

Structure from Fleeting Illumination of Faint Spinning Objects in Flight

*Russell Fung, Valentin Shneerson, Dilano K. Saldin, and Abbas Ourmazd**

Dept. of Physics

University of Wisconsin Milwaukee

1900 E. Kenwood Blvd, Milwaukee, WI 53211

** Author to whom correspondence should be addressed. Email: ourmazd@uwm.edu*

Supplementary Information

TABLE OF CONTENTS

- I. [Nomenclature](#)
- II. [Test Molecule](#)
- III. [Simulation of Diffraction Patterns](#)
 - a. [Latent space nodes & random molecular orientations](#)
 - b. [Noise-free diffraction patterns](#)
 - c. [Diffraction patterns with shot noise](#)
- IV. [Orientation Determination](#)
 - a. [The GTM algorithm](#)
 - b. [Use of information](#)
 - c. [Orientalional freedom about one axis](#)
 - d. [Orientalional freedom about three axes](#)
- V. [Placing the Diffraction Data on a Regular Cartesian Grid \(“Gridding”\)](#)
- VI. [Structure Recovery by Iterative Phasing](#)
- VII. [References](#)

I. Nomenclature

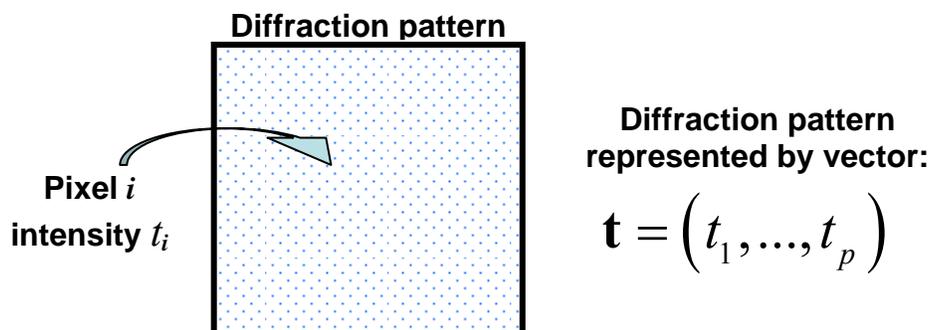
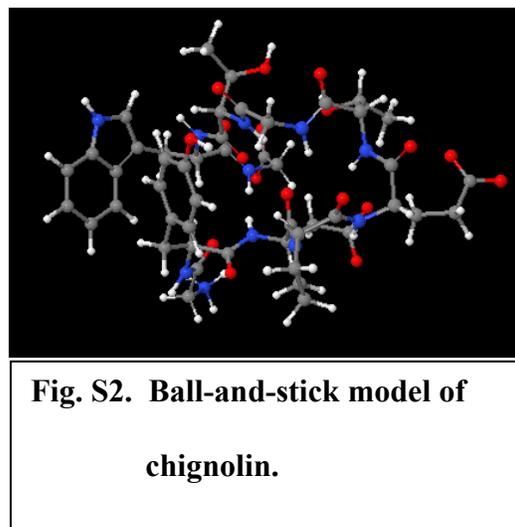


Fig. S1. Schematic of data representation. A diffraction pattern is represented by a vector, whose components are the measured intensity values at the p pixels of the diffraction pattern.

II. Test Molecule

We use chignolin, a 10-residue, 16Å diameter, 1kD synthetic protein with 77 non-hydrogen atoms. (PDB reference: 1UAO.) A ball-and-stick model of the protein is shown in Fig. S2.



III. Simulation of Diffraction Patterns

a. Latent space nodes and random molecular orientations

For the case, where the molecule is free to assume any orientation about one axis, the angular range of 360° was divided into 3° "bins," each representing a latent space node. Within each bin, 300 angles were randomly selected and used to generate diffraction patterns.

For the case, where the molecule is free to assume any orientation about three axes, each orientation can be represented by a point on the surface of the unit 4-sphere¹. The generation of an arbitrary number of points on a hypersphere with a uniform distribution is an open field of research². Algorithms for generating nearly uniform distributions are available. We have used that given in Lovisollo and da Silva³ to generate a nearly uniform distribution of nodes in latent space. A modification of this code was used to generate randomly distributed points on a contiguous patch of the unit 4-sphere, as follows. Choose a number $\ell : \{0 < \ell \leq 1\}$ and generate random numbers

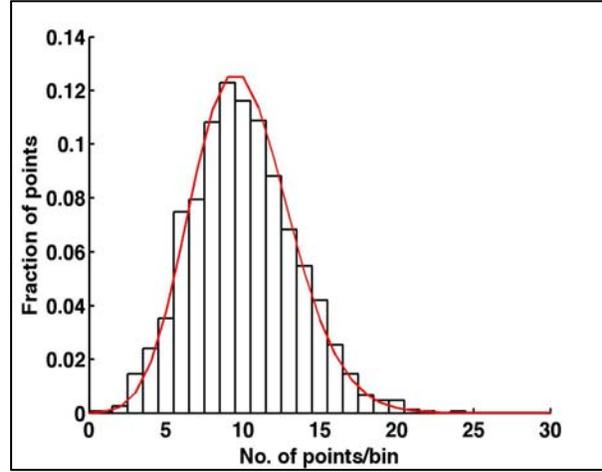
$$G, H, \omega_3 : \left\{ 0 \leq G \leq \frac{\pi}{2} \ell, -\ell \leq H \leq \ell, 0 \leq \omega_3 \leq 2\pi \ell \right\}. \text{ Triplets } (\omega_1, \omega_2, \omega_3), \text{ where}$$

