



# Skopi: a simulation package for diffractive imaging of noncrystalline biomolecules

Ariana Peck,<sup>a,‡</sup> Hsing-Yin Chang,<sup>a,‡</sup> Antoine Dujardin,<sup>a</sup> Deeban Ramalingam,<sup>a</sup> Monarin Uervirojnangkoorn,<sup>a</sup> Zhaoyou Wang,<sup>a</sup> Adrian Mancuso,<sup>b,c</sup> Frédéric Poitevin<sup>a</sup> and Chun Hong Yoon<sup>a\*</sup>

Received 14 December 2021

Accepted 3 June 2022

Edited by A. Barty, DESY, Hamburg, Germany

‡ These authors contributed equally to this work.

**Keywords:** single-particle imaging; fluctuation X-ray scattering; holography; free-electron lasers; simulation.

**Supporting information:** this article has supporting information at journals.iucr.org/j

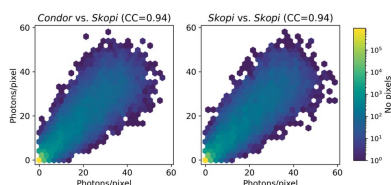
<sup>a</sup>Linac Coherent Light Source, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA,

<sup>b</sup>European XFEL, Holzkoppel 4, 22869 Schenefeld, Germany, and <sup>c</sup>Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia. \*Correspondence e-mail: yoon82@slac.stanford.edu A. Barty, DESY, Hamburg, Germany

X-ray free-electron lasers (XFELs) have the ability to produce ultra-bright femtosecond X-ray pulses for coherent diffraction imaging of biomolecules. While the development of methods and algorithms for macromolecular crystallography is now mature, XFEL experiments involving aerosolized or solvated biomolecular samples offer new challenges in terms of both experimental design and data processing. *Skopi* is a simulation package that can generate single-hit diffraction images for reconstruction algorithms, multi-hit diffraction images of aggregated particles for training machine learning classifiers using labeled data, diffraction images of randomly distributed particles for fluctuation X-ray scattering algorithms, and diffraction images of reference and target particles for holographic reconstruction algorithms. *Skopi* is a resource to aid feasibility studies and advance the development of algorithms for noncrystalline experiments at XFEL facilities.

## 1. Introduction

The unique capabilities of X-ray free-electron laser (XFEL) sources have led to significant advances in structural biology since the first hard X-ray laser began operation at the Linac Coherent Light Source (LCLS) in 2009 (Emma *et al.*, 2010; Jamison, 2010). The XFEL technique of serial femtosecond crystallography has proven particularly transformative for protein crystallographers (Chapman *et al.*, 2011; Boutet *et al.*, 2012; Spence, 2018). Ultra-bright X-rays enable the study of crystals too small or radiation sensitive for synchrotron sources, while the femtosecond pulses have permitted time-resolved studies of enzyme catalysis at atomic resolution (Schlichting, 2015; Stauch & Cherezov, 2018). By contrast, coherent X-ray diffractive imaging (CXDI) of noncrystalline biological samples has remained a largely nascent application, despite these experiments being the original intent of XFEL technology (Neutze *et al.*, 2000; Bogan *et al.*, 2008). Several studies have demonstrated successful single-particle imaging (SPI) of large targets such as viruses and organelles (Seibert *et al.*, 2011; van der Schot *et al.*, 2015; Brändén *et al.*, 2019; Hantke *et al.*, 2014), but 3D reconstruction has been limited to nanoscale resolution (Starodub *et al.*, 2012; Ekeberg *et al.*, 2015; Rose *et al.*, 2018; Kurta *et al.*, 2017). Multi-particle techniques, including fluctuation X-ray scattering (FXS) (Mendez *et al.*, 2014, 2016; Kurta *et al.*, 2017; Doniach, 2018; Pande *et al.*, 2018) and Fourier transform holography (FTH) (Eisebitt *et al.*, 2004; Marchesini *et al.*, 2008; Shintake, 2008; Gorkhover *et al.*, 2018), have also been explored for structure determination. However, CXDI of noncrystalline biological



OPEN ACCESS

Published under a CC BY 4.0 licence

samples remains far from routine, in contrast to the success of XFEL experiments on inorganic materials that provide significantly more signal (Ayyer *et al.*, 2020). Improvements in resolution are needed before these techniques can address novel questions in structural biology.

The principal challenges facing CXDI studies of non-crystalline samples are the low signal relative to instrumental background, the low hit rate due to difficulties with sample delivery and reliable hit finding, and the intrinsic heterogeneity of the imaged particles, which limits the resolution (Bielecki *et al.*, 2020; Daurer *et al.*, 2017). Improvements in sample delivery and in the dynamic range of detectors have increased the quality of the measured signal (Bielecki *et al.*, 2019), but both experimental and algorithmic advances are needed to achieve high-resolution reconstructions from these data. In particular, accurate classification is required both to identify the useful single-particle hits from shots of aggregates or no particles and to cluster images on the basis of conformational heterogeneity (Maia *et al.*, 2009; Yoon *et al.*, 2011; Reddy *et al.*, 2017; Cruz-Chú *et al.*, 2021). In addition, more sophisticated background subtraction methods are needed to isolate the low-intensity particle scattering prior to reconstruction. While methods of performing classification and background removal have been explored and applied to the few experimental data sets available (Assalauova *et al.*, 2020; Shi *et al.*, 2019), it is unclear how well these algorithms will generalize to new experiments.

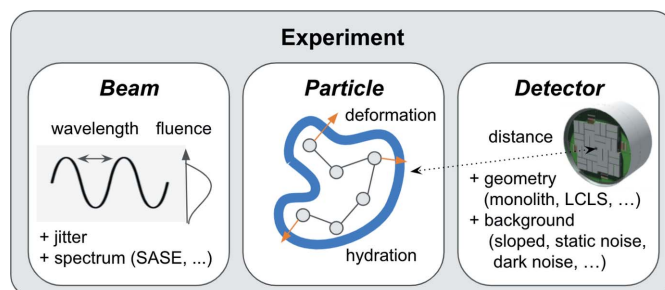
Realistic simulations of CXDI experiments could accelerate the maturation of this field by providing a test bed for advanced data pre-processing and reconstruction algorithms. *Condor* is an existing open-source program already available for CXDI simulations and was created to facilitate planning of experiments at first-generation XFELs (Hantke *et al.*, 2016). However, the ongoing development of different types of SPI experiments with higher repetition rates and more sophisticated detectors poses new challenges and thus calls for tools capable of simulating these technological advances at scale. Here we present *Skopi*, a software package designed for rapid and high-throughput simulations of SPI, FXS and FTH experiments on modern detectors. Additional features of *Skopi* include the ability to simulate single-particle aggregates to facilitate single-particle classification algorithm development using machine learning (Yoon *et al.*, 2011; Shi *et al.*, 2019; Ignatenko *et al.*, 2021) and the modeling of the solvent shell and conformational heterogeneity inherent in biomolecules. The simulated experiments are highly customizable and readily support modeling the beam characteristics and gain auto-ranging detectors available at LCLS as well. *Skopi* also makes it convenient to include diverse sources of noise, including fluctuating dark noise, beam miscentering, a static sloped background and fluence jitter, all of which can impact reconstruction. We anticipate that *Skopi* will be a useful resource for developing new data processing algorithms, guiding experimental design, and in time aiding on-the-fly feedback during CXDI experiments of noncrystalline biological samples by incorporating live beamline measurements to build a digital twin of the experiment.

