

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, 22761 Hamburg, Germany*

²*Center for Free Electron Laser Science, 22761 Hamburg, Germany*

³*The Hamburg Center for Ultrafast Imaging, 22761 Hamburg, Germany*

Single particle imaging (SPI) at X-ray free electron lasers (XFELs) is a technique to determine the 3D structure of nanoscale objects like biomolecules from a large number of diffraction patterns of copies of these objects in random orientations. Millions of low signal-to-noise diffraction patterns with unknown orientation are collected during an X-ray SPI experiment. The patterns are then analyzed and merged using a reconstruction algorithm to retrieve the full 3D-structure of particle. The resolution of reconstruction is limited by background noise, signal-to-noise ratio in diffraction patterns and total amount of data collected. We recently introduced a reference-enhanced holographic single particle imaging methodology [Optica **7**,593-601(2020)] to collect high enough signal-to-noise and background tolerant patterns and a reconstruction algorithm to recover missing parameters beyond orientation and then directly retrieve the full Fourier model of the sample of interest. Here we describe a phase retrieval algorithm based on maximum likelihood estimation using pattern search dubbed as MaxLP, with better scalability for fine sampling of latent parameters 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-nm resolution biomolecule imaging.

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 3D structure of the object [3].

This analysis often broadly follows the following three steps: First, diffraction patterns containing signal 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 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 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 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 substantial signal from 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 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 were a spherical gold nanoparticle (AuNP) and a 2D crystal lattice with unit cell size comparable to target object. The total scattered signal is

* kartik.ayyer@mpsds.mpg.de

increased for each shot when a reference is attached and the signal now becomes more tolerant to background. A reconstruction algorithm was 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 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-EMC algorithm [24] which 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 Section 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 pattern search dubbed as MaxLP. MaxLP scales efficiently with sampling of latent parameters and total number of diffraction patterns in comparison with 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 Section II and its performance is shown with simulated data in Section III A.

We also discuss the performance of the holographic-SPI method with the MaxLP 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 MaxLP algorithm with the holographic SPI experimental setup shows substantial promise in pushing the SPI technique to sub-nm resolution.

II. RECONSTRUCTION ALGORITHM

In a holographic SPI experiment, the sample of interest (target object) is in the vicinity of a strongly scattering

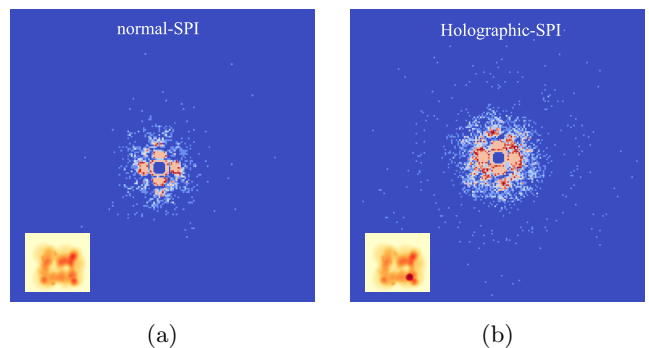


FIG. 1. (a) Simulated diffraction pattern for a random object in a normal single-particle imaging (SPI) case with 4810 photons. (b) Diffraction pattern from the same particle with a small AuNP attached (10,000 photons). If background is present, structural information can be discerned at higher scattering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

reference, which in this work, we take to be a spherical gold nanoparticle (AuNP). 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: unknown orientation θ , diameter of AuNP d and relative positions of AuNP and target object in 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 $\rho_o(\mathbf{r})$

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

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

$$I(\mathbf{q}, D, \mathbf{t}) = |F_o(\mathbf{q}) + F_s(\mathbf{q}, D)e^{2\pi i \mathbf{q} \cdot \mathbf{t}}|^2, \quad (2)$$

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 of 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. 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 which improve the SNR in the presence of background at these resolutions. 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 diameter of the AuNP and its relative position. In order to retrieve the target object's structure which 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 EMC algorithm (Ref-EMC) was developed for holographic SPI [24]. The conventional EMC algorithm involves three steps in each iteration: *Expand*, *Maximize* and *Compress* which iteratively update the 3D intensity model to one which 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 Ref-EMC, 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 these 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 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 which quickly becomes computationally expensive. Additionally, this method is composed of projection operations minimizing a Euclidean error metric. This implicitly assumes a Gaussian error distribution which 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 pattern search technique dubbed as Maximum Likelihood Phaser (MaxLP) 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 was tested for different intensity signal levels in diffraction datasets with 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 MaxLP 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 to 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 Eqn 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

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

Here F_o is the complex number which 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 ($\Delta s = \mathbf{q} \cdot \mathbf{t}$) between the AuNP and the target object.

