

Single-particle imaging by x-ray free-electron lasers—How many snapshots are needed?

Cite as: Struct. Dyn. 7, 024102 (2020); doi: 10.1063/1.5144516

Submitted: 31 December 2019 · Accepted: 2 March 2020 ·

Published Online: 20 March 2020



View Online



Export Citation



CrossMark

I. Poudyal,  M. Schmidt,  and P. Schwander^{a)} 

AFFILIATIONS

Department of Physics, University of Wisconsin-Milwaukee, 3135 N. Maryland Ave., Milwaukee, Wisconsin 53211, USA

^{a)} Author to whom correspondence should be addressed: pschwan@uwm.edu

ABSTRACT

X-ray free-electron lasers (XFELs) open the possibility of obtaining diffraction information from a single biological macromolecule. This is because XFELs can generate extremely intense x-ray pulses that are so short that diffraction data can be collected before the sample is destroyed. By collecting a sufficient number of single-particle diffraction patterns, the three-dimensional electron density of a molecule can be reconstructed *ab initio*. The quality of the reconstruction depends largely on the number of patterns collected at the experiment. This paper provides an estimate of the number of diffraction patterns required to reconstruct the electron density at a targeted spatial resolution. This estimate is verified by simulations for realistic x-ray fluences, repetition rates, and experimental conditions available at modern XFELs. Employing the bacterial phytochrome as a model system, we demonstrate that sub-nanometer resolution is within reach.

© 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/1.5144516>

INTRODUCTION

X-rays have been used for more than sixty years¹ to determine the structures of proteins and other biologically important macromolecules. Protein structures are determined by the interpretation of electron density maps obtained from measured structure factors. Since the interaction of x-rays with matter is weak, crystals are widely used to determine these structure factors. When the crystals are exposed to x-ray radiation, diffraction is amplified along specific directions that are determined by Bragg's law. In this way, structure factor amplitudes can be measured with sufficient precision. However, the phase of the structure factors is not experimentally accessible and needs to be retrieved from additional experiments.² X-ray free-electron lasers (XFELs)³ can provide enormous incident intensities so that diffraction is sufficiently strong to access structure factor amplitudes from single particles. The particles are destroyed shortly after exposure with the incident radiation. However, XFEL pulses are short enough to collect a diffraction signal before the object is damaged. This is the so-called "diffraction-before-destruction" principle.^{4–6} In a single-particle imaging (SPI) experiment, a large number of two-dimensional (2D) diffraction patterns (snapshots) of single molecules are recorded by a pixel area detector. These snapshots are extremely noisy and taken in random and unknown orientations. Therefore, the orientations of the molecules relative to each other have to be determined from the snapshots. Multiple algorithmic methods of orientation recovery have been

developed to assign orientations to single-particle x-ray diffraction patterns.^{7–11} The oriented patterns are merged into a three-dimensional (3D) diffraction volume with phases initially unknown. The phases can be retrieved from the diffraction volume by iterative phasing.^{12,13} Finally, from the phased diffraction volume, the electron density map is determined.

Electron density reconstructions from experimental SPI datasets collected at XFELs have achieved resolutions in the regime of a few tens of nanometers.^{14,15} Diffraction up to a resolution of 5.9 Å was already observed,¹⁶ but a reconstruction of a 3D diffraction volume was not attempted due to the small number of diffraction patterns. This immediately raises the central question of how many diffraction patterns must be collected for a 3D reconstruction of the electron density for a given resolution. In other words, how many diffraction patterns are required to obtain a diffraction volume at a signal-to-noise ratio (SNR) sufficient to reach the desired resolution by iterative phasing? Clearly, this depends on the molecule under investigation and experimental conditions such as the wavelength, beam size, and incident x-ray fluence. To answer this important question is the primary focus of this paper.

Since the beamtime at XFELs is expensive, and sparsely available, it is important to have a sound estimate of the required number of diffraction patterns to design such an experiment. In contrast to XFELs, electron microscopes are now ubiquitously available. Using cryogenic

electron microscopy (cryo-EM), the structures of single molecules have been determined at near-atomic resolution,¹⁷ which surpasses the resolution reached at XFELs to date. However, the duration of the ultrashort x-ray pulses is faster than the molecular fluctuations. Accordingly, the XFEL provides a snapshot of a molecule “frozen in time” during x-ray exposure. Ambient temperatures are necessary to keep the molecules alive and enable protein dynamics. Using single-particle imaging at the XFEL, such dynamics can be probed with unprecedented time resolution down to the femtosecond regime. In contrast, cryo-EM requires quenching the sample to cryogenic temperatures, which may alter the structure,¹⁸ and with cryo-EM, the time resolution is limited to a few milliseconds.¹⁹

So far, the SPI techniques at XFELs were applied to large biological assemblies, primarily viruses.^{14,15,20} Here, we estimate by simulation how the SPI approach could be applied to a much smaller protein at a more relevant, molecular resolution. We selected the phytochrome, a light regulated enzyme, as a suitable model system. Phytochromes are red light photoreceptors characterized in plants, fungi, and bacteria and undergo large structural changes after red light absorption. The full-length, functional bacterial phytochromes (BphPs) consist of multiple domains. The PAS (Period ARNT Sim), GAF (cGMP phosphodiesterase/adenylyl cyclase/FhIA), and PHY (phytochrome-specific) domains form the photosensory core module (PCM).^{21–25} An effector domain has enzymatic activity, which is covalently linked to the PHY domain. The PHY domain has a

tongue-like structure, which contacts the GAF domain to seal the biliverdin (BV) chromophore pocket^{22,26} as shown in Fig. 1. Upon photoexcitation, phytochromes interconvert between a dark-adapted red-light absorbing state, Pr, and a photoactivated far red-light absorbing state, Pfr. The sensory tongue probes the configuration of the BV chromophore and transmits the signal to the PHY domain. The structure of the tongue undergoes substantial changes between the Pr and Pfr states.^{27,28} In the Pr state, the tongue assumes a loop to β -strand conformation, whereas in the Pfr state, it assumes a loop to α -helix conformation. Accordingly, large-scale conformational changes with amino acid displacements across several tens of Å between the Pr and Pfr states are required.^{28–30} However, the molecular details of the structural changes during the Pr to Pfr transition and their long-range effects on the effector domains are not well understood. Such changes may not be accommodated by the crystal lattice and thus hidden in crystallographic methods so that SPI is required. With the advancement in x-ray technology, we anticipate that single-particle x-ray experiments on the full-length phytochromes can be conducted at sub-nanometer resolution. With this, the structural dynamics of the Pr to Pfr transition in the full-length functional BphPs could be observed. Here, we estimate how many diffraction patterns are required to determine a 3D diffraction volume for an intact (full-length) phytochrome in the Pr form that can be successfully phased to calculate the three-dimensional electron density map at a targeted resolution.

