Holographic Single-Particle Imaging for Weakly Scattering, Heterogeneous Nanoscale Objects

Abhishek Mall[1,2] and Kartik Ayyer[1,2,3,*]

[1] *Max Planck Institute for the Structure and Dynamics of Matter, Hamburg 22761, Germany*
[2] *Center for Free Electron Laser Science, Hamburg 22761, Germany*
[3] *The Hamburg Center for Ultrafast Imaging, Hamburg 22761, Germany*

(Received 21 October 2022; accepted 13 March 2023; published 8 May 2023)

Single-particle imaging (SPI) at x-ray free-electron lasers is a technique to determine the three-dimensional structure of nanoscale objects like biomolecules from a large number of diffraction patterns of copies of these objects in random orientations. The technique has been limited to relatively low resolution due to background noise and heterogeneity of the target particles. A recently introduced reference-enhanced holographic SPI methodology uses strongly scattering holographic references to improve background tolerance, and, thus, the achievable resolution, at the cost of additional latent variables beyond orientation. Here, we describe an improved reconstruction algorithm based on maximum likelihood estimation, which scales better, enabling fine sampling of latent parameters to reach high resolutions, and much better performance in the low signal limit. Furthermore, we show that structural variations within the target particle are averaged in real space, significantly improving robustness to conformational heterogeneity in comparison to conventional SPI. With these computational improvements, we believe reference-enhanced SPI is capable of reaching sub-nanometer resolution biomolecule imaging.

DOI: 10.1103/PhysRevApplied.19.054027

I. INTRODUCTION

Single-particle imaging (SPI) experiments leverage x-ray free-electron lasers (XFELs) to investigate the structure and dynamics of biological entities in their near-native state [1]. The ultrashort x-ray pulses enable the study of ultrafast structural dynamics in biological samples such as proteins and viruses at room temperature via diffraction before destruction [2], and without the need to freeze samples. Many XFEL pulses with high flux and spatial coherence diffract from unknown target objects one at a time, in random orientations to collect millions of diffraction patterns. These are then computationally aligned and merged to reconstruct the three-dimensional (3D) structure of the object [3].

This analysis often broadly follows the following three steps. First, diffraction patterns containing scattering from spurious contaminants, multiple particle aggregates, and other outliers are identified and removed using some machine learning framework [4,5]. The rest of the patterns each represent a tomographic slice through the target object's 3D Fourier transform, albeit without the Fourier phases. The second step in the analysis pipeline is to align and average these patterns to obtain the 3D distribution of Fourier magnitudes [6,7]. Finally, iterative phase-retrieval methods are used to recover the structure of the particle from these oversampled magnitudes, possibly with some background subtraction [8–11].

SPI has been used successfully in imaging samples in the 100-nm size range since they have a high scattering cross section in comparison to the background, hence yielding high-quality diffraction patterns [10,12,13]. However, for smaller particles, signal levels are much lower and extraneous background photons can severely hinder the alignment process. And while reconstruction algorithms are remarkably tolerant to low signal levels [6,11,14], background often poses a fundamental limit on the achievable resolution since the diffraction signal from a compact object falls off very quickly with increasing momentum transfer, but the background usually does not [15].

Various sample delivery methods have been used to deliver samples to the x-ray beam focus, each balancing the requirements of maximizing efficiency (hit rate) with minimizing background. Aerosol sample delivery has low

*kartik.ayyer@mpsd.mpg.de

background signal but has a relatively low particle density, leading to low hit fractions [16,17]. Liquid jets and solid substrates can be used as carrier media to increase the hit ratio [18–22], but each hit now also has a substantial signal from the medium that obscures signal from the sample and limits orientation determination. With the advent of high repetition rate XFELs [23], even with the low hit rates of aerosol methods, one can still collect millions of patterns in a day [5]. Nevertheless, even here, the background from the carrier gas of the aerosol and from the detector false positives still produces sufficient background to limit the resolution to around 2 nm.

Recently, a reference-enhanced SPI [24] technique was introduced, based on the holographic principle. The approach suggests attaching a strongly scattering particle to the target object to form a composite object. The references considered are a spherical gold nanoparticle (AuNP) and a two-dimensional (2D) crystal lattice with a unit-cell size comparable to the target object. The total scattered signal is increased for each shot when a reference is attached and the signal now becomes more tolerant to background. A reconstruction algorithm is also developed to analyze the diffraction patterns generated from such a system in order to recover unknown parameters beyond orientation resulting from shot-to-shot variations in the composition of the composite object.

This holographic SPI technique alleviates the problem of high sensitivity to extraneous background, but at the cost of additional computational complexity in recovering the structure from the data. The technique introduces additional degrees of freedom in the composite system, which are unknown parameters for each pattern in the diffraction dataset. The algorithm now has to recover not only the unknown orientation of the target object but also these hidden (latent) parameters characterizing the properties of the reference and the relative displacement between the reference and the target object. For the composite system where the reference is a spherical AuNP, the reference can be specified by one parameter: its diameter D.

The previous work also introduced a reference expand–maximize–compress (EMC) algorithm [24] that is a modified version of the EMC algorithm [6,7] that recovers these additional latent parameters and directly reconstructs the target object's complex Fourier transform and not just the intensities. However, as we will discuss in Sec. II, this method has problems dealing with very weak patterns as well as scaling to fine parameter sampling required to reach a high resolution. In this paper we propose a phase-retrieval algorithm for holographic SPI based on maximum likelihood estimation via a pattern search dubbed as MLP. MLP scales efficiently with the sampling of latent parameters and the total number of diffraction patterns in comparison with the previously introduced *divide* and *concur* approach in the reference EMC algorithm [25]. Additionally, we observe that it performs much better when the signal level is low. The algorithm is described in Sec. II and its performance is shown with simulated data in Sec. III A.

We also discuss the performance of the holographic SPI method with the MLP algorithm in the case where the target object itself is heterogeneous. In conventional SPI this heterogeneity would have to be classified shot by shot before averaging since averaging intensities from variable structures can lead to meaningless results. Here, we show that in our approach, even without classification, we are able to reconstruct the average structure, making the process of handling conformational variations much more tractable. Thus, we find that this combination of the MLP algorithm with the holographic SPI experimental setup shows substantial promise in pushing the SPI technique to sub-nanometer resolution.

