

# A polar Fourier geometric approach to volume-free single-particle 3D reconstruction

Cong T. S. Van, Cyril F. Reboul, Joseph J. E. Caesar, Rubén Meana-Pañeda and Hans Elmlund\*

National Cancer Institute (NCI), National Institutes of Health (NIH), Frederick, MD 21701, USA. \*Correspondence e-mail: hans.elmlund@nih.gov

Received 8 January 2026

Accepted 7 April 2026

Edited by S. Raunser, Max Planck Institute of Molecular Physiology, Germany

This article is part of the special issue *CryoEM in the Fast Lane of Structural Biology*.

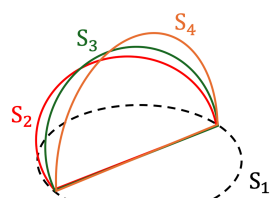
**Keywords:** single-particle 3D reconstruction; polar Fourier coordinates; orientation-recovery problem.

**Supporting information:** this article has supporting information at [www.iucrj.org](http://www.iucrj.org)

We introduce a compact mathematical formulation for the inverse single-particle 3D reconstruction problem, a high-dimensional inverse problem in which millions of parameters are estimated from extremely noisy experimental measurements. Given a collection of noisy 2D projection images (particles) of an unknown 3D charge-density distribution, the objective is to infer the unknown particle orientations and thereby enable *ab initio* 3D reconstruction via tomographic methods. We develop a method for generating regularized reprojections directly from the noisy particles that does not rely on explicit 3D density reconstruction. Instead, we recast the *ab initio* orientation-recovery problem in polar Fourier coordinates through discretization of the rotation group  $SO(3)$ . The directions of projection are mapped onto slices intersecting the origin of the 3D Fourier transform. The rotations in the plane normal to a projection direction are mapped onto radial lines in the 2D Fourier transforms of the particles. An optimization procedure jointly estimates particle projection directions, in-plane rotation angles and rotational origin offsets. Regularized reprojections are computed by averaging along lines in the polar Fourier representation, exploiting data redundancy to suppress noise and improve stability. We present the mathematical framework in detail and provide initial benchmarks demonstrating the performance and robustness of the approach.

## 1. Introduction

The fundamental theory for 3D reconstruction from projection images in electron microscopy was developed by De Rosier & Klug (1968) and used to reconstruct the tomato bushy stunt virus (Crowther *et al.*, 1970). The theory is based on the projection-slice theorem (Bracewell, 1956), which states that the Fourier transform of the projection of a 3D charge-density distribution corresponds to a central section through the 3D volume's Fourier transform. The projection-slice theorem underpins convolution interpolation in electron microscopy, both for extracting reprojections from volumes at known 3D orientations (Yang & Penczek, 2008; Penczek, 2010; Penczek *et al.*, 2004) and for calculating 3D reconstructions from particles with assigned 3D orientations (Penczek *et al.*, 2004; O'Sullivan, 1985; Beatty *et al.*, 2005; Jackson *et al.*, 1991). Another consequence of the projection-slice theorem is that any two projection images with non-parallel directions will share a common line in the Fourier domain, and their relative 3D orientations will be fixed to a rotation around the axis defined by this line. This concept of 'common lines' provided the mathematical foundation for angular reconstitution (Van Heel, 1987), where the relative 3D orientations of three class averages are determined unambiguously and used as an anchor set to determine the relative 3D orientations of larger sets of class averages. Angular



$$\zeta(S_2) + \zeta(S_3) + \zeta(S_4) = \text{one line of } S_1$$

reconstitution was the first *ab initio* 3D reconstruction approach used to successfully reconstruct particles with low point-group symmetry (Stark *et al.*, 1995; Schatz *et al.*, 1995). However, no *ab initio* 3D reconstruction approach can directly resolve the ambiguity of enantiomorphism. Therefore, the handedness of the 3D density map must be determined from the handedness of resolved secondary structure elements, as it was done for the bacteriorhodopsin electron crystallography structure (Henderson *et al.*, 1990), or tilt experiments if the resolution of the 3D map is insufficient (Rosenthal & Henderson, 2003).

The theory of common lines has remained in use for icosahedral particles (Fuller *et al.*, 1996) but modern probabilistic *ab initio* 3D reconstruction approaches (Zivanov *et al.*, 2018; Punjani *et al.*, 2017; Van *et al.*, 2025) instead rely on projection matching (Grigorieff, 2007; Penczek *et al.*, 1992), where particles are iteratively matched with volume reprojections to generate probabilities that direct the orientation search. Much of the recent work on 3D reconstruction theory utilizing common lines has focused on improving their error-prone detection (Hall *et al.*, 2007; Singer *et al.*, 2010; Singer & Shkolnisky, 2011; Herman & Kalinowski, 2008; Wang *et al.*, 2024) or the use of algebraic constraints for denoising prior to detection (Muller *et al.*, 2024).