$G(\omega_1) = (2\omega_1 - \sin 2\omega_1) / 4$ and $H(\omega_2) = \cos \omega_2$ are distributed randomly over a contiguous patch of the unit 4-sphere, with the size of the patch determined by ℓ .

The random distribution of the generated points was verified as follows.

- i. Using the method of Lovisollo and da Silva, 1500 latent space nodes on the surface of the unit 4-sphere were generated.
- ii. Using the algorithm above with $\ell = 1$, 15000 random points were generated on the surface of the unit 4-sphere.
- iii. For each of the points, the distances to all the nodes were calculated and the point assigned to the closest node.
- iv. The number of points N assigned to each node was tallied. The distribution of the numbers N was plotted as a histogram.

- v. The resulting histogram closely follows a Poisson distribution with mean 10, as expected from 15000 random points assigned to 1500 "bins" (Fig. S3).



b. Noise-free diffraction patterns

Diffraction patterns were generated from the square moduli of the molecular

$$F_{\mathbf{q}} = \sum_j f_j(q) \exp(i\mathbf{q} \cdot \mathbf{r}_j),$$

where \mathbf{q} is the scattering wavevector, and f_j the form factor of the atom at the position \mathbf{r}_j . The form factors were calculated from their parameterization in terms of Cromer-Mann coefficients⁴.

With the direction of incidence of the X-

rays parallel to the z -axis of a Cartesian coordinate system in the laboratory frame of

reference, the intensities $I_{\mathbf{q}} = |F_{\mathbf{q}}|^2$ were calculated for components of \mathbf{q} , $q_x = n\pi/L_x$,

$q_y = m\pi/L_y$, and $q_z = \left(\sqrt{k^2 - q_x^2 - q_y^2}\right) - k$, where k is the wavenumber of the X-rays, L_x and

L_y the corresponding expected dimensions of the scattering object, and m and n integers

Fig. S3. Distribution of generated molecular orientations. Histogram of 15000 points on the surface of the unit 4-sphere, generated as described in the text, and assigned to 1500 uniformly spaced sampling "bins." The histogram closely follows a Poisson distribution with mean 10, confirming the random nature of the points.

in the range -19 to 20. The resolution corresponding to the largest values of q_x (or q_y) is 1.8 Å. Each diffraction pattern stems from a random orientation of the molecule.

c. Diffraction patterns with shot noise

Shot noise was modeled by Poisson statistics. For a distribution with a mean λ , the probability that a non-negative integer k is observed is $f(k, \lambda) = e^{-\lambda} \lambda^k / k!$.

A random number generator drawing from a uniform distribution was used to generate random numbers from a Poisson distribution with mean λ as follows:

- i. Generate a random number $r : \{0 \leq r \leq 1\}$;
- ii. If $r < f(0, \lambda)$, select 0;
- iii. Otherwise, select p , where p is such that $\sum_{k=0}^{p-1} f(k, \lambda) \leq r < \sum_{k=0}^p f(k, \lambda)$.

To generate diffraction patterns with a mean photon count \bar{I}_{ref} at some specific reference resolution $|q_{ref}|$, proceed as follows:

- i. Generate noise-free diffraction patterns as described in section IIIb.
- ii. Calculate the mean photon count at the reference resolution $|q_{ref}|$ by averaging the intensities of all pixels within a thin shell with radius $|q_{ref}|$. This defines \bar{I}_{ref}^{calc} .
- iii. Rescale all diffraction pattern pixel intensities by the factor $\bar{I}_{ref} / \bar{I}_{ref}^{calc}$. Each pixel now contains the appropriate "expected" photon count.
- iv. Replace the expected photon count at each pixel by a number randomly drawn from a Poisson distribution with mean equal to the expected signal for the pixel.

We have verified that this approach generates the correct distribution by examining the signal statistics in simulated diffraction patterns. Specifically, we have established that in low-signal, high-diffraction vector regions, where Wilson statistics and shot noise dominate, the signal, averaged over a shell of constant diffraction vector modulus, is accurately described by Poisson statistics.