II. RECONSTRUCTION ALGORITHM

In a holographic SPI experiment the sample of interest (target object) is in the vicinity of a strongly scattering reference, which in this work, we take to be a spherical gold nanoparticle (AuNP). A large number of diffraction patterns of these conjugates are collected in order to reconstruct the 3D structure of the target object. This reconstruction process consists broadly of two steps, the first being the determination of the unknown, or latent, parameters for each pattern, followed by a step to recover the structure given the data with the estimated parameters, which include the orientation of the object and the relative position and structure of the reference. Fortunately, the structure of the spherical AuNP can be described by just one number, namely its diameter D. For illustrative purposes, we limit ourselves to a two-dimensional object with only one in-plane rotational degree of freedom. Within this space, for a homogeneous reproducible object, we have four latent parameters to solve for, i.e., the unknown orientation θ, the diameter of the AuNP d, and the relative positions of the AuNP and the target object in the x and y directions (t_x, t_y). In the general case, there are three orientational and translational parameters each, but the structure of the problem remains unchanged. The electron density of the composite object $\rho(\mathbf{r})$ is the sum of electron densities of the spherical AuNP $\rho_s(\mathbf{r}, D)$ and the unknown target object $\rho_o(\mathbf{r})$,

$$\rho(\mathbf{r}) = \rho_o(\mathbf{r}) + \rho_s(\mathbf{r} - \mathbf{t}, D), \tag{1}$$

where $\mathbf{t}$ is the relative shift of the centers of the two objects and D represents the diameter of the spherical AuNP. The total intensity distribution on the far-field detector in each frame is then

$$I(\mathbf{q}, D, \mathbf{t}) = \left|F_o(\mathbf{q}) + F_s(\mathbf{q}, D)e^{2\pi i \mathbf{q}.\mathbf{t}}\right|^2, \tag{2}$$

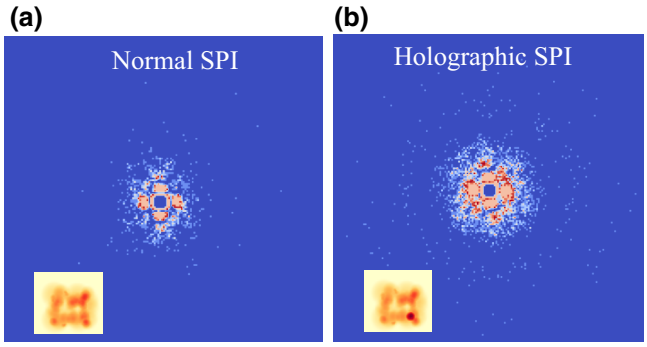

FIG. 1. (a) Simulated diffraction pattern for a random object in a normal SPI case with 4810 photons. (b) Diffraction pattern from the same particle with a small AuNP attached (10 000 photons). If background is present, structural information can be discerned at higher scattering angles in the holographic case. Insets show the true projected electron density of the objects.

where the F terms represent the Fourier transform of electron densities, $F(\mathbf{q}) = \mathcal{F}[\rho(\mathbf{r})](\mathbf{q})$, and frame-by-frame shifts of the sphere transform to a phase ramp.

The simulated diffraction dataset contains a large number of patterns defined by Eq. (2), each with random D and $\mathbf{t}$ parameters and rotated in plane by a uniform random angle θ. These holographic intensities are then Poisson sampled with a given mean intensity level to generate the simulated photon counts per pixel. In Fig. 1 we see the effect of the spherical AuNP attached as a reference to the target object on a simulated pattern for a given set of latent parameters with the same effective incident fluence. In both patterns, the single-frame signal-to-noise ratio (SNR) is very low beyond the first few speckles. In conventional SPI this is improved by aligning and averaging a large number of patterns as long as this alignment can be performed in the presence of background and sample heterogeneity. For the chosen AuNP size, the conjugates have on average twice as many total scattered photons with much of the excess in the higher order rings that improve the SNR in the presence of background at these resolutions, making hit detection significantly more effective. Larger AuNPs will scatter more strongly, but with lower contrast along the diffraction rings. The radii and intensity modulations in these rings are relevant to solve for the diameter of the AuNP and its relative position. In order to retrieve the target object's structure that is assumed to be common to all the intensities, the first step is to solve for the set of latent parameters for each diffraction pattern and retrieve $F_o(\mathbf{q})$.

A modified version of the EMC algorithm (reference EMC algorithm) was developed for holographic SPI [24]. The conventional EMC algorithm involves three steps in each iteration: *expand*, *maximize*, and *compress* that iteratively update the 3D intensity model to one that has a higher likelihood of generating the observed diffraction patterns. When the reference is attached, the optimal final model is not the intensity distribution of the composite assembly, but rather the complex Fourier transform of just the target object, $F_o(\mathbf{q})$. In the E step of the reference EMC algorithm, the latent parameter space is grid sampled and predicted intensities are generated for each sample using the current estimate of $F_o(\mathbf{q})$. In the M step the noisy diffraction patterns are compared with these predicted intensities and a probability distribution over parameters for each pattern is calculated. This distribution is then used to update the predicted intensities for each sampled parameter vector. The final C step needs to recover the optimal $F_o(\mathbf{q})$ consistent with this stack of intermediate predicted intensities, one for each sampled parameter set. This separate reconstruction problem is somewhat reminiscent of ptychography where multiple intensities are generated from a common object by varying translations [26]. In this case however, the intensities are generated by coherent addition with a reference [Eq. (2)] rather than by multiplication with a probe function. In Ref. [24] a *divide* and *concur* iterative phase-retrieval approach was implemented to solve this problem. However, the requirement to generate the intermediate intensities limits its scalability of this algorithm. For a finer sampling of diameters and relative shifts, which is necessary for a high resolution, one ends up with many realizations of intermediate intensities that quickly becomes computationally expensive. Additionally, this method is composed of projection operations minimizing a Euclidean error metric. This implicitly assumes a Gaussian error distribution that becomes increasingly incorrect at lower signal levels [27].

A. Maximum likelihood phase retrieval

Taking the above mentioned considerations into account, we implement a maximum likelihood estimation strategy using a pattern search technique dubbed as MLP to retrieve the full $F_o(\mathbf{q})$ of an unknown target object. The E and M steps of the EMC iteration are the same as before, with the only difference that we only use the most likely set of parameters rather than the whole probability distribution. The C step is changed from the phase-retrieval-like approach using intermediate intensities to a direct search for the most likely complex Fourier amplitudes of the target object given the diffraction data and the current estimate of the latent parameters. The optimization is performed using a pattern search procedure implemented for each model pixel independently. The approach is tested for different intensity signal levels in diffraction datasets with the low signal having as few as 2000 photons per pattern. We find that the algorithm is more noise tolerant with better fidelity at lower signal levels where the previously introduced *divide* and *concur* approach yields poor results.

In addition, the MLP approach scales more favorably as the sampling of latent parameters is made finer. The computational complexity of the C step is now independent of the sampling, while in the previous approach, finer sampling would result in a large number of intermediate intensities from which to perform the "phase retrieval". Since the final resolution is ultimately determined by the accuracy of the estimated latent parameters, this algorithm can be efficiently scaled and the parameters can be refined by local searches.