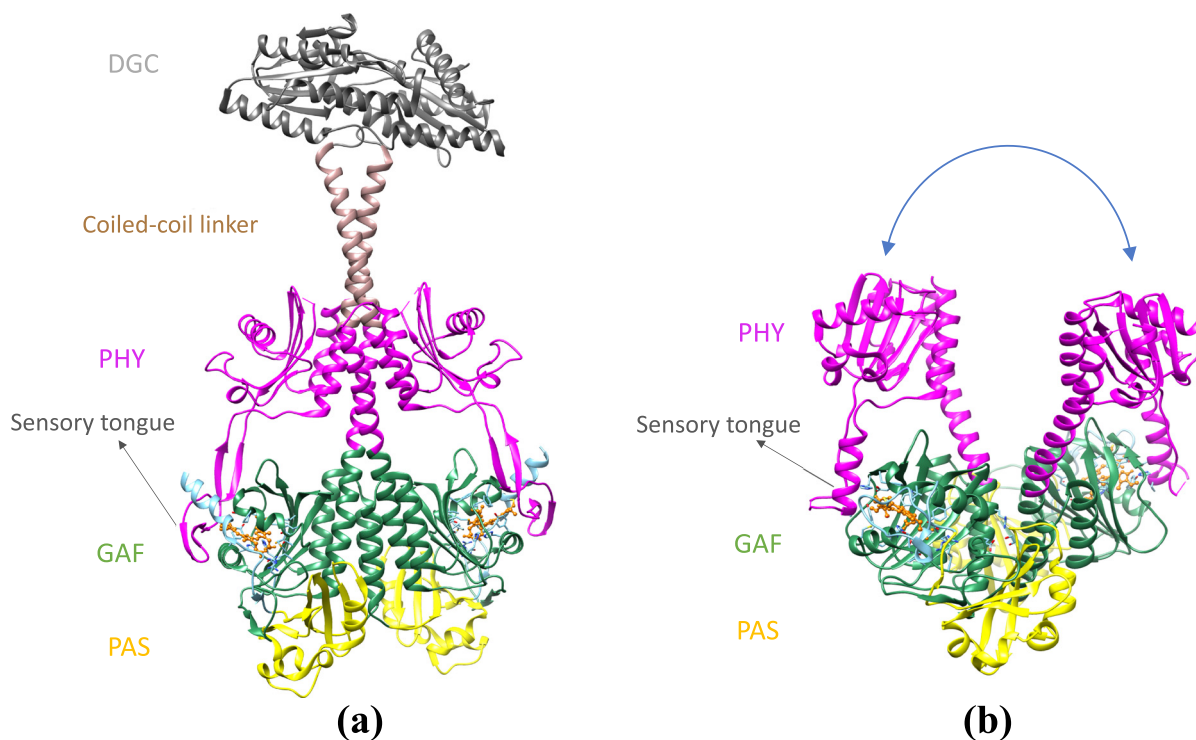


FIG. 1. Pr and Pfr structures of *Idiominarina sp.* and *D. radiodurans* phytochrome. (a) Full-length dark-adapted red-light absorbing Pr state²⁹ (pdb code 5llw) of the *Idiominarina sp.* phytochrome. Individual domains are colored in yellow, green, magenta, brown, and gray for PAS, GAF, PHY, and coiled-coil and di-guanylyl cyclase (DGC) effector domains, respectively. (b) The photosensory core module (PCM) from the *D. radiodurans* phytochrome in the photoactivated far red-light absorbing Pfr state²⁸ (pdb code 5c5k). The PAS, GAF, and PHY domain constitute the PCM and are represented with the same color code as in (a). The PHY domains are displaced substantially in the Pfr structure [blue curved arrow (b)]. The sensory tongue is marked in both structures, and the biliverdin (BV) chromophores (orange) are shown as ball-and-stick models.

METHODS

The simulations reported in this paper were done for the full-length *Idiomarina sp.* phytochrome molecule [protein data bank (PDB) entry 5llw] whose structure was recently determined.²⁹ This structure has an approximate diameter D of 164 Å and consists of about 11 000 atoms. Diffraction patterns were simulated according to the formalism in Appendix A and implemented in Python. The simulations were done for a photon energy of 2.48 keV corresponding to a wavelength of 5.0 Å. The resolution at the edge of the detector was 10 Å. The phases needed to recover the electron density can be retrieved by sampling the continuous diffraction pattern at sufficiently small intervals in reciprocal space. To retain the phase information, these intervals must be smaller than $1/(2D)$,^{31,32} i.e., oversampled at least twice with respect to the molecular diameter. This determines the size of the Shannon pixels in the diffraction pattern. For phytochrome, 73×73 detector pixels are required to reach a resolution of 10 Å. The simulated signal for each detector pixel was converted to the expected number of photons for an incident x-ray fluence of 10^{20} photons/cm² achievable at an XFEL. The measured photon counts follow Poisson statistics.³³ Accordingly, diffraction patterns were simulated by adding Poisson noise (“shot noise”) to the calculated diffraction signal.

For the simulation, randomly oriented diffraction patterns were generated using uniform random rotation quaternions as described in Appendix B. The (3D) diffraction volume of the molecule was obtained by orienting the noisy diffraction patterns relative to each other. To retrieve the electron density, the entire diffraction volume (reciprocal space) needs to be covered. Only scattering vectors ending on the Ewald sphere contribute to the diffraction pattern of a molecule in a particular orientation. Accordingly, each diffraction pattern accesses a spherical cap of the diffraction volume centered at the origin of the reciprocal space. Consequently, a large number of diffraction patterns (snapshots) from many different molecular orientations are required to fully sample the reciprocal space. The ensemble of all snapshots is then merged into a diffraction volume using a (cone-gridding) algorithm, which has been previously used for SPI at the XFEL.³⁴ The diffraction volume is phased by an iterative phasing algorithm.¹³ As a result, the phases of the structure factors are recovered, and the electron density is determined. The resolution is validated using Fourier Shell Correlation (FSC),³⁵ a method now widely accepted in cryo-EM single-particle imaging.

NUMBER OF SNAPSHOTS

Let D be the particle diameter and d be the aimed resolution. We define a dimensionless quantity $R =$ “number of resolution elements” as follows:

$$R = \frac{D}{d}. \quad (1)$$

The number of Shannon voxels in the outermost shell covered by a single diffraction pattern for oversampling by a factor of two is

$$nV_{\text{snapshot}} = \frac{2\pi(1/d)}{1/2D} = \frac{4\pi D}{d} = 4\pi R. \quad (2)$$

Accordingly, the number of Shannon voxels in the resolution shell is

$$nV_{\text{shell}} = \frac{4\pi(1/d)^2}{(1/2D)^2} = \frac{16\pi D^2}{d^2} = 16\pi R^2. \quad (3)$$

The probability of an outermost shell voxel hit by a randomly oriented diffraction pattern is therefore

$$p = \frac{nV_{\text{snapshot}}}{nV_{\text{shell}}} = \frac{1}{4R}. \quad (4)$$

Let $\langle n \rangle$ denote the mean number of expected photons per Shannon pixel of a diffraction pattern at the resolution shell. As single photons are counted by the detector, the signal follows Poisson statistics, $\text{var}(n^2) = \langle n \rangle$. Accordingly, the signal-to-noise ratio (SNR) is

$$\text{SNR} = \frac{\langle n \rangle}{\sqrt{\text{var}(n^2)}} = \frac{\langle n \rangle}{\sqrt{\langle n \rangle}} = \sqrt{\langle n \rangle}. \quad (5)$$

Due to the weak scattering of x-rays from a single molecule, the SNR of a single diffraction snapshot is way too low for any high-resolution information. It is therefore necessary to obtain information from many snapshots by averaging. The number of times M a voxel must be hit by a diffraction pattern in order to reach a desired SNR is

$$M \times \langle n \rangle = \text{SNR}^2. \quad (6)$$

As an example, for $\langle n \rangle = 0.002$ (phytochrome at 10 Å resolution), each voxel must be visited at least 500 times to achieve a SNR of 1.

The probability P for a single voxel visited at least M times by an ensemble of nS snapshots is estimated using the following sum of Binomial distributions:

$$P(p, M, nS) = 1 - \sum_{k=0}^{M-1} \binom{nS}{k} p^k (1-p)^{nS-k}. \quad (7)$$

Under the assumption that individual Shannon voxels are visited independently (justification given in Appendix C), the joint probability \tilde{P} to observe all voxels at least M times is

$$\tilde{P}(p, M, nS, nV_{\text{shell}}) = P(p, M, nS)^{\frac{nV_{\text{shell}}}{2}}. \quad (8)$$

The factor $\frac{1}{2}$ in front of nV_{shell} in the equation is due to Friedel’s symmetry. Using these relations, we can estimate the total number of diffraction patterns needed to cover the entire diffraction volume at any desired probability \tilde{P} . For the special case when $M = 1$, i.e., at very high signal levels, an estimation of the number of snapshots was proposed previously¹⁴ and is in agreement with our formulation.

Equation (7) cannot be analytically solved for nS . Instead, one can calculate the right-hand side with increasing nS until the desired probability P is reached. An implementation of an efficient algorithm in Python is listed in Appendix D. However, an analytical formula for the number of snapshots can be obtained by using the de Moivre–Laplace theorem, which approximates the binomial distribution by a Gaussian,

$$\binom{nS}{k} p^k (1-p)^{nS-k} \simeq \frac{1}{\sqrt{2\pi p(1-p)nS}} e^{-\frac{(k-pnS)^2}{2p(1-p)nS}},$$

$$\begin{aligned}
 P(p, M, nS) &= 1 - \sum_{k=0}^{M-1} \binom{nS}{k} p^k (1-p)^{nS-k} \\
 &= \sum_{k=M}^{\infty} \binom{nS}{k} p^k (1-p)^{nS-k} \\
 &\simeq \frac{1}{\sqrt{2\pi p(1-p)nS}} \int_M^{\infty} e^{-\frac{(k-pnS)^2}{2p(1-p)nS}} dk \\
 &= \frac{1}{2} \operatorname{erfc}\left(\frac{M-pnS}{\sqrt{p(1-p)nS}}\right).
 \end{aligned}$$

This approximation allows us to write nS explicitly,

$$nS(p, M, P) = \left(\frac{\sqrt{E(P)^2 + 4pM} - E(P)}{2p} \right)^2,$$

where $E(P) = \operatorname{erfc}^{-1}(2P) \sqrt{p(1-p)}$.