IV. Orientation Determination

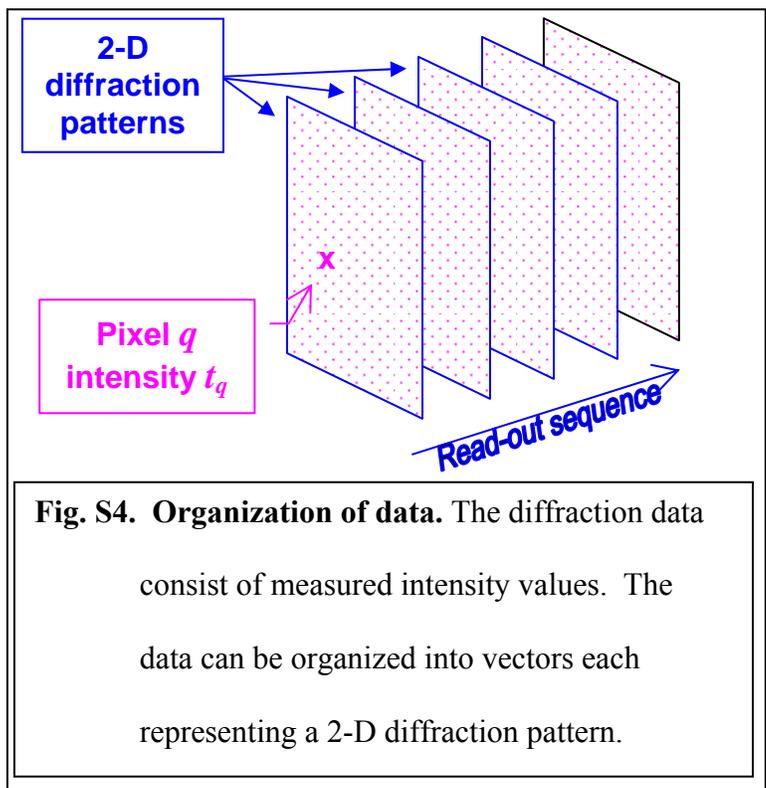
a. The Generative Topographic Mapping (GTM) algorithm

We use nonlinear Factor Analysis in its "Generative Topographic Mapping" (GTM)

implementation^{5,6} to embed an appropriate manifold in manifest space. Technical details on GTM can be found in the references

cited in our paper, especially those listed again here for convenience. Below, we provide a short overview of the approach.

Like Principal Components Analysis (PCA), Factor Analysis (FA) seeks to describe a multidimensional dataset in terms of a small number of variables. However, whereas



PCA identifies the "principal components" accounting for the *variances* in the data, FA identifies the "latent variables" responsible for the *covariances* ("correlations") in the data^{7,8}. In more detail, PCA analyzes the data distribution (cloud) to determine a series of orthogonal principal components, each describing the extension of the cloud of data in an orthogonal direction. In essence, PCA describes the variance of the data as an ellipsoid defined by orthogonal, i.e., uncorrelated principal components. FA, on the other hand, attempts to fit the correlations of the data given by the off-diagonal elements of the covariance matrix. Both PCA and FA are linear analytical approaches.

The relationship between the latent and manifest spaces in this work is nonlinear. We therefore need a nonlinear FA algorithm to reconstruct the 3-D diffracted intensity distribution (diffraction volume) from the correlations in the ensemble of scattered photons. More generally, the approach used in this work relies on embedding a low-dimensional manifold in a higher dimensional space. A number of powerful manifold embedding techniques are available, each with particular advantages. Here, we use GTM, primarily for three reasons:

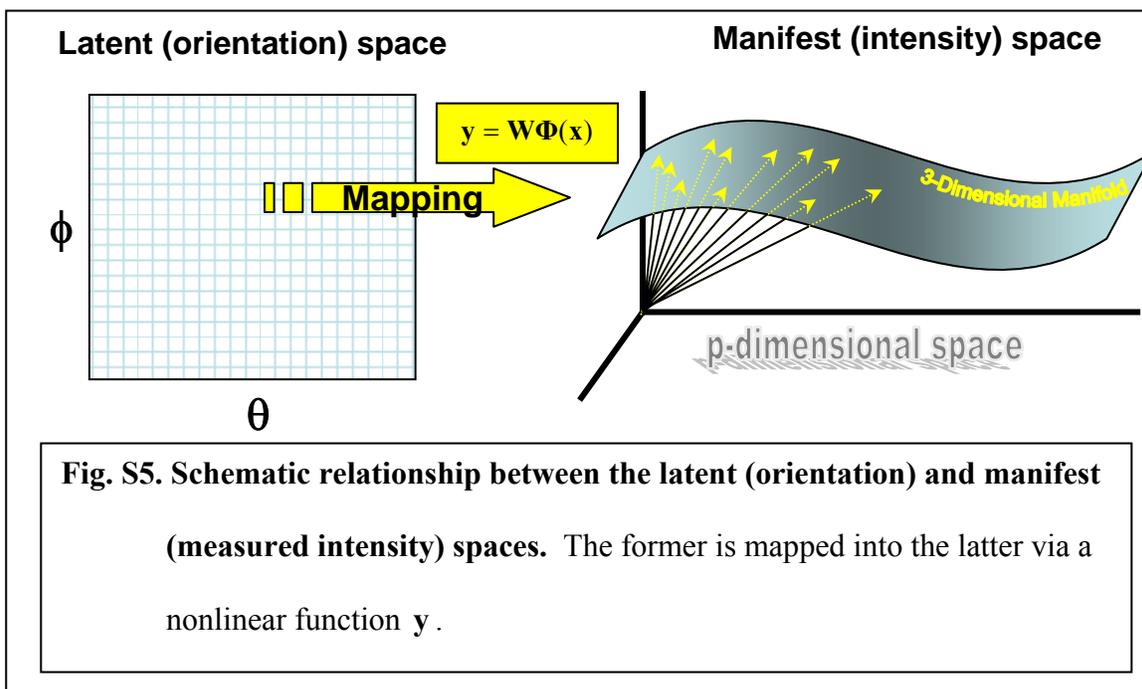
1. It naturally incorporates noise through a principled Bayesian approach;
2. It does not require the calculation of potentially noise-sensitive high-order derivatives; and
3. It allows one to specify a particular latent space (i.e., define the variables of interest) and a set of analytic basis functions.