B. Single-pixel behavior

The optimal $F_o(\mathbf{q})$ is determined at every model pixel independently. To illustrate the working of the reconstruction algorithm, we now discuss the behavior for a single-pixel detector at a given $\mathbf{q}$ assuming the orientations are known. In holographic or reference-enhanced SPI, the measured intensity at a given detector pixel is described by Eq. (2). The estimated orientation of the object then relates the detector pixel to some pixel in the Fourier representation of the target object's electron density. A single model pixel intensity at a fixed $\mathbf{q}$ can be written as

$$I_{\text{pix}} = \left|F_{o,\text{pix}} + F_{s,\text{pix}}(D)e^{2\pi i \mathbf{q}_{\text{pix}}.\mathbf{t}}\right|^2. \tag{3}$$

Here F_o is the complex number that we need to solve from the diffraction dataset, I_{pix}, with varying shifts, $\mathbf{t}$, and AuNP diameters, D, at the given model pixel from multiple realizations in different patterns.

Figure 2 depicts the behavior of the true intensity and measured photon counts for different phase shifts ($\Delta s = \mathbf{q}.\mathbf{t}$) between the AuNP and the target object. The photon distribution at a given phase-shift value is Poisson distributed with the mean being the true intensity shown in orange. When the average photon count at the pixel is high enough, the intensity follows a sinusoidal form with respect to the phase shifts, as seen in Fig. 2(a). The relative phase, offset, and amplitude of this sinusoidal curve tells us the phase and magnitude of the object's Fourier transform at that $\mathbf{q}$. As the range of phase shifts becomes small, one is effectively zooming into the sinusoidal curve for the same amount of average photons [see Fig. 2(c)], and the fitting becomes more challenging. Similarly, when the average photon count on the detector falls, as seen in Fig. 2(b), the photon distribution becomes too sparse to immediately see the true intensity and in the hardest case, for smaller phase shifts and low counts, the true intensity plot becomes almost flat [Fig. 2(d)].

The log likelihood of such a diffraction dataset at any model pixel can be calculated by

$$Q(F_{o,\text{pix}}) = \sum_d K_{\text{pix},d} log(I_{\text{pix},d}) - I_{\text{pix},d}, \tag{4}$$

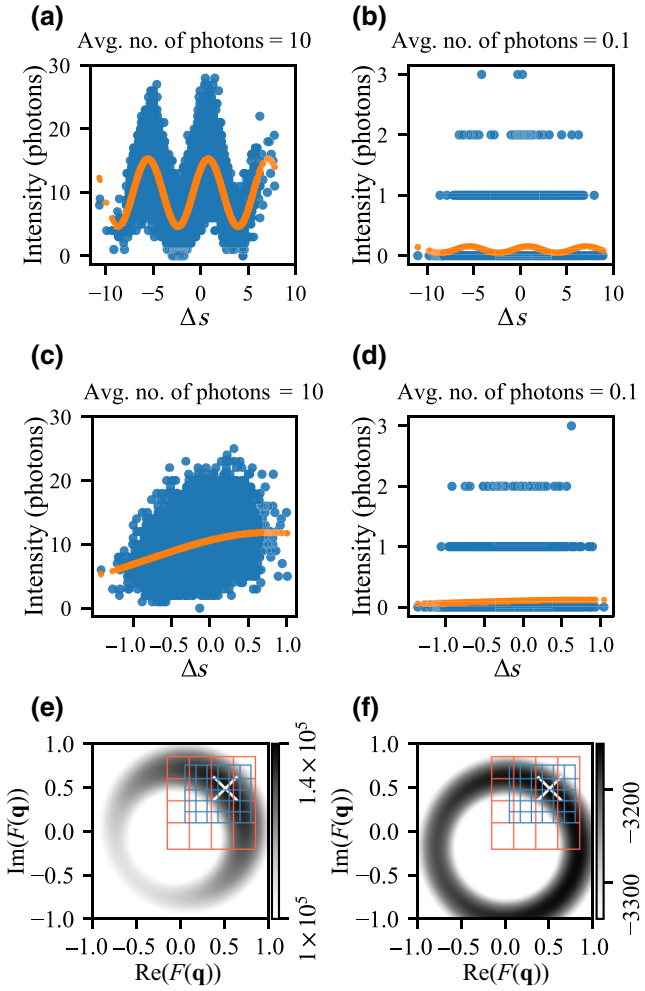

FIG. 2. Intensity behavior at a single detector pixel for holographic SPI diffraction data. (a)–(d) Average number of photons versus relative phase shifts of a gold sphere with respect to the target object. The ideal intensity on the same pixel is represented by the orange line. (e),(f) Log-likelihood landscapes obtained from diffraction data for cases (a),(d), respectively. The '×' denotes the true value of $F(\mathbf{q})$ maximizing the log likelihood. The grid lines (first grid is *orange*, second grid is *blue*) depict the size of the pattern search grid in consecutive iterations.

where $I_{\text{pix},d}$ is the predicted intensity given by Eq. (3) and $K_{\text{pix},d}$ is the number of observed photons in pattern number d. The $K_{\text{pix},d}!$ term is neglected since it does not depend upon the model F_o. Figures 2(e) and 2(f) show this log-likelihood function distribution in the complex plane of $F_{o,\text{pix}}$ for the diffraction pattern data given from Figs. 2(a) and 2(d), respectively. The likelihood has a well-behaved landscape with a sharp maxima with different values of the real and imaginary parts of F_o, when the average photon count is high enough with a large shift range. When the average photon count is low with small shifts, the likelihood landscape becomes much flatter along the visible ridge.

Furthermore, this likelihood distribution is independent of the measured photon counts at the neighboring detector pixel. As a result, one can implement an algorithm for each pixel individually as a maximum likelihood estimation problem and try to converge for an optimal F_o for each pixel and trivially parallelize the algorithm over all model pixels. Of course, the Fourier transform is oversampled and, thus, its values at neighboring pixels are strongly correlated. This information will be taken into account later in Sec. III.

C. Finding the most likely solution

The likelihood function given in Eq. (4), can not be solved analytically. While one can perform this 2D optimization in the complex plane with many methods, we find that derivative-based approaches are not robust, with maximization of likelihood failing for low photon count cases and small shifts. We find reasonable success with a pattern search approach [28] for two reasons: first, it can be efficiently parallelized on graphical processing units (GPUs) and, second, it is quite robust in low signal and small shift cases.