## 2. Modular implementation

The design of *Skopi* is highly modular, with the three main components of a CXDI experiment – the particle, beam and detector – agnostic to the type of experiment being simulated (Fig. 1). Each component can be modeled with a range of complexity, from the ideal case to more sophisticated representations that mimic experimental errors and noise. We briefly describe these components below.

The particle object stores the biomolecule's atomic coordinates and associated atomic form factors. Coordinates can be supplied as either a PDB or an HDF5 file; the latter format permits coarse-graining by allowing the user to specify 'atom' positions and their corresponding form factors. Conformational heterogeneity can be modeled by sampling along the particle's normal modes to generate alternate states; this is accomplished using the anisotropic network model from the *ProDy* package (Bakan *et al.*, 2011) [Fig. S1(a) in the supporting information]. Although unsuitable for certain types of dynamics like side-chain flips and local unfolding, normal modes analysis offers a general way of modeling large-scale collective motions (Tama & Sanejouand, 2001). Further, the amplitude of these motions can be tuned by choosing distinct modes and the scale of coordinate deformation along each mode, making it easy to generate data sets with varying amounts of heterogeneity. For types of dynamics not accounted for by normal modes, the user could employ external programs to build a library of discrete conformations, which *Skopi* could then sample from to produce a heterogeneous data set.

Another source of sample heterogeneity encountered in CXDI experiments is the disordered solvent shell surrounding aerosolized particles, which helps to preserve structural integrity in a vacuum but increases shot-to-shot variation (Hau-Riege *et al.*, 2007; Mandl *et al.*, 2020). To account for this effect, *Skopi* enables modeling of a hydration layer with a user-defined thickness that follows the particle's contour [Fig. S1(b)]. This solvent mask represents a continuum water model rather than a collection of solvent molecules, and a uniform water density function is sampled at discrete positions within this mask during the diffraction calculation. For



**Figure 1**  
The modular architecture of *Skopi*. The three principal components of each experiment – the beam, particle and detector – are initialized independently of each other. Once these components are set up, diffraction patterns from a range of CXDI experiments can be efficiently simulated.

particles that contain large cavities, the interior portion of the solvent shell may not entirely fill the particle. In these cases, the remaining void can be modeled using a uniform function of the desired scattering density, which we anticipate will be useful for modeling nucleic acid-containing viruses. A library of particles adopting different conformations or with variable degrees of hydration can be created and then sampled from to generate heterogeneous data sets.

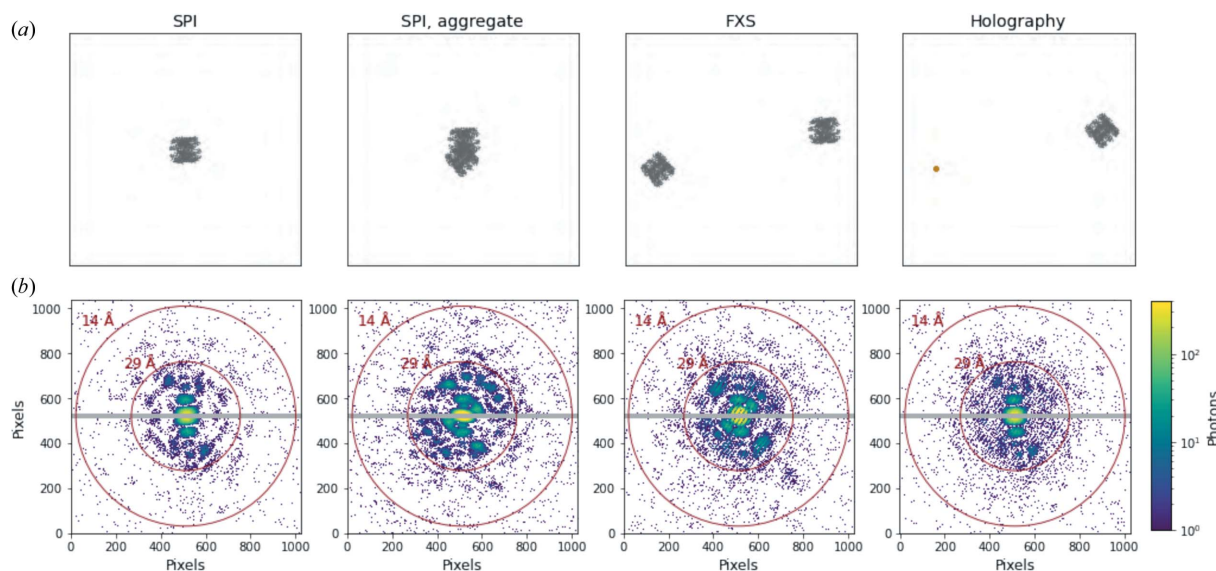
The particle object can be duplicated to form aggregates or clusters. In an experiment, the exact distribution of the number of particles would depend on the particle concentration and parameters governing the injection method (Bielecki *et al.*, 2019; Welker *et al.*, 2021). In *Skopi*, we introduce a parameter *sticking*; when turned on, all the particles in the beam aggregate to form a single cluster. The feature is implemented using a collision model based on the *N*-body simulations of ballistic particle–cluster aggregation (Meakin, 1991; Okuzumi *et al.*, 2009), which is a sequence of successive collisions between a cluster and a single particle. The present model uses a ‘hit-and-stick’ collision model, where the particle is randomly rotated and positioned with respect to the cluster and then translated towards the cluster center until they collide. We neglect any effects from compression or fragmentation of the particle during the collision. Fig. S2 illustrates the particle aggregation procedure.

The beam object contains information about the beam’s dimensions, fluence and wavelength spectrum. The simplest case models a single monochromatic spike with spatially uniform fluence. A more advanced option accounts for self-amplified spontaneous emission (SASE), which produces ultra-bright pulses but broadens the beam’s energy spectrum (Gelsoni *et al.*, 2017). *Skopi* models SASE spectra as trains of uncorrelated spikes using a Gaussian kernel density estimate

to approximate their energy distribution (Fig. S3). Noise can be added by modeling spatial variation of the fluence in the plane of the beam to follow a Gaussian profile and by introducing jitter, in which shot-to-shot changes in total fluence are Gaussian distributed.

The simplest detector is a monolithic square pixel array of user-defined dimensions. The diverse detectors in use at LCLS for CXDI experiments are also supported [Fig. S4(a)]. Recently developed detectors, such as the Jungfrau and ePix10k, are implemented with an auto-ranging feature that avoids saturation to increase their dynamic range (Blaj *et al.*, 2015; van Driel *et al.*, 2020; Redford *et al.*, 2018) [Fig. S4(b)]. LCLS detectors also have access to the calibration constants and fluctuating dark noise from specific experiments [Fig. S5(a)]. Once initialized with a beam, the detector object is populated with information about the pixels’ position in reciprocal space, solid angle subtended and polarization. The detector object also supports modeling various sources of error in addition to fluctuating dark noise. One of these is beam miscentering (Loh *et al.*, 2013), in which the direct beam position is displaced relative to the detector’s center from shot to shot. These offsets are assumed to be Gaussian distributed and are accounted for by mapping the displaced pixels’ coordinates to their corresponding vectors in reciprocal space [Fig. S5(b)]. The user can also supply a custom sloped background to model parasitic scattering or shadows from the experiment’s optics, which are sources of correlated noise that cannot be averaged away by merging more data [Fig. S5(c)].

