Structure of a single particle from scattering by many particles randomly oriented about an axis: towards structure solution without crystallization?

D. K. Saldin, V. L. Shneerson
Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, WI 53211

M. R. Howells, S. Marchesini
Lawrence Berkeley National Laboratory, Berkeley, CA 94720

H. N. Chapman
Center for Free Electron Laser Science, DESY, and Universität Hamburg, 22607 Hamburg, Germany

M. Bogan
SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025

D. Shapiro
Brookhaven National Laboratory, Upton, NY 11973, and Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Department of Physics, Arizona State University, Tempe, AZ 82851

Abstract. We demonstrate a complete algorithm for determining the electron density of an individual particle from diffraction patterns of many particles, randomly oriented about a single axis. The algorithm operates on angular correlations amongst the measured intensity distributions. We also demonstrate the ability to recover the angular correlation functions of a single particle from measured diffraction patterns.

E-mail: dksaldin@uwm.edu

PACS numbers: 61.05.cp,87.64.Bx,82.53.Ps

Submitted to: New Journal of Physics
1. Introduction

The reconstruction of a high-resolution image of a single particle from scattering by several symmetrically equivalent ones has been demonstrated recently [1]. We present here a solution to the more general case of scattering by many particles, randomly oriented about an axis. The sustained x-ray or electron fluence needed for image reconstruction would destroy all high resolution detail if focused on a single molecule, since resolution varies inversely as the third or fourth power of fluence [2]. By sharing this fluence over many identical molecules, crystallography both amplifies the scattered intensities, and allows reconstruction of the periodically averaged undamaged charge density of a molecule. Part of the price for this crystallographic redundancy is that the scattered signal is sampled only at the Shannon "frequency” of the complex amplitudes, resulting in loss of phase information. If the scattered intensities from a single molecule could be measured at their (finer) Shannon angular sampling rate, the structure of the scatterer could be determined by iterative phasing algorithms [3]. These advantages of signal amplification, damage reduction for high resolution, and access to oversampled intensities (allowing solution of the phase problem) may be combined if the structure of a single particle may be determined from diffraction patterns from many identical particles in random orientations. This is possible if scattering is recorded either from stationary particles or for a recording time less than the rotational diffusion time.

Consider x-ray scattering from identical molecules in solution. If the scattering is from radiation of pulse length longer than the rotational diffusion time, \( \tau \), of the molecules, the signal from each molecule will be its rotational average and signal amplification is provided by the addition of such rotational averages from all illuminated molecules, as exploited by the technique of small-angle x-ray scattering (SAXS). Although it was initially thought that only angularly averaged quantities such as the average size, and radius of gyration may be extracted from an angularly averaged \( I(q) \) curve, it has been demonstrated [4, 5] that, with some reasonable additional assumptions, even anisotropic details of the molecular shape may be found from measured SAXS data. Nevertheless, variations in a SAXS signal are greatest in the region of very small magnitude scattering wavevector, \( q \), where the signal is sensitive largely to the overall shape of the molecule, rather than its internal structure [6].

The increasing availability of intense short-pulse x-ray sources allows a measurement of a type not possible before, namely fluctuations of intensity about mean SAXS values from instantaneous snapshots of a relatively small number of identical particles, differing only in random positions and random orientations. This becomes possible with intense radiation pulses shorter than \( \tau \). The resulting instantaneous diffraction patterns will contain intensities which are not quite constant on rings of constant \( q \). The angular correlation functions of these intensity fluctuations contain information about the diffraction volume of the individual scatterers [7, 8]. More intense x-ray sources also allow a useful signal to be measured from higher values of \( q \), thus permitting access to higher resolution structure. Completely new possibilities for structure determination of
individual non-crystallized molecules become possible if such single-particle diffraction volumes may be deduced reliably.