Pattern search optimization is essentially a nonderivative technique that does not require a gradient calculation in its update of parameters. The search begins with a 2D grid of values of **F** (see Figs. 2(e) and 2(f), orange grid lines). At each iteration, this grid moves to a set of values that best maximizes the likelihood function. If it finds a set of values of **F** that does not have better likelihood then it stays at the current value and the grid shrinks and becomes denser with smaller steps between the grid values (blue grid lines). A search is run till a threshold error between the current and the previous estimate is reached. This is performed for all the model pixels and then the final model is used to get updated latent parameters for the next EMC iteration. The pseudocode for a single iteration is described in Algorithm 1.

Input:
$F_o(q)$ ▷ *Current estimate of target object Fourier transform at pixel q*
$K_{\mathrm{pix},d}$ ▷ *Dataset: photons at pixel,* pix *and frame, d*
Ω_d ▷ *Current enstimate of latent parameters:* $\mathbf{t}, D, \theta$ *for each frame d*

grid ▷ *Square grid of points in the complex plane with unit spacing*

function RunIteration($F_o(q)$, $K_{\mathrm{pix},d}$, Ω_d)
  for q in 1..N_{pix} **do**
    $F'_o(q) \leftarrow$ OptimizePixel($q, F_o(q), K_{\mathrm{pix}}, \Omega$)
  end for
end function

function OptimizePixel(q, F_o, K, Ω)
  step $\leftarrow |F_o|/4$
  for n in 1..10 **do**
    for i in 1..N_{grid} **do**
      $F_i \leftarrow F_o +$ grid$_i \times$ step
      $Q_i \leftarrow$ Likelihood(K, Ω) ▷ *(Eq. 4)*
    end for
    $i_{\mathrm{opt}} \leftarrow \mathrm{argmin}_i\, Q_i$
    if $i_{\mathrm{opt}} = 0$ **then** ▷ *Center of grid*
      step $\leftarrow$ step / $\sqrt{N_{\mathrm{grid}}}$
    end if
    if step $< |F_o| \times 10^{-3}$ **then**
      break
    end if
  end for
end function

Algorithm 1. Pseudocode for one iteration of the MLP. The latent parameters are determined as in the EMC algorithm.

III. RESULTS

We test the performance of the algorithm described above on 2D simulated data with a single angular degree of freedom. The target object is a randomly generated agglomeration of small spheres, approximating the electron density distribution of a biological sample. Each diffraction pattern is a Poisson-sampled distribution of scattered photons from a conjugate object consisting of the target attached to a spherical gold nanoparticle (AuNP) reference. The average diameter of the AuNP is roughly one fifth the size of the target and the contrast is 11 times higher to reflect the electron density ratio of gold to organic matter. The conjugate objects varied from pattern to pattern in multiple ways. The relative shifts in the x and y directions between their centers are sampled from a normal distribution with a standard deviation of 1 pixel. The diameter of AuNP is also normally distributed with a mean and standard deviation of 7 and 0.5 pixels, respectively, and the target object is 35 pixels across. The intensity distribution from this composite object is rotated in plane by a uniform random angle before Poisson sampling. The diffraction patterns are collected on a circular detector with a diameter of 185 pixels and a central hole with a 4 pixel radius.

Multiple datasets are simulated with varying signal levels with 10 000 patterns in each dataset. A common randomly generated object is used as the target in all of the simulations. The signal levels range from 2×10^3 to 10^5 photons/frame at the low and high extremes.

Figures 3(a) and 3(b) depict the absolute magnitude of the reconstructed complex Fourier model for the highest and lowest signal levels. The algorithm is run for 100 iterations and Fig. 3(d) shows the convergence behavior. Convergence is achieved after 15 iterations for the high signal level and 40 iterations for the low signal level. For each reconstruction, the iteration times range from 200 s for the low signal case and 600 s for the high signal case on a single V100 GPU. At the final reconstruction iteration,

both Fourier models have a few poorly reconstructed pixels, primarily at low q. This is due to the missing data in the central hole as well as regions where the scattering from the reference is significantly weaker than from the target. The intensity at these pixels is only minimally affected by the frame-to-frame variation in shifts or diameters, resulting in the Fourier phase being poorly constrained. These pixels are filled in using the following approach. First, a support mask is calculated using data from moderate q using the dark-field diffractive imaging approach [29]. This mask is then used as a real-space constraint and the difference map iterative phase-retrieval algorithm [9] is used to fill in the missing region as well as obtain a real-space image.

A. Signal level dependence of reconstruction quality

The fidelity of the reconstructed model is calculated using the Fourier ring correlation (FRC) metric, shown in Fig. 3(c). We compare the performance of the MLP method with the previously published method of *divide* and *concur* in [24] at low and high signal levels. For high signals, both algorithms perform well, but the *divide* and *concur* approach fails for the low signal case, whereas the MLP reconstructs the Fourier model with FRC > 0.5 up to $q = 55$ pixels. This can be attributed to the MLP method correctly accounting for the Poisson noise process rather than the Gaussian noise model implicit in the *divide* and *concur* method.

We then investigate the ability of the algorithm to correctly estimate the latent parameters given the current model. The sampling rate of all the latent parameters is kept identical across all datasets to facilitate comparison. Figure 4 shows the error distribution for the diameter of the AuNP, in-plane orientation and relative shifts in the x and y directions, respectively. The sampling of the in-plane orientation is in the range of 0° and 180° with a step size of 2°. In Fig. 4(a) one can see a distribution of errors between the retrieved and true in-plane angles for the high and low signal levels, respectively. The centers are shifted since the reconstruction has an arbitrary overall orientation. As expected, the distribution is narrow for a high signal, with $\sigma = 0.85^\circ$ that is below the sampling rate, and significantly broader for the low signal, with $\sigma = 11.4^\circ$.

Similarly, the AuNP diameter and x and y shifts are sampled with a step size of 0.5 and 1 pixel, respectively. Figure 4(b) shows that the diameter of the AuNP for each frame is close to the true values with uncertainties for the high and low signal being 0.15 and 0.18 pixels. Even for the lowest signal level the diameter can be accurately estimated below half the sampling rate, which is consistent with the high precisions that can be obtained in small-angle x-ray scattering. Figures 4(c) and 4(d) show the error distribution of shifts in the x and y directions for the two