Once the particle, beam and detector are set up, any of the three principal CXDI experiments can be easily and efficiently simulated. *Skopi* provides convenient interfaces for each of the SPI, FXS and FTH experiment types, from which small- or wide-angle X-ray scattering profiles can be derived through



**Figure 2** An overview of noncrystalline CXDI experiments. *Skopi* supports simulations of SPI, FXS and FTH experiments. (a) Projections of the particle(s) in the plane of the beam and (b) the corresponding diffraction patterns, shown for each experiment type. In the case of SPI, either an individual particle or an aggregate can be simulated. In FXS and FTH experiments, multiple particles are in the beam; for FTH, one of these particles serves as a reference, in this case a small cluster of gold atoms. The biomolecule used in these simulations is a chaperonin (PDB 3iyf; Zhang *et al.*, 2010)). In row (b) the gray region marks the gap between panels of the PnCCD detector. The beam fluence has been artificially inflated to aid visualization.

**Table 1**  
Sources of noise that can be modeled with *Skopi*.

Type	Origin	Implementation
Poisson noise	Counting statistics	Quantization of diffraction intensities such that photon counting follows a Poisson distribution [equation (9)]
Aggregation	Sample delivery	Multi-particle clusters are generated using a ballistic aggregation model, with each particle randomly oriented and positioned with respect to the others
Heterogeneity	Sample dependent	A library of conformational states is generated by sampling along the particle's normal modes
Hydration layer	Sample delivery	A solvent shell that follows the particle's contours is represented using a continuum water model and is discretized to calculate the solvent contribution to diffraction (Liu <i>et al.</i> , 2012)
Beam miscentering	Beam characteristics	Displacements in the direct beam position relative to the detector center are assumed to be independent along each axis of the detector and Gaussian distributed
Fluence jitter	Beam characteristics	Shot-to-shot variation in the beam fluence is drawn from a Gaussian distribution
Fluctuating dark noise	Detector characteristics	Representative pedestal-subtracted dark shots from past LCLS experiments contribute incoherently
Static background	Parasitic scattering	The contribution from a custom background model is added incoherently to the diffraction intensities

radial averaging. In the case of FTH, *Skopi* supports the variation in which the hologram is produced by interference between the target particle and a reference specimen, rather than by modifying the beam shape using structured apertures. For the reference particle, *Skopi* provides a basic gold nanoparticle to facilitate setting up the experiment, while external libraries like the *Atomic Simulation Environment* software package can be employed to generate objects of different shapes (Hjorth Larsen *et al.*, 2017). The SPI, FXS and FTH interfaces act as wrappers for a common experiment class, recognizing that CXDI experiments are at their core very similar, differing only in the number, type and relative position(s) of particle(s) in the beam [Fig. 2(a)]. For each shot, the particle or set of particles is randomly displaced and oriented in the volume intersected by the sample-delivery jet and beam, and the coherent diffraction pattern is computed [Fig. 2(b)]. The incorporation of random displacements accounts for the variable distance between the particles and detector that can result from aerosol injection. To facilitate incremental testing of reconstruction algorithms, either the ideal intensities or quantized photons can be saved, in addition to the number of particles, particle orientations and particle positions at each shot. Beyond Poisson error, the additional sources of noise described above can be defined in configuring the experiment and individually tuned to achieve the desired signal-to-noise ratio (Table 1).

### 3. Benchmarking

*Skopi* achieves rapid and high-throughput simulations of CXDI experiments through graphical processing unit (GPU) acceleration and parallelization using a message passing interface (MPI). For a reciprocal-space grid of fixed dimensions, the diffraction calculation is performed by direct summation over each atom's contribution to the intensity and thus scales linearly with the number of atoms in the particle. Consequently, the full 3D diffraction volume for each particle is computed when an experiment is first configured and thus represents a one-time cost. Diffraction images are then efficiently calculated by slicing through this pre-computed volume at a rate independent of the number of atoms (Fig. S6). The resolution of this volume is determined by the detector distance and the X-ray wavelength, while its default dimensions are sufficiently large that interpolation error during

slicing is negligible. The disadvantage of this approach is the inability to account for the stochastic effects of radiation damage during each interaction between the sample and beam (Nass, 2019); however, radiation damage is anticipated to be minimal for biological samples, particularly in comparison to instrumental background (Neutze *et al.*, 2000; Östlin *et al.*, 2019; Spence, 2017). On a single NVIDIA GeForce RTX 2080 Ti GPU, an SPI data set acquired on a one megapixel PnCCD detector could be simulated at a rate of 0.6 diffraction patterns per second after pre-computing the diffraction volume. This rate was reduced to 0.4 images per second for the SPI aggregate, FXS and FTH experiments due to the need to distribute clustered or dispersed particles in the beam for each shot. Data set generation can be easily parallelized across an arbitrary number of GPU nodes to increase throughput.

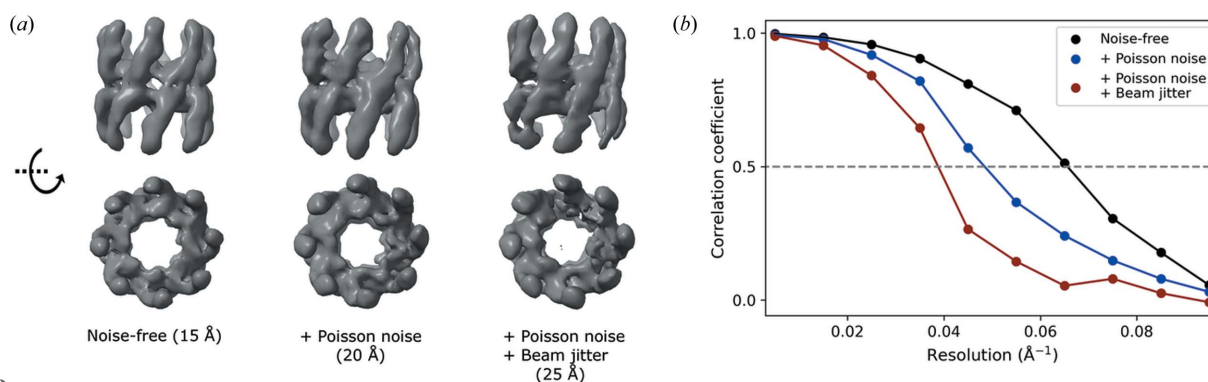
### 4. Validation

Comparison of SPI diffraction images simulated using *Skopi* and another CXDI software package, *Condor*, show that the two yield identical results within error from Poisson noise (Fig. S7). We then validated *Skopi* by recovering the protein structure from simulated diffraction images of a chaperonin (PDB 3iyf). SPI data sets consisting of 5000 images each were generated in the absence or presence of Poisson noise on a PnCCD detector positioned for a resolution limit of 14 Å at the edge of the detector. On average, the noisy data sets featured 0.0076 photons per pixel at the edge. The theoretical number of snapshots required for reconstructing a 70 Å particle at 14 Å resolution at this signal level is 3380 (Poudyal *et al.*, 2020; Ekeberg *et al.*, 2015). Reconstruction was performed using a Cartesian implementation of the multi-tiered iterative phasing (*MTIP*) algorithm (Donatelli *et al.*, 2017; Dujardin *et al.*, 2020). The resolution of each reconstruction was assessed by computing the Fourier shell correlation (FSC) between the recovered density and the ground-truth map,

$$\text{FSC}(k) = \frac{\sum_{|k_i| \in k} F_1(\mathbf{k}) F_2^*(\mathbf{k})}{\left[ \sum_{|k_i| \in k} |F_1(\mathbf{k})|^2 \sum_{|k_i| \in k} |F_2^*(\mathbf{k})|^2 \right]^{1/2}}, \quad (1)$$

where  $F_1$  and  $F_2$  are the Fourier components of the density maps being compared and \* denotes the complex conjugate. When compared with a noise-free reference map, the spatial