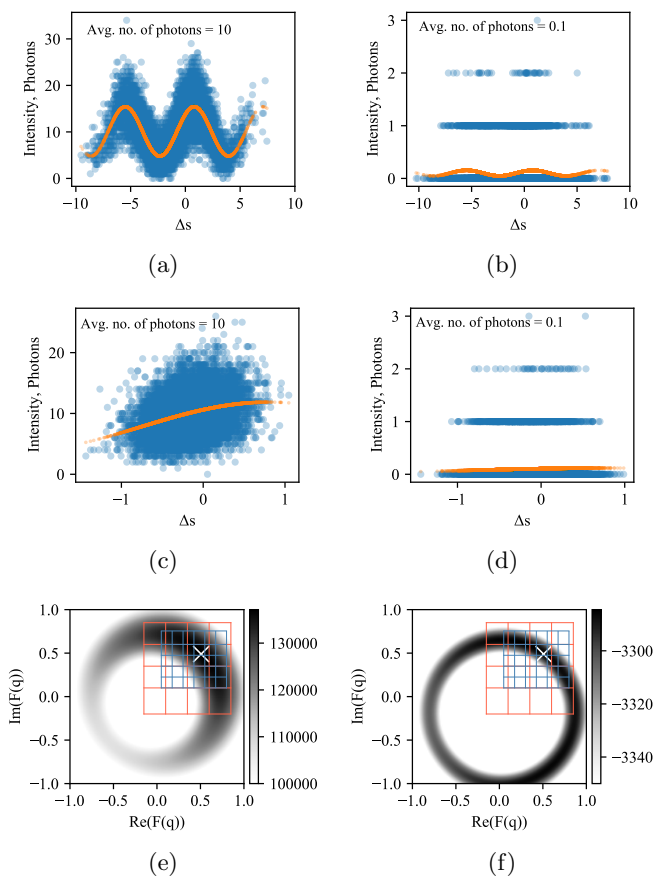


FIG. 2. Intensity behavior at a single detector pixel for a holographic-SPI diffraction data. (a)-(d) Average number of photons *vs* relative phase shifts of gold sphere with respect to the target object. The ideal intensity on the same pixel is represented by the orange line. (e) & (f) Log-likelihood landscapes obtained from diffraction data for case (a) & (d), respectively. The ‘x’ denotes the true value of $F(\mathbf{q})$ maximizing the log-likelihood. The grid lines (1st - orange and 2nd - blue) depict the size of pattern search grid in consecutive iterations.

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

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

where $I_{\text{pix},d}$ is 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 . Figure 2(e-f) show this log-likelihood function distribution in the complex plane of $F_{o,\text{pix}}$ for the diffraction pattern data given from Fig. 2(a) and (d), respectively. The likelihood has a well behaved landscape with a sharp maxima with different values of real and imaginary part of F_o , when the average photon count is high enough with 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 the 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, the model at neighboring pixels are strongly correlated. This information will be taken into account later in Section 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 found that derivative based approaches were not robust, with maximization of likelihood failing for low photon count cases and small shifts. We found reasonable success with a pattern search approach [28] for two reasons: it could be efficiently parallelized on the graphical processing units (GPUs) and secondly it was quite robust in low signal and small shift cases.

Pattern search optimization is essentially a non-derivative technique that does not require a gradient calculation in its update of parameters. The search begins with a 2D grid of values of \mathbf{F} (see Fig. 2(e-f), orange grid lines). At each iteration, this grid moves to a set of values which best maximizes the likelihood function. If it finds a set of values of \mathbf{F} which does not have better likelihood then it stays at 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.

III. RESULTS

We tested the performance of the algorithm described above on 2D simulated data with a single angular degree of freedom. The target object was a randomly generated agglomeration of small spheres, approximating the electron density distribution of a biological sample. Each diffraction pattern was 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 was roughly 1/5th the size of the target and the contrast was 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 x - and y - directions between their centers were sampled from a normal distribution with a standard deviation of 1 pixel. The diameter of AuNP was also normally distributed with a mean and standard deviation of 7 and 0.5 pixels, respectively. The intensity distribution from this composite object was rotated in-plane by a uniform random angle before Poisson sampling. The diffraction patterns were collected on a circular detector with a diameter of 185 pixels and a central hole with a 4 pixel radius.

