Theses and Dissertations

December 2016

Reconstruction of 3D Image for Particles By the Method of Angular Correlations from XFEL Data

Sung Soon Kim
University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Condensed Matter Physics Commons

RECONSTRUCTION OF 3D IMAGE FOR PARTICLES BY THE METHOD OF ANGULAR CORRELATIONS FROM XFEL DATA

by

Sung Soon Kim

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin – Milwaukee

December 2016

ABSTRACT

RECONSTRUCTION OF 3D IMAGE FOR PARTICLES BY THE METHOD OF ANGULAR CORRELATIONS FROM XFEL DATA

by

Sung Soon Kim

The University of Wisconsin-Milwaukee, 2016
Under the Supervision of Professor Dilano K. Saldin

The world's first X-ray Free Electron Laser (XFEL), the Linac Coherent Light Source (LCLS) at the Stanford Linear Accelerator Center (SLAC), is now generating X-ray pulses of unprecedented brilliance (one billion times brighter than the most powerful existing sources), and at the amazing rate of only a few femtoseconds [2]. The first such experiments are being performed on relatively large objects such as viruses, which produce low resolution, low-noise diffraction patterns on the basis of the so called "diffraction before destruction" principle. Despite the promise of using XFEL for the determination of the structures of viruses, the results so far from experimental data present difficulties in working to reconstruct 3D images for the viruses by our method. One of the rare instances in which images are reconstructed from experimental data is the mimi virus work of Hajdu et al, [56]. In this present paper, we examine the capabilities of the method that is based on the angular momentum decomposition of scattered intensities, which enables us to overcome common problems such as missing or imperfect data that are inevitable in experiments. This angular momentum decomposition method helps to avoid the effect of a

finite beam size, and existing gap size. In addition to the problem caused by the finite panels of detectors used when the data are collected, the effect of noise, curved Ewald Sphere, shot to shot variations of incident X-ray pulse intensities and shots to multiple nano particles are also studied.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables and Boxes

Appendices

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my academic advisor Professor Dilano K. Saldin for the patient and continuous instructions of my Ph.D. study, research and supplying variety of knowledge. I would like to thank him for encouraging me to go forward to get the degree. I would also like to thank my committee members, professor Peter Schwander who gave me short and concrete answers for my questions anytime, professor Michael Weinert who showed the importance of assumptions in the thesis, professor Daniel Agterberg who gave distinguished lectures in the class and professor Wilfred Tysoe who did the proof reading for my thesis and corrected all detailed mistakes. I thank all of them for serving as my committee members even at hard times in summer session through sharp questioning, indicating deficiencies and insightful discussions. I have been fully moved by their dedications and cares for students through all my years in the department of physics. I also thank graduate students, Manoj Kashyap, Sandi Wibowo, and Ashish Mishra with whom I took several classes and shared ideas of how to get through the difficult times that we face in a tough environment. Hin-Chunck Poon has been a good friend who worked as a postdoc and helped me a lot with his long year research experience. Ahmad Hosseinizadeh also used to give me short comments how to understand some physical problems which are not directly related to my area but are very important to enrich the field of my research. Lastly, I really appreciate my mother, my lovely wife and two children for their support and unconditional love through all these hard times of my graduate school years. I dedicate this thesis to my family, professors, and all graduate students in the department of physics who pursue their goals.

Chapter 1

INTRODUCTION

The advent of X-ray free electron lasers (XFELs) such as Linac Coherent Light Source (LCLS) [1,2] produces X-rays of a brightness some ten billion times [3,4] greater than any pre-existing X-ray source (including synchrotrons), and a few femtoseconds duration. It has been suggested that such pulses could avoid the damage process and allow structure determinations without crystallization of bioparticles such as proteins, viruses and amino-acids. Reconstructing 3D image of a microscopic entity such as a virus from randomly oriented many ultrashort diffraction patterns (DPs) generated by XFEL as 2D data has been proposed in different ways. There are some ways to reconstruct 3D structures from completely randomly and uniformly oriented diffraction snapshots such as Manifold Embedding by Diffusion Map [5][6][7], EMC (Expand, Maximize and Compress)[8] or Angular Correlation method [9,10]. Manifold Embedding by Diffusion Map is an alternative approach in order to extract the structure of biological particles such as viruses from diffraction patterns with random orientations. This method is basically a nonlinear PCA (Principle Component Analysis) technique, which generates a manifold from a cloud of points where each point is a diffraction snapshot in the dataset. In fact, diffusion map provides a mathematical link to the symmetrized icosahedral Wigner D-functions [24], which are applied to extract the orientations of snapshots, and then those orientations are used to recover 3D structure of the object. EMC is another algorithm for 3D structural recovery from diffraction patterns. Through the EMC algorithm, a 3D intensity volume is iteratively updated by Expansion, Maximization and Compression. These three iterative steps map the 3D intensity model to a tomographic representation and the reconstructed

intensity that is used for phasing and extract the structure. Angular correlations that we prefer to use here are the products of two or three vectors, whose elements represent intensities in diffraction patterns. These correlations are independent of particle orientation and therefore ideal for the XFEL problem. It has been found possible to reconstruct a 3D image of the particle in an arbitrary orientation from these orientation independent quantities. In most cases, the analysis of XFEL diffraction pattern proposed so far would be so-called single particle methods. Thus, methods have to be developed for eliminating diffraction patterns from both zero and multiple particle hits. In contrast with the method of angular correlations, all that is necessary is to eliminate hits on zero particles and as the method is capable of more efficiently utilizing available data. Another advantage of the method working with the angular momentum representation is that it is possible to devise a simple test on whether the measured signal is coming mainly from an icosahedral particle, which is a symmetry associated with many viruses [11]. The point is that many of the low-order angular momenta (e.g. $l=2$, $l=4$) cannot be summed to produce signals consistent with an icosahedral scattering (the odd values of l are forbidden by Friedel's Law [12,13]). If the measured signal is really from an icosahedral object, the angular momenta of the representation of the diffraction volume have to be dominated by $l=0$, and $l=6$ terms, and the ratio of these term with those of $l=2$ and $l=4$, may help quantify how much of the signal is from an icosahedral scatterings and how much from stray scattering from e.g. apertures or non-icosahedral parts of the structure, etc. What is more, the low q data is dominated by isotropic scattering corresponding to $l=0$, a fact that may allow us to fill in the missing data in the beam stop by an appropriate analytical model. Through this angular correlation method, we developed a series of algorithms to reconstruct the 3D image of the objects

such as STNV (Satellite Tobacco Necrosis Virus) [14], PBCV-1 (Paramecium Bursaria Chlorella Virus1) [15], RDV (Rice Dwarf Virus) [16] or Nanorice particles [17] using computer simulated diffraction patterns from those models. The developed algorithms have shown the successful procedures in reconstructing 3D images of the models. However, even though LCLS has been producing real experimental diffraction patterns for some years now, the effort to reconstruct the 3D images of those models turned out to be a bit disappointing. Thus in this dissertation, we look for the discrepancies between the simulated data and the real experimental data. In addition, we also try to find the remedies to resolve the discrepancies by adding some artifacts to the simulated data such as the effects of noise, a curved Ewald Sphere [18], the absence of central intensities, or shot-to-shot variations in the incident X-ray intensities when we access the algorithms that have already developed and tested. The appropriate methods adjusting these effects should be adopted to achieve the goal. Especially for the small sized sample such as STNV with comparatively large scattering angles from an X-ray shot, more detailed treatments calculating intensities on a curved Ewald Sphere in the reciprocal space must be considered to generate the simulated data, which will help handling the real experimental diffraction patterns. We develop this packet of treatments for icosahedral viruses by taking a twelve atom cluster, STNV virus and PBCV-1 as samples to generate diffraction patterns testing through laptop computers, AVI or MORTIMER many core cluster supercomputer in UWM or Sherlock minicomputer in the department. In the case of noise, we assume Poisson noise in the intensities of Cartesian grid points where the detectors are placed. Once the noise effect appears not to deform the original shape of the sample, we lower the intensities on the outer most polar grid points to check on how effect of Poisson noise

affects the 3D imaging in the real space and how much we can lower the intensities to keep the original sample shape in the 3D images. Crucial to the treatment of Poisson noise is that the average of the correlations, which is what one evaluates from the experimental data, is the same as the correlation of the average intensities. Here there is some confusion in the literature. While in general there is no guarantee that this is so, in practice it does seem to be obeyed for intensities dominated by Poisson noise as is the case for any practical correlations, as we have demonstrated by computer to about 4 significant figures. Next, the curved Ewald Sphere (ES) in the reciprocal space is examined to testify that if the scattering angles are big enough to change the calculations for the angular correlations of intensities on the Cartesian grid points, the curvature of the ES will play an important role of the treatments. For the multi particle coherent and incoherent scattering case, two nanorice particle samples that are azimuthally symmetric are adopted to verify if the simulated diffraction patterns come into play along the algorithms that have been already developed. Finally, we put some bad artifacts in to the simulated data to resemble the experimental data so that these calculation methods will be useful in 3D reconstructions of nanoparticles. Here, comparisons are introduced between the real experimental diffraction patterns and simulated patterns where those bad artifacts are included, which will help understand how to deal with the data in reality. What we really focus on in this paper is the methodology how to overcome the discrepancies appearing in handling real experimental data with the algorithms that have been developed through simulated diffraction patterns. Computer generated simulated data showed well behaved tendencies following mathematical formats and expected styles of the results. One of the main stream of the methods to overcome the stated discrepancies is to test the simulated data after making

4

them the same as the real ones. For one instance, in the case of shot to shot variations of incident X-ray pulse intensities, we give Gaussian intensity variations to the incident X-ray by multiplying a Gaussian amplitude factor instead of using flat plain incident X-ray pulses when we generated computer simulated data. After adding this artifact, we tested it to make the 3D reconstructed structural image of the object acceptable. Another example is the simulated data with Poisson noise to mimic the noise that has always appeared in real experimental data. Assuming that we can only have ten or twenty thousand of experimental DPs with very small number of photons counted on the detector, we examined if the 3D reconstructed imaging procedures can successfully achieved with these many DPs of low level of photon counts. In the caption of each image, R-factor [19] is calculated to check the discrepancies between the input amplitudes and output amplitudes in the reciprocal space. Also in the reality, all DPs have lost central intensities. To overcome this deficiency, we extrapolated the DPs into the beam stops and checked if those complemented DPs can help reconstructing 3D structural images of objects. In chapter 7, we used RDV experimental diffraction patterns from LCLS to reconstruct the 3D image through the algorithms stated in the previous chapters in which the remedies for the artifacts are introduced. The difference between simulated data and experimental ones of RDV are so big that every single step of whole calculating procedures has been treated very carefully to attain a decent 3D reconstructed structural image. Once all the procedures had been done, we could get a decent 3D image of RDV (Rice Dwarf Virus) with icosahedral symmetry which is checked by B_l-values [20,21]. We found that the size of RDV from the reconstructed image is almost the same as we expected even though the shape is not a perfect icosahedron. There are still important limits to find 3D structural

image of nanoparticles through XFEL without using symmetric property or some constraints in performing the algorithm. In our algorithm, we also impose icosahedral symmetry in getting a diffraction volume consisting of intensities in the reciprocal space. Being free from imposing symmetries and constraints to find the original images of the objects are to be discussed further more with all researchers. In the final chapter, we list a couple of issues for the future, to pursue in developing new algorithms to handle real experimental data more efficiently and successfully.

Chapter 2

Formulation of Angular Correlation

2.1 Correlation Theory

A data set consisting of *N* Cartesian Diffraction Patterns of random particle orientations are needed for the reconstruction of the full 3D structure. We follow the approach based on the analysis of the average angular correlation $J(q,\phi;\ q,\phi+\Delta\phi)$ over all diffraction patterns of the measured intensities at two pixels specified by (q,ϕ) and $(q,\phi+\Delta\phi)$ on each diffraction pattern, as illustrated in Fig 1. The 3D distribution of scattered intensity in the reciprocal space may be expanded as a sum of spherical harmonics [22], $Y_{lm}(\theta,\phi)$, namely

$$I(\bar{q}) = \sum_{lm} I_{lm}(q)Y_{lm}(\phi,\theta) \tag{1}$$

The intensity in a particular diffraction pattern is a saucer shaped slice through this distribution, representing a part of the Ewald sphere S1 corresponding to an incident X-ray wave vector antiparallel to the Z axis as in Fig 2. Any point on this sphere may be specified by the polar coordinates defined with respect to the [x, y, z] frame. Since the Ewald sphere is a 2D manifold [23], only two of the three coordinates are independent. Indeed, geometrical considerations show that the intensity on the red Ewald sphere in Fig 2 may be written as

$$I(q,\phi) = \sum_{lm} I_{lm}(q)Y_{lm}(\theta,\phi) \tag{2}$$
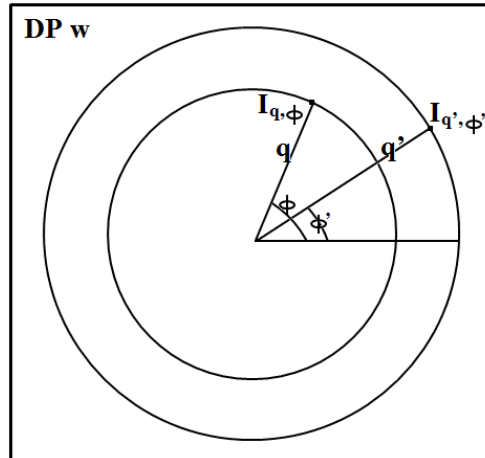
Fig 1. A diffraction pattern pixel may be labeled by the magnitude q of the scattering vector, and an azimuthal angle ϕ in the frame of reference attached to the diffraction pattern. A set of intensity cross-correlations may be constructed by multiplying the intensities $I_{q\phi}$ and $I_{q'\phi'}$ on each diffraction pattern (w) and summing over all diffraction patterns.

.
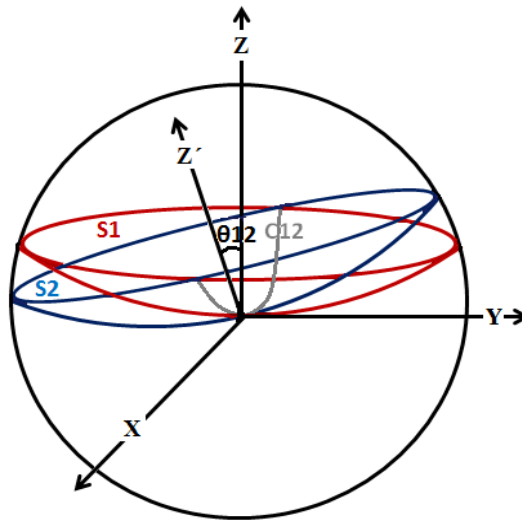
Fig 2. Construction of a 3D diffraction volume from Ewald spheres of random orientation is illustrated. Two Ewald Sphere sections, corresponding to incident X-ray antiparallel to z and z' axis are denoted by S1 and S2 above. A large number of such randomly oriented diffraction patterns allow the assignment of scattered intensity to all points in a 3D volume of reciprocal space.

The diffraction intensities would lie on a portion of the Ewald sphere (S1) of radius equal to the wave number $K = 2\pi/\lambda$, where λ is the wave length of the x-rays. Sets of points on each Ewald sphere may be specified by polar, azimuthal angles and radial distance q. According to Fig 3, polar angle θ depends on the radial distance q:

$$\theta(q) = \pi/2 - sin^{-1}(q/2K) \tag{3}$$
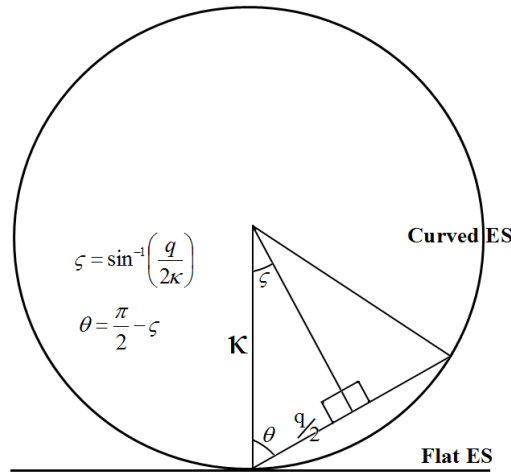
Fig 3. Section through the Ewald sphere S1, viewed antiparallel to X-axis

Thus the measured intensities in a diffraction pattern arising from radiation incident antiparallel to the z-axis is

$$I_z(q,\phi) = \sum_{lm} I_{lm}(q)Y_{lm}(\theta(q),\phi) \tag{4}$$

Now, consider the cross-correlation $J(q,\phi; q',\phi')$ with N number of diffraction patterns with each Wth molecular orientation may be written as

$$J(q,\phi;q'\phi') = \frac{1}{N}\sum_{W} I^W(q,\phi)I^W(q',\phi') \tag{5}$$

which may be computed from the experimental data without knowledge of the orientation of the individual particle. Considering the random orientations of the different particles, Eq. (5) can be rewritten as

$$J(q,\phi;q',\phi') = \frac{1}{N} \sum_w \sum_{lmm'} D_{lmm'}^{w*} I_{lm'}^*(q) Y_{lm}^*(\theta(q),\phi) \sum_{l'm''m'''} D_{l'm''m'''}^{w} I_{l'm'''}(q) Y_{l'm''}(\theta'(q'),\phi') \quad (6)$$

The Wigner D functions D^w [24] which are functions of three Euler angles specifying the orientations of the molecule. Performing the sum over all W, effectively a sum over the space of all the elements of the SO(3) group [25], and applying the orthogonality theorem [26] (see e.g. Tinkham 2003), we find

$$\frac{1}{N} \sum_w D_{lmm'}^{w*} D_{lm''m'''}^{w} = \frac{1}{2l+1} \delta_{ll'} \delta_{mm''} \delta_{m'm'''} \quad (7)$$

Performing the sum (6) using (7) with $\Delta\phi = |\phi'-\phi|$ will lead to

$$J(q,q',\Delta\phi) = \sum_l F_l(q,q',\Delta\phi) B_l(q,q') \quad (8)$$

where

$$F_l(q,q',\Delta\phi) = \frac{1}{2l+1} \sum_m Y_{lm}^*(\theta(q),\phi) Y_{lm}(\theta'(q'),\phi')$$
$$= \frac{1}{4\pi} P_l[\cos\theta(q)\cos\theta'(q') + \sin\theta(q)\sin\theta'(q')\cos\Delta\phi] \quad (9)$$

From Fig3, $\theta(q) \approx \frac{\pi}{2} \approx \theta'(q')$. For a flat Ewald Sphere [27], therefore (9) becomes

$$F_l(q,q',\Delta\phi) = \frac{1}{4\pi} P_l(\cos\Delta\phi). \quad (10)$$

Where P_l is a Legendre Polynomial of order l, and

$$J_{qq'\Delta\phi} = \sum_{l} F_{qq'\Delta\phi} B_l \qquad (11)$$

Angular cross correlations of the intensities can be found from Eq. (5) and all elements of the matrix F consists of real valued Legendre Polynomials. Thus, the Eq. (8) is purely real, and may be solved for the real coefficients by B_l through matrix inversion;

$$B_l = \sum_{\Delta\phi} \{F^{-1}\}_{l,\Delta\phi} J_{qq'\Delta\phi} \qquad (12)$$
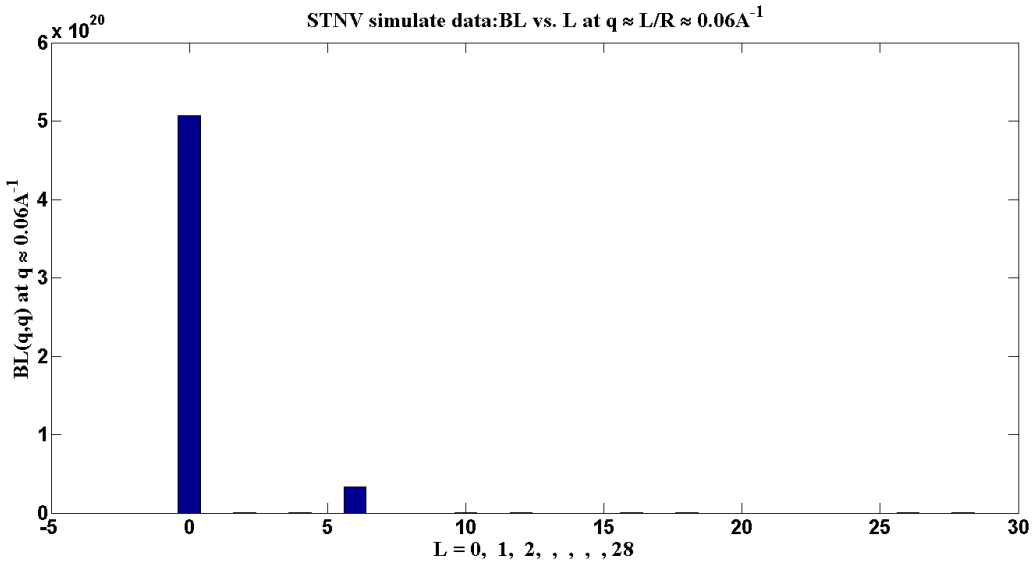
Fig 4. Plot Bl vs. l at $q \approx 0.06\text{Å}^{-1}$ from the simulated 10000 diffraction patterns for STNV virus extracted from PDB (Protein Data Bank) file (Entry No.= 2BUK) whose radius is $R \approx 100A$. At $l = 0$ and 6, the peaks appear the strongest.

2.2 Numerical Tests for the Angular correlation parameters

We tested our approach of 3D reconstruction of icosahedral virus to a realistic model from PDB entry 2BUK. First we calculated the amplitude $A(q)$ of the scattered X-ray on the flat ES using PDB file consisting of atoms and their coordinates via Eq. (13),

$$A(q) = \sum_{j} f_j(q) \exp(i\bar{q} \cdot \bar{r}_j) \quad \text{with} \quad f(q) = \sum_{i=1}^{4} a_i \exp\left(-b_i \left(\frac{q}{4\pi}\right)^2\right) + c \qquad (13)$$

where a_i, b_i and c are Cromer-Mann's coefficients [28] that are specified by each atom in the sample, by squaring them to get the intensities on the Cartesian grids. Next, converting these intensities $I(q) = |A(q)|^2$ into the intensities on the polar grids to get the angular correlation, $J_{q,q',\Delta\phi}$ (we will refer to this $C2$ $(q,q',\Delta\phi)$ hereafter) by Eq. (10), enables us to find the B_l's by Eq. (12) through F matrix of Legendre polynomials inversion by Eq. (9), which we need to get the diffraction volume representing intensities on the reciprocal space. To get B_l, one can use the orthogonality of Legendre polynomials for convenience. Since we do not have to use large square matrix, which becomes the integral Eq. (14), since each element in the inversion matrix of F^{-1} is generated by Eq. (14).

$$B_l = \left(\frac{2l+1}{2}\right)\int_0^\pi C2(\Delta\phi)P_l(\cos\Delta\phi)\sin\Delta\phi\,d(\Delta\phi) \qquad (14)$$

Our work done so far is that we generated 8 times oversampling [29], 141x141x500 DPs from STNV model (pdb-file:entry#:2buk) for the interpolation to work accurately. We produced scattering amplitudes and square them to get the intensities on each Cartesian grid points. This would simulate the nicely behaved experimental data for the future use from the detector that is based on the Cartesian grids. Next, the interpolation from Cartesian to polar grid points is performed to use the angular correlation property.

Example of Interpolation from Cartesian Grid Points to Polar Grid Points
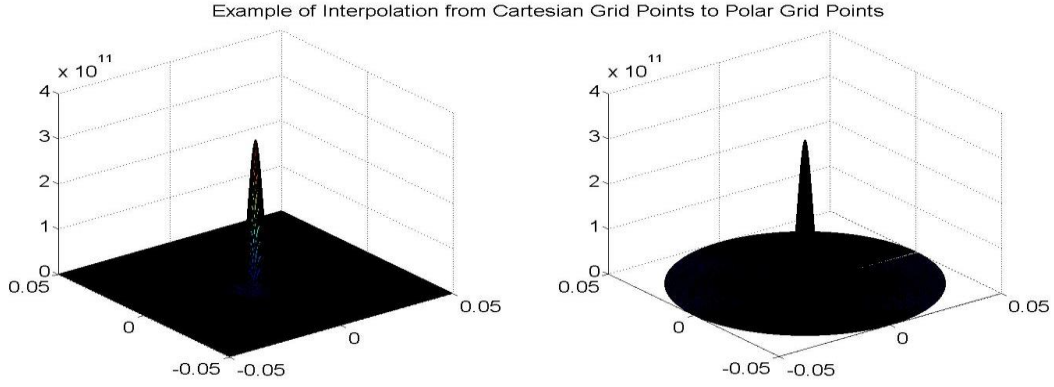
Fig 5. Plot of the intensities on Cartesian coordinates and Polar coordinate after interpolation of STNV simulated diffraction patterns from pdb-file (pdb entry#: 2BUK)

Once the intensities on the polar grid points come out, the angular correlation $J(q,\phi,q',\phi')$ are found by Eq. (5). Fig 6 is a plot for $J(q,\phi,q',\phi')$ vs. $\Delta\phi$'s. Then, using the Legendre polynomial matrix $P_l(\cos\Delta\phi)$, Eq. (10), one can also find $B_l(q,q')$ using Eq. (12) with the inversion of $F_l(q,q',\Delta\phi)$ by Eq. (9).
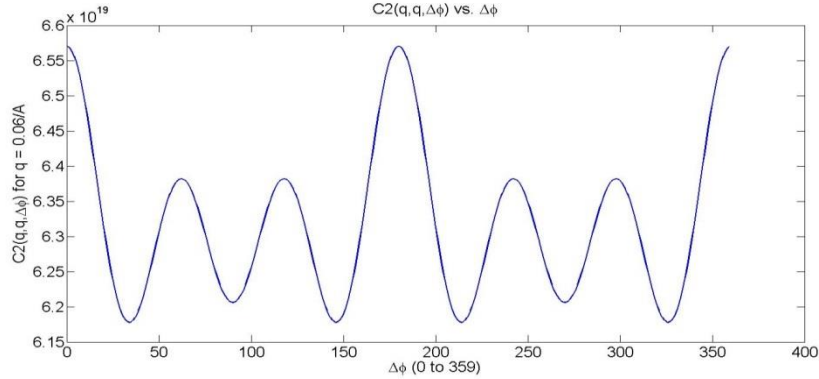
Fig 6. Plot of the intensity correlations $J(q,q',\Delta\phi)$ where $\Delta\phi = \phi\text{-}\phi'$ for $q = q' \approx$ 0.06Å^{-1} indicated. These quantities are calculated from the simulated diffraction patterns of randomly oriented Satellite Tobacco Necrosis Virus (pdb entry: 2BUK). Each of these plots a linear combination of Legendre polynomials, exactly as predicted by the theory. The structural information resides in the magnitudes of the expansion coefficients $B_l(q,q')$ of the Legendre polynomials.

13
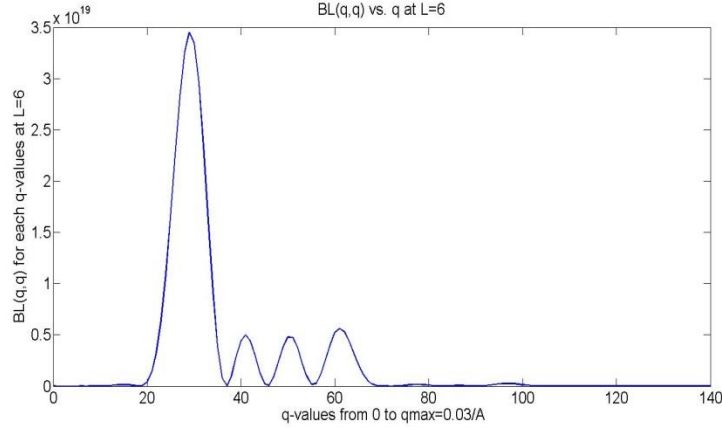
Fig 7. Plot of $B_l(q,q)$ vs q at $l = 6$. X-axis shows that qmax $\approx 0.3\text{Å}^{-1}$ is divided into 140. As we expended, at $q \approx 0.06\text{Å}^{-1}$ (28th or 29th value from the figure), which is corresponding q for $l = 6$ by $q = l/R$, the peak appearing at this q represents that the icosahedral symmetry holds for STNV virus.