Although this idea may be applied to e.g. completely randomly oriented molecules in a liquid [7], we examine here a case, where the orientations of the individual particles are assumed random only about a single axis parallel to the incident beam (or perpendicular the plane of the diffraction pattern, if a flat Ewald sphere is assumed). A possible application of this geometry is for the structure determination of e.g. a membrane bound protein *in situ* with the membrane perpendicular to the incident beam, where the individual proteins may be assumed identical in structure, but randomly oriented about an axis perpendicular to the membrane. This is a reasonable model of e.g. K channel proteins embedded in membrane, since their function depends on the presence of a pore (the K channel) that joins the interior of a cell to its outside.

2. Diffraction pattern from a rod with axis perpendicular to an incident x-ray beam

We illustrate the principle with an artificial object consisting of a right cylinder of uniform refractive index $n$, radius $a$, length $d$, and director, i.e. axis, $\mathbf{d}$. Under the Born approximation, the differential scattering cross section is

$$
\frac{d\sigma}{d\Omega} = \frac{\omega^4 a^2}{c^4} \frac{|n^2 - 1|^2}{(q \cdot \mathbf{d})^2} \left[q^2 - (q \cdot \mathbf{d})^2\right] (1 - \sin^2 \theta \cos^2 \phi) \times \sin^2 \left(\frac{q \cdot \mathbf{d}}{2}\right) J_1^2 \left(\frac{a}{c} \sqrt{q^2 - (q \cdot \mathbf{d})^2}\right).
$$

(1)

The beam was assumed incident along the $\hat{z}$ axis, with $\mathbf{q} = \frac{\omega}{c} [\sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + (\cos \theta - 1) \hat{z}]$, where $\theta$ and $\phi$ are the polar and azimuthal angles of scattering, and incident polarization along the $\hat{x}$ axis. The polarization term $1 - \sin^2 \theta \cos^2 \phi$ was summed over the two orthogonal polarizations of the far-field diffracted wave. Although, on our simple model, a typical membrane protein may be better approximated with $\mathbf{d}$ parallel to $\hat{z}$, for this simplified object, our principle is better illustrated with this axis taken in the $\hat{x} - \hat{y}$ plane. If such a cylinder is assumed to have its director $\mathbf{d}$ along the $y$ axis, its projection onto a plane perpendicular to the incident beam will be a rectangle with longer side parallel to the $y$ axis. The x-ray diffraction pattern from such an object would be expected to be a superposition of mutually perpendicular sinc functions, giving rise to sets of fringes parallel to the $x$ and $y$ axes around a central high-intensity region. This is precisely what is seen in the calculated diffraction pattern of Fig. 1. Note the particularly high-intensity band through the center of the diffraction pattern and parallel to the $q_x$ axis.
Structure of a single particle from scattering by many particles randomly oriented about an axis: towards

3. Diffraction pattern from many rods, randomly oriented about an axis perpendicular to the diffraction pattern

We next simulate a diffraction pattern resulting from x-rays incident on 10 such identical rods, related only by random orientations about an axis perpendicular to the diffraction pattern.
The pattern (Fig. 2) is clearly a superposition of diffraction patterns of the form of Fig. 1 with 10 bright horizontal central bands as seen in that figure randomly oriented about the center of the pattern. In this paper we investigate whether it is possible to develop an algorithm for recovering the single-particle diffraction pattern of Fig. 1 from many short-pulse multiparticle diffraction patterns of the form of Fig. 2 representing instantaneous positions of rods, randomly oriented about the z axis.

Although we have performed our simulations for 10 particles per diffraction pattern, essentially the same results are obtained for any (relatively small) number of particles. The only practical concern is that measurements of the angular fluctuations about the mean SAXS value on a particular resolution ring \( q \) are easier for a relatively small number of particles per diffraction pattern. However, to obtain convergent values for the correlations, it is necessary to average them over a relatively large number of diffraction patterns (100 in the case of our simulations). The experimental results of section 8, bear out these conclusions.