However, other manifold embedding techniques may turn out to be more powerful in specific instances, for example when additional unknown dimensions are involved.

The ensemble of photons can be represented as a collection of vectors, each of which has components consisting of the measured intensity values at the pixels of a 2-D diffraction pattern snapshot (Fig. S4). Specifically, a snapshot can be represented by a p -dimensional vector $\mathbf{t} = (t_1, \dots, t_p)$, where t_q is the recorded intensity at pixel q , and p is the number of detector pixels. The entire scattered photon ensemble is then $\mathbf{t}_N = \{\mathbf{t}_1, \dots, \mathbf{t}_N\}$, where N is the number of snapshots in the dataset.

The latent variables (which specify the molecular orientations) cannot be directly accessed. However, the cloud of points $\mathbf{t}_N = \{\mathbf{t}_1, \dots, \mathbf{t}_N\}$ defines a three-dimensional manifold in p -dimensional space, which constitutes a nonlinear mapping of the latent orientation space into the manifest intensity space. This manifold reflects the dimensionality and geometry of the latent space (Fig. 2 of the article, reproduced here for convenience as Fig. S5).

The nonlinear mapping function \mathbf{y} can be expanded in terms of a basis set. In matrix notation, $\mathbf{y} = \mathbf{W}\Phi(\mathbf{x})$, where \mathbf{x} represents the set of latent variables, $\Phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \dots, \phi_M(\mathbf{x}))^T$, $\{\phi_m(\mathbf{x})\}$, $1 \leq m \leq M$ are the basis set, and the matrix \mathbf{W} contains the so-called loading factors, in effect the expansion coefficients describing the nonlinear mapping function \mathbf{y} in terms of the basis set.



For the present problem, we use a basis set of Gaussians of the form:

$$\phi_m(x) = \exp\left\{-\frac{1}{2}\left(\frac{x - \mu_m}{\sigma}\right)^2\right\}, \quad m = 1, \dots, M, \quad (1)$$

with a bias term of 1. When the molecule is free to assume any orientation about one axis, the distance metric takes the form:

$$|x_k - \mu_m| \equiv \min\left(|\theta_k - \tilde{\theta}_m|, 2\pi - |\theta_k - \tilde{\theta}_m|\right), \quad (2)$$

with θ representing the angular position about the axis of rotation. When the molecule is free to assume any orientation in 3-D space, we use a quaternion representation. This is because angles (e.g., Euler angles) do not provide a well-behaved distance metric. The latent space is that of unit quaternions, with molecular orientations ("latent space nodes") spanning the surface of the unit 4-sphere. Note that the latent space and basis functions are both periodic.

The latent space "node spacing" is chosen to insure that the largest minimum orientational difference, measured in terms of the internodal quaternion distance, does not exceed the appropriate sampling angle. Specifically, we insure that

$|q_1 - q_2| \leq |q_0 q_1 - q_1|$, where $q_{1,2}$ represent neighboring points on the surface of the unit 4-sphere, and q_0 the quaternion corresponding to the appropriate sampling angle. The GTM distance metric then becomes:

$$|x_k - \mu_m| \equiv \cos^{-1}\left(|q_k \cdot \tilde{q}_m|\right), \quad (3)$$

where q_k refers to quaternions representing the latent space nodes, and \tilde{q}_m the centers of the basis functions in latent space.

Once the mapping function \mathbf{y} is determined, the distribution of the measured vectors \mathbf{t}_N in manifest space can be related to its counterpart in latent space, the distribution of diffraction patterns in the latent space of molecular orientations, whence the diffraction volume can be reconstructed. GTM determines the nonlinear mapping function $\mathbf{y} = \mathbf{W}\Phi(\mathbf{x})$ by using the expansion coefficients ("loading factors") \mathbf{W} as fitting parameters to reproduce the internal correlations in the data. This is done by fitting a 3-D manifold to the data points ("vector tips") in the p -dimensional manifest space.