Finally, we could get the diffraction volume through these B_l's, and reconstructed the original model (STNV: pdp entry# 2BUK). The algorithm to find diffraction volume will be discussed in section 4.1 and 6.2. The diffraction volume consisting of intensities in the reciprocal space is used to find the phases using the phasing algorithm proposed by Oszlanyi and Süto [30][31]. Fig 8 shows the brief illustration of an iterative method for finding phases.
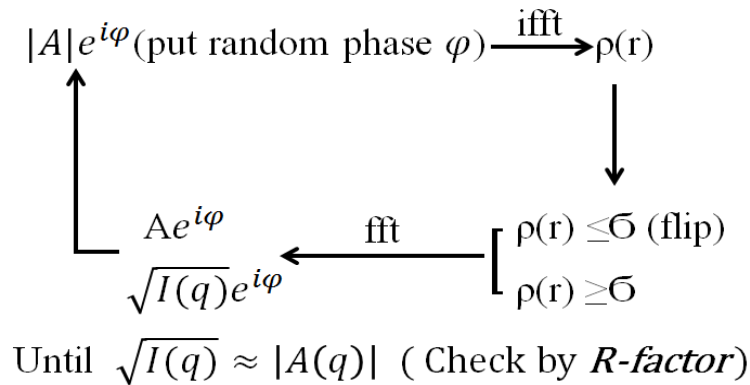
Fig 8. The iterative phasing algorithm of charge flipping method proposed by Oszlanyi and Süto [30][31].

14

Fig 9 and Fig 10 shows the reconstructed STNV model through the algorithm mentioned above [10].
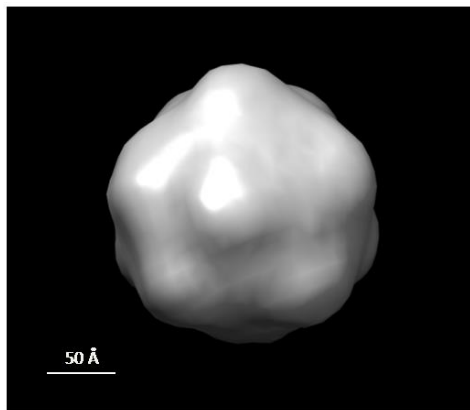
Fig 9. Reconstructed image from the diffraction volume of a single STNV particle computed directly from a structure factor calculation. STNV is about 20 nm in diameter. The figure depicts a view of the icosahedron approximately its 5-fold rotation axis. The reconstruction assumed a maximum value of q, $q_{max} \approx$ 0.314Å^{-1}, implying a resolution $\approx$ 20Å. $R_f = 0.042$.
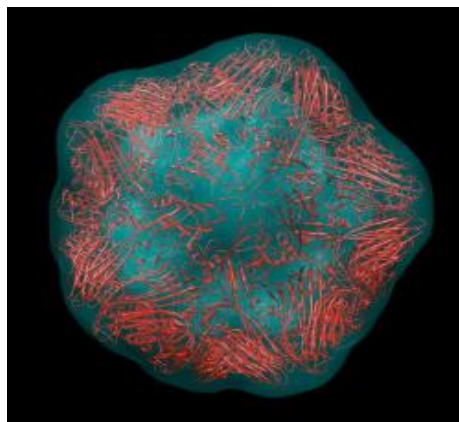
Fig 10. Same as Fig 9, except that the diffraction volume was reconstructed from the average of angular correlation on 10000 diffraction patterns of STNV from uniformly distributed directions over SO(3). A ribbon diagram of the STNV structure from pdb entry#= 2BUK is superimposed on the reconstructed electron density illustrates the accuracy of the reconstruction [10].

2.3 Summary of Procedures For Reconstructing 3D images of objects

Three dimensional reconstructed structural imaging process starts from collecting diffraction patterns of an object. Either simulated DPs or experimental DPs have 2D matrices form. Simulated data can be collected after the calculations according to Eq. (13) while experimental data are collected by running XFEL equipment. Once these Cartesian DPs are collected, interpolation from the intensities on the Cartesian grid points to polar grid points should be done to get $C2$, B_l. To find signs of B_l elements, we will discuss the method in section 4.1. Using B_l elements, diffraction volume consisting of intensities in the reciprocal space could be obtained. However, intensities in the reciprocal space have no phases that can give us the information of charge densities of the object. Thus, we use the phasing algorithm to find the charge densities in the real space. The Fig 11 shows the brief diagram of the 3D imaging process starting from Cartesian DPs for an object.

Cartesian DPs
↓
Polar DPs
↓
C2, BL (**Sign Determination using C3, TL**)
↓
Intensities (Reciprocal Space)
↓
Charge Densities (Real Space)
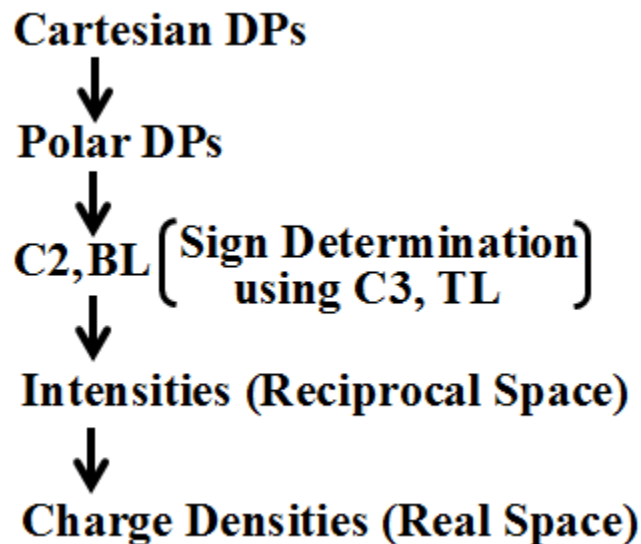
Fig 11. The 3D reconstructed structural imaging process after collecting DPs of an object. The data are either simulated DPs or experimental DPs.

Chapter 3

Apply Poisson Noise to Simulated DPs and Its effects

3.1 Poisson Noise

First, we can consider Poisson distribution function [32],

$$P(n,\lambda) = \frac{\lambda^n e^{-\lambda}}{n!} \qquad (15)$$

where n is the random variable and λ is the expectation value of n. In Poisson distribution, $\lambda = E(n)$ is the relation between the mean value and random variables. Using this formula, we adopt Knuth Algorithm [33] for generating a Poisson random number.

Box 1. The algorithm for generating Poisson random number by Knuth

Initialize n=0, p=1

Do :

       $n = n + 1$

       $p = p * rnd\ [0,\ 1]$

       while $p > e^{-\lambda}$

Return n - 1

Box 1. This algorithm indicates that increasing n value up to $p > e^{-\lambda}$ implies p is greater probability of appearance than nonappearance, since if n=0 is plugged into the formula Eq (15), then the probability will be $P(0, \lambda) = e^{-\lambda}$. In the matlab code, poissrnd(λ) is a very convenient tool for generating random variables. That is $p > P(0, \lambda) = e^{-\lambda}$.

3.2 Apply Poisson Noise on Simulated DPs and their effects

Simulated diffraction patterns are ideal to follow the scattering rules through X-ray shots. However, in reality, the real experimental data always contain noise that is discretely added on the detectors in each shot. The reasonable guess is to assume that these discrete items would be intensities modified by Poisson noise, since its random variables are integers (shot noise). Here in this section we examine the effect of Poisson noise [34] added on the simulated diffraction patterns, and of lowering the outermost polar ring intensities to see how the added noise affects the original shape of the sample as the noise are applied. The overview of the process for the Effect of the Poisson Noise on the 3D structural image is briefly illustrated in Box 2.
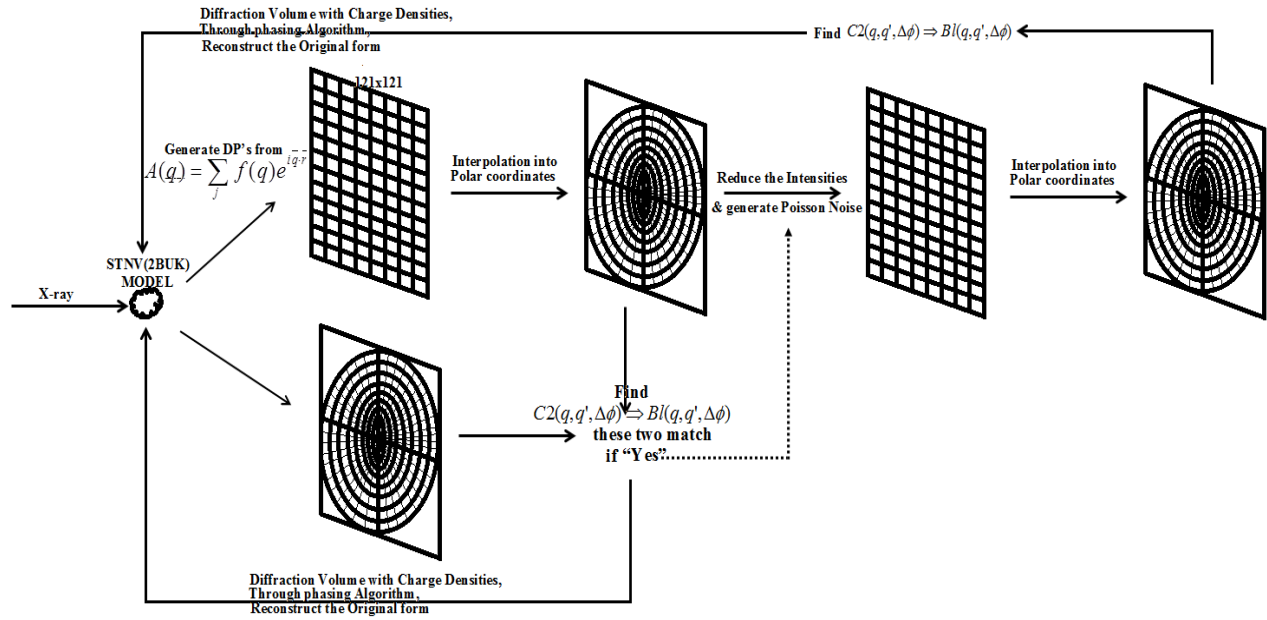
Box 2. Above diagram shows the algorithm to add Poisson noise on the simulated diffraction patterns of randomly oriented STNV virus and reduce the outermost circular intensities. This procedure shows the reconstruction of the sample model to examine the effect of discretely added noise on the diffraction patterns.

We chose some sample models, such as a twelve atom cluster, STNV (Satellite Tobacco Necrosis Virus) and PBCV-1(Paramecium Bursaria Chlorella Virus 1) to test the effect of Poisson noise added on the randomly oriented simulated data that consist of intensities on Cartesian grid points. Fig12, Fig13, and Fig14 show the original 12 atom cluster, STNV and PBCV-1 how their original icosahedral images are affected after adding Poisson noise of 0.05 ph/SP is applied to their simulated diffraction patterns.
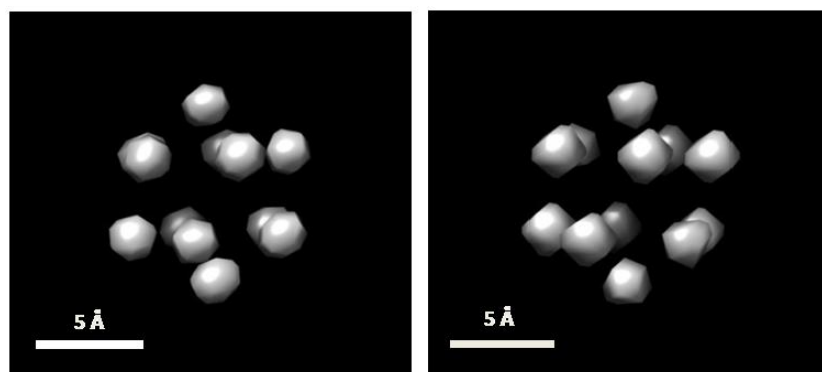
Fig 12. By taking 20000 DPs and adding Poisson noise on the simulated 12 atom cluster diffraction patterns with resolution ~ 3Å, we compare the noise-free and noised cases. These two reconstructed images do not make noticeable difference between noise-free (left) and noised with the noise level of 0.05 ph/SP (right). $R_f = 0.6083$. These two images show both icosahedral shapes.
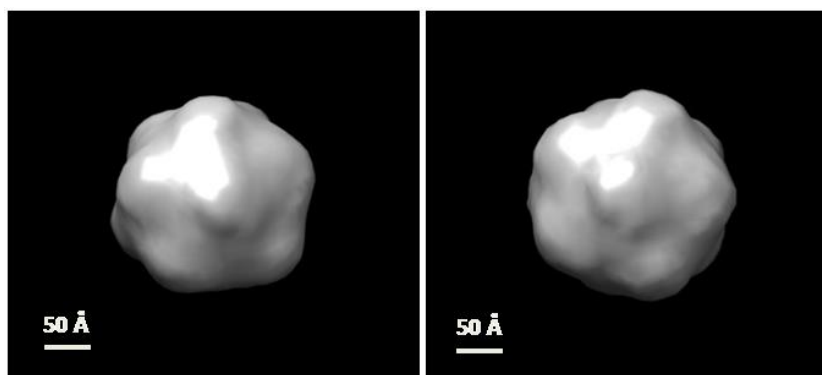
Fig 13. After taking simulated 1000 DPs (randomly oriented) of STNV (pdb entry#: 2BUK), access the procedures to make the images. The left image is for noise-free DPs and the right image is for noised (0.05 ph/SP) DPs with resolution ≈ 20Å. $R_f = 0.0579$.
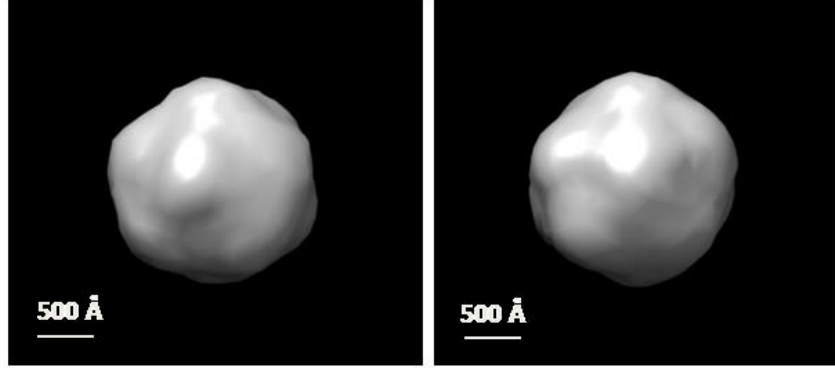
Fig 14. After taking simulated 1000 DPs (randomly oriented) of Chlorella (pdb entry#: PBCV-1), access the procedures to make the images. The left image is for noise-free DPs and the right image is for noised (0.05 ph/SP) DPs with resolution ≈ 200Å. $R_f = 0.1923$.

Now, we want to reduce the intensities with noises down to 0.05 photons/Shannon pixels to see how the original images of 12 atom cluster, STNV and Chlorella will change. We add Poisson noise on the Cartesian grid points of DPs and interpolate them. Then, find $C2(q,q',\Delta\phi)$, $C3(q,q',\Delta\phi)$ and $Bl(q,q')$, $Tl(q,q')$ [35]. by

$$C2(q,q';\Delta\phi) = < \int I(q,\phi)I(q,\phi+\Delta\phi)d\phi > \tag{16}$$

$$C3(q,q';\Delta\phi) = < \int [I(q,\phi)]^2 I(q,\phi+\Delta\phi)d\phi > \tag{17}$$

$$Bl(q,q') = (l+0.5)\int_0^\pi P_l(x)\sin\Delta\phi\, C2(q,q';\Delta\phi)d\phi \tag{18}$$

$$Tl(q,q') = (l+0.5)\int_0^\pi P_l(x)\sin\Delta\phi\, C3(q,q';\Delta\phi)d\phi \tag{19}$$

where $x=\cos\theta\cos\theta'+\sin\theta\sin\theta'\cos\Delta\phi$.

Also, the average number of photon counts per Shannon Pixel is introduced [36] as

$$< N_p >= F \cdot P \cdot r_e^2 \cdot N_{atom} \cdot |f| \cdot \Omega_p \tag{20}$$

where F is the photon fluence, N_{atom} is the number of non-hydrogen atoms in the molecule, P is the polarization factor, and f is the average atomic scattering factor (e.g. between C and O, $f \sim 7$) with for a particle of width w, solid angle $\Omega_p \sim \lambda^2/(2\pi)^2 (\Delta q_s)^2 = \lambda^2/(2w)^2$ where $\Delta q_s = \pi/w$ is the Shannon interval in the reciprocal space.

Thus,

$$< N_p >= F \cdot P \cdot r_e^2 \cdot N_{atom} \cdot |f|^2 \left(\frac{\lambda}{2\pi}\right)^2 \tag{21}$$

Eq. (21) is used frequently used while the intensities become lower and examine how the image of icosahedral samples would differ according to this intensity level.

3.3 Lower limit of the intensity strength with Poisson noise

Fig 15 and Fig 16 illustrate Poisson-noised DPs, how lowered intensities on the grid points of diffraction patterns would change the 3D images of samples, and what would be the low limit of the strength of those intensities to keep the original image shape.
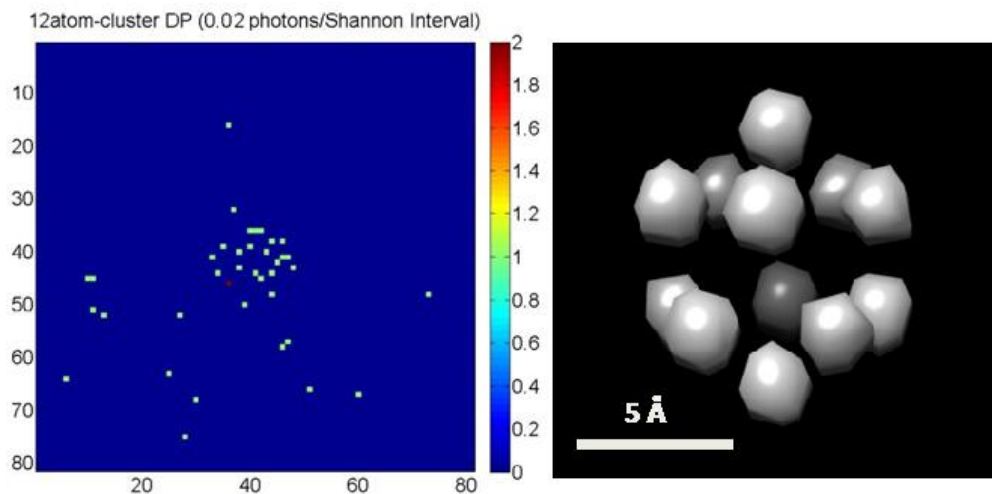
Fig 15. One diffraction pattern of 12 atom cluster with the intensity of 0.02 ph/SP is shown. 20000dps of this intensity level with Poisson-noise added will make the icosahedral 3D image. $R_f = 0.6838$.
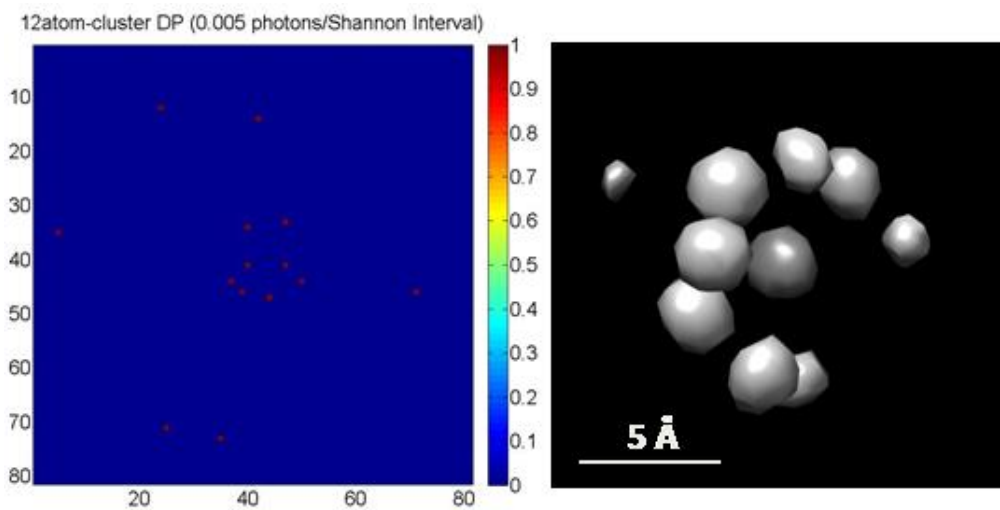
Fig 16. One diffraction pattern of 12 atom cluster with 0.005 ph/SP. 20000dps with Poisson noise will make the original icosahedral 3D image deformed. $R_f = 0.7174$

Seen from Fig 15 and Fig 16, one can conclude that if the intensities get lowered down below the certain level, it is impossible to reconstruct the original structure through angular correlation method. Fig 17 illustrate that we generated simulated 1000 diffraction patterns

of STNV and performed all procedures to get the final image . If we lowered the intensities of 1000 DPs down to 0.01 ph/SP and add Poisson noise on the Cartesian grid points of this STNV model, the original shape does not remain the same. Probably, more many DPs should be used to maintain the original icosahedral shape.
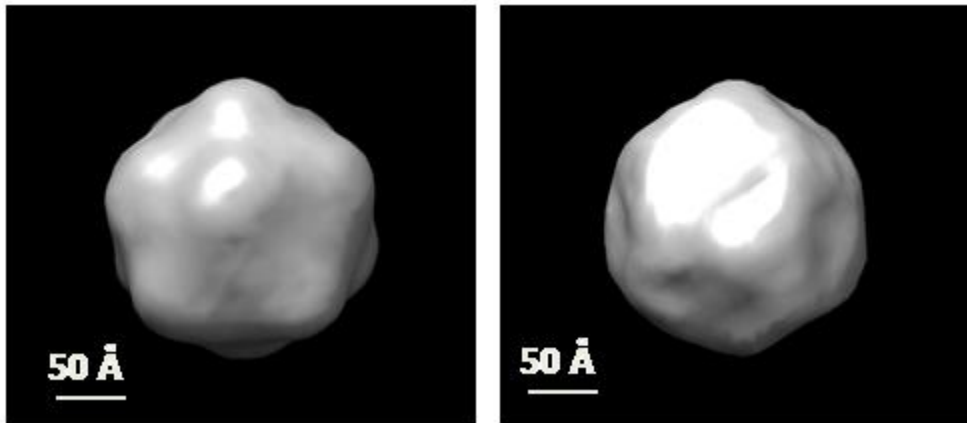
Fig 17. The left side image shows the icosahedral shape after performing 1000 noise free DPs of STNV(pdb entry#:2BUK). The right side image shows the image after performing 1000 DPs of Cartesian intensities lowered down to 0.01 ph/SP with Poisson noise added. $R_f = 0.0487$

When the intensities on the Cartesian grid points are lowered down to 0.005 photon/Shannon interval, we could not get the original icosahedral image. The same phenomenon happened in PBCV-1 case. Thus, even though the intensities on thousands of diffraction patterns of simulated data that are weakened do not seem to be a bad artifact in icosahedral 3D imaging procedures.
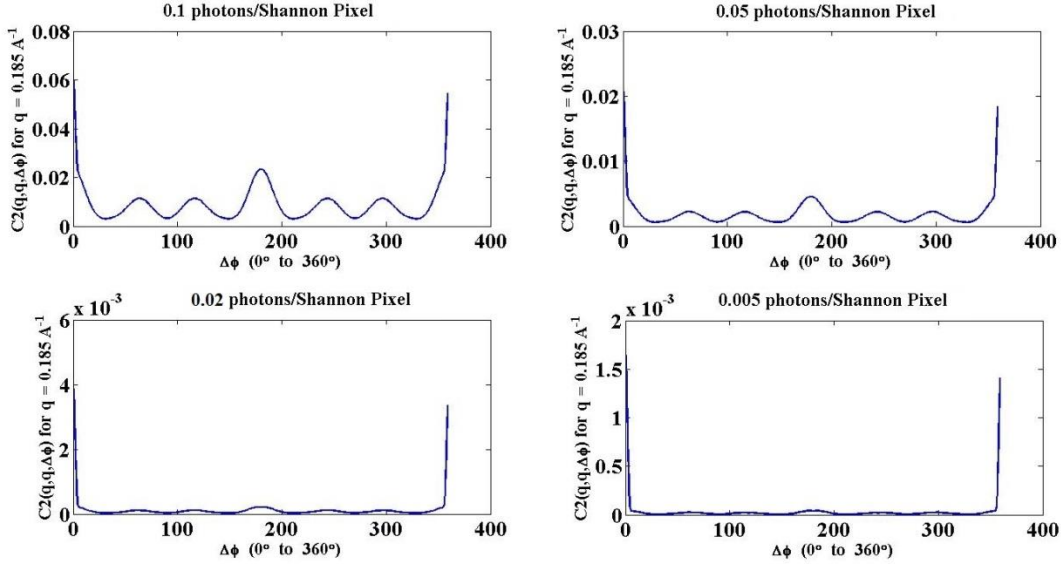
Fig 18. $C2(q,q,\Delta\phi)$ vs. $\Delta\phi$ plots from 20000 DPs for 12 atom cluster are illustrated. When the Diffraction Patterns of 0.005 ph/SP are accessed, the expected 3D image does not appear as in Fig 16.

In our many trials, it is observed that taking many more simulated DPs than 10000 DPs of icosahedral sample under incident X-ray shots of intensity level (0.05 ph/SP ~ 0.01 ph/SP) gives us more chances to achieve the original icosahedral reconstructed 3D images. Even though in processing simulated 20000 DPs of 12 atom cluster with 0.02 ph/SP intensity level gave us the original icosahedral image, with the X-ray shot intensity level of 0.01ph/SP in 20000 DPs are unstable to give the original icosahedral image. In STNV or PBCV-1 case, 1000 DPs with the Poisson noise of 0.05ph/SP mostly give the reconstructed icosahedral 3D images. For accessing the DPs of X-ray shots of low intensity level with 0.05 ph/SP ~ 0.005 ph/SP, many more than 100000 DPs should be taken to make the reconstructed 3D image stably possible.

Chapter 4

Formulation of Parameters for Curved Ewald Sphere

Saldin et. al. showed in 2011 that it is possible to recover images of icosahedral viruses from the average angular correlations amongst diffraction patterns of unknown orientations. We repeat the calculations here, taking account of the effects of a curved Ewald Sphere, the fact that the detectors form Cartesian grids from which it is necessary to interpolate onto circles on each diffraction pattern, and assuming the particle orientations are random in SO(3). We find excellent reconstructions in most cases with as few as 300 diffraction patterns (DPs). From Eq. (18) and (19), $x = cos\theta cos\theta' + sin\theta sin\theta' cos\Delta\phi$ is an important parameter to get $C2, C3, Bl, Tl$. If we handle large viruses such as Mimi virus or Chlorella virus, we can ignore the curved Ewald Sphere effects, since θ close to $\pi/2$ by Eq. (3), so x would become $cos\Delta\phi$. However, when the sample particles are small enough to make scattering angles big, then those minute effects are not negligible. Thus we are supposed to take a curved ES. The following is a brief summary of the method used to calculate the angular correlations $C2$, $C3$ and their angular Fourier transforms [37] leading to the reconstruction of the final 3D image for the object. The diagram followed by the formulation will help understanding of how to get the simulated diffraction patterns on a curved Ewald Sphere and 2D flat pixelated Cartesian grids.