4. Circular harmonic expansion and angular correlation functions

We begin by representing the single-particle diffraction pattern (Fig. 1) by the circular-harmonic expansion

\[
I(q_x, q_y) = \sum_m I_m(q)e^{im\phi}
\]  

(2)

where \( I(q_x, q_y) \) represents the intensity of a point on the diffraction pattern of Cartesian reciprocal-space coordinates \((q_x,q_y)\) or equivalent polar coordinates \((q,\phi)\). In general, the coefficients \( I_m(q) \) are complex valued. It should be noted that since the circular harmonics \( e^{im\phi} \) form a complete orthonormal set of basis functions for the expansion of an arbitrary intensity distribution around a resolution ring of magnitude \( q \), the set of coefficients \( I_m(q) \) for appropriately-spaced resolution rings, \( q \), can represent any planar diffraction pattern. The reality of \( I(q) \) dictates that \( I_{-m}(q) = [I_m(q)]^\ast \). If the scattering angles are small enough that a flat Ewald sphere may be assumed, Friedel’s rule \( I(-q) = I(q) \) implies that only even values of \( m \) are nonzero in the sum on the RHS of (2).

Angular pair-correlation functions spanning pairs of resolution rings specified by \( q \) and \( q' \) are calculated in two steps. First, the data \( I(q,\phi) \) of each resolution ring \( q \) is mean subtracted to find

\[
I'(q,\phi) = I(q,\phi) - \langle I(q,\phi) \rangle_{\phi}.
\]  

(3)

For small values of \( q \), \( \langle I(q,\phi) \rangle_{\phi} \) may be identified with the usual small-angle x-ray scattering (SAXS) signal \( I_{SAXS}(q) \). We then define an angular pair correlation taken between the data on two resolution rings of radii \( q \) and \( q' \) of a diffraction pattern by \[7, 8, 9\]

\[
C_2(q, q', \Delta\phi) = \frac{1}{N_\phi} \sum_{\phi} I'(q,\phi)I'(q',\phi + \Delta\phi).
\]  

(4)
Likewise, one may define a triple correlation \([11]\) over the resolution ring \(q\) by

\[
C_3(q, q, \Delta\phi) = \frac{1}{N\phi} \sum_\phi [I'(q, \phi)]^2 I'(q, \phi + \Delta\phi). \tag{5}
\]

Their angular discrete Fourier transforms are defined by

\[
B_m(q, q') = \frac{1}{N\phi} \sum_{\Delta\phi} C_2(q, q', \Delta\phi)e^{-im\Delta\phi} \tag{6}
\]

and

\[
T_m(q, q) = \frac{1}{N\phi} \sum_{\Delta\phi} C_3(q, q, \Delta\phi)e^{-im\Delta\phi}. \tag{7}
\]

5. Algorithm for recovering a single-particle diffraction pattern from multi-particle ones

Correlation functions of the form \(C_2\) and \(C_3\) would be expected to be identical regardless of the orientation of a particle about an axis parallel to the incident radiation \([12]\). If multiple particles, randomly oriented about this axis, scatter incident radiation under conditions where interparticle interference is negligible, the resulting diffraction pattern would be a superposition of those of the individual particles, and thus \(C_2\) and \(C_3\) calculated from this composite diffraction pattern and averaged over many such diffraction patterns (we included data from 100 such simulated patterns) would be expected to converge on the corresponding quantities for a single particle. We may write

\[
I'(q, \phi) = \sum_j \sum_{m \neq 0} I_m(q)e^{im(\phi - \eta_j)} \tag{8}
\]

where the subtraction of \(\langle I(q, \phi)\rangle_\phi\) in (3) is equivalent to omitting the term \(m = 0\) in this summation, and where \(\eta_j\) is the (random) orientation of particle \(j\) about \(\hat{z}\).