In slightly more technical terms, GTM is a Bayesian, nonlinear factor analytical algorithm, which:

- i. Establishes a link (a "mapping") between the latent and manifest spaces \mathbf{x} and \mathbf{t} in terms of a set of (nonlinear) basis functions;
- ii. Uses this mapping to embed the lower-dimensional space \mathbf{x} as a manifold in the higher-dimensional space \mathbf{t} ;
- iii. Adjusts the parameters of the mapping function to obtain a maximum likelihood fit to the distribution $\mathbf{t}_N = \{\mathbf{t}_1, \dots, \mathbf{t}_N\}$, and hence determine $p(\mathbf{t})$ (specifically $p(\mathbf{t} | \mathbf{x}; \mathbf{W}, \beta)$, when Gaussians of inverse variance β are used, with \mathbf{W}, β as the fitting parameters);
- iv. Uses Bayes' theorem to deduce $p(\mathbf{x} | \mathbf{t})$, which represents the 3-D diffracted intensity distribution in terms of the latent variables \mathbf{x} , given the measured dataset \mathbf{t}_N .

b. Use of information

A primary advantage of this approach is its use of the information in the entire data set, not the information contained in one, or a few low-signal diffraction patterns. This can be seen as follows. GTM uses the matrix \mathbf{T} , which contains the entire scattered photon ensemble:

$$\begin{aligned} \mathbf{T} &= (\mathbf{t}_1, \dots, \mathbf{t}_N)^T \\ &= \begin{pmatrix} t_{11} & \cdots & t_{1p} \\ \vdots & \ddots & \vdots \\ t_{N1} & \cdots & t_{Np} \end{pmatrix}, \end{aligned} \quad (4)$$

where \mathbf{t} represents a data vector. Using the nomenclature of Svensen, the key steps in GTM are, as follows.

- i. Fit the elements of the loading matrix $\mathbf{W}^T = (\Phi^T \mathbf{G} \Phi + \lambda \mathbf{I})^{-1} \Phi^T \mathbf{R} \mathbf{T}$.
- ii. Compute the responsibilities $r_{kn} = p(\mathbf{x}_k | \mathbf{t}_n; \mathbf{W}, \beta) = \frac{p(\mathbf{t}_n | \mathbf{x}_k; \mathbf{W}, \beta)}{\sum_{k'} p(\mathbf{t}_n | \mathbf{x}_{k'}; \mathbf{W}, \beta) p(\mathbf{x}_{k'})}$.
- iii. Compute the matrix elements $g_{kj} = \delta_{kj} \sum_{n=1}^N r_{kn}$, which are a sum over the entire scattered photon ensemble.

In this way, the information in the entire photon ensemble is used in every step of the Expectation-Maximization (EM) algorithm. Eq. **Error! Reference source not found.** also shows that it is immaterial whether, for example, the matrix \mathbf{T} with columns

representing diffraction patterns is used, or its transpose with columns representing the data at a given pixel in all the diffraction patterns (Fig. S4). Any other representation related to \mathbf{T} by a unitary operation may also be adopted. In our analysis, square roots of the pixel intensities formed the components of the data vectors.

c. Orientational freedom about one axis

To determine the effect of noise on the accuracy with which the molecular orientations can be deduced, two sets of 3000 diffraction patterns were generated with the molecule free to rotate about the x -axis, and the incident X-ray beam along the positive z -axis. The first set contained no noise, the second Poisson noise corresponding to a mean photon count (MPC) of 4×10^{-2} per pixel at 1.8 Å. The following GTM parameters were used to determine the orientations. Noise-free case: 120 uniformly spaced nodes, 60 uniformly spaced Gaussian basis functions, 600 pixels with highest variance across the dataset. Noisy case: 120 uniformly spaced nodes, 60 uniformly spaced Gaussian basis functions, 1200 pixels with highest variance across the dataset.

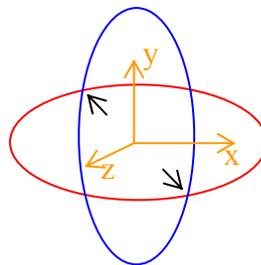


Fig. S6. Schematic of possible orientations with molecule able to rotate about one of two orthogonal axes. The two sets of diffraction patterns stemming from the molecule able to assume any orientation about the x - or y -axis contain identical diffraction patterns at the orientations indicated by the arrows. These can be identified by correlation analysis after an appropriate 90° rotation of the diffraction pattern.

To determine the 3-D structure of the molecule, two sets of 36000 diffraction patterns were generated at an MPC of 4×10^{-2} per pixel, with the molecule assuming random orientations about the x - or the y -axis, respectively, and the incident beam along the positive z -axis. Each of the x - and y -sets of diffraction patterns was analyzed by GTM in batches of 3000 randomly selected diffraction patterns, using the following parameters: 120 uniformly spaced nodes, 60 uniformly spaced Gaussian basis functions, 1200 pixels with highest variance across the dataset. The resultant diffraction node assignments were then used to average the diffraction patterns in the same orientational class, and the outcomes of the 12 batches provided to GTM again to determine the orientations of the class averages from the different batches with respect to each other. (With sufficient computer memory, all 36000 diffraction patterns can be aligned at once.) The same procedure was used separately to orient the 36000 diffraction patterns in the y -set. The absolute orientation ("common origin") of the x - and y -sets with respect to each other was determined by identifying the two pairs of most similar diffraction patterns in the x - and y -sets. Ideally, these pairs should be 180° apart (Fig. S6). The determined pair orientations were to within about one latent space node of this value. The values thus determined were used to identify a common origin and combine the x - and y -data sets.

