



A workflow for single-particle structure determination via iterative phasing of rotational invariants in fluctuation X-ray scattering

Tim B. Berberich,^{a,b} Serguei L. Molodtsov^{a,c,d} and Ruslan P. Kurta^{a*}

^aEuropean XFEL, Holzkoppel 4, 22869 Schenefeld, Germany, ^bl. Institute of Theoretical Physics, University of Hamburg, Notkestraße 9-11, 22607 Hamburg, Germany, ^cInstitute of Experimental Physics, TU Bergakademie Freiberg, Leipziger Straße 23, 09599 Freiberg, Germany, and ^dCenter for Efficient High Temperature Processes and Materials Conversion (ZeHS), TU Bergakademie Freiberg, Winklerstrasse 5, 09599 Freiberg, Germany. *Correspondence e-mail: ruslan.kurta@xfel.eu

Received 18 September 2023

Accepted 29 January 2024

Edited by A. Barty, DESY, Hamburg, Germany

This article is part of a collection of articles from the IUCr 2023 Congress in Melbourne, Australia, and commemorates the 75th anniversary of the IUCr.

Keywords: fluctuation X-ray scattering; single-particle imaging; iterative phasing; X-ray free-electron lasers.

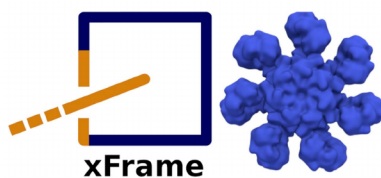
Supporting information: this article has supporting information at journals.iucr.org/j

Fluctuation X-ray scattering (FXS) offers a complementary approach for nano- and bioparticle imaging with an X-ray free-electron laser (XFEL), by extracting structural information from correlations in scattered XFEL pulses. Here a workflow is presented for single-particle structure determination using FXS. The workflow includes procedures for extracting the rotational invariants from FXS patterns, performing structure reconstructions via iterative phasing of the invariants, and aligning and averaging multiple reconstructions. The reconstruction pipeline is implemented in the open-source software *xFrame* and its functionality is demonstrated on several simulated structures.

1. Introduction

Advances in X-ray sources and instrumentation over recent decades (Jaeschke *et al.*, 2016) have been accompanied by an extensive development of techniques and methods for X-ray diffraction and imaging (Chapman *et al.*, 2006; Rodenburg, 2008; Nugent, 2010; Nakasako *et al.*, 2020). The emergence of high-power X-ray free-electron lasers (XFELs) (Ueda, 2018) opened new horizons for crystallographic studies of biological materials (Chapman *et al.*, 2011; Boutet *et al.*, 2012; Wiedorn *et al.*, 2018), most importantly in the time domain (Pandey *et al.*, 2020; Orville, 2020). At the same time, intense and ultrashort X-ray pulses produced by an XFEL made it possible to carry out ‘diffraction before destruction’ experiments on individual bioparticles (Bogan *et al.*, 2008; Mancuso *et al.*, 2010; Seibert *et al.*, 2011; Hantke *et al.*, 2014; Kimura *et al.*, 2014; Ekeberg *et al.*, 2015; Rose *et al.*, 2018), as it was previously envisioned (Solem & Baldwin, 1982; Neutze *et al.*, 2000; Gaffney & Chapman, 2007). Such experiments open up the possibility of imaging particles for which it is difficult or impossible to obtain crystals. At the same time, while serial crystallography with an XFEL allows for high-resolution structure determination of proteins and macromolecules, single-particle imaging (SPI) is a developing way to provide biologically significant structural information (Aquila *et al.*, 2015; Chapman, 2019; Bielecki *et al.*, 2020).

Several approaches have been proposed so far for structure determination of bioparticles from scattering measurements with an XFEL. The most common approach is based on iterative phasing of the measured single-particle intensity patterns (Fienup, 1982; Marchesini, 2007), which enables *ab initio* high-throughput imaging of 2D structure projections (Seibert *et al.*, 2011; Hantke *et al.*, 2014; Kimura *et al.*, 2014).



Published under a CC BY 4.0 licence

Since, generally, a complete 3D structure of a particle is of interest, it is necessary to assemble the 3D scattered intensity distribution from 2D diffraction patterns, measured from reproducible particles in unknown orientations (Ekeberg *et al.*, 2015; Rose *et al.*, 2018; Nakano *et al.*, 2018; Assalauova *et al.*, 2020). The latter is known as the orientation determination problem and is often solved using Bayesian methods (Loh & Elser, 2009; Flamant *et al.*, 2016) or other approaches (Bortel & Tegze, 2011; Yefanov & Vartanyants, 2013; Fung *et al.*, 2009; Giannakis *et al.*, 2012; Kassemeyer *et al.*, 2013; Nakano *et al.*, 2017). Such a two-step SPI approach, involving orientation determination and iterative phasing, can only be applied to the diffraction snapshots measured from individual particles. This might be challenging to accomplish in practice for arbitrary small bioparticles such as proteins, for which the individual snapshots are weak and noisy (Ekeberg *et al.*, 2022).

One of the possible ways to study weakly scattering noncrystalline particles is to perform X-ray measurements on a multiparticle system, as realized in biological small-angle X-ray scattering (SAXS) (Chaudhuri *et al.*, 2017; Vela & Svergun, 2020). Such solution scattering measurements are, however, associated with loss of information caused by the rotational averaging of intensities from individual particles in the ensemble, and typically result in a low-resolution fit of the particle structure. At the same time, taking solution scattering snapshots by XFEL pulses which are shorter than the characteristic rotational diffusion time of the particles allows one to measure structural information that is usually inaccessible in traditional SAXS at synchrotron sources. This additional information is hidden in the scattered intensity fluctuations defined by an instantaneous configuration of the ensemble of particles, and can be extracted by means of angular cross-correlation functions (CCFs). The fluctuation X-ray scattering (FXS) approach thus seeks to determine the structure of a single particle, by using statistically averaged CCFs measured from a dilute multiparticle system (Kam, 1977, 1980; Kam *et al.*, 1981). Therefore, FXS can potentially bridge the gap between conventional imaging and crystallographic methods.

FXS is a natural extension of SAXS because it also relies on rotationally invariant descriptions of the 3D single-particle intensity distribution (Kurta *et al.*, 2017). Similarly to SAXS, forward modeling approaches are also applicable to FXS data, where the reciprocal-space constraints are expressed by CCFs (Liu *et al.*, 2013; Malmerberg *et al.*, 2015; Kurta *et al.*, 2017). In fact, the information content of FXS measurements is substantially higher as compared with SAXS. For instance, in the case of 2D structure determination, it has been shown that the information accessible via FXS is equivalent to complete knowledge of the 2D single-particle intensity pattern (Kurta *et al.*, 2013; Pedrini *et al.*, 2013). Moreover, despite the limited information content of two-point CCFs (Elser, 2011), they are sufficient to produce successful *ab initio* 3D structure reconstructions (Donatelli *et al.*, 2015; Kurta *et al.*, 2017; Pande *et al.*, 2018).

Although the idea of biological FXS was formulated almost half a century ago (Kam, 1977), it was first put into practice only recently with the advent of XFELs (Kurta *et al.*, 2017;

Pande *et al.*, 2018). Progress in the development of X-ray instrumentation and sample delivery systems has led to a recent surge in FXS-related activity (Wochner *et al.*, 2009; Altarelli *et al.*, 2010; Saldin *et al.*, 2010, 2011; Kirian *et al.*, 2011; Kurta *et al.*, 2012; Starodub *et al.*, 2012; Mendez *et al.*, 2016; Martin, 2017). At the same time, practical applications of FXS are still quite rare compared with the more traditional SPI or SAXS (Kurta *et al.*, 2016). The availability of relevant practical algorithms and open-source software that implement the quite involved and often obscure mathematical apparatus of FXS may help to advance in this direction. Here we present a workflow for single-particle structure determination via iterative phasing based on rotational invariants which are accessible in FXS. The workflow is implemented in the open-source software suite *xFrame*, which includes methods for computing the CCF, extracting rotational invariants and performing structure reconstructions, as well as subsequent alignment and averaging of multiple reconstructions.

2. Theoretical background

2.1. Fluctuation X-ray scattering

We first define the real-space single-particle electron density by $\rho(\mathbf{r})$ and the corresponding scattered X-ray intensity distribution in reciprocal space by $I(\mathbf{q})$, where $\mathbf{r}$ and $\mathbf{q}$ are the real- and reciprocal-space vectors, respectively. Within the kinematic X-ray scattering approximation the density $\rho(\mathbf{r})$ is related to the scattered intensity $I(\mathbf{q})$ in the far field via the absolute square of its Fourier transform $\hat{\rho}(\mathbf{q})$,

$$I(\mathbf{q}) = K(\mathbf{q}) \left| \int \rho(\mathbf{r}) \exp(-i\mathbf{q}\mathbf{r}) d\mathbf{r} \right|^2 = K(\mathbf{q}) |\hat{\rho}(\mathbf{q})|^2, \quad (1)$$

where $K(\mathbf{q})$ is a $\mathbf{q}$ -dependent term which encompasses relevant experimental factors, *e.g.* polarization of X-rays, incident intensity fluctuations *etc.* [see, for instance, a review of possible intensity corrections in SAXS experiments (Pawn, 2013)]. Hereafter, we assume that each experimental image can be properly corrected for $K(\mathbf{q})$, so that the resulting scaled $I(\mathbf{q})$ is defined only by the electron density of the sample. We also assume that any background scattering present in realistic measurements (*e.g.* solvent scattering or parasitic scattering from beamline components), which is neglected in equation (1), can also be properly corrected. We can then consider realizations of an ensemble of $N_p \geq 1$ reproducible particles, randomly positioned and oriented in space. Similarly to conventional SAXS from dilute solutions of biological particles (Vela & Svergun, 2020), we assume scattering conditions such that interference scattering between different particles can be neglected. Furthermore, let us denote the electron density of a particular instance of the dilute multiparticle system by $\rho_\omega(\mathbf{r})$ and its scattered intensity by $I_\omega(\mathbf{q})$, where ω stands for the orientation states of all contained N_p particles. In a typical FXS experiment (Fig. 1), the instantaneous scattered intensity $I_\omega^{E_\lambda}(\mathbf{q})$ measured on a large-area detector represents a 2D cut of the 3D scattered intensity $I_\omega(\mathbf{q})$ defined by a portion of the Ewald sphere E_λ , where λ denotes the wavelength of the incident X-ray beam, indicating the

dependence of the Ewald sphere radius on the photon energy. This geometric condition can be formulated in spherical coordinates ($q \geq 0, 0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi$) as a q dependence of the polar angle θ (Saldin *et al.*, 2009),

$$\theta(q) = \pi/2 - \arcsin[q/(2\kappa)], \quad (2)$$

where $\kappa = 2\pi/\lambda$ is the angular wavenumber, $q = |\mathbf{q}| = 2\kappa \sin(\alpha/2)$ is the magnitude of the scattering vector, α is the scattering angle, and thus $I_{\omega}^{E_{\lambda}}(q, \phi) = |\hat{\rho}_{\omega}[q, \theta(q), \phi]|^2$ [see Fig. 1(b)]. By slight abuse of notation, we shall from now on use I_{ω} to denote both the full 3D intensity distribution of a sample and its 2D cut $I_{\omega}^{E_{\lambda}}$ along the Ewald sphere.

The central idea of FXS is that from a collection of M scattering images $I_{\omega_i}(q, \phi)$, $i = 1, \dots, M$, corresponding to M random realizations of the multiparticle system, information about the single-particle electron density $\rho(\mathbf{r})$ [equation (1)] can be extracted. This can be accomplished using the angular CCFs (Kam, 1977, 1980). In the present study we employ the average two-point CCF defined at distinct momentum transfer magnitudes q and q' as (Kam, 1977)

$$C_M(q, q', \Delta) = \frac{1}{2\pi M} \sum_{i=1}^M \int_0^{2\pi} I_{\omega_i}(q, \phi) I_{\omega_i}(q', \Delta + \phi) d\phi, \quad (3)$$

where $0 \leq \Delta < 2\pi$ is the angular coordinate and statistical averaging is performed over M scattering patterns.

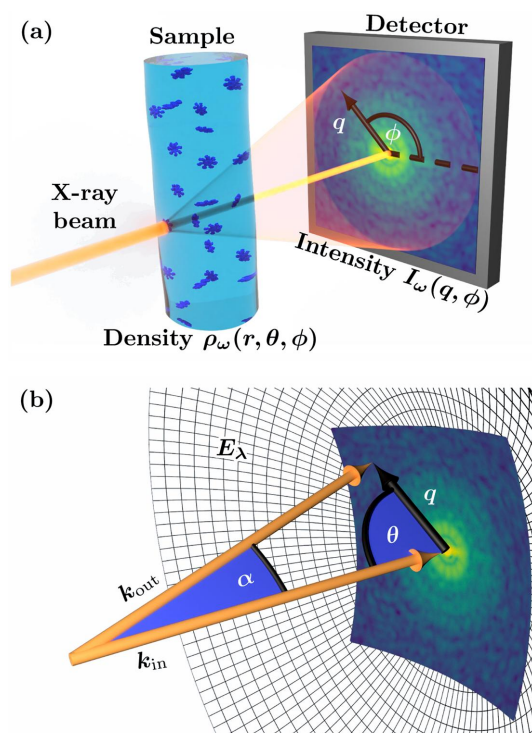


Figure 1

Scattering geometry of an FXS experiment. (a) The incident X-ray beam is diffracted from a sample solution (or aerosol) and recorded in the far field on a 2D detector. Here ρ_{ω} corresponds to the illuminated portion of the dilute sample and I_{ω} denotes the scattered intensity. (b) Measured diffraction pattern mapped on the Ewald sphere E_{λ} , where $\mathbf{q} = \mathbf{k}_{\text{out}} - \mathbf{k}_{\text{in}}$ is the scattering vector, $\mathbf{k}_{\text{in}}$ is the wavevector of the incident X-ray pulse, $\mathbf{k}_{\text{out}}$ is the wavevector of the scattered pulse and $|\mathbf{k}_{\text{in}}| = |\mathbf{k}_{\text{out}}| = \kappa$.

2.2. Rotational invariants in fluctuation X-ray scattering

In order to establish the connection between the single-particle intensity $I(\mathbf{q})$ [equation (1)] and $C_M(q, q', \Delta)$ [equation (3)], it is customary to express $I(\mathbf{q})$ using a suitable orthonormal basis. In this work we consider two cases of practical interest, which we shall call the 2D and 3D cases: they correspond to uniform distributions of particle orientations ω_i over (a) the rotation group SO(2) in two dimensions or (b) the rotation group SO(3) in three (see Fig. 2). Using circular harmonics in the 2D case and spherical harmonics in the 3D case, it is possible to show that $I(\mathbf{q})$ and $C_M(q, q', \Delta)$ can be related via the rotational invariants B_n and B_l , respectively (Kam, 1977; Saldin *et al.*, 2009; Altarelli *et al.*, 2010, 2012; Donatelli *et al.*, 2015; Kurta *et al.*, 2016).