4.1 Angular Correlation $C2, C3$ and B_l, T_l leading to Diffraction Volume

From the above expression Eq. (16) and Eq. (17) will provide the average of paired the intensity product over all polar grid points with an angular separation. Among the discrete intensities, Eq. (16) and Eq. (17) can be described as [38]

$$C2(q,q';\Delta\phi) = \frac{1}{N} \sum_{\Delta\phi=0} I(q,\phi) \cdot I(q',\phi+\Delta\phi) \tag{22}$$

$$C3(q,q';\Delta\phi) = \frac{1}{N} \sum_{\Delta\phi=0} [I(q,\phi)]^2 \cdot I(q',\phi+\Delta\phi) \tag{23}$$

And Eq. (18) and Eq. (19) become

$$B_l(q.q') = \frac{2l+1}{2} \sum_{\Delta\phi=0} P_l(x)\sin\Delta\phi \cdot C2(q,q';\Delta\phi) \tag{24}$$

$$T_l(q,q') = \frac{2l+1}{2} \sum_{\Delta\phi=0} P_l(x)\sin\Delta\phi \cdot C3(q,q';\Delta\phi) \tag{25}$$

where $x=cos\theta cos\theta'+sin\theta sin\theta' cos\Delta\phi$ with Eq. (3).

By a different derivation as we show

$$B_l(q,q') = \sum_{m=-l}^{l} I_{lm}(q) \cdot I_{lm}(q') \tag{26}$$

$$T_l(q,q') = \sum_{\substack{l_1, l_2 \\ m_1,m_2,m}} G(l_1,m_1,l,m,l_2,m_2) I_{l_1 m_1}(q) I_{lm}^*(q') I_{l_2 m_2}(q) \tag{27}$$

Here, G's are the Gaunt's coefficients and $I_{lm}(q)$ are the spherical harmonics expansion coefficients of the 3D diffraction volume [39] of a single particle in the reciprocal space. For an icosahedral object, $I_{lm}(q)$ can be written as

$$I_{lm}(q) = a_{lm} g_l(q) \tag{28}$$

26

The **z**-axis is chosen as the 5-fold symmetry axis. $g_l(q)$'s denote the expansion of an icosahedral harmonics needed to specify the 3D diffraction volume of an icosahedral particle as described well in [35]. $g_l(q)$ is a real. $I_{lm}(q)$ is known up to a sign, since $\sum a_{lm}^2 = 1$, with $B_l(q,q) = g_l(q)g_l(q)$, $g_l(q) = \pm\sqrt{B_l(q,q)}$. To determine the signs with an equivalent expression of Eq. (27),

$$T_l(q,q) = \sum_{l_1,l_2} b(l_1,l,l_2)g_{l_1}(q)g_l(q)g_{l_2}(q) \qquad (29)$$

where

$$b(l_1,l,l_2) = \sum_{m_1,m,m_2} G(l_1m_1,lm,l_2m_2)a_{l_1m_1}a_{lm}a_{l_2m_2} \qquad (30)$$

To find the diffraction volume, $B_l(q,q)$ and $T_l(q,q)$ are used with l values. In this test, STNV (pdb-entry#: 2buk) is used. Because of its icosahedral symmetry, only the following l values are allowed for l_{max}=28. l = 0, 6, 10, 12, 16, 18, 20, 22, 24, 26, 28. These l values give only 2^{11} sign combinations. The best-fit combination can be searched exhaustively by comparing theoretical and experimental *B's* and *T's*.

4.2 Interpolations from the intensities on enlarged polar grids down to a Curved Ewald Sphere

When X-ray Free Electron Laser (XFEL) gives shots to a small particle such as STNV whose diameter is about 20 nm, the shots tend to diffract more widely than large particles like Mimi virus whose size is about 450 nm. This will lead us to consider the curvature of

the Ewald sphere in the reciprocal space and change the calculation done in the flat Ewald sphere. After generating simulated data on enlarged polar grid points in the reciprocal space, interpolation is to be performed so as to get the intensities on Cartesian grids to mimic the experimental data which is based on the Cartesian grids. Fig 19 shows the formation of parameters of a curved Ewald Sphere and how interpolations will be done through the different parameters from the calculation on the flat Ewald Sphere.
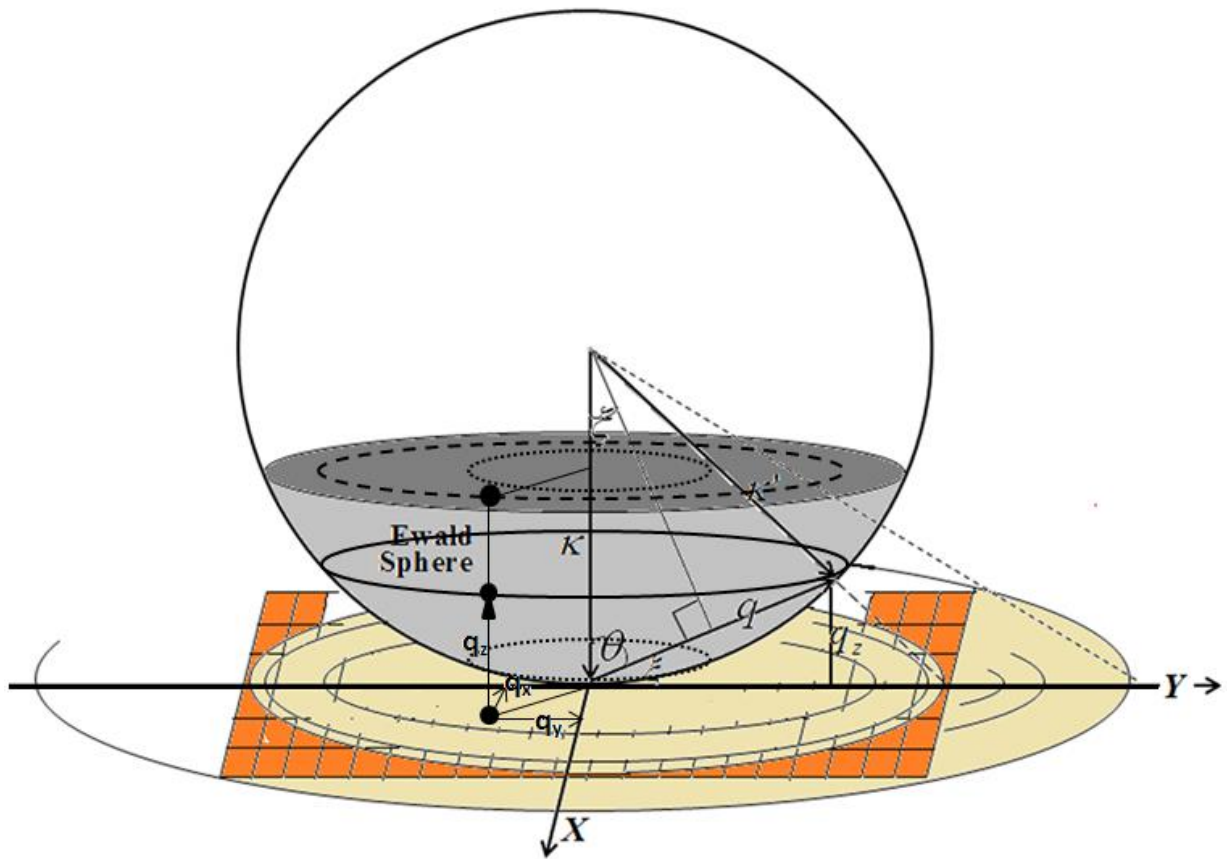
Fig 19. The figure illustrates that the parameters used in dealing with flat Ewald Sphere should be changed when X-ray shots on a small particle such as STNV give large scattering angles. The figure also shows how the interpolations will be done to make the angular correlations work for the 3D imaging process using icosahedral particles exposed on 10Å wave length of X-ray.

We carefully examined the effect of these curved Ewald Sphere [40] parameters with the curvature and grabbing the idea of reconstruction of the particle through these parameters rather than those of flat Ewald Sphere. We recognized that the curvature of the Ewald Sphere will help us understand the whole range of spectra for all imaging procedures we developed through simulated diffraction patterns. Using Eq. (13), one could generate many simulated diffraction patterns to do numerical tests for the flat Ewald Sphere. Different from a flat Ewald Sphere, we have carefully examined the effect of a curved Ewald Sphere on the 3D imaging process and noticed that Eq. (13) still functioning well while we apply the new variables through the curved Ewald Sphere set up into all the imaging calculations.

Eq. (13) gives simply coherent X-ray diffraction amplitudes and leads to Eq. (32), and the wave number defined by

$$K = 2\pi / \lambda \tag{31}$$

Using $\bar{q} = \bar{k}' - \bar{k}$ with Eq. (13) to get amplitudes on a curved Ewald Sphere, new sets of parameters come arise as

$$i\bar{q} \cdot \bar{r}_j = q_x r_{jx} + q_y r_{jy} + q_z r_{jz} \tag{32}$$

with

$$\theta(q) = \frac{\pi}{2} - \sin^{-1} \xi = \frac{\pi}{2} - \sin^{-1}\left(\frac{\lambda q}{4\pi}\right) \tag{33}$$

Thus, q_x, q_y, q_z become

$$q_x = q \cos\left(\sin^{-1}\left(\frac{\lambda q}{4\pi}\right)\right) \cos\phi \tag{34}$$

$$q_y = q\cos\left(\sin^{-1}\left(\frac{\lambda q}{4\pi}\right)\right)\sin\phi \qquad (35)$$

$$q_z = \frac{\lambda q^2}{4\pi} \qquad (36)$$

Taking the above formulation, we examined the reconstruction procedures and apply X-ray pulse of wavelength, $\lambda=10\text{Å}$ and $q_{max} \approx 0.3\text{Å}^{-1}$ in the simulated diffraction patterns from STNV pdb file to explore all the calculations for 3D imaging process. Fig 20 describes the different categories of diffraction patterns and how these patterns are to be interpolated considering the curvature of the Ewald Sphere. The number of photons per Shannon Pixel can be counted as one or two that are supposed to hit the outermost circular (polar) grids of ES. Considering the wave length of 10Å, a curved Ewald Sphere is adopted.
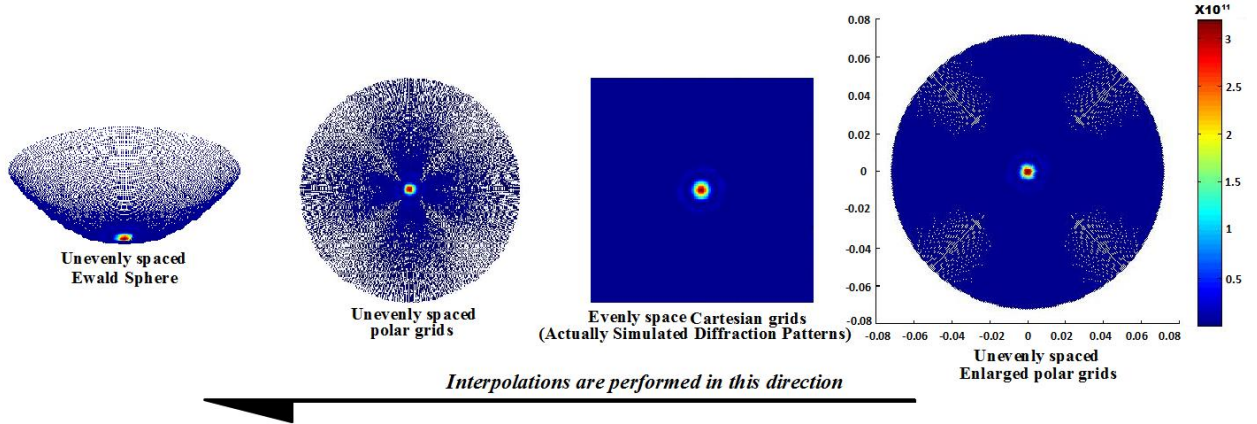
Fig 20. Different category of diffraction patterns in the reciprocal space is shown. From the right, the enlarged simulated diffraction patterns on polar grids are generated and interpolated to intensities on the square grid points which mimic the real experimental DPs based on rectangular detector grids.

4.3 C2, C3 and *Bl*, *Tl* with Sample Model and reconstructed image

In this section, we generated 500 simulated diffraction patterns of STNV (pdb entry:2buk) [41] using the curved Ewald Sphere formulation introduced in the previous section and examined all the calculations from Eq. (31) to Eq. (37) to see if these worked. First, intensities through Eq. (13) and Eq. (31) on the enlarged polar grid points are calculated. The enlarged grids points are not radially uniform. The interpolation from unevenly spaced polar grids to evenly spaced Cartesian grids provides the simulated diffraction patterns can mimic the real experimental DPs. Once interpolated, Cartesian diffraction patterns contain the information of three parameters q_x, q_y and q_z that are important in dealing with a curved Ewald Sphere case. Big samples (particles or molecules) produce generally small scattering angles ζ such that the complement angles θ get closer to $\pi/2$, which leads to $q_z \approx$ 0. However, if the scattering angles are large enough that we cannot ignore the curvature, then the parameter q_z's come into play. Different from the flat Ewald Sphere, obtaining $I(q,\phi)$ on a polar grid with not only q_x and q_y but also q_z with r_j's provided from pdb file as atom coordinates come in handy. Then we could get C2, C3, B_l, T_l by Eq. 22, 23, 24, 25. Fig 21 shows how quantity relations look like.
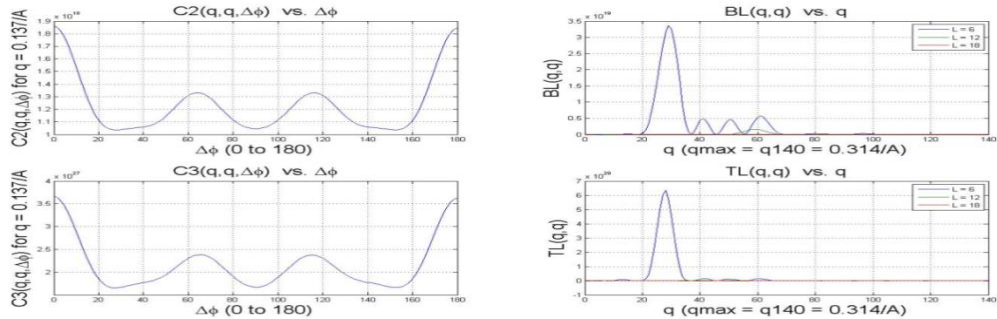
Fig 21. Obtaining intensities on polar grids from Cartesian grids after performing interpolations, one could get C2, C3 through Eq.(22), Eq.(23). As seen in the figure, $B_l(q, q)$ vs. q at $q=l$ /R where l=6, R=100A, the peak appears as we expected.

31

Same as in Fig4, Fig18 describe that at q values corresponding to l=0, 6, 10, 12, 16, 18, 20, 22, 24, 26, 28 through $q = l/R$ give peaks in angular momentum decomposition analysis [9]. It is of interest to know whether one can reconstruct a nearby icosahedral structure by ignoring non-icosahedral l's. If the diffraction volume has the icosahedral symmetry, one would expect the low angular momentum quantum numbers to be primarily $l = 0$ and $l = 6$ as in Fig 4. In our test, we could recognize all other icosahedral symmetry components are relatively small $Bl(q,q)$ values. In the curved Ewald Sphere case, we examined 40 sets of 100 DPs, 200 DPs, 300 DPs, 500 DPs to investigate how these different number of diffraction patterns make 3D icosahedral image after all the calculation process. We found that even 100 simulated DPs will make a good icosahedral 3D image of more than 85% of chances (See Fig 22). Fig 22 shows the ribbon model of STNV (satellite tobacco necrosis virus) provided by pdb (entry#: 2BUK).

Fig 22. The 3D image of Satellite Tobacco Necrosis Virus (STNV) through all the calculations using pdb-file consisting positions of atoms and the distance of each atom from the center.

Our main goal in this section is to reconstruct the 3D icosahedral image along the different number of diffraction patterns and to examine how the small number of diffraction patterns affects imaging processing. We tested first with forty sets of different 500 randomly oriented diffraction patterns, processing all calculation procedures for each set and never failed to get the icosahedral 3D images as in Fig 23. Judging the probability to get the right images according to different number of DP sets is pretty much time consuming labor. The probability of getting the right images of STNV could be obtained by empirical tests. These successful probabilities may vary if we take the different DP sets. The Fig 25 shows the approximate chances to get icosahedral images in the trials with different number of DP sets.
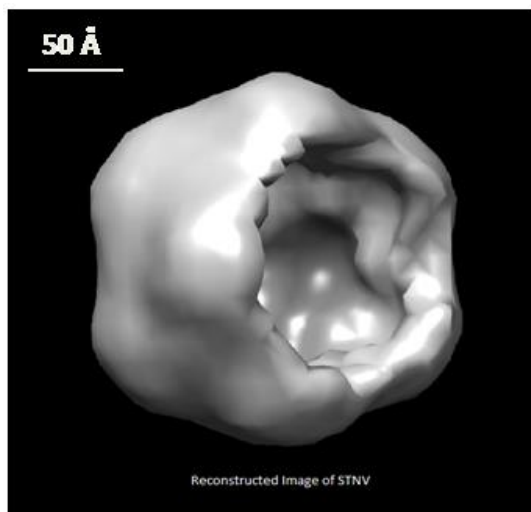
Fig 23. The image of STNV through the all calculations with curved Ewald Sphere parameters. Forty trials with 500 randomly oriented simulated DPs never failed to get the right image. By eliminating a part of the surface from the reconstructed image using the Chimera "Volume Eraser" command the internally hollowed structure appears as expected. Perhaps any internal structure is less than the capsid. $R_f = 0.0432$.

The diffraction patterns for a randomly oriented particle on a curved Ewald Sphere in the reciprocal space tend to generate a little clearer image as in Fig 23. The behavior of intensities on a curved Ewald Sphere fluctuate the same as for the case of those on a flat ES. Fig 24 shows the typical behavior of intensities along difference q values.
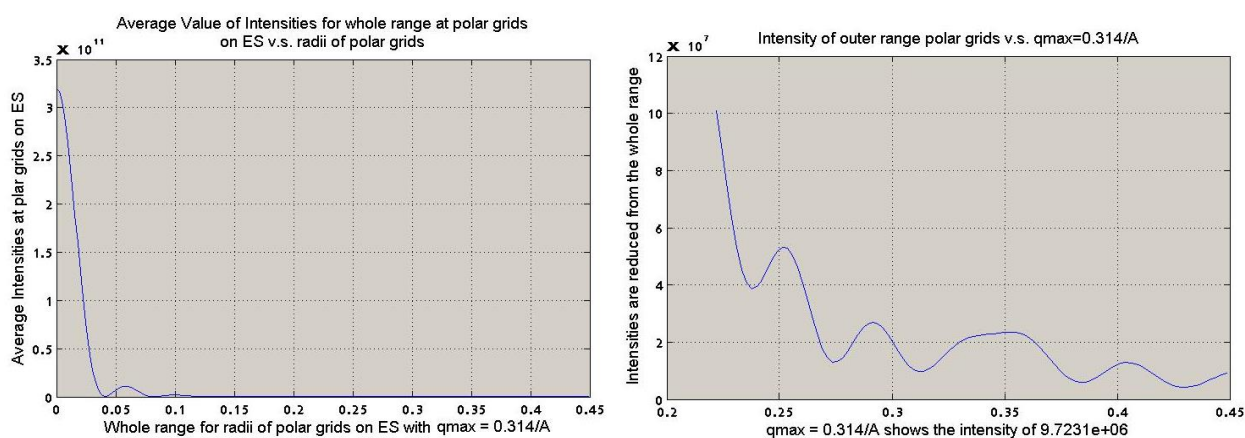
Fig 24. Plot of Intensity vs. q values for whole range of enlarged polar grids where we take only a part while the interpolation is being done according to $q\text{max}=l\text{max}/R$. $q\text{max}\approx0.3Å$, Radius(R)=100Å, $l\text{max}=28$, res $\approx 20Å$.

We made many trials to do all calculation procedures of making images using different number of simulated diffraction patterns that were randomly oriented STNV using curved Ewald Sphere parameters. When we tried with forty 300 DP sets, only once we failed getting a good image out of forty trials. This probability may slightly vary according to DP sets, since it is probabilistic approach. Comparing to flat Ewald Sphere cases, the result illustrated in Fig 25 does not change much. These empirical examinations cautiously lead us to assume that more than 500 simulated DPs generated by randomly oriented particle

seem to be irrelevant in our work. In general, the more DPs of randomly oriented sample can give the better chances of the successful performance to make good reconstructed images. However, due to the icosahedral symmetry, more than randomly oriented simulating 500 DPs seem to be redundant except for some statistically biased cases.
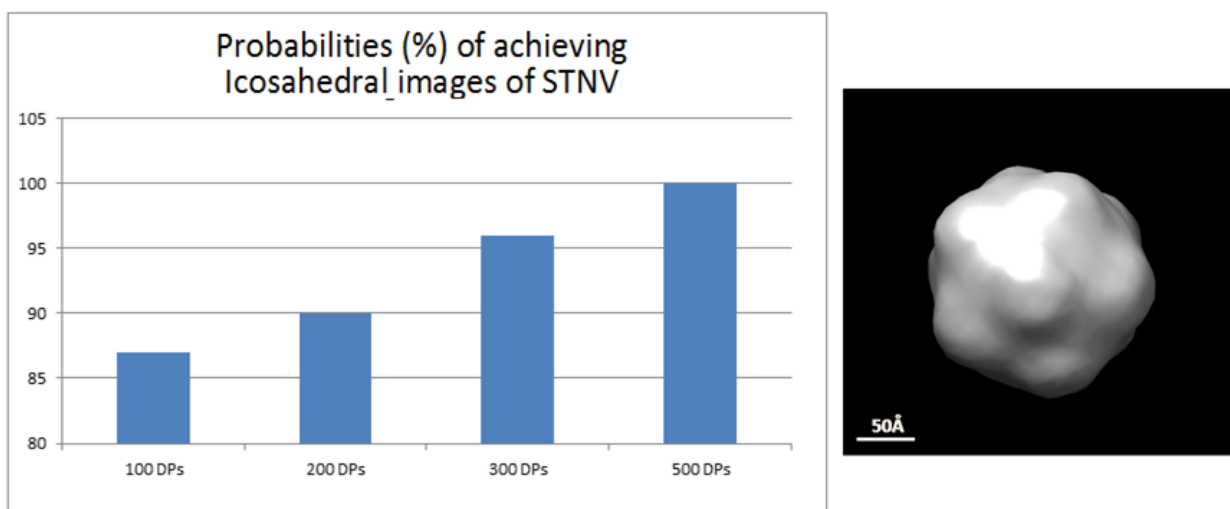
Fig 25. Empirical list for the chances to produce icosahedral images of STNV when all the procedures of calculations with 40 sets for each categorized DP are performed. In this work, curved Ewald Sphere parameters are used. (This statistical data may vary in other trials with different number of DP sets). $R_f = 0.0432$.

As Eq. (36) indicates, curved Ewald Sphere parameters will be applied to a small sample that gives rise to a large q_{max} on the detector in the reciprocal space, and a low incident X-ray photon energy that provides long wavelength λ. Then the q_z values in Eq. (36) are not negligible. Empirically, $\theta = \sin^{-1}(q_z/q) \approx 8°$ seems to be the threshold angle to decide for applying curved Ewald Sphere algorithm in all 3D imaging procedures.

Chapter 5

Shot to Shot Variations of Incident X-ray and Lost Central Intensities

In the experimental equipment set up for XFEL, the strength of an incident X-ray pulse at the center is assumed higher than the edge and a certain variations [42] of the strength in all incident X-ray shots occur. The consequences of these variations can cause some calculation difficulties in image processing. As for the inconsistent shot variations, some adjustments are needed to enable calculation procedures to reach their final goals of 3D imaging process. In this chapter through the simulated 4000 DPs of a randomly oriented and located sample of STNV with the addition of Gaussian variations (4%) to the incident X-ray pulses, a remedy for this problematic issue could be understood. Understanding the deficiency of collecting DPs under the shot to shot variation from Gaussian Incident X-ray pulse can be achievable by using DPs from different region where the sample are placed. A Gaussian pulse does not give same influence in all regions of the locations of sample. We collect DPs from differently separated by σ(standard deviation of the pulse variation). While taking all DPs from the whole region of sample location, the 3D imaging process failed, if only the DPs generated by the sample located within one unit of σ of the Incident Gaussian pulse, the reconstruct 3D image of the sample (STNV) appeared, not as good as a perfect icosahedral shape though. STNV is used as a sample in this simulating procedure of 4000 DPs with the sample size ≈ 200Å, randomly rotated and located in $0 \leq R_C$(polar coordinate for the Center of sample) $\leq$ 600Å , qmax ≈ 0.3Å^{-1}, *res* ≈ 20Å, lmax = 28, incident Gaussian beam diameter = 1000Å, σ = 166Å, energy of the incident beam ≈ 1.4 KeV, photons of incident X-ray pulse ≈ 10^6 photons/Å^2. With these specifications of the simulating process, *C2(q,q',Δϕ)*, *C3(q,q',Δϕ)* which will lead to obtain *B$_l$(q,q')* and

$T_l(q,q')$ via Eq. (24) and (25) are achievable and effective for pursuing 3D reconstructed image of STVN. When we access the real data, experimental DPs always contain the lost central intensities. These lost central intensities in DPs give quite a big problem to work on angular correlation method. In section 2 of this chapter will discuss some remedies for the problem.

5.1 Shot to Shot Variation of Gaussian Incident X-ray pulse on a particle and its remedy

The strength of incident X-ray pulse may vary. In this section, we simulated shot to shot variations that occur when the incident X-ray hits a particle. Considering a circular X-ray pulse shot hits a particle, the particle may be positioned at the center or the edge or somewhere in the middle. Then the influence of the X-ray shot may differ by the location of the particle. The scattered X-ray photons that pass through the particle from the center are counted more than from the edge. This variation gives the number of photons on the detector inconsistently. We provide the remedy to adjust this artifact so that the calculations for reconstructing 3D image of the particle under the shot variation will be available. Fig 26 illustrates the skimming view for the X-ray pulse shot on a particle. When the incident X-ray shot hits the sample as in Fig 26, the influence of the strength of the X-ray affects differently according to the position of the sample.
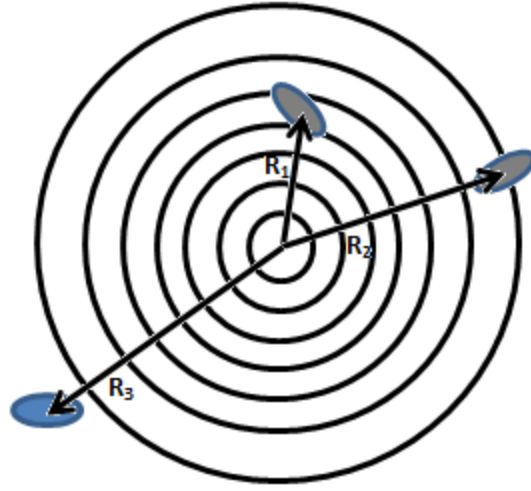
Fig 26. Incident X-ray has a strongest intensity at the center and attenuated as it is away from the center. The intensity follows the Gaussian along the radius of the X-ray pulse front. First sample is under the strong influence of the shot, second sample is weakly influenced and the third sample has almost no influence of the shot.