Since $\operatorname{erfc}(0) = 1$, the probability P becomes 0.5 for $nS_c = \frac{M}{p}$. We call nS_c the characteristic number of snapshots, the number required for a single voxel being visited at least M times with probability 50%. Together with Eqs. (4) and (6), this yields

$$nS_c = 4RM = \frac{4R \operatorname{SNR}^2}{\langle n \rangle}. \quad (9)$$

To verify the approximation, we calculated the exact probability P according to Eq. (7) as a function of $nS/4RM$ for different values of M . The result is depicted in Fig. 2(a). All curves admit a value of $P = 0.5$ for $nS/4RM = 1$, in close agreement with the approximation (9).

Now, for all voxels being jointly visited at least M times, the number of snapshots must be larger of course. With a joint probability

$\tilde{P} = 0.5$ in Eq. (8), a single-voxel probability $P = (1/2)^{2/nV}$ must be reached instead, a value much closer to one. We express the ratio of the exact number of snapshots nS_e , derived from Eq. (8), to the characteristic number of snapshots from Eq. (9) by a correction factor $C_f = nS_e/nS_c$. A plot of C_f as a function of the mean number of photons $\langle n \rangle$ for a SNR value of 1.0 is shown in Fig. 2(b). From this plot, we observe that nS_e approaches nS_c as the number of photons is lowered. The exact number nS_e can easily be estimated from the characteristic number nS_c by multiplication with the proper correction factor taken from the graph, without need to solve Eq. (8).

According to Eq. (9), the most important parameter for estimating the number of diffraction patterns is the mean number of expected photons per Shannon pixel $\langle n \rangle$ at the desired resolution d . Different methods to calculate $\langle n \rangle$ are given in Appendix E. An estimate of the number of snapshots for the full-length phytochrome molecule to reach a SNR of 1.0 at different resolutions and various experimental conditions is tabulated in Table I. The conditions are similar to the experimental conditions available at the existing XFELs such as the Linac Coherent Light Source (LCLS) or the European XFEL (EuXFEL).

RESULTS

Simulated diffraction patterns of the full-length phytochrome are presented in Fig. 3. Figure 3(a) shows a noise-free diffraction pattern, and Fig. 3(b) shows a diffraction pattern corresponding to an incident photon fluence of 10^{20} photons/cm² and a photon energy of 2.48 keV (wavelength 5 Å). Only ~200 photons are scattered from a single phytochrome molecule. Using our formalism [Eq. (8)], 38 000 diffraction snapshots are required to reach a SNR of 1.0 at 10 Å resolution. Accordingly, we simulated noisy 38 000 patterns that were subsequently merged into a 3D diffraction volume. Central slices through the reconstructed 3D volume are shown in Fig. 4(a), while Fig. 4(b) shows a section through the noise-free volume, derived from the

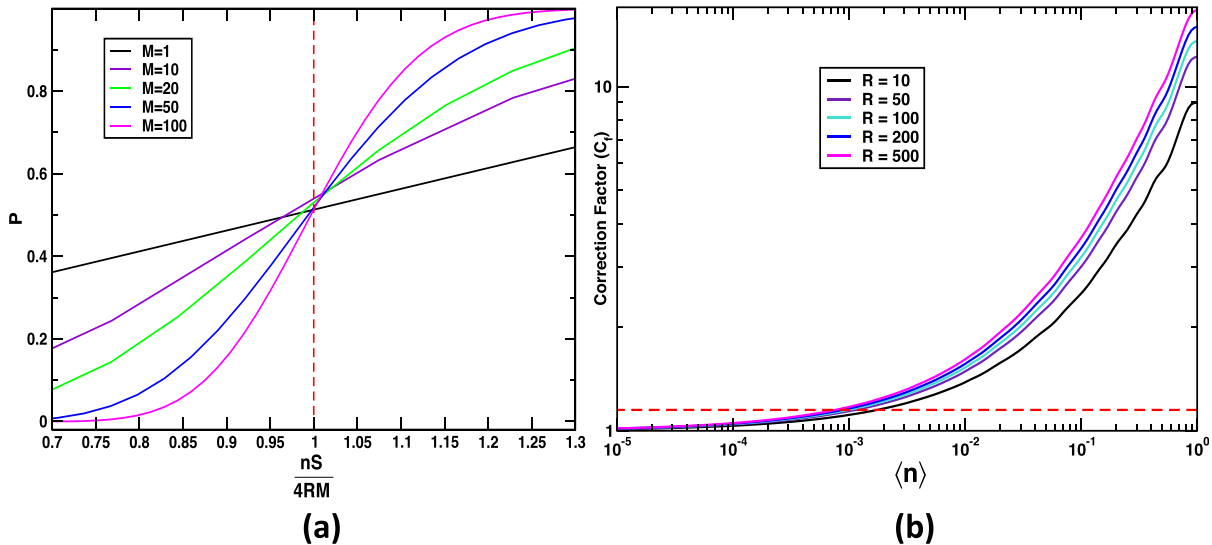


FIG. 2. Probability P and correction factor C_f . (a) Probability P of visiting a single voxel at least M times as a function of $\frac{nS}{4RM}$ ($R = 16$). For the characteristic number of snapshots $nS_c = 4RM$, all curves admit a probability P close to 0.5. (b) Log-log plot of the correction-factor $C_f = nS_e/nS_c$ as a function of $\langle n \rangle$ for different numbers of resolution elements R and SNR = 1.0. The horizontal dashed line corresponds to a value of $C_f = 1.15$. For $\langle n \rangle < 10^{-3}$, the exact number of snapshots nS_e is within 15% of the characteristic number nS_c .

TABLE I. Estimated number of snapshots required to reach SNR = 1 at different resolutions for the full-length phytochrome molecule at various experimental conditions. $\langle n \rangle$ is the mean number of photons per Shannon pixel at the desired resolution d .

Resolution	Soft x-ray energy = 2.48 keV, $\lambda = 5 \text{ \AA}$, beam size = $1.0 \mu\text{m} \times 1.0 \mu\text{m}$, and fluence = 10^{20} ph/cm^2		Hard x-ray energy = 8.0 keV, $\lambda = 1.5 \text{ \AA}$, beam size = $0.1 \mu\text{m} \times 0.1 \mu\text{m}$, and fluence = 10^{22} ph/cm^2		Hard x-ray (larger beam) energy = 6.0 keV, $\lambda = 2.07 \text{ \AA}$, beam size = $0.5 \mu\text{m} \times 0.5 \mu\text{m}$, and fluence = $4.0 \times 10^{20} \text{ ph/cm}^2$	
	$\langle n \rangle$	# Snapshots	$\langle n \rangle$	# Snapshots	$\langle n \rangle$	# Snapshots
30 \AA	1.8×10^{-2}	1774	1.0×10^{-1}	488	8.4×10^{-3}	3394
25 \AA	6.7×10^{-2}	5000	4.8×10^{-2}	1007	3.8×10^{-3}	8302
10 \AA	2.0×10^{-3}	38 244	2.6×10^{-2}	4323	2.2×10^{-3}	35 063
5 \AA	a	a	6.6×10^{-3}	26 978	5.0×10^{-4}	284 978
3 \AA	a	a	4.3×10^{-3}	66 075	3.5×10^{-4}	672 010

^aNot accessible due to the wavelength or unpractically high scattering angle.

atomic model. In the simulation, the central three voxels of the diffraction volume were set to zero, which takes into account the experimentally inaccessible central area of the detectors used at the XFELs. Iterative phasing was used to recover the electron density from the diffraction volume using the combination of the hybrid-input-output (HIO)¹² algorithm and shrink-wrap algorithm.¹³ The HIO algorithm was applied for the first fifty iterations with feedback parameter $\beta = 0.9$. After this, the shrink-wrap algorithm was used with an adaptive support constraint determined anew for each iteration cycle. For that, the present electron density was convoluted with a Gaussian of width σ . Voxels that contain electron densities larger than 14% of the maximum were assigned to the new support constraint. The initial width σ was set to six voxels and reduced by 5% after each iteration until a minimum of one voxel was reached. This algorithm converged

after a few hundred iterations. The reconstructed electron density at 10 \AA is shown in Fig. 5(a).