**Figure 3** Reconstruction from simulated SPI data sets. SPI data sets from a chaperonin were simulated in the absence or presence of noise. A multi-tiered iterative phasing algorithm was used to recover the protein structure from 5000 images of the indicated data set. (a) Isosurfaces of the density map reveal the loss of eightfold symmetry with increasing noise. The resolution of each reconstruction is noted in parentheses. (b) The resolution was measured as the spatial frequency at which the FSC between the reconstructed and reference maps dropped to 0.5.

frequency  $k$  at which the FSC falls to 0.5 is considered to estimate the resolution of the reconstruction (Liao & Frank, 2010). The protein structure was accurately recovered from noise-free diffraction images simulated on a PnCCD detector (Fig. 3), with a data set size comparable to what has previously been used to evaluate the *MTIP* algorithm (Chang *et al.*, 2021). As expected, the addition of Poisson noise degraded the quality of the recovered structure (Donatelli *et al.*, 2017), reducing the resolution from 15 to 20 Å. Introducing beam jitter further diminished reconstruction quality, as evident in the loss of the protein’s eightfold symmetry (Fig. 3). *Skopi* thus provides a useful tool to assess the tolerance of reconstruction algorithms to different types of noise.

### 5. Conclusions

We have presented *Skopi* as a convenient tool for the efficient simulation of CXDI experiments of noncrystalline biological samples. The modular design of this package makes it easy to represent each component of an XFEL experiment with a range of complexity, including modeling the most up-to-date features of current LCLS detectors (van Driel *et al.*, 2020). Another focus of *Skopi* is to produce simulated data with realistic noise. The different sources of error that typify CXDI experiments – including fluence jitter, beam miscentering and fluctuating dark noise – can be readily incorporated into the simulation without additional scripting.

The paucity of experimental CXDI data sets of noncrystalline biological samples has hindered algorithmic growth in the field. We anticipate that *Skopi* could help fill this gap through its capability of rapidly providing diverse simulated data sets with realistic noise. This may prove particularly valuable for the development of machine learning algorithms to perform classification, since labeled data can easily be generated in bulk. Classification algorithms would benefit data pre-processing to separate single-particle hits from aggregate shots and better assess hit rates, and could also be used during reconstruction to sort diffraction images based on the particle’s conformation (Ignatenko *et al.*, 2021), as done in cryo-electron microscopy (Punjani & Fleet, 2021; Zhong *et al.*, 2021; Chen & Ludtke, 2021).

In addition to aiding the development of reconstruction algorithms, we envisage *Skopi* as a potential tool for experimental design. By supporting both the advanced detectors currently in use at LCLS and the inclusion of fluctuating dark noise from past experiments, *Skopi* enables the simulation of highly realistic diffraction patterns that might be obtained at this facility. These simulated data sets, in turn, provide an estimate of the available signal under different experimental conditions, potentially allowing settings to be optimized and offering an estimate of the volume of data required for reconstruction in advance of XFEL experiments. The use of the *simS2E* software package to anticipate data collection needs at the European XFEL offers a precedent for optimizing CXDI experiments using simulation (Yoon *et al.*, 2016). Looking ahead, we are working towards making *Skopi* available within the larger *SIMEX* framework to facilitate large-scale start-to-end simulations of noncrystalline CXDI experiments (Fortmann-Grote *et al.*, 2017). In these frameworks, simulations are chained together starting from source parameters and progressing to propagation of the coherent X-rays through beamline optics, interaction of the photons with matter, and their subsequent detection and analysis. Such large-scale start-to-end simulations will be valuable for the development of noncrystalline XFEL experiments.

### 6. Code availability

*Skopi* is available through the Python Package Index (PyPi) at <https://pypi.org/project/skopi/>. The version referenced in this work is 0.5.1. More information about installation can be found at <https://github.com/chuckie82/skopi/blob/main/docs/installation.md> and additional example usage scripts at <https://github.com/chuckie82/skopi/tree/main/examples/scripts>.

## APPENDIX A

### Theory, implementation and usage

#### A1. Elements of theory

In this section, we summarize the main elements of coherent X-ray optics taken from Paganin (2006) that allow us

to derive the diffraction model implemented in *Skopi*. Propagation of electromagnetic waves  $\Psi(\mathbf{r}, t)$  in non-magnetic matter follows the wave equation, where the local velocity is modulated by the local electric permittivity  $\epsilon$  while the magnetic permeability  $\mu_0$  is assumed to be that of free space everywhere,

$$\left[ \epsilon(\mathbf{r})\mu_0 \frac{\partial^2}{\partial t^2} - \nabla^2 \right] \Psi(\mathbf{r}, t) = 0. \quad (2)$$

Separation of temporal and spatial variables is assumed through spectral decomposition,

$$\Psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{1/2}} \int_0^\infty d\omega \phi_\omega(\mathbf{r}) \exp(-i\omega t), \quad (3)$$

where the frequency  $\omega$  of each monochromatic wave is related to its wavevector modulus  $k$  through the speed of light in free space,  $\omega = kc$ . The effect of local perturbations can be factored in the refractive index  $n_\omega(\mathbf{r}) = c[\epsilon(\mathbf{r})\mu_0]^{1/2}$ , which measures the relative speed of light in the medium compared with free space. Inserting the spectral decomposition of the electromagnetic wave into the wave equation yields an inhomogeneous Helmholtz equation for each monochromatic wave,

$$[\nabla^2 + k^2 n_\omega^2(\mathbf{r})] \phi_\omega(\mathbf{r}) = 0. \quad (4)$$

In free space, equation (4) simplifies to the homogeneous Helmholtz equation, which can be solved and yields the following solution to the wave equation (2):  $\Psi(\mathbf{r}, t) = \exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$ . The spatial wavelength  $\lambda$  of this wave is such that  $\cos(k\lambda) = 1$ , so the wavevector modulus is related to it through the relation  $k = 2\pi/\lambda$ .

More generally, it can be shown that equation (2) is equivalent to its integral equation formulation, which involves the incident wave  $\phi^{(0)}$  and a convolution with Green's function  $G_\omega = \exp(ikr)/r$ ,

$$\phi_\omega(\mathbf{r}) = \phi_\omega^{(0)}(\mathbf{r}) - \frac{k^2}{4\pi} G_\omega * [(1 - n_\omega^2) \phi_\omega](\mathbf{r}). \quad (5)$$

In the Born approximation, the wavefunction on the right-hand side of equation (5) is assumed to be identical to the incident wavefunction, thus trivializing the integral equation formulation. Further assuming the case of Fraunhofer (far-field) diffraction and of an incident plane wave  $\phi_\omega^{(0)}(\mathbf{r}) = \exp(i\mathbf{k}_0 \cdot \mathbf{r})$ , we can introduce the scattering amplitude  $f(\Delta\mathbf{k})$ , where  $\Delta\mathbf{k} = k(\mathbf{r}/r) - \mathbf{k}_0$ ,

$$\begin{aligned} \phi_\omega(\mathbf{r}) &= \exp(i\mathbf{k}_0 \cdot \mathbf{r}) + \frac{\exp(ikr)}{r} f(\Delta\mathbf{k}), \\ f(\Delta\mathbf{k}) &= -\frac{k^2}{4\pi} \int d\mathbf{r}' [1 - n_\omega^2(\mathbf{r}')] \exp(-i\Delta\mathbf{k} \cdot \mathbf{r}'). \end{aligned} \quad (6)$$