Each atom has its own radial position and interact the Gaussian X-ray [43], thus the generated amplitude by each atom follows the Eq. (37).

$$A(q) = f(q)\exp(iq \cdot r) \cdot Amp \cdot \exp\left(-\frac{(r+R_C)^2}{2\sigma^2}\right) \tag{37}$$

For convenience, we can take Amplitude of the incident X-ray varies from 1 to 0.98 (2%) and the standard deviation of the ray has σ = 1/3R where R is the radius of the X-ray pulse front considering the 4% of intensity variation of the incident X-ray pulse. All amplitudes coming from each atom are added together to form an intensity on a detector grid in the reciprocal space by squaring the amplitude sum according to the Eq. (13). Here, atomic scattering factor $f(q)$ is defined as

$$f(q) = \sum_{i=1}^{4} a_i \exp\left(-b_i\left(\frac{q}{4\pi}\right)^2\right) + c \tag{38}$$

38

where a_i , b_i , c are Cromer-Mann's coefficients [44] that are specified by each atom in the sample. When generating simulation DPs, the center of the sample STNV, will take up a random position in a circular space of radius $0 \sim 600A^o$ while the X-ray shot influence is committed up to 500Å.
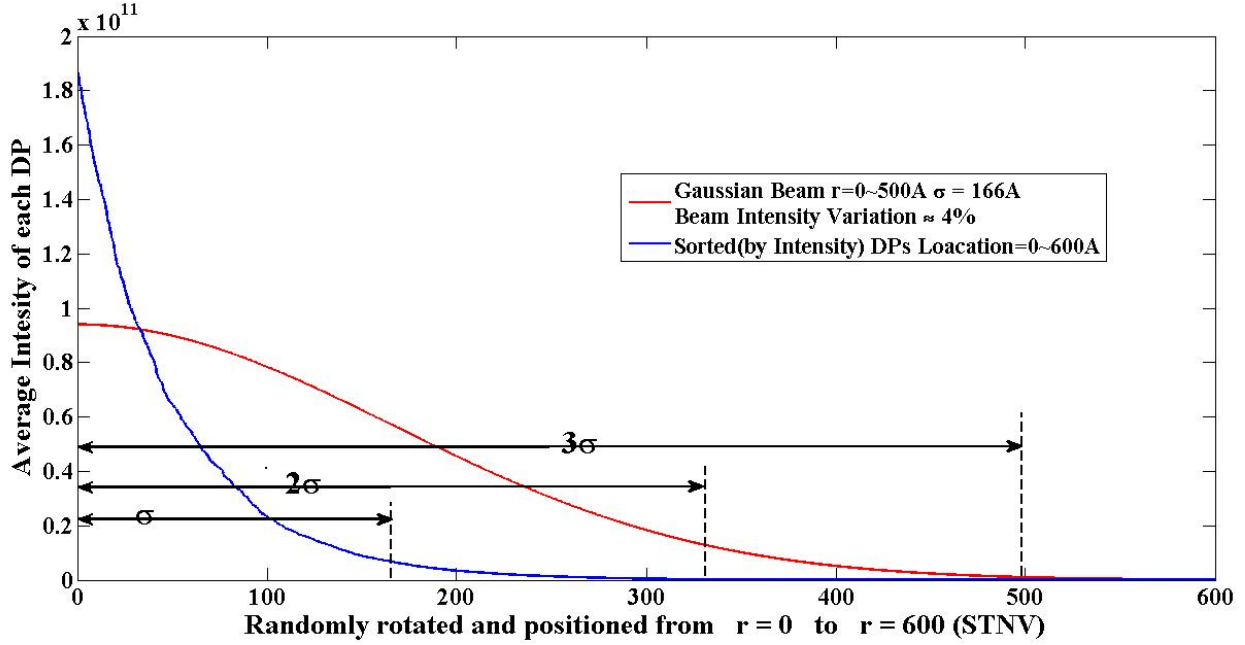
Fig 27. 4000 Cartesian DPs from randomly oriented and located within 0~600Å away from the beam center in each X-ray shot of intensity variations are generated. A STNV sample is supposed to be positioned at one of the region as in Fig 23. The plot shows how many DPs should be discarded to get the right information of the sample in 3D imaging.

Using Eq. (13), (37) and (38), one can generate Cartesian DPs of a randomly oriented sample and interpolate intensities on the Cartesian grid points into polar grid points. Since the intensity of incident X-ray shot varies on each DP as in Fig 27, the shot influence plays a role to generate DPs according to the sample position. Fig 27 shows in our test that at most 7% (280 DPs/4000 DPs) from the location of the sample ($0 \leq R_C \leq 166=1\sigma$) enable

the calculation of $C2(q,q',\Delta\phi)$ and $C3(q,q',\Delta\phi)$ which will lead to obtain $BL(q,q')$ and $TL(q,q')$ using square matrix inversion as in Eq. (12).
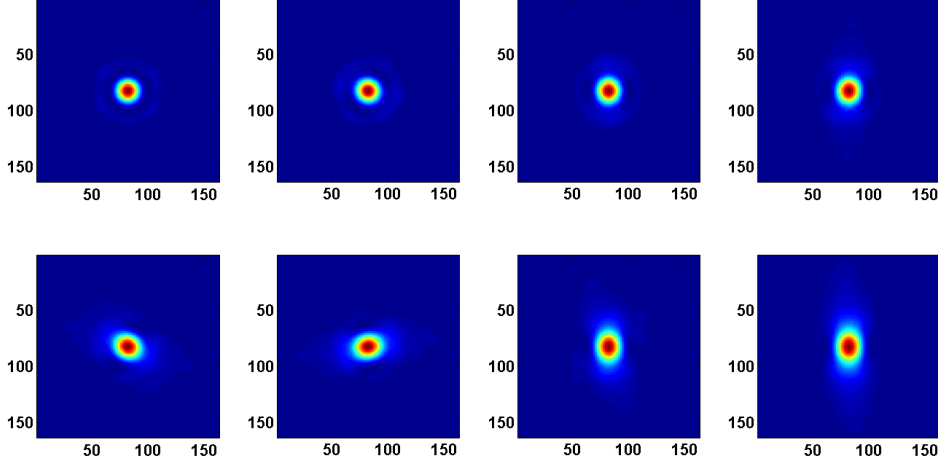
Fig 28. Images of Cartesian DPs by varying Intensity X-ray shot on a sample (STNV) in different region. These 2D images show the diffraction patterns as the center of STNV located gradually away from the X-ray beam center. Most of DPs are not usable except high 7% of all DPs when they are sorted by average intensities.

For the images of STNV DPs from the shots in Region1 generate patterns of strong central brightness and hexagonal fringes. Even though it is possible to select DPs by probability under the shots onto the region of $\leq 3\sigma$, visional differentiation has many limits that can cause quite a sizable calculation errors in each step. This is the reason why the sorted average intensity level at a certain low q should be checked to distinguish the DPs under the shots of X-ray onto the region ($\leq \sigma$) from ones under the shots onto the sample located away from the center by farther than one unit of σ. Among about 4000 DPs, 280 DPs of them are selected considering safe functioning of angular correlation method. These selected Cartesian DPs are interpolated to polar DPs. Under the influence of no incident X-ray shot variations, the intensities on polar grid are the same at low q. However if there are

Gaussian X-ray intensity variations are applied, even at a low q position on polar grids, the intensity levels are not consistent along the different angular positions shown in Table1. Table2 shows that at low q_1 ~ q_7, the intensities along the different angular positions appeared the same irrespective of the orientation of the sample. At least in 2D diffraction patterns, it is clear [45] that there exist quantities independent of the orientation of the particle. Thus under the incident X-ray shot variations, the average intensities in polar DPs can be equalized to access the angular correlation method just like the uniform strength of incident X-ray shots. By the reasons mentioned above, DPs generated by the sample in the region ($\geq$ σ) in Fig 27 where only very weak influence of shots performed by varying Gaussian pulse will be discarded.

					Angular Position on Polar Grids (0° ~ 359°)						
	196°	197°	198°	199°	200°	201°	202°	203°	204°	205°	206°
DP1 q_7	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11	2.1761e+11
q_6	1.8830e+11	1.8830e+11	1.8830e+11	1.8830e+11	1.8830e+11	1.8829e+11	1.8829e+11	1.8829e+11	1.8829e+11	1.8829e+11	1.8829e+11
q_5	1.5849e+11	1.5849e+11	1.5849e+11	1.5849e+11	1.5849e+11	1.5848e+11	1.5848e+11	1.5848e+11	1.5848e+11	1.5848e+11	1.5848e+11
DP2 q_7	4.7163e+10	4.7323e+10	4.7480e+10	4.7633e+10	4.7781e+10	4.7925e+10	4.8064e+10	4.8199e+10	4.8329e+10	4.8453e+10	4.8686e+10
q_6	4.3459e+10	4.3662e+10	4.3861e+10	4.4054e+10	4.4243e+10	4.4426e+10	4.4603e+10	4.4774e+10	4.4939e+10	4.5098e+10	4.5250e+10
q_5	3.9496e+10	3.9740e+10	3.9978e+10	4.0211e+10	4.0438e+10	4.0658e+10	4.0872e+10	4.1079e+10	4.1278e+10	4.1470e+10	4.1653e+10

Table 1. Polar diffraction patterns containing intensities generated on polar grid points under varying intensities of incident X-ray shots are illustrated. From the above table, DP1 from uniform incident X-ray shot shows no intensity varying at q_7. DP2 shows varying intensities even at q_7 as the angular position changes.

Fig 27 shows us a hint how to adjust different average intensity levels proportionally at q_7 for each DP. By equalizing the average intensity levels at a designated q could be a future help to overcome the problem of DPs containing lost central intensities. Once all polar DPs are synchronized by proportional adjustment, it is ready to obtain $C2(q,q',\Delta\phi)$ and $C3(q,q',\Delta\phi)$ using Eq. (22) and (23) proposed by Z. Kam [38] as a cross correlation function, which is equivalent to Eq. (5). $Bl(q,q')$ and $T_l(q,q')$ are independent of angular positions and can obtained from the Eq. (24) and (25) which will lead to get the diffraction

volume in reciprocal space. The obtained diffraction volume, whose elements are intensities in reciprocal space, will be employed to get the charge densities in real space by iterative phasing algorithm [30][31]. As illustrated in Eq. (26) ~ (30), $I_{lm}(q)$ are known up to signs. Because of icosahedral symmetry, only the following l values are allowed for l_{max} = 28 with l = 0, 6, 10, 12, 16, 18, 20, 22, 24, 26, 28. Since L has eleven values for icosahedral symmetry, there are only 2^{11} sign combinations exhaustively performed to get the best-fit combinations by comparing theoretical and experimental $T(q,q')$'s. To understand how this method works, simulated experimental data for B and T were obtained directly from PDB file of STNV (entry#: 2BUK). The combination of sings which gives the best-fit between the theoretical and experimental diagonal elements of T's for a given reference shell was found. Then the signs can propagate to other shells by the non-diagonal $B(q,q')$. Hence with $I_{lm}(q)$, the diffraction volume can be calculated. The iterative charge flipping method is used for phasing the diffraction volume and gives the charge densities for STNV and the icosahedral image of STNV is shown up as in Fig 17.

| | Angular Position on Polar Grids (0° ~ 359°) |
	196°	197°	198°	199°	200°	201°	202°	203°	204°	205°	206°
q_0	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11	3.1972e+11
q_1	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11	3.1647e+11
q_2	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11	3.0687e+11
q_3	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11	2.9141e+11
q_4	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11	2.7084e+11
q_5	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11	2.4613e+11
q_6	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11	2.1844e+11
q_7	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11	1.8901e+11
q_8	1.5907e+11	1.5907e+11	1.5907e+11	1.5907e+11	1.5907e+11	1.5907e+11	1.5906e+11	1.5906e+11	1.5906e+11	1.5906e+11	1.5906e+11
q_9	1.2983e+11	1.2982e+11	1.2982e+11	1.2982e+11	1.2982e+11	1.2982e+11	1.2981e+11	1.2981e+11	1.2981e+11	1.2981e+11	1.2981e+11

Table 2. Intensities on polar grid points interpolated from a Cartesian DP shows that from q_0 up to $q_7 \approx 0.016\text{Å}^{-1} \approx 0.05q_{max}$ where $q_{max} = q_{140} \approx 0.314\text{Å}^{-1}$ the intensity variations do not appear along the different angular positions where uniform strength of incident X-ray shots are performed onto the sample STNV.
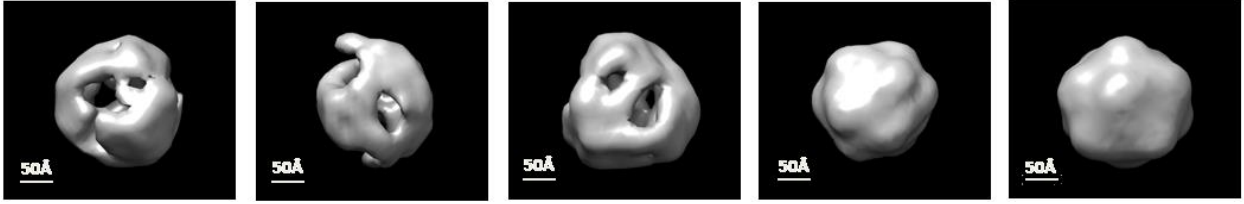
Fig 29. These reconstructed images from Gaussian Intensity varying X-ray shots referring to Fig 24. From the left, the first image is from all 4000 DPs ($0 \leq$ location of sample ≤ 600Å). The second image is from 2720 DPs (68%, locations of sample $\leq 3\sigma$), the third image is from 1200 DPs (30%, locations of sample $\leq 2\sigma$), the fourth image is from 280 DPs (7%, locations of sample $\leq \sigma$). The last image is from 200 simulated DPs with no varying incident X-ray. From the left, $R_f = 0.0576$, 0.0598, 0.0532, 0.0414, 0.0419.

5.2 How to overcome the lost central intensities in DPs and the remedies for simulated data under shot to shot incident X-ray intensity variations.

In reality, the experimental DPs containing lost central intensities are observed in detector grid points. Fig 30 shows how the real DP data look like and supposed to be overcome in each calculation procedures. The left figure in Fig 30 is the one that the central intensities are eliminated artificially from the generated DPs using PDB file of PBCV1 (pdb entry#:1M4X) [46] and the right one is from an experimental PBCV1 DP. As one can see central part of DPs are eliminated to let the scattered X-ray pass through to protect detectors from the strong influence of the rays. This can cause some difficulties in performing interpolations, getting $C2$, $C3$ and B_l, T_l and hereafter up to find reconstructing 3D images. To overcome these lost information and shot to shot variation, simulated data with eliminated central intensities and shot to shot variations are applied to develop the algorithm to help the calculation procedures go on. The black line in the figure indicates the connecting part of detectors where the scattering rays are not recorded. By the

advantage of $C2(q,q',\Delta\phi)$ and $C3(q,q',\Delta\phi)$ through point by point calculation, this problematic issue can be resolved easily.
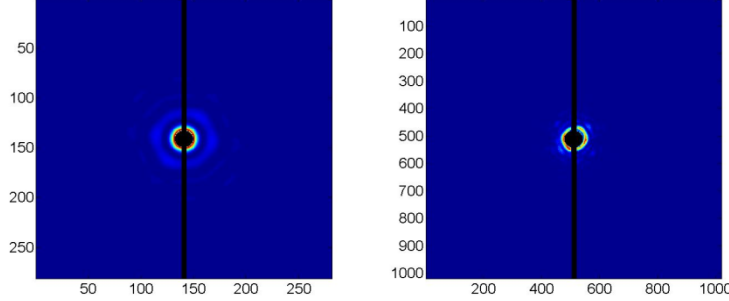
Fig 30. Two figures show the lost part of the intensities in DPs of PBCV-1. The left is a simulated DP with artificially removed intensities to resemble the one on the right to overcome the issue of the lost intensities. Also the shot to shot incident X-ray variations are applied here.

However, the shot to shot variations of incident X-ray strength should be considered to pass through all calculation procedures to achieve the goal of 3D imaging. At this point, so called Intensity fitting using Atomic form factor [47][48], which is scattering amplitude, sometimes we refer to Rayleigh fitting, should be adopted to extrapolate the lost central intensities into the beam stop of all DPs. The reason why the extrapolations are needed into the beam stops is that as Fig 31, even though there are lost radial central intensities at low q, from q_0 up to q_6 in experimental DPs, one could easily observe that imaging process is still successful. For example, where $q_{max} = q_{140}$, Fig 31 indicates that if we lost intensities at $q_0 \sim q_9$, then the polar intensities at q_9, q_8 and q_7 should be resumed by some appropriate method.
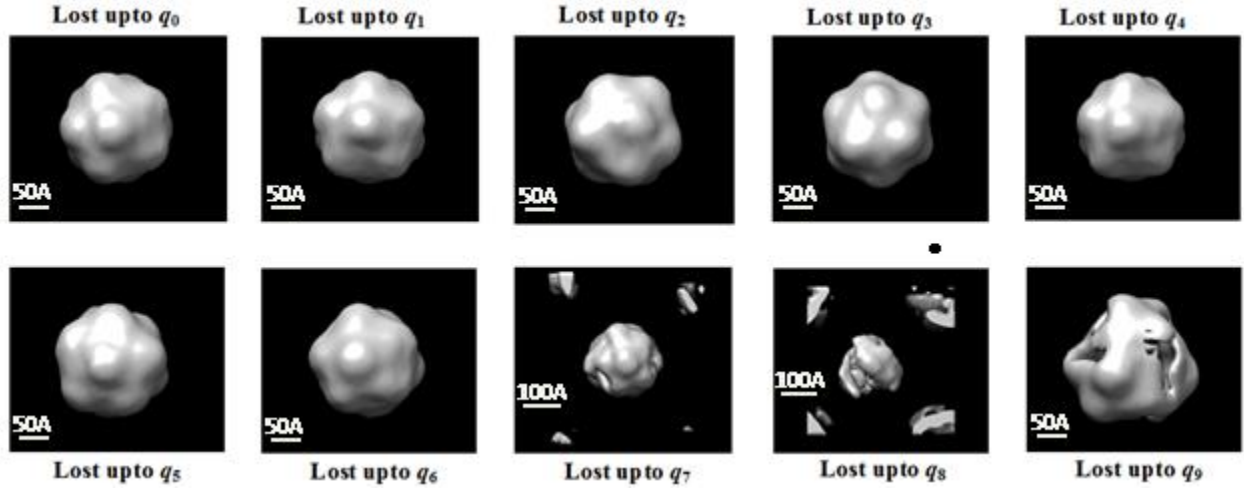
Fig 31. Through STNV 3D imaging process with 500 DPs, where $q_{max} = q_{140} \approx$ 0.3Å^{-1}, if the radial intensities at $q_0 \sim q_8$, q_9 and the more are lost, then the 3D imaging process fail for icosahedral viruses. Thus, the extrapolation down into q_7, which is within 5% of q_{max}, is needed for the 3D imaging process to be successful. From the top left, R_f = 0.0440, 0.0778, 0.0458, 0.0789, 0.0453, 0.0786, 0.0463, 0.0797, 0.0775, 0.0984.

Now, a fitting method, sometimes we refer to analytical fitting, using the scattering amplitude is introduced as the following to resume the lost central intensities by the extrapolation into the beam stops. Amplitude in the reciprocal space can be written as

$$A(q) = \int \rho(r) \exp(iq \cdot r) d^3 r \qquad (39)$$

Considering 99.9% of atoms in STNV are C, N and O, for convenience we can take $\rho(r) \approx$ average charge density ρ_o, then the Eq. (39) under the spherical coordinates will be

$$A(q) = \rho_0 \int \exp(iq \cdot r) r^2 \sin\theta \, dr \, d\theta \, d\phi \qquad (40)$$

with $x = \cos\theta$ *and* $dx = -\sin\theta \, d\theta$, (40) will lead to

$$A(q) = \rho_0 \int_0^{2\pi} d\phi \int_{-1}^{1} \exp(iqrx) \, dx \int r^2 dr = 2\pi\rho_0 \int \left[\frac{\exp(iqrx)}{iqr} \right]_{-1}^{1} r^2 dr$$

$$= 2\pi\rho_0 \int \left[\frac{e^{iqr} - e^{-iqr}}{iqr} \right] r^2 \, dr = 2\pi\rho_0 \left\{ \frac{1}{iq} \left[\int re^{iqr} dr - \int re^{-iqr} dr \right] \right\}$$

$$= \frac{2\pi\rho_0}{iq} \left[\int re^{qr} dr - \int re^{-iqr} dr \right]. \tag{41}$$

Using the Integral Table, $\int re^{\alpha r} dr = \frac{e^{\alpha r}}{\alpha^2} (\alpha r - 1)$, Eq.(41) becomes to

$$A(q) = \frac{2\pi\rho}{iq(-q^2)} \left[e^{iqr}(iqr - 1) - e^{-iqr}(-iqr - 1) \right]$$

$$= \frac{2\pi\rho_0}{-iq^3} \left[iqr(e^{iqr} + e^{-iqr}) - (e^{iqr} - e^{-iqr}) \right]$$

$$= -\frac{2\pi\rho_0 r}{q^2} \left[e^{iqr} + e^{-iqr} \right] + \frac{2\pi\rho_0}{iq^3} \left[e^{iqr} - e^{-iqr} \right]$$

$$= -\frac{2\pi\rho_0 r}{q^2} [2\cos(qr)] + \frac{2\pi\rho_0}{iq^3} [2i\sin(qr)]$$

$$= -\frac{4\pi\rho_0}{q^3} qr\cos(qr) + \frac{4\pi\rho_0}{q^3} \sin(qr). \tag{42}$$

Therefore, the resuming amplitudes in DPs for the lost central intensities into beam stops would follow the scattering amplitude is written as

$$A(q) = \frac{4\pi\rho_0}{q^3} [\sin(qr) - qr\cos(qr)]. \tag{43}$$

Fig 32 shows how the lost central intensities are extrapolated by the scattering intensity fitting down to the expected low q-value. The blue and red line in the plot shows that the intensity fitting resumes the central lost intensities with quite a good precision (3~4%) comparing to the whole range of intensities, which leads to a successful 3D imaging process as in Fig 31.
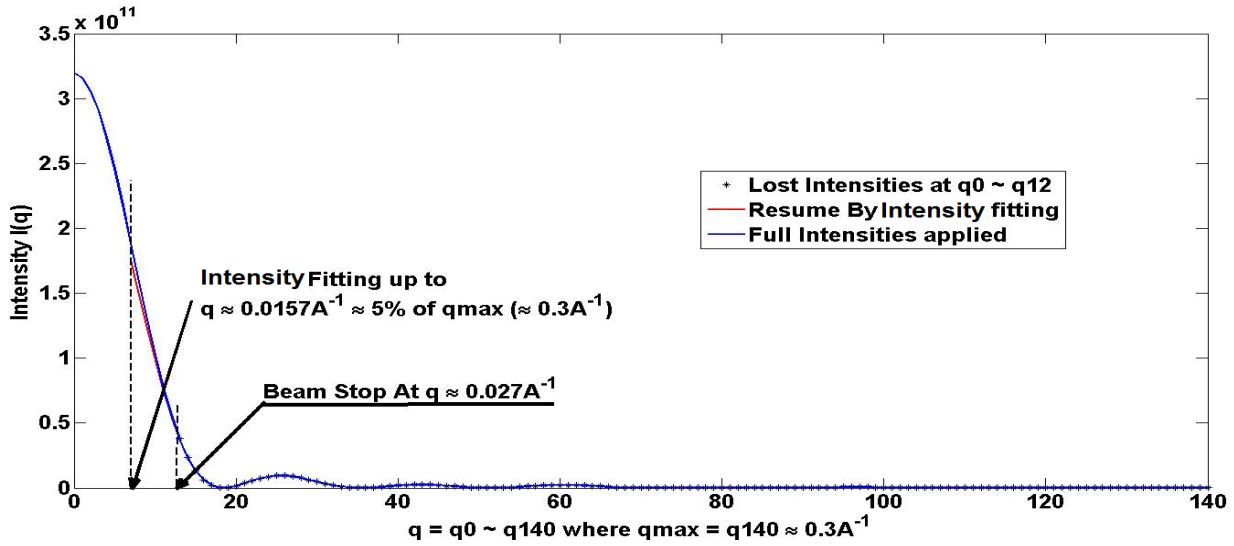
Fig 32. Three intensity plots along the angular position at 47° from a polar diffraction pattern vs. whole range of q values is illustrated. Blue line shows the intensities at all q values. Red line shows the extrapolations down to $q \approx 0.016\text{Å}^{-1}$. The *'s indicate the lost central intensities from the center ($q = 0$) to $q = 0.027\text{Å}^{-1}$.

Intensity fitting as in Eq. (43) is used to extrapolate the lost central intensities around the center of polar DPs so that the calculations for $C2$, $C3$, B_l, T_l are possible to make 3D imaging process successful. Fig 31 and Fig 32 give a hint how to handle and what portion of lost intensities from the whole range in experimental DPs are allowed to get 3D imaging calculation procedures working well. Table 3 shows the comparison of intensities on a polar DP grid points with three categories in Fig 32.

47

Lost $q_0 \sim q_{12}$

Trace $q_{12} \rightarrow q_7$

No Lost

(At angle 46° in a polar diffraction pattern where $q_{max} = q_{140}$)

	Lost $q_0 \sim q_{12}$	Resume $q_{12} \sim q_7$	No Lost
q_0	0	0	3.1972e+11
q_1	0	0	3.1570e+11
q_2	0	0	3.0577e+11
q_3	0	0	2.9038e+11
q_4	0	0	2.7020e+11
q_5	0	0	2.4594e+11
q_6	0	0	2.1812e+11
q_7	0	1.7573e+11	1.8660e+11
q_8	0	1.4901e+11	1.5881e+11
q_9	0	1.2293e+11	1.2975e+11
q_{10}	0	9.8406e+10	1.0233e+11
q_{11}	0	7.6163e+10	7.7472e+10
q_{12}	0	5.6714e+10	5.5914e+10
q_{13}	3.7993e+10	3.7993e+10	3.7993e+10
q_{14}	2.3764e+10	2.3764e+10	2.3764e+10
q_{15}	1.3230e+10	1.3230e+10	1.3230e+10
q_{16}	6.3150e+09	6.3150e+09	6.3150e+09
q_{17}	1.9615e+09	1.9615e+09	1.9615e+09
q_{18}	2.6518e+08	2.6518e+08	2.6518e+08
q_{19}	3.2078e+08	3.2078e+08	3.2078e+08
q_{20}	1.5989e+09	1.5989e+09	1.5989e+09
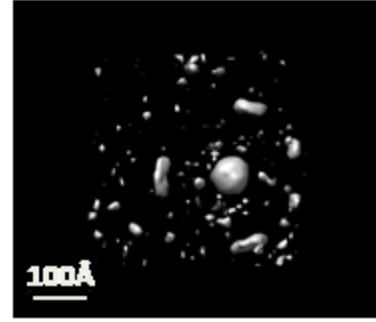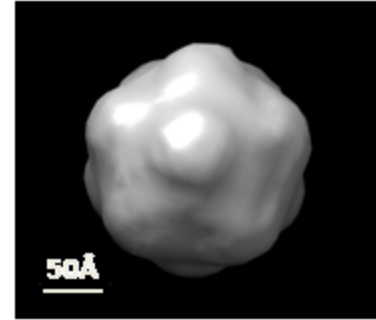
Table 3. First column shows lost intensities up to $q_{12} \approx 0.016\text{Å}^{-1}$. Second column shows six intensities at $q_{12} \sim q_7$ are extrapolated by the Intensity fitting. Third column shows no lost intensities. The images from the top to bottom on the right side correspond to 500 DPs of first column, second column and third column respectively. From the top, $R_f = 0.0958$, 0.0510, 0.0440.

In this case, three category of $C2(q,q',\Delta\phi)$, $C3(q,q',\Delta\phi)$, $B_l(q,q')$, $T_l(q,q')$ will coincide with each other over the $q \approx 0.016\text{Å}^{-1}$. Nevertheless, if DPs lost more than certain portion of central intensities, then those DPs will not produce nicely shaped 3D images as in Fig 31 and Table 3.