Single-particle 3D reconstruction algorithms based on probabilistic inference typically define the model being refined as the real-space volumetric density sought (Zhong *et al.*, 2021; Punjani *et al.*, 2017; Jaitly *et al.*, 2010; Scheres, 2012a; Lyumkis *et al.*, 2013). However, the model required for generating reprojections with maximally enhanced signal-to-noise ratio (SNR) does not have to match the model that we seek to estimate from a structural point of view. In this study, we therefore asked whether the concept of common lines could be used to create an abstract mathematical model that replaces volumetric 3D reconstruction, while preserving the desired signal enhancement along equivalent lines in the dataset. All computations related to generating and manipulating the 3D reconstruction (masking, filtering, reprojection, *etc.*) are removed, and an algorithm with computational complexity similar to 2D multi-reference alignment (Scheres, Valle & Carazo, 2005; Scheres, Valle, Nuñez *et al.*, 2005; Reboul *et al.*, 2016; Yang *et al.*, 2012; Sorzano *et al.*, 2010; Romanov *et al.*, 2021) is provided that expresses the geometrical constraints inherent to the system – that non-parallel pairs of reprojections share a common line in the Fourier domain. Our mathematical model embeds a search space of  $N_p$  projection directions, which are equidistant points on the unit sphere. The  $N$  particles are assigned to these  $N_p$  projection directions, while optimizing the in-plane degrees of freedom. When creating an average from each projection direction, the common lines from particles in the other  $N_p - 1$  projection directions each contribute a single intersecting line to the generation of the ‘class average’, which is more appropriately referred to as a ‘reprojection estimate’ in this setting. Unlike classical common-line approaches, which attempt to explicitly detect and validate individual common lines in noisy data, our method assumes a discretized orien-

tation manifold and uses common-line geometry implicitly as a deterministic averaging operator. Hence, common lines are not detected but enforced geometrically through SO(3) discretization. We simply assume that the projection directions in the dataset data are distributed according to some distribution on the unit sphere that defines the geometrical relationships needed for averaging. Our formulation replaces explicit 3D Fourier gridding with an operator defined directly on discretized central sections of reciprocal space, preserving the projection-slice geometry while eliminating the need for volumetric interpolation. Optimization of the model can be carried out exactly as it is done for projection-matching-based implementations (Grigorieff, 2007; Scheres, 2012a; Punjani *et al.*, 2017; Van *et al.*, 2025) or multi-reference 2D alignment (Reboul *et al.*, 2016; Scheres, Valle, Nuñez *et al.*, 2005; Sorzano *et al.*, 2010; Yang *et al.*, 2012; Sigworth, 1998; Romanov *et al.*, 2021). We implemented the method in polar coordinates; the particles are converted into 2D polar Fourier transforms and then reprojection estimates are generated through averaging lines using pre-calculated data structures that express the geometrical relationships. We outline the mathematical theory in detail, describe tests that validate the accuracy of our implementation, and provide benchmarks that demonstrate how the method can be used to accelerate projection-matching-based approaches for iterative *ab initio* 3D reconstruction.

## 2. Methods

### 2.1. Statistical formulation of the orientation-recovery problem

*Ab initio* 3D reconstruction in single-particle cryo-EM requires estimating unknown particle 3D orientations from extremely noisy 2D projection images. We consider a set of  $N$  noisy particles  $\{X_i\}_{i=1}^N$ , modeled in Fourier space as central sections of the Fourier transform of an unknown 3D density  $V$ . For particle  $i$ , the forward model is

$$X_i(k) = H_i(k)[\wp_{\phi_i} V](k) + \varepsilon_i(k), \quad (1)$$

where  $k$  indexes 2D Fourier coordinates,  $k \in \Omega_{2D}$ ;  $H_i(k)$  is the contrast transfer function (CTF);  $\wp_{\phi_i}$  is the projection operator at pose  $\phi_i \in \text{SE}(3)$  (3D rotation + in-plane shift);  $v \in \mathbb{R}^3$  is the real-valued 3D density sought and  $V$  its Fourier transform; and  $\varepsilon_i(k) \simeq N(0, \sigma_i^2)$  is additive Gaussian noise.

We seek to estimate orientations  $\{\phi_i\}$  and volume  $V$  using the following regularized objective:

$$\mathcal{L}(V, \{\phi_i\}) = \mathcal{L}_{\text{data}} + \lambda \mathcal{L}_{\text{prior}}, \quad (2)$$

where  $\lambda$  is a regularization parameter and

$$\mathcal{L}_{\text{data}} = \sum_{i=1}^N \sum_{k \in \Omega_{2D}} \frac{|X_i(k) - H_i(k)[\wp_{\phi_i} V](k)|^2}{\sigma_i^2(|k|)}. \quad (3)$$