Assuming a specific relationship between the refractive index and the electron density of the object  $\rho$ , the scattering amplitude is scaled by the Thomson scattering length  $r_0$ ,

$$f(\Delta\mathbf{k}) = -r_0 \int d\mathbf{r}' \rho(\mathbf{r}') \exp(-i\Delta\mathbf{k} \cdot \mathbf{r}'). \quad (7)$$

Finally, denoting by  $2\theta$  the angle between the incident and resulting wavevectors, we introduce the notation  $\mathbf{s} = \Delta\mathbf{k}$ , where  $s = (4\pi/\lambda) \sin \theta$ . In the superposition approximation, the molecular electron density is approximated as a sum of individual atomic contributions. In diffraction space, the atomic form factors  $f_i$  are tabulated as a function of  $\sin \theta/\lambda$ ,

$$f(\mathbf{s}) = -r_0 \sum_i f_i \left( \frac{s}{4\pi} \right) \exp(-i\mathbf{s} \cdot \mathbf{r}_i). \quad (8)$$

The diffracted intensity  $I$  is scaled by the number of incident photons per beam focus area, or fluence  $F$ , which can be modeled as a constant or varying shot to shot. For a given detector pixel, the signal will be scaled by the solid angle  $\Omega$  from the interaction point. The non-corrected diffracted intensity at pixel  $i$  thus reads  $I_i = F\Omega_i |f(\mathbf{s}_i)|^2$ . However, polarization  $\mathbf{u}_p$  of the incoming beam will result in non-isotropic Thomson scattering which further weights the intensity by  $|\mathbf{u}_i \times \mathbf{u}_p|^2$ , where  $\mathbf{u}_i$  is the unit vector pointing to pixel  $i$ . Optionally, to account for the multiple-spike nature of the X-ray beam spectrum due to SASE, the diffracted intensity is incoherently summed over a discrete set of varying frequencies weighted by the frequency content  $w_\omega$ . It is also possible to add different types of corrupting terms incoherently to the diffracted intensity, such as static noise or a sloped background.

Finally, encompassing the various terms listed above into functions  $\alpha$  and  $\beta$ , the diffracted intensity is quantized by drawing from the corresponding Poisson distribution  $\mathcal{P}$  to simulate the photon counts  $C$  measured on each pixel  $i$  of the detector,

$$C_i = \mathcal{P} \left[ \alpha(\mathbf{s}_i) \sum_\omega^{\text{SASE}} w_\omega |f_\omega(\mathbf{s}_i)|^2 + \beta(\mathbf{s}_i) \right]. \quad (9)$$

## A2. Implementation

In practice, equation (8) is pre-computed on a discrete 3D cubic grid whose reciprocal extent is determined by the detector and beam parameters. Once the resulting diffraction volume is available in memory, it can be readily sliced at an arbitrary orientation and curvature in two steps. First, the detector pixels are mapped to their corresponding positions in the reciprocal volume. Second, the complex structure factor at each pixel is estimated through trilinear interpolation. The resulting diffraction intensities are then adjusted to account for beam polarization, solid-angle distortion and diverse sources of noise.

## A3. Usage

Python Script 1 illustrates how the modularity of *Skopi* is implemented in practice. Classes for the beam, particle and detector are instantiated and given to the experiment class, which in turn generates diffraction patterns. In the first example, files from a prior LCLS experiment are used to define the diffraction geometry. In general, a geometry file is not required.

```

import skopi

# define path to input files
input_dir = './input/'
beamfile = input_dir+'beam/amo86615.beam'
geom = input_dir
      +'lcls/amo86615/'
      +'PNCCD::CalibV1/Camp.0:pnCCD.1/'
      +'geometry/0-end.data'
pdbfile = input_dir+'pdb/3iyf.pdb'

# create beam object
beam = skopi.Beam(beamfile)
beam.set_photons_per_pulse(1e14)

# initialize a PnCCD detector
from skopi import Pnccddetector as Detector
detector = Detector(
    geom=geom,
    beam=beam)
detector.distance = 0.1 # (m)

# set up particle from a PDB file
particle = skopi.Particle()
particle.read_pdb(pdbfile, ff='WK')

# generate an SPI-style Experiment object
experiment = skopi.SPIExperiment(
    det=detector,
    beam=beam,
    particle=particle)

# simulate one diffraction shot
img = experiment.generate_image()

# visualize shot
viz = skopi.Visualizer(
    experiment,
    diffraction_rings="auto",
    log_scale=True)
viz.imshow(img)

```

Script 2 illustrates how to generate FXS images from five chaperone particles on a simple square detector.

```

# define simple square detector
from skopi.detector \
    import SimpleSquareDetector as Detector
num_pix_per_side = 1000
physical_size_side = 0.1 # (m)
detector_distance = 0.1 # (m)

detector = Detector(
    num_pix_per_side,
    physical_size_side,
    detector_distance,
    beam=beam)

# generate an FXS-style Experiment object
experiment = skopi.FXSEExperiment(
    det=detector,
    beam=beam,
    particles=[particle],
    jet_radius=1e-4,
    n_part_per_shot=5)

# simulate one diffraction shot
img = experiment.generate_image()

# visualize shot
fig = plt.figure(figsize=(6,6),dpi=300)
viz = skopi.Visualizer(
    experiment,
    diffraction_rings="auto",
    log_scale=True)
viz.imshow(img)

```

The outputs from the two code examples above for SPI and FXS are shown in Fig. S8.

## Acknowledgements

We thank Haoyuan Li and Juncheng E for their contributions to the code.

## Funding information

This research was supported by the Exascale Computing Project (grant No. 17-SC-20-SC), a collaborative effort of the US Department of Energy Office of Science and the National Nuclear Security Administration. Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences (contract No. DE-AC02-76SF00515).