Chapter 6

Multi Particle Scattering Test By Positioning two nanorice particles in a Cylindrical Space

When the injector in an experiment in reality ejects aerosol [49] containing nano sized objects, these sample objects are supposed to get a shot of the incident X-ray pulse with the duration of a few femtoseconds. Then the illuminated region in the aerosol will form a cylindrical shape in which nano particles such as viruses, proteins, or other objects are placed with random orientations. In spite of all background checks much like the influences of solvent, multi particle interference, intensity variations of incident X-ray pulses, noise, chemical affinity between particles, dipole moments of each particle, etc. Here we examine how multiple particles affect performing the calculation procedures to achieve the reconstruction of 3D images of samples, and how different multi particle scatterings from single particle ones. First, we placed two ellipsoidal shapes of nanorice particles whose short radii are about 8Å and long radii are about 25Å in the cylindrical space of which the radius is about 1000Å and length of 3000Å [50] that resembles the illuminated region in the aerosol by incident X-ray pulses (See Fig 33). And then generating 8000 simulated DPs for two randomly oriented nanorice particles are processed using Eq. (13). Fig 33 shows the simple diagram of equipment set up in performing an X-ray pulse shot on to two nanorice partices in the solvent. We are going to start developing a simple theory of how to form intensities on the detector grid points through Eq. (13) and Eq. (32) on a flat ES with $q_z = 0$. Fig 34 shows one nanorice particle as a sample model having azimuthal symmetry. We randomly rotated and placed these two particles in the illuminated area, and examined if the algorithm that we developed works properly for obtaining the 3D reconstructed structural image.
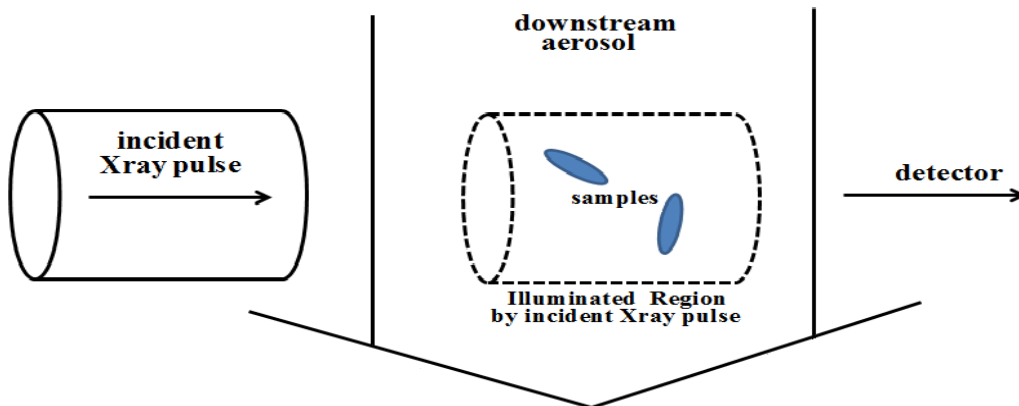
Fig 33. When an incident X-ray shot pulse moves from the right to the left, shortly after it will illuminate a cylindrical space in the sheet of downstream solvent containing two nanorice particles and scattered into the detector and form intensities on the detector grid points. Cartesian Diffraction Patterns are generated by the scattered X-rays from the two nanorice particles (No overlaps).

2D : projected onto xy-plane

2D : projected onto xz-plane

2D : projected onto yz-plane

3D One Nanorice Model(ellipsoid)
LongRadius=25A, ShortRadius=8A
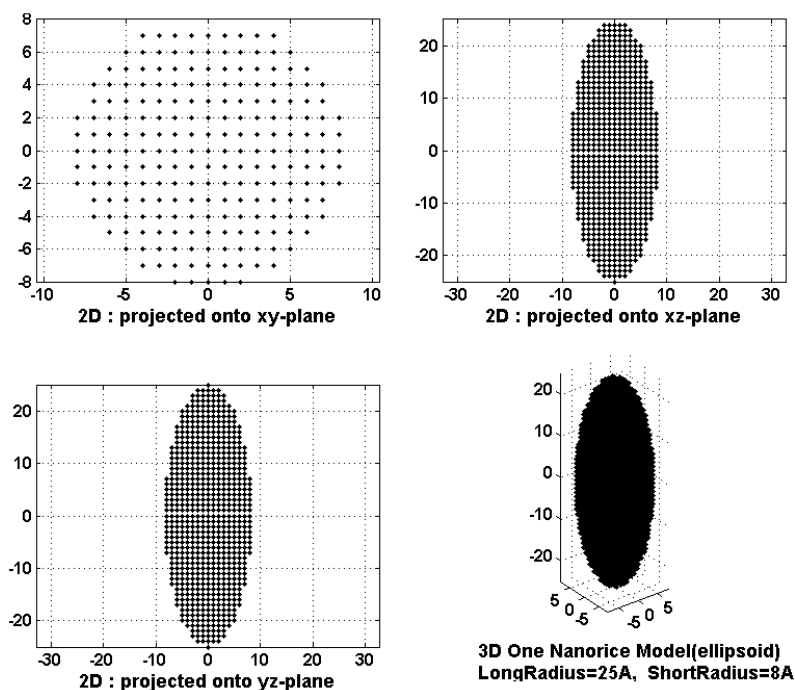
Fig 34. One nanorice particle as a sample model with the azimuthal symmetry is adopted to examine if the algorithm for achieving the 3D reconstructed structural image works properly.

6.1 Calculate $C2\ (q,q',\Delta\phi)$ from adding intensities on polar DPs generated by randomly oriented two particles (Incoherent).

For single particle, Eq. (13) enables us to calculate an intensity on a polar grid points when the center of the sample moved by **R**. Assuming the sample "a" contains two atoms (1, 2) as in Fig 31, then the amplitude at q becomes

$$I_a(q) = A_a(q)A_a^*(q) = \left(\sum_{j=1}^{2} f_j(q)\exp[iq\cdot(r_j + R)]\right)\left(\sum_{k=1}^{2} f_k(q)\exp[-iq\cdot(r_k + R)]\right)$$

$$= \left(f_1\exp[iq\cdot(r_1 + R)] + f_2\exp[iq\cdot(r_2 + R)]\right)\left(f_1\exp[-iq\cdot(r_1 + R)] + f_2\exp[-iq\cdot(r_2 + R)]\right)$$

$$= f_1^2 + f_2^2 + f_1 f_2\left(\exp[iq\cdot(r_2 - r_1)] + \exp[-iq\cdot(r_2 - r_1)]\right)$$

$$= f_1^2 + f_2^2 + 2f_1 f_2\cos[q\cdot(r_2 - r_1)]. \tag{44}$$

Here $I(q)$ is the same as the one whose center is moved. Even though for the sample containing many atoms $(r_2 - r_1)$ will be possible pairs of all atom-positions and independent of positions of the centers. Now, let us adopt one more same sample "b" containing same two atoms (3, 4). Fig 35 shows the diagram for different procedures of $C2$ calculation between adding intensities and adding amplitudes [51].
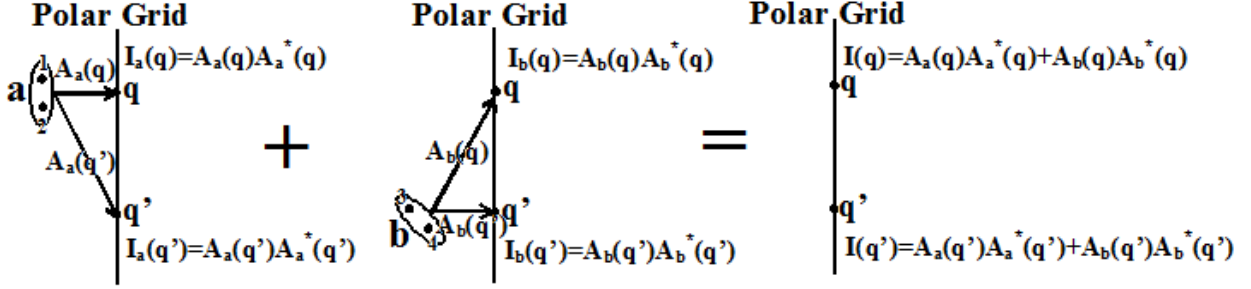
Fig 35. Diagram of adding intensities (incoherent) used in calculating $C2(q,q'\Delta\phi)$ is illustrated. The average of products $< I(q)I(q')>$ becomes one of $C2$ elements.

By the same way, using Eq. (44),

$$I_b(q) = A_b(q)A_b^*(q) = f_1^2(q)f_2^2(q) + 2f_1(q)f_2(q)\cos[q \cdot (r_4 - r_3)]$$

$$= f_1^2 f_2^2 + 2f_1 f_2 \cos[q \cdot (r_4 - r_3)]. \tag{45}$$

$$I_a(q') = A_a(q')A_a^*(q') = f'^2_1 f'^2_2 + 2f_1' f_2' \cos[q' \cdot (r_2 - r_1)]$$

$$I_b(q') = A_b(q')A_b^*(q') = f'^2_1 f'^2_2 + 2f_1' f_2' \cos[q' \cdot (r_4 - r_3)]. \tag{46}$$

Thus,

$$I(q) = I_a(q) + I_b(q) = A_a(q)A_a^*(q) + A_b(q)A_b^*(q)$$

$$\tag{47}$$

$$I(q') = I_a(q') + I_b(q') = A_a(q')A_a^*(q') + A_b(q')A_b^*(q')$$

From (46) and (47), $C2\ (q,q',\Delta\phi)$ can be written as

$$C2(q,q',\Delta\phi) = < \left(I_a(q) + I_b(q)\right)\left(I_a(q') + I_b(q')\right) > \tag{48}$$

If we forget average bracket for a moment, then (44) will be

52

$$2\left(f_1^2 + f_2^2 + f_1 f_2 \{\cos[q \cdot (r_2 - r_1)] + \cos[q \cdot (r_4 - r_3)]\}\right)$$

$$(49)$$

$$\times 2\left(f'^2_1 + f'^2_2 + f_1' f_2' \{\cos[q' \cdot (r_2 - r_1)] + \cos[q' \cdot (r_4 - r_3)]\}\right)$$

6.2 Calculate $C2$ $(q,q',\Delta\phi)$ from adding amplitudes on polar DPs generated by randomly oriented two particles (Coherent).

Adding amplitudes would be a bit complicated in coherent case. Fig 36 shows the diagram that illustrates adding amplitudes at different q.
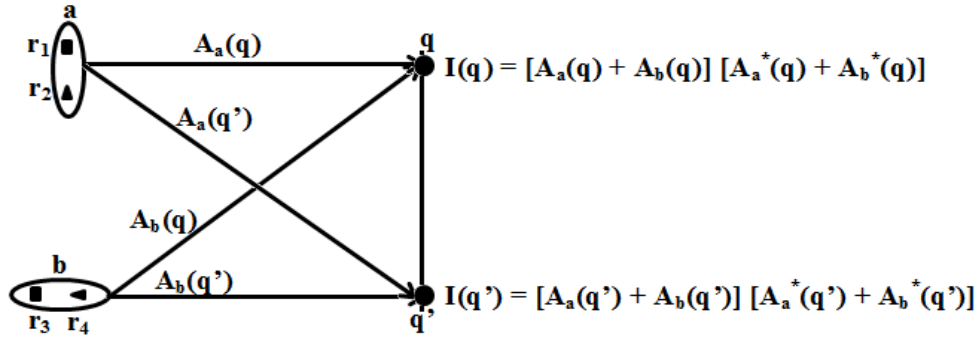
Fig 36. Diagram of adding amplitudes (Coherent) used in calculating $C2(q,q'\,\Delta\phi)$ is illustrated. The average of products $< I(q)I(q')>$ becomes one of $C2$ elements.

$$C2(q,q',\Delta\phi) = < I(q)I(q') >$$

$$(50)$$

$$= \left([A_a(q) + A_b(q)][A_a^*(q) + A_b^*(q)]\right)\left([A_a(q') + A_b(q')][A_a^*(q') + A_b^*(q)]\right)$$

The multiplication of conjugate amplitudes in Eq. (50) becomes

53

$$I(q) = [f_1 \exp(iq \cdot r_1) + f_2 \exp(iq \cdot r_2) + f_1 \exp(iq \cdot r_3) + f_2 \exp(iq \cdot r_4)]$$

$$\times [f_1 \exp(-iq \cdot r_1) + f_2 \exp(-iq \cdot r_2) + f_1 \exp(-iq \cdot r_3) + f_2 \exp(-iq \cdot r_4)]$$

$$= 2 \left(\begin{array}{l} f_1^2 + f_2^2 + f_1 f_2 \cos[q \cdot (r_2 - r_1)] + f_1 f_2 \cos[q \cdot (r_4 - r_1)] + f_1^2 \cos[q \cdot (r_3 - r_1)] \\ + f_1 f_2 \cos[q \cdot (r_3 - r_2)] + f_2^2 \cos[q \cdot (r_4 - r_2)] + f_1 f_2 \cos[q \cdot (r_4 - r_3)] \end{array} \right)$$

$$= 2 \left(\begin{array}{l} f_1^2 + f_2^2 + f_1^2 \cos[q \cdot (r_3 - r_1)] + f_2^2 \cos[q \cdot (r_4 - r_2)] \\ + f_1 f_2 \left[\cos[q \cdot (r_2 - r_1)] + \cos[q \cdot (r_4 - r_1)] + \cos[q \cdot (r_3 - r_2)] + \cos[q \cdot (r_4 - r_3)] \right] \end{array} \right) \quad (51)$$

By the same way,

$$I(q') = 2 \left(\begin{array}{l} f_1'^2 + f_2'^2 + f_1'^2 \cos[q' \cdot (r_3 - r_1)] + f_2'^2 \cos[q' \cdot (r_4 - r_2)] \\ + f_1' f_2' \left[\cos[q' \cdot (r_2 - r_1)] + \cos[q' \cdot (r_4 - r_1)] + \cos[q' \cdot (r_3 - r_2)] + \cos[q' \cdot (r_4 - r_3)] \right] \end{array} \right) \quad (52)$$

If we put the subscript "C" for Amplitude Addition (Coherent), "I" for Intensity Addition (Incoherent) and $\delta I(q)$ for the difference of intensities between $I_C(q)$ and $I_I(q)$.

$$I_C(q) = I_I(q) + \delta I(q) \quad \text{and} \quad I_C(q') = I_I(q') + \delta I(q') \quad (53)$$

Where

$$\delta I(q) = f_1^2 \cos[q \cdot (r_3 - r_1)] + f_2^2 \cos[q \cdot (r_4 - r_2)]$$

$$+ f_1 f_2 \left(\cos[q \cdot (r_4 - r_1)] + \cos[q \cdot (r_3 - r_2)] \right),$$

$$\delta I(q') = f_1'^2 \cos[q' \cdot (r_3 - r_1)] + f_2'^2 \cos[q' \cdot (r_4 - r_2)]$$

$$+ f_1' f_2' \left(\cos[q' \cdot (r_4 - r_1)] + \cos[q' \cdot (r_3 - r_2)] \right). \quad (54)$$

It is easily observed that these are all cosine functions so that as we take many DPs, $\delta I_C(q)$ and $\delta I_C(q')$ will internally cancel each other. That is,

$$I_C(q)I_C(q') = I_I(q)I_I(q') + I_I(q)\delta I(q') + I_I(q')\delta I(q) + \delta I(q)\delta I(q')$$

gives rise to the same as

$$C_C(q,q',\Delta\phi) = C_I(q,q',\Delta\phi) + <I_I(q)\delta I(q') + I_I(q')\delta I(q) + \delta I(q)\delta I(q')> \qquad (55)$$

In Eq. (55), the second, third and fourth terms will disappear at $q = q'$ when enough DPs are taken since the positions (r_j) give many combinations cosine values ($\pm$). However, only these terms will not disappear even at $q = q'$ by canceling each other only if $\Delta\phi = 0^o$ and $180^o(\pi)$ since $\delta I(q)\delta I(q') \rightarrow [\delta I(q)]^2$ is just only one value squared comparing to at $\Delta\phi \neq 0^o$. $\delta I(q)\delta I(q')$ is the product of two values even at $q = q'$. For $iq \cdot r = iqr\cos\theta$, θ changes even though at the same q (radial) and r (atom positions). For example, on a polar grid at same q and r, let us assume there are two intensities $I_1(q,r)$ and $I_2(q,r)$. Then,

$$C2(q,q,\Delta\phi \neq 0) = I_1(q,r)I_2(q,r) \quad \text{and} \quad C2(q,q,\Delta\phi = 0) = \frac{[I_1(q,r)]^2 + [I_2(q,r)]^2}{2} \qquad (56)$$

This result gives us $C_2(q,q,\Delta\phi = 0) > C_2(q,q,\Delta\phi \neq 0)$. The left hand side of Eq. (56) can be $\pm$ values while right hand side of Eq. (56) is always positive. The consequence indicates that $C2$ from intensity addition (Incoherent) is almost the same as $C2$ from amplitude addition (Coherent) except at $\Delta\phi = 0$, π and 2π as in Fig 38. Thus, in this case, $q \cdot r = qr\cos\theta$ and θ has an important role.
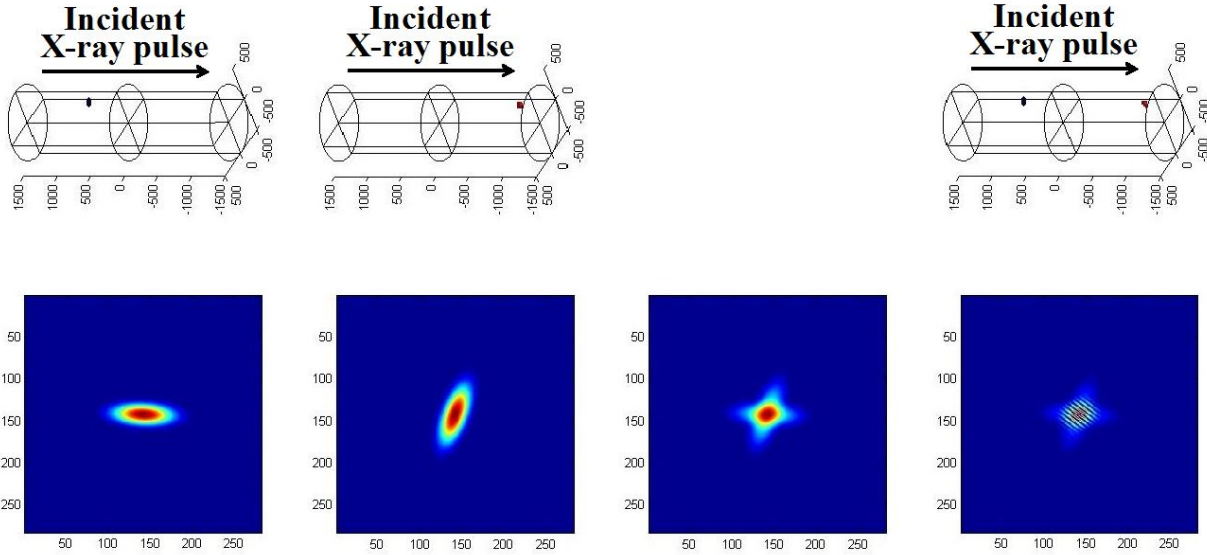
Fig 37. Assuming the experimental set up on above, putting in two nanorices in a cylindrical space illuminated in the aerosol by the incident X-ray pulse with the diameter of 1000A° and the length of 3000A° with 8000 different random orientations are adopted in generating DPs. DPs are shown on the bottom. From the left, the first and the second are the single particle shot and the third is the addition of the two DPs. The fourth DP is coherent scattering from the two nanorices whose positions are exactly the same as the first and the second at the same time.

Fig 38 proves the consequence from Eq. (56). Two $C2(q,q,\Delta\phi)$'s for Incoherent case (intensity addition) and Coherent case (amplitude addition) are the same except $\Delta\phi = 0, \pi$ and 2π.
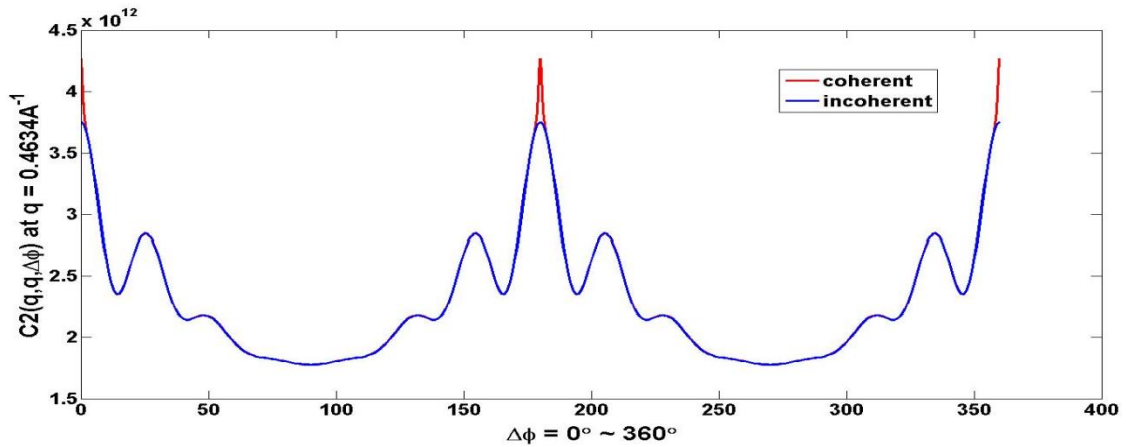
Fig 38. Two $C2(q,q,\Delta\phi)$'s coherent and incoherent cases are compared. As one can see that only the difference occurs at 0°, 180° and 360° at high q. The reason is explained with Eq. (57).

At the lower q values, the distinctions come clearer as in Fig 39.
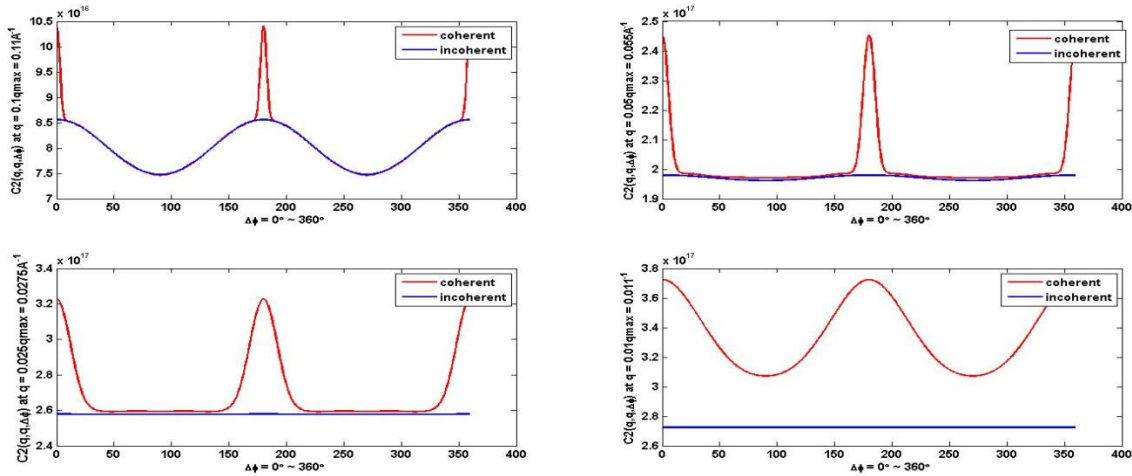
Fig 39. By lowering q values, the difference between coherent case and incoherent case appears clear. As increase the q values, the differences between two cases at $\Delta\phi=0^o$, 180°, 360° become smaller.

As one can easily observe in Fig 40, the average intensities in polar grids of incoherent and coherent cases, the difference between them becomes larger as q gets lowered. This is why the plot in Fig 39 shows the big difference of $C2$'s values at low q.
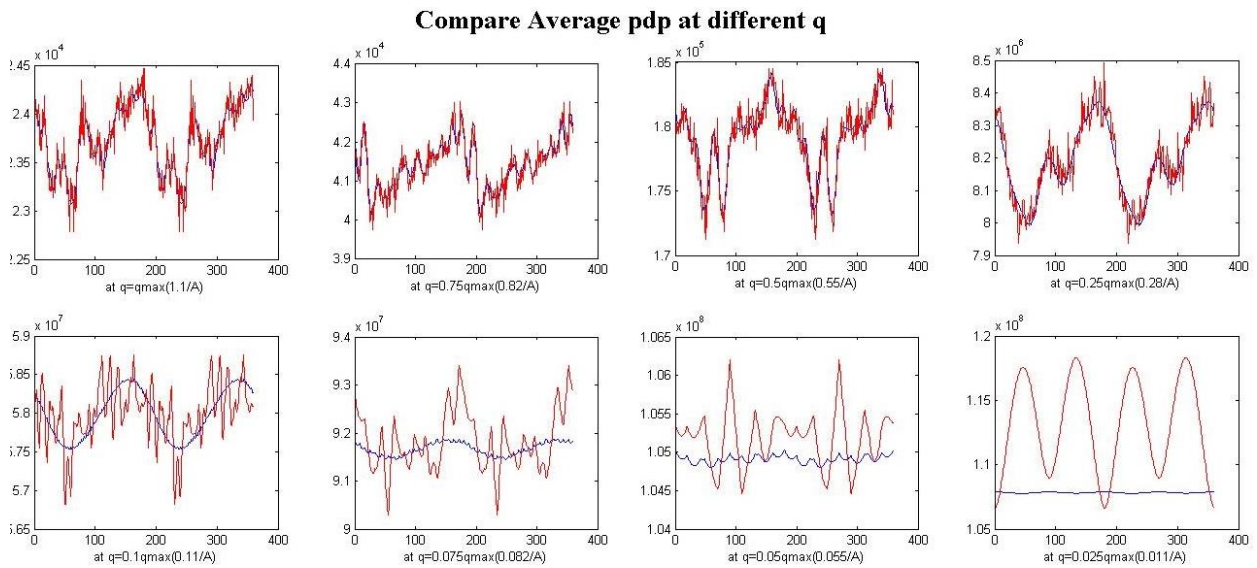
Fig 40. The average intensity difference on polar grids along the different angular positions becomes larger as q gets lowered. This shows that $C2(q,q,\Delta\phi)$ becomes wider as in Fig 39.

Now, once we got $C2(q,q',\Delta\phi)$ and $C3(q,q',\Delta\phi)$, we can calculate $B_l(q,q')$ and $T_l(q,q')$. For azimuthally symmetric particles, one may take $m = 0$ in Eq. (26) and (27). This immediately suggests

$$| I_{l,m=0}(q) | = \sqrt{B_l(q,q)} \tag{57}$$

Also, as $I_{l,m=0}(q)$ is determined by the integral of a real intensity with a real Legendre polynomial, it is real and the only remaining task is to determine its sings. We can determine these signs from the triple correlations by assuming, as it must be for a nanorice particle that azimuthal symmetry of the amplitudes implies azimuthal symmetry of the intensities from the usual Clebsch-Gordon rules [52] for adding angular momenta. In this case, all magnetic quantum numbers are equal to zeros, and the triple correlation reduces to a sum over only the angular momentum quantum number l. Thus,

$$T_l(q,q) = N\sum_{l_1,l_2} I_{l0}(q)I_{l_10}(q)I_{l_20}(q)\ G(l0;l_10;l_20) \tag{58}$$

where G is a Gaunt's coefficient [53]. Note that the two point triple correlations are scaled for multiple particles by exactly the same factor of N as the quantities of B_l. Since an ellipsoid has azimuthal symmetry about a particle axis we can choose the long particle axis as z-axis, thus eliminating any other components of the magnetic quantum number except $m = 0$. $| I_{l0}(q) |$ can be directly obtained by Eq. (57).