The resolution and reproducibility of the reconstructed electron density were accessed using Fourier Shell Correlation (FSC). For this, we randomly split the diffraction patterns into two disjoint sets “1” and “2” and processed each set independently, resulting in two electron density maps. The FSC is calculated from the Fourier transformation of the two maps using

$$FSC(q) = \frac{\sum_q F_1(\mathbf{q}) \cdot F_2^*(\mathbf{q})}{\sqrt{\sum_q |F_1(\mathbf{q})|^2} \sqrt{\sum_q |F_2(\mathbf{q})|^2}}, \quad (10)$$

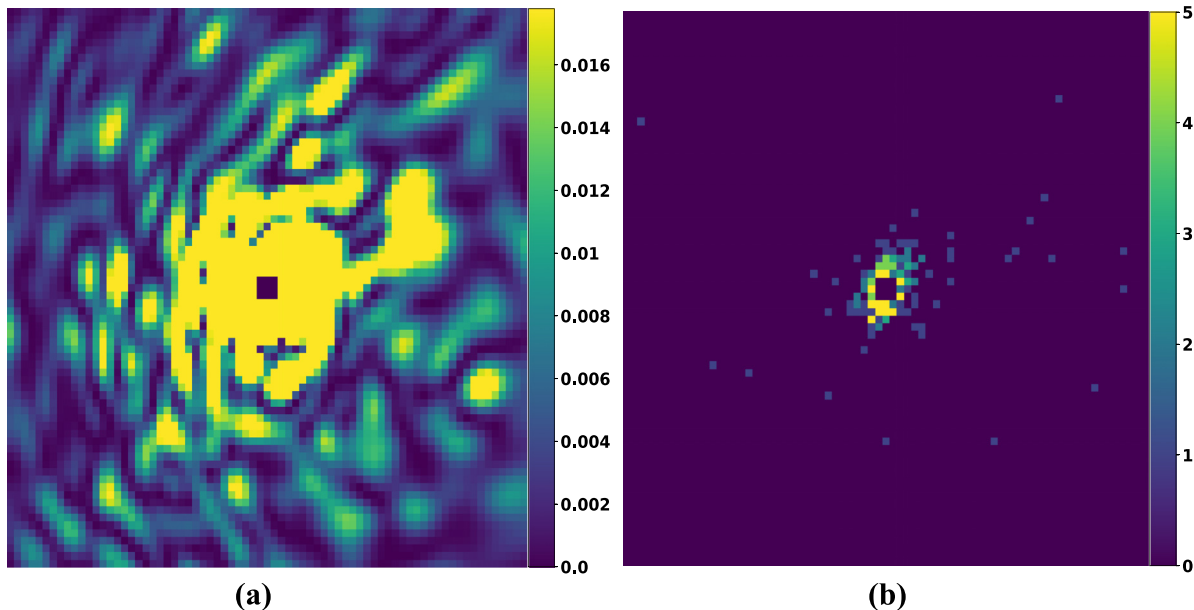


FIG. 3. Simulated diffraction patterns. (a) Noise-free diffraction pattern with 10 \AA resolution at the edge of the detector and 5 \AA wavelength. (b) Diffraction pattern for a photon fluence of $10^{20} \text{ photons/cm}^2$. The total number of scattered photons is 196, and the average photon count $\langle n \rangle$ at 10 \AA resolution is 0.002. The color code corresponds to the number of photons per pixel.

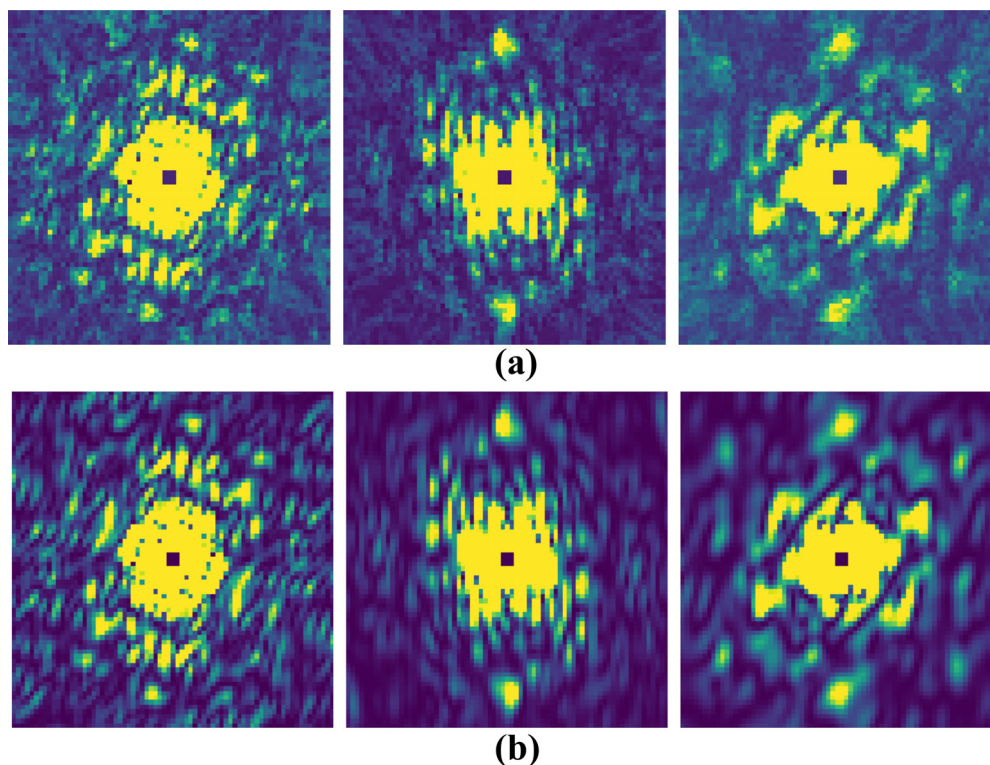


FIG. 4. Reconstructed and exact diffraction volume. (a) Central slices of the reconstructed volume, reconstructed from 38 000 noisy patterns. (b) Central slices through the exact, noise-free diffraction volume. The slices correspond to the xy , yz , and zx planes.

where $F_1(\mathbf{q})$ and $F_2(\mathbf{q})$ are the Fourier transforms of maps 1 and 2, respectively, q is the magnitude of the scattering vector, and $*$ denotes the complex conjugate. The resolution limit is defined by the value where the FSC drops below a certain threshold. Conventional thresholds used by the cryo-EM community are 0.143 and 0.5.^{36,37} These thresholds are represented by horizontal dashed lines in Fig. 5(b). The FSC drops below 0.5 at around 10 Å [Fig. 5(b), blue line]. This demonstrates that the number of diffraction patterns estimated by Eq. (8), with $\tilde{P} = 0.5$ and $\text{SNR} = 1$, is sufficient to reach the targeted resolution. Now, we want to test if a smaller number of snapshots could be sufficient to obtain the electron density at the same resolution. To address this, the reconstruction workflow is repeated with half the number of patterns (19 000), which corresponds to a SNR of 0.67. The FSC [Fig. 5(b), red line] reveals that instead of 10 Å, only about 20 Å is reached in this case.

To evaluate whether near-atomic resolution could be realistically reached by an XFEL experiment, the same pipeline is repeated for a target resolution of 3 Å. For that, a higher photon energy of 8.27 keV is used instead, which corresponds to a wavelength of 1.5 Å. A smaller x-ray focal spot with 100 nm diameter is chosen, which yields a photon fluence of 1×10^{22} photons/cm². A total of ~ 2000 photons/pattern are scattered per phytochrome molecule. According to our formalism, 66 000 noisy patterns are required to reach a resolution of 3 Å. The reconstructed electron density at 3 Å is shown in Fig. 5(c). Details of the structure can be identified from the inset of Fig. 5(c). We also validated the resolution of 3 Å by FSC [Fig. 5(d)]. With half the number

of patterns ($\text{SNR} = 0.65$), the resolution reaches only about 8 Å [Fig. 5(d), red line].