2.2.1. 2D case: rotational invariants B_n . In the 2D case we are interested in a 2D projection of the 3D particle structure, under the constraint that orientations of the particles composing a dilute system can only differ from each other by rotations around axes parallel to the incident X-ray beam [Figs. 2(a) and 2(b)]. Experimentally such situations have been realized in a study of nanoparticles deposited on a membrane (Pedrini *et al.*, 2013). According to the projection-slice theorem, the 2D projection of the structure is related to the scattered intensity distribution measured in a plane that cuts reciprocal space orthogonal to the incident-beam direction and passes through the reciprocal-space origin. Such measurements can only be performed at ‘flat Ewald sphere’

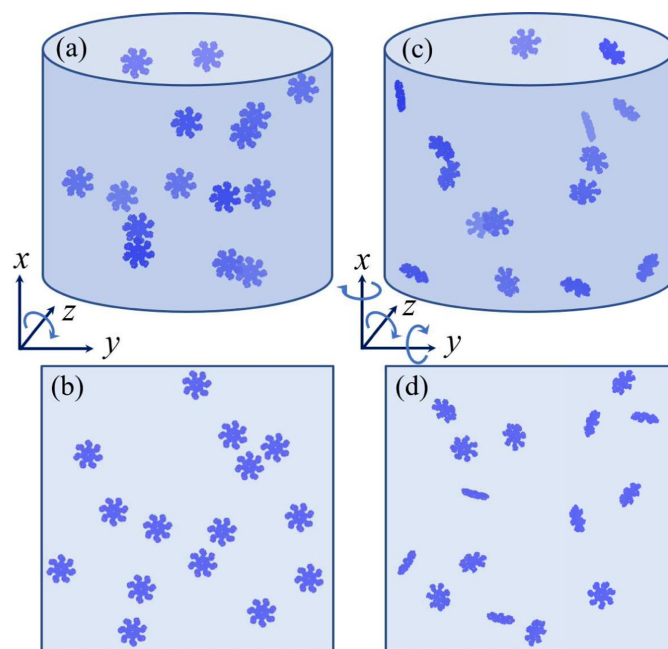


Figure 2

Snapshots of the samples corresponding to the (a), (b) 2D and (c), (d) 3D cases (see Section 2.2). Particle orientations are uniformly distributed (a), (b) over the rotation group SO(2) about the z axis (which is parallel to the incident X-ray beam direction) and (c), (d) over the rotation group SO(3). Here (a) and (c) are the bulk 3D samples, while (b) and (d) are the planar 2D samples.

conditions, *e.g.* in SAXS geometry, when $\theta(q) \simeq \pi/2$ [see equation (2)].

Considering the 2D single-particle scattering intensity [equation (1)] in polar coordinates (q, ϕ) , the circular harmonic expansion (Fourier series expansion) of $I(q, \phi)$ can be specified as

$$I(q, \phi) = \sum_{n=-\infty}^{\infty} I_n(q) \exp(in\phi), \quad (4a)$$

$$I_n(q) = \frac{1}{2\pi} \int_0^{2\pi} I(q, \phi) \exp(-in\phi) d\phi, \quad (4b)$$

where $I_n(q)$ are circular harmonic expansion coefficients of the single-particle scattering intensity.

Using equation (4a) in (3), the average CCF at $M \rightarrow \infty$ can be written as (Kurta *et al.*, 2013)

$$\mathcal{C}(q, q', \Delta) \equiv \lim_{M \rightarrow \infty} C_M(q, q', \Delta) = \sum_{n=-\infty}^{\infty} B_n(q, q') \exp(in\Delta), \quad (5)$$

which identifies the invariants $B_n(q, q')$ as circular harmonic expansion coefficients of $\mathcal{C}(q, q', \Delta)$. Similar to the 3D case, this result is valid for a dilute system of particles ($N_p \geq 1$), with $B_n(q, q')$ expressed as

$$B_n(q, q') = I_n(q) I_n^*(q') \begin{cases} N_p^2, & \text{if } n = 0, \\ N_p, & \text{otherwise,} \end{cases} \quad (6)$$

where the asterisk ‘*’ denotes complex conjugation.

Note that equation (6) provides a direct connection between the experimentally accessible invariants B_n and the harmonic expansion coefficients $I_n(q)$ of the 2D single-particle intensity. The rotational invariance of B_n is a direct consequence of the Fourier shift theorem. It implies that a rotation $\mathcal{R}(\varphi)$ by an angle φ acts on the harmonic coefficients $I_n(q)$ by multiplication with a phase factor, *i.e.*

$$I_n(q) \xrightarrow{\mathcal{R}(\varphi)} I_n(q) \exp(in\varphi). \quad (7)$$

2.2.2. 3D case: rotational invariants B_l . In the 3D case we are interested in the 3D structure of particles, while the orientations of the particles composing a dilute system are uniformly distributed over SO(3) [Figs. 2(c) and 2(d)]. This situation corresponds to typical conditions in conventional biological SAXS measurements. The spherical harmonic expansion of the single-particle scattered intensity $I(q, \theta, \phi)$ [equation (1)] can be specified as

$$I(q, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l I_m^l(q) Y_m^l(\theta, \phi), \quad (8a)$$

$$I_m^l(q) = \frac{2\pi}{0} \int_0^\pi I(q, \theta, \phi) Y_m^{l*}(\theta, \phi) \sin \theta d\theta d\phi, \quad (8b)$$

where $Y_m^l(\theta, \phi)$ are spherical harmonics and $I_m^l(q)$ denote the expansion coefficients. By substituting equation (8a) into (3), the average CCF at $M \rightarrow \infty$ can be expressed via (Kam, 1977; Saldin *et al.*, 2009)

$$\begin{aligned} \mathcal{C}(q, q', \Delta) &\equiv \lim_{M \rightarrow \infty} C_M(q, q', \Delta) \\ &= \sum_{l=0}^{\infty} B_l(q, q') F_l(q, q', \Delta), \end{aligned} \quad (9)$$

where $B_l(q, q')$ represent the rotational invariants and $F_l(q, q', \Delta)$ is defined using Legendre polynomials P_l via

$$F_l(q, q', \Delta) = P_l(\sin \theta \sin \theta' \cos \Delta + \cos \theta \cos \theta'), \quad (10)$$

in which the angles θ and θ' are related to q and q' using equation (2). Importantly, equation (9) is also valid when averaging $C_M(q, q', \Delta)$ over scattering intensities $I_{\omega_i}(q, \phi)$ from a dilute system of ($N_p \geq 1$) particles and the invariants $B_l(q, q')$ can be expressed in terms of the expansion coefficients $I_m^l(q)$ by

$$B_l(q, q') = \sum_{m=-l}^l I_m^l(q) I_m^{l*}(q') \begin{cases} N_p^2, & \text{if } l = 0, \\ N_p, & \text{otherwise.} \end{cases} \quad (11)$$

The rotational invariance of $B_l(q, q')$ is a direct consequence of the fact that for a given order l the spherical harmonics $Y_m^l(\theta, \phi)$ satisfy

$$\sum_{m=-l}^l Y_m^l(\theta, \phi) Y_m^{l*}(\theta, \phi) = (2l+1)/4\pi$$

which, as a constant, is invariant under rotations. Alternatively, consider that a rotation $\mathcal{R}(\alpha, \beta, \gamma)$ in SO(3) defined by Euler angles (α, β, γ) acts on the expansion coefficients $I_m^l(q)$ via the Wigner D matrices $D_{nm}^l(\alpha, \beta, \gamma)$, *i.e.*

$$I_m^l(q) \xrightarrow{\mathcal{R}(\alpha, \beta, \gamma)} \sum_{n=-l}^l I_n^l(q) D_{nm}^l(\alpha, \beta, \gamma), \quad (12)$$

which satisfy the orthogonality condition

$$\sum_{m=-l}^l D_{nm}^{l*}(\alpha, \beta, \gamma) D_{n'm}^l(\alpha, \beta, \gamma) = \delta_{nn'},$$

where $\delta_{nn'}$ is the Kronecker delta (Rose, 1957).

2.3. Information content of B_n and B_l

For simplicity, we shall proceed by considering the single-particle case ($N_p = 1$). Before making use of the invariants B_n and B_l [equations (6) and (11)] for single-particle structure recovery, it is instructive to understand how much information they retain about the single-particle intensity $I(\mathbf{q})$. Considering a discretization of the momentum transfer variable q [see equation (47) in Appendix A], it is possible to treat the harmonic coefficients as complex matrices, that is $\mathbf{I}_n$ for fixed n is a matrix of size $N \times 1$ (*i.e.* a column vector) with elements $I_n(q_k)$, and $\mathbf{I}_l$ for fixed l is a matrix of size $N \times (2l+1)$ with elements $I_m^l(q_k)$. This allows one to express the invariants for fixed orders n and l as matrix products,

$$\mathbf{B}_n = \mathbf{I}_n \mathbf{I}_n^\dagger, \quad \mathbf{B}_l = \mathbf{I}_l \mathbf{I}_l^\dagger, \quad (13)$$

where the symbol ‘†’ denotes the conjugate transpose. By construction, these are positive semi-definite Hermitian matrices, and thus diagonalizable with positive eigenvalues (Saldin *et al.*, 2009; Donatelli *et al.*, 2015). The maximal rank of $\mathbf{B}_n$ is 1, while the maximal rank of $\mathbf{B}_l$ is $N_l = \min(2l+1, N)$.

Together this means that there exists a complex vector $\mathbf{v}_n$ of length N and a positive eigenvalue λ_n , as well as a complex $N \times N_l$ matrix $\mathbf{V}_l$ and a diagonal matrix $\mathbf{\Lambda}_l$ of eigenvalues $\lambda_{l,1}, \dots, \lambda_{l,N_l}$, such that

$$\mathbf{B}_n = \mathbf{v}_n \lambda_n \mathbf{v}_n^\dagger, \quad \mathbf{B}_l = \mathbf{V}_l \mathbf{\Lambda}_l \mathbf{V}_l^\dagger. \quad (14)$$

Equations (13) and (14) show two different decompositions of the same positive semi-definite matrices $\mathbf{B}_n$ and $\mathbf{B}_l$ in the 2D and 3D cases, respectively. Since such decompositions are unique up to unitary transformations (Kam, 1977), there exists a complex phase factor $\mathbf{u}_n$ such that $\mathbf{u} \mathbf{u}^\dagger = 1$ (here the phase factor is a complex number with an absolute value of 1 or, equivalently, a complex unitary matrix $\mathbf{u}$ of size 1×1 , i.e. a singleton matrix), as well as a complex matrix $\mathbf{U}_l$ of size $N_l \times (2l + 1)$ such that $\mathbf{U}_l \mathbf{U}_l^\dagger = \text{id}$ (hereafter 'id' stands for the identity matrix), which satisfy

$$\mathbf{I}_n = \mathbf{v}_n \sqrt{\lambda_n} \mathbf{u}_n = \tilde{\mathbf{v}}_n \mathbf{u}_n, \quad \mathbf{I}_l = \mathbf{V}_l \sqrt{\mathbf{\Lambda}_l} \mathbf{U}_l = \tilde{\mathbf{V}}_l \mathbf{U}_l, \quad (15)$$

where for brevity we defined $\tilde{\mathbf{v}}_n = \mathbf{v}_n \sqrt{\lambda_n}$ and $\tilde{\mathbf{V}}_l = \mathbf{V}_l \sqrt{\mathbf{\Lambda}_l}$. One may notice the formal analogy of the expressions for $\mathbf{I}_n$ and $\mathbf{I}_l$. The information contained in each invariant is thus enough to specify the corresponding intensity harmonic coefficients, $I_n(q)$ or $I_m^l(q)$, up to a unitary matrix for each expansion order. In the present context, the problem of determining these unknown unitary matrices is analogous to solving the orientation determination problem in SPI.

In coherent X-ray diffraction imaging (CXDI), and particularly in SPI, to reconstruct the real-space structure one seeks to solve the phase problem using the measured scattering intensities as constraints (Chapman *et al.*, 2006). The inverse problem in FXS can be solved in a similar way, by employing the measured invariants B_n and B_l as constraints. The main approaches to solving the inverse problem in FXS include analytical phasing (Kurta *et al.*, 2013; Pedrini *et al.*, 2013), iterative phasing (Donatelli *et al.*, 2015; Kurta *et al.*, 2017; Pande *et al.*, 2018) and optimization (Saldin *et al.*, 2011; von Ardenne *et al.*, 2018). The single-particle structure reconstruction workflow presented below is based on the multitiered iterative phase retrieval algorithm (MTIP) (Donatelli *et al.*, 2015). The MTIP algorithm represents a generalization of conventional iterative phasing schemes employed in CXDI and enables *ab initio* 2D and 3D single-particle structure recovery (e.g. without symmetry constraints) from the rotational invariants B_n and B_l .

3. Single-particle structure reconstruction workflow

A complete workflow for single-particle structure determination from diffraction patterns measured in an FXS experiment includes a number of procedures:

- (i) Statistical averaging of the angular two-point CCF.
- (ii) Extraction of rotational invariants (B_n or B_l) from the CCF.
- (iii) Reconstruction of the single-particle density $\rho(\mathbf{r})$ and intensity $I(\mathbf{q})$ via iterative phasing using the invariants.
- (iv) Alignment and averaging of the reconstruction results.

A detailed description of the procedures implemented in our reconstruction workflow is provided in the following subsections.

3.1. Calculation of the average two-point CCF

In practical calculations of the angular CCF (3) we consider a uniform polar grid. The angular grid points, Δ_t and ϕ_t , are given by $\Delta_t = \phi_t = t2\pi/N_\phi$, where N_ϕ is the number of angular grid points, and the radial sampling points are defined in equation (47). The average two-point CCF can be determined on this grid as

$$C_M(q_k, q_p, \Delta_t) = \frac{1}{M} \sum_{i=1}^M \left[\sum_{j=0}^{N_\phi-1} I_{\omega_i}(q_k, \phi_j) W_i(q_k, \phi_j) \times I_{\omega_i}(q_p, \Delta_t + \phi_j) W_i(q_p, \Delta_t + \phi_j) \right] / \left[\sum_{j=0}^{N_\phi-1} W_i(q_k, \phi_j) W_i(q_p, \Delta_t + \phi_j) \right], \quad (16)$$

where $W_i(q_k, \phi_j)$ is a binary mask for the i th image (Zaluzhnyy *et al.*, 2017). The mask has the value of 0 for all sampling points (q_k, ϕ_j) for which image data should be excluded (masked) from the analysis, and the value of 1 otherwise. Equation (16) suggests that, in practice, the CCF can be successfully determined even if many pixels are masked on individual diffraction patterns (see Section S4 of the supporting information), in the limiting case allowing measurements to be made using two-point detectors [see, for instance, Clark *et al.* (1983)].

3.2. Extraction of the rotational invariants

In the 2D case, equation (5) directly identifies the invariants $B_n(q, q')$ as the coefficients of the circular harmonic expansion of the averaged CCF $\mathcal{C}(q, q', \Delta)$, and thus their determination is straightforward. In the 3D case, the relation between the invariants $B_l(q, q')$ and $\mathcal{C}(q, q', \Delta)$, given by equations (9) and (10), is more complicated. At flat Ewald sphere conditions (small-angle approximation) equation (10) simplifies to $F_l(q, q', \Delta) = P_l(\cos \Delta)$. In this case equation (9) represents the Legendre series expansion of $\mathcal{C}(q_1, q_2, \Delta)$, and the invariants $B_l(q, q')$ can be extracted by applying the inverse Legendre transform to $\mathcal{C}(q_1, q_2, \Delta)$. In general, at curved Ewald sphere conditions $[\theta(q) \neq \pi/2]$, the inverse Legendre transform extraction is not applicable since $F_l(q, q', \Delta)$ do not satisfy the orthogonality property of Legendre polynomials; therefore other approaches need to be applied.

A common way of extracting the invariants $B_l(q, q')$ is to consider equation (9), for each pair of fixed q and q' and discretized angular coordinate Δ [see $C_M(q_k, q_p, \Delta_t)$ in equation (16)], as a system of N_Δ linear equations. By treating the CCF $\mathcal{C}(q, q', \Delta)$ at fixed q and q' as a vector $\mathbf{C}_\Delta$ of size N_Δ , the invariant $B_l(q, q')$ as a vector $\mathbf{B}_l$ of size $L + 1$, and $F_l(q, q', \Delta)$ as an $N_\Delta \times (L + 1)$ matrix $\mathbf{F}_\Delta^l$, it is possible to specify this system of linear equations as

$$\mathbf{C}_\Delta = \mathbf{F}_\Delta^l \mathbf{B}_l, \quad (17)$$

where L is the maximal considered invariant order, defined by the complexity of the particle structure and extent of the measured correlation data in reciprocal space. In practical applications, usually $N_\Delta > L$; therefore, the matrix $\mathbf{F}_\Delta^l$ is not square, and the overdetermined system of linear equations (17) may only be solved approximately. Usually least-squares methods such as the pseudo-inversion (Ford, 2014), based on a singular value decomposition of $\mathbf{F}_\Delta^l$, are used to solve such linear systems.