The only unknown here is the signs of $I_{l0}(q)$. Those signs can be determined by fitting all possible sings of $I_{l0}(q)$ to the values $T_l(q,q)$ of the triple correlations computed directly from DPs of randomly oriented particles using Eq. (25). It should be stressed that the number of equations is equal to the number of distinct $T_l(q,q)$ values, namely the numbers

of q and l values, as is the number of unknown $I_{l0}(q)$, so there will not be a short information conflict. It should be pointed out that, for a single particle object, the structure of such a simple particle had been determined even experimentally by a different method [54]. After the determination for the signs of $I_{l0}(q)$, one could get intensities in the reciprocal space using Eq. (1). Briefly, we explicit simulations of the expected diffraction patterns with Poisson noises and calculated from them, the angular correlations $C2(q,q';\Delta\phi)$ and $C3(q,q';\Delta\phi)$ leading to $B_l(q,q')$ and $T_l(q,q')$ from which the diffraction volume is calculated after the sign determination. Through an iterative phasing process, the image we found is shown in Fig 41.
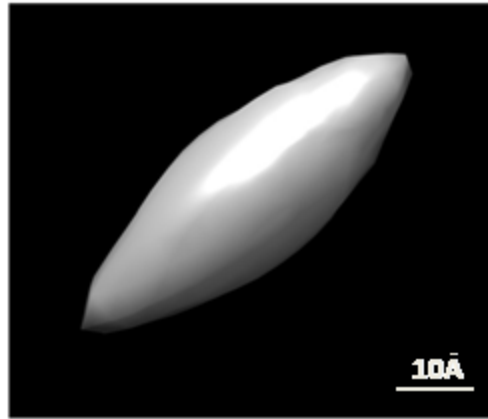
Fig 41. The charge density of single nanorice particle recovered from 8000 simulated diffraction patterns for two nanorice particles including shot noise of 10 ph/SP. $R_f = 0.1666$.

A nanorice particle is composed of Fe atoms with its density $7.7874 g/cm^3$, atom number 26, and assuming photon fluence 10^{14} photons/μm^2, photon counts per Shannon pixel $\approx$ 0.05 photons/Shannon Pixel and charge density $\rho(r) = 2.2$ electrons/A^3. As observed through the calculation by Eq. (20), photon counts/Shannon pixel $\approx$ 0.05 indicates that

adding noises onto the simulated DPs would barely affect the parameters we need. The procedures done by our developed algorithms show that there is no need for a hit-finder program [55] to reject multiple particle hits. In fact the particle-number-dependent scaling factors are the same for the pair and two-point triple correlations. Consequently it is not necessary to know at the outset exactly how many particles there are in the ensemble since both the pair correlation and two-point triple correlations are derived from diffraction patterns with exactly the same number of particles. The possibility using diffraction patterns possibly from multiple particles adds considerably to the capability of the use for XFEL for structure determination of individual particles as it will add greatly to the "hit-rate". Our simulation suggests that here maybe even advantages to considering diffraction patterns from ensembles of multiples in terms of convergence. We believe there is no other method that has been proposed for XFEL structure determination that has this ability from multiple particles per shot with independently random particle orientations.

6.3 Some other examples from different configurations and models result in successful
    Images.

We examine a bit more complicated azimuthally symmetric model and by placing two nanorice particles in a relatively small confined spherical space where these two particles easily touch without overlapping. The procedures led to the successful 3D images of the particles. Fig 41 and Fig 42 show the models that are used in 3D imaging process and those reconstructed final images. The simulations in Fig 41 and Fig 42 have been performed beforehand to justify the one in Fig 37.   The specification of an ellipsoidal

shape of nanorice particle used here that has long radius of 1000Å and short radius of 333Å are placed where the centers of the particles are place in the spherical region of radius = 1000Å. Fig 43 shows the more complicated shape of single particle of azimuthally symmetric model resembling d-orbital, is also adopted for this whole 3D imaging procedures with 2000 DPs to verify to get the right reconstructed image.
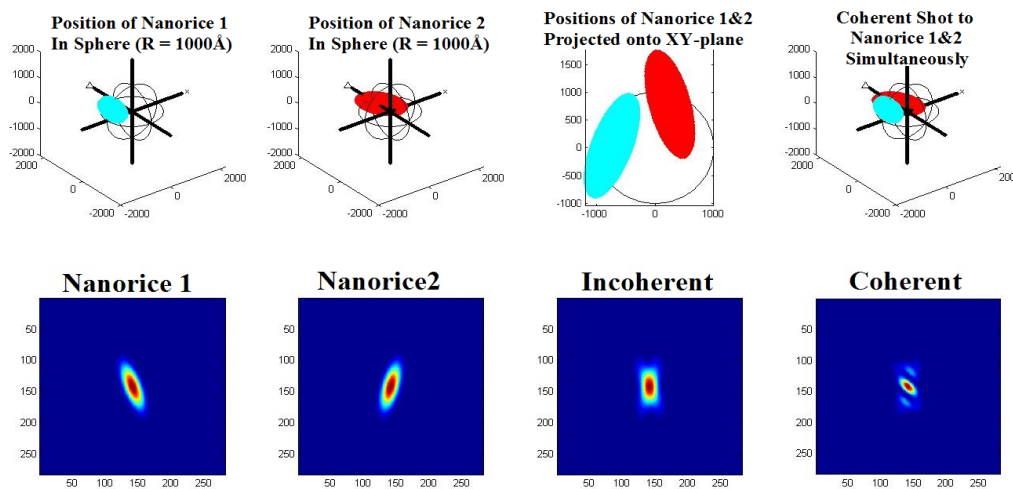
Fig 42. The centers two nanorices with long radius = 1000Å, short radius = 333Å are placed in a spherical region of radius = 1000Å and an X-ray shot is simulated onto these particles. The difference of images from diffraction patterns for incoherent and coherent cases is shown. The final image after processing all the work from this set up is the same as the one on Fig 41.
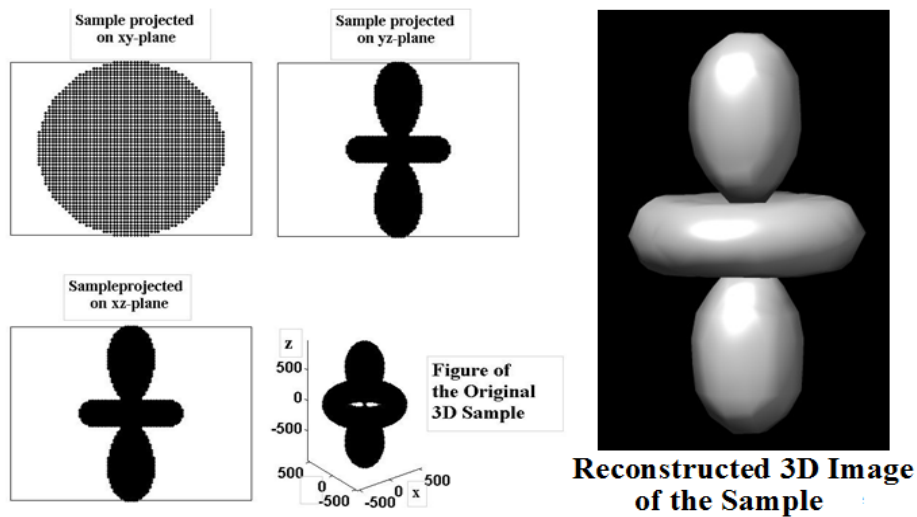
Fig 43. Another azimuthally symmetrical model has been adopted in one of our simulations for a test to verify if our algorithm works.

6.4 Photon-Counts/Shannon-Pixel Treatment

Now, we want to testify if the two nanorice particle 3D imaging procedure is possible with diffraction patterns generated by LCLS by examining the number of photons in a Shannon pixel. By lowering down the number of photons in the diffraction patterns with Poisson noise added, it is observable how the 3D reconstructed images of two nanorice particles change according to the photon-count/Shannon pixel. After accessing 100000 DPs of randomly oriented two nanorice particles in a cylindrical space as Fig 44 shows that 0.05 ph/SP seems to be the marginal threshold noise level under the assumption that the photon fluence in LCLS is approximately 10^6 photons/Å^2.
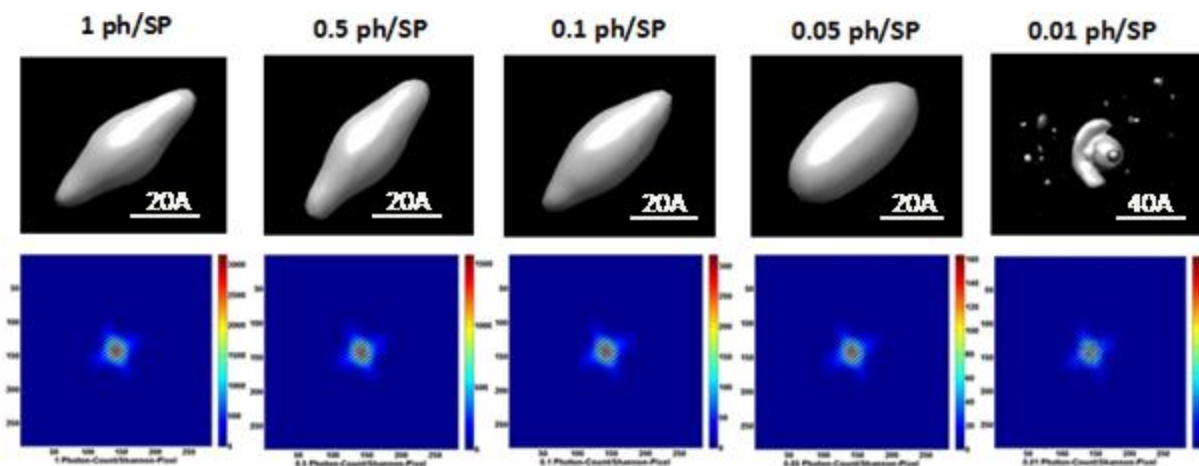
Fig 44. The top figures indicate that the noise level of 0.01 ph/SP failed the reconstructing image. 0.05 ph/SP seems marginal threshold noise level to get the decent reconstructed 3D image of two nanorice particles from 100000 DPs. From the left, R_f = 0.1731, 0.1737, 0.1726, 0.1650, 0.1629.

The above figures are from the one hundred thousand simulated diffraction patterns of two nanorice particles. If this performance is held in LCLS, probably more many diffraction patterns should be collected with the current photon fluence and noise, since the quality of

experimental data far worse than that of simulated ones. Photon fluence may be the important factor to perform the 3D reconstruction imaging process with small sized particles (size $< 50\text{Å}$).

Chapter 7

The 3D imaging process with real experimental data and those results

In this chapter, we are going to deal with real experimental data and show all the steps by which the data are treated. Brief diagram for equipment set up is shown in Fig 45. An incident X-ray pulse probably has some variations of its strength, more as strong in the center and weak at the edge. In RDV data collection from an SPI group [56], the photon energy E = 1.6keV was used, so the wavelength λ = hc/E = 7.756Å where h is the Plank's constant and c is the speed of light. The distance between two adjacent pixels in the detector is about 7.5×10^{-5}m, but in RDV case 4 times down sampling is achieved. Thus Δp = $4(7.5 \times 10^{-5}$m$) = 3.0 \times 10^{-4}$m. Fig 45 shows how to get q_{max} by setting up equations as

$$\tan\theta = \frac{N\Delta p/2}{Zd}, \qquad \text{thus } \theta = \tan^{-1}\left(\frac{N\Delta p/2}{Zd}\right). \tag{59}$$

$$q_{max} = K\sin\theta = \frac{2\pi}{\lambda}\sin\theta. \tag{60}$$

From (59) and (60),

$$q_{max} = \frac{2\pi}{\lambda}\sin\left[\tan^{-1}\left(\frac{N\Delta p/2}{Zd}\right)\right] \tag{61}$$

Therefore, Eq. (61) gives us $q_{max} \approx 0.054\text{Å}^{-1}$. As we take $q_{max} \approx 0.054\text{Å}^{-1}$, we want to take l_{max}=20 by $l_{max} \approx q_{max}R$ with the radius of RDV ≈ 375Å. Once q_{max} value is obtained, we need to interpolate intensities in Cartesian grid points into polar grid points. There are some difficulties in the interpolation process. Since the raw data of intensities in Cartesian grids as in Fig 46 contain zeros and negative numbers, those negative numbers are changed

to zeros. After converting all intensities on Cartesian grids to only zeros and positive numbers, finding centers of each DP will be an important issue to make already developed algorithms work.
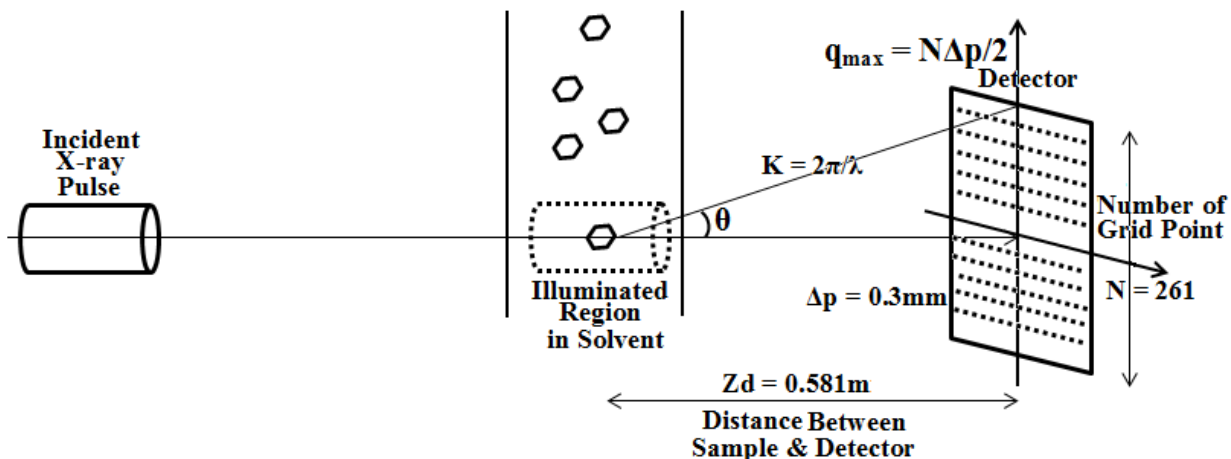
Fig 45. This diagram shows a brief equipment set up to collect Cartesian diffraction patterns for Rice Dwarf Virus (RDV).

Through the equipment set up as in Fig 45, the collected Cartesian raw diffraction patterns have many zeros in the middle columns where a couple of detectors are connected. Also there are lost intensities around the center, since the detectors are designed to let the incident X-ray pulse pass through the center to protect them from the strong intensities. Fig 46 shows the DPs of lost central intensities. These lost intensities are supposed to be resumed by the Rayleigh scattering amplitude function to make the image processing possible.
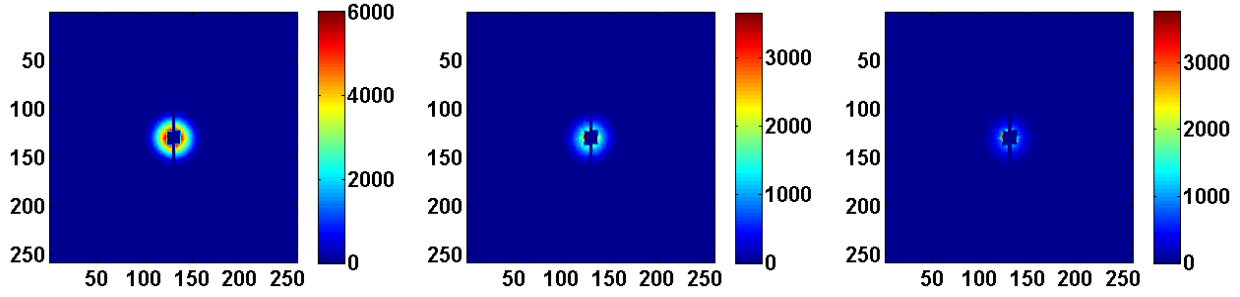
Fig 46. 2D images from DPs of RDV raw data. The central parts of the images indicate that there are lost central intensities in Cartesian DPs and that these lost should be resumed in some appropriate methods introduced in Chapter 5.

7.1 Find Centers in Each Diffraction Pattern

DPs of RDV raw data consisting of intensities on Cartesian grid points do not give us the exact locations of centers for each DP. Thus, we need to find them by using symmetric characteristics observed in most of simulated data. Our work for finding each center of DP is based on Friedel's Law [57][58] that intensities in DPs are distributed Centro-symmetrically about the center. The Freidel's Law defined as

$$F(q) = \int_{-\infty}^{+\infty} f(q) \exp(iq \cdot r) \, dr \qquad (62)$$

and has the following property

$$F(q) = F^*(-q) \implies |F(q)|^2 = |F^*(-q)|^2. \qquad (63)$$

From all simulated DPs on flat Ewald Spheres, one can observe that all intensities are Centro-symmetric. However, real experimental data do not follow those nicely distributed form. Thus we want to focus only on the central region where strong intensities are

66

recorded. Fig 47 shows how to find a center by comparing the strong central intensities around a point that we assume the center. If we take all pairs of intensities in experimental DPs to apply Friedel's Law for finding a center, that would be a crucial mistake since all experimental data are not Centro-symmetric while simulated ones are.
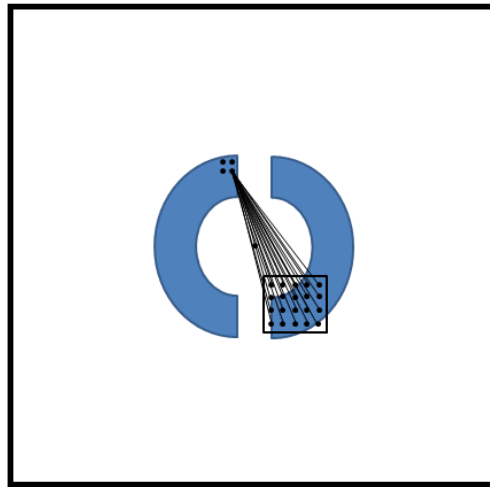
Fig 47. When Friedel's Law applied in experimental DPs for finding a center, only the central region where strong intensities are recorded should be considered while in simulated DPs, all possible pairs in the whole region is considered. Since the experimental data do not follow the Centro-symmetric property precisely.

7.2 Interpolation of Intensities from Cartesian grids to Polar grids

After converting all negative intensities in raw DPs into zeros, we could find centers from the central regions of Cartesian DPs using Friedel's Law through comparing pair intensities as in Fig 47 by taking pairs of minimum intensity differences between the two grid points. Taking average coordinate values of these chosen pairs give us the center of a

DP. Once the centers of DPs are found, it is necessary to rearrange the DP matrices by adjusting column and row sizes so that the new centers are supposed to be located at the middle of each square DP. Now, we interpolate these Cartesian DPs into Polar DPs as shown in Fg 48.

	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0.0991	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0.0194	0.0057	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	689.4361	394.6081	83.4555	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	2.4010e+03	1.5648e+03	697.7360	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	468.0261
13	2.9404e+03	2.0229e+03	1.1096e+03	199.6454	0	0	0	0	0	0	0	0	0	0	0	0	0	0	338.3860	1.1757e+03
14	3.3664e+03	2.4273e+03	1.4890e+03	551.3851	0	0	0	0	0	0	0	0	0	0	0	0	0	132.6666	1.0150e+03	1.8948e+03
15	3.0480e+03	2.2454e+03	1.4598e+03	688.1119	0	0	0	0	0	0	0	0	0	0	0	0	0	659.9987	1.5066e+03	2.3633e+03
16	3.1810e+03	2.6353e+03	1.8097e+03	980.2972	147.9539	0	0	0	0	0	0	0	0	0	0	0	284.3715	1.1269e+03	1.9706e+03	2.8156e+03
17	2.9339e+03	2.7277e+03	1.9231e+03	1.1243e+03	329.9549	0	0	0	0	0	0	0	0	0	0	0	647.6827	1.4456e+03	2.2528e+03	2.9174e+03
18	2.6386e+03	2.6242e+03	2.0326e+03	1.2568e+03	483.9738	0	0	0	0	0	0	0	0	0	0	179.6824	920.6802	1.6684e+03	2.4240e+03	2.5833e+03
19	2.3716e+03	2.4311e+03	2.1770e+03	1.4023e+03	629.3291	0	0	0	0	0	0	0	0	0	0	429.0066	1.1249e+03	1.8277e+03	2.2584e+03	2.2858e+03
20	1.8790e+03	1.9016e+03	1.8951e+03	1.2483e+03	614.1488	0	0	0	0	0	0	0	0	0	0	651.1909	1.3334e+03	2.0185e+03	2.1082e+03	2.1388e+03
21	1.6672e+03	1.6737e+03	1.6785e+03	1.2444e+03	659.5289	77.7814	0	0	0	0	0	0	0	0	135.3363	664.2734	1.2120e+03	1.6167e+03	1.7472e+03	1.8708e+03
22	1.4369e+03	1.4004e+03	1.3623e+03	1.1093e+03	616.4407	130.5687	0	0	0	0	0	0	0	0	253.8737	731.0239	1.2123e+03	1.3680e+03	1.4454e+03	1.5351e+03
23	1.2017e+03	1.1858e+03	1.1705e+03	1.0641e+03	617.4850	174.2750	0	0	0	0	0	0	0	0	371.2257	845.4384	1.2443e+03	1.2527e+03	1.2671e+03	1.2874e+03
24	1.0188e+03	1.0046e+03	968.3707	931.0734	558.0397	186.9212	0	0	0	0	0	0	0	4.6282	412.0116	827.5072	1.0431e+03	1.0565e+03	1.0703e+03	1.1099e+03
25	814.4667	800.3948	784.7529	773.3364	517.4394	195.5467	0	0	0	0	0	0	0	70.0856	393.6729	726.6038	864.0079	946.9281	983.8226	981.2554
26	653.5994	668.8634	647.4418	620.1675	450.5098	185.4259	0	0	0	0	0	0	0	119.3613	411.2489	674.5758	686.5192	716.3056	743.1827	755.5505
27	509.1752	459.8923	454.1414	484.8196	410.7818	181.1487	0	0	0	0	0	0	0	126.7007	348.1558	484.1285	466.1329	465.4286	546.7176	619.8825
28	369.1335	403.2879	427.3741	408.7144	344.3463	159.6236	0	0	0	0	0	0	0	100.9303	242.0192	319.9130	349.2946	384.9442	442.3732	480.0414
29	228.2672	267.6096	294.6507	269.3481	236.3871	113.4215	0	0	0	0	0	0	0	116.3719	245.0162	289.3651	320.4513	335.3833	349.6116	360.0494
30	167.3689	201.4188	238.1427	237.0060	234.9500	121.5584	2.8365	0	0	0	0	0	5.4810	126.6150	241.4591	246.8952	254.8106	245.1627	240.4903	244.0833
31	143.9453	153.2811	163.0989	164.1258	163.6195	92.1694	7.5148	0	0	0	0	0	8.8871	62.7779	122.9558	177.6936	192.9566	147.0190	125.6884	129.0442
32	121.5483	119.6245	105.7237	99.0495	100.0206	61.6598	7.9919	0	0	0	0	0	14.4349	62.2298	92.6423	107.4924	98.6062	64.5493	62.0462	74.8920
33	61.2206	61.8960	51.4669	51.8805	68.5772	51.0201	8.7685	0	0	0	0	0	9.7134	34.7797	54.3723	68.0443	65.2134	55.0617	60.4064	61.8187

Fig 48. Polar DPs are obtained by interpolating Cartesian DPs. The red rectangular spots show the maximum column values. Blue rectangular spots show also the maximum column values, but considering the values on the same rows, it is far less than the row averages.

As illustrated in Fig 48, each column contains very weak intensities at low q, even at lower q's than q's of column max intensities. Thus, we need to remove these weak intensities to prepare for the extrapolation into the beam stop shown as in Fig 49. The columns of red highlighted intensities will be used to extrapolate into the region of lost central intensities.

The columns of blue highlighted intensities would be ignored for Rayleigh tracing, since these numbers are too low as column maxima so that the errors could occur sizably in calculation procedures to achieve a 3D image. Almost 85% of experimentally recorded intensities are less than 10 and 35% of zeros while the maximum intensities of each column are about order of $10^3 \sim 10^4$. This is why we are bound to focus on the strong intensities around centers, since these much strength of recorded intensities could give us some trust for precisions of calculations. Even though we consider noise added to the intensities, strong intensities are affected much less in the calculations of average values such as C2 and C3 as in Eq. (22) and (23). Fig 49 shows that the red highlighted column maxima are to be employed by a fitting method using analytical scattering amplitude while the blue highlighted column maxima are not.

	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	3.3664e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	3.0480e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	3.1810e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	2.9339e+03	2.7277e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.9174e+03
18	2.6386e+03	2.6242e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.4240e+03	2.5833e+03
19	2.3716e+03	2.4311e+03	2.1770e+03	1.4023e+03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.2584e+03	2.2858e+03
20	1.8790e+03	1.9016e+03	1.8951e+03	1.2483e+03	0	0	0	0	0	0	0	0	0	0	0	0	1.3334e+03	2.0185e+03	2.1082e+03	2.1388e+03
21	1.6672e+03	1.6737e+03	1.6785e+03	1.2444e+03	659.5289	0	0	0	0	0	0	0	0	0	0	0	1.1220e+03	1.6167e+03	1.7472e+03	1.8708e+03
22	1.4369e+03	1.4004e+03	1.3623e+03	1.1093e+03	616.4407	0	0	0	0	0	0	0	0	0	0	0	1.2123e+03	1.3680e+03	1.4454e+03	1.5351e+03
23	1.2017e+03	1.1858e+03	1.1705e+03	1.0641e+03	617.4850	0	0	0	0	0	0	0	0	0	0	845.4384	1.2443e+03	1.2527e+03	1.2671e+03	1.2874e+03
24	1.0188e+03	1.0046e+03	968.3707	931.0734	558.0397	0	0	0	0	0	0	0	0	0	412.0116	827.5072	1.0431e+03	1.0565e+03	1.0703e+03	1.1099e+03
25	814.4667	800.3948	784.7529	773.3364	517.4394	195.5467	0	0	0	0	0	0	0	0	393.6729	726.6038	864.0079	946.9281	983.8226	981.2554
26	653.5994	668.8634	647.4418	620.1675	450.5098	185.4259	0	0	0	0	0	0	0	0	411.2489	674.5758	686.5192	716.3056	743.1827	755.5505
27	509.1752	459.8923	454.1414	484.8196	410.7818	181.1487	0	0	0	0	0	0	0	126.7007	348.1558	484.1285	466.1329	465.4286	546.7176	619.8825
28	369.1335	403.2879	427.3741	408.7144	344.3463	159.6236	0	0	0	0	0	0	0	100.9303	242.0192	319.9130	349.2946	384.9442	442.3732	480.0414
29	228.2672	267.6096	294.6507	269.3481	236.3871	113.4215	0	0	0	0	0	0	0	116.3719	245.0162	289.3651	320.4513	335.3833	349.6116	360.0494
30	167.3689	201.4188	238.1427	237.0060	234.9500	121.5584	0	0	0	0	0	0	0	126.6150	241.4591	246.8952	254.8106	245.1627	240.4903	244.0833
31	143.9453	153.2811	163.0989	164.1258	163.6195	92.1694	0	0	0	0	0	0	0	62.7779	122.9558	177.6936	192.9566	147.0190	125.6884	129.0442
32	121.5483	119.6245	105.7237	99.0495	100.0206	61.6598	0	0	0	0	0	0	0	62.2298	92.6423	107.4924	98.6062	64.5493	62.0462	74.8920
33	61.2206	61.8960	51.4669	51.8805	68.5772	51.0201	0	0	0	0	0	0	0	34.7797	54.3723	68.0443	65.2134	55.0617	60.4064	61.8187

Fig 49. The intensities at lower q's than at q's of column maximums are eliminated to prepare for the intensity extrapolations into the region of lost central intensities.