d. Orientational freedom about three axes

Sets of up to 10^4 diffraction patterns with an MPC of 4×10^{-2} per pixel at 1.8 \AA were simulated for arbitrary molecular orientations over a $30^\circ \times 30^\circ \times 30^\circ$ patch of the surface of the unit 4-sphere. The following GTM parameters were used to determine the

orientations: 200 uniformly spaced nodes, 45 uniformly spaced Gaussian basis functions, 1044 pixels from an annulus excluding the central pixels and the pixels outside the radius defined by the width of the diffraction pattern.

As only relative orientations of the diffraction patterns are important, we plot the cumulative fraction of determined pair-wise orientational distances for the ensemble of diffraction patterns vs. the error in this determination. (The orientational error is measured in terms of the Euclidean distance in the space of unit quaternions, a well-behaved metric for orientational (dis)similarity.) We also compute an RMS error to characterize the orientational accuracy by a single number. Fig. S7 shows the orientational accuracy achieved for a set of 10^3 diffraction patterns from the test molecule in random orientations over a $30^\circ \times 30^\circ \times 30^\circ$ patch of the surface of the unit 4-sphere, at a signal level corresponding to 4×10^{-2} photons per pixel at 1.8 Å. The RMS error for the distribution is 5.2° .

GTM uses a Gaussian noise model. It is therefore at first sight surprising that it is able to determine the orientation of diffraction patterns to high fidelity in the presence of severe shot noise. We believe this follows from the Central Limit Theorem, but postpone a discussion to a later publication.

V. Placing the Diffraction Data on a Regular Cartesian Grid ("Gridding")

For each of the x - and y -series, the first round of GTM (on batches of 3000 "raw" diffraction patterns) produced 12 sets of 120 bin-averaged diffraction patterns. A second

application of GTM consolidated these further into 120 class-averaged diffraction patterns. Once the absolute origins between these two series were determined, momentum transfer vectors were assigned to each pixel of each of these 240 class-averaged diffraction patterns. These constitute the 3-D volume of diffraction intensities. Delaunay triangulation was used to obtain interpolated diffraction intensities on a uniform Cartesian grid. This substantially facilitates phasing, which involves forward and inverse Fast Fourier Transforms.

VI. Structure Recovery by Iterative Phasing

The phases associated with the gridded diffraction data, and hence the 3-D molecular electron density was determined by an iterative algorithm⁹, which successively imposes constraints/modifications of the electron density in real space through *object domain operations* (ODO)^{10,11} and in reciprocal space¹².

The 3-D Fourier transform of the gridded diffraction *intensities* yields the 3-D autocorrelation function of the molecular electron density. Since the extent of the autocorrelation map is twice that of the electron density map, the approximate spatial extent of the molecular electron density can be found directly from the diffraction intensities¹³.

A flow chart and pseudo code for our iterative phasing algorithm is shown in Fig. S8. The