The regularization objective

$$\mathcal{L}_{\text{prior}}(V) = \sum_{l \in \Omega_{3\text{D}}} \frac{|V(l)|^2}{\tau^2(|l|)}, \quad (4)$$

where  $l$  indexes 3D Fourier coordinates and  $\tau^2(|l|)$  is a shell-wise signal variance, imposes a Gaussian prior on the reconstructed 3D Fourier coefficients

$$V(l) \simeq N[0, \tau^2(|l|)]. \quad (5)$$

## 2.2. Volume-based fixed-point optimization

Standard fixed-point approaches alternate between estimating orientations by projection matching and reconstructing an intermediate 3D volume. While effective, this strategy is computationally expensive and enforces a tight coupling between orientation estimation and 3D volume interpolation. In contrast to the approach implemented in the software package *RELION* (Scheres, 2012a,b), which uses weighted orientation assignment, we use a hard orientation assignment optimization scheme, alternating between orientation update

$$\phi_i = \arg \min_{\phi} \sum_{k \in \Omega_{2\text{D}}} \frac{|X_i(k) - H_i(k)[\wp_{\phi_i} V](k)|^2}{\sigma_i^2(|k|)} \quad (6)$$

and volume update for fixed orientations:

$$V(l) = \frac{\sum_{i=1}^N \sum_{k \in \Omega_{2\text{D}}} \wp_{\phi_i}^T(l, k) \frac{H_i(k)X_i(k)}{\sigma_i^2(|k|)}}{\sum_{i=1}^N \sum_{k \in \Omega_{2\text{D}}} \wp_{\phi_i}^T(l, k) \frac{H_i^2(k)}{\sigma_i^2(|k|)} + \frac{\lambda}{\tau^2(|l|)}}, \quad (7)$$

where  $\wp_{\phi_i}^T(l, k)$  is the reconstruction operator that maps  $X_i(k)$  onto the 3D Fourier grid through convolution interpolation with a Kaiser–Bessel kernel (Penczek *et al.*, 2004; Beatty *et al.*, 2005). The summation over  $k$  in equation (7) represents summation over neighboring components of a central section of the Fourier volume. The result of the expression  $\wp_{\phi_i}^T(l, k)[H_i(k)X_i(k)/\sigma_i^2(|k|)]$  is a 3D window of Fourier components. Although a value of  $<1$  was originally proposed for the regularization parameter  $\lambda$  (Scheres, 2012a,b), we used  $\lambda = 1$  for all subsequent calculations. We previously described the probabilistic optimization algorithm used for solving equation (6) in the expectation-maximization iterations (Van *et al.*, 2025).

## 2.3. Proposed volume-free fixed-point reconstruction framework

Here we introduce an alternative formulation in which orientation estimation is carried out entirely in the polar Fourier domain, without explicitly forming intermediate 3D volumes. The key idea is to replace volume-based reprojection with a linear operator that directly estimates reprojections from the set of noisy particles using discretized SO(3) geometry and polar common-line constraints. The SO(3) manifold is the mathematical space that represents all possible ways an object can be rotated in 3D space around a fixed center, whereas the 2-sphere ( $S^2$ ) is the surface of this 3D space at a defined radius. Furthermore, the  $S^1$  manifold denotes a circle representing all possible rotations about the

origin in a single 2D plane. Our discretization of the rotation group SO(3) maps the 2-sphere,  $S^2$ , onto slices passing through the origin of the 3D Fourier transform, and the circle manifold,  $S^1$ , onto radial lines in the 2D Fourier transforms of the particles. We discretize SO(3) to a uniform grid of size  $N_m \times N_{\theta}$  over  $S^2 \times S^1$ , where  $N_m$  is the number of slices (references) on  $S^2$  and  $N_{\theta}$  is the number of lines (in-plane rotations) on  $S^1$ , so each orientation  $o_i \in \text{SO}(3)$  corresponds to  $o_{m,\theta} \in S^2 \times S^1$ . An Archimedean spiral on  $S^2$  is used to generate evenly distributed projection directions (Baldwin & Penczek, 2007). Particles, CTF and  $\sigma^2$  factors are converted to polar Fourier coordinates:

$$X_i(k) \rightarrow \hat{X}_i(r, \theta) \equiv \hat{X}_i, \quad (8)$$

$$H_i(k) \rightarrow \hat{H}_i(r, \theta) \equiv \hat{H}_i \quad (9)$$

and

$$\sigma_i^2(|k|) \rightarrow \hat{\sigma}_i^2(r) \equiv \hat{\sigma}_i^2. \quad (10)$$

Equations (8)–(10) define the short-hand notation used from here on. In the polar representation, projection directions correspond to Fourier slices, in-plane rotations correspond to angular shifts, and common lines correspond to identical radial lines between slices. Instead of computing  $\wp_{\alpha} V$  we construct reprojection estimates directly from particles assigned to orientations. We define the index set  $I_{\alpha}$  comprising all particles assigned to projection direction  $\alpha$ . We also define the index set  $I_{\beta}$  comprising all particles assigned to all orientations  $\beta$  other than  $\alpha$ :

$$I_{\alpha} = \{i : \phi_i = \alpha\} \quad (11)$$

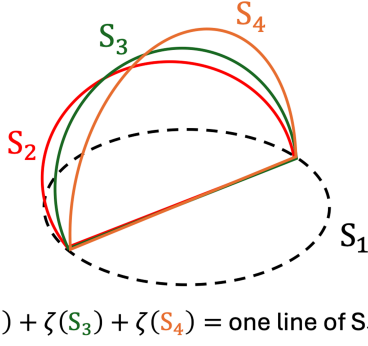
and

$$I_{\beta} = \{i : \phi_i \neq \alpha\}. \quad (12)$$

Next, we define a linear operator that estimates the reprojection at  $\alpha$  from the contributions of particles from both sets  $I_{\alpha}$  and  $I_{\beta}$ . Formally, we define the operator as