Here we propose a different extraction method for the invariants $\mathbf{B}_l$ which is based on the circular harmonic expansion coefficients of the averaged CCF. It is possible to express $F_l(q, q', \Delta)$ in equation (9) in terms of spherical harmonics $Y_m^l(\theta, \phi)$ as (Saldin *et al.*, 2009)

$$F_l(q, q', \Delta) = \frac{1}{2l+1} \sum_{m=-l}^l Y_m^l(\theta, \phi) Y_m^{l*}(\theta', \phi + \Delta), \quad (18)$$

while the spherical harmonics can be specified using the associated Legendre polynomials $P_l^{|m|}$ as

$$Y_m^l(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^{|m|}(\cos \theta) \exp(im\phi). \quad (19)$$

By combining equations (18) and (19) with (9), the averaged CCF can be expressed as

$$\begin{aligned} \mathcal{C}(q, q', \Delta) &= \sum_{l=0}^{\infty} B_l(q, q') \sum_{m=-l}^l \frac{(l-m)!}{4\pi(l+m)!} P_l^{|m|}(\cos \theta) \\ &\quad \times P_l^{|m|}(\cos \theta') \exp(-im\Delta). \end{aligned} \quad (20)$$

From this equation one may see that the circular harmonic expansion coefficients $C_n(q, q') = (1/2\pi) \int_0^{2\pi} \mathcal{C}(q, q', \Delta) \times \exp(-in\Delta) d\Delta$ take the form

$$C_n(q, q') = \sum_{l \geq |n|} B_l(q, q') \bar{P}_l^{|n|}(\cos \theta) \bar{P}_l^{|n|}(\cos \theta') \quad (21)$$

with

$$\bar{P}_l^{|n|}(\cos \theta) = \sqrt{\frac{(l-n)!}{4\pi(l+n)!}} P_l^{|n|}(\cos \theta). \quad (22)$$

Since the associated Legendre polynomials $P_l^{|n|}$ vanish for $l < |n|$, the summation in equation (21) involves only orders $l \geq |n|$. This means that $\mathbf{B}_l$ for a given l is completely determined by the circular harmonic expansion coefficients $C_0, \dots, C_l$, up to the order $n = l$. Considering a cutoff order L we obtain a linear system of equations that can be written in matrix form as

$$\mathbf{C}_n = \mathbf{P}_n^l \mathbf{B}_l, \quad (23)$$

where $n, l \leq L$, and $\mathbf{P}_n^l$ is the $(L+1) \times (L+1)$ upper-triangular matrix whose elements are

$$\mathbf{P}_n^l = \bar{P}_l^{|n|}(\cos \theta) \bar{P}_l^{|n|}(\cos \theta'). \quad (24)$$

The upper-triangular linear system (23) can be directly solved using back-substitution (Ford, 2014). The computational complexity of the proposed method of extraction of $\mathbf{B}_l$,

including calculation of the Fourier coefficients $C_n(q, q')$, is given by $O[N_\Delta \log(N_\Delta) + (L+1)^2]$, as opposed to $O[N_\Delta^2(L+1)]$ when solving equation (17) directly using singular value decomposition.

Note that $B_0(q, q')$ can be directly determined from the measured SAXS intensity profiles, $I_{\text{SAXS}}(q) = [1/(2\pi M)] \times \sum_{i=1}^M \int_0^{2\pi} I_{\omega_i}(q, \phi) d\phi$ [see equation (3)], as $B_0(q, q') = 4\pi I_{\text{SAXS}}(q) I_{\text{SAXS}}(q')$, where the 4π factor is due to the chosen normalization of the spherical harmonics [see equation (19)]. This way of determining $B_0(q, q')$ should be preferred in the case of noisy diffraction patterns, and is unavoidable if more complex forms of the CCFs are applied to mitigate background scattering [see, for instance, Kurta *et al.* (2017)].

3.3. Iterative phasing using the MTIP algorithm

For *ab initio* single-particle structure determination from the extracted invariants $\mathbf{B}_n$ or $\mathbf{B}_l$, we employ the MTIP algorithm (Donatelli *et al.*, 2015; Kommera *et al.*, 2021) with certain modifications (see Sections 3.3.3 and 3.3.4). Similarly to other iterative phasing methods used in CXDI/SPI, the single-particle structure is recovered by iteratively enforcing constraints in real and reciprocal spaces. Details of the implemented iterative phasing loop are described in the following subsections.

3.3.1. Real-space constraints. A finite size of a particle leads to the formulation of the commonly used real-space support constraint, which defines a region of space where the electron density is expected to have nonzero values. Each reconstruction run starts with an initial random guess for the electron density $\rho(\mathbf{r})$ of a particle. In our implementation the initial support function $S(\mathbf{r})$ is defined as a sphere, which can be optimally set to the expected size of the reconstruction target. As the reconstruction progresses the support constraint is systematically updated according to the shrinkwrap (SW) algorithm (Marchesini *et al.*, 2003), that is the updated support is determined as an isosurface at a specified threshold of the convolution between the current density guess $\rho'(\mathbf{r})$ and a Gaussian function, *i.e.*

$$S \rho'(\mathbf{r}) = \begin{cases} 1, & \text{if } \left\{ |\rho'(\mathbf{r})| \star \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{|\mathbf{r}|^2}{2\sigma^2}\right) \right] \right\} > \gamma, \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

Here $\star$ denotes convolution, and the free parameters σ and γ define the standard deviation of the Gaussian function and a relative threshold, respectively. Usually, the initial value of σ is chosen to be slightly above the expected full-period resolution of the reconstructed density and gradually decreases as the reconstruction progresses. The threshold γ is defined relative to the maximum density value $\max[\rho'(\mathbf{r})]$ in the current phasing iteration and remains constant.

An often encountered set of real-space constraints used in X-ray imaging can be formulated as a (density) value projection P_V , implemented here in the following general form:

$$P_V \rho'(\mathbf{r}) = \rho_{\text{Re}}(\mathbf{r}) + i\rho_{\text{Im}}(\mathbf{r}), \quad (26a)$$

with

$$\rho_{\text{Re}}(\mathbf{r}) = \begin{cases} V_{\text{Re}}^{\min}, & \text{if } \text{Re}[\rho'(\mathbf{r})] < V_{\text{Re}}^{\min}, \\ V_{\text{Re}}^{\max}, & \text{if } \text{Re}[\rho'(\mathbf{r})] > V_{\text{Re}}^{\max}, \\ \text{Re}[\rho'(\mathbf{r})], & \text{otherwise} \end{cases} \quad (26b)$$

and

$$\rho_{\text{Im}}(\mathbf{r}) = \begin{cases} V_{\text{Im}}^{\min}, & \text{if } \text{Im}[\rho'(\mathbf{r})] < V_{\text{Im}}^{\min}, \\ V_{\text{Im}}^{\max}, & \text{if } \text{Im}[\rho'(\mathbf{r})] > V_{\text{Im}}^{\max}, \\ \text{Im}[\rho'(\mathbf{r})], & \text{otherwise.} \end{cases} \quad (26c)$$

$\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ define the real and imaginary parts of the corresponding arguments, and $V_{\text{Re}}^{\min}$, $V_{\text{Re}}^{\max}$, $V_{\text{Im}}^{\min}$ and $V_{\text{Im}}^{\max}$ are free parameters. Depending on a particular choice of these four parameters, it is possible to set arbitrary bounds on the real and imaginary parts of $\rho(\mathbf{r})$, particularly to impose reality or non-negativity. The value projection of $\rho'(\mathbf{r})$ on the support function can then be defined as

$$P_{\text{SV}} \rho'(\mathbf{r}) = S P_V \rho'(\mathbf{r}). \quad (27)$$

The real-space domain incorporates the well known X-ray imaging algorithms, such as error reduction (ER) (Gerchberg & Saxton, 1972) and hybrid input–output (HIO) (Fienup, 1982). Iterative update of $\rho(\mathbf{r})$ by these methods can be expressed using the projection P_{SV} . In the ER scheme, the input electron density in the $(i + 1)$ th iteration is defined using the output of the i th iteration as (see Fig. 3)

$$\rho_{(i+1)}(\mathbf{r}) = P_{\text{SV}} \rho'_{(i)}(\mathbf{r}), \quad (28)$$

while for the HIO algorithm we have

$$\rho_{(i+1)}(\mathbf{r}) = \begin{cases} \rho'_{(i)}(\mathbf{r}), & \text{if } P_{\text{SV}} \rho'_{(i)}(\mathbf{r}) = \rho'_{(i)}(\mathbf{r}), \\ \rho_{(i)}(\mathbf{r}) - \beta[\rho'_{(i)}(\mathbf{r}) - P_{\text{SV}} \rho'_{(i)}(\mathbf{r})], & \text{otherwise,} \end{cases} \quad (29)$$

where $\beta \in (0, 1]$ is a free parameter that regulates the strength of the negative feedback (Fienup, 1982; Donatelli *et al.*, 2015).

3.3.2. Reciprocal-space constraints. The main reciprocal-space constraint is realized by means of the correlation projection P_C (see Fig. 3), which takes the current approx-

imation of the single-particle intensity harmonic coefficients I_n or I_m^l and determines the closest function, in the discrete L^2 norm, whose harmonic coefficients comply with equation (15).

In the 2D case this corresponds to finding the complex phase factors u_n (with $|u_n| = 1$) for which the discrete L^2 distance between $\mathbf{I}_n$ and $\tilde{\mathbf{v}}_n u_n$ becomes minimal (Donatelli *et al.*, 2015), that is

$$u_n = \arg \min_{u_n \text{ with } |u_n|=1} \|\mathbf{I}_n - \tilde{\mathbf{v}}_n u_n\|_q^2 = \arg \max_{u_n \text{ with } |u_n|=1} \text{Re}[\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n u_n \rangle_q] \\ = \frac{\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n \rangle_q}{|\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n \rangle_q|} = \frac{\sum_{k=0}^{N-1} I_n(q_k) \tilde{\mathbf{v}}_n^*(q_k) q_k}{|\sum_{k=0}^{N-1} I_n(q_k) \tilde{\mathbf{v}}_n^*(q_k) q_k|}. \quad (30)$$

$\|\cdot\|_q$ and $\langle \cdot, \cdot \rangle_q$ are the L^2 norm and scalar product, respectively, weighted by q_k in order to comply with a continuous L^2 norm defined on a spherical grid. In the derivation of equation (30) it was considered that the maximal real part of the scalar product $\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n u_n \rangle_q$ is obtained for the phase factor u_n that forces it to become real. Using the latter result, it is possible to formulate the correlation projection P_C in the 2D case as

$$P_C \mathbf{I}_n = \tilde{\mathbf{v}}_n \frac{\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n \rangle_q}{|\langle \mathbf{I}_n, \tilde{\mathbf{v}}_n \rangle_q|} \equiv \mathbf{I}'_n. \quad (31)$$

Analogously, in the 3D case one seeks to minimize (Donatelli *et al.*, 2015)

$$\arg \min_{\mathbf{U}_l \in \mathbf{U}(2l+1)} \|\mathbf{I}_l - \tilde{\mathbf{V}}_l \mathbf{U}_l\|_{F_q} = \arg \min_{\mathbf{U}_l \in \mathbf{U}(2l+1)} \sqrt{\sum_{m,k} q_k^2 (\mathbf{I}_l - \tilde{\mathbf{V}}_l \mathbf{U}_l)_{m,k}^2} \quad (32)$$

over all unitary matrices $\mathbf{U}_l$ of size $2l + 1$, where $\|\cdot\|_{F_q}$ is the Frobenius norm (Ford, 2014) weighted by the square of the radial points q_k^2 . The weighting factors q_k^2 are again present to comply with the L^2 norm defined on a spherical grid. Such a minimization problem is known as a unitary Procrustes problem (Gower & Dijkstra, 2004). Instead of zero padding $\tilde{\mathbf{V}}_l$ in the case of $N_l \leq (2l + 1)$, we alter the minimization constraint of equation (32) in requiring $\mathbf{U}_l$ to be a semi-unitary matrix, by which we mean an $N_l \times (2l + 1)$ matrix that satisfies $\mathbf{U}_l \mathbf{U}_l^\dagger = \text{id}$. A solution to the minimization problem (32) is found using the singular value decomposition

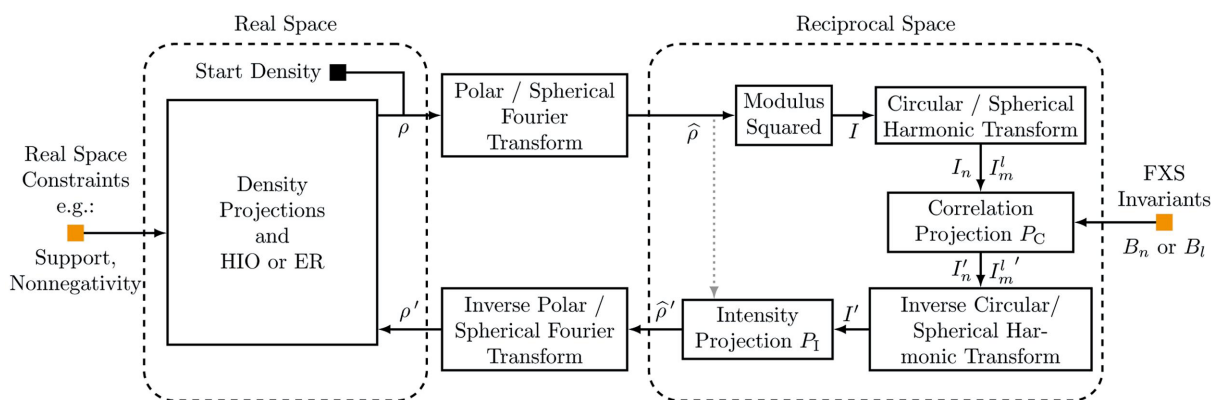


Figure 3

A scheme of the implemented MTIP loop. The filled orange squares mark the entry points for constraints, and the black square denotes the initial density guess. The quantities labeled on the scheme (ρ , I , I_n , I_m^l etc.) should be interpreted as iterative estimates of the corresponding theoretical quantities defined in equations (4)–(11).

$\mathbf{X}_l \mathbf{\Sigma}_l \mathbf{T}_l^\dagger$ of the $N_l \times (2l + 1)$ matrix $\tilde{\mathbf{V}}_l^\dagger \mathbf{D}^2 \mathbf{I}_l$, where $\tilde{\mathbf{V}}_l^\dagger$ is the conjugate transpose of $\tilde{\mathbf{V}}_l$ defined in equation (15), $\mathbf{D}$ is the $N \times N$ diagonal matrix of radial grid points $\mathbf{D} = \text{diag}(q_0, \dots, q_{N-1})$, while $\mathbf{X}_l$ and $\mathbf{T}_l^\dagger$ are unitary matrices of sizes $N_l \times N_l$ and $N_l \times (2l + 1)$, respectively, and $\mathbf{\Sigma}_l$ is an $N_l \times N_l$ diagonal matrix of non-negative singular values. The minimizing matrix $\mathbf{U}_l$ is then given by

$$\mathbf{U}_l = \mathbf{X}_l \mathbf{T}_l^\dagger, \quad (33)$$

and consequently the correlation projection P_C in the 3D case can be specified as

$$P_C \mathbf{I}_l = \tilde{\mathbf{V}}_l \mathbf{X}_l \mathbf{T}_l^\dagger = \mathbf{I}_l. \quad (34)$$

The other projection applied in reciprocal space is the intensity projection P_I defined as

$$P_I \hat{\rho}(\mathbf{q}) \equiv \hat{\rho}'(\mathbf{q}) = \begin{cases} \frac{\hat{\rho}(\mathbf{q})}{|\hat{\rho}(\mathbf{q})|} \sqrt{\text{Re}[I'(\mathbf{q})]}, & \text{if } \text{Re}[I'(\mathbf{q})] > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (35)$$

This is formulated similarly to the Fourier modulus projection, which serves as the main reciprocal-space constraint in conventional CXDI/SPI. Note that the described formalism differs from conventional SPI in that here $I'(\mathbf{q})$ is not the experimentally measured scattered intensity but rather its current approximation, which is, along with the real-space density $\rho(\mathbf{r})$, iteratively refined using the measured invariants B_n and B_l as constraints.