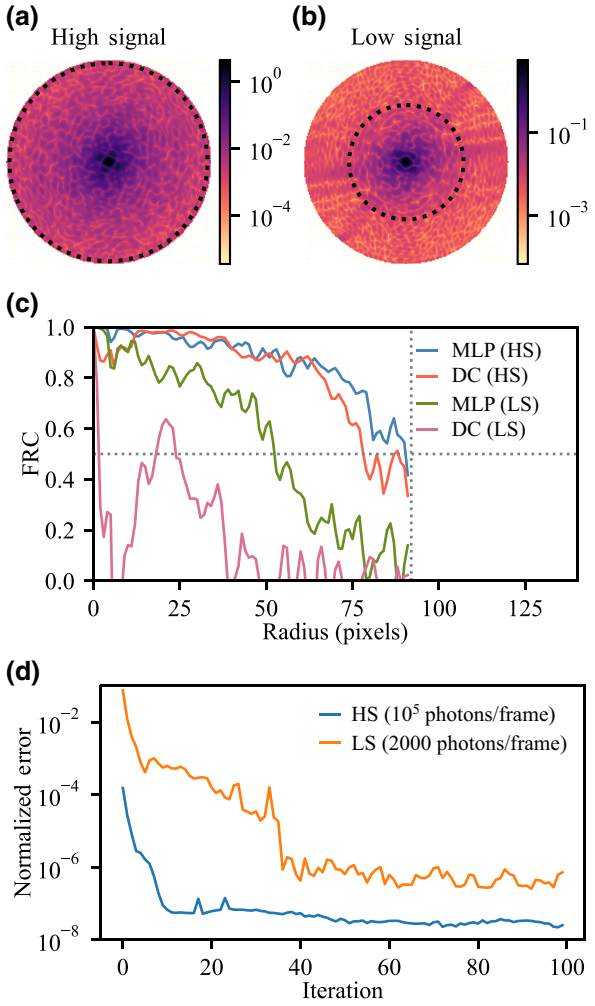

FIG. 3. Simulation results for a homogeneous target object attached to a spherical AuNP as reference. (a) Magnitude of reconstructed Fourier model for high signal diffraction data with the MLP method (10^5 photons/frame). (b) Same for low signal diffraction data (2000 photons/frame). Both datasets have 10^4 frames. Dashed rings highlight the FRC $= 0.5$ cutoff. (c) Comparison of Fourier ring correlation (FRC) between reconstructions and ground truth with different signal level diffraction data. Results are shown for the MLP, divide and concur (DC), high signal (HS), and low signal (LS) cases. (d) Normalized error (difference between successive iterations) versus iteration number showing faster convergence for higher signal levels as observed in conventional SPI as well [6].

signal levels. The error is confined for many of the patterns within half the sampling bin size of 1 pixel. However, the error distribution for the low signal case is relatively broad. This is strongly correlated to the error in the in-plane orientations. A large error in orientation leads to wrongly estimated shifts because the diffraction pattern is obtained by adding the AuNP to the target object and rotating the entire composite object. The predicted shift values

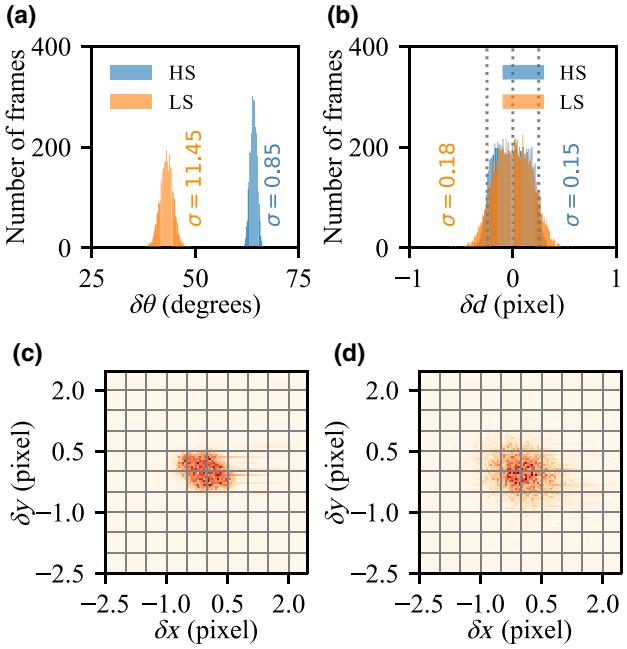

FIG. 4. Error distribution for reconstructed latent parameters for a homogeneous target object attached with a spherical AuNP as reference for low (LS) and high signal (HS) levels. (a) In-plane orientation, (b) diameter of gold sphere, (c),(d) relative displacements between the center of the target object and reference in the x and y directions. The vertical lines in (b) and the grid lines in (c),(d) show half the sampling rate, which sets a lower bound on the width of the distributions.

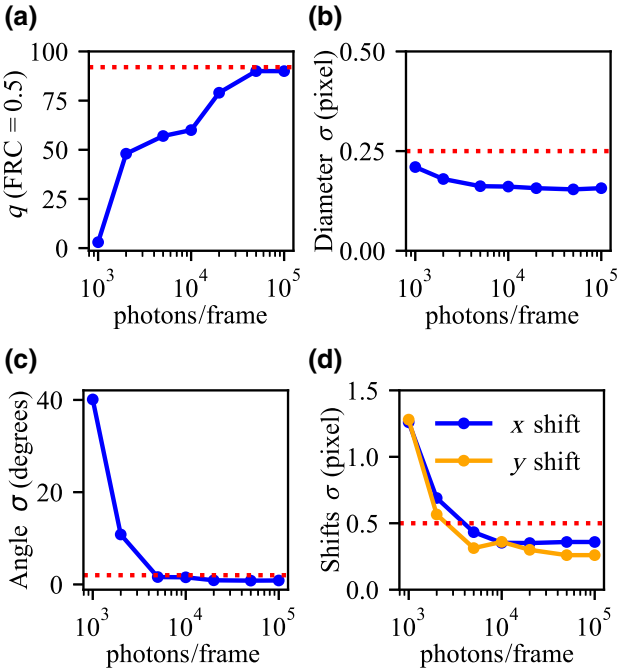

FIG. 5. Evaluation metrics of MLP for different photon signal levels. (a) The q value when FRC $= 0.5$ versus number of photons/frame. The horizontal dashed line corresponds to a resolution of 1 real-space pixel. Standard deviation values for (b) diameter errors, (c) orientation errors, and (d) shift errors in the x and y directions. The horizontal dashed lines in (b)–(d) indicate half the sampling rate.

are rotation corrected using the predicted orientations and then compared with true values.

The behavior of the reconstruction metrics discussed above is evaluated for multiple intermediate signal levels in Fig. 5. Figures 5(b)–5(d) show the dependence of the latent parameter errors on the signal level. Here we see that beyond 5000 photons/frame, the errors are all below the sampling rate indicated by the dotted red line. Figure 5(a) shows the dependence of the resolution on signal level, which depends both on an accurate estimation of latent parameters as well as the total signal from all the frames in the aggregate. Thus, we expect the reconstruction FRC to improve significantly with more patterns beyond this threshold of 5000 photons/frame.