## References

- Assalauova, D., Kim, Y. Y., Bobkov, S., Khubbutdinov, R., Rose, M., Alvarez, R., Andreasson, J., Balaur, E., Contreras, A., DeMirci, H., Gelisio, L., Hajdu, J., Hunter, M. S., Kurta, R. P., Li, H., McFadden, M., Nazari, R., Schwander, P., Teslyuk, A., Walter, P., Xavier, P. L., Yoon, C. H., Zaare, S., Ilyin, V. A., Kirian, R. A., Hogue, B. G., Aquila, A. & Vartanyants, I. A. (2020). *IUCr*, **7**, 1102–1113.
- Ayyer, K., Xavier, P. L., Bielecki, J., Shen, Z., Daurer, B. J., Samanta, A. K., Awel, S., Bean, R., Barty, A., Bergemann, M., Ekeberg, T., Estillore, A. D., Fangohr, H., Giewekemeyer, K., Hunter, M. S., Karnevskiy, M., Kirian, R. A., Kirkwood, H., Kim, Y., Koliyadu, J., Lange, H., Letrun, R., Lübke, J., Michelat, T., Morgan, A. J., Roth, N., Sato, T., Sikorski, M., Schulz, F., Spence, J. H., Vagovic, P., Wollweber, T., Worbs, L., Yefanov, O., Zhuang, Y., Maia, F. R. N. C., Horke, D. A., Küpper, J., Loh, N. D., Mancuso, A. P. & Chapman, H. N. (2020). *Optica*, **8**, 15–23.
- Bakan, A., Meireles, L. M. & Bahar, I. (2011). *Bioinformatics*, **27**, 1575–1577.
- Bielecki, J., Hantke, M. F., Daurer, B. J., Reddy, H. K. N., Hasse, D., Larsson, D. S. D., Gunn, L. H., Svenda, M., Munke, A., Sellberg, J. A., Flueckiger, L., Pietrini, A., Nettelblad, C., Lundholm, I., Carlsson, G., Okamoto, K., Timneanu, N., Westphal, D., Kulyk, O., Higashiura, A., van der Schot, G., Loh, N. D., Wysong, T. E., Bostedt, C., Gorkhover, T., Iwan, B., Seibert, M. M., Osipov, T., Walter, P., Hart, P., Bucher, M., Ulmer, A., Ray, D., Carini, G., Ferguson, K. R., Andersson, I., Andreasson, J., Hajdu, J. & Maia, F. R. N. C. (2019). *Sci. Adv.* **5**, eaav8801.
- Bielecki, J., Maia, F. R. N. C. & Mancuso, A. P. (2020). *Struct. Dyn.* **7**, 040901.
- Blaj, G., Caragiulo, P., Carini, G., Carron, S., Dragone, A., Freytag, D., Haller, G., Hart, P., Hasi, J., Herbst, R., Herrmann, S., Kenney, C., Markovic, B., Nishimura, K., Osier, S., Pines, J., Reese, B., Segal, J., Tomada, A. & Weaver, M. (2015). *J. Synchrotron Rad.* **22**, 577–583.
- Bogan, M. J., Benner, W. H., Boutet, S., Rohner, U., Frank, M., Barty, A., Seibert, M. M., Maia, F., Marchesini, S., Bajt, S., Woods, B., Riot, V., Hau-Riege, S. P., Svenda, M., Marklund, E., Spiller, E., Hajdu, J. & Chapman, H. N. (2008). *Nano Lett.* **8**, 310–316.
- Boutet, S., Lomb, L., Williams, G. J., Barends, T. R., Aquila, A., Doak, R. B., Weierstall, U., DePonte, D. P., Steinbrener, J., Shoeman, R. L., Messerschmidt, M., Barty, A., White, T. A., Kassemeyer, S., Kirian, R. A., Seibert, M. M., Montanez, P. A., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S. M., Philipp, H. T., Tate, M. W., Hromalik, M., Koerner, L. J., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M. J., Caleman, C., Fromme, R., Hampton, C. Y., Hunter, M. S., Johansson, L. C., Katona, G., Kupitz, C., Liang, M., Martin, A. V., Nass, K., Redecke, L., Stellato,



- F., Timneanu, N., Wang, D., Zatsepin, N. A., Schafer, D., Defever, J., Neutze, R., Fromme, P., Spence, J. C., Chapman, H. N. & Schlichting, I. (2012). *Science*, **337**, 362–364.
- Brändén, G., Hammarin, G., Harimoorthy, R., Johansson, A., Arnlund, D., Malmerberg, E., Barty, A., Tångefjord, S., Berntsen, P., DePonte, D. P., Seuring, C., White, T. A., Stellato, F., Bean, R., Beyerlein, K. R., Chavas, L. M. G., Fleckenstein, H., Gati, C., Ghoshdastider, U., Gumprecht, L., Oberthür, D., Popp, D., Seibert, M., Tilp, T., Messerschmidt, M., Williams, G. J., Loh, N. D., Chapman, H. N., Zwart, P., Liang, M., Boutet, S., Robinson, R. C. & Neutze, R. (2019). *Nat. Commun.* **10**, 2589.
- Chang, H.-Y., Slaughter, E., Mirchandaney, S., Donatelli, J. & Yoon, C. H. (2021). *arXiv:2109.05339*.
- Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., DePonte, D. P., Weierstall, U., Doak, R. B., Maia, F. R., Martin, A. V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmeß, M., Coleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. (2011). *Nature*, **470**, 73–77.
- Chen, M. & Ludtke, S. J. (2021). *Nat. Methods*, **18**, 930–936.
- Cruz-Chú, E. R., Hosseinizadeh, A., Mashayekhi, G., Fung, R., Ourmazd, A. & Schwander, P. (2021). *Struct. Dyn.* **8**, 014701.
- Daurer, B. J., Okamoto, K., Bielecki, J., Maia, F. R. N. C., Mühlig, K., Seibert, M. M., Hantke, M. F., Nettelblad, C., Benner, W. H., Svenda, M., Timneanu, N., Ekeberg, T., Loh, N. D., Pietrini, A., Zani, A., Rath, A. D., Westphal, D., Kirian, R. A., Awel, S., Wiedorn, M. O., van der Schot, G., Carlsson, G. H., Hasse, D., Sellberg, J. A., Barty, A., Andreasson, J., Boutet, S., Williams, G., Koglin, J., Andersson, I., Hajdu, J. & Larsson, D. S. D. (2017). *IUCrJ*, **4**, 251–262.
- Donatelli, J. J., Sethian, J. A. & Zwart, P. H. (2017). *Proc. Natl Acad. Sci. USA*, **114**, 7222–7227.
- Doniach, S. (2018). *X-ray Free Electron Lasers*, edited by S. Boutet, P. Fromme & M. S. Hunter, pp. 427–439. Cham: Springer.
- Driel, T. B. van, Nelson, S., Armenta, R., Blaj, G., Boo, S., Boutet, S., Doering, D., Dragone, A., Hart, P., Haller, G., Kenney, C., Kwaitowski, M., Manger, L., McKelvey, M., Nakahara, K., Oriunno, M., Sato, T. & Weaver, M. (2020). *J. Synchrotron Rad.* **27**, 608–615.