Finally, we analyze the effect of different types of backgrounds on the quality of the reconstruction. For that, we consider a uniform background with the same magnitude as the phytochrome signal at 10 Å and a uniform background three times the magnitude. Additionally, we take a q -dependent background assuming a helium gas, using the atomic scattering factors of He. The magnitude of the helium background is set equal to the signal of the phytochrome at 10 Å. The diffraction patterns, including the background, are converted to photon counts by addition of Poisson noise as described above. We repeat the merging and reconstruction processes to obtain 3D electron density maps. The FSC with a background is depicted in Fig. 6 and compared with the FSC without any background. Addition of a uniform background equal to the molecular signal (Fig. 6, red line) and the helium background (magenta line) has a small effect, and the target resolution of 10 Å can still be reached with 38 000 diffraction patterns. For a background three times the molecular signal, however, the resolution is reduced to 12 Å (blue line). We conclude that a background comparable to the molecular signal does not need substantially more diffraction patterns to reach the target resolution.

SUMMARY

We estimated the appropriate number of snapshots required in order to reconstruct the three-dimensional electron density of a biological molecule at any desired resolution. Quantitative results are

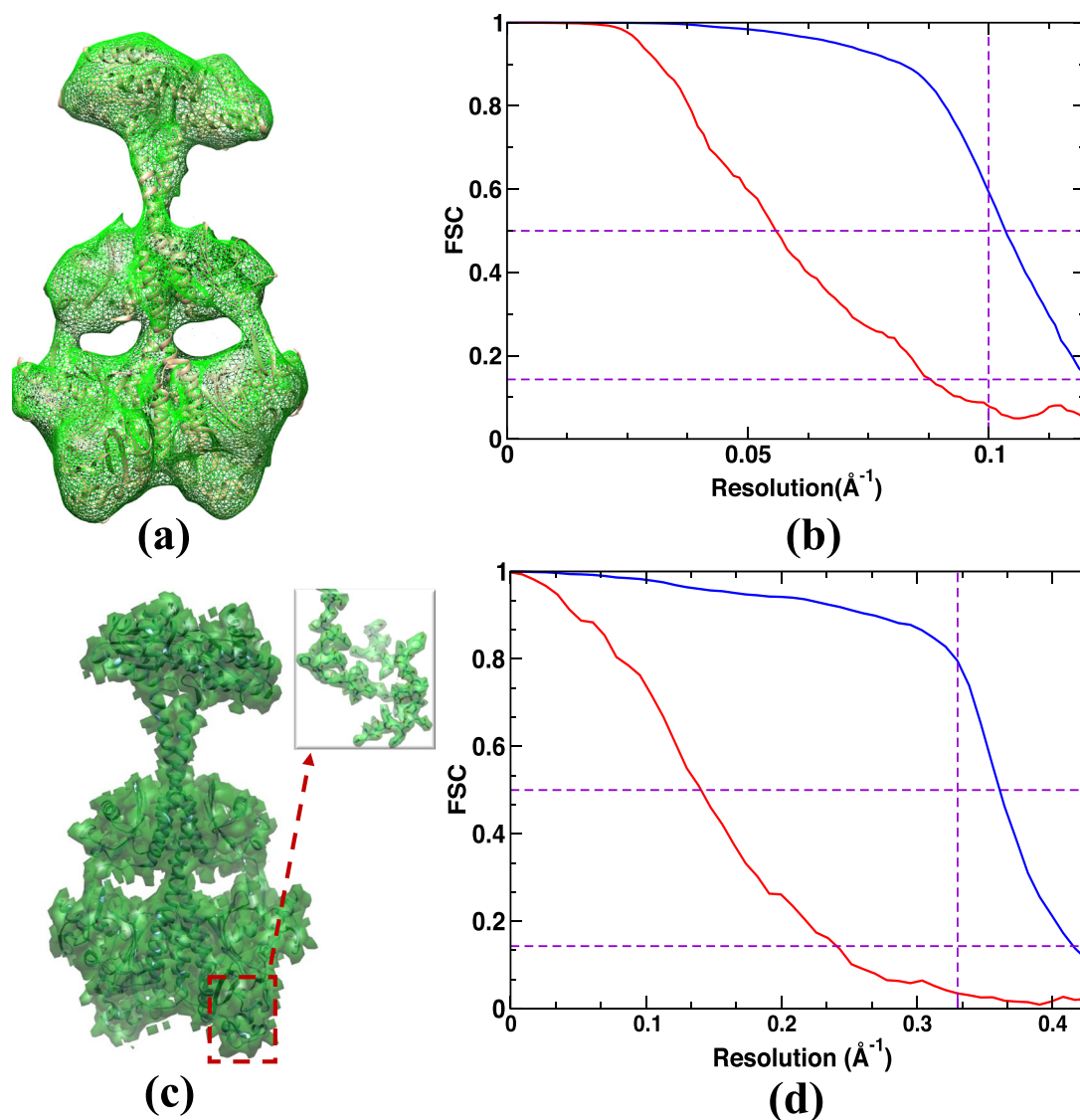


FIG. 5. Electron density obtained by iterative phasing and resolution validation by Fourier Shell Correlation (FSC). (a) Reconstructed electron density from 38 000 noisy diffraction patterns of the full-length phytochrome targeted at 10 Å resolution displayed at the 3σ contour level with atomic model superimposed using Chimera.⁴⁷ (b) FSC from (a) for all diffraction patterns (blue) and half the number of patterns (red). (c) Reconstructed electron density from 66 000 noisy diffraction patterns targeted at 3 Å resolution and beam parameters different from (a), as explained in the text. (d) FSC from (c) for all diffraction patterns (blue) and half the number of patterns (red). The horizontal dashed lines show the established thresholds for FSC (0.143 and 0.5). The vertical dashed line corresponds to the target resolution.

derived as a function of the desired resolution, the molecular size, the expected average number of photons per Shannon pixel, and the SNR in the resolution shell. Using this formalism, we demonstrated that a SNR of 1.0 and a joint probability $\tilde{P} = 0.50$ are sufficient to reach the desired resolution by iterative phasing validated by Fourier Shell Correlation (FSC). The derivation assumed that orientation recovery does not require additional snapshots. Indeed, a study of single-particle diffraction imaging recently reported successful reconstruction of electron density from diffraction patterns at a signal level of less than 100 photons on the average pattern.³⁸ We, therefore, conclude that for a protein like the phytochrome, where

200 photons per diffraction pattern can be expected in an XFEL experiment (see Fig. 3), additional snapshots for orientation recovery are not required.

Our formalism can be extended to incorporate other experimental conditions. In this work, we studied the effect of the uniform background and q -dependent background from a gas of helium. However, any form of background or other nuisances such as x-ray streaks and variations in the detector response^{39–41} can be incorporated. By including such signals in the simulations, the impact on the required number of snapshots can be evaluated following the same framework outlined in this paper.

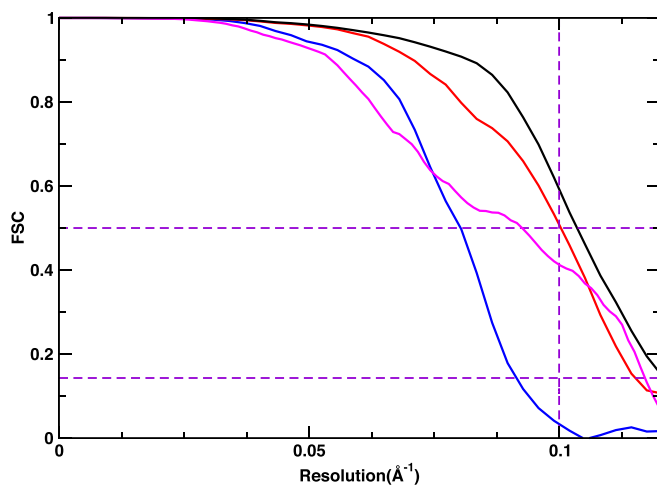


FIG. 6. Effect of different types of backgrounds on the resolution. Fourier Shell Correlation (FSC) for a uniform background with the same magnitude as the phytochrome signal at 10 Å (red) and a uniform background three times the magnitude (blue). The FSC of a q -dependent background representing a gas of helium with a magnitude equal to the signal at 10 Å (magenta). The FSC without a background is included for comparison (black).

