of error accumulation becomes significant as each step depends on the previous result and serves as the reference for the future calculations. Numerical simulation is conducted for data generated from a set of preassigned spherical harmonics expansions, and it is found that if preassigned data is used as the solved part of each $I_{\ell m}(q_i)$ the algorithm works effectively. However, if the calculated value is used, and if the slightly different data is substituted back for further calculation provided that the error at the beginning is small, the error accumulates very fast and error is seen above 10% after less than 100
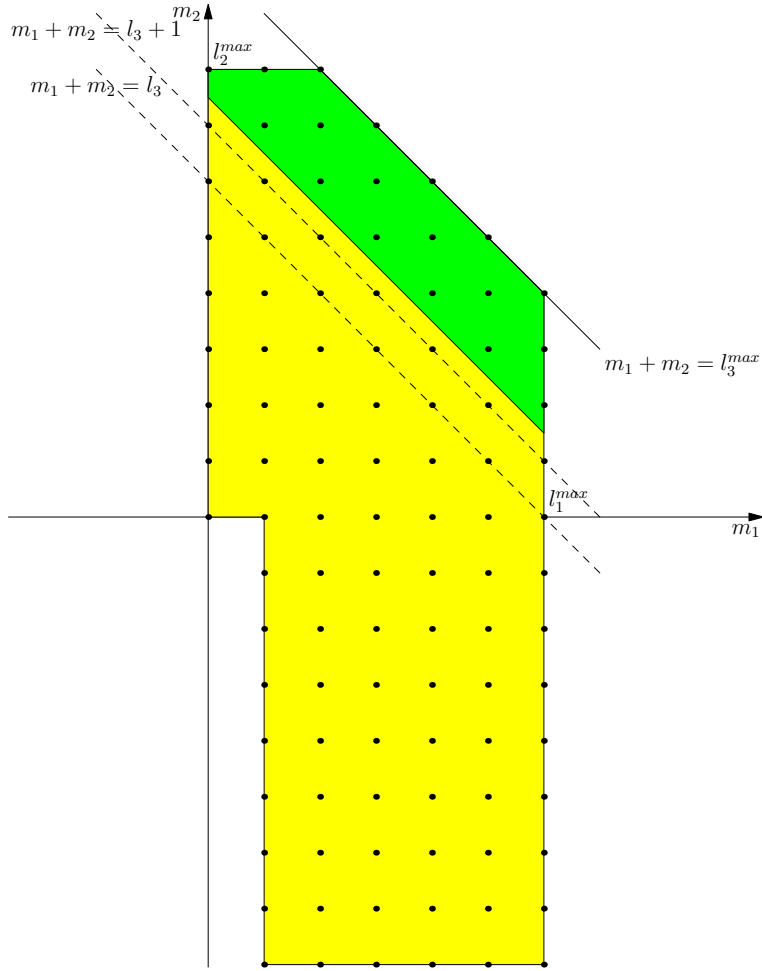
Figure 5.1: This figure shows the bootstraping method in m space. Each dot is an independent $T(m_1, m_2)$. Those within the green zone have been used to calculate spherical harmonics coefficients for higher order $\ell + 3$, and those along the dashed lines are used to form the equations for the current ℓ_3. Then the dashed lines move down for two units to the next set of unused dots.

iterations, which is far from enough for real world problem.

The fast exponential error accumulation indicates that the decoupling conducted in the previous subsection is too aggressive. The current computer is capable of doing least linear fitting at much larger scale than $2\ell+1$ we encountered at low resolution (thus small ℓ). Combining a bunching of different ℓs at the same q_i in the bootstraping stage reduces the number of iterations significantly, and numerical calculation shows that the error increases much slower and it is possible to propagate from the q_is calculated from cross correlation to at least 20 more outer q_is, therefore make it a practical algorithm.

5.5   Conclusion

The method above can be applied to the data with high signal-noise ratio and highly incomplete (due to beam stop, gap or dead zone in the detector, or streak from water buffer) and yield the 3D intensity in reciprocal space without explicit sorting. The only requirement is the diffraction patterns be taken from uniformly random orientation.

The bootstrapping algorithm is linear, analytical and determinant in theory. The correctness is verified with numerical calculation with error-free data.