Substituting (8) into (6) and (7), it may be shown that

\[
B_m(q, q') = N_p I_m(q)I_m^*(q') \quad \text{for} \quad m \neq 0, \tag{9}
\]

where \(N_p\) is the number of particles illuminated, and that

\[
T_m(q, q) = N_p I_m^*(q) \sum_{M \neq 0, m} I_M(q)I_{m-M}(q) \quad \text{for} \quad m \neq 0. \tag{10}
\]

The quantities on the LHS’s of (9) and (10) may be found from the angular Fourier transforms of pair- and triple-correlation functions computable from measured data of diffraction patterns from multiple particles. Since this is a quantity from measurement, and the quantities from the RHS are from theory, and since we seek only relative magnitudes of the circular harmonic expansion coefficients, \(I_m(q)\), we could simply remove the common scaling factor, \(N_p\), from Eqs. (6 and (10), and estimate

\[
|I_m(q)| = \sqrt{B_m(q, q)} \quad \text{for} \quad m \neq 0 \tag{11}
\]

\([B_m(q, q)\) is real and non-negative from (9)]. A full determination of these expansion coefficients requires also an estimation of the phases of \(I_m(q)\).
From (9), with the factor of $N_p$ omitted,
\[ I_m(q) = B_m(q, q')/I'_m(q'), \quad \text{for } m \neq 0. \tag{12} \]
Thus, the determination of the phase of $I_m$ for a single resolution ring (say $q'$) uniquely determines the phases of $I_m$ for the same value of $m$ for all other rings $q'$.

Indeed, substitution of (12) into (10) (with $N_p$ omitted) leads to
\[ T_m(q, q) = \frac{B_m^*(q, q')}{|I_m(q')|e^{i\phi_m(q, q')}} \times \sum_{M \neq 0, M \neq m} \frac{B_M(q, q')}{|I_M(q')|e^{-i\phi_M(M, q')}} \times \frac{B_{m-M}(q, q')}{|I_{m-M}(q')|e^{-i\phi_{m-M}(m-M, q')}}; \tag{13} \]
for $m \neq 0$. The only unknowns in (13) are the (real) phases $\phi'(M, q')$ of the circular-harmonic expansion coefficients, $I_M(q')$, for $M \neq 0$.

Once again, the quantity on the LHS of (13) may be found from experiment. In general there are $\mu$ unknown phases $\phi'(M, q')$ for a single resolution ring $q'$, where $\mu$ is the number of values of $M$ which give rise to distinct values of $I_M(q')$. In our simulation, we took circular-harmonic expansion coefficients $M$ up to a maximum value 38, so given the fact that $I^*_m(q') = I_M(q')$ and that $M$ takes up only even values, $\mu=19$. In general, these $\mu$ real numbers may be found by optimizing the function
\[ \sum_{m \neq 0} |T_m^{(obs)} - T_m^{(calc)}|^2 \tag{14} \]
by means of a global optimization algorithm such as simulated annealing [13] (where $T_m^{(obs)}$ are the quantities found from the angular Fourier transforms of the triple correlations $C_3$ from the measured diffraction patterns, and $T_m^{(calc)}$ their values estimated from the RHS of (13)). In the present simulation, for an individual object which gives rise to a diffraction pattern with mirror line, the origin of the angular coordinate may be taken about this line and the coefficients $I_M(q)$ taken to be real. Then, the only uncertainties are their signs. In this case, even an exhaustive search through $2^{19}$ sign combinations to optimize (14) took no more than a few minutes on a laptop computer. The value of $I_0(q)$ is estimated from the same data via $\langle I(q, \phi) \rangle_\phi$, the angular average of the intensities of resolution rings $q$. [Since $I_0(q)$ is real and positive, there is no need to determine its phase (or sign).]

6. Simulations

The importance of the determination of the phases $\phi'(m, q')$ (or in this case, the signs $s(m, q')$) via the triple correlations $C_3$ and (13) may be judged by the diffraction pattern (Fig. 3) simulated via (2) with the magnitudes of $I_m(q)$ determined by (11) and random signs and (2). This is seen to bear no relation whatsoever to the simulated single particle diffraction pattern of Fig. 1.

The step of determining the signs of $I_m(q)$ via the angular triple correlations, is crucial, as the reconstructed pattern bears no relation to even the central part of Fig. 1 for random signs of $I_m(q)$.
Figure 3. Part of the single particle diffraction pattern within the red resolution ring of Fig. 1 reconstructed from the magnitudes $|I_m(q)|$ of the circular-harmonic expansion coefficients determined from the pair correlations $C_2$ and random signs of the coefficients.