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

Figure 3(a) and (b) depicts the absolute magnitude of the reconstructed complex Fourier model for the highest and lowest signal levels. 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 were filled in using the following approach. First, a support mask was calculated using data from moderate q using the dark-field diffractive imaging approach [29]. This mask was then used as a real-space constraint and the difference map iterative phase retrieval algorithm [9] was 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 Figure 3(c). We compare the performance of MaxLP 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 *di-*

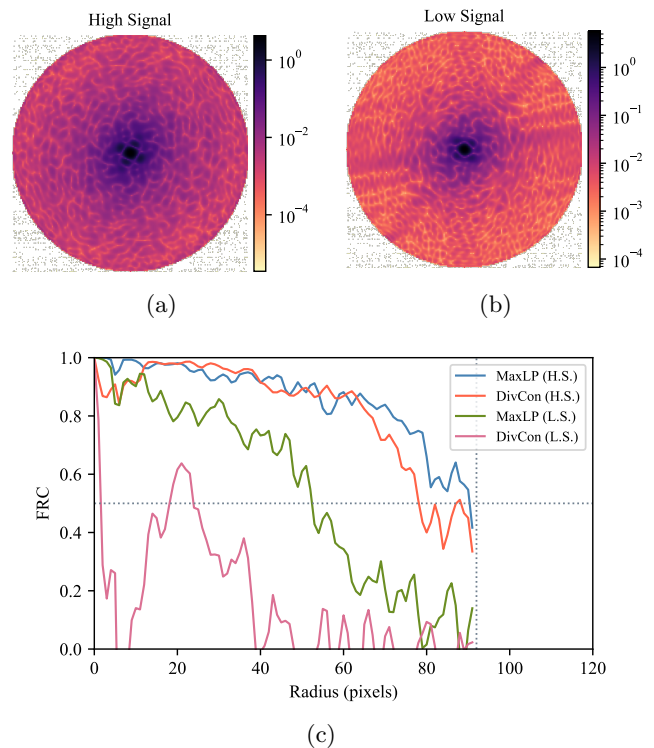


FIG. 3. Simulation results for homogeneous target object attached to a spherical AuNP as reference. (a) Magnitude of reconstructed Fourier model for high signal diffraction data with MaxLP method (10^5 ph/frame). (b) Same for low signal diffraction data (2000 ph/frame). Both datasets had 10^4 frames. (c) Comparison of Fourier ring correlation (FRC) between reconstructions and ground truth with different signal level diffraction data. Maximum-likelihood Phaser (MaxLP), Divide and Concur (DivCon), High Signal (H.S.) and Low Signal (L.S.)

vide and *concur* approach fails for the low signal case, whereas the MaxLP reconstructs the Fourier model with $FRC > 0.5$ up to $q = 55$ pixels. This can be attributed to the MaxLP 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 AuNP, in-plane orientation and relative shifts in x - and y - directions, respectively. The sampling of the in-plane orientation is in range of 0 and 180° with step size of 2° . In Fig. 4(a) one can see distribution of errors between the retrieved and true in-plane angles for high and low signal level, respectively. The centers are shifted since the reconstruction has an arbitrary overall orientation. As expected, the distribution is narrow for high signal, with $\sigma = 0.85^\circ$ which is below the sampling rate, and significantly broader for low signal with $\sigma = 11.4^\circ$.

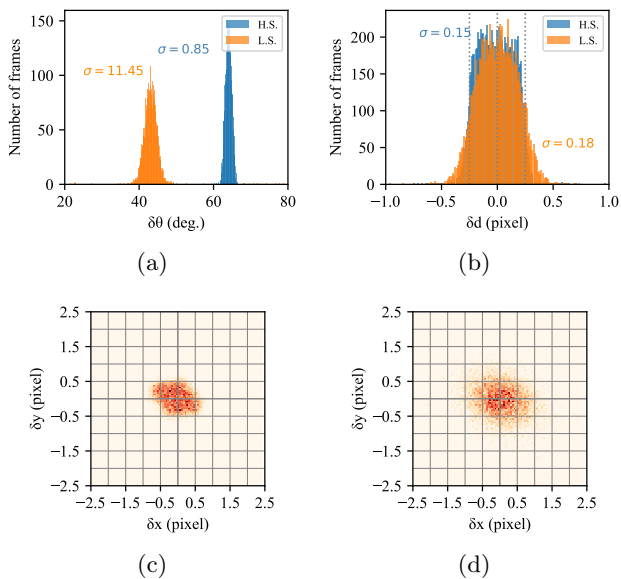


FIG. 4. Error distribution for reconstructed latent parameters for homogeneous target object attached with a spherical AuNP as reference for low (L.S.) and high signal (H.S.) levels. (a) In-plane orientation, (b) diameter of gold sphere, (c) & (d) relative displacements between center of target object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