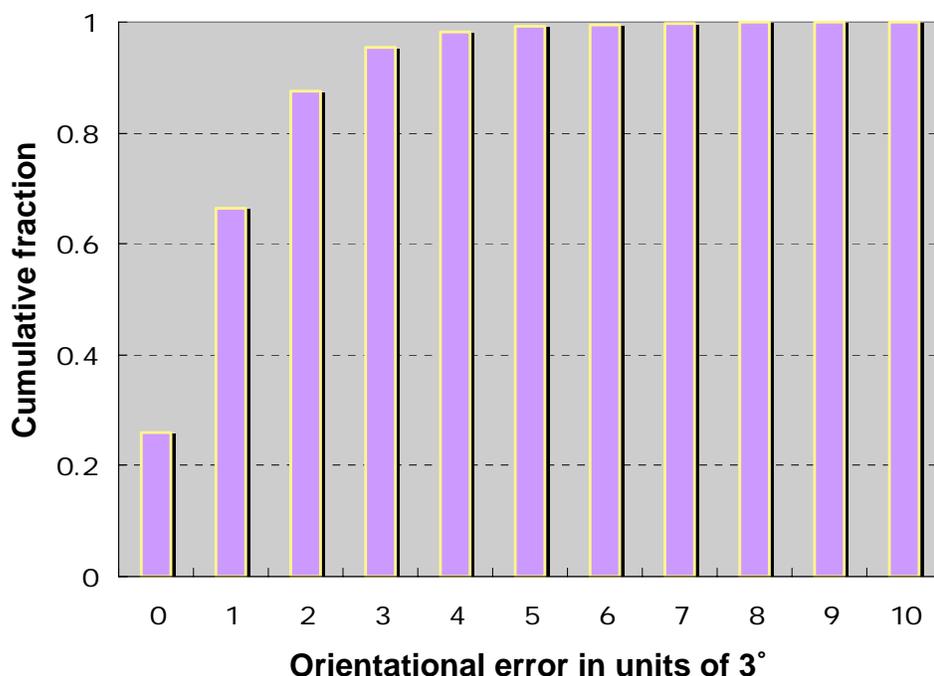


Fig. S7. Orientation recovery with molecule free to assume any orientation in 3-

D. Histogram of the accuracy with which the molecular orientations have been determined from a set of diffraction patterns, with the molecule free to assume any orientation over a $30^\circ \times 30^\circ \times 30^\circ$ patch of the surface of the unit 4-sphere representing all possible orientations in 3-D. The vertical axis shows the cumulative fraction of pair-wise orientational distances between the diffraction patterns aligned to within the orientational error shown on the horizontal axis. The mean photon count is 4×10^{-2} per pixel at 1.8 \AA , with shot noise. The root-mean-square orientational error is 5.2° , corresponding to $1.04 \times$ the orientational accuracy needed for structure recovery to 1.8 \AA resolution.

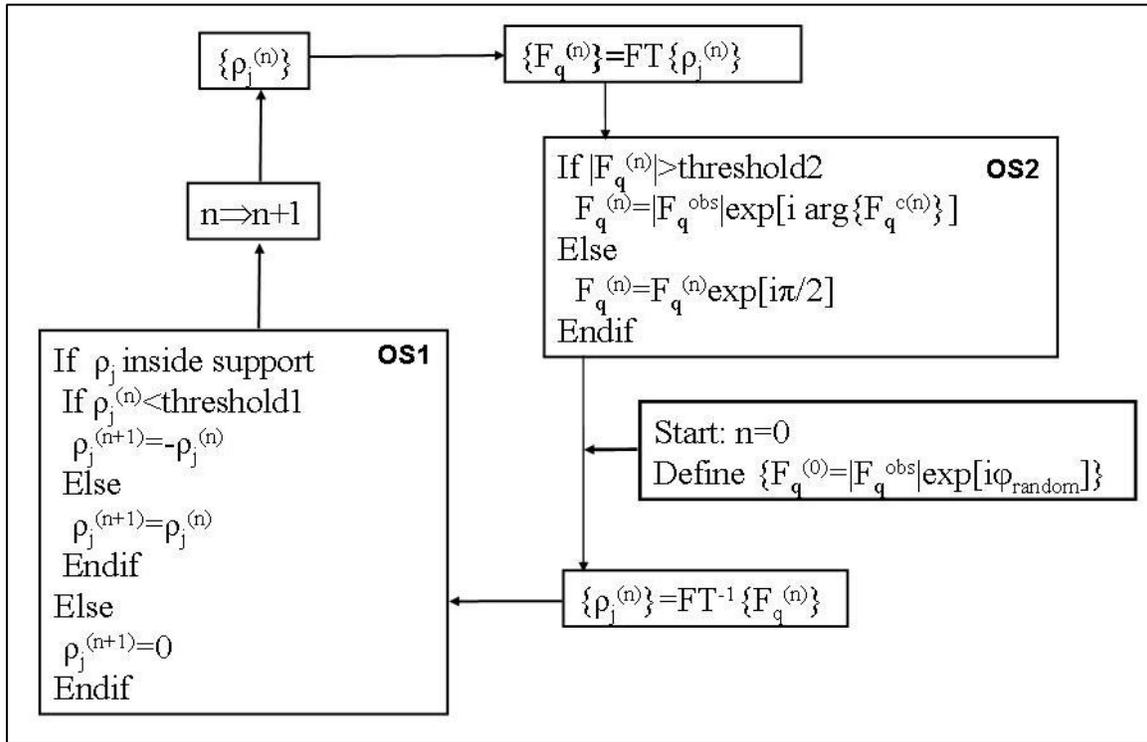


Fig. S8. Flow chart and pseudo code of the iterative phasing algorithm.

square roots of the gridded diffraction intensities are assumed to be proportional to the protein structure factors $F_{\mathbf{q}}$, where a reciprocal-space vector \mathbf{q} is defined by $\mathbf{q} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$ and the unit cell vectors \mathbf{b}_i ($i=1,2,3$) of the reciprocal space are defined by the usual relationships $\mathbf{b}_i \cdot \mathbf{a}_j = \delta_{ij}$ with respect to real-space unit cell vectors \mathbf{a}_j , so chosen as to define the 3-D volume expected to contain the molecule. Since the phases associated with these structure factors are initially unknown, the process begins by assigning random phases to those structure factors $F_{\mathbf{q}}$ corresponding to values of the Laue index $l \geq 0$. Assumption of Friedel's Law, $F_{-\mathbf{q}} = F_{\mathbf{q}}^*$ then allows the assignment of complex structure factors for $l < 0$. An (inverse) fast Fourier transform (FFT) algorithm

calculates an initial 3-D electron density distribution, whose reality (in the mathematical sense) is assured by the Friedel relationship among the structure factors. In general, the computed electron density is spread over a real-space volume larger than that of the molecule.