The result of reconstructing the single-particle diffraction pattern up to the resolution ring marked in Fig. 1 with the quantities $I_m(q)$ from the correlations $C_2$ and $C_3$ and SAXS data of 100 multi-particle diffraction patterns like Fig. 2 is shown in Fig. 4. A remarkable similarity to the model single-particle diffraction pattern of Fig. 1 is apparent. Simulated Poisson noise produces an unimportant peak at the origin of the correlation functions $C_2$ and $C_3$. For the parameters of our simulations, interference fringes from interparticle interference are finer than a detector pixel, and are thus irrelevant.

7. Reconstruction of projected electron density of a single rod from the recovered single-particle diffraction pattern

Finally, we recover (Fig. 5) from the diffraction pattern of Fig. 4 the projected electron density of the individual particle, after $\sim 15$ iterations of the charge-flipping algorithm [14]. An unambiguous reconstruction of the $80 \text{ nm} \times 25 \text{ nm}$ rod projection to the resolution of the reconstructed diffraction pattern ($\sim \pi/0.45 \simeq 7 \text{ nm}$) is found, without any assumption about compact support of the particle. The fidelity of the reconstruction may be judged by the fact that in the direction of the rod diameter (the short direction of the projected electron density, parallel to the $x$ axis) even a tapering off of the electron density towards the edge is seen, exactly as expected for the projection of a solid cylinder.
Structure of a single particle from scattering by many particles randomly oriented about an axis: towards structure

Figure 4. The single particle diffraction pattern within the red ring of Fig. 1 reconstructed from the magnitudes of $I_m(q)$ determined from the mean pair correlations $C_2$, and signs from the mean triple correlations $C_3$ from 100 multi-particle diffraction patterns like that of Fig. 2.

Figure 5. Real-space image of the 80×25 nm projection of the cylindrical object recovered to ~7 nm resolution from the reconstructed single-particle diffraction pattern of Fig. 4 after ~15 iterations of the charge-flipping algorithm [14].

8. Experimental

In order to test these ideas, we have collected soft-Xray transmission diffraction patterns at the Advanced Light Source (ALS) from gold nanorods (25 nm x 80 nm) lying on their side on a transparent silicon nitride membrane. To prepare individual isolated gold nanorods on the surface, a solution of nanorods was aerosolized via charge-reduced electrospray and electrostatically captured onto the silicon nitride membrane using
previously described methods [10].

Fig. 6 shows a transmission electron microscope (TEM) image of these rods, which differ only by rotation about the normal to the membrane. Rotation of one rod about the beam direction therefore rotates its diffraction pattern by the same amount. Using 750 eV highly coherent X-rays (1.65 nm wavelength), hundreds of diffraction patterns were collected from different 15-micron diameter regions, each containing 20-100 rods. In these patterns, the distance between rods is so large that coherent interference fringes resulting from different particles are much finer than the features from individual rods. The resulting patterns were analyzed by dividing the patterns into annular rings and forming the ring autocorrelation (RAC) functions defined in Eq. (4) with $q' = q$. We may now test the hypothesis that the sum of many such rings for a given $q$ over many patterns converges to the RAC functions for a single rod, using this experimental data. If so, the retrieval of a real-space image of one rod requires only solution of the phase problem twice - first to recover the diffracted intensity from its RAC functions, and second to recover the object charge-density from its diffraction pattern intensity distribution.

Fig. 7 shows a typical diffraction pattern, with the path of a ring indicated at L. The resolution of this ring at $q$ is about 34 nm. Figs. 8 and 9 show RAC functions computed around this path, with the abscissa range from 0 to 180°. Fig. 8 shows the result for the first RAC (lower panel, dashed line), obtained from a diffraction pattern which includes tens of randomly oriented rods. The continuous line (lower panel) shows

Figure 6. TEM image of gold nanorods, 25×80 nm in projection, lying on a SiN membrane 100 nm thick. In the scattering experiment whose results are reported here, the rods were not clumped, as shown here, but dispersed to be randomly oriented, with random interparticle distances.
a simulation of the RAC function for a single rod. The upper panel shows the intensity in the diffraction pattern plotted using a logarithmic intensity scale. In Fig. 9, 121 of these RAC functions (from different regions of sample) have been summed, and the result is compared again with the simulated RAC function from a single rod (lower panel). The experimental sum is seen to converge well to the model curve.