BIBLIOGRAPHY

[1] J. C. H. Spence. Two-wavelength inversion of multiply scattering soft x-ray intensities to charge density. *Acta Crystallographica Section A*, 65:28–38, 2008.

[2] John C. H. Spence, Richard A. Kirian, Xiaoyu Wang, Uwe Weierstall, Kevin E. Schmidt, Thomas White, Anton Barty, Henry N. Chapman, Stefano Marchesini, and James Holton. Phasing of coherent femtosecond x-ray diffraction from size-varying nanocrystals. *Opt. Express*, 19(4):2866–2873, Feb 2011.

[3] D. Sayre. Prospects for long-wavelength x-ray microscopy and diffraction. In M. Schlenker, M. Fink, J. P. Goedgebuer, C. Malgrange, J. C. Viénot, and R. H. Wade, editors, *Imaging Processes and Coherence in Physics*, volume 112 of *Lecture Notes in Physics*, pages 229–235. Springer-Verlag, Berlin, 1980.

[4] J.R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21:2758, 1982.

[5] R.W. Gerchberg and W.O. Saxton. *Optik*, 34:275, 1971.

[6] V. Elser, I. Rankenburg, and P. Thibault. Searching with iterated maps. *Proceedings of the National Academy of Sciences*, 104(2):418–423, 2007.

[7] AckermannW., AsovaG., AyvazyanV., AzimaA., BaboiN., BahrJ., BalandinV., BeutnerB., BrandtA., BolzmannA., BrinkmannR., BrovkoO. I., CastellanoM., CastroP., CataniL., ChiadroniE., ChorobaS., CianchiA., CostelloJ. T., CubaynesD., DardisJ., DeckingW., Delsim-HashemiH., DelserieysA., Di PirroG., DohlusM., DustererS., EckhardtA., EdwardsH. T., FaatzB., FeldhausJ., FlottmannK., FrischJ., FrohlichL., GarveyT., GenschU., GerthCh., GorlerM., GolubevaN., GraboschH.-J., GreckiM., GrimmO., HackerK., HahnU., HanJ. H., HonkavaaraK., HottT., HuningM., IvanisenkoY., JaeschkeE., JalmuznaW., JezynskiT., KammeringR., KatalevV., KavanaghK., KennedyE. T., KhodyachykhS., KloseK., KocharyanV., KorferM., KolleweM., KoprekW., KorepanovS., KostinD., KrassilnikovM., KubeG., KuhlmannM., LewisC. L. S., LiljeL., LimbergT., LipkaD., LohlF., LunaH., LuongM., MartinsM., MeyerM., MichelatoP., MiltchevV., MollerW. D., MonacoL., MullerW. F. O., NapieralskiO., NapolyO., NicolosiP., NolleD., NunezT., OppeltA., PaganiC., PaparellaR., PchalekN., Pedregosa-GutierrezJ., PetersenB., PetrosyanB., PetrosyanG., PetrosyanL., PflugerJ., PlonjesE., Polet-

toL., PozniakK., PratE., ProchD., PucykP., RadcliffeP., RedlinH., RehlichK., RichterM., RoehrsM., RoenschJ., RomaniukR., RossM., RossbachJ., RybnikovV., SachwitzM., SaldinE. L., SandnerW., SchlarbH., SchmidtB., SchmitzM., SchmuserP., SchneiderJ. R., SchneidmillerE. A., SchneppS., SchreiberS., SeidelM., SertoreD., ShabunovA. V., SimonC., SimrockS., SombrowskiE., SorokinA. A., SpanknebelP., SpesyvtsevR., StaykovL., SteffenB., StephanF., StulleF., ThomH., TiedtkeK., TischerM., ToleikisS., TreuschR., TrinesD., TsakovI., VogelE., WeilandT., WeiseH., WellhoferM., WendtM., WillI., WinterA., WittenburgK., WurthW., YeatesP., YurkovM. V., ZagorodnovI., and ZapfeK. Operation of a free-electron laser from the extreme ultraviolet to the water window. *Nature Photonics*, 1(6):336–342, June 2007.