OUTLOOK

The primary challenge of Single-Particle Imaging (SPI) at XFELs is to reach sub-nanometer resolution to visualize the atomic details of biological macromolecules. Most importantly, a sufficiently large number of single-particle diffraction patterns must be collected during the allocated experimental beamtime. With the new generation of high repetition-rate XFELs, which deliver pulse energies of 10 mJ and higher, improvement in sample delivery technology,⁴² and specialized detectors,⁴³ we expect to collect tens of millions of snapshots during a single shift. Combined with noise-robust data analysis algorithms for single-particle detection and orientation recovery,^{11,15} the goal may be reached in the near future. According to our estimates (see Table I), it should be possible to reach sub-nanometer resolution even for smaller proteins with molecular masses similar to that of the phytochrome. However, the main advantage of SPI unfolds in the presence of structural variability, associated with the biological function. Manifold-based machine learning algorithms applied to a large ensemble of single-particle snapshots allow us to reveal the concerted structural changes exercised by the sample.¹⁵ This enables us to map conformational spectra together with energy landscapes, determine possible functional pathways, and compile 3D molecular movies. Performing single-particle diffraction in time-resolved mode can further advance these promising opportunities for understanding the biological function and bring structural biology to a new level.

The relevance of our formalism is that it establishes a sound mathematical formalism to determine the number of snapshots required to answer a specific biological question at any resolution. These estimates will be useful for beamtime proposals to assert that datasets of sufficient quality can be collected. Similarly, it can help beamline scientists to decide whether enough snapshots can be collected during the allocated beamtime. In this way, the feasibility of single-particle experiments can be judged in an objective way and can

be used to guide experiments at new and existing XFELs for a broad class of biological macromolecules.

ACKNOWLEDGMENTS

We acknowledge the valuable discussions with E. Stojkovic, A. Ourmazd, and members of the UWM single-particle data analysis group. This research was supported by the U.S. National Science Foundation under Award No. STC1231306 and by the UWM Research Growth Initiative.

APPENDIX A: SIMULATION OF DIFFRACTION PATTERNS

For the simulation of diffraction patterns, we construct a regular two-dimensional Cartesian grid, which represents the detector pixel positions. The geometric relation of a detector pixel to the diffraction space is shown in Fig. 7. The coordinates (x, y) denote the position of a detector pixel, which has a corresponding point (q_x, q_y, q_z) in reciprocal space located on the Ewald sphere, where q is the scattering vector. From this geometry, the relationship between the two-dimensional detector and the reciprocal space is given by

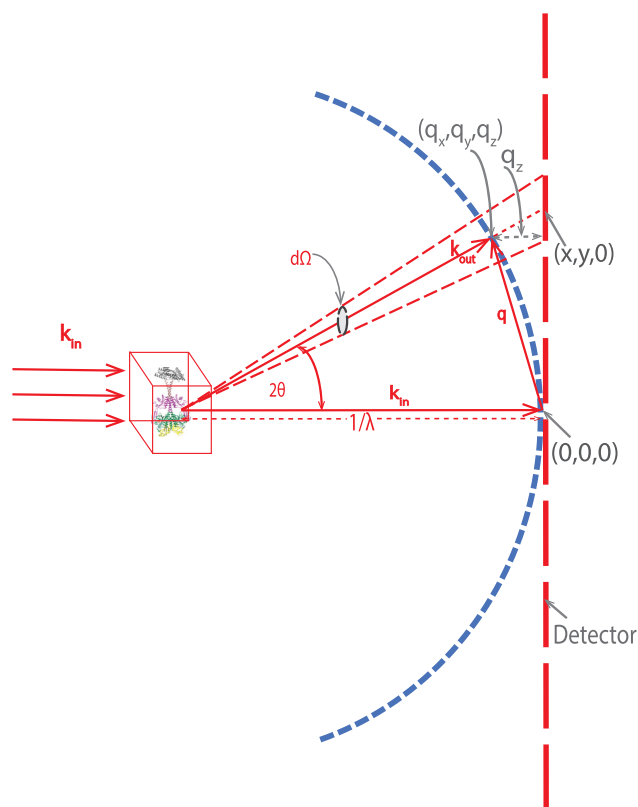


FIG. 7. Elastic scattering in forward geometry. 2θ denotes the angle between the incident k_{in} and scattered k_{out} wave vector. The scattered vector k_{out} ends on the Ewald sphere due to energy conservation ($|k_{out}| = |k_{in}|$). The scattering vector $q = k_{out} - k_{in}$ denotes the resulting momentum transfer.

$$\begin{aligned}
 q &= \frac{2}{\lambda} \sin(\theta), \\
 q_x &= q \left(1 - \frac{\lambda^2 q^2}{4}\right)^{\frac{1}{2}} \sin(\arctan 2(x, y)), \\
 q_y &= q \left(1 - \frac{\lambda^2 q^2}{4}\right)^{\frac{1}{2}} \cos(\arctan 2(x, y)), \\
 q_z &= -\frac{\lambda}{2} q^2.
 \end{aligned}$$

The expected photon count for a Shannon pixel on the detector with the scattering vector (\mathbf{q}) is given by

$$n(\mathbf{q}) = \Phi r_e^2 |\mathbf{F}(\mathbf{q})|^2 d\Omega,$$

where Φ is the incident flux (“photons/area” = “pulse_energy/photon_energy/area”), “ o ” is the oversampling ratio, $o \geq 2$, and “ r_e ” is the classical electron radius ($2.8179 \times 10^{-5} \text{ \AA}$). $\mathbf{F}(\mathbf{q})$ is the structure factor of the molecule and $d\Omega$ the solid angle subtended by the area of the Shannon pixel

$$d\Omega = \left(\frac{\lambda}{oD}\right)^2 \cos^3(2\theta).$$

This yields an expected photon count of

$$n(\mathbf{q}) = \Phi r_e^2 |\mathbf{F}(\mathbf{q})|^2 \left(\frac{\lambda}{oD}\right)^2 \cos^3(2\theta).$$

For small scattering angles (2θ), the formula can be approximated by

$$n(\mathbf{q}) = \Phi r_e^2 |\mathbf{F}(\mathbf{q})|^2 \left(\frac{\lambda}{oD}\right)^2.$$

APPENDIX B: RANDOMLY ORIENTED DIFFRACTION PATTERNS USING QUATERNIONS

To generate uniform random oriented diffraction patterns, we use unit quaternions $\mathbf{Q} = (a + bi + cj + dk)$ as follows:⁴⁴

$$\begin{aligned}
 \mathbf{Q} &= (\sqrt{1 - u_1} \sin(2\pi u_2), \sqrt{1 - u_1} \cos(2\pi u_2), \\
 &\quad \sqrt{u_1} \sin(2\pi u_3), \sqrt{u_1} \cos(2\pi u_3)).
 \end{aligned}$$

The three numbers u_1 , u_2 , and u_3 are chosen at random, uniformly distributed in the interval $[0, 1]$.

These quaternions can be expressed as a rotation matrix⁴⁵

$$\mathbf{R} = \begin{bmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2bd + 2ac \\ 2bc + 2ad & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2cd + 2ab & a^2 - b^2 - c^2 + d^2 \end{bmatrix}.$$

This matrix rotates the molecule, represented by the three-dimensional atomic coordinates (x_j, y_j, z_j) to a new position with coordinates (x'_j, y'_j, z'_j) , written in matrix form,

$$\mathbf{X}'_j = \mathbf{R}\mathbf{X}_j.$$

These coordinates are then used to calculate the structure factor of the molecule consisting of N atoms, according to

$$\mathbf{F}(\mathbf{q}) = \sum_{j=1}^N f_j(q) \exp(2\pi i \mathbf{q} \cdot \mathbf{X}'_j),$$

where $f_j(q)$ is the atomic scattering factor of atom j and q is the magnitude of the scattering vector \mathbf{q} (see also [Appendix A](#)).