$$\mathcal{R}_{\alpha}(\{\hat{X}_i, \hat{H}_i, \hat{\sigma}_i^2\}, I_{\alpha}, I_{\beta}) = \frac{\sum_{i \in I_{\alpha}} \frac{\hat{H}_i \hat{X}_i}{\hat{\sigma}_i^2} + \sum_{j \in I_{\beta}} \zeta_{\alpha\beta} \frac{\hat{H}_j \hat{X}_j}{\hat{\sigma}_j^2}}{\sum_{i \in I_{\alpha}} \frac{\hat{H}_i^2}{\hat{\sigma}_i^2} + \sum_{j \in I_{\beta}} \zeta_{\alpha\beta} \frac{\hat{H}_j^2}{\hat{\sigma}_j^2} + \frac{\lambda}{\tau_{\alpha}^2(r)}}, \quad (13)$$

where  $\zeta_{\alpha\beta}$  extracts and interpolates the common radial line. Hence, the result of the summation  $\sum_{i \in I_{\alpha}} (\hat{H}_i \hat{X}_i / \hat{\sigma}_i^2)$  is the unnormalized 2D polar Fourier transform that would typically be referred to as a ‘class average’ in multi-reference 2D alignment, whereas the result of the summation  $\sum_{j \in I_{\beta}} \zeta_{\alpha\beta} (\hat{H}_j \hat{X}_j / \hat{\sigma}_j^2)$  is the unnormalized 2D polar Fourier transform constructed from line contributions from particles assigned to different orientations. The  $\mathcal{R}_{\alpha}$  operator ties together the projection-slice theorem and the discretized angular search space through the discrete polar Fourier transform (Fig. 1). We replace the volume-based projection in the objective function to obtain the orientation-update equation in the volume-free framework. Thus,



**Figure 1**  
Illustration of reciprocal-space line contributions induced by the projection-slice theorem on discretized SO(3). One line of the common-line contribution on Fourier slice  $S_1$  associated with orientation  $\alpha$  does not contain the line of  $\alpha$  itself but contains the sum of all common lines between plane  $\alpha$  and slices  $S_2, S_3, S_4, \dots$  of  $I_\beta$ .

$$\phi_i^{(t+1)} = \arg \min_{\alpha \in G} \sum_{(r, \theta)} \frac{|\hat{X}_i(r, \theta) - \mathcal{R}_\alpha^{(i)}(r, \theta)|^2}{\hat{\sigma}_i^2(r)}, \quad (14)$$

where  $G$  is the discretized SO(3) grid. This preserves the optimization objective but changes the representation of the reprojection operator. In contrast to volume-based projection matching, the common-lines formulation separates the contribution of lines from particles with parallel projection directions from the contribution of lines from particles with non-parallel directions in the reprojection estimation. This allows us to address the problem of directional bias (Reuter & Fischl, 2011) with the updated reprojection estimation operator

$$\mathcal{R}_\alpha = \frac{\sum_{j \in I_\beta} \zeta_{\alpha\beta} \frac{\hat{h}_j \hat{X}_j}{\hat{\sigma}_j^2}}{\sum_{j \in I_\beta} \zeta_{\alpha\beta} \frac{\hat{h}_j^2}{\hat{\sigma}_j^2} + \frac{\lambda}{\tau_\alpha^2(r)}}. \quad (15)$$

Equation (15) corresponds to replacing the full reprojection estimator in equation (13) with a leave-one-direction-out estimator, mirroring other common-lines-based approaches (Elmlund & Elmlund, 2009; Elmlund *et al.*, 2010, 2008). Since particles assigned to identical projection directions share no geometrically independent common lines, excluding them does not remove independent constraints. The final volume is reconstructed in Cartesian coordinates by equation (7) after all  $\{\phi_i\}$  have been estimated using equations (14) and (15). Fig. 2 provides a schematic summary of our method for polar Fourier *ab initio* 3D reconstruction.

#### 2.4. Interpolation of polar common lines

Since each slice is discretized over  $N_\theta$  in-slice rotations on  $S^1$ , the line-extraction operation  $\zeta_{\alpha\beta}(\hat{X}_\beta)$  extracts the common radial line  $L_{\alpha\beta}$  of slice  $\hat{X}_\beta$ . The common line lies at a continuous angular position  $\phi_{\beta \rightarrow \alpha}$  in slice  $\beta$ . In general, this angle does not coincide with a discrete angular sample. Let  $\theta_\gamma$  and  $\theta_{\gamma+1}$  be the two adjacent angular coordinates in slice  $\beta$  such that

$$\theta_\gamma \leq \phi_{\beta \rightarrow \alpha} < \theta_{\gamma+1}. \quad (16)$$