[8] Qun Shen, Ivan Bazarovc, and Pierre Thibault. Diffractive imaging of nonperiodic materials with future coherent x-ray sources. *Journal of Synchrotron Radiation*, 11:432–438, 2004.

[9] D. Starodub, R. B. Doak, K. Schmidt, U. Weierstall, J. S. Wu, J. C. H. Spence, M. Howells, M. Marcus, D. Shapiro, A. Barty, and H. N. Chapman. Damped and thermal motion of laser-aligned hydrated macromolecule beams for diffraction. *The Journal of Chemical Physics*, 123(24):244304, 2005.

[10] D. Starodub, P. Rez, G. Hembree, M. Howells, D. Shapiro, H. N. Chapman, P. Fromme, K. Schmidt, U. Weierstall, R. B. Doak, and J. C. H. Spence. Dose, exposure time and resolution in serial x-ray crystallography. *Journal of Synchrotron Radiation*, 15:62–73, 2008.

[11] Richard Neutze, Remco Wouts, David van der Spoel, Edgar Weckert, and JanosHajdu. Potential for biomolecular imaging with femtosecond x-ray pulses. *Nature*, 406:752, 2000.

[12] K. J. Gaffney and H. N. Chapman. Imaging atomic structure and dynamics with ultrafast x-ray scattering. *Science*, 316:1444–1448, 2007.

[13] J. Miao, K. O. Hodgson, and D. Sayre. A new approach to 3-d structures of biomolecules utilizing single molecule diffraction images. *Proc. Natl Acad. Sci. USA*, 98:6641–6645, 2001.

[14] V. L. Shneerson, A. Ourmazd, and D. K. Saldin. Crystallography without crystals. i. the commonline method for assembling a three-dimensional diffraction volume from single-particle scattering. *Acta Crystallographica Section A*, A64:303–315, 2008.

[15] J. C. H. Spence and R. B. Doak. Single molecule diffraction. *Physical Review Letters*, 92:198102, 2004.

[16] D. Sayre. Some implications of a theorem due to shannon. *Acta Cryst.*, 5:843, 1952.

[17] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney. Reconstruction of the shapes of gold nanocrystals using coherent x-ray diffraction. *Phys. Rev. Lett.*, 87(19):195505, Oct 2001.

[18] Henry N. Chapman, Anton Barty, Stefano Marchesini, Aleksandr Noy, Stefan P. Hau-Riege, Congwu Cui, Malcolm R. Howells, Rachel Rosen, Haifeng He, J.C.H. Spence, Uwe Weierstall, Tobias Beetz, Chris Jacobsen, and David Shapiro. High resolution three-dimensional x-ray diffraction microscopy. *J. Opt. Soc. Am. A*, 23:1179–1200, 2006.

[19] J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara. Atomic resolution imaging of a carbon nanotube from diffraction intensities. *Science*, 300:1419, 2003.

[20] R. Barakat and G. Newsam. *J. Math. Phys.*, 25:3190–3193, 1984.

[21] R. P. Millane. Iterative projection algorithms for solving inverse problems. In *Proc. Oceans 2003*, pages 2714–2719, 2003.

[22] J.P. Abrahams and A.W.G. Leslie. *Acta Cryst.*, D52:30–42, 1996.

[23] V. Elser. Phase retrieval by iterated projections. *J. Opt. Soc. Am. A*, 20:40, 2003.

[24] H. H. Bauschke, P. L. Combettes, and D. R. Luke. *J. Opt. Soc. Am. A*, 19:1334–1345, 2002.

[25] H. H. Bauschke, P. L. Combettes, and D. R. Luke. *J. Opt. Soc. Am. A*, 20:1025–1034, 2003.

[26] D. R. Luke. *Inverse Problems*, 42:576–595, 2005.

[27] S. Marchesini. A unified evaluation of iterative projection algorithms for phase retrieval. *Rev. Sci, Inst.*, 78:011301, 2007.

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

[29] Uwe Weierstall, Q.Chen, J.C.H. Spence, M. Howells, M. Isaacson, and R. Panepucci. *Ultramicroscopy*, 90:171, 2001.

[30] S. Eisebett, M. Lorgen, and W. Eberhart. *Appl. Phys. Lett.*, 84:3373, 2004.