B. Heterogeneity of target object

Many biological entities have flexible subunits that can move following a continuous landscape of various conformational states. The study of these conformation changes can help us understand the function of such biomolecules. However, there are significant computational challenges to retrieve conformational states and the structure of a target object from diffraction data. In conventional SPI the averaging of patterns from variable structures has an undetermined effect on the retrieved structure. This is because even though the Fourier transform is a linear operator, it is the intensities (squared magnitudes) that are averaged across different frames. In general, the resultant intensity distribution has lower contrast and is not the Fourier transform of any compact object. In some cases with an uncorrelated random motion of atoms, one could expect the background-subtracted intensities to represent the scattering from the average structure, but this is in no way guaranteed, especially with correlated motions. Thus, extensive computational efforts must be employed to discover a subset of data from a homogeneous ensemble or to solve for additional latent parameters associated with structural variations.

In the current holographic regime, we solve for the complex Fourier transform, which is linearly related to the real-space structure. Thus, we may expect the reconstruction to be that of the average structure. To examine whether this is indeed the case, we perform two computational experiments below where the target object varies from frame to frame but the algorithm still attempts to reconstruct a single structure.

1. Two-state heterogeneity

We investigate the scenario in which the target object is composed of a homogeneous rigid part (the large unit)

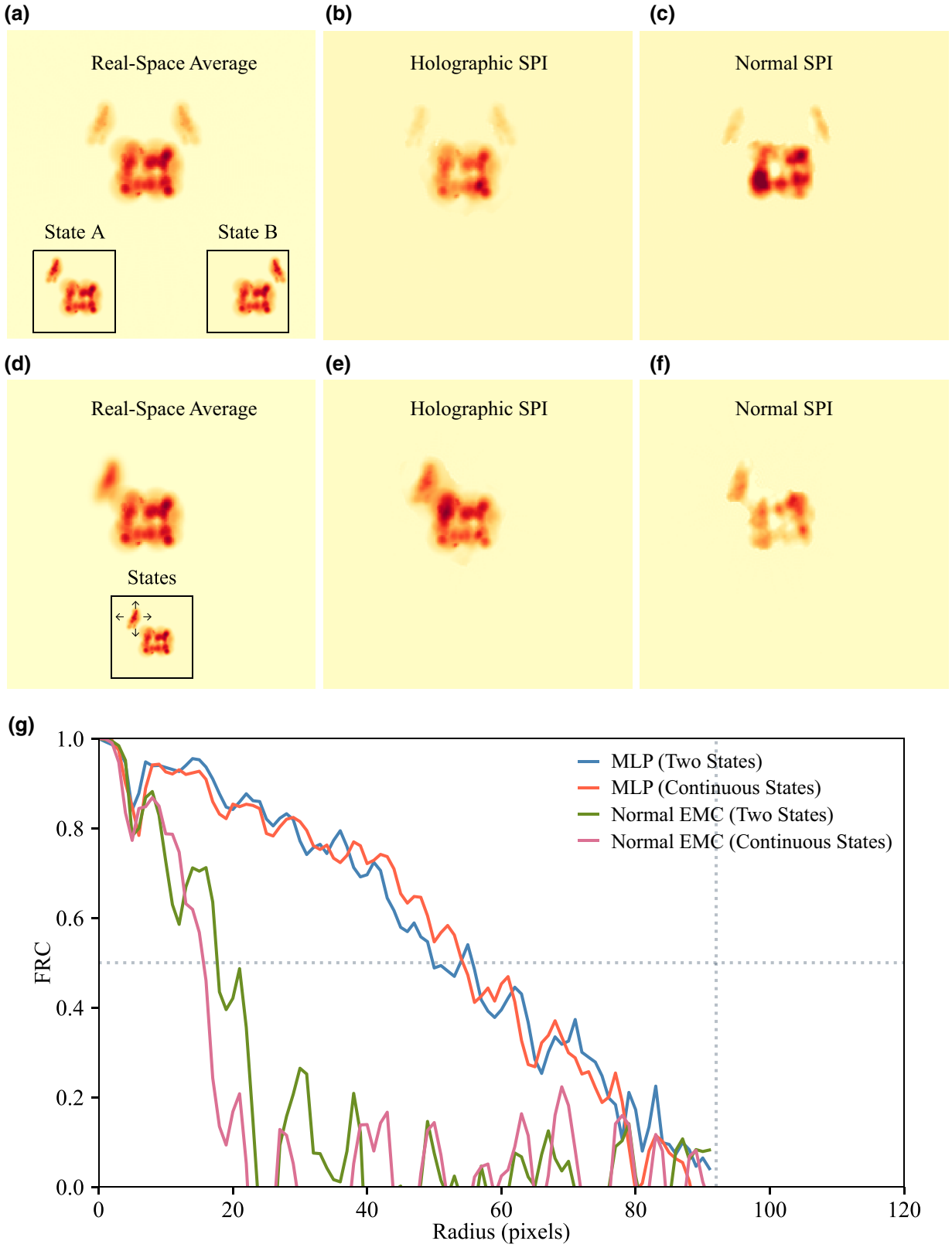

FIG. 6. Simulation results for structural heterogeneity in the target object. (a) Average structure of target object that exists in one of two discrete states (state A and state B, shown in inset). (b) Reconstructed real-space target object with an AuNP sphere attached using MLP. (c) Reconstructed real-space target object without reference attached. (d) Average structure of target object that exists in a continuous distribution of states. Inset shows the target object with a random subunit where the arrow depicts the free direction of motion for a subunit in each shot. (e) Reconstructed real-space target object with an AuNP sphere attached using MLP. (f) Reconstructed real-space target object without reference attached. (g) Comparison of FRC between the holographic SPI and normal SPI for the different scenarios.

and a moving subunit that exists in different states shot to shot. Without explicitly modifying the MLP algorithm or providing any information about such heterogeneity, the algorithm analyzes the diffraction dataset simulated from these target objects. In the first case, the moving subunit object exists in two distinct discrete states: state A and state B, as shown in the inset of Fig. 6(a). The larger subunit remains in the same position, whereas the smaller subunit occupies one of two opposite states about the center of the larger one.

The diffraction patterns are now generated from this heterogeneous ensemble in two ways, one with the AuNP attached and the other with just the target object (as in conventional SPI). In the former case, we still have the shot-to-shot variations in the AuNP diameter and relative position as before, which are recovered with the MLP procedure developed above. For the latter case, conventional intensity reconstruction is performed with the EMC algorithm, followed by phase retrieval. Figures 6(b) and 6(c) show the reconstructions from the holographic and normal SPI scenarios. In the holographic case the reconstruction is an object very similar to the real-space average of structure shown in Fig. 6(a). However, in the absence of a reference, as shown in Fig. 6(c), the reconstruction after phase retrieval is quite poor. This is also seen in the FRC comparisons with the real-space average in Fig. 6(g).