Rather than interpolating in angular coordinates, we perform interpolation in the embedded Cartesian Fourier plane. Let  $d_l$  and  $d_r$  denote the Euclidean distances to the two neighboring angular samples. The interpolation weights are defined as

$$w_l(k) = k \left[ \frac{d_r(k)}{d_l(k) + d_r(k)} \right], \quad w_r(k) = k \left[ \frac{d_l(k)}{d_l(k) + d_r(k)} \right], \quad (17)$$

where the radial weighting proportional to the frequency index  $k$  corresponds to the Jacobian of the polar-to-Cartesian coordinate transformation, which ensures that contributions from higher-frequency rings are appropriately weighted. The interpolated radial line is given by

$$\zeta_{\alpha\beta}(\hat{X}_\beta) = w_l(r, \phi) L_l + w_r(r, \phi) L_r. \quad (18)$$

The extracted interpolated radial line is subsequently inserted into slice  $\alpha$  at its corresponding angular location using the same distance-based weighting scheme.

#### 2.5. Point-group symmetry

For point-group symmetry  $N_{\text{sym}} > 1$ , we incorporate symmetry directly in reprojection estimation by including symmetry-induced common-line contributions for every particle. For each particle orientation, all non-identity symmetry operators generate additional valid common lines with each reference slice; these radial lines are extracted and inserted in polar Fourier space during the averaging step. Symmetry is therefore enforced during the iterative reciprocal-space estimation, without constructing or symmetrizing a 3D volume.

#### 2.6. Estimation of SSNR from polar reprojection correlations

The shell-wise prior variance  $\tau^2(k)$  is estimated from an even/odd split of the dataset. Twofold cross-validation in our algorithmic context yields two sets of polar reprojection estimates,  $\mathcal{R}_\alpha^{\text{even}}$  and  $\mathcal{R}_\alpha^{\text{odd}}$ , for each discretized projection direction  $\alpha$ . For each resolution shell  $k$  (corresponding to radial coordinate  $r$ ), we compute a ring-wise Fourier correlation (Saxton & Baumeister, 1982) across all projection directions:

$$\text{FRC}_\alpha(r) = \frac{\mathbb{R}\{\sum \mathcal{R}_\alpha^{\text{even}}(r) [\mathcal{R}_\alpha^{\text{odd}}(r)]^*\}}{\left[ \sum |\mathcal{R}_\alpha^{\text{even}}(r)|^2 \sum |\mathcal{R}_\alpha^{\text{odd}}(r)|^2 \right]^{1/2}}. \quad (19)$$

This quantity measures the reproducible signal content of the reprojection model at each spatial frequency. Under the assumption of independence between half-sets (Scheres & Chen, 2012; Sorzano *et al.*, 2022), the corresponding spectral signal-to-noise ratio (SSNR) is estimated using the standard relationship (Unser *et al.*, 2005)

$$\text{SSNR}_\alpha(r) = \frac{\text{FRC}_\alpha(r)}{1 - \text{FRC}_\alpha(r)}. \quad (20)$$

The ring-wise prior variance  $\tau_\alpha^2(r)$  is then obtained directly from this SSNR estimate. Because the reprojection model is constructed entirely in the polar Fourier domain, this proce-

**Table 1**  
Overall polar Fourier *ab initio* 3D reconstruction algorithm.

Stage	Orientation search	Regularization	Shift search	Maximum No. of projection directions
1	Hybrid SHC/probabilistic	Gaussian	No	500
2	Hybrid SHC/probabilistic	Gaussian	No	500
3	Hybrid SHC/probabilistic	ML	Yes	1000
4	Hybrid SHC/probabilistic	ML	Yes	1000
5	Probabilistic	ML	Yes	1000
6	Probabilistic	ML	Yes	1000
7	Probabilistic	ML	Yes	1500
8	Probabilistic	ML	Yes	1500

ture provides a resolution-dependent estimate of signal variance without requiring the formation of intermediate 3D volumes.