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

[32] J. C. H. Spence, B. Calef, and J. M. Zuo. Dynamic inversion by the method of generalized projections. *Acta Crystallographica Section A*, 55:112–118, 1999.

[33] J. C. H. Spence. Direct inversion of dynamical electron diffraction patterns to structure factors. *Acta Crystallographica Section A*, 54(1):7–18, Jan 1998.

[34] Christoph T Koch and J.C.H. Spence. A useful expansion of the exponential of the sum of two non-commuting matrices, one of which is diagonal. *Journal of Physics A: Mathematical and General*, 36(3):803, 2003.

[35] J.C.H. Spence. *High Resolution Electron Microscopy*. Oxford University Press, 3rd edition, 2003.

[36] Russell Fung, Valentin Shneerson, Dilano K. Saldin, and Abbas Ourmazd. Structure from fleeting illumination of faint spinning objects in flight. *Nature Physics*, 5:64, 2009.

[37] K. E. Schmidt, J. C. H. Spence, U. Weierstall, R. Kirian, X. Wang, D. Starodub, H. N. Chapman, M. R. Howells, and R. B. Doak. Tomographic femtosecond x-ray diffractive imaging. *Physical Review Letters*, 101(11):115507, 2008.

[38] D. K. Saldin, V. L. Shneerson, R. Fung, and A. Ourmazd. Structure of isolated biomolecules from ultrashort x-ray pulses: exploiting the symmetry of random orientations. *Journal of Physics: Condensed Matter*, 21:134014, 2009.

[39] V. L. Shneerson and D. K. Saldin. Molecular shapes from small angle x-ray scattering: Extension of the theory to higher scattering angles. *Acta Crystallographica Section A*, A65:128–134, 2009.

[40] J. C. H. Spence, K. Schmidt, J. S. Wu, G. Hembree, U. Weierstall, B. Doak, and P. Fromme. Diffraction and imaging from a beam of laser-aligned proteins: resolution limits. *Acta Crystallographica Section A*, A61(2):237–245, Mar 2005.

[41] Detlef-M. Smilgies. Molecular axes and planes as an eigenvalue problem. *Journal of Applied Crystallography*, 41(2):363–365, Apr 2008.

[42] H. Goldstein. *Classical Mechanics.* Addison-Wesley, 1950.

[43] Joachim Frank. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies.* Oxford University Press, 2007.

[44] Zvi Kam. The reconstruction of structure from electron micrographs of randomly oriented particles. *Journal of Theoretical Biology*, 82:15–39, 1980.

[45] Stephen W. Provencher and Robert H. Vogel. Three-dimensional reconstruction from electron micrographs of disordered specimens i. method. *Ultramicroscopy*, 25:209–222, 1988.

[46] N. Boisset, J.C. Taveau, and J.N.Lamy. An approach to the architecture of scutigera coleoptrata haemocyanin by electron microscopy and image processing. *Biol. Cell.*, 68:73–84, 1990.

[47] Christopher M. Bishop and Markus Svensen. The generative topographic mapping. *Neural Computation*, 10:215–234, 1998.

[48] Ne-Te Duane Loh and Veit Elser. Reconstruction algorithm for single-particle diffraction imaging experiments. *Phys. Rev. E*, 80(2):026705, Aug 2009.

[49] Uwe Weierstall, H Chapman, R Doak, J.C.H.Spence, D Starodub, D Shapiro, P Kennedy, J Warner, G Hembree, and P Fromme. Droplet streams for serial crystallography of proteins. *Experiments in Fluids*, 44:675689, 2008.

[50] D. DePonte, U. Weirstall, D. Starodub, J. Warner, J.C.H.Spence, and R.B.Doak. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys D: Appl Phys*, 41:195505, 2008.

[51] F. Wang, E. Weckert, B.Ziaja, D.S.D.Larsson, and D. van der Spoel. Coherent diffraction of a single virus particle: The impact of a water layer on the available orientational information. *Physical Review E*, 83:031907, 2011.

[52] B. L. Henke, E. M. Gullikson, and J. C. Davis. X-ray interactions: Photoabsorption, scattering, transmission, and reflection at e=50-30,000 ev, z=1-92. *At. Data Nucl. Data Tables*, 54:181–342, 1993.