APPENDIX C: JUSTIFICATION FOR CALCULATING JOINT PROBABILITY

For the calculation of the joint probability \tilde{P} , we regarded the statistics for each voxel to be independent of all other voxels, as it is the case for independent dice. However, as a diffraction pattern samples a set of voxels, entirely determined by the orientation of the molecule, the assumption of independence is not granted. To estimate the effect on our formalism, we calculated the joint probability \tilde{P} by simulation as follows: We recorded the actual number of visits for each voxel by merging 4323 diffraction patterns of the phytochrome at 10 \AA (see [Table 1](#)), where the number of patterns was estimated by Eqs. (6) and (8), based on independent voxels with $M = 39$. Of a total of 27 trials, we found 15 instances with all voxels visited at least M times. This corresponds to a probability of 0.55. As the statistical error expected from the number of trials is about ± 0.1 , this estimation is within the value predicted by the assumption of independent voxels. We therefore conclude that our calculation of the joint probability, as given by Eq. (8), is sufficient for the purpose of the present work.

APPENDIX D: PYTHON CODE TO ESTIMATE THE EXACT NUMBER OF SNAPSHOTS

```

import numpy as np

def logFactorial(n):
    if n < 20:
        value = np.log(np.math.factorial(n))
    else:
        value = 0.5*np.log(2*np.pi*n) + n*np.log(n/np.e)
    return value

def pnm(p,N,M):
    #this function gives the probability to observe a voxel at least M
    times
    #from an ensemble of N snapshots
    # p = probability to hit a voxel for a single snapshot
    # N = Number of Snapshots
    # M = Redundancy
    if N < M:
        s = 1
    else:
        s = 0
        lp = np.log(p)
        lmp = np.log(1-p)
        for k in np.arange(M):

```

```

s = s + np.exp(logFactorial(N) - logFactorial(N-k) -
logFactorial(k) + k*lp + (N-k)*lmp)
return np.maximum(1-s,0)

def numberOfSnapShots(d,D,nPhotons,SNR,P_tilde):
#nS number of Snapshots
#d is the Resolution
# D is the Diameter of a Molecule
#nPhotons is the Number of Photons
#P_tilde is the combined probability
#Number of Resolution elements
R = D/d
#number of voxels at Resolution Shell
nV_Shell = 16*np.pi*R**2

#probability per Shannon voxel
p = 1./(4*R)
M = np.ceil(SNR**2/nPhotons)

# P -> Probability to observe a voxel at least M times from an
ensemble of nS snapshot
# obtained from given P_tilde
P = np.exp(2*np.log(P_tilde)/nV_Shell)

nSmax = 1e12#
step = 2**10#

nS0 = M
while step > 1:
for nS in np.arange(nS0,nSmax,step):
if pnm(p,nS,M) > P:
break
nS0 = nS - step
step = step /2
return nS

```

APPENDIX E: ESTIMATING THE AVERAGE NUMBER OF PHOTONS IN THE OUTERMOST SHANNON PIXELS

The average number of photons $\langle n \rangle$ can be calculated from the molecular structure factor $F(\mathbf{q})$ as described in Appendix A. However, this requires calculations based on an atomic model, usually not available. Fortunately, there are useful approximations that only require the molecular weight and the approximate size of the molecule. Representing all non-hydrogen atoms by carbon atoms, the number of atoms N_c can be estimated from the molecular weight MW ,

$$N_c = MW/12 \text{ Da.}$$

At low resolution, one can approximate the molecule by a sphere with diameter D .⁴⁶ The squared structure factor of a sphere filled with N_c carbon atoms of atomic number $Z_c = 6$ is

$$|F(\mathbf{q})|^2 = N_c^2 Z_c^2 \left(3 \frac{\sin(\pi q D) - \pi q D \cos(\pi q D)}{(\pi q D)^3} \right)^2.$$

Accordingly, the average number of photons becomes

$$\langle n \rangle = \phi r_e^2 N_c^2 Z_c^2 \left(3 \frac{\sin(\pi q D) - \pi q D \cos(\pi q D)}{(\pi q D)^3} \right)^2 d\Omega.$$

With the number of resolution elements $R = D/d$ and $d = \frac{1}{q}$, this can also be written as

$$\langle n \rangle = \phi r_e^2 N_c^2 Z_c^2 \left(3 \frac{\sin(\pi R) - \pi R \cos(\pi R)}{(\pi R)^3} \right)^2 d\Omega.$$

Hence, for the low-resolution approximation, the geometric part of the structure factor depends only on the number of resolution elements.

For the high-resolution regime, a different approximation is quite useful. With N_c carbon atoms, the squared structure factor is

$$|F(\mathbf{q})|^2 = \left| f_C(\mathbf{q}) \sum_{j=1}^{N_c} \exp(2\pi i \mathbf{q} \cdot \mathbf{X}_j) \right|^2 = |f_C(\mathbf{q})|^2 \left| \sum_{j=1}^{N_c} \exp(2\pi i \mathbf{q} \cdot \mathbf{X}_j) \right|^2,$$

where $f_C(\mathbf{q})$ denotes the atomic scattering factor of carbon. Taking the average over a shell with radius q yields

$$\langle |F(\mathbf{q})|^2 \rangle = |f_C(q)|^2 \left\langle \left| \sum_{j=1}^{N_c} \exp(2\pi i \mathbf{q} \cdot \mathbf{X}_j) \right|^2 \right\rangle.$$

If the complex number $\exp(2\pi i \mathbf{q} \cdot \mathbf{X}_j)$ can be approximated by a random phasor, the quantity $\langle |\sum_{j=1}^{N_c} \exp(2\pi i \mathbf{q} \cdot \mathbf{X}_j)|^2 \rangle$ becomes equal to N_c ³³ so that

$$\langle n \rangle = \phi r_e^2 N_c |f_C(q)|^2 d\Omega.$$

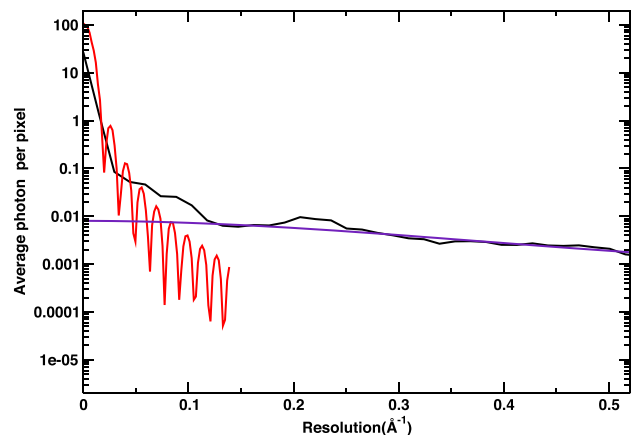


FIG. 8. Average number of photons per pixel as a function of resolution. Estimation for the full-length phytochrome using the atomic model (pdb code 5llw) and experimental parameters from Table I, "Hard x-rays" (black). Approximation assuming a spherical molecule filled with carbon atoms (red) and an approximation based on random phases (magenta).

For this, high-resolution approximation $\langle n \rangle$ is directly proportional to N_c and $|f_c(q)|^2$.

Both approximations are compared with the exact calculation using the atomic model of the full-length phytochrome (pdb code 5llw) and experimental parameters from Table I “Hard x-rays.” As shown in Fig. 8, the low-resolution approximation is reasonable up to a resolution of ~ 20 Å, whereas the high-resolution approximation is in excellent agreement for sub-nanometer resolution and below.

APPENDIX F: COMPUTATIONAL RESOURCES

The computational workflow was implemented with custom-written Python programs. All calculations were performed on a Linux workstation with AMD Processor FX-8350 (Eight-Core) and 16 GB RAM. Computational times were approximately as follows: 6 h for simulation of 30 000 diffraction patterns and 1 h to merge 30 000 diffraction patterns to a 3D diffraction volume using the cone-gridding algorithm.