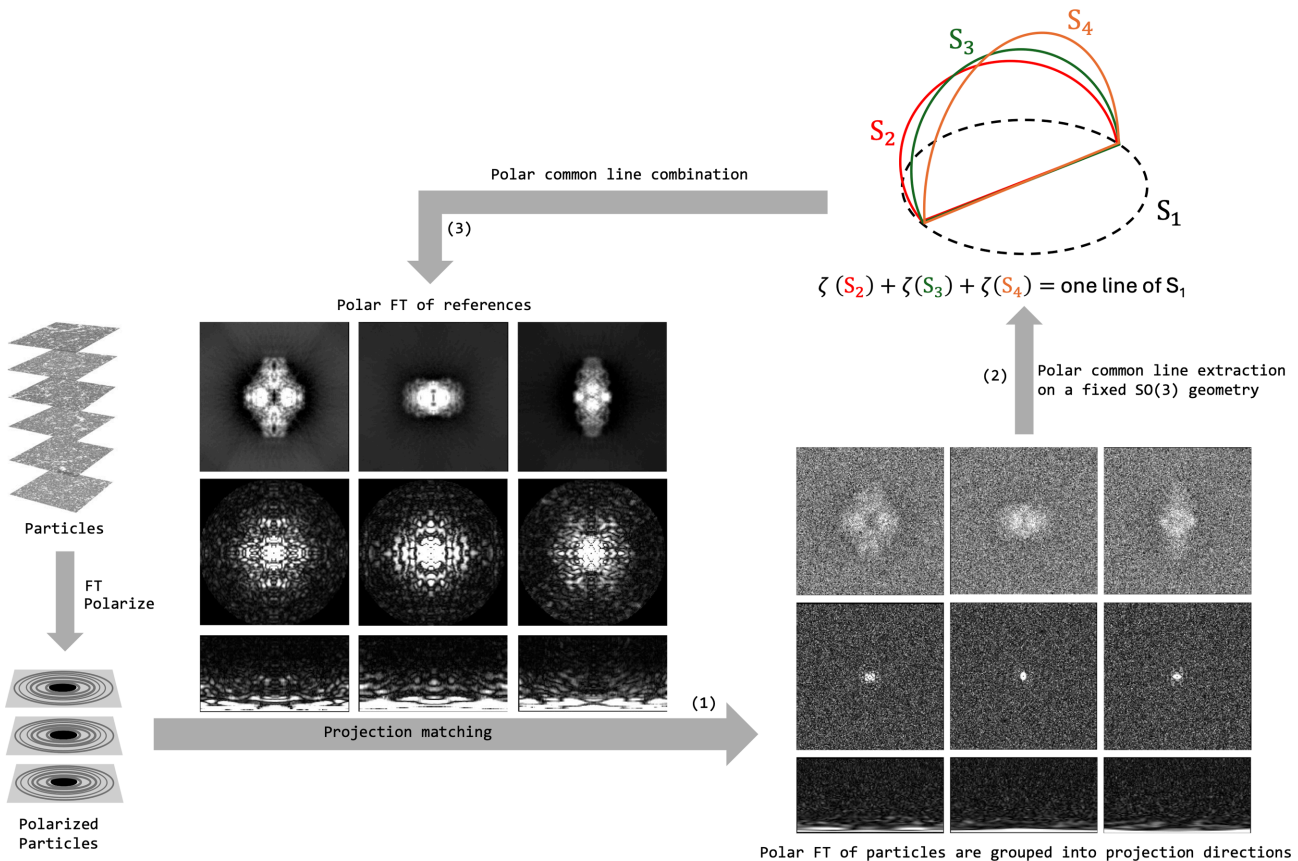
### 2.7. Polar Fourier *ab initio* 3D reconstruction algorithm

We use the same optimization scheme as in our recently published *ab initio* 3D reconstruction algorithm based on probabilistic projection matching, described in detail by Van *et al.* (2025). The search is divided into eight stages (see Table 1). The degree of down sampling decreases with increasing stage number. The initial 3D orientations subjected to optimization

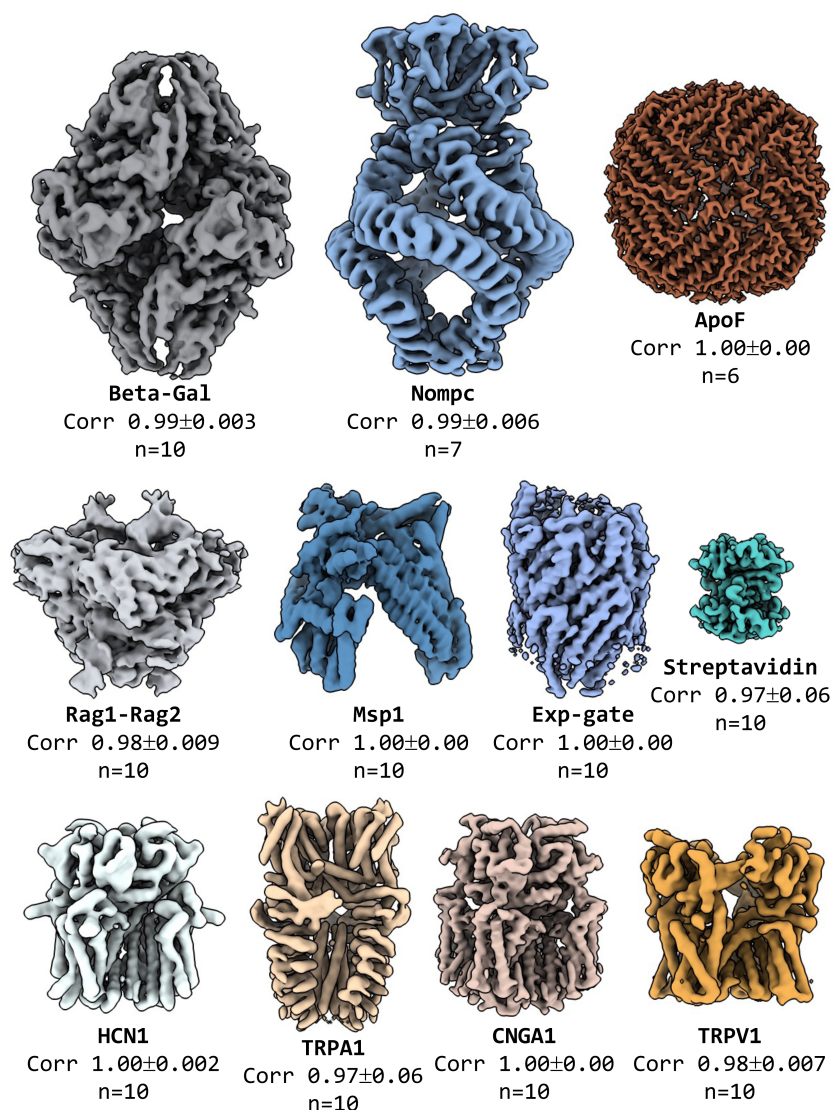
are generated randomly. In the first two stages, regularization with a Gaussian low-pass filter is applied. In the following stages, maximum-likelihood (ML) regularization is applied to denoise the reprojection model. The regularization is limited to uniform Fourier filtering approaches (Gaussian filter and ML regularization). Non-uniform real-space regularization approaches are not applicable without additional transformation steps, since we operate exclusively within the polar Fourier domain. The optimization algorithm outlined in equations (14) and (15) is applied in each stage.