- Dujardin, A., Slaughter, E., Donatelli, J., Zwart, P., Perazzo, A. & Yoon, C. H. (2020). *Proceedings of the 19th Python in Science Conference, July 6–July 12, Austin, Texas*, edited by M. Agarwal, C. Calloway, D. Niederhut & D. Shupe, pp. 18–23.
- Eisebitt, S., Lüning, J., Schlotter, W., Lörger, M., Hellwig, O., Eberhardt, W. & Stöhr, J. (2004). *Nature*, **432**, 885–888.
- Ekeberg, T., Svenda, M., Abergel, C., Maia, F. R., Seltzer, V., Claverie, J. M., Hantke, M., Jönsson, O., Nettelblad, C., van der Schot, G., Liang, M., DePonte, D. P., Barty, A., Seibert, M. M., Iwan, B., Andersson, I., Loh, N. D., Martin, A. V., Chapman, H., Bostedt, C., Bozek, J. D., Ferguson, K. R., Krzywinski, J., Epp, S. W., Rolles, D., Rudenko, A., Hartmann, R., Kimmel, N. & Hajdu, J. (2015). *Phys. Rev. Lett.* **114**, 098102.
- Emma, P., Akre, R., Arthur, J., Bionta, R., Bostedt, C., Bozek, J., Brachmann, A., Bucksbaum, P., Coffee, R., Decker, F. J., Ding, Y., Dowell, D., Edstrom, S., Fisher, A., Frisch, J., Gilevich, S., Hastings, J., Hays, G., Hering, P., Huang, Z., Iverson, R., Loos, H., Messerschmidt, M., Miahnahri, A., Moeller, S., Nuhn, H. D., Pile, G., Ratner, D., Rzepiela, J., Schultz, D., Smith, T., Stefan, P., Tompkins, H., Turner, J., Welch, J., White, W., Wu, J., Yocky, G. & Galayda, J. (2010). *Nat. Photon.* **4**, 641–647.
- Fortmann-Grote, C., Buzmakov, A., Jurek, Z., Loh, N.-T. D., Samoylova, L., Santra, R., Schneidmiller, E. A., Tschentscher, T., Yakubov, S., Yoon, C. H., Yurkov, M. V., Ziaja-Motyka, B. & Mancuso, A. P. (2017). *IUCrJ*, **4**, 560–568.
- Geloni, G., Huang, Z. & Pellegrini, C. (2017). *X-ray Free Electron Lasers: Applications in Materials, Chemistry and Biology*, edited by U. Bergmann, V. Yachandra & J. Yano, pp. 1–44. Cambridge: Royal Society of Chemistry.
- Gorkhover, T., Ulmer, A., Ferguson, K., Bucher, M., Maia, F. R. N. C., Bielecki, J., Ekeberg, T., Hantke, M. F., Daurer, B. J., Nettelblad, C., Andreasson, J., Barty, A., Bruza, P., Carron, S., Hasse, D., Krzywinski, J., Larsson, D. S. D., Morgan, A., Mühlig, K., Müller, M., Okamoto, K., Pietrini, A., Rupp, D., Sauppe, M., van der Schot, G., Seibert, M., Sellberg, J. A., Svenda, M., Swiggers, M., Timneanu, N., Westphal, D., Williams, G., Zani, A., Chapman, H. N., Faigel, G., Möller, T., Hajdu, J. & Bostedt, C. (2018). *Nat. Photon.* **12**, 150–153.
- Hantke, M. F., Ekeberg, T. & Maia, F. R. N. C. (2016). *J. Appl. Cryst.* **49**, 1356–1362.
- Hantke, M. F., Hasse, D., Maia, F. R. N. C., Ekeberg, T., John, K., Svenda, M., Loh, N. D., Martin, A. V., Timneanu, N., Larsson, D. S. D., van der Schot, G., Carlsson, G. H., Ingelman, M., Andreasson, J., Westphal, D., Liang, M., Stellato, F., DePonte, D. P., Hartmann, R., Kimmel, N., Kirian, R. A., Seibert, M. M., Mühlig, K., Schorb, S., Ferguson, K., Bostedt, C., Carron, S., Bozek, J. D., Rolles, D., Rudenko, A., Epp, S., Chapman, H. N., Barty, A., Hajdu, J. & Andersson, I. (2014). *Nat. Photon.* **8**, 943–949.
- Hau-Riege, S. P., London, R. A., Chapman, H. N., Szoke, A. & Timneanu, N. (2007). *Phys. Rev. Lett.* **98**, 198302.
- Hjorth Larsen, A., Jørgen Mortensen, J., Blomqvist, J., Castelli, I. E., Christensen, R., Dulak, M., Friis, J., Groves, M. N., Hammer, B., Hargus, C., Hermes, E. D., Jennings, P. C., Bjerre Jensen, P., Kermode, J., Kitchin, J. R., Leonhard Kolsbjerg, E., Kubal, J., Kaasbjerg, K., Lysgaard, S., Bergmann Maronsson, J., Maxson, T., Olsen, T., Pastewka, L., Peterson, A., Rostgaard, C., Schiøtz, J., Schütt, O., Strange, M., Thygesen, K. S., Vegge, T., Vilhelmsen, L., Walter, M., Zeng, Z. & Jacobsen, K. W. (2017). *J. Phys. Condens. Matter*, **29**, 273002.
- Ignatenko, A., Assalauova, D., Bobkov, S. A., Gelisio, L., Teslyuk, A. B., Ilyin, V. A. & Vartanyants, I. A. (2021). *Mach. Learn. Sci. Technol.* **2**, 025014.
- Jamison, S. (2010). *Nat. Photon.* **4**, 589–591.
- Kurta, R. P., Donatelli, J. J., Yoon, C. H., Berntsen, P., Bielecki, J., Daurer, B. J., DeMirci, H., Fromme, P., Hantke, M. F., Maia, F. R. N. C., Munke, A., Nettelblad, C., Pande, K., Reddy, H. K. N., Sellberg, J. A., Sierra, R. G., Svenda, M., van der Schot, G., Vartanyants, I. A., Williams, G. J., Xavier, P. L., Aquila, A., Zwart, P. H. & Mancuso, A. P. (2017). *Phys. Rev. Lett.* **119**, 158102.
- Liao, H. Y. & Frank, J. (2010). *Structure*, **18**, 768–775.
- Liu, H., Morris, R. J., Hexemer, A., Grandison, S. & Zwart, P. H. (2012). *Acta Cryst.* **A68**, 278–285.
- Loh, N. D., Starodub, D., Lomb, L., Hampton, C. Y., Martin, A. V., Sierra, R. G., Barty, A., Aquila, A., Schulz, J., Steinbrener, J., Shoeman, R. L., Kassemeyer, S., Bostedt, C., Bozek, J., Epp, S. W., Erk, B., Hartmann, R., Rolles, D., Rudenko, A., Rudek, B., Foucar, L., Kimmel, N., Weidenspointner, G., Hauser, G., Holl, P., Pedersoli, E., Liang, M., Hunter, M. S., Gumprecht, L., Coppola, N., Wunderer, C., Graafsma, H., Maia, F. R., Ekeberg, T., Hantke, M., Fleckenstein, H., Hirsemann, H., Nass, K., White, T. A., Tobias, H. J., Farquar, G. R., Benner, W. H., Hau-Riege, S., Reich, C., Hartmann, A., Soltau, H., Marchesini, S., Bajt, S., Barthelmeß, M., Strueder, L., Ullrich, J., Bucksbaum, P., Frank, M., Schlichting, I., Chapman, H. N. & Bogan, M. J. (2013). *Opt. Express*, **21**, 12385.