A support constraint is now applied in real space by setting to zero the electron density outside the volume expected to be occupied by the protein. In addition, the electron density within the expected volume of the protein is modified according to the *charge flipping* prescription of Oszlányi and Sütö, which was shown by Wu et al.¹⁴ to be a special case of Fienup's output-output algorithm¹⁵ with feedback parameter $\beta = 2$.

According to the charge flipping prescription, electron density values exceeding a certain threshold δ are unmodified, while the signs of those below this threshold are reversed.

The value of this threshold is chosen to optimize the progress of the algorithm, as monitored by an *R*-factor between the gridded "experimental" structure factors and those calculated from a Fourier transform of the electron density recovered by the algorithm. (In our calculations, we take δ to be 5% of the maximum electron density.) A Fourier transform of the modified electron density specifies the same distribution in reciprocal space. The continued reality of this modified electron density insures the resulting calculated structure factor phases satisfy Friedel's Law.

A different threshold is employed to divide the reciprocal-space amplitudes into *strong* and *weak* reflections. The magnitude of the threshold amplitude was again monitored by the same *R*-factor as for the real-space threshold above. The optimum division was found

when 25% of the weakest reflections were classified as weak. A reciprocal-space constraint is applied to the strong reflections: their amplitudes (or moduli) are replaced by the square roots of the corresponding *measured* intensities, while retaining the phases from the Fourier transform operation. As for the weak reflections, their moduli are left unchanged, but their phases are shifted by $\pi/2$. The resulting set of complex structure factors is then subjected to an inverse Fourier transformation, which yields another real-space electron distribution. This is modified in the same way as before, and the whole process repeated for several iterations.

This algorithm constrains the solution to be consistent with the measured intensities of the strong reflections in reciprocal space, and to the expected size of the object in real space. Subject to these constraints, it allows a thorough exploration of configuration space by iteratively modifying the phases of the weak reflections in reciprocal space and the signs of the small electron densities in real space.

When recovering an unknown structure from a set of measured diffracted amplitudes, the only way to obtain some indication of the progress of the algorithm is to monitor a quantity (e.g. a reliability-, or R-factor) characteristic of the agreement between the measured amplitudes and those calculated from the current estimate of the electron density. During the course of the iterations with our simulated data, the R-factor typically dropped rapidly within the first few iterations, reached a plateau around 0.37, remained there for about the next 100 iterations, and then suddenly dropped to a final value around 0.24. No further change occurred, even after 80 iterations. A similar behavior was

observed for most initial sets of random phases. When the R-factor did not display this behavior, e.g., when the R-factor remained high for 200 iterations, the resulting electron density was discarded.

VII. References

1. Kuipers, J.B., *Quaternions and Rotation Sequences* (Princeton University Press, Princeton, NJ, 1999).
2. See, e.g., Sloane, N. J. A., Hardin, R. H. & Cara, P., Spherical designs in four dimensions, *Proceedings of the Information Theory Workshop*, 253-257 (IEEE Press, 2003).
3. Lovisolo, L. & da Silva, E.A.B., Uniform distribution of points on a hyper-sphere with applications to vector bit-plane encoding, *IEEE Proc.* **148**, 187-193 (2001).
4. Cromer, D. T. & Mann, J. B., Atomic scattering factors computed from numerical Hartree-Fock wavefunctions, *Acta Cryst.* **A24**, 321-324 (1968).
5. Bishop, C.M., in *Learning in Graphical Models*, ed. M.I. Jordan, (MIT Press, 1999) 371-403.
6. Svensen, J. F. M., *The Generative Topographic Mapping* (Thesis, Aston University, 1998) and references therein.
7. Basilevsky, A., *Statistical Factor Analysis and Related Methods* (John Wiley & Son, New York, 1994).
8. Bartholomew, D. J. & Knott, M., *Latent Variable Models & Factor Analysis* (Oxford University Press, New York, 1999).
9. Miao, J., Hodgson, K., and Sayre, D., An approach to three-dimensional structures of biomolecules by using single-molecule diffraction images, *Proc. Natl. Acad. Sci., USA*, **98**, 6641-6645 (2001).
10. Fienup, J. R., Reconstruction of an object from the modulus of its Fourier transform, *Opt. Lett.* **3**, 27-29 (1978).

-
11. Oszlányi, G. & Sütő, A., *Ab initio* structure solution by charge flipping, *Acta Cryst.* **A60**, 134-141 (2004).
 12. Oszlányi, G. & Sütő, A., *Ab initio* structure solution by charge flipping. II. Use of weak reflections, *Acta Cryst.* **A61**, 147-152 (2005).
 13. Marchesini, S. et al., X-ray image reconstruction from a diffraction pattern alone, *Phys. Rev. B* **68**, 140101-1 - 140101-4 (2003).
 14. Wu, J.S. et al., Iterative phase retrieval without support, *Opt. Lett.* **29**, 2737-2739 (2004).
 15. Fienup, J. R., Phase retrieval algorithms: a comparison, *Appl. Opt.* **21**, 2758-2769 (1982).