Similarly, AuNP diameter and x - & 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 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 which can be obtained in small-angle X-ray scattering. Figure 4(c) & (d) show the error distribution of shifts in x - and y - direction for the two 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 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-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 accurate estimation of latent parameters as well as total signal from all the

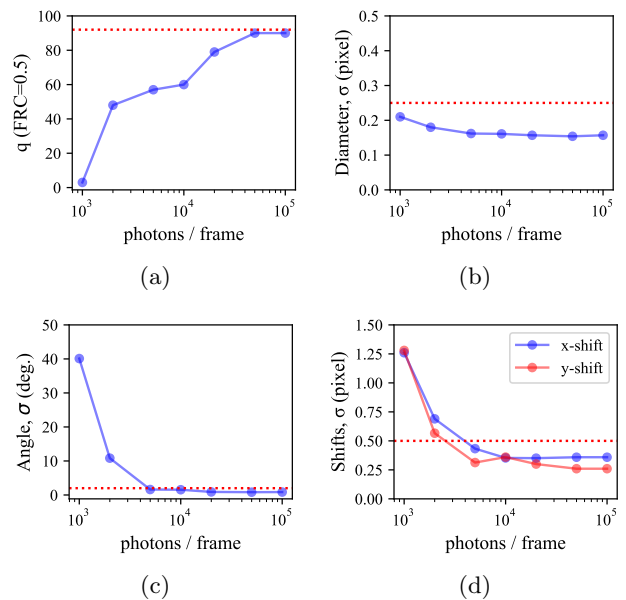


FIG. 5. Evaluation metrics of Maximum-likelihood phaser (MaxLP) for different photon signal levels. (a) q value when FRC = 0.5 vs number of photons/frame. The horizontal dashed line corresponds to a resolution of 1 real-space pixel. Standard deviation values for (b) diameter errors, (c) orientation errors,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

frames in 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 sub-units that can move following a continuous landscape of various conformational states. The study of these conformational changes can help us understand the function of such biomolecules. However, there are significant computational challenges to retrieve conformational states and structure of 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's 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 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