We would like to stress that this result, while expected, is not trivial. First, we are still measuring real-valued intensities and not complex Fourier amplitudes. Second, each pattern has a different set of latent variables, leading to different intensity distributions on the detector. The difference is that the MLP algorithm does not average these intensities directly, but rather fits a common complex Fourier transform of the target object to all these intensities, which turns out to reflect the average object over all patterns.

2. Continuous heterogeneity

The previous example addressed the extreme case of having the moving subunit in two diametrically opposite locations. We now investigate a more realistic scenario where the moving subunit occupies a continuous local distribution of positions. The displacements are sampled from a normal distribution with $\sigma = 0.5$ pixels, as represented by arrows in the inset of Fig. 6(d). The dataset consists of diffraction patterns generated from the target object where the smaller subunit is in a different position every frame. Figures 6(e) and 6(f) show the MLP and normal SPI reconstructions as before. In the holographic case the homogeneous part is reconstructed to a good resolution while the wobbling subunit part has been reconstructed as a blurry object, as expected in a real-space average. This is in contrast to the normal SPI case, where the resolution of the reconstruction is globally affected due to the averaging in intensity space, leading to loss of contrast at higher q. Once again, Fig. 6(g) shows that the holographic SPI reconstruction compares much more favorably with the average structure.

The critical aspect of this robustness to heterogeneity seems to be the presence of translational latent parameters. For the continuous translation case, one can envision sharpening the small subunit reconstruction by choosing the AuNP shifts relative to that subunit rather than the whole object. However, the current method will struggle to solve for a conjugate of two unknown objects with a shot-by-shot variation in their relative position. This is because of the assumption that the signal at any given q is composed of a known reference and the target object, and if part of the object itself is used as the reference, its Fourier transform at high q is not known *a priori* before sharpening.

IV. CONCLUSIONS

In conventional SPI experiments the biggest challenge in reaching sub-nanometer resolutions is to collect a large number of diffraction patterns with sufficiently low background in order to enable orientation determination and averaging. The introduced holography-based imaging methodology overcomes this issue by attaching a strong scattering object (a reference) to the sample of interest. This significantly improves background tolerance, but adds a large amount of computational complexity by introducing more unknown parameters associated with each pattern defining the conjugate system. In this paper we describe an algorithm based on maximum likelihood estimation using a pattern search called MLP that enables one to scale the method to high resolution and to weakly scattering objects. The results above show how it significantly outperforms the previously published method in the cases of low signal and without the need for a large number of intermediate average intensity models.

We extend the application of MLP to retrieve the structure of the target object for the scenario where it has shot-to-shot structural heterogeneity. Two cases of heterogeneity are investigated: one in which a subunit is in two discrete conformational states and one in which the subunit is in a continuous distribution of states. In both cases the algorithm reconstructs an object that is the real-space average of structure in all the conformation states, without prior knowledge that the target object is varying from shot-to-shot. On the contrary, conventional SPI without an added holographic reference generates an average over the intensities in Fourier space that on phase retrieval yields a poorly reconstructed object not just in the vicinity of the moving subunit, but globally due to the delocalized nature of the Fourier transform.

The algorithm can also be applied when the reference is a 2D crystal lattice and the target object is located in one of the unit cells. In such scenarios, one solves for relative shifts between the center of the unit cell and target object in each shot and the size of the unit cell.

In future work we will explore the possibility of adjusting the *xy* shifts in order to selectively reconstruct different regions of the particle to varying degrees of sharpness. We also plan to explicitly incorporate structural heterogeneity either as multiple discrete classes [5] or as continuous latent variables [30]. While intensity-space classifications have been performed previously, the holographic approach may enable a more fine-grained approach similar to that applied in cryogenic electron microscopy [31].

The data generation and algorithm codes used in this work are open access and are available at https://github.com/AyyerLab/RefSPI-MaxLP.

ACKNOWLEDGMENTS

The authors acknowledge valuable discussions with Tamme Wollweber, Yulong Zhuang, Zhou Shen, and Parichita Mazumder. This work is supported by the Cluster of Excellence 'CUI: Advanced Imaging of Matter' of the Deutsche Forschungsgemeinschaft (DFG) - EXC 2056 - project ID 390715994.

[1] A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. DePonte, P. Drell, S. Doniach, K. H. Downing, and T. Earnest, *et al.*, The linac coherent light source single particle imaging road map, Struct. Dyn. **2**, 041701 (2015).

[2] Henry N. Chapman, Anton Barty, Michael J. Bogan, Sébastien Boutet, Matthias Frank, Stefan P. Hau-Riege, Stefano Marchesini, Bruce W. Woods, Saša Bajt, and W. Henry Benner, *et al.*, Femtosecond diffractive imaging with a soft-x-ray free-electron laser, Nat. Phys. **2**, 839 (2006).

[3] Egor Sobolev, Sergei Zolotarev, Klaus Giewekemeyer, Johan Bielecki, Kenta Okamoto, Hemanth K. N. Reddy, Jakob Andreasson, Kartik Ayyer, Imrich Barak, and Sadia Bari, *et al.*, Megahertz single-particle imaging at the European XFEL, Commun. Phys. **3**, 1 (2020).

[4] Chun Hong Yoon, Peter Schwander, Chantal Abergel, Inger Andersson, Jakob Andreasson, Andrew Aquila, Saša Bajt, Miriam Barthelmess, Anton Barty, and Michael J. Bogan, *et al.*, Unsupervised classification of single-particle x-ray diffraction snapshots by spectral clustering, Opt. Express **19**, 16542 (2011).

[5] Kartik Ayyer, P. Lourdu Xavier, Johan Bielecki, Zhou Shen, Benedikt J. Daurer, Amit K. Samanta, Salah Awel, Richard Bean, Anton Barty, and Martin Bergemann, *et al.*, 3D diffractive imaging of nanoparticle ensembles using an x-ray laser, Optica **8**, 15 (2021).

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

[7] Kartik Ayyer, T.-Y. Lan, Veit Elser, and N. Duane Loh, Dragonfly: An implementation of the expand–maximize–compress algorithm for single-particle imaging, J. Appl. Crystallogr. **49**, 1320 (2016).

[8] James R. Fienup, Reconstruction of an object from the modulus of its fourier transform, Opt. Lett. **3**, 27 (1978).

[9] Veit Elser, Phase retrieval by iterated projections, JOSA A **20**, 40 (2003).