3.3.3. Polar and spherical Fourier transforms. Since the main reciprocal projection P_C is formulated in terms of harmonic coefficients of the scattered intensity, the need arises to implement the complete phasing loop (see Fig. 3) on a polar/spherical grid (including the Fourier transforms), in order to avoid the inaccuracies and performance limitations which would be imposed otherwise by repeated interpolations between Cartesian and polar/spherical grids. There is, however, no discrete Fourier transform in polar or spherical coordinates which would allow for repeated forward and inverse transforms. The approach applied here, as proposed by Donatelli *et al.* (2015), is to numerically approximate Hankel transforms, which connect the harmonic expansion of a function to the harmonic expansion of its Fourier transform. Consider $\rho_n(r)$ and $\rho_m^l(r)$ to be the harmonic expansion coefficients of an electron density in polar and spherical coordinates, and let $\hat{\rho}_n(q)$ and $\hat{\rho}_m^l(q)$ be the expansion coefficients of the respective scattering amplitudes (Fourier transformed densities). The connection between $\rho_n(r)$ and $\hat{\rho}_n(r)$ in the 2D case is then given by the Hankel transform

$$\hat{\rho}_m(q) = (-i)^m \int_0^\infty \rho_m(r) J_m(qr) r dr, \quad (36a)$$

$$\rho_m(r) = (i)^m \int_0^\infty \hat{\rho}_m(q) J_m(qr) q dq, \quad (36b)$$

where J_m are Bessel functions of the first kind on integer order m . In the spherical (3D) case one finds

$$\hat{\rho}_m^l(q) = \sqrt{\frac{2}{\pi}} (-i)^l \int_0^\infty \rho_m^l(r) j_l(qr) r^2 dr, \quad (37a)$$

$$\rho_m^l(r) = \sqrt{\frac{2}{\pi}} (i)^l \int_0^\infty \hat{\rho}_m^l(q) j_l(qr) q^2 dq, \quad (37b)$$

where j_l are spherical Bessel functions. One approach to numerically approximate the continuous Hankel transforms given in equations (36) and (37) is to expand the harmonic coefficients $\rho_m(r)$ or $\rho_m^l(r)$ (and their reciprocal-space counterparts) in some orthogonal basis, thereby shifting the Hankel integral to the expansion functions (see Appendix B). In the original version of MTIP this is accomplished using the cosine/sine series expansions (Donatelli *et al.*, 2015) (see Appendix B2). We also developed another approximation based on Zernike polynomial expansions, which allowed us to obtain closed-form expressions for the quadrature weights of the discretized Hankel transform (see Appendix B3). Further investigation, however, showed that both approaches converge to direct approximations of the integrals in equations (36) and (37) using a Riemann sum (see Appendices B1 and B4). We therefore employ the midpoint rule as a default approximation scheme for the Hankel integrals in our reconstruction workflow (see Appendix B4). In the 2D case, the Hankel transform (36) can thus be approximated on a discrete polar grid as

$$\hat{\rho}_m(q_k) \simeq \frac{(-i)^m}{Q_{\max}^2} \sum_{p=0}^{N-1} \rho_m(r_p) \omega_m(p, k), \quad (38a)$$

$$\rho_m(r_p) \simeq \frac{i^m}{R_{\max}^2} \sum_{k=0}^{N-1} \hat{\rho}_m(q_k) \omega_m(k, p), \quad (38b)$$

with the quadrature weights $\omega_m(p, k)$ being defined by

$$\omega_m(p, k) = \frac{\pi^2(1+2p)}{2} J_m \left[\frac{\pi}{4N} (1+2p)(1+2k) \right]. \quad (38c)$$

In the 3D case, the spherical Hankel transform (37) is approximated by

$$\hat{\rho}_m^l(q_k) \simeq \frac{(-i)^l}{Q_{\max}^3} \sum_{p=0}^{N-1} \rho_m^l(r_p) \omega_l(p, k), \quad (39a)$$

$$\rho_m^l(r_p) \simeq \frac{i^l}{R_{\max}^3} \sum_{k=0}^{N-1} \hat{\rho}_m^l(q_k) \omega_l(k, p), \quad (39b)$$

using the quadrature weights

$$\omega_l(p, k) = \frac{\pi^2 \sqrt{\pi}}{2\sqrt{2}} (1+2p)^2 j_l \left[\frac{\pi}{4N} (1+2p)(1+2k) \right]. \quad (39c)$$

Note that the weights in the inverse transforms (38b) and (39b) are determined by transposing the parameters p and k in the weight functions specified for the forward transforms in equations (38c) and (39c), respectively.

3.3.4. Fourier transform stabilization. We empirically found that stabilizing the Fourier transforms in the iterative loop by the following procedure may improve the convergence of reconstructions. The basic idea behind this operation is to reduce possible errors due to approximating the continuous Fourier transform in each MTIP iteration (see Section 3.3.3). Using the notation in Fig. 3 this correction can be expressed by modifying the definition of $\rho' = \text{FT}^{-1}(\hat{\rho}')$ as

$$\rho' = \text{FT}^{-1}(\hat{\rho}') + [\rho - \text{FT}^{-1}(\hat{\rho})], \quad (40)$$

where FT^{-1} denotes the inverse Fourier transform. In the limit of a completely converged MTIP reconstruction, *i.e.* when $\hat{\rho} = \hat{\rho}'$, meaning that the reciprocal-space projections do not change the intensity anymore, this definition ensures that the modified density ρ' coincides with the input density ρ of the current iteration. Without this procedure ρ' and ρ would differ due to the applied Fourier transform approximations.

3.3.5. Error metrics. The evolution of the iterative phasing process can be tracked using several metrics, which may serve as convergence and error estimates. In analogy to error metrics commonly used in conventional X-ray imaging (Fienup, 1978, 1982), we define the relative normalized errors in reciprocal and real space as (see Fig. 3)

$$E_{\text{reciprocal}} = \frac{\|I(\mathbf{q}) - I'(\mathbf{q})\|_{L^2}}{\|I(\mathbf{q})\|_{L^2}}, \quad (41a)$$

$$E_{\text{real}} = \frac{\|\rho'(\mathbf{r}) - P_{\text{SV}}\rho'(\mathbf{r})\|_{L^2}}{\|\rho'(\mathbf{r})\|_{L^2}}, \quad (41b)$$

where P_{SV} is the density projection defined in equation (27). Here $\|\cdot\|_{L^2}$ denotes the L^2 norm in polar/spherical coordinates, that is

$$\|f\|_{L^2} = \sqrt{\int_0^{2\pi} \int_0^\infty f(r, \phi) f(r, \phi)^* r \, dr \, d\phi}$$

and

$$\|g\|_{L^2} = \sqrt{\int_0^{2\pi} \int_0^\pi \int_0^\infty g(r, \theta, \phi) g(r, \theta, \phi)^* r^2 \sin \theta \, dr \, d\theta \, d\phi},$$

for square integrable functions $f(r, \phi)$ and $g(r, \theta, \phi)$. Since, as previously mentioned, the single-particle scattered intensity is initially unknown in FXS and reconstructed during the phasing process, the metrics $E_{\text{reciprocal}}$ and E_{real} can only serve as convergence indicators and do not directly estimate the deviation of the current solution from experimental observables. For this reason, we also define metrics for determining the relative difference in the L^2 norm on the level of the invariants as

$$E_n = \frac{\int \int q q' |B_n(q, q') - B_n^l(q, q')|^2 \, dq \, dq'}{\int \int q q' |B_n(q, q')|^2 \, dq \, dq'}, \quad (42a)$$

$$E_l = \frac{\int \int q^2 (q')^2 |B_l(q, q') - B_l^l(q, q')|^2 \, dq \, dq'}{\int \int q^2 (q')^2 |B_l(q, q')|^2 \, dq \, dq'}, \quad (42b)$$

where $B_n(q, q')$ and $B_l(q, q')$ denote the input invariants employed as constraints, while $B_n^l(q, q')$ and $B_l^l(q, q')$ are the invariants calculated from the harmonic coefficients $I_n(q)$ and $I_l^l(q)$ corresponding to the current phasing loop iteration (see Fig. 3).

3.4. Alignment and averaging of reconstructions

The invariants B_n or B_l , employed as input data in the phasing process, do not contain (by definition) any information about the absolute position, orientation and point inversion of a particle in space. Therefore, individual reconstructions $\rho(\mathbf{r})$ initiated from a random density guess may vary in these properties. Similarly to conventional iterative phasing schemes, the MTIP algorithm may also produce nonunique solutions (Donatelli *et al.*, 2015). Therefore, it is customary to present the final solution as an average of selected and aligned individual reconstructions.

Combining the action of rotations on the intensity harmonic coefficients in two [equation (7)] and three dimensions [equation (12)], in their matrix form, with equation (15) allows one to examine their action on the level of the unknowns $\mathbf{u}_n$ and $\mathbf{U}_l$ via

$$\mathbf{I}_n \exp(in\varphi) = \tilde{\mathbf{v}}_n \mathbf{u}_n \exp(in\varphi) = \tilde{\mathbf{v}}_n \mathbf{u}'_n, \quad (43a)$$

$$\mathbf{I}^l \mathbf{D}^l(\alpha, \beta, \gamma) = \tilde{\mathbf{V}}_l \mathbf{U}_l \mathbf{D}^l(\alpha, \beta, \gamma) = \tilde{\mathbf{V}}_l \mathbf{U}'_l, \quad (43b)$$

where we interpret $\mathbf{D}^l(\alpha, \beta, \gamma)$ for each l as a $(2l+1) \times (2l+1)$ matrix. Since in the 2D case $\mathbf{u}_n$ is itself a phase factor, the rotational freedom in φ allows us to freely specify $\mathbf{u}'_n$ for a single chosen order n during the iterative phasing process. This condition causes the number of possible orientation states an individual reconstruction can attain to become finite. This, in turn, enables *a posteriori* algebraic orientation determination on the level of individual 2D reconstructions (see Section S2 of the supporting information).

In the 3D case, the restriction posed by equation (43b) is not strong enough to fix any of the unknown matrices $\mathbf{U}'_l$ during the reconstruction process. Therefore, orientational alignment of 3D reconstructions is performed after completing the iterative phasing as follows. First, all reconstructions are centered at their respective centers of density and a reference reconstruction ρ_{ref} is selected. All reconstructions are then orientationally aligned with respect to this reference using fast Fourier transforms on the special orthogonal group $\text{SO}(3)$ as described by Kostelec & Rockmore (2008). This procedure enables efficient calculations of the rotational cross-correlation $C(\mathcal{R})$ between the reference $\rho_{\text{ref}}(r, \theta, \phi)$ and any other reconstructed density $\rho(r, \theta, \phi)$, which is given by

$$C(\mathcal{R}) = \int_{r_{\min}}^{r_{\max}} dr \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\phi \, \rho_{\text{ref}}(r, \theta, \phi) \mathcal{R} \rho(r, \theta, \phi). \quad (44)$$

$\mathcal{R}$ is a rotation in $\text{SO}(3)$ and $\mathcal{R} \rho(r, \theta, \phi)$ is a rotated version of the reconstructed density $\rho(r, \theta, \phi)$.

The cross-correlation $C(\mathcal{R})$ is maximal at the rotation $\mathcal{R} = \mathcal{R}_{\text{opt}}$ for which the rotated density $\mathcal{R}_{\text{opt}} \rho$ optimally

matches the corresponding reference ρ_{ref} . To facilitate structure alignment it is helpful to limit the range ($r_{\text{min}}, r_{\text{max}}$) to regions of the reconstructed densities that are not spherically symmetric. In order to correct for a possible point inversion in the reconstructions, we apply this alignment procedure to each 3D reconstruction ρ , as well as its point-inverse ρ_{inv} , resulting in two aligned candidates per reconstruction. Finally, we determine the relative distance of the two candidates $\rho_{\text{rot}} \in \{\mathcal{R}_{\text{opt}}\rho, \mathcal{R}_{\text{opt}}\rho_{\text{inv}}\}$ from the reference density ρ_{ref} using the L^2 norm,

$$L(\rho_{\text{rot}}, \rho_{\text{ref}}) = \frac{\|\rho_{\text{rot}} - \rho_{\text{ref}}\|_{L^2}}{\|\rho_{\text{ref}}\|_{L^2}}, \quad (45)$$

and select the candidate ρ_{rot} with the lowest distance for subsequent averaging.

Note that, in the 2D case, after centering and aligning the reconstructions according to the procedure described in Section S2 of the supporting information, we also use equation (45) to correct for point inversion.

The presented algorithm allows one to select the reconstructions to be used in the final average on the basis of their error metrics (41) and (42), as well as their distance (45) from the reference structure. Finally a resolution estimate of the average can be computed using a generalized version of the phase retrieval transfer function (PRTF) (Kurta *et al.*, 2017),

$$\text{PRTF}(\mathbf{q}) = \frac{|\langle \text{FT}[\rho_i(\mathbf{r})] \rangle_i|}{|\langle \hat{\rho}'_i(\mathbf{q}) \rangle_i|}, \quad (46)$$

where $\langle \cdot \rangle_i$ denotes averaging over the selected collection of aligned reconstructions and $\text{FT}[\rho_i(\mathbf{r})]$ is the Fourier transform of the i th aligned electron density (see Fig. 3). If we assume that $|\hat{\rho}'_i(\mathbf{q})|$ are identical in all individual reconstructions, as is the case in conventional CXDI [where $|\hat{\rho}'_i(\mathbf{q})| = \sqrt{I(\mathbf{q})}$ and $I(\mathbf{q})$ is the experimentally determined intensity], expression (46) reduces to the conventional PRTF formula [see *e.g.* Chapman *et al.* (2006)].

4. *xFrame*: a Python implementation of the reconstruction workflow

The reconstruction workflow described in Section 3 is implemented in the open-source software suite *xFrame* available at <https://github.com/European-XFEL/xFrame>. The software consists of the back-end framework, which takes care of technical details unrelated to the reconstruction process (*e.g.* multiprocessing, GPU access, data storage, input settings *etc.*), and the *fxs* project which implements various routines of the reconstruction pipeline (calculations of the CCF, extraction of invariants, iterative phasing, averaging of reconstructions).

4.1. Dependencies

Table 1 lists *xFrame* dependencies and their usage. For computationally expensive operations such as the Fourier and harmonic transforms, we use existing software that references to C or Fortran routines wherever possible. In all other cases we rely on *numpy* vectorization and GPU acceleration using

Table 1

List of dependencies of *xFrame*.

Package	Usage
<i>numpy</i> (Harris <i>et al.</i> , 2020)	All parts of <i>xFrame</i>
<i>scipy</i> (Virtanen <i>et al.</i> , 2020)	Invariants extraction
<i>pyOpenCL</i>	GPU access
<i>shms</i> (Schaeffer, 2013)	Spherical harmonic transforms
<i>pysofft</i> (SOFT) (Kostelec & Rockmore, 2008)	3D alignment of reconstructions
<i>matplotlib</i> (Hunter, 2007)	2D plots
<i>openCV</i> (Bradski, 2000)	2D plots
<i>vtk</i> (Schroeder <i>et al.</i> , 2006)	3D plots
<i>h5py</i> (The HDF Group <i>et al.</i> , 2020)	Data storage
<i>ruamel.yaml</i>	Software settings
<i>click</i>	Command-line interface
<i>psutil</i>	Hardware information

OpenCL. Although *xFrame* depends exclusively on cross-platform packages, it has currently only been tested on Linux-based operating systems.

4.2. Input/output data formats

xFrame requires input data in the form of a set of diffraction patterns in binary format or a statistically averaged two-point CCF $C_M(q, q', \Delta)$ in HDF5 format. Human-readable YAML files are used to specify the input settings for different *xFrame* routines. The output data produced by *xFrame* are stored in four standard formats, which are HDF5, YAML, VTK and PNG. The HDF5 format is used for general-purpose data storage, *e.g.* to save calculated metrics and reconstruction results, and the output YAML files are used to store the input settings associated with a particular reconstruction. Finally, VTK files and PNG images target visualization of reconstruction results. Specifically, the open-source VTK file format allows one to examine the reconstructed densities on their native spherical or polar grid without any further post-processing.