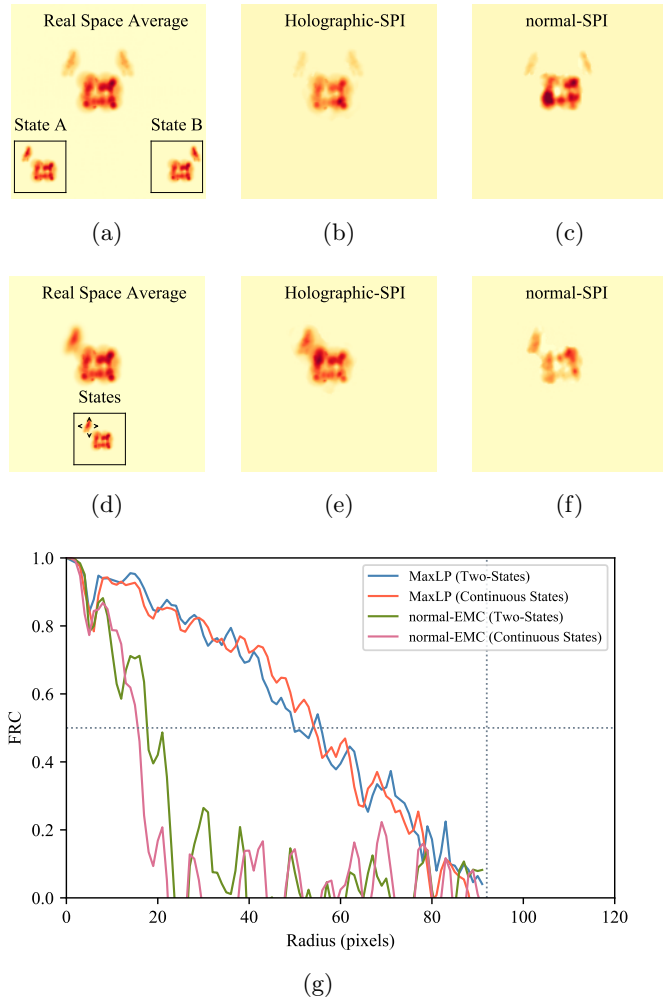


FIG. 6. Simulation results for structural heterogeneity in the target object (a) Average structure of target object which exists in one of two discrete states (State A & State B, shown in inset). (b) Reconstructed real-space target object with an AuNP sphere attached using MaxLP. (c) Reconstructed real-space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

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) and a moving sub-unit that exists in different states shot-to-shot. Without explicitly modifying the MaxLP 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 sub-unit object exists in two distinct discrete states: State A & State B, as shown in inset of Fig. 6(a). The larger sub-unit remains in same position, whereas the smaller sub-unit 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 MaxLP procedure developed above. For the latter case, conventional intensity reconstruction is performed with the EMC algorithm, followed by phase retrieval. Figure 6(b-c) shows 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. Secondly, each pattern has a different set of latent variables, leading to different intensity distributions on the detector. The difference is that the MaxLP 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 sub-unit in two diametrically opposite locations. We now investigate a more realistic scenario where the moving sub-unit 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 sub-unit is in a different position every frame. Figure 6(e-f) show the MaxLP and normal SPI reconstructions as before. In the holographic case, the homogeneous part is reconstructed to a good resolution while the wobbling sub-unit part has been re-

constructed 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, the 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 sub-unit reconstruction by choosing the AuNP shifts relative to that sub-unit rather than whole object. Such an option is not possible in conventional SPI where the diffraction pattern is invariant to translating the object within the X-ray focus.

IV. CONCLUSIONS

In conventional SPI experiments, the biggest challenge in reaching sub-nm 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 described an algorithm based on maximum-likelihood estimation using pattern search called MaxLP 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 extended the application of MaxLP to retrieve the structure of the target object for the scenario where it has shot-to-shot structural heterogeneity. Two cases of heterogeneity were investigated: one in which a sub-unit

is in two discrete conformational states and one in which the sub-unit is in a continuous distribution of states. In both the cases, the algorithm reconstructs an object that is real space average of structure in all the conformation states, without prior knowledge that the target object was varying from shot-to-shot. On the contrary, the 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 sub-unit, 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 unit cell and target object in each shot and the size of 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 Github.

V. ACKNOWLEDGMENTS

The author acknowledges the extremely valuable discussions with Tamme Wollweber, Yulong Zhuang, Zhou Shen and Parichita Mazumder. This research is funded by Max Planck Society.

VI. DISCLOSURES

The authors declare no conflict of interest.

-
- [1] A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. DePonte, P. Drell, S. Doniach, K. Downing, T. Earnest, *et al.*, *Structural Dynamics* **2**, 041701 (2015).
 - [2] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. Bajt, W. H. Benner, *et al.*, *Nature Physics* **2**, 839 (2006).
 - [3] E. Sobolev, S. Zolotarev, K. Giewekemeyer, J. Bielecki, K. Okamoto, H. K. Reddy, J. Andreasson, K. Ayer, I. Barak, S. Bari, *et al.*, *Communications Physics* **3**, 1 (2020).
 - [4] C. H. Yoon, P. Schwander, C. Abergel, I. Andersson, J. Andreasson, A. Aquila, S. Bajt, M. Barthelmeß, A. Barty, M. J. Bogan, *et al.*, *Opt. Express* **19**, 16542 (2011).
 - [5] K. Ayer, P. L. Xavier, J. Bielecki, Z. Shen, B. J. Daurer, A. K. Samanta, S. Awel, R. Bean, A. Barty, M. Bergemann, *et al.*, *Optica* **8**, 15 (2021).
 - [6] N.-T. D. Loh and V. Elser, *Physical Review E* **80**, 026705 (2009).
 - [7] K. Ayer, T.-Y. Lan, V. Elser, and N. D. Loh, *Journal of applied crystallography* **49**, 1320 (2016).
 - [8] J. R. Fienup, *Optics letters* **3**, 27 (1978).
 - [9] V. Elser, *JOSA A* **20**, 40 (2003).
 - [10] I. V. Lundholm, J. A. Sellberg, T. Ekeberg, M. F. Hantke, K. Okamoto, G. van der Schot, J. Andreasson, A. Barty, J. Bielecki, P. Bruza, *et al.*, *IUCrJ* **5**, 531 (2018).