[53] J. M. Cowley and A. F. Moodie. The scattering of electrons by atoms and crystals. i. a new theoretical approach. *Acta Cryst.*, 10:609–619, 1957.

[54] William Humphrey, Andrew Dalke, and Klaus Schulten. Vmd – visual molecular dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.

[55] Greg Hura, Daniela Russo, Robert M. Glaeser, Teresa Head-Gordon, Matthias Krackd, and Michele Parrinello. Water structure as a function of temperature from x-ray scattering experiments and ab initio molecular dynamics. *Phys. Chem. Chem. Phys.*, 5:1981–1991, 2003.

[56] A. H. Narten and H. A. Levy. Liquid water: Molecular correlation functions from x-ray diffraction. *Journal Chemical Physics*, 55:2263, 1971.

[57] Henry Chapman, Petra Fromme, Anton Barty, Thomas White, Richard Kirian, Andrew Aquila, Mark Hunter, Joachim Schulz, Daniel De-Ponte, Uwe Weierstall, R Doak, Filipe Maia, Andrew Martin, Ilme Schlichting, Lukas Lomb, Nicola Coppola, Robert Shoeman, Sascha Epp, Robert Hartmann, Daniel Rolles, Artem Rudenko, Lutz Foucar, Nils Kimmel, Georg Weidenspointner, Peter Holl, Mengning Liang, Miriam Barthelmess, Carl Caleman, Sebastien Boutet, Michael Bogan, Jacek Krzywinski, Christoph Bostedt, Sasa Bajt, Lars Gumprecht, Benedikt Rudek, Benjamin Erk, Carlo Schmidt, Andre Homke, Christian Reich, Daniel Pietschner, Lothar Struder, Gunter Hauser, Hubert Gorke, Joachim Ullrich, Sven Herrmann, Gerhard Schaller, Florian Schopper, Heike Soltau, Kai-Uwe Kuhnel, Marc Messerschmidt, John Bozek, Stefan Hau-Riege, Matthias Frank, Christina Hampton, Raymond Sierra, Dmitri Starodub, Garth Williams, Janos Hajdu, Nicusor Timneanu, M Seibert, Jakob Andreasson, Andrea Rocker, Olof Jonsson, Martin Svenda, Stephan Stern, Karol Nass, Robert Andritschke, Claus-Dieter Schroter, Faton Krasniqi, Mario Bott, Kevin Schmidt, Xiaoyu Wang, Ingo Grotjohann, James Holton, Thomas Barends, Richard Neutze, Stefano Marchesini, Raimund Fromme, Sebastian Schorb, Daniela Rupp, Marcus Adolph, Tais Gorkhover, Inger Andersson, Helmut Hirsemann, Guillaume Potdevin, Heinz Graafsma, Bjorn Nilsson, and John Spence. Femtosecond x-ray protein nanocrystallography. *Nature*, 470(7332):73–77, Feb 2011. 10.1038/nature09750.

[58] Richard A. Kirian, Xiaoyu Wang, Uwe Weierstall, Kevin E. Schmidt, John C. H. Spence, Mark Hunter, Petra Fromme, Thomas White, Henry N. Chapman, and James Holton. Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express*, 18(6):5713–5723, Mar 2010.

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

[60] 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 scat-

tering fluctuations of many copies. *Phys. Rev. Lett.*, 106(11):115501, Mar 2011.

[61] Richard A. Kirian, Thomas A. White, James M. Holton, Henry N. Chapman, Petra Fromme, Anton Barty, Lukas Lomb, Andrew Aquila, Filipe R. N. C. Maia, Andrew V. Martin, Raimund Fromme, Xiaoyu Wang, Mark S. Hunter, Kevin E. Schmidt, and John C. H. Spence. Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. *Acta Crystallographica Section A*, 67(2):131–140, Mar 2011.

[62] R. Kakarala. A group-theoretic approach to the triple correlation. In *Higher-Order Statistics, 1993., IEEE Signal Processing Workshop on*, pages 28 –32, 1993.

[63] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum Theory of Angular Momentum*. World Scientific Pub Co Inc, 1988.