4.3. *xFrame* usage

A typical reconstruction pipeline using the command-line interface of *xFrame* is shown in Fig. 4. It is possible to enter the

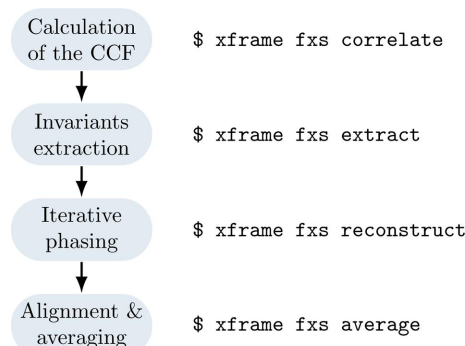


Figure 4

Diagram of a typical reconstruction workflow (left) using the command-line tools (right) of *xFrame*. Each of the *xFrame* commands takes as an argument a human-readable settings file that specifies all relevant options.

workflow at different points, by running `xframe fxs correlate` to compute the CCF (16) from a set of input diffraction patterns, or extracting the rotational invariants B_n or B_l from the two-point CCF (by running `xframe fxs extract` with a specified input CCF in HDF5 format), or directly running reconstructions using the extracted invariants (`xframe fxs reconstruct`), which can then be aligned and averaged in a final step using `xframe fxs average`. Apart from the command-line interface it is also possible to use *xFrame* directly as a Python module. Details on the installation process as well as tutorials can be found at <https://xframe-fxs.readthedocs.io>.

The iterative phasing process is implemented in a way that allows one to run a single reconstruction per available CPU

core using the Python *multiprocessing* module, while at the same time access to the GPU resources is shared among all parallel reconstructions (see Section S1 in the supporting information).

5. Reconstructions from simulated data using *xFrame*

Here we demonstrate single-particle structure recovery from simulated FXS data using *xFrame*. The scattering intensities (1) were simulated assuming ideal kinematic X-ray scattering without noise. We considered dilute limit approximation, where inter-particle interference can be neglected, and simulated 10^5 diffraction patterns for each of the considered model structures presented in Fig. 5. Diffraction patterns were computed up to a maximum momentum transfer

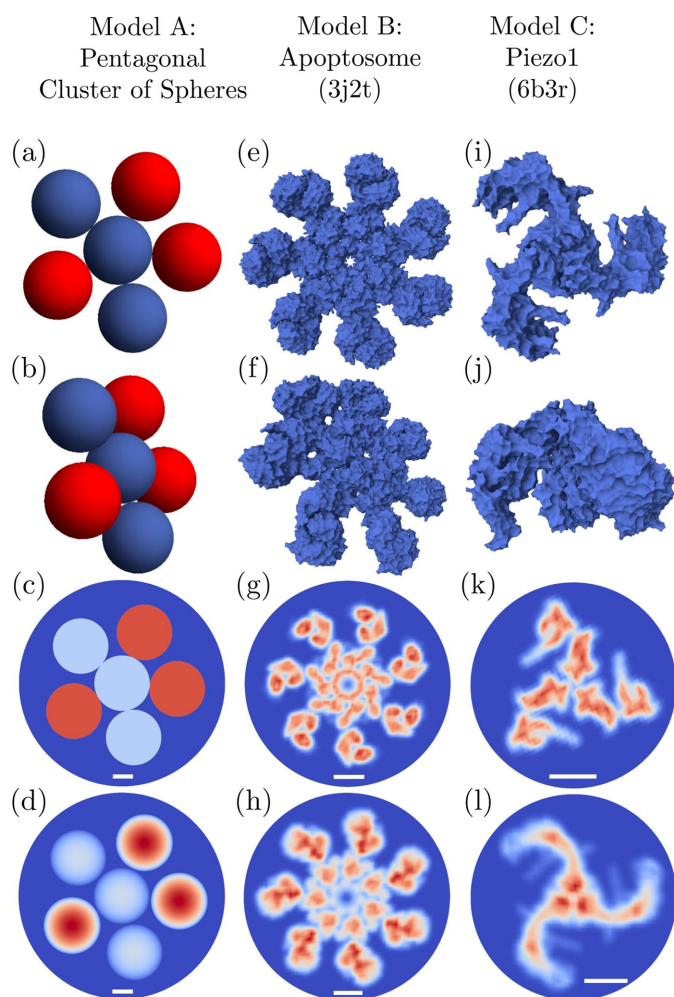


Figure 5

Three model structures (models A–C) considered for reconstructions using *xFrame*: (a)–(d) a pentagonal cluster consisting of spheres of uniform density with a diameter of 140 nm, with the red spheres being of doubled density as compared with the blue spheres; (e)–(h) the human apoptosome (PDB entry 3j2t; Yuan *et al.*, 2013); (i)–(l) the mechano-sensitive ion channel Piezo1 (PDB entry 6b3r; Guo & MacKinnon, 2017). The two upper rows show distinct views of the 3D structures. The third row displays 2D slices through the centers of the respective 3D models, and the bottom row displays 2D projections on the image plane, which were produced using the electron-density maps generated in *UCSF Chimera* (Pettersen *et al.*, 2004) for the corresponding 3D models. The white scale bars shown in the two bottom rows correspond to 5 nm.

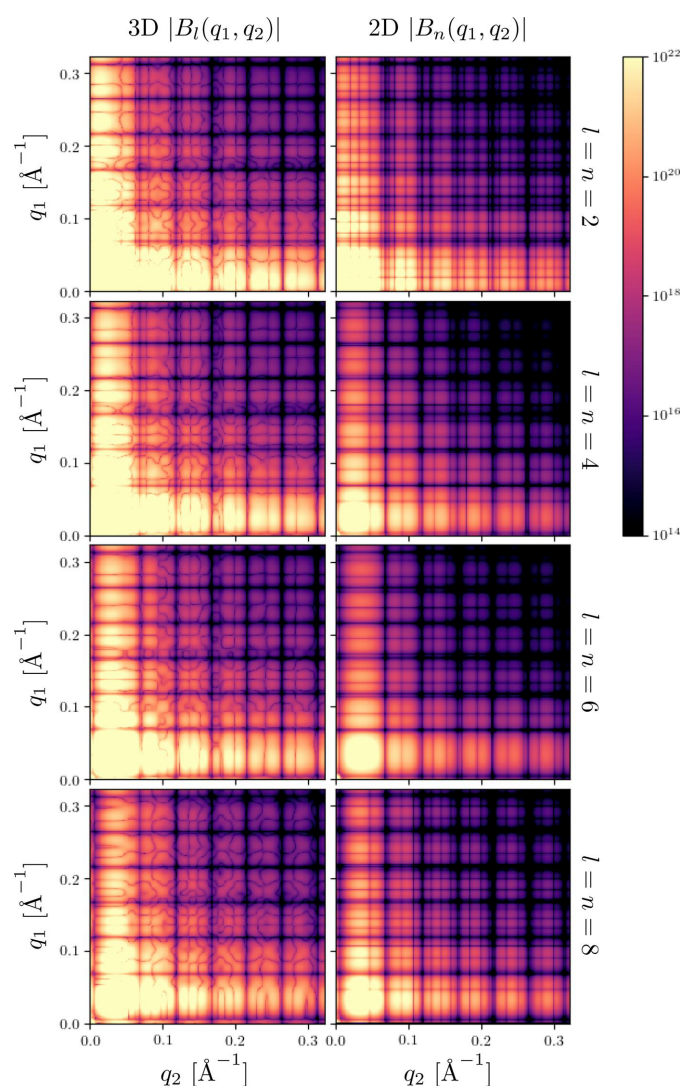


Figure 6

Absolute values of the rotational invariants $|B_n(q_1, q_2)|$ and $|B_l(q_1, q_2)|$ of orders $n, l = 2, 4, 6$ and 8 , determined for model A in the 2D case (right) and 3D case (left). The invariants $B_n(q_1, q_2)$ show features in the form of straight lines, which is a direct consequence of the fact that each B_n is a matrix of rank 1. The invariants $B_l(q_1, q_2)$ display more complex features since the respective matrix B_l can have a rank higher than 1 (see Section 2.3).

$Q_{\max} = 0.32 \text{ \AA}^{-1}$ for model A, and up to $Q_{\max} = 0.42 \text{ \AA}^{-1}$ for models B and C. The diffraction patterns were then used to determine the averaged CCF (16), and subsequently the invariants B_n and B_L .

We first consider the reconstruction results for the single-particle case ($N_p = 1$), where the input set of diffraction patterns was simulated from single particles in random orientations in two or three dimensions. For instance, the invariants extracted [by solving equation (23)] from the 2D and 3D FXS data sets simulated for a single pentagonal cluster of spheres (model A) are shown in Fig. 6.

The complete iterative phasing process in *xFrame* is divided into the main and refinement stages, where the electron density with the lowest error metric obtained during the main phasing stage is further optimized in the refinement stage. *xFrame* allows one to separately set up the number, sequence and parameters of ER, HIO and SW procedures (see Section

3.3.1) in the main and refinement stages. For the 3D reconstructions shown in Fig. 7 the main stage consisted of blocks of $60 \times$ HIO, followed by $1 \times$ SW and $40 \times$ ER steps. The number of used iteration blocks varied from five in the case of model A to 30 for model B and model C. The 3D refinement part consisted of a single block of $1 \times$ SW followed by $200 \times$ ER steps for all considered models.

The obtained 2D reconstructions [Figs. 7(d), 7(h) and 7(l)] were produced using ten main stage iteration blocks consisting of $500 \times$ HIO followed by $1 \times$ SW and $200 \times$ ER steps, while the refinement part consisted of $1 \times$ SW step followed by $200 \times$ ER iterations. The HIO parameter β was determined in the i th iteration as $\beta(i) = a \exp(bi) + c$, with parameters a , b and c chosen in such a way that $\beta(i)$ was exponentially decreasing during the reconstruction process from 0.5 down to 0.14 for the 3D reconstructions, and from 0.1 down to 0.01 for the 2D reconstructions. The SW threshold value γ [equation (25)] was set to 0.11, and the standard deviation σ was linearly decreasing from 2 nm to 1.5 nm during the reconstruction for

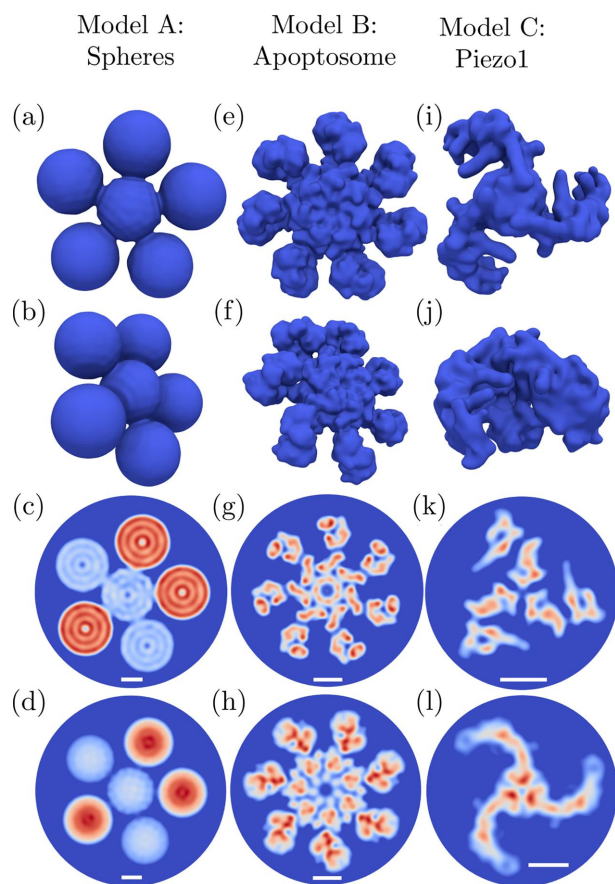


Figure 7

Averaged 2D and 3D reconstructions obtained using *xFrame* for three different model structures shown in Fig. 5. The two upper rows show two distinct views of the reconstructed 3D structures, and the third row displays 2D slices through the centers of the respective 3D reconstructions. The bottom row displays averaged 2D reconstructions, which correspond to the 2D projections shown in the bottom row in Fig. 5. The first two rows show isosurfaces at 15% of the maximal reconstructed electron-density value, and the last two rows display density values higher than 15% of the maximal density value. The 2D slices in (c), (g) and (k) are taken at approximately the same regions of the electron density as given for the model structures in (c), (g) and (k) of Fig. 5, respectively. The white scale bars shown in the two bottom rows correspond to 5 nm.

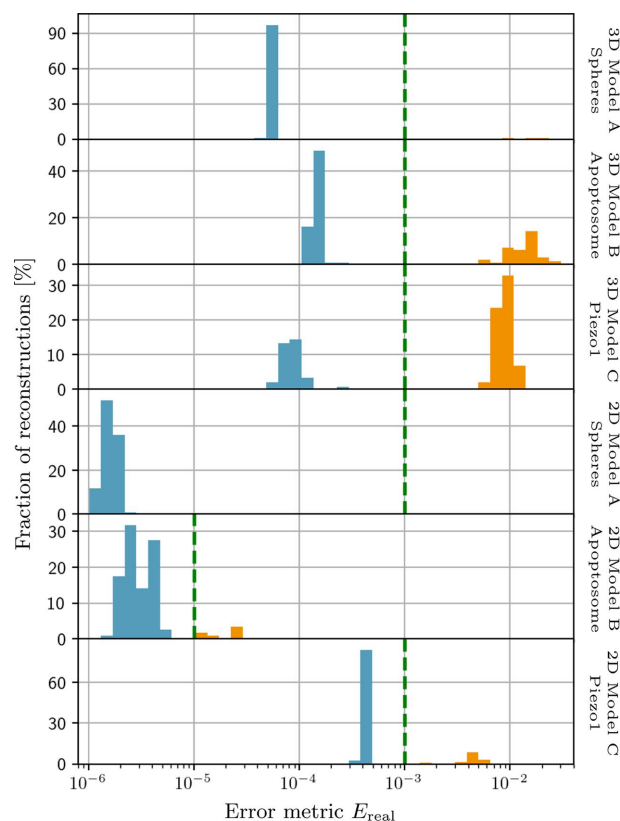


Figure 8

Normalized histograms of the final error metric values E_{real} [equation (41b)] shown for all individual 2D and 3D reconstructions listed in Table 2. For most of the models the reconstructions cluster into two groups separated by at least half an order of magnitude in their final error value, which allows one to identify converged reconstructions by introducing a threshold. The thresholds for each model are signified by the dashed green lines, placed at 10^{-5} for the 2D model B, and at 10^{-3} for all other models. Thus, the parts of the histograms shown in light-blue and orange colors correspond to converged and not converged reconstructions, respectively. For the 2D model A all computed reconstructions have approximately similar values of E_{real} and were considered to be converged. The 3D model A has a total of three reconstructions with error values around 10^{-2} that did not converge.

Table 2
Reconstruction statistics using *xFrame*.

Model	Reconstructions performed	Reconstructions converged (% of total)
3D model A	113	110 (97)
3D model B	168	111 (66)
3D model C	340	116 (34)
2D model A	120	120 (100)
2D model B	120	113 (94)
2D model C	120	103 (86)

all models. The density value projection [equations (26)] was parameterized as $V_{\text{Re}}^{\text{min}} = 0$, $V_{\text{Re}}^{\text{max}} = \infty$, $V_{\text{Im}}^{\text{min}} = -2$ and $V_{\text{Im}}^{\text{max}} = 2$. We empirically found that allowing small nonzero values of $V_{\text{Im}}^{\text{min}}$ and $V_{\text{Im}}^{\text{max}}$ results in improved convergence rates of reconstructions. The Fourier transform stabilization has been applied as described in Section 3.3.4. The presented 3D reconstructions were obtained considering spherical harmonic expansion orders up to $l_{\text{max}} = 127$, while the 2D reconstructions used circular harmonic orders up to $n_{\text{max}} = 255$. The reciprocal projection used invariants B_n and B_l up to the same maximal orders while setting all odd-order invariants to 0. The considered number of radial steps was $N = 256$ for all models, and the angular sampling was chosen such that the maximal harmonic order could be resolved (see Appendix A). To follow the reconstruction progress we used the error metric E_{real} defined in equation (41b).