- [11] K. Ayer, A. J. Morgan, A. Aquila, H. DeMirci, B. G. Hogue, R. A. Kirian, P. L. Xavier, C. H. Yoon, H. N. Chapman, and A. Barty, *Optics Express* **27**, 37816 (2019).
- [12] T. Ekeberg, M. Svenda, C. Abergel, F. R. N. C. Maia, V. Seltzer, J.-M. Claverie, M. Hantke, O. Jönsson, C. Nettelblad, G. van der Schot, *et al.*, *Phys. Rev. Lett.* **114**, 098102 (2015).
- [13] M. Rose, S. Bobkov, K. Ayer, R. P. Kurta, D. Dzhigaev, Y. Y. Kim, A. J. Morgan, C. H. Yoon, D. Westphal, J. Bielecki, *et al.*, *IUCrJ* **5** (2018).
- [14] H. T. Philipp, K. Ayer, M. W. Tate, V. Elser, and S. M. Gruner, *Optics express* **20**, 13129 (2012).
- [15] I. Poudyal, M. Schmidt, and P. Schwander, *Structural Dynamics* **7**, 024102 (2020).
- [16] A. Munke, J. Andreasson, A. Aquila, S. Awel, K. Ayer, A. Barty, R. J. Bean, P. Berntsen, J. Bielecki, S. Boutet, *et al.*, *Scientific Data* **3**, 160064 (2016), data Descriptor.
- [17] J. Bielecki, M. F. Hantke, B. J. Daurer, H. K. N. Reddy, D. Hasse, D. S. D. Larsson, L. H. Gunn, M. Svenda, A. Munke, J. A. Sellberg, *et al.*, *Science Advances* **5** (2019), 10.1126/sciadv.aav8801.
- [18] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, *et al.*, *Nature* **470**, 73 (2011).
- [19] R. G. Sierra, H. Laksmono, J. Kern, R. Tran, J. Hattné, R. Alonso-Mori, B. Lassalle-Kaiser, C. Glöckner, J. Hellmich, D. W. Schafer, *et al.*, *Acta Crystallographica Section D* **68**, 1584 (2012).
- [20] M. S. Hunter, B. Segelke, M. Messerschmidt, G. J. Williams, N. A. Zatsepin, A. Barty, W. H. Benner, D. B. Carlson, M. Coleman, A. Graf, *et al.*, *Scientific reports* **4**, 6026 (2014).
- [21] D. Nam, C. Kim, Y. Kim, T. Ebisu, M. Gallagher-Jones, J. Park, S. Kim, S. Kim, K. Tono, M. Yabashi, T. Ishikawa, and C. Song, *Journal of Physics B: Atomic, Molecular and Optical Physics* **49**, 034008 (2016).
- [22] C. Seuring, K. Ayer, E. Filippaki, M. Barthelmess, J.-N. Longchamp, P. Ringler, T. Pardini, D. H. Wojtas, M. A. Coleman, K. Dörner, *et al.*, *Nature Communications* **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, D. Mross, *et al.*, *Nature Photonics* , 391 (2020).
- [24] K. Ayer, *Optica* **7**, 593 (2020).
- [25] S. Gravel and V. Elser, *Physical Review E* **78**, 036706 (2008).
- [26] F. Pfeiffer, *Nature Photonics* **12**, 9 (2018).
- [27] P. Thibault and M. Guizar-Sicairos, *New Journal of Physics* **14**, 063004 (2012).
- [28] V. Torczon, *SIAM Journal on optimization* **7**, 1 (1997).
- [29] A. V. Martin, N. Loh, C. Y. Hampton, R. G. Sierra, F. Wang, A. Aquila, S. Bajt, M. Barthelmess, C. Bostedt, J. D. Bozek, *et al.*, *Optics express* **20**, 13501 (2012).
- [30] Y. Zhuang, S. Awel, A. Barty, R. Bean, J. Bielecki, M. Bergemann, B. J. Daurer, T. Ekeberg, A. D. Estillore, H. Fangohr, *et al.*, *IUCrJ* **9** (2022).
- [31] E. D. Zhong, T. Bepler, B. Berger, and J. H. Davis, *Nature methods* **18**, 176 (2021).