### 2.8. Ensemble *ab initio* 3D volume analysis

In the next section, we describe numerous repeated benchmarking runs of our method on a variety of single-particle datasets. To simplify the presentation of our results and to avoid having to show all the reconstructed density maps obtained from a given dataset, we implemented a tool for analysis of an ensemble of 3D volumes. The 3D orientation of an *ab initio* 3D map as well as its absolute hand (Rosenthal & Henderson, 2003) is arbitrary. Our method for 3D registration of arbitrarily oriented 3D density maps with unknown handedness relies on pairwise *ab initio* docking of all possible pairs of inputted volumes in their two respective hands, followed by calculation of a correlation-based metric of similarity, calcu-



**Figure 2**  
Schematic summary of our method for polar Fourier *ab initio* 3D reconstruction. Particles are Fourier transformed and converted to polar coordinates. At each iteration of expectation maximization, polar reprojection estimates are generated from the particles using the discrete SO(3) configuration obtained in the previous iteration. No 3D volume is needed or computed during the iterative process. A final 3D volume in Cartesian coordinates is reconstructed using the final relative 3D orientations assigned to the particles.



**Figure 3** Three-dimensional maps obtained with the discrete polar Fourier *ab initio* 3D reconstruction approach. Number of independent repeats ( $n$ ) are indicated together with the average correlation to the medoid of the ensemble, calculated to a resolution of 6 Å, and its standard deviation.

lated in the resolution range [6, 100] Å, and stored in a matrix. Next, we analyze all pairwise correlation-based scores to determine which of the 3D volumes shows the best agreement to all the other ones. The identified medoid volume is the representative volume of the ensemble – the one with the largest sum of similarities to all other members. No averaging of volumes is done. Instead, the volumes of the ensemble are compared with the medoid volume, and possible outliers are identified through analysis of the correlation-based scores.

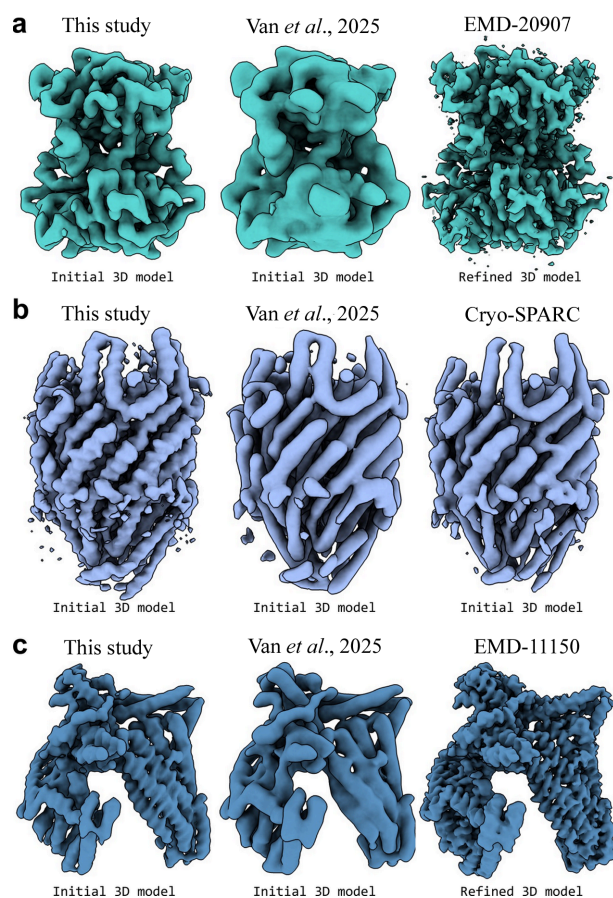
### 3. Results

#### 3.1. *Ab initio* 3D map quality assessment

We sought to validate the ability of our algorithm to recover approximate relative particle 3D orientations and to reconstruct *ab initio* 3D volumes suitable for subsequent high-resolution refinement. Fig. 3 and Table S1 of the supporting

information summarize the results for all analyzed datasets. Following particle import, 2D clustering was performed using the *SIMPLE* program *abinitio2D*, after which classes were manually selected. Table S1 reports, for each dataset, the number of retained particles used for 3D analysis, the success rate of the repeated *ab initio* 3D reconstruction runs and the resolution obtained.