Individual reconstructions were classified as ‘converged’ or ‘not converged’ according to the histograms of the final values

of the error metric E_{real} (see Fig. 8 and Table 2). Note the different convergence rates for different structures in Table 2. The reconstructed structures shown in Fig. 7 were obtained by aligning and averaging 100 converged reconstructions for each model using *xframe* `fxs average`. The corresponding PRTF curves [equation (46)] computed for the single-particle reconstructions are shown in Fig. 9, indicating that the resolution is Fourier limited, *i.e.* limited by the extent of the simulated FXS data in reciprocal space. The impact of the maximum considered spherical harmonic expansion order l_{max} on the resolution of the obtained 3D reconstructions is illustrated in Fig. S4 of the supporting information.

The ability to perform X-ray scattering measurements from just individual particles ($N_p = 1$) in solution at near-physiological conditions represents an ideal scenario for FXS analysis, although it might be challenging to achieve in practice for weakly scattering bioparticles. The invariant-based FXS approach offers the possibility to perform reconstructions based on multiparticle ($N_p > 1$) X-ray scattering (see Sections 2.1 and 2.2), in which the total scattering from bioparticles is enhanced compared with single-particle measurements. According to equations (6) and (11), a scaled version of the single-particle invariants can be extracted from such multiparticle FXS data. Under the assumption that N_p is known, the invariants extracted from the multiparticle scattering data can be normalized, *i.e.* the zero harmonic order by N_p^2 and all higher orders by N_p , and used to perform single-particle structure recovery as described above. Reconstruction results for the multiparticle scattering case ($N_p = 10$) presented in Fig. S5 of the supporting information look very similar to those obtained in the single-particle case (Fig. 7). The sensitivity of the reconstruction results to the accuracy of the determined scaling factors N_p is demonstrated in Fig. S6 of the supporting information.

6. Summary and conclusions

In this work, we presented a workflow for single-particle structure determination from FXS measurements. The workflow consists of several steps, including calculation of the average two-point CCF from a set of diffraction patterns, extraction of rotational invariants from the CCF, iterative phasing of rotational invariants using the MTIP algorithm, and selection, alignment and averaging of individual reconstructions. We proposed a new method for extracting rotational invariants from the angular Fourier spectra of the CCF (see Section 3.2). We also introduced several modifications to the original version of the MTIP algorithm published by Donatelli *et al.* (2015), including discrete versions of the Hankel transform (Section 3.3.3) and additional measures to improve phasing stability (Sections 3.3.4 and 3.4).

We considered different approximations of the Hankel transform (see Appendix B) using orthogonal basis expansions, including formulations based on the cosine/sine series expansion (Appendix B2) and Zernike polynomial expansion (Appendix B3), as well as direct approximation of Hankel integrals with Riemann sums using the midpoint rule

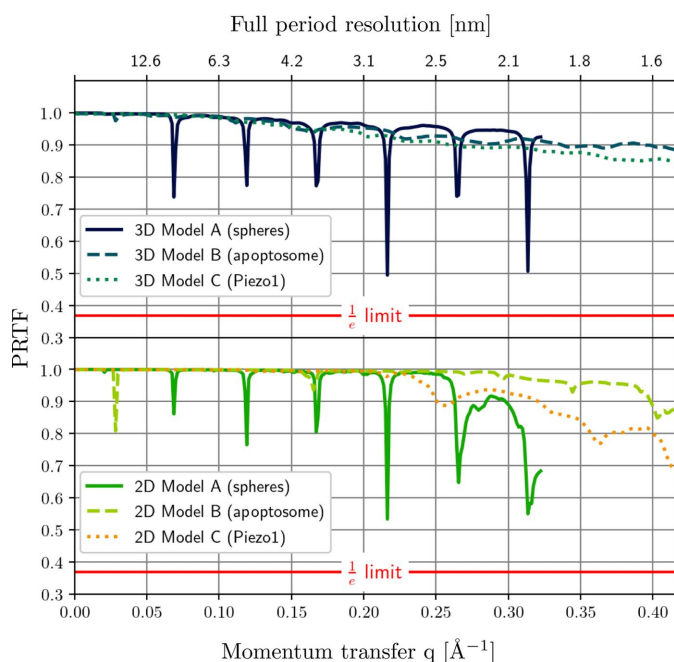


Figure 9
PRTF curves determined for the averaged 2D (bottom) and 3D (top) reconstructions shown in Fig. 7. The red lines define a cutoff value of $1/e$, used to estimate the reconstruction resolution by PRTF. Since all PRTF curves are above this threshold, the resolution of the reconstructed structures is limited by the momentum transfer cutoffs in the respective input data, which are $0.32 \text{ \AA}^{-1}$ for model A and $0.42 \text{ \AA}^{-1}$ for models B and C.

(Appendix B4). As a byproduct, we derived a closed-form expression for the Hankel transform of the radial part of the 3D Zernike polynomials (Appendix B5). Our results show that the Hankel transform, defined via orthogonal basis expansion of density, produces results that converge to the results of direct approximation of the continuous Hankel integral with a Riemann sum (Appendix B1).

The proposed workflow has been implemented in the open-source software suite *xFrame*. *xFrame* features a multi-processing scheme that allows for parallel reconstruction runs on a multi-core CPU while at the same time enabling GPU acceleration of time-consuming steps within each of the parallel reconstructions (see Section S1 in the supporting information). Successful reconstructions can be identified using distinct error metrics (Section 3.3.5), and subsequently aligned and averaged (Section 3.4). In the 2D case, the alignment routine is partially incorporated into the MTIP loop and completed *a posteriori* (see Section 3.4 and Section S2 of the supporting information). We demonstrated the functionality of *xFrame* by performing 2D and 3D reconstructions from simulated single-particle scattering data for several structures (Section 5). The results show successful *ab initio* recovery of the particle shapes and internal density distribution without the need to apply symmetry constraints. Reconstructions from multiparticle scattering data are also possible, while the feasibility of such reconstructions relies on accurate knowledge of the number of particles N_p contributing to X-ray snapshots (see Section S3 of the supporting information).

FXS was originally proposed for single-particle structure determination from multiparticle solution X-ray scattering; thus, it complements traditional SAXS and SPI techniques. The inverse problem in FXS is solved using iterative phase retrieval, in which two phase problems are tackled simultaneously. The first involves finding unknown unitary matrices to determine the single-particle scattered intensity, which is equivalent to solving the orientation determination problem in conventional SPI. The second is related to finding optimal phases of the single-particle scattering amplitudes, similar to conventional iterative phase retrieval used in SPI or CXDI. At the same time, FXS may naturally expand the information content of traditional solution SAXS experiments if X-ray scattering measurements are performed on timescales faster than the rotation diffusion time of particles in solution. Such multiparticle measurements, however, require very precise detector corrections to be able to detect weak intensity fluctuations about orientationally averaged SAXS (see Section S4 of the supporting information for a brief summary of challenges related to experimental measurements and data processing). If such requirements can be fulfilled, FXS may potentially close the gap between conventional SPI, SAXS and crystallographic structure determination, particularly in time-resolved studies with an XFEL (Kurta *et al.*, 2023).

Although the single-particle structure reconstruction workflow is presented here in the context of biological applications, it may also serve as an alternative way for 2D and 3D structure determination of arbitrary molecules, nanoparticles, engineered nanostructures *etc.*, provided that FXS

data of sufficient quality can be measured. We hope that the presented open-source software *xFrame* can facilitate efforts in this direction.

APPENDIX A

Discrete polar and spherical grids

In practical applications data acquisition and processing are realized on a finite discrete grid, and hence our reconstruction workflow is implemented on a discrete polar/spherical grid. We define the extent and sampling of the radial coordinates in real and reciprocal space as

$$r_p = (p + \frac{1}{2})R_{\max}/N, \quad q_k = (k + \frac{1}{2})Q_{\max}/N, \quad (47)$$

$$p, k = 0, \dots, N - 1,$$

where the sets r_p and q_k of N radial points sample real and reciprocal space (in spherical or polar coordinates) up to a maximum extent $R_{\max}$ and $Q_{\max}$, respectively. In practice, $Q_{\max}$ is usually determined by the maximum extent of the measured scattered intensity on a detector (or quality of the CCF), while the real-space cutoff $R_{\max}$ can be justified by the finite dimensions of the particle, so that the particle density centered at the origin of the coordinate system satisfies the condition $\rho(R_{\max}) = 0$. The chosen discretization is compatible with the midpoint rule applied in this work for approximating integrals with Riemann sums (Hughes-Hallett *et al.*, 2012). We use the reciprocity relation $R_{\max}Q_{\max} = \pi N$ in computations of the discrete Fourier transforms.

As for the angular coordinates, in the 2D polar grid we consider the angle ϕ to be uniformly sampled from 0 to 2π . In the case of the 3D spherical grid, this also applies to the azimuthal angle ϕ , while the polar angle θ is sampled from 0 to π on Gauss–Legendre nodes, *i.e.* $\cos(\theta_i^L) = x_i^L$, where x_i^L is the i th zero of the Legendre polynomial of the order $L = l_{\max}$, and $l_{\max}$ is the maximal spherical harmonic order considered in the expansion (8a).

APPENDIX B

Hankel transform approximations

In this appendix we consider different methods for approximating the continuous Hankel transforms in equations (36) and (37).

B1. Expansion via orthogonal polynomials

One of the approaches to derive a discrete version of the Hankel transforms in equations (36) and (37) is to expand $\rho_m(r)$ or $\rho_m^l(r)$ using an orthogonal basis $\xi_i(r)$ of all square integrable functions on the interval $[0, R_{\max}]$, *e.g.* in the 2D case

$$\rho_{m,i} = \int_0^{R_{\max}} \rho_m(r) \xi_i^*(r) dr, \quad (48a)$$

$$\rho_m(r) = \sum_{i=0}^{\infty} \rho_{m,i} \xi_i(r), \quad (48b)$$

where $\rho_{m,i}$ are the expansion coefficients. This allows one to shift the Hankel integration from $\rho_m(r)$ to the expansion functions $\xi_i(r)$, i.e. using equation (48b) in (36a) we get

$$\hat{\rho}_m(q) = (-i)^m \sum_{i=0}^{\infty} \rho_{m,i} \int_0^{R_{\max}} \xi_i(r) J_m(qr) r dr. \quad (49)$$

The integral in equation (49) can be precomputed once independently of the considered function $\rho_m(r)$, and then used for all subsequent Hankel transform computations. The remaining integral in (48a) in the determination of the expansion coefficients $\rho_{m,i}$ can then be approximated by using one of the available methods, e.g. the trapezoidal rule (Donatelli *et al.*, 2015) or midpoint rule (present work), to derive a discrete form of equations (36) and (37). In the present case, equation (49) can be approximated in the following general form,

$$\hat{\rho}_m(q_k) \simeq A_m \sum_p \rho_m(r_p) \omega_m(k, p), \quad (50)$$

where A_m are constants, $\omega_m(k, q)$ are quadrature weights defined by the integral in equation (49), and r_p and q_k are discrete coordinates in real and reciprocal space [see equation (47)]. Such an approach has been implemented by Donatelli *et al.* (2015) using a cosine/sine series expansion. An example of the cosine/sine approach, using the midpoint rule to approximate the expansion coefficients, is presented in Appendix B2. In this work we also developed an approximation of the Hankel transforms by employing Zernike polynomials as the orthogonal basis functions $\xi_i(r)$. The advantage of the obtained expressions is that, in this case, the integral in (49) can be evaluated analytically (see Appendix B3) and thus does not require numerical approximation as in the case of a cosine/sine series expansion (see Appendix B2).

At the same time, analysis of expression (49) defined in terms of an arbitrary orthogonal basis $\xi_i(r)$ reveals further aspects of such an expansion. By applying the midpoint rule to approximate the integral in (48b) on a discrete grid defined in Appendix A, and substituting the result into (49), we obtain

$$\hat{\rho}_m(q) \simeq (-i)^m \sum_{i=0}^{\infty} \frac{R_{\max}}{N} \sum_{p=0}^{N-1} \rho_m(r_p) \xi_i^*(r_p) \int_0^{R_{\max}} \tilde{J}_m(qr) \xi_i(r) dr, \quad (51)$$

where $\tilde{J}_m(qr) = J_m(qr)r$. Considering that $\tilde{J}_m(qr)$ is a real function, the integral in (51) defines, in fact, the complex conjugated coefficients $\tilde{J}_{m,i}^*(q)$ of expansion of $\tilde{J}_m(qr)$ in the basis $\xi_i(r)$, that is [see equation (48a)]

$$\tilde{J}_{m,i}^*(q) = \left[\int_0^{R_{\max}} \tilde{J}_m(qr) \xi_i^*(r) dr \right]^*. \quad (52)$$

Using the latter result in (51) and rearranging the terms we get

$$\hat{\rho}_m(q) \simeq (-i)^m \frac{R_{\max}}{N} \sum_{p=0}^{N-1} \rho_m(r_p) \left[\sum_{i=0}^{\infty} \tilde{J}_{m,i}(q) \xi_i(r_p) \right]^*. \quad (53)$$

One may recognize that the expression in square brackets is exactly the series expansion of $J_m(qr_p)$ in terms of $\xi_i(r_p)$ [see

equation (48b)]. In the limit of infinite expansion orders ($i \rightarrow \infty$) such an approximation scheme is, therefore, independent of the chosen orthogonal basis ξ_i and results in the following expression:

$$\hat{\rho}_m(q_k) \simeq (-i)^m \frac{R_{\max}}{N} \sum_{p=0}^{N-1} \rho_m(r_p) J_m(q_k r_p) r_p. \quad (54)$$

Note that equation (54) corresponds exactly to simple numeric evaluation of the continuous Hankel transform using a Riemann sum in the integration bounds $[0, R_{\max}]$ [see equation (69) in Appendix B4]. Similar derivations can be performed for the inverse Hankel transform, and also in the 3D case. This conclusion is supported numerically for the 3D case in Fig. S7 of the supporting information, where it is shown that the Zernike [equation (68)] and cosine/sine [equations (56)] weights approach the weights obtained by directly approximating the Hankel integrals using the midpoint rule [equation (39c)]. In our reconstruction workflow, we therefore employ the quadrature weights from the midpoint formulation (see Appendix B4) as a default option, while the Zernike approximation is available for testing.

B2. Cosine/sine series expansion approximation

The quadrature weights given by Donatelli *et al.* (2015) are determined by employing a cosine/sine series expansion of $\rho_m(r)$ or $\rho_m^l(r)$, and their reciprocal-space counterparts. Using the definitions in Appendix A and applying the midpoint rule for approximating the integrals, the weights for the 2D case take the following form for odd orders m ,

$$\omega_m(p, k) = 2\pi^2 N \sum_{s=0}^{s_{\max}} \sin \left[\frac{\pi s(1+2p)}{2N} \right] \times \int_0^1 \sin(\pi s x) J_m \left[\frac{\pi(1+2k)}{2} x \right] x dx, \quad (55a)$$

and for even orders m ,

$$\omega_m(p, k) = 2\pi^2 N \sum_{s=0}^{s_{\max}} c_s \cos \left[\frac{\pi s(1+2p)}{2N} \right] \times \int_0^1 \cos(\pi s x) J_m \left[\frac{\pi(1+2k)}{2} x \right] x dx, \quad (55b)$$

where $s_{\max}$ defines the maximum expansion order in the cosine/sine series, $c_s = 1/2$ for $s = 0$, and $c_s = 1$ otherwise.