REFERENCES

- J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, “A three-dimensional model of the myoglobin molecule obtained by x-ray analysis,” *Nature* **181**, 662 (1958).
- B. Rupp, *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology* (Garland Science, 2009).
- B. J. McNeil and N. R. Thompson, “X-ray free-electron lasers,” *Nat. Photonics* **4**(12), 814 (2010).
- R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Haidu, “Potential for bio-molecular imaging with femtosecond X-ray pulses,” *Nature* **406**, 752–757 (2000).
- J. Hajdu, “Single-molecule X-ray diffraction,” *Curr. Opin. Struct. Biol.* **10**, 569 (2000).
- K. J. Gaffney and H. N. Chapman, “Imaging atomic structure and dynamics with ultrafast X-ray scattering,” *Science* **316**, 1444 (2007).
- V. L. Shneerson, A. Ourmazd, and D. K. Saldin, “Crystallography without crystals. I. The common-line method for assembling a three-dimensional diffraction volume from single-particle scattering,” *Acta Crystallogr., Sect. A* **64**(2), 303–315 (2008).
- R. Fung, V. Shneerson, D. K. Saldin, and A. Ourmazd, “Structure from fleeting illumination of faint spinning objects in flight,” *Nat. Phys.* **5**, 64 (2009).
- N. T. D. Loh and V. Elser, “Reconstruction algorithm for single-particle diffraction imaging experiments,” *Phys. Rev. E* **80**, 026705 (2009).
- M. Tegze and G. Bortel, “Atomic structure of a single large biomolecule from diffraction patterns of random orientations,” *J. Struct. Biol.* **179**, 41 (2012).
- K. Ayer, T. Y. Lan, V. Elser, and N. D. Loh, “Dragonfly: An implementation of the expand-maximize-compress algorithm for single-particle imaging,” *J. Appl. Crystallogr.* **49**, 1320 (2016).
- J. R. Fienup, “Phase retrieval algorithms: A comparison,” *Appl. Opt.* **21**, 2758 (1982).
- S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells *et al.*, “X-ray image reconstruction from a diffraction pattern alone,” *Phys. Rev. B* **68**(14), 140101 (2003).
- T. Ekeberg, M. Svenda, C. Abergel, F. Maia, V. Seltzer, J. M. Claverie *et al.*, “Three-dimensional reconstruction of the giant mimivirus particle with an X-ray free-electron laser,” *Phys. Rev. Lett.* **114**(9), 098102 (2015).
- A. Hosseinizadeh, G. Mashayekhi, J. Copperman, P. Schwander, A. Dashti, R. Sepehr *et al.*, “Conformational landscape of a virus by single-particle X-ray scattering,” *Nat. Methods* **14**(9), 877–881 (2017).
- A. Munke, J. Andreasson, A. Aquila, S. Awel, K. Ayer, A. Barty *et al.*, “Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source,” *Sci. Data* **3**, 160064 (2016).
- Y. Cheng, “Single-particle Cryo-EM at crystallographic resolution,” *Cell* **161**, 450 (2015).
- R. M. Glaeser, “Proteins, interfaces, and cryo-EM grids,” *Curr. Opin. Colloid Interface Sci.* **34**, 1 (2018).
- J. Frank, “Time-resolved cryo-electron microscopy: Recent progress,” *J. Struct. Biol.* **200**, 303 (2017).
- M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, O. Jönsson *et al.*, “Single mimivirus particles intercepted and imaged with an X-ray laser,” *Nature* **470**, 78 (2011).
- Z. Y. Jiang, L. R. Swem, B. C. Rushing, S. Devanathan, G. Tollin, and C. E. Bauer, “Bacterial photoreceptor with similarity to photoactive yellow protein and plant phytochromes,” *Science* **285**, 406 (1999).
- J. R. Wagner, J. S. Brunzelle, K. T. Forest, and R. D. Vierstra, “A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome,” *Nature* **438**, 325 (2005).
- E. S. Burgie, A. N. Bussell, J. M. Walker, K. Dubiel, and R. D. Vierstra, “Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome,” *Proc. Natl. Acad. Sci.* **111**(28), 10179–10184 (2014).
- N. C. Woitowich, A. S. Halavaty, P. Waltz, C. Kupitz, J. Valera, G. Tracy *et al.*, “Structural basis for light control of cell development revealed by crystal structures of a myxobacterial phytochrome,” *IUCrj* **5**, 619 (2018).
- J. C. Sanchez, M. Carrillo, S. Pandey, M. Noda, L. Aldama, D. Feliz *et al.*, “High-resolution crystal structures of a myxobacterial phytochrome at cryo and room temperatures,” *Struct. Dyn.* **6**, 054701 (2019).
- L. O. Essen, J. Mailliet, and J. Hughes, “The structure of a complete phytochrome sensory module in the Pr ground state,” *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14709–14714 (2008).
- H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke *et al.*, “Signal amplification and transduction in phytochrome photosensors,” *Nature* **509**, 245 (2014).
- E. S. Burgie, J. Zhang, and R. D. Vierstra, “Crystal structure of *Deinococcus* phytochrome in the photoactivated state reveals a cascade of structural rearrangements during photoconversion,” *Structure* **24**(3), 448–457 (2016).
- G. Gourinchas, S. Etzl, C. Göbl, U. Vide, T. Madl, and A. Winkler, “Long-range allosteric signaling in red light-regulated diguanylyl cyclases,” *Sci. Adv.* **3**(3), e1602498 (2017).
- A. Björling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman *et al.*, “Structural photoactivation of a full-length bacterial phytochrome,” *Sci. Adv.* **2**, e1600920 (2016).
- D. Sayre, “Some implications of a theorem due to Shannon,” *Acta Crystallogr.* **5**, 843 (1952).
- D. Sayre, H. N. Chapman, and J. Miao, “On the extendibility of x-ray crystallography to noncrystals,” *Acta Crystallogr., Sect. A* **54**, 232–239 (1998).
- W. Goodman, *Statistical Optics* (John Wiley & Sons, 2000).
- A. Hosseinizadeh, P. Schwander, A. Dashti, R. Fung, R. M. D’Souza, and A. Ourmazd, “High-resolution structure of viruses from random diffraction snapshots,” *Philos. Trans. R. Soc. B* **369**(1647), 20130326 (2014).
- R. Henderson, A. Sali, M. L. Baker, B. Carragher, B. Devkota, K. H. Downing *et al.*, “Outcome of the first electron microscopy validation task force meeting,” *Structure* **20**, 205 (2012).
- M. Van Heel and M. Schatz, “Fourier shell correlation threshold criteria,” *J. Struct. Biol.* **151**(3), 250–262 (2005).
- P. B. Rosenthal and R. Henderson, “Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryo-microscopy,” *J. Mol. Biol.* **333**(4), 721–745 (2003).
- K. Ayer, A. J. Morgan, A. Aquila, H. DeMirici, B. G. Hogue, R. A. Kirian *et al.*, “Low-signal limit of X-ray single particle diffractive imaging,” *Opt. Express* **27**(26), 37816 (2019).
- A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. Deponce *et al.*, “The linac coherent light source single particle imaging road map,” *Struct. Dyn.* **2**(4), 041701 (2015).
- A. Hosseinizadeh, A. Dashti, P. Schwander, R. Fung, and A. Ourmazd, “Single-particle structure determination by X-ray free-electron lasers: Possibilities and challenges,” *Struct. Dyn.* **2**(4), 041601 (2015).
- B. Ziaja, Z. Jurek, N. Medvedev, V. Saxena, S. K. K. Son, and R. Santra, “Towards realistic simulations of macromolecules irradiated under the conditions of coherent diffraction imaging with an x-ray free-electron laser,” *Photonics* **2**(1), 256–269 (2015).

- ⁴²J. Bielecki, M. F. Hantke, B. J. Daurer, H. K. N. Reddy, D. Hasse, D. S. D. Larsson *et al.*, “Electrospray sample injection for single-particle imaging with x-ray lasers,” *Sci. Adv.* **5**, eaav8801 (2019).
- ⁴³A. Allahgholi, J. Becker, A. Delfs, R. Dinapoli, P. Goettlicher, D. Greiffenberg *et al.*, “The adaptive gain integrating pixel detector at the European XFEL,” *J. Synchrotron Radiat.* **26**, 74 (2019).
- ⁴⁴K. Shoemake, “Uniform random rotations,” *Graphics Gems III (IBM Version)* (Elsevier, 1992).
- ⁴⁵J. B. Kuipers *et al.*, *Quaternions and Rotation Sequences* (Princeton University Press, Princeton, 1999), Vol. 66.
- ⁴⁶J. S. Pedersen, “Analysis of small-angle scattering data from colloids and polymer solutions: Modeling and least-squares fitting,” *Adv. Colloid Interface Sci.* **70**, 171 (1997).
- ⁴⁷E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng *et al.*, “UCSF Chimera—A visualization system for exploratory research and analysis,” *J. Comput. Chem.* **25**, 1605 (2004).