- Maia, F. R., Ekeberg, T., Timneanu, N., van der Spoel, D. & Hajdu, J. (2009). *Phys. Rev. E*, **80**, 031905.
- Mandl, T., Östlin, C., Dawod, I. E., Brodmerkel, M. N., Marklund, E. G., Martin, A. V., Timneanu, N. & Coleman, C. (2020). *J. Phys. Chem. Lett.* **11**, 6077–6083.
- Marchesini, S., Boutet, S., Sakdinawat, A. E., Bogan, M. J., Bajt, S., Barty, A., Chapman, H. N., Frank, M., Hau-Riege, S. P., Szöke, A., Cui, C., Shapiro, D. A., Howells, M. R., Spence, J. C. H., Shaevitz, J. W., Lee, J. Y., Hajdu, J. & Seibert, M. M. (2008). *Nat. Photon.* **2**, 560–563.
- Meakin, P. (1991). *Rev. Geophys.* **29**, 317–354.
- Mendez, D., Lane, T. J., Sung, J., Sellberg, J., Levard, C., Watkins, H., Cohen, A. E., Soltis, M., Sutton, S., Spudich, J., Pande, V., Ratner, D. & Doniach, S. (2014). *Phil. Trans. R. Soc. B*, **369**, 20130315.
- Mendez, D., Watkins, H., Qiao, S., Raines, K. S., Lane, T. J., Schenk, G., Nelson, G., Subramanian, G., Tono, K., Joti, Y., Yabashi, M., Ratner, D. & Doniach, S. (2016). *IUCrJ*, **3**, 420–429.
- Nass, K. (2019). *Acta Cryst.* **D75**, 211–218.
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature*, **406**, 752–757.
- Okuzumi, S., Tanaka, H. & Sakagami, M. (2009). *Astrophys. J.* **707**, 1247–1263.
- Östlin, C., Timneanu, N., Coleman, C. & Martin, A. V. (2019). *Struct. Dyn.* **6**, 044103.
- Paganin, D. M. (2006). *Coherent X-ray Optics*. Oxford University Press.
- Pande, K., Donatelli, J. J., Malmerberg, E., Foucar, L., Bostedt, C., Schlichting, I. & Zwart, P. H. (2018). *Proc. Natl Acad. Sci. USA*, **115**, 11772–11777.
- Poudyal, I., Schmidt, M. & Schwander, P. (2020). *Struct. Dyn.* **7**, 024102.
- Punjani, A. & Fleet, D. J. (2021). *J. Struct. Biol.* **213**, 107702.
- Reddy, H. K., Yoon, C. H., Aquila, A., Awel, S., Ayer, K., Barty, A., Berntsen, P., Bielecki, J., Bobkov, S., Bucher, M., Carini, G. A., Carron, S., Chapman, H., Daurer, B., DeMirici, H., Ekeberg, T., Fromme, P., Hajdu, J., Hanke, M. F., Hart, P., Hogue, B. G., Hosseinizadeh, A., Kim, Y., Kirian, R. A., Kurta, R. P., Larsson, D. S., Duane Loh, N., Maia, F. R., Mancuso, A. P., Mühlhig, K., Munke, A., Nam, D., Nettelblad, C., Ourmazd, A., Rose, M., Schwander, P., Seibert, M., Sellberg, J. A., Song, C., Spence, J. C., Svenda, M., Van der Schot, G., Vartanyants, I. A., Williams, G. J. & Xavier, P. L. (2017). *Sci. Data*, **4**, 170079.
- Redford, S., Andrä, M., Barten, R., Bergamaschi, A., Brückner, M., Dinapoli, R., Fröjd, E., Greiffenberg, D., Lopez-Cuenca, C., Mezza, D., Mozzanica, A., Ramilli, M., Ruat, M., Ruder, C., Schmitt, B., Shi, X., Thattil, D., Tinti, G., Vetter, S. & Zhang, J. (2018). *J. Instrum.* **13**, C01027.
- Rose, M., Bobkov, S., Ayer, K., Kurta, R. P., Dzhigaev, D., Kim, Y. Y., Morgan, A. J., Yoon, C. H., Westphal, D., Bielecki, J., Sellberg, J. A., Williams, G., Maia, F. R. N. C., Yefanov, O. M., Ilyin, V., Mancuso, A. P., Chapman, H. N., Hogue, B. G., Aquila, A., Barty, A. & Vartanyants, I. A. (2018). *IUCrJ*, **5**, 727–736.
- Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
- Schot, G. van der, Svenda, M., Maia, F. R., Hantke, M., DePonte, D. P., Seibert, M. M., Aquila, A., Schulz, J., Kirian, R., Liang, M., Stellato, F., Iwan, B., Andreasson, J., Timneanu, N., Westphal, D., Almeida, F. N., Odić, D., Hasse, D., Carlsson, G. H., Larsson, D. S., Barty, A., Martin, A. V., Schorb, S., Bostedt, C., Bozek, J. D., Rolles, D., Rudenko, A., Epp, S., Foucar, L., Rudek, B., Hartmann, R., Kimmel, N., Holl, P., Englert, L., Duane Loh, N. T., Chapman, H. N., Andersson, I., Hajdu, J. & Ekeberg, T. (2015). *Nat. Commun.* **6**, 5704.
- Seibert, M. M., Ekeberg, T., Maia, F. R., Svenda, M., Andreasson, J., Jönsson, O., Odić, D., Iwan, B., Rocker, A., Westphal, D., Hantke, M., DePonte, D. P., Barty, A., Schulz, J., Gumprecht, L., Coppola, N., Aquila, A., Liang, M., White, T. A., Martin, A., Coleman, C., Stern, S., Abergel, C., Seltzer, V., Claverie, J. M., Bostedt, C., Bozek, J. D., Boutet, S., Miahnahri, A. A., Messerschmidt, M., Krzywinski, J., Williams, G., Hodgson, K. O., Bogan, M. J., Hampton, C. Y., Sierra, R. G., Starodub, D., Andersson, I., Bajt, S., Barthelmeß, M., Spence, J. C., Fromme, P., Weierstall, U., Kirian, R., Hunter, M., Doak, R. B., Marchesini, S., Hau-Riege, S. P., Frank, M., Shoeman, R. L., Lomb, L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Schmidt, C., Foucar, L., Kimmel, N., Holl, P., Rudek, B., Erk, B., Hömke, A., Reich, C., Pietschner, D., Weidenspointner, G., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Schlichting, I., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B., Chapman, H. N. & Hajdu, J. (2011). *Nature*, **470**, 78–81.
- Shi, Y., Yin, K., Tai, X., DeMirici, H., Hosseinizadeh, A., Hogue, B. G., Li, H., Ourmazd, A., Schwander, P., Vartanyants, I. A., Yoon, C. H., Aquila, A. & Liu, H. (2019). *IUCrJ*, **6**, 331–340.
- Shintake, T. (2008). *Phys. Rev. E*, **78**, 041906.
- Spence, J. C. H. (2017). *IUCrJ*, **4**, 322–339.
- Spence, J. C. H. (2018). *IUCrJ*, **5**, 236–237.
- Starodub, D., Aquila, A., Bajt, S., Barthelmeß, M., Barty, A., Bostedt, C., Bozek, J. D., Coppola, N., Doak, R. B., Epp, S. W., Erk, B., Foucar, L., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hartmann, R., Holl, P., Kassemeyer, S., Kimmel, N., Laksmono, H., Liang, M., Loh, N. D., Lomb, L., Martin, A. V., Nass, K., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schulz, J., Shoeman, R. L., Sierra, R. G., Soltau, H., Steinbrener, J., Stellato, F., Stern, S., Weidenspointner, G., Frank, M., Ullrich, J., Strüder, L., Schlichting, I., Chapman, H. N., Spence, J. C. & Bogan, M. J. (2012). *Nat. Commun.* **3**, 1276.
- Stauch, B. & Cherezov, V. (2018). *Annu. Rev. Biophys.* **47**, 377–397.
- Tama, F. & Sanejouand, Y. H. (2001). *Protein Eng.* **14**, 1–6.
- Welker, S., Amin, M. & Küpper, J. (2021). *arXiv:2102.05955*.
- Yoon, C. H., Schwander, P., Abergel, C., Andersson, I., Andreasson, J., Aquila, A., Bajt, S., Barthelmeß, M., Barty, A., Bogan, M. J., Bostedt, C., Bozek, J., Chapman, H. N., Claverie, J. M., Coppola, N., DePonte, D. P., Ekeberg, T., Epp, S. W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C. Y., Hartmann, A., Hartmann, E., Hartmann, R., Hauser, G., Hirsemann, H., Holl, P., Kassemeyer, S., Kimmel, N., Kiskinova, M., Liang, M., Loh, N. T., Lomb, L., Maia, F. R., Martin, A. V., Nass, K., Pedersoli, E., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schulz, J., Seibert, M., Seltzer, V., Shoeman, R. L., Sierra, R. G., Soltau, H., Starodub, D., Steinbrener, J., Stier, G., Strüder, L., Svenda, M., Ullrich, J., Weidenspointner, G., White, T. A., Wunderer, C. & Ourmazd, A. (2011). *Opt. Express*, **19**, 16542–16549.
- Yoon, C. H., Yurkov, M. V., Schneidmiller, E. A., Samoylova, L., Buzmakov, A., Jurek, Z., Ziaja, B., Santra, R., Loh, N. D., Tschentscher, T. & Mancuso, A. P. (2016). *Sci. Rep.* **6**, 24791.
- Zhang, J., Baker, M. L., Schröder, G. F., Douglas, N. R., Reissmann, S., Jakana, J., Dougherty, M., Fu, C. J., Levitt, M., Ludtke, S. J., Frydman, J. & Chiu, W. (2010). *Nature*, **463**, 379–383.
- Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. (2021). *Nat. Methods*, **18**, 176–185.