In the 3D case the weights for odd orders l are

$$\omega_l(p, k) = 2\sqrt{2}\pi^2 \sqrt{\pi} N^2 \sum_{s=0}^{s_{\max}} \sin \left[\frac{\pi s(1+2p)}{2N} \right] \times \int_0^1 \sin(\pi s x) j_l \left[\frac{\pi(1+2k)}{2} x \right] x^2 dx, \quad (56a)$$

and for even orders l ,

$$\omega_l(p, k) = 2\sqrt{2}\pi^2\sqrt{\pi}N^2 \sum_{s=0}^{s_{\max}} c_s \cos\left[\frac{\pi s(1+2p)}{2N}\right] \times \int_0^1 \cos(\pi s x) j_l\left[\frac{\pi(1+2k)}{2}x\right] x^2 dx. \quad (56b)$$

The quadrature weights (55) and (56) can be directly used in forward transforms (38a) and (39a), respectively, while for the inverse transforms (38b) and (39b) the corresponding weight functions should be transposed with respect to p and k .

B3. Zernike polynomial expansion approximation

Here we derive the quadrature weights $\omega_m(p, k)$ and $\omega_l(p, k)$ by employing the radial parts of the Zernike polynomials (Zernike, 1934; Born & Wolf, 2019) as the basis functions $\xi_i(r)$ in the expansion (48).

The radial parts $R_{D,s}^h(r)$ of the D -dimensional Zernike polynomials (Zernike & Brinkman, 1935) can be defined using Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ as

$$R_{D,s}^h(r) = (-1)^{(s-h)/2} r^h P_n^{(\alpha,\beta)}(1-2r^2), \quad (57)$$

$$\alpha = h + \frac{D}{2} - 1, \quad \beta = 0, \quad n = \frac{s-h}{2}$$

for even $s-h$, and $R_{D,s}^h(r) = 0$ otherwise, where s and h are non-negative integers.

The radial polynomials $R_{D,s}^h(r)$ form a set of orthogonal polynomials on the interval $[0, 1]$, with the orthogonality condition given for an arbitrary h by

$$\int_0^1 R_{D,s}^h(r) R_{D,s'}^h(r) r^{D-1} dr = \frac{\delta_{s,s'}}{2s+D}. \quad (58)$$

This implies that any sufficiently smooth function $f(r)$ that is defined on a finite interval $[0, R_{\max}]$ has a series expansion in the polynomials $R_{D,s}^h(r)$, i.e.

$$f(r) = \sum_{s=0}^{\infty} f_s R_{D,s}^h\left(\frac{r}{R_{\max}}\right), \quad (59a)$$

$$f_s = \frac{2s+D}{R_{\max}^D} \int_0^{R_{\max}} f(r) R_{D,s}^h\left(\frac{r}{R_{\max}}\right) r^{D-1} dr. \quad (59b)$$

The advantage of Zernike expansions is that the Hankel transform of $R_{D,s}^h$ can be evaluated exactly. In the 2D case ($D = 2$) for $h = m$ one finds

$$\int_0^1 R_{2,s}^m(r) J_m(qr) r dr = (-1)^{(s-m)/2} \frac{J_{s+1}(q)}{q}, \quad (60)$$

whereas in the 3D case ($D = 3$) for $h = l$ we have

$$\int_0^1 R_{3,s}^l(r) j_l(qr) r^2 dr = (-1)^{(s-l)/2} \frac{j_{s+1}(q)}{q}, \quad (61)$$

for $q \neq 0$. A proof of equation (60) can be found in Appendix VII of Born & Wolf (2019), and equation (61) is derived in this work in Appendix B5. Since we are interested in approximating the Hankel transforms of a function defined on a finite interval $[0, R_{\max}]$, for an arbitrary positive $R_{\max}$, we use scaled versions of the integrals (60) and (61),

$$\int_0^{R_{\max}} R_{2,s}^m\left(\frac{r}{R_{\max}}\right) J_m(qr) r dr = (-1)^{(s-m)/2} \frac{J_{s+1}(qR_{\max})}{q} R_{\max} \quad (62)$$

and

$$\int_0^{R_{\max}} R_{3,s}^l\left(\frac{r}{R_{\max}}\right) j_l(qr) r^2 dr = (-1)^{(s-l)/2} \frac{j_{s+1}(qR_{\max})}{q} R_{\max}^2. \quad (63)$$

Let us now consider the Hankel transform for the 2D case, as specified in equation (49), and obtain its discrete version in the form of equation (50). By considering the expansion (59) of $f(r) = \rho_m(r)$ up to the maximum order $s = s_{\max}$, with each expansion coefficient $\rho_{m,s}$ approximated using the midpoint rule, we can write

$$\begin{aligned} \hat{\rho}_m(q) &= (-i)^m \int_0^{R_{\max}} \sum_{s=0}^{\infty} \rho_{m,s} R_{2,s}^m\left(\frac{r}{R_{\max}}\right) J_m(qr) r dr \\ &\simeq (-i)^m \sum_{s=0}^{s_{\max}} \frac{2s+2}{NR_{\max}} \sum_{p=0}^{N-1} \rho_m(r_p) R_{2,s}^m\left(\frac{r_p}{R_{\max}}\right) r_p \\ &\quad \times \int_0^{R_{\max}} R_{2,s}^m\left(\frac{r}{R_{\max}}\right) J_m(qr) r dr. \end{aligned} \quad (64)$$

Since $\rho_m(r)$ is finitely supported, the integration range reduces from $[0, \infty)$ to $[0, R_{\max}]$. Note that equation (64) is already represented as a weighted sum with weights that are independent of the function $\rho_m(r)$. Using the integral relation from equation (62), we arrive at (for $q \neq 0$)

$$\begin{aligned} \hat{\rho}_m(q) &\simeq \frac{(-i)^m}{N} \sum_{p=0}^{N-1} \rho_m(r_p) \frac{r_p}{q} \sum_{s=0}^{s_{\max}} (-1)^{(s-m)/2} (2s+2) \\ &\quad \times R_{2,s}^m\left(\frac{r_p}{R_{\max}}\right) J_{s+1}(qR_{\max}). \end{aligned} \quad (65)$$

Considering the definitions in equation (47), we finally obtain the discrete version of the Hankel transform in the 2D case in the form of (50),

$$\hat{\rho}_m(q_k) \simeq \frac{(-i)^m}{Q_{\max}^2} \sum_{p=0}^{N-1} \rho_m(r_p) \omega_m(p, k), \quad (66)$$

with the quadrature weights determined as

$$\begin{aligned} \omega_m(p, k) &= \frac{\pi(1+2p)}{(1+2k)} \sum_{s=0}^{s_{\max}} (-1)^{(s-m)/2} (2s+2) \\ &\quad \times R_{2,s}^m\left(\frac{1+2p}{2N}\right) J_{s+1}\left[\frac{\pi(1+2k)}{2}\right]. \end{aligned} \quad (67)$$

Following a similar procedure, the Zernike weights for the 3D case can be specified as

$$\omega_l(p, k) = \frac{\pi\sqrt{2\pi}(1+2p)^2}{2(1+2k)} \sum_{s=0}^{s_{\max}} (-1)^{(s-l)/2} (2s+3) \times R_{3,s}^l \left(\frac{1+2p}{2N} \right) j_{s+1} \left[\frac{\pi(1+2k)}{2} \right]. \quad (68)$$

The quadrature weights (67) and (68) can be directly used in the forward transforms (38a) and (39a), respectively, while for the inverse transforms (38b) and (39b) the corresponding weights are obtained by transposing $\omega_m(p, k)$ and $\omega_l(p, k)$ with respect to p and k .

B4. Direct approximation of the Hankel integrals using the midpoint rule

Here we consider direct approximations of the Hankel transforms given in equations (36) and (37) with Riemann sums.

In the 2D case the integral in (36a) can be approximated using the midpoint rule as

$$\hat{\rho}_m(q_k) \simeq (-i)^m \frac{R_{\max}}{N} \sum_{p=0}^{N-1} \rho_m(r_p) J_m(q_k r_p) r_p. \quad (69)$$

Using the definitions in Appendix A, we can present equation (69) in the form

$$\hat{\rho}_m(q_k) \simeq A_m \sum_{p=0}^{N-1} \rho_m(r_p) \omega_m(p, k), \quad (70)$$

where $A_m = (-i)^m / Q_{\max}^2$ and the quadrature weights $\omega_m(p, k)$ are given in equation (38c).

Similarly, in the 3D case the integral in (37a) can be approximated as

$$\hat{\rho}_l(q_k) \simeq A_l \sum_{p=0}^{N-1} \rho_l(r_p) \omega_l(p, k), \quad (71)$$

where $A_l = (-i)^l / Q_{\max}^3$ and the quadrature weights $\omega_l(p, k)$ are given in equation (39c). Approximations for the inverse Hankel transforms (36b) and (37b) can be obtained in a similar way, producing the discrete forward and inverse Hankel transforms given in equations (38) and (39), respectively.

B5. Hankel transform of the radial part of the 3D Zernike polynomial

Here we provide a proof of equation (61) for the 3D Zernike polynomials, which closely follows the derivation of equation (60) given by Born & Wolf (2019). We shall need the Rodrigues' formula for Jacobi polynomials $P_n^{(\alpha, \beta)}(z)$ and the series expansion of spherical Bessel functions $j_l(r)$, which can be obtained from 10.2.2, 10.47.3 and 18.5(ii) of NIST (2023):

$$j_l(z) = \sqrt{\frac{\pi}{2z}} \sum_{p=0}^{\infty} \frac{(-1)^p}{p! \Gamma(p+l+\frac{3}{2})} \left(\frac{z}{2} \right)^{2p+l+1/2}, \quad (72)$$

$$P_n^{(\alpha, \beta)}(z) = \frac{(-1)^n}{2^n n!} (1-z)^{-\alpha} (1+z)^{-\beta} \times \left(\frac{d}{dz} \right)^n (1-z)^{\alpha+n} (1+z)^{\beta+n}, \quad (73)$$

where Γ denotes the gamma function and $n!$ is the factorial of a non-negative integer number n .

By substituting $z = 1 - 2r^2$ in equation (73) and considering in equation (57) Jacobi polynomials for $\beta = 0$, the radial part of the Zernike polynomials for $D = 3$ yields

$$R_{3,s}^l(r) = \frac{(-1)^n}{n!} r^{-[\alpha+(1/2)]} \left[\frac{d}{d(r^2)} \right]^n (r^2)^{\alpha+n} (1-r^2)^n, \quad (74)$$

with $\alpha = l + 1/2$ and $n = (s - l)/2$. By expressing the spherical Bessel function in equation (72) in terms of the argument $z = qr$, and using equations (72) and (74) in the left-hand side of equation (61), we find the following expansion for the integral:

$$\int_0^1 R_{3,s}^l(r) j_l(qr) r^2 dr = \sqrt{\frac{\pi}{16}} \frac{(-1)^n}{n!} \sum_{p=0}^{\infty} \frac{(-1)^p}{p! \Gamma(p+\alpha+1)} \times \left(\frac{q}{2} \right)^{2p+\alpha-1/2} f(p, \alpha, n), \quad (75)$$

with

$$f(p, \alpha, n) = 2 \int_0^1 (r^2)^{p+1/2} \left[\frac{d}{d(r^2)} \right]^n (r^2)^{\alpha+n} (1-r^2)^n dr. \quad (76)$$

The integral in (76) can be reformulated by introducing the variable $u = r^2$, which yields

$$f(p, \alpha, n) = \int_0^1 u^p \left(\frac{d}{du} \right)^n u^{\alpha+n} (1-u)^n du. \quad (77)$$

We shall now perform integration by parts in equation (77) for two cases, $p \geq n$ and $p < n$. Notice that a single application of integration by parts results in

$$\begin{aligned} f(p, \alpha, n) &= \left[u^p \left(\frac{d}{du} \right)^{n-1} u^{\alpha+n} (1-u)^n \right]_0^1 \\ &\quad - p \int_0^1 u^{p-1} du \left(\frac{d}{du} \right)^{n-1} u^{\alpha+n} (1-u)^n \\ &= -p \int_0^1 u^{p-1} du \left(\frac{d}{du} \right)^{n-1} u^{\alpha+n} (1-u)^n, \end{aligned} \quad (78)$$

taking into account that $n-1 < n$ and $\alpha > 0$.

For $p \geq n$ we can then perform integration by parts n times in (77), which results in

$$f(p, \alpha, n) = (-1)^n \frac{p!}{(p-n)!} \int_0^1 u^{p+\alpha} (1-u)^n du. \quad (79)$$

One may recognize the remaining integral as the beta function [see 5.12.1 in NIST (2023)], which finally yields for $p \geq n$

$$f(p, \alpha, n) = (-1)^n \frac{p!}{(p-n)!} \frac{\Gamma(p+\alpha+1)\Gamma(n+1)}{\Gamma(p+n+\alpha+2)}. \quad (80)$$

In the case of $p < n$, it is possible to perform integration by parts only p times in (77), which gives

$$f(p, \alpha, n) = (-1)^p p! \left[\left(\frac{d}{du} \right)^{n-p-1} u^{\alpha+n} (1-u)^n \right]_0^1 = 0. \quad (81)$$

Using equations (80) and (81) in (75), the latter can be rewritten as

$$\begin{aligned} \int_0^1 R_{3,s}^l(r) j_l(qr) r^2 dr &= (-1)^n \frac{1}{q} \sqrt{\frac{\pi}{2q}} \\ &\times \sum_{p=n}^{\infty} \frac{(-1)^{p-n}}{(p-n)! \Gamma(p-n+2n+\alpha+2)} \left(\frac{q}{2} \right)^{2(p-n)+2n+\alpha+1} \\ &= (-1)^n \frac{1}{q} \sqrt{\frac{\pi}{2q}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma[k+(s+1)+\frac{3}{2}]} \left(\frac{q}{2} \right)^{2k+(s+1)+1/2}, \end{aligned} \quad (82)$$

where in the last step we considered that $2n+\alpha=s+\frac{1}{2}$ and introduced a variable $k=p-n$. Note that the sum in equation (82), including the prefactor $\sqrt{\pi/(2q)}$, is precisely the series expansion of a spherical Bessel function $j_{s+1}(q)$ of the order $s+1$ [see equation (72)]. This finally yields

$$\int_0^1 R_{3,s}^l(r) j_l(qr) r^2 dr = (-1)^n \frac{j_{s+1}(q)}{q} = (-1)^{(s-n)/2} \frac{j_{s+1}(q)}{q}, \quad (83)$$

which completes the proof of equation (61).

Acknowledgements

We are grateful to Alexander Lichtenstein for valuable discussions and supporting the project. We also thank Patrick Adams and Steffen Hauf for helpful discussions. Open access funding enabled and organized by Projekt DEAL.