During these many hours of automated data recording (with computer-controlled sample stage motions for each new region) the intensity of the X-ray source steadily decreased.

9. Discussion and Conclusions

Although demonstrated in this paper only for an artificial object of rectangular projection, these results suggest the possibility of deducing the diffraction pattern of a single particle from the angular correlations of a diffraction pattern from the scattering of radiation from multiple identical particles in random orientations. If this were possible, it should also be possible to reconstruct an image of an individual particle via an iterative phasing algorithm applied to an oversampled reconstructed single-particle diffraction pattern.

Such a demonstration opens up the possibility of the following future extensions:
(a) The generalization of this method to particles of no particular projected symmetry, such as a general protein molecule.
Figure 8. Upper panel: diffracted intensity from a single x-ray diffraction pattern from a group of rods in random orientations. The ring L of Fig. 7 is represented by the red horizontal line in this plot, in which the diffraction pattern intensities are plotted on a logarithmic gray scale, with brighter regions representing higher intensities. The dark "finger" around $\phi=180^\circ$ is a projection of the beamstop holder (see Fig. 7). Lower panel: ring autocorrelation (RAC) function (dotted curve) from the experimental diffraction pattern of the upper panel, and from a simulated diffraction pattern of a single rod (more continuous curve).

(b) an application [11] to particles entirely randomly oriented in 3D, allowing the structure determination of a difficult-to-crystallize molecule from short-pulse diffraction patterns of such molecules in solution, provided scattering by solution atoms may be minimized or isolated.

c) The results are relevant also to the kind of “diffract and destroy” approach to biomolecular structure determination [15] proposed for the x-ray free electron laser (XFEL), where measurements of diffraction from single molecules have been proposed. In this case, the SAXS background, which arises from uncorrelated scattering by different particles, is entirely absent, as is scattering from solution atoms [8]. Alternatively, the ability to extract meaningful information from multi-particle diffraction patterns may help overcome intensity issues from single particle scattering.

d) The method is not restricted to diffraction patterns from scattering of externally incident radiation. A diffraction pattern may be measured also from x-rays arising from the fluorescence of a specific atom buried deeply within a large molecule [16] (e.g. the
Structure of a single particle from scattering by many particles randomly oriented about an axis: towards structure

Figure 9. Upper panel: intensity of diffraction pattern # 121, plotted as in Fig. 8. Lower panel: average RAC function from the data of many such diffraction patterns (dotted line). This is seen to approximate well to the corresponding calculated RAC function for a single rod.

Fe atom in haemoglobin). Such a pattern would be most sensitive to intramolecular scattering, less to that from its surroundings. This may allow the reconstruction of the 3D diffraction volume of an individual molecule from that of many randomly oriented ones, and hence of the molecular structure, without too much interference from solvent (or membrane) scattering.

Advances in technology, such as fast column read-out area detectors, brighter sources, shorter pulses, and zone-plate focusing, have greatly improved the experimental conditions for such experiments. When coupled with new theoretical advances, such as those reported here, entirely new vistas are opened up for the structure determination of microscopic particles.

10. Acknowledgment

DKS and JCHS jointly acknowledge support for this work from DOE Grant DE-SC0002141, UW, KES, and JCHS from NSF grant MCB-0919195, DKS from DE-FG02-06-ER46277 and the UWM Research Growth Initiative, JCHS from DE-FG03-02ER45996 and the NSF Center for Biophotonics Science and Technology. Part of the work was funded by the DOE at the SLAC National Accelerator Laboratory under
contract number DE-AC02-76F00515. The Lawrence Berkeley National Laboratory and
Advanced Light Source facility are supported by the Director, Office of Energy Research,
Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of
Energy, under Contract No. DE-AC03-76SF00098.

References

   Dosch H 2009 PNAS 106 11511.