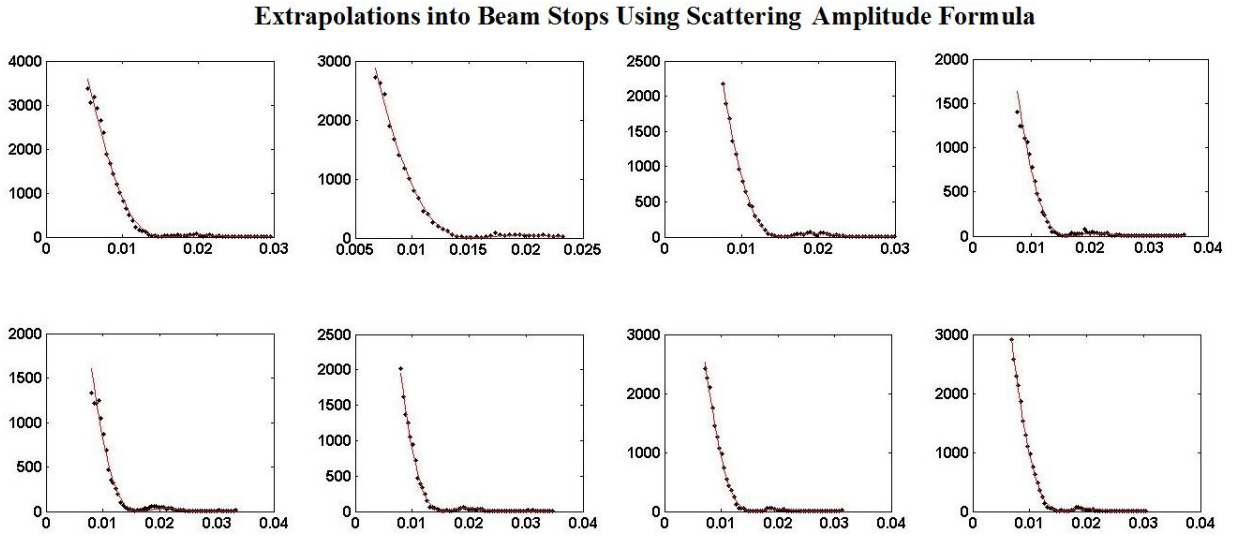## 7.3 Scattering Amplitude is used for Extrapolations into Beam Stops

As we test in section 5.2, even though DPs with lost intensities at low 5% of q_{max}, it is still possible to do all procedures for achieving a 3D image of a sample. Thus, this technique is used to extrapolate into the region of lost intensities of at low $0.05q_{max}$. Fig 50 shows that the dashed arrows indicate how many lost intensities are resumed in a column using the square of scattering amplitude function.

	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1.3298e+04	1.7983e+04	2.1717e+04	1.8291e+04	0	0	0	0	0	0	0	0	0	0	0	0	2.2773e+04	2.4910e+04	2.0535e+04	1.9069e+04
7	1.2549e+04	1.6970e+04	2.0494e+04	1.7261e+04	0	0	0	0	0	0	0	0	0	0	0	0	2.1490e+04	2.3507e+04	1.9378e+04	1.7995e+04
8	1.1711e+04	1.5837e+04	1.9125e+04	1.6108e+04	0	0	0	0	0	0	0	0	0	0	0	0	2.0055e+04	2.1937e+04	1.8084e+04	1.6793e+04
9	1.0803e+04	1.4609e+04	1.7643e+04	1.4859e+04	0	0	0	0	0	0	0	0	0	0	0	0	1.8500e+04	2.0237e+04	1.6682e+04	1.5491e+04
10	9.8474e+03	1.3317e+04	1.6082e+04	1.3545e+04	0	0	0	0	0	0	0	0	0	0	0	0	1.6864e+04	1.8446e+04	1.5206e+04	1.4121e+04
11	8.8649e+03	1.1988e+04	1.4477e+04	1.2193e+04	0	0	0	0	0	0	0	0	0	0	0	0	1.5181e+04	1.6606e+04	1.3689e+04	1.2712e+04
12	7.8765e+03	1.0652e+04	1.2863e+04	1.0834e+04	0	0	0	0	0	0	0	0	0	0	0	0	1.3488e+04	1.4754e+04	1.2163e+04	1.1294e+04
13	6.9020e+03	9.3337e+03	1.1272e+04	9.4934e+03	0	0	0	0	0	0	0	0	0	0	0	0	1.1820e+04	1.2929e+04	1.0658e+04	9.8970e+03
14	3.3664e+03	8.0589e+03	9.7321e+03	8.1968e+03	0	0	0	0	0	0	0	0	0	0	0	0	1.0205e+04	1.1163e+04	9.2024e+03	8.5453e+03
15	3.0480e+03	6.8485e+03	8.2704e+03	6.9657e+03	0	0	0	0	0	0	0	0	0	0	0	0	8.6725e+03	9.4864e+03	7.8202e+03	7.2618e+03
16	3.1810e+03	5.7202e+03	6.9079e+03	5.8181e+03	0	0	0	0	0	0	0	0	0	0	0	0	7.2438e+03	7.9236e+03	6.5319e+03	6.0655e+03
17	2.9339e+03	2.7277e+03	5.6614e+03	4.7683e+03	0	0	0	0	0	0	0	0	0	0	0	0	5.9367e+03	6.4938e+03	5.3533e+03	2.9174e+03
18	2.6386e+03	2.6242e+03	4.5429e+03	3.8262e+03	0	0	0	0	0	0	0	0	0	0	0	0	4.7637e+03	5.2108e+03	2.4240e+03	2.5833e+03
19	2.3716e+03	2.4311e+03	2.1770e+03	1.4023e+03	0	0	0	0	0	0	0	0	0	0	0	0	3.7323e+03	4.0826e+03	2.2584e+03	2.2858e+03
20	1.8790e+03	1.9016e+03	1.8951e+03	1.2483e+03	0	0	0	0	0	0	0	0	0	0	0	0	1.3334e+03	2.0185e+03	2.1082e+03	2.1388e+03
21	1.6672e+03	1.6737e+03	1.6785e+03	1.2444e+03	659.5289	0	0	0	0	0	0	0	0	0	0	0	1.2120e+03	1.6167e+03	1.7472e+03	1.8708e+03
22	1.4369e+03	1.4004e+03	1.3623e+03	1.1093e+03	616.4407	0	0	0	0	0	0	0	0	0	0	0	1.2123e+03	1.3680e+03	1.4454e+03	1.5351e+03
23	1.2017e+03	1.1858e+03	1.1705e+03	1.0641e+03	617.4850	0	0	0	0	0	0	0	0	0	0	845.4384	1.2443e+03	1.2527e+03	1.2671e+03	1.2874e+03
24	1.0188e+03	1.0046e+03	968.3707	931.0734	558.0397	0	0	0	0	0	0	0	0	0	412.0116	827.5072	1.0431e+03	1.0565e+03	1.0703e+03	1.1099e+03
25	814.4667	800.3948	784.7529	773.3364	517.4394	195.5467	0	0	0	0	0	0	0	0	393.6729	726.6038	864.0079	946.9281	983.8226	981.2554
26	653.5994	668.8634	647.4418	620.1675	450.5098	185.4259	0	0	0	0	0	0	0	0	411.2489	674.5758	686.5192	716.3056	743.1827	755.5505
27	509.1752	459.8923	454.1414	484.8196	410.7818	181.1487	0	0	0	0	0	0	0	126.7007	348.1558	484.1285	466.1329	465.4286	546.7176	619.8825
28	369.1335	403.2879	427.3741	408.7144	344.3463	159.6236	0	0	0	0	0	0	0	100.9303	242.0192	319.9130	349.2946	384.9442	442.3732	480.0414
29	228.2672	267.6096	294.6507	269.3481	236.3871	113.4215	0	0	0	0	0	0	0	116.3719	245.0162	289.3651	320.4513	335.3833	349.6116	360.0494
30	167.3689	201.4188	238.1427	237.0060	234.9500	121.5584	0	0	0	0	0	0	0	126.6150	241.4591	246.8952	254.8106	245.1627	240.4903	244.0833
31	143.9453	153.2811	163.0989	164.1258	163.6195	92.1694	0	0	0	0	0	0	0	62.7779	122.9558	177.6936	192.9566	147.0190	125.6884	129.0442
32	121.5483	119.6245	105.7237	99.0495	100.0206	61.6598	0	0	0	0	0	0	0	62.2298	92.6423	107.4924	98.6062	64.5493	62.0462	74.8920
33	61.2206	61.8960	51.4669	51.8805	68.5772	51.0201	0	0	0	0	0	0	0	34.7797	54.3723	68.0443	65.2134	55.0617	60.4064	61.8187

Fig 50. The dashed arrows indicate how many lost intensities are extrapolated into the region of lost central intensities using scattering amplitude formula. Considering $q_{max} = 121^{st}$ row, 6th row where the extrapolations into the lost region are ended, is about a bit lower than $0.05q_{max}$.

In Fig 50, if the maximum intensity in a column is too low, then these columns are not participated in intensity extrapolations, since the low column maxima can cause somewhat

big errors in calculation procedures so that it would be difficult to make the appropriate images. In simulation DPs, the intensities do not change very much along the angular positions unlike the row q_7 of DP1 in Table 1.

Extrapolations into Beam Stops Using Scattering Amplitude Formula

Fig 51. From Fig 50, each column is fitted using scattering amplitude formula as in Eq (43) function that is used for extrapolations into beam stops. Each column has different beam stops as shown in this figure.

Fig 51 indicates that if we select quite trustful columns of strong intensity maxima and use them to extrapolate into beam stops, then those selected ones would help us to go on to the next step calculations. After filling into beam stops with resumed intensities using scattering amplitude function, we are supposed to check the intensities at $q \approx 0.05 q_{max}$ and recognize that we need to normalize the average intensities at that q.

7.4 Sort and Normalize the Average of the Intensities at 0.05 $q_{\max}$.

As we can see in Table 2, at $q \approx 0.05q_{\max}$, the intensities are consistent irrespective of angular positions or orientations in simulated DPs. Considering shot to shot incident X-ray pulse variations, the average of intensities at $q \approx 0.05q_{\max}$ varies as in Table 1. We followed the steps introduced in this chapter to overcome the artifacts appeared in 494 experimental diffraction patterns of Rice Dwarf Virus from SPI [59], with the specification as in Fig 45, and compared $C2(q,q,\Delta\phi)$ and $B_l(q,q)$ at $q \approx 0.018\text{Å}^{-1}$ with those from pdb-file (entry# : 1uf2) as in Fig 52. After performing all the procedures explained in this chapter, we could get the 3D reconstructed image of RDV with decent icosahedral shape, not perfectly though.
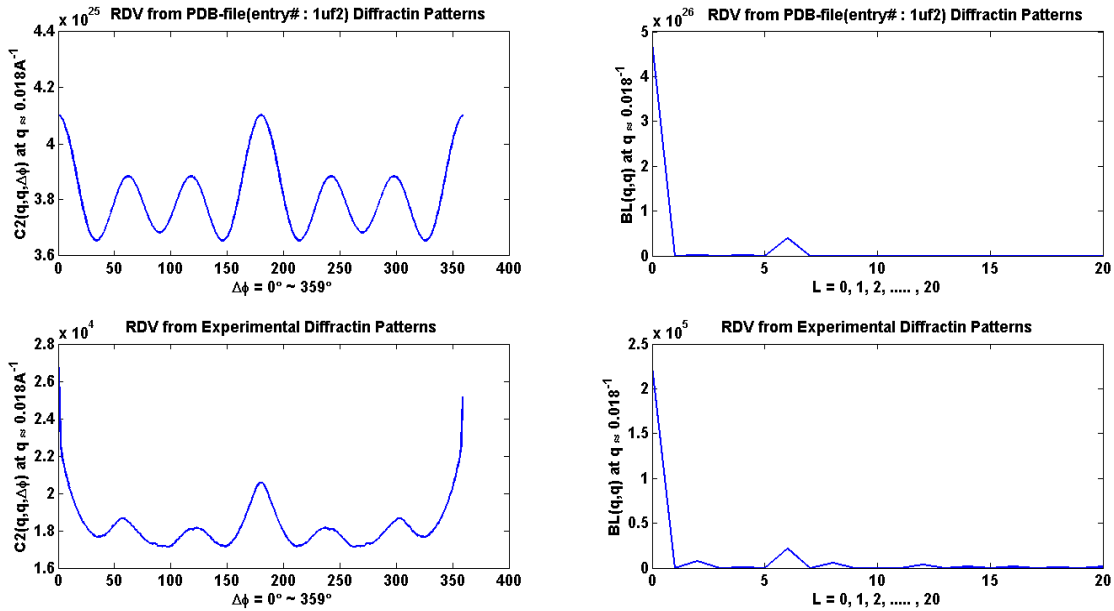
Fig 52. When 494 RDV experimental diffraction patterns are accessed through the procedures introduced in this chapter, the similarity between $C2(q,q,\Delta\phi)$ and $B_l(q,q)$ from PDB file and those from the experimental data is found. The peaks in $B_l(q,q)$ at l=0, 6, 10, 12, 16, 18, 20 from the experimental data give a hint to make the 3D reconstruction image possible.

Similar procedures have been performed with 198 mimi virus experimental diffraction patterns [60]. The peaks as in Fig 53 appeared at the allowed *l*-values for icosahedral symmetry in some $B_l(q,q)$ vs. L plots.
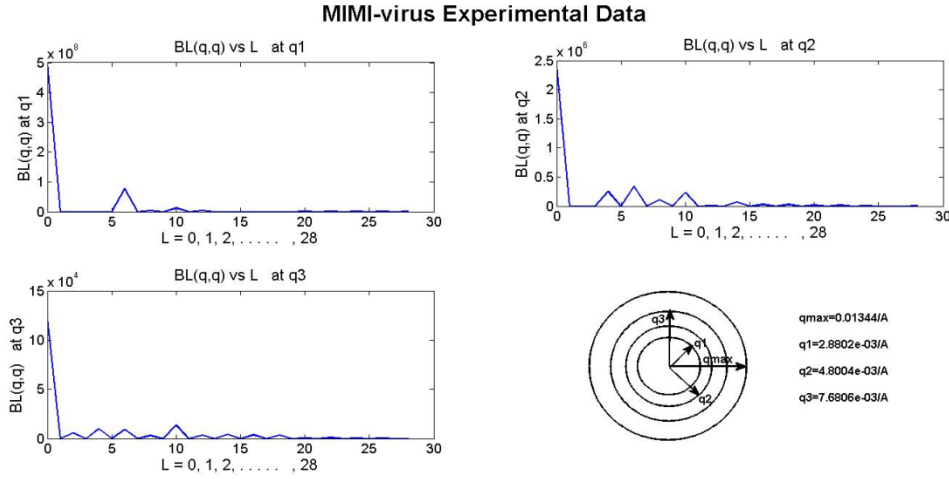
Fig 53. One hundred ninety eight experimental diffraction patterns of mimi virus are accessed. The size is approximately 4500Å, incident photon energy ≈ 1.2KeV, detector pixel size ≈ 75μm, distance between the detector and the sample ≈ 0.74m, the number of detector pixels ≈ 1000x1000. It is observed that the peaks appeared at the allowed *l*-values for icosahedral symmetry.

With $B_l(q,q)$ and $T_l(q,q)$, one could get the diffraction volume through the algorithm using Eq. (26)~(30). And the iterative phasing algorithm would generate the 3D reconstructed images as in Fig 54 for RDV and mimi viruses.
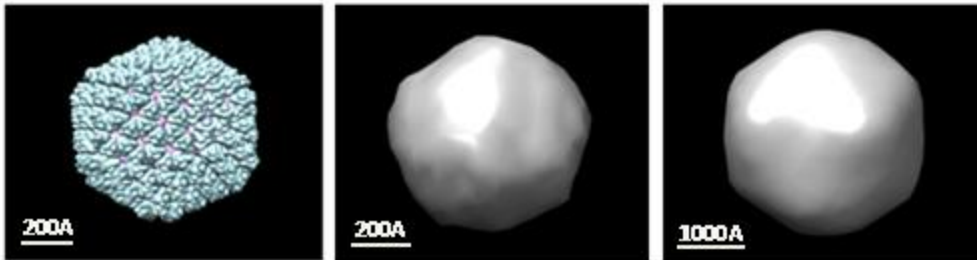
Fig 54. The left side shows 3D image of RDV(entry#:1uf2_pdb1), the middle image is from 494 RDV experimental DPs with *res* ≈ 120Å and the right side image is from 198 experimental DPs of mimi virus *res* ≈ 500Å through the procedures that we have developed in this research. However, the right side image for the mimi virus should be reexamined since its size appeared a bit smaller than expected.

73

To find the resolution for the 3D imaging process of RDV experimental data, we used FSC (Fourier Shell Correlation) [61][62] method by dividing 494 diffraction patterns into two sets of 247 DPs. Once the charge densities of each set are obtained after performing the phasing process, one could get 3D amplitude maps in the reciprocal space. Fig 55 shows that two sets of each 247 DPs of RDV diffraction patterns are performed separately to get 3D reconstructed images.
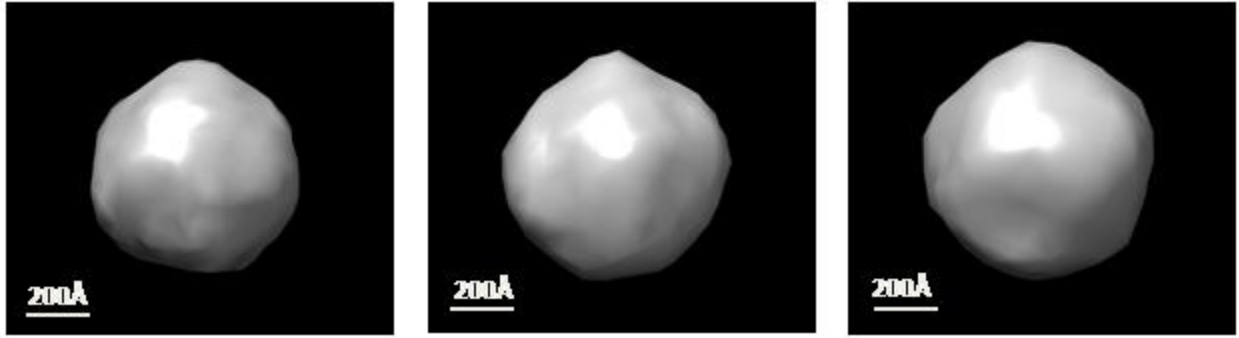
Fig 55. The first two images from the left, two sets of each 247 DPs of RDV diffraction patterns are performed separately to get 3D reconstructed images. The right image is for 494 DPs. The resolution by FSC shows ≈ 220Å.

Fourier Shell Correlation is a method to compute over successive shells of certain radius and width, but it directly compares the two Fourier transforms as in Eq. (64). Through the plot of FSC(q) vs. q, one could get the resolution by selecting the q value at FSC(q) = 0.5.

$$FSC(q) = \frac{\sum_{q_i} A_1(q_i) \cdot A_2^{*}(q_i)}{\sqrt{\left[\sum_{q_i} |A_1(q_i)|^2\right] \cdot \left[\sum_{q_i} |A_2(q_i)|^2\right]}} \tag{64}$$

$A_1(q_i)$ and $A_2(q_i)$ are amplitudes of the first and second shell-elements of radius q in the reciprocal space.

Fig 56 shows the Fourier Shell Correlations through the plot for FSC(q) vs. q.
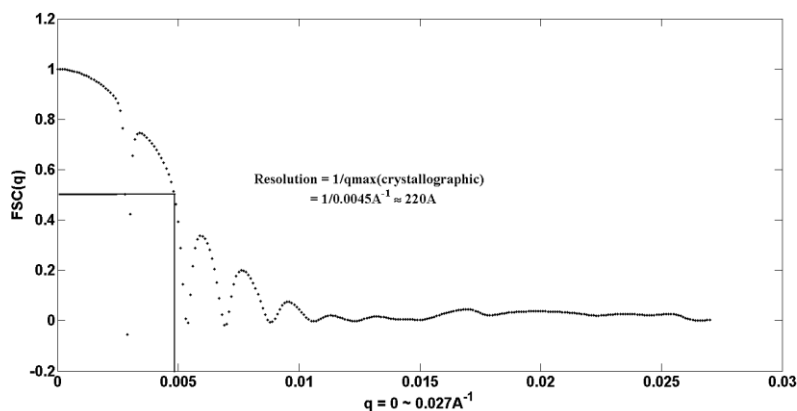
Fig 56. Charge Densities after separating 494 RDV experimental DPs into two sets of 247 DPs are obtained. FSC is performed to get the resolution. Resolution $\approx$ 220Å.

It is observed that the resolution from FSC for the experimental data is greater than the one, $1/q_{max}$ used for simulated data. In RDV experimental data, the resolution $\approx 1/q_{max} \approx$ 110Å. By FSC, the resolution $\approx 1/0.0045\text{Å}^{-1} \approx$ 220Å.

Chapter 8

Discussion and Conclusion

Approximately 40% of all kinds of biomolecules cannot be crystalized [9]. "Diffract before Destroy" experiments with ultrashort durations and ultra-bright XFEL pulses generating diffraction patterns on digital detectors give us good chances to reconstruct the 3D images of uncrystallized individual biomolecules [63]. Some membrane proteins or viruses are difficult to be crystalized. Although half a million proteins have been sequenced, their structures of about 10% have been determined so far (www.pdb.org). In spite of crystallizing deficiencies for those nanoparticles, the "Diffract before Destroy" experimental technique provides an unprecedented opportunity to determine the structures of uncrystallized particles. Using this technique, at the LCLS (Linac Coherent Light Sources) [64] in Stanford, CA. performed a few experimental procedures for relatively large objects such as viruses that produce low resolution diffraction patterns, amongst which we obtained the data of mimivirus [58], RDV, PR772 and PBCV1. One of the methods for reconstructing 3D images of the objects is angular correlations between intensities in DPs. The method of angular correlations recovers quantities from diffraction patterns of randomly oriented particles, as expected to be measured at an X-ray Free Electron Laser (XFEL) proportional to quadratic functions of the spherical harmonic expansion coefficients of the diffraction volume of a single particle. We have already illustrated that it is possible to reconstruct a randomly oriented icosahedral virus, helical virus or some other nanoparticles such as nanorice particles from the average over all measured diffraction patterns of such correlations through computer generating simulated data. However, when we deal with real experimental data, many artifacts should be

considered, for example, the effect of water, noise, shot to shot variations of strength of incident X-ray pulses, curved Ewald Sphere, lost central intensities, off set of centers in each diffraction pattern, involvement of a lot of zeros on the Cartesian grid points of the detectors, chemical affinity between biomolecules, or multi particle scattering, etc. We introduced a couple of methods how to refine the raw data of diffraction patterns from experiments, to overcome those artifacts and perform the calculation procedures that have already been developed for enabling 3D imaging process. For one instance, shot noise can be overcome by taking many more diffraction patterns so that the averaging the product of intensities would significantly reduce the noise effect. Applying the methods introduced in this paper has many small difficulties in handling the real experimental data though, since the obtained diffraction patterns have their unknown irregularities in recorded intensities that can cause bad precisions of interpolations for intensities from Cartesian grids to polar grids, which is an important initial step to go on the next level of calculations. This may be resolved by eliminating some untrustworthy features such as extremely low intensities in beam stop area and filling in the region of lost central intensities with the extrapolated values by an intensity fitting [48] mentioned in section 7.3. The problem created by shot-to-shot variations in incident strength of X-ray pulses can be accommodated by selecting full-hit diffraction patterns from those of partial hits and by normalizing the intensities on polar grid points. We also showed possibilities to handle multi particle scatterings by adopting two nanorice particles as a sample when we simulated X-ray shot and make diffraction patters. The procedures succeeded in finding charge densities of single nanorice particle through simulated diffraction patterns from two nanorice particles. It is found that the curves of $C2(q,q,\Delta\phi)$ coincide between the coherent and incoherent scatterings except

when $\Delta\phi$ is zeros and π. We have suggested how to deal with this difference. These simulations were performed assuming an incoherent source of X-rays as the intensity from many particles is assumed to be the sum of intensities from single particles. This is possible due to the random positions of the particles. Thus, the distinction with the coherent case is not very important over most of the range of the correlations, as the randomness of particle positions give the scattering a kind of incoherence [65]. Furthermore, it would be worth trying more than two particles in the aerosol that will be illuminated by the incident X-ray pulse when the diffraction patterns are collected through the shots. Nevertheless, the two independently randomly oriented particles are more than can be done with most methods from the XFEL structural problem. The angular correlation method leaves a cautiously optimistic anticipation to attain single particle charge density from the diffraction patterns of multiple particle (more than two) scatterings. Moreover we showed that, even if a small object (of the size of a few Angstroms) exposed in X-ray shots gives very small number of photons on the detector (0.01 ph/SP), these diffraction patterns can achieve the 3D imaging of the object even for a curved Ewald Sphere using twelve atom cluster model. When we access the RDV experimental data, there are two distinctive features that are very different from the simulated ones. The first is that there are lost central intensities and shot to shot variations in each diffraction pattern with different centers. Second is that the distribution of photons on the 2D detector is not similar as we observed in the simulated data. However, the extrapolations into the beam stop using an analytic scattering amplitude after eliminating the irregular intensities around the center could resolve the difficulties in recovering 3D image of RDV. Even though we tried

methods introduced in previous chapters, there are still many problems to be solved such as phasing or finding angular correlations precisely even when all artifacts are eliminated.

To sum up what we have done through the research, we showed three major achievements.

- First, in the presence of Poisson noise, using one million simulated DPs of viruses and a twelve atom cluster model, we reconstructed a 3D structural image at atomic resolution (at least in the latter case) at photon counts of 0.01 photons/Shannon Pixel successfully through the angular correlation method.

- Second, angular correlation method also enables the diffraction patterns from multiple particle scattering to be performed to get a 3D reconstructed structure of a single particle. We achieved this using simulated diffraction patterns with two nanorice particles of 50Å length each.

- Lastly, we reconstructed the 3D electron density of a RDV virus at the resolution of 220Å for the first time with an angular correlation method from the real experimental data collected from LCLS.

The incipient future endeavor would be the elimination of solvent scattering, perhaps through a Babinet argument since all living biomolecules contain large portion of water. Intensities in diffraction patterns generated by the scattering from water and biomolecules should be distinguished. Observations for the artifacts that hinder the reconstruction of the structure of nanoparticles are to be followed and the remedies have to be studied more for the better resolution of 3D images of samples of interest. We would also like to be free from imposing symmetry that is used to get diffraction volumes. In our algorithm, we impose icosahedral or azimuthal symmetry in getting diffraction volumes even though we

do not impose any constraints and symmetry in phasing. The brightness of XFEL is also to be studied when we consider the scattering from a small sized object (few Angstroms) containing small number of atoms that will leave very small number of photons on the detector. More realistically, we can simulate diffraction patterns with all possible artifacts, such as solvent, weak intensities of XFEL for small objects, chemical affinity between objects (as used in multi particle scattering), shot-to-shot variations of incident X-ray pulses, curved Ewald Spheres, lost central intensities, off-set of centers in each diffraction pattern. Developing algorithms with all possible artifacts adopted at the same time can give us some directions how to handle the real experimental data. Later on, some other XFEL research centers will open soon in Europe, South Korea, and Switzerland [66]. If we have chances to get the experimental data from those recent XFEL research centers, there will be more expectable opportunities to realize 3D imaging process more clearly and precisely for nanoparticles of our interest. The following Table shows the contingent schedule of beam operations for some XFEL research centers and mostly will open one year later than these operations with beams as the following diagram.