The 2D analysis stage in *SIMPLE* (Caesar *et al.*, 2020) is a critical component of the workflow. Information extracted at this stage determines the degree of downsampling, the low-pass filter cutoff and the stochastic particle-update sampling used in the subsequent *ab initio* 3D reconstruction step, as previously described (Van *et al.*, 2025). During the *ab initio* 3D stage, particles were masked with a soft-edged spherical mask prior to computation of polar Fourier transforms. No *B*-factor sharpening or additional masking was applied to the reconstructed volumes derived from 3D orientations obtained using the polar Fourier *ab initio* method.



**Figure 4**

Comparison of *ab initio* 3D reconstructions. (a) 52 kDa streptavidin reconstructed using the method described in this study (left), volume-based probabilistic projection matching (middle) (Van *et al.*, 2025) and a map refined in *RELION* for comparison (right). (b) Asymmetric export gate reconstructed using the method described in this study (left), volume-based probabilistic projection matching (middle) (Van *et al.*, 2025) and an *ab initio* model generated with *cryoSPARC* for comparison (right). (c) Asymmetric main conformation of merozoite surface protein 1 reconstructed using the method described in this study (left), volume-based probabilistic projection matching (middle) (Van *et al.*, 2025) and a map refined in *cryoSPARC* for comparison (right).

Resolution estimates reported in Table S1 were computed using the  $FSC = 0.143$  criterion and are expected to underestimate the resolution relative to volume-based Fourier shell correlation (FSC) estimates that employ envelope masking. Direct comparison with maps generated using volume-based probabilistic projection matching in our recent study (Van *et al.*, 2025) demonstrates equivalent or improved map quality for the polar Fourier *ab initio* approach. The most significant improvements are observed for lower molecular weight and lower symmetry targets, like streptavidin [Fig. 4(a)], export gate [Fig. 4(b)] and msp1 [Fig. 4(c)]. Critical methodological differences between the two approaches, in addition to the reprojection model, include the absence of envelope masking, the lack of real-space regularization and the removal of directional bias in the polar Fourier *ab initio* approach. The improved map quality is likely attributable to excessive real-space smoothing in the volume-based method (Van *et al.*,

2025), which was designed to generate an initial 3D model. In future work, we will re-design the non-uniform regularization method used in the volume-based approach and investigate opportunities for hybridizing the two methods.

### 3.2. Performance benchmarking

*SIMPLE* uses a stratified parallel execution model that is applied consistently across deployment environments. Fine-grained shared-memory parallelism (OpenMP) is implemented inside numerical kernels and domain operations, while coarse-grained multi-process parallelism is confined to application-level commanders that partition work, launch workers and aggregate results. Importantly, the same commander/worker model can be executed either (i) on high-performance computing systems using an abstracted job-queue interface (e.g. *SLURM*, *PBS*, *LSF*) or (ii) on a single node – such as a standard PC or Mac workstation or even a laptop – using a local backend that launches multiple worker processes directly. In all cases, the scientific workflow and control logic are identical; only the mechanism used to start worker processes differs.

The volume-based and volume-free 3D reconstruction applications considered here consist of three phases: probability-table generation and pose determination with partial reconstruction are done in distributed memory, whereas global aggregation/reduction of the distributed parts is done in shared memory (Van *et al.*, 2025). Both distributed-memory phases are embarrassingly parallel over the particle stack: the dataset is split into balanced subsets, each subset is processed independently by a worker process (using OpenMP threads within that process), and the outputs are then reduced into global structures used by subsequent steps. To quantify end-to-end performance, we measured the global wall-clock breakdown per iteration into these three phases, as well as the corresponding total time.

In the first stochastic search phase, each particle is processed independently, whereas in the second probabilistic phase, orientation assignments are coupled: assigning a particle to a given reference affects the probability landscape for all references (Van *et al.*, 2025). That coupling is what makes the method a truly global optimization scheme. Across all phases, most of the compute is spent on matching particles with the reprojection model. The second major bottleneck in the volume-based approach is the reconstruction and volume reprojection stages, which are considered jointly as ‘reprojection estimation’. The major performance advantage of the volume-free approach is in the reconstruction and volume reprojection stages: the 3D Fourier gridding and volume reprojection steps are replaced by a reprojection estimation directly on the polar grid. It is therefore of interest to compare directly the ‘reprojection estimation’ contribution with the total per-iteration wall-clock time, since the compute due to matching projections is identical between the two approaches.

In the msp1 benchmark [Fig. 5(a)], reprojection estimation in the volume-based approach corresponds to 17% of the total per-iteration wall-clock time on average, whereas the volume-

**Table 2**

The total number of iterations ( $N_{iters}$ ) and total runtimes in hours and minutes for representative datasets.