[10] Ida V. Lundholm, Jonas A. Sellberg, Tomas Ekeberg, Max F. Hantke, Kenta Okamoto, Gijs van der Schot, Jakob Andreasson, Anton Barty, Johan Bielecki, and Petr Bruza, *et al.*, Considerations for three-dimensional image reconstruction from experimental data in coherent diffractive imaging, IUCrJ **5**, 531 (2018).

[11] Kartik Ayyer, Andrew J. Morgan, Andrew Aquila, Hasan DeMirci, Brenda G. Hogue, Richard A. Kirian, P. Lourdu Xavier, Chun Hong Yoon, Henry N. Chapman, and Anton Barty, Low-signal limit of x-ray single particle diffractive imaging, Opt. Express **27**, 37816 (2019).

[12] Tomas Ekeberg, Martin Svenda, Chantal Abergel, Filipe R. N. C. Maia, Virginie Seltzer, Jean-Michel Claverie, Max Hantke, Olof Jönsson, Carl Nettelblad, and Gijs van der Schot, *et al.*, Three-Dimensional Reconstruction of the Giant Mimivirus Particle with an X-Ray Free-Electron Laser, Phys. Rev. Lett. **114**, 098102 (2015).

[13] Max Rose, Sergey Bobkov, Kartik Ayyer, Ruslan P. Kurta, Dmitry Dzhigaev, Young Yong Kim, Andrew J. Morgan, Chun Hong Yoon, Daniel Westphal, and Johan Bielecki, *et al.*, Single-particle imaging without symmetry constraints at an x-ray free-electron laser, IUCrJ **5**, 727 (2018).

[14] Hugh T. Philipp, Kartik Ayyer, Mark W. Tate, Veit Elser, and Sol M. Gruner, Solving structure with sparse, randomly-oriented x-ray data, Opt. Express **20**, 13129 (2012).

[15] I. Poudyal, M. Schmidt, and P. Schwander, Single-particle imaging by x-ray free-electron lasers–how many snapshots are needed?, Struct. Dyn. **7**, 024102 (2020).

[16] Anna Munke, Jakob Andreasson, Andrew Aquila, Salah Awel, Kartik Ayyer, Anton Barty, Richard J. Bean, Peter Berntsen, Johan Bielecki, and Sébastien Boutet, *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). data Descriptor,

[17] Johan Bielecki, Max F. Hantke, Benedikt J. Daurer, Hemanth K. N. Reddy, Dirk Hasse, Daniel S. D. Larsson, Laura H. Gunn, Martin Svenda, Anna Munke, and Jonas A. Sellberg, *et al.*, Electrospray sample injection for single-particle imaging with x-ray lasers, Sci. Adv. **5**, eaav8801 (2019).

[18] Henry N. Chapman, Petra Fromme, Anton Barty, Thomas A. White, Richard A. Kirian, Andrew Aquila, Mark S. Hunter, Joachim Schulz, Daniel P. DePonte, and Uwe Weierstall, *et al.*, Femtosecond x-ray protein nanocrystallography, Nature **470**, 73 (2011).

[19] Raymond G. Sierra, Hartawan Laksmono, Jan Kern, Rosalie Tran, Johan Hattne, Roberto Alonso-Mori, Benedikt Lassalle-Kaiser, Carina Glöckner, Julia Hellmich, and Donald W. Schafer, *et al.*, Nanoflow electrospinning serial femtosecond crystallography, Acta Crystallogr. Sect. D **68**, 1584 (2012).

[20] Mark S. Hunter, Brent Segelke, Marc Messerschmidt, Garth J. Williams, Nadia A. Zatsepin, Anton Barty, W. Henry Benner, David B. Carlson, Matthew Coleman, and Alexander Graf, *et al.*, Fixed-target protein serial microcrystallography with an x-ray free electron laser, Sci. Rep. **4**, 6026 (2014).

[21] Daewoong Nam, Chan Kim, Yoonhee Kim, Tomio Ebisu, Marcus Gallagher-Jones, Jaehyun Park, Sunam Kim, Sangsoo Kim, Kensuke Tono, Makina Yabashi, Tetsuya Ishikawa, and Changyong Song, Fixed target single-shot imaging of nanostructures using thin solid membranes at sacla, J. Phys. B: At. Mol. Opt. Phys. **49**, 034008 (2016).

[22] Carolin Seuring, Kartik Ayyer, Eleftheria Filippaki, Miriam Barthelmess, Jean-Nicolas Longchamp, Philippe Ringler, Tommaso Pardini, David H. Wojtas, Matthew A. Coleman, and Katerina Dörner, *et al.*, Femtosecond x-ray coherent diffraction of aligned amyloid fibrils on low background graphene, Nat. Commun. **9**, 1836 (2018).

[23] W. Decking, S. Abeghyan, P. Abramian, A. Abramsky, A. Aguirre, C. Albrecht, P. Alou, M. Altarelli, P. Altmann, K. Amyan, V. Anashin, M. Mommerz, L. Monaco, C. Montiel, M. Moretti, I. Morozov, P. Morozov, and D. Mross, *et al.*, A MHz-repetition-rate hard x-ray free-electron laser driven by a superconducting linear accelerator, Nat. Photon. **14**, 391 (2020).

[24] Kartik Ayyer, Reference-enhanced x-ray single-particle imaging, Optica **7**, 593 (2020).

[25] Simon Gravel and Veit Elser, Divide and concur: A general approach to constraint satisfaction, Phys. Rev. E **78**, 036706 (2008).

[26] Franz Pfeiffer, X-ray ptychography, Nat. Photon. **12**, 9 (2018).

[27] Pierre Thibault and Manuel Guizar-Sicairos, Maximum-likelihood refinement for coherent diffractive imaging, New J. Phys. **14**, 063004 (2012).

[28] Virginia Torczon, On the convergence of pattern search algorithms, SIAM J. Optim. **7**, 1 (1997).

[29] Andrew V. Martin, N. D. Loh, Christina Y. Hampton, Raymond G. Sierra, Fenglin Wang, Andrew Aquila, Saša Bajt, Miriam Barthelmess, Christoph Bostedt, and John D. Bozek, *et al.*, Femtosecond dark-field imaging with an x-ray free electron laser, Opt. Express **20**, 13501 (2012).

[30] Yulong Zhuang, Salah Awel, Anton Barty, Richard Bean, Johan Bielecki, Martin Bergemann, Benedikt J. Daurer, Tomas Ekeberg, Armando D. Estillore, and Hans Fangohr, *et al.*, Unsupervised learning approaches to characterizing heterogeneous samples using x-ray single-particle imaging, IUCrJ **9**, 204 (2022).

[31] Ellen D. Zhong, Tristan Bepler, Bonnie Berger, and Joseph H. Davis, CryoDRGN: Reconstruction of heterogeneous cryo-EM structures using neural networks, Nat. Methods **18**, 176 (2021).