Schedule hard X-ray FELs worldwide
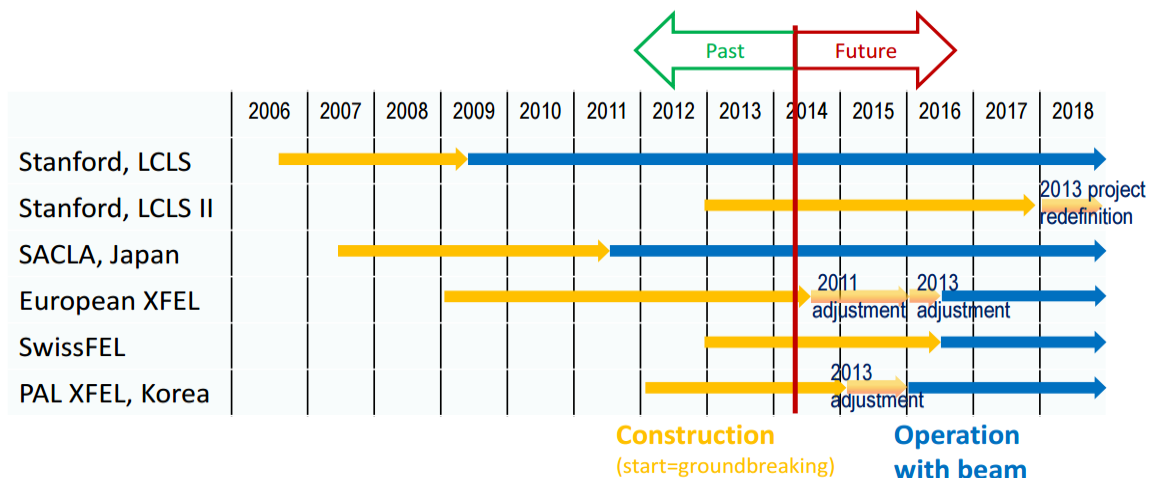
References

[1] W. A. Barletta, J.Bisognano, J. N. Corlett, P. Emma, Z. Huang, K. J. Kim, R. Lindberg, J. B. Murphy G. R. Neil, D. C. Nguyen, C. Pelligrini, R. A. Rimmer, F. Sannibale, G. Stupakov, R. P. Walker, and A. A. Zholents, "Free electron lasers: Present status and Future challenges," Nucl. Instrum. Methods Phys doi: 10.1016/j.nima.2010.02.274, Res, Sect. A 618, 69 (2010).

[2] P. Emma et al, "First lasing and operation of an angstrom-wavelength free-electron laser", Nat. Photonics 4, 641-647,doi:10.1038/nphoton.2010.176, (2010).

[3] DESY/XFEL Status, Matthias Clausen, DESY, Cryogenic Controls Group-MKS-2 P5(2012), SLAC(April, 25, 2012).

[4] D.M. Mills, J.R. Heliwell, A. Kwick, T. Ohta, I.A. Robinson, A. Authier, J. "Report of the working group on synchrotron radiation nomenclature: brightness, spectral brightness or brilliance?" Synchrotron Radiation. 12, 385, doi:10.1107. (2005).

[5] Dimitrios Giannakis, Peter Schwander and Abbas Ourmazd, "The symmetries of image Formation by scattering. I. Theoretical framework", Optics Express, Vol 20, pp.12799-12826, doi:10.1364/E.20.012799, (2012).

[6] Peter Schwander, Dimitrios Giannakis, Chun Hong Yoon and Abbas Ourmazd, "The symmetries of image formation by scattering. II. Applications", Optics Express, Vol 20. pp.12827-12849, doi:10.1364/OE.20.012827, (2012).

[7] A. Hosseinizadeh, P. Schwander, A. Dashti, R. Fung, R. M. D'Souza and A. Ourmazd, "High-Resolution Structure of Viruses from Random Sanpshots", Philosophical Transactions of the Royal Society, B369, doi:10.1098/rstb.2013.0326, (2014).

[8] Ne-Te Duane Loh and Veit Elser, "Reconstruction Algorithm for Single-Particle Diffraction Imaging Experiments", Phys. Rev. E 80, Vol 80, doi:10.1103/PhysRevE.80. 026705, (2009).

[9] D.K. Saldin, V.L. Shneerson, R Fung, A Ourmazd.: "Structure of isolated biomolecules Obtained from ultrashort X-ray pulses: exploiting the symmetry of random orientatin", Journal of Physics: Condensed Matter, Vol 21, 12, doi:10.1088/0953-8984/21/13/134014, Mar, (2009).

[10] D.K. Saldin, H.C. Poon, P. Schwander, M. Uddin, and M. Schmidt, "Fiber Diffraction without Fibers", Optics Express, A19, 17318-17335, doi: 10.1103/PhysRevLett. 110. 265505, (2011).

[11] Zvi. Kam, "The reconstruction of structure from electron micrographs of randomly Oriented particles", Journal of Theretical Biology, doi:10.1016/0022-5193(80)90088-0, (1980).

[12] M. Nespolo, and Giovanni Ferraris, "Applied geminograph-symmetry analysis of twin Crystals and definition of twinning by reticular polyholohedry", (Acta Crystallographica A). doi : 10.1107/S0108767303025625, Nov 5, (2003).

[13] Crystal Structure Analysis 5.062 (Spring/2008): http://ocw.mit.edu/terms, MIT Open Course Ware.

[14] T. Alwyn Jones, Lars Liljas. "Structure of satellite tobacco necrosis virus after Crystallographic refinement at 2.5Å resolution". Journal of Molecular Biology, Voume 177, Issue 4, P735-765. doi:10.1016/0022-2836(84)90047-0, 25 August 1984.

[15] Yu G. Kuznetsov, James R. Gurnon, James L. Van Etten, Alexander McPhereson. "Atomic force microscopy investigation of a Chlorella virus, PBCV-1". Journal of Structural Biology, Volume 149, Issue 3, P256-263. doi:10. 1016/m.jsb 10.07. March 2005.

[16] Hong Zhou, Matthew L. Baker, Wen Jiang, Matthew Dougherty, Joanita Jakana, Gang Dong, Guangying Lu and Wah Chiu. "Electron cryomicroscopy and bioinformatics Suggest protein fold models for rice dwarf virus". Nature Structural Biology 8, 868-873

doi:10.1038/nsb1001-868. (2001).

[17] Andrew V. Martin, Jakob Andreassson, Andrew Aquila, Saša Bajt, Thomas R. M.

Barends, Miriam Barthelmess, Anton Barty, W. Henry Benner, Chritoph Bostedt,

John D. Bozek, Phillip Bucksbaum, Carl Caleman, Nicola Coppola, Daniel P. Deponte,

Tomas Ekeberg, Sascha W. Epp, Benjamin Erk, Gerge R. Farquar, Holger Fleckenstein,

Lutz Foucar, Matthias Frank, Lars Gumprecht, Christina Y. Hampton, Max Hanktke,

Andreas Hartmann, Elisabeth Hartmann, Robert Hartmann, Stephan P. Hau-Riege,

Günther Hauser, Peter Holl, André Hoemke, Olof Jönsson, Stephan Kassemeyer,

Nils Kimmel, Lukas Lomb, Filipe R. N. C Maia, Stefano Marchesini, Marc Messerschmidt,

Karol Nass, Duško Odic, Carlo Schmidt, Joachim Schultz, M. Marvin Seibert, Robert L.

Shoeman, Raymond G. Sierra, Heike Soltau, Dmitri Starodub, Jan Steinbrener, Fracesco

Stellato, Lothar Strüder, Martin Svenda, Herbert Tobias, Joachim Ullrich, Geog

Weidenspointneroll. "Single particle imaging with soft X-ray at Linac Coherent Light

Source". Proc. SPIE8078, Advances in X-ray Free-Electron Lasers: Radiation Schemes,

X-ray Optics, and Instrumentation, 807809; doi:10.1117/12.886754. (May 20, 2011).

[18] Ayhan Bingölbali and C. A. MacDonald, "Curved Crystal X-ray Optics for Monochromatic

Imaging with a clinical source" Medical Physics, Vol 36, No. 4, doi:10.1118/1.3083568,

(2009).

[19] Axel T. Brünger, "Free R-values: a novel statistical quantity for assessing the accuracy of

Crystal Structure". Letters to Nature, Nature 355, 472-475; doi:10.1038/355472a0, (1992).

[20] Haiguang Liu, Billy K. Poon, D. K. Saldin, John C. H. Spence and Peter H. Zwart. "3D

Single particle imaging using angular correlations from X-ray data", International

Union of Crystallography printed in Singapore; dio:10.1107/S0108767313006016, ISSN

0108-7673. (2011).

[21] S. S. Kim, S. Wibowo and D. K. Saldin. "Structure of Single Particles from randomly oriented ensembles", Internal Medicine Review, May 2016.

[22] D.A. Varshalovich, A.N. Moskalev, V.K. Khersonskii : "Quantum Theory of Angular Momentum", P72-P163, World Scientific Publishing Co. Pte. Ltd ISBN-13 978-9971-50-107-5. (2008).

[23] Alessio Del Bue, Marko Stostic, Marija Dodig and João Xavier. "2D-3D Registration of Deformable Shapes with Manifold Projection", Snstituto Superior Tecnoco Institute for Systems and Robotics. Av Povisco Pais 1-1049-001 Lisboa-Portugal. ISBN:978-1-4244-5655-0,doi:10.1109/ICIP.2009.5413705.(2009).

[24] J. Pagaran, S. Fritzsche and G. Gaigalas. "Maple procedures for the coupling of angular Momenta IX. Wigner D-functions and rotation matrices", Computer Physics Communications, ELESEVIER, doi:10.1016/j.cpc.2015.12.008, (2005).

[25] L. Lovisolo and E. A. B. Da Silva, "Uniform distribution of Points on a hyper-sphere with Applications to vector bit-plane encoding", IEE Proc. Vision Image Signal Process, doi:10.1049/ip-vis:200110361, (2001).

[26] D. L. D. Casper and A. Klug Cold. "Physical Principles in the construction of regular viruses". Spring Harbor Sym. Qunt. Biol. 27. 1-24, PMID:14019094, (1962).

[27] Yang Leng. "Materials Characterization. Introduction to Microscopic and Spectroscopic Methods", P54, 55. John Wiley & Sons (Asia) Pte Ltd. Copyright © (2008).

[28] www.iucr.org/dictionary/Cromer_Mann_Coefficients.

[29] K. Pande, M. Schmidt, P. Schwander, and D. K. Saldin. "Simulation on time-resolved Structure determination of uncrystallized biomolecules in the presence of shot noise",

Structural Dynamics 2. 024103;doi:10. 1063/1.4916980, (2015).

[30] Oszlanyi G. and Süto A. "An alternative structure determination procedure for high resolution (small molecule) structures", Acta Cryst. A61, 147-152, doi: 10.1107/S0108767304027746, (2005).

[31] Gabor Oszlanyi and Andras Süto. "Ab initio structure solution by charge flipping", Acta Cryst. A60, 134-141, doi: 10.1107/S0108767303027569. (2004).

[32] Statistics for Management and Economics 7th Ed. 184-189, Mathematical Statistics with Applications 4th Ed. 116-121.

[33] Knuth Algorithm for generating Poisson random number: Wikipedia.

[34] T. A. White, A. Barty, F. Stellato, J. M. Holton, R. A. Kirian, N. A. Zatsepin and H. N. Chapman. "Crystallographic data processing for free-electron laser sources", Acta Cryst. D69, 1231-1240; doi:10.1107/S0907444413013620, (2013).

[35] H. C. Poon, D. K. Saldin, "Use of triple correlation for the sign determinations of expansion Coefficients of symmetric approximations to the diffraction volumes of regular viruses", Structural Dynamics 2. 041716; doi: 10.1603/1.4922476, (2015).

[36] Perspective of Imaging of Single Protein Molecules with Present Design of European XFEL, NOTKESTRASSE 85 – 22607 HAMBURG: ISSN 0418-9833. (Aug 1, 2014).

[37] Luís B. Almeida. "An Introduction to the Angular Fourier Transform", Acoustics, Speech, And Signal Processing, 1993. ICASSP-93 IEEE International Conference on (Volume 3), ISBN:0-7803-7402-9, doi:10.1109/ICASSP.1993.319481, April (1993).

[38] Zvi Kam. "Determination of Macromolecule structure in Solution by Spatial Correlation of Scattering Fluctuations ", Macromolecules, doi:10.1021/ma60059q009, Sept (1977).

[39] V. L. Shneerson, A. Ourmazd and D. K. Saldin, "Crystallography without crystals. I. The

common-line method for assembling a three-dimensional diffraction volume from single-particle scatterint", Acta Crystallographica Section A, Volume 64, Part 2, dio:10.1107/S 01087673070676621, March (2008).

[40] David J. DeRosier, "Correction of high-resolution data for curvature of Ewald Sphere", Ultramicroscopy, Volume 81, dio:10.1016/S0304-3991(99)00120-5, March (2000).

[41] http://www.rscb.org/pdb/explore.do?structureId=2BUK.

[42] T. R. M. Barends, L. Fourcar, R. L. Shoeman, S. Bari, S. W. Epp, R. Hartmann, G. Hauser, M. Huth, C. Kieser, L. Lomb, K. Motomura, K. Nagaya, C. Schmidt, R. Strecker, D. Anielski, R. Boll, B. Erk, H. Fukuzawa, E. Hartmann, T. Hatsui, P. Poll, Y. Inubushi, T. Ishikawa, S. Kassemeyer, C. Kaiser, F. Koeck, N. Kunishma, M. Kurka, D. Rolles, B. Rudek, A. Rudenko, T. Sato, C.-D. Schroeter, H.Soltau, L. Strueder, T. Tanaka, T. Togashi, K. Tono, J. Ullrich, S. Yase, S. Wada, M. Yao, M. Yabashi, K. Ueda and I. Schlichting. "Anomalous signal from S atoms in protein crystallographic data from an X-ray Free-Electron laser", Acta Cryst, D69, 838-842, doi:10.1107/S0907444913002448, (2013).

[43] S. M. Vinko, O. Ciricosta, B. I. Cho, K. Engelhorn, H.-K. Chung, C. R, Brown, T. Burian, J. Chalupsky, R. W. Falcone, C. Graves, V. Hajkova, A.Higginbotham, L. Juha, J. Krzywinski, H. J. Lee, M. Messerschmidt, C. D. Murphy, Y. Ping, A. Scherz, W. Schlotter, S. Toeikis, J. J. Turner, L. Vysin, T and Wang, B. Wu. "Creation and diagonosis of a solid-density plasma with X-ray free-electron laser", Naure 482, 59-62; doi:10.1038/ nature 10746, 25 Jan (2012).

[44] International Tables for X-ray Crystallography Vol. IV, (Table 2.2B, P100), Cromer-Mann Coefficients.

[45] H. C. Poon and D. K. Saldin, "Beyond the crystallization paradigm: structure determination

from diffraction patterns of ensembles of randomly oriented particles", Ultramicroscopy 111, 788, doi: 10.1016/j.ultramic.2010.11.1003, (2011).

[46] www.rcsb.org/pdb/results/results.do?grid=72D2759B&tabtoshow=Current.

[47] Douglas L. Dorset Hauptman : Structural Electron Crystallography P96: Woodward Medical Research Institute Inc. Buffalo, New York ISBN 978-4757-6623-3, (1955).

[48] https://en.wikipedia.org/wiki/Atomic_form_factor.

[49] Michael J. Bogan, W. Henry Benner, Stefan P. Hau-Riege, Henry N. Chapman and Matthias Frank, "Aerosol sample preparation methods for X-ray diffractive imaging: Size-selected Spherical nanoparticles on silicon nitride foils", Journal of Aerosol Science, doi:10.1016/j. Jaerosci.2007.07.007,Vol 38, Nov (2007).

[50] D. P. Deponte, J. T. McKeown, U. Weirstall, R. B. Doak, and J. C. H. Spence, "Toward ETEM serial crystallography: electron diffraction from liquid jets", Utramicroscopy. 111, 824-7; doi:10.1016/j.ultramic. (2011).

[51] Michael M Woolfson,"An Introduction to X-ray Crystallography", 2nd Ed, Cambridge University Press, P48-50, ISBN:0 521 42359 7, (1997).

[52] J. R. Driscoll, D. M. Healy, "Computing Fourier Transform and Convolutions on the 2-Sphere", Advanced in Applied Mathematics, Vol 15, P202-250, doi:10.1006/aama. (1994).

[53] J. B. Pendry, "Low Energy Electron Diffraction: The Theory and Its Application to Determination of Surface Structure (Techniques of physics)", Imperial College, London, (Academic Press), ISBN-10:0125505507, ISBN-13:978-0125505505, (1974).

[54] N. D. Loh, M. J. Bogan, V. Elser, A. Barty, S. Boutet, S. Bajt, J. hajdu, T. Ekeberg, F. R. N. C. Maia, J. Schulz, M. M. Seibert, B. Iwan, N. Tlmneanu, S. Marchesini, I. Schlicting, R. L. Schoeman, L. Lomb, M. Frank, M. Liang, and H. N. Chapman. "Reconstructed 3D Fourier

intensities from randomly oriented single shot diffraction pattern", Phys. Rev. Lett 104, 239902; doi:10.1103/PhysRevLett, 104.225501, (2010).

[55] http://github.com/antonbarty/cheetah-old/wiki.

[56] LCLS beamtime SPI, proposal number L866.

[57] http://www6.slac.stanford.edu/news/2012-08-13-worlds-most-powerful-xray.aspx

[58] Tomas Ekeberg. M. Svenda, C. Abergel, F. R. C. Maia, V. Selzer, J. M. Claverie, M. Hantke, O. Jönsson, C. Nettelblad, G. van de Schot, M. Liang, D. P. DePonte, A. Barty, M. M. Siebert, B. Iwan, I. Andersson, N. D. Loh, A. V. Martin, H. Chapman, C. Bostedt, J. D. Bozek, K. R. Fergusson, J. Krzywinski, S. W. Epp, D. Rolles. A, Rudenko, R Hartmann, N. Kimmel, and J. Hajdu, "Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser", Phys. Rev. Lett 114, 098102, doi: 10.1103/PhyRevLett.114.098102, (2015).

[59] LCLS beamtime SPI, proposal number L866.

[60] https://cxidb.org/browse.html/cxidb-30.cxi/entry_1/image_1/data.

[61] van Heel M and Schatz M, "Fourier Shell Correlation threshold criteria", J Struct Biol, PMID: 16125414, doi:10.1016/j.jsb.2005.05.009, (2005).

[62] https://en.wikipedia.org/wiki/Fourier_shell_correlation.

[63] R. Neutze, R. Wouts, D. van de Spoel, E. Weckert and J. Hadju, "Potential for Biomolecular Imaging with femtosecond X-ray pulses", Nature 406, 752-757, doi:10.1038/35021099, (2000).

[64] https://www6.slac.stanford.edu/news/2012-08-13-worlds-most-powerful-xray.aspx

[65] F. Hajdu, "Analytic approximation for incoherent scattering X-ray intensities", Acta Cryst. A27, doi:10.1107/S0567739471000159, (1971).

[66] https://cas.web.cern.ch/cas/Switzerland-2014/Lectures/Braun.pdf.

Appendices

Appendix A

Find The Center of Each Diffraction Pattern

```matlab
function getCenterByKim(prd)
minI = 100 ; ps = size(prd,3) ; rcI = [] ; GC = [] ; % GC=Grand-Center ;
for p=1:ps
    PC = [] ;
    for c=1:30
        for r=1:63
            if(prd(r,c,p) > minI)
                %========================================================
                rcI = [] ; % rcI(1827x3)
                for co=35:63
                    for ro=1:63 rcI = [rcI;ro co prd(ro,co,p)] ; end
                end
                %========================================================
                mA = [] ;
                dI = abs(rcI(:,3) - prd(r,c,p)) ; % dI = difference
                mA = find(dI == min(dI)) ; % min Diffence Array

                cent = [] ; rc = 0 ; cc = 0 ; mcr = 0 ; mcc = 0 ;
                for chk=1:length(mA)
                    rc = (r + rcI(mA(chk),1))/2 ;
                    cc = (c + rcI(mA(chk),2))/2 ;
                    cent = [cent;rc cc] ;
                end
                % mcr=mean-center-row, mcc=mean-center-col
                mcr = mean(cent(:,1)) ; mcc = mean(cent(:,2)) ;
                PC = [PC;mcr mcc] ; clear cent ;
            end
        end
    end
    mPCr = 0 ; mPCc = 0 ; mPCr = mean(PC(:,1)) ; mPCc = mean(PC(:,2)) ;
    GC = [GC;mPCr mPCc] ; clear mPCr mPCc PC ;
end
save('GC.mat','GC') ;
```

Appendix B

Find C2, C3 through Point by Point

```matlab
function getC2C3bypts(Cpdp)
% Be careful ! ==> in the middle of a row, there are many 0's
[r c p] = size(Cpdp) ;    % r = 33, c = 360, p= 182
C2 = zeros(360, r, r); C3 = zeros(360, r, r) ; % C2,C3 = 360x33x33

for dphi = 0 : 359
    for q1 = 1 : r
        for q2 = 1 : r % decide dphi, q2, q1.
            % s1, s2 = 1x360 row-vectors : s3, s4 = 1x360 row-vectors :
            s12N = 0 ; s12 = [] ; s34 = [] ;
            %================================================
            for page = 1 : p
                s1 = 0 ; s2 =0 ; s3 = 0 ; s4 = 0 ; q = zeros(3,360) ;
                q(2,:) = Cpdp(q1,:,page) ;
                q(1,:) = circshift(Cpdp(q2,:,page),[0,-dphi]) ;
                q(3,:) = circshift(Cpdp(q2,:,page),[0, dphi]) ;
                s1 = q(1,:).*q(2,:) ; s2 = q(3,:).*q(2,:) ;
                s3 = q(1,:).*(q(2,:).^2) ; s4 = q(3,:).*(q(2,:).^2) ;
                s12 = [s12 s1 s2] ; s34 = [s34 s3 s4] ;
            end
            s12A = find(s12 ~= 0) ; s12N = 0 ;
            s12N = length(s12A)   ; clear s1 s2 s3 s4 s12A ;
            if(s12N > 0)
                C2(dphi+1,q1,q2) = sum(s12)/s12N ;
                C3(dphi+1,q1,q2) = sum(s34)/s12N ;
            end
            %================================================
        end
    end
end

save('c2.mat','C2');
save('c3.mat','C3');
```

Appendix C

Extrapolations into Beam Stops

```matlab
function [Rdp, Rmean] = RayTrace3(maxR)
% In this program, the ultimate goal is
% to find radius-optimal & const-optimal in Rayleigh-Fomula.(rop, cop)
minI = 10 ; qmax = 0.05076 ; rmx = 2 ; Rdp = maxR ;
rs = size(Rdp,1) ; cs = size(Rdp,2) ; ps = size(Rdp,3) ; Rmean = [] ;
dq = qmax/(rs-1) ; q = (0:rs-1)'*dq ;  E = zeros(rs,1) ; R = E ;

% VERY IMPORTANT ::: How many cells in A column do you use as references?
% VERY IMPORTANT ::: ex) fin = 30, fin = 20, fin =rs ... (Total qPts = 121)
% HERE : 30/121, 20/121, rs(=121)/121, ....(for Least-Squae-Fit-difference)
fin = 50 ;

% r(:,1)=Radius , r(:,3) = 16pi^2*rho = C
% r(:,2) = DistanceSquare From Rayleigh-values (Using Least Square)
r = zeros(101,3) ; r(:,1) = (270:370)' ;


for p=1:ps
    Rmean0 = [] ; Rmean1 = [] ;
    for c=1:cs

%=============================================================================
%*****************************************************************************
%=============================================================================
    E(:) = Rdp(:,c,p) ;
    if(max(E) > minI)
        a = find(E>0) ; st = a(1)   ; clear a ;
    for ri=1:length(r(:,1))
        %=================================================================
        % Rayleigh Fill-In
        for i=st:fin
            R(i)=(sin(q(i)*r(ri,1)) -
q(i)*r(ri,1)*cos(q(i)*r(ri,1)))^2/q(i)^6;
        end
        qER = [q(st:fin) E(st:fin) R(st:fin)] ;

        % tc = temporary column
        tc = E(st:fin) ; ed=length(tc) ;
        nzA = find(tc~=0) ; zA = find(tc==0);
        if(tc(end)==0) nzA=[nzA ; ed] ; zA(end) = [] ; end % ed=end
        y = interp1(nzA, tc(nzA), zA) ; tc(zA) = y ; clear nzA zA ;
        EA = 0 ; EA = dq*sum(tc) - 0.5*dq*(tc(1)+tc(ed)) ;
        %=================================================================
        % Find Area of Rc
        xmin = qER(1,1) ; xmax = qER(end,1) ;
        fun=@(x) ((sin(x.*r(ri,1))-(x.*r(ri,1)).*cos(x.r(ri,1))).^2./(x.^6);
```

```matlab
        RA = integral(fun,xmin,xmax) ; C = EA/RA ; qER(:,3) = C*qER(:,3) ;
        %=================================================================
        sum0 = 0 ; for j=1:ed sum0=sum0+(qER(j,2)-qER(j,3))^2; end
        r(ri,2) = sum0 ; r(ri,3) = C ;
    end

    rA = find(r(:,2) == min(r(:,2))) ; r_idx = rA(1) ; clear rA ;
    rop = r(r_idx,1) ; cop = r(r_idx,3) ; %radius-optimal, constant-optimal


%=========================================================================
% In the above, I found r(radius), C(constant) for Rayleigh-Formula
% Rayleigh-Fill-In <<AGAIN>> Using rop(radius-optimal),cop(const-optimal)
%=========================================================================
        % Tracing Starts : cf = column filling index
        % fill in upto rmx : Here, rmx = 2
        for cf=rmx:st-1
         E(cf) = cop*(sin(q(cf)*rop) - q(cf)*rop*cos(q(cf)*rop))^2/q(cf)^6;
        end
        Rdp(:,c,p) = E ;

    Rmean0 = [Rmean0; rop] ;
    end
%=========================================================================
%*************************************************************************
%=========================================================================
    if(length(Rmean0) ~= 0) Rmean1 = [Rmean1 ; mean(Rmean0)] ; end

    end
    Rmean = [Rmean ; mean(Rmean1)] ;
end
```

CURRICULUM VITAE

9/2008 – 12/2016   **University of Wisconsin Milwaukee**
- **Ph.D. Physics**
  Concentration in Condensed Matter Physics. BioXFEL imaging from Experimental Data

  Thesis: Reconstruction of 3D Image For Particles By The Method Of Angular Correlations From XFEL Data

1/2003 – 12/2006   **University of Wisconsin Milwaukee**
- **M.S. Mathematics**
  Concentration in Numerical Analysis

  Thesis: Solving PDE Model Problems using DOLFIN and MATLAB

9/1993 – 5/1998   **University of Wisconsin Whitewater**
- **B.S. Mathematics (Major)**
  Concentration in Statistics.

- **Computer Science (Minor)**
  Concentration in Simulations.

3/1982 – 2/1989   **SungKyunKwan University**
- **B.S. Physics (Major)**
  Concentration in Energy Band Solution in a Potential Well

- **5/1982-7/1983 (Military Service)**

EXPERIENCE

6/2013 – 12/2016	**University of Wisconsin Milwaukee**
	▪ **Research Assistant**
	(Develop Algorithms for Imaging Particles from the LCLS Experimental Data)
8/2008 – 5/2013	**University of Wisconsin Milwaukee**
	▪ **Teaching Assistant**
	(Teach Calculus and Non-calculus based General Physics, Heath Science, and All the related Labs)
9/2007 – 1/2008	**University of Wisconsin Washington County**
	▪ **Lecturer**
	(Teach Calculus based General Physics, Mechanics)
9/2003 – 1/2005	**University of Wisconsin Milwaukee**
	▪ **Teaching Assistant**
	(Intermediate Algebra for freshmen and Sophomores)
5/1989 - 1/1993	**Jail College Preparatory Institute**
	▪ **High School Math-II**
	(Algebra, Geometry, Numbers, Vector, Calculus I, II, Statistics)

More than 20 years of tutoring for math and physics.

USABLE COMPUTER LANGUAGE

- **Basic**
- **C++, C**
- **Assembly**
- **Cobol**
- **Fortran**
- **Pascal**
- **Python**

SOFTWARE

- **Matlab**
- **Mapple**
- **DOLFIN**