References

- Altarelli, M., Kurta, R. P. & Vartanyants, I. A. (2010). *Phys. Rev. B*, **82**, 104207.
- Altarelli, M., Kurta, R. P. & Vartanyants, I. A. (2012). *Phys. Rev. B*, **86**, 179904.
- Aquila, A., Barty, A., Bostedt, C., Boutet, S., Carini, G., dePonte, D., Drell, P., Doniach, S., Downing, K. H., Earnest, T., Elmlund, H., Elser, V., Gühr, M., Hajdu, J., Hastings, J., Hau-Riege, S. P., Huang, Z., Lattman, E. E., Maia, F. R. N. C., Marchesini, S., Ourmazd, A., Pellegrini, C., Santra, R., Schlichting, I., Schroer, C., Spence, J. C. H., Vartanyants, I. A., Wakatsuki, S., Weis, W. I. & Williams, G. J. (2015). *Struct. Dyn.* **2**, 041701.
- Ardenne, B. von, Mechelke, M. & Grubmüller, H. (2018). *Nat. Commun.* **9**, 2375.
- Assalauova, D., Kim, Y. Y., Bobkov, S., Khubbutdinov, R., Rose, M., Alvarez, R., Andreasson, J., Balaur, E., Contreras, A., DeMirici, 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). *IUCrJ*, **7**, 1102–1113.
- Bielecki, J., Maia, F. R. N. C. & Mancuso, A. P. (2020). *Struct. Dyn.* **7**, 040901.
- 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.
- Born, M. & Wolf, E. (2019). *Principles of Optics*, 60th Anniversary Ed. Cambridge University Press.
- Bortel, G. & Tegze, M. (2011). *Acta Cryst. A* **67**, 533–543.
- Boutet, S., Lomb, L., Williams, G. J., Barends, T. R. M., 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. H., Chapman, H. N. & Schlichting, I. (2012). *Science*, **337**, 362–364.
- Bradski, G. (2000). *Dr Dobbs's J. Softw. Tools*, **25**(11), 122–125.
- Chapman, H. N. (2019). *Annu. Rev. Biochem.* **88**, 35–58.
- Chapman, H. N., Barty, A., Marchesini, S., Noy, A., Hau-Riege, S. P., Cui, C., Howells, M. R., Rosen, R., He, H., Spence, J. C. H., Weierstall, U., Beetz, T., Jacobsen, C. & Shapiro, D. (2006). *J. Opt. Soc. Am. A*, **23**, 1179–1200.
- 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. N. C., 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., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S. L., 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. D., Starodub, J., Williams, G., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Røcker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andrichke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X. I., Grotjohann, I., Holton, J. M., Barends, T. R. M., 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. H. (2011). *Nature*, **470**, 73–77.
- Chaudhuri, B., Muñoz, I. G., Qian, S. & Urban, V. S. (2017). Editors. *Biological Small Angle Scattering: Techniques, Strategies and Tips*, Advances in Experimental Medicine and Biology, Vol. 1009. Singapore: Springer.
- Clark, N. A., Ackerson, B. J. & Hurd, A. J. (1983). *Phys. Rev. Lett.* **50**, 1459–1462.
- Da Vela, S. & Svergun, D. I. (2020). *Curr. Res. Struct. Biol.* **2**, 164–170.
- Donatelli, J. J., Zwart, P. H. & Sethian, J. A. (2015). *Proc. Natl Acad. Sci. USA*, **112**, 10286–10291.
- Ekeberg, T., Assalauova, D., Bielecki, J., Boll, R., Daurer, B. J., Eichacker, L. A., Franken, L. E., Galli, D. E., Gelisio, L., Gumprecht, L., Gunn, L. H., Hajdu, J., Hartmann, R., Hasse, D., Ignatenko, A., Koliyadu, J., Kulyk, O., Kurta, R., Kuster, M., Lugmayr, W., Lübke, J., Mancuso, A. P., Mazza, T., Nettelblad, C., Ovcharenko, Y., Rivas, D. E., Rose, M., Samanta, A. K., Schmidt, P., Sobolev, E., Timneanu, N., Usenko, S., Westphal, D., Wollweber, T., Worbs, L., Xavier, P. L., Yousef, H., Ayyer, K., Chapman, H. N., Sellberg, J. A., Seuring, C., Vartanyants, I. A., Küpper, J., Meyer, M. & Maia, F. R. N. C. (2022). *bioRxiv*:2022.03.09.483477.

- Ekeberg, T., Svenda, M., Abergel, C., Maia, F. R. N. C., 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.
- Elser, V. (2011). *Ultramicroscopy*, **111**, 788–792.
- Fienup, J. R. (1978). *Opt. Lett.* **3**, 27–29.
- Fienup, J. R. (1982). *Appl. Opt.* **21**, 2758–2769.
- Flamant, J., Le Bihan, N., Martin, A. V. & Manton, J. H. (2016). *Phys. Rev. E*, **93**, 053302.
- Ford, W. (2014). *Numerical Linear Algebra with Applications*, 1st ed. Academic Press.
- Fung, R., Shneerson, V., Saldin, D. K. & Ourmazd, A. (2009). *Nat. Phys.* **5**, 64–67.
- Gaffney, K. J. & Chapman, H. N. (2007). *Science*, **316**, 1444–1448.
- Gerchberg, R. W. & Saxton, W. O. (1972). *Optik*, **35**, 227–246.
- Giannakis, D., Schwander, P. & Ourmazd, A. (2012). *Opt. Express*, **20**, 12799–12826.
- Gower, J. C. & Dijksterhuis, G. B. (2004). *Procrustes Problems*, Oxford Statistical Science Series, Vol. 30. Oxford University Press.
- Guo, Y. R. & MacKinnon, R. (2017). *eLife*, **6**, e33660.
- 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.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. E. (2020). *Nature*, **585**, 357–362.
- Hughes-Hallett, D., Gleason, A. M., McCallum, W. G., Flath, D. E., Lock, P. F., Gordon, S. P., Lomen, D. O., Lovelock, D., Osgood, B. G., Pasquale, A., Quinney, D., Tecosky-Feldman, J., Thrash, J., Rhea, K. R. & Tucker, T. W. (2012). *Calculus: Single and Multi-variable*, 6th ed. John Wiley & Sons.
- Hunter, J. D. (2007). *Comput. Sci. Eng.* **9**, 90–95.
- Jaeschke, E. J., Khan, S., Schneider, J. R. & Hastings, J. B. (2016). Editors. *Synchrotron Light Sources and Free-Electron Lasers*. Springer Cham.
- Kam, Z. (1977). *Macromolecules*, **10**, 927–934.
- Kam, Z. (1980). *J. Theor. Biol.* **82**, 15–39.
- Kam, Z., Koch, M. H. J. & Bordas, J. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 3559–3562.
- Kasemeyer, S., Jafarpour, A., Lomb, L., Steinbrener, J., Martin, A. V. & Schlichting, I. (2013). *Phys. Rev. E*, **88**, 042710.
- Kimura, T., Joti, Y., Shibuya, A., Song, C., Kim, S., Tono, K., Yabashi, M., Tamakoshi, M., Moriya, T., Oshima, T., Ishikawa, T., Bessho, Y. & Nishino, Y. (2014). *Nat. Commun.* **5**, 3052.
- Kirian, R. A., Schmidt, K. E., Wang, X., Doak, R. B. & Spence, J. C. H. (2011). *Phys. Rev. E*, **84**, 011921.
- Kommer, P. R., Ramakrishnaiah, V., Sweeney, C., Donatelli, J. & Zwart, P. H. (2021). *J. Appl. Cryst.* **54**, 1179–1188.
- Kostelec, P. J. & Rockmore, D. N. (2008). *J. Fourier Anal. Appl.* **14**, 145–179.
- Kurta, R. P., Altarelli, M. & Vartanyants, I. A. (2016). *Adv. Chem. Phys.* **161**, 1–39.
- Kurta, R. P., Altarelli, M., Weckert, E. & Vartanyants, I. A. (2012). *Phys. Rev. B*, **85**, 184204.
- Kurta, R. P., Donatelli, J. J., Yoon, C. H., Berntsen, P., Bielecki, J., Daurer, B. J., DeMirici, 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.
- Kurta, R. P., Dronyak, R., Altarelli, M., Weckert, E. & Vartanyants, I. A. (2013). *New J. Phys.* **15**, 013059.
- Kurta, R. P., van Driel, T. B., Dohn, A. O., Berberich, T. B., Nelson, S., Zaluzhnyy, I. A., Mukharamova, N., Lapkin, D., Zederkof, D. B., Seaberg, M., Pedersen, K. S., Kjaer, K. S., Rippey, G. I., Biasin, E., Møller, K. B., Gelisio, L., Haldrup, K., Vartanyants, I. A. & Nielsen, M. M. (2023). *Phys. Chem. Chem. Phys.* **25**, 23417–23434.
- Liu, H., Poon, B. K., Saldin, D. K., Spence, J. C. H. & Zwart, P. H. (2013). *Acta Cryst.* **A69**, 365–373.
- Loh, N. D. & Elser, V. (2009). *Phys. Rev. E*, **80**, 026705.
- Malmerberg, E., Kerfeld, C. A. & Zwart, P. H. (2015). *IUCrJ*, **2**, 309–316.
- Mancuso, A. P., Gorniak, T., Staier, F., Yefanov, O. M., Barth, R., Christophis, C., Reime, B., Gulden, J., Singer, A., Pettit, M. E., Nisius, T., Wilhein, T., Gutt, C., Grübel, G., Guerassimova, N., Treusch, R., Feldhaus, J., Eisebitt, S., Weckert, E., Grunze, M., Rosenhahn, A. & Vartanyants, I. A. (2010). *New J. Phys.* **12**, 035003.
- Marchesini, S. (2007). *Rev. Sci. Instrum.* **78**, 011301.
- Marchesini, S., He, H., Chapman, H., Hau-Riege, S., Noy, A., Howells, M., Weierstall, U. & Spence, J. (2003). *Phys. Rev. B*, **68**, 140101.
- Martin, A. V. (2017). *IUCrJ*, **4**, 24–36.
- 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.
- Nakano, M., Miyashita, O., Jonic, S., Song, C., Nam, D., Joti, Y. & Tama, F. (2017). *J. Synchrotron Rad.* **24**, 727–737.
- Nakano, M., Miyashita, O., Jonic, S., Tokuhisa, A. & Tama, F. (2018). *J. Synchrotron Rad.* **25**, 1010–1021.
- Nakasako, M., Kobayashi, A., Takayama, Y., Asakura, K., Oide, M., Okajima, K., Oroguchi, T. & Yamamoto, M. (2020). *Biophys. Rev.* **12**, 541–567.
- Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature*, **406**, 752–757.
- NIST (2023). *NIST Digital Library of Mathematical Functions*. <https://dlmf.nist.gov/>, release 1.1.9 of 2023-03-15. Edited by F. W. J. Olver, A. M. F. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl & M. A. McClain. NIST, Gaithersburg, MD, USA.
- Nugent, K. A. (2010). *Adv. Phys.* **59**, 1–99.
- Orville, A. M. (2020). *Curr. Opin. Struct. Biol.* **65**, 193–208.
- 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.
- Pandey, S., Bean, R., Sato, T., Poudyal, I., Bielecki, J., Cruz Villarreal, J., Yefanov, O., Mariani, V., White, T. A., Kupitz, C., Hunter, M., Abdellatif, M. H., Bajt, S., Bondar, V., Echelmeier, A., Doppler, D. M., Emons, M., Frank, M., Fromme, R., Gevorkov, Y., Giovannetti, G., Jiang, M., Kim, D., Kim, Y., Kirkwood, H., Klimovskaia, A., Knoska, J., Koua, F. H. M., Letrun, R., Lisova, S., Maia, L., Mazalova, V., Meza, D., Michelat, T., Ourmazd, A., Palmer, G., Ramilli, M., Schubert, R., Schwander, P., Silenzi, A., Sztuk-Dambietz, J., Tolstikova, A., Chapman, H. N., Ros, A., Barty, A., Fromme, P., Mancuso, A. P. & Schmidt, M. (2020). *Nat. Methods*, **17**, 73–78.
- Pawn, B. R. (2013). *J. Phys. Condens. Matter*, **25**, 383201.
- Pedriani, B., Menzel, A., Guizar-Sicairos, M., Guzenko, V. A., Gorelick, S., David, C., Patterson, B. D. & Abela, R. (2013). *Nat. Commun.* **4**, 1647.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Rodenburg, J. M. (2008). *Adv. Imaging Electron. Phys.* **150**, 87–184.
- Rose, M. (1957). *Elementary Theory of Angular Momentum*, Structure of Matter Series. New York: John Wiley & Sons.

- 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.
- Saldin, D. K., Poon, H. C., Schwander, P., Uddin, M. & Schmidt, M. (2011). *Opt. Express*, **19**, 17318–17335.
- Saldin, D. K., Poon, H. C., Shneerson, V. L., Howells, M., Chapman, H. N., Kirian, R. A., Schmidt, K. E. & Spence, J. C. H. (2010). *Phys. Rev. B*, **81**, 174105.
- Saldin, D. K., Shneerson, V. L., Fung, R. & Ourmazd, A. (2009). *J. Phys. Condens. Matter*, **21**, 134014.
- Schaeffer, N. (2013). *Geochem. Geophys. Geosyst.* **14**, 751–758.
- Schroeder, W., Martin, K. & Lorensen, B. (2006). *The Visualization Toolkit: an Object-Oriented Approach to 3D Graphics*. Kitware.
- Seibert, M. M., Ekeberg, T., Maia, F. R. N. C., 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. H., 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., Andrich, 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.
- Solem, J. & Baldwin, G. (1982). *Science*, **218**, 229–235.
- 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. N., 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. H. & Bogan, M. J. (2012). *Nat. Commun.* **3**, 1276.
- The HDF Group, Koziol, Q. & US DOE Office of Science (2020). HDF5-Version 1.12.0, <https://www.osti.gov/biblio/1631295>.
- Ueda, K. (2018). Editor. *Applied Sciences. Special Issue 'X-ray Free-Electron Laser'*. MDPI.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., de Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T. J., Robitaille, T. P., Spura, T., Jones, T. R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y. O. & Vázquez-Baeza, Y. (2020). *Nat. Methods*, **17**, 261–272.
- Wiedorn, M. O., Oberthür, D., Bean, R., Schubert, R., Werner, N., Abbey, B., Aepfelbacher, M., Adriano, L., Allahgholi, A., Al-Qudami, N., Andreasson, J., Aplin, S., Awel, S., Ayer, K., Bajt, S., Barák, I., Bari, S., Bielecki, J., Botha, S., Boukhelef, D., Brehm, W., Brockhauser, S., Cheviakov, I., Coleman, M. A., Cruz-Mazo, F., Danilevski, C., Darmanin, C., Doak, R. B., Domaracký, M., Dörner, K., Du, Y., Fangohr, H., Fleckenstein, H., Frank, M., Fromme, P., Gañán-Calvo, A. M., Gevorkov, Y., Giewekemeyer, K., Ginn, H. M., Graafsma, H., Graceffa, R., Greiffenberg, D., Gumprecht, L., Göttlicher, P., Hajdu, J., Hauf, S., Heymann, M., Holmes, S., Horke, D. A., Hunter, M. S., Imlau, S., Kaukher, A., Kim, Y., Klyuev, A., Knoška, J., Kobe, B., Kuhn, M., Kupitz, C., Küpper, J., Lahey-Rudolph, J. M., Laurus, T., Le Cong, K., Letrun, R., Xavier, P. L., Maia, L., Maia, F. R. N. C., Mariani, V., Messerschmidt, M., Metz, M., Mezza, D., Michelat, T., Mills, G., Monteiro, D. C. F., Morgan, A., Mühlig, K., Munke, A., Münnich, A., Nette, J., Nugent, K. A., Nuguid, T., Orville, A. M., Pandey, S., Pena, G., Villanueva-Perez, P., Poehlsen, J., Previtali, G., Redecke, L., Riekehr, W. M., Rohde, H., Round, A., Safenreiter, T., Sarrou, I., Sato, T., Schmidt, M., Schmitt, B., Schönherr, R., Schulz, J., Sellberg, J. A., Seibert, M. M., Seuring, C., Shelby, M. L., Shoeman, R. L., Sikorski, M., Silenzi, A., Stan, C. A., Shi, X., Stern, S., Sztuk-Dambietz, J., Szuba, J., Tolstikova, A., Trebbin, M., Trunk, U., Vagovic, P., Ve, T., Weinhausen, B., White, T. A., Wrona, K., Xu, C., Yefanov, O., Zatsepin, N., Zhang, J., Perbandt, M., Mancuso, A. P., Betzel, C., Chapman, H. & Barty, A. (2018). *Nat. Commun.* **9**, 4025.
- Wochner, P., Gutt, C., Autenrieth, T., Demmer, T., Bugaev, V., Ortiz, A. D., Duri, A., Zontone, F., Gröbel, G. & Dosch, H. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 11511–11514.
- Yefanov, O. M. & Vartanyants, I. A. (2013). *J. Phys. B At. Mol. Opt. Phys.* **46**, 164013.
- Yuan, S., Topf, M., Reubold, T. F., Eschenburg, S. & Akey, C. W. (2013). *Biochemistry*, **52**, 2319–2327.
- Zaluzhnyy, I. A., Kurta, R. P., André, A., Gorobtsov, O. Y., Rose, M., Skopintsev, P., Besedin, I., Zozulya, A. V., Sprung, M., Schreiber, F., Vartanyants, I. A. & Scheele, M. (2017). *Nano Lett.* **17**, 3511–3517.
- Zernike, F. (1934). *Physica*, **1**, 689–704.
- Zernike, F. & Brinkman, H. C. (1935). *Proc. Akad. Wet. Amsterdam*, **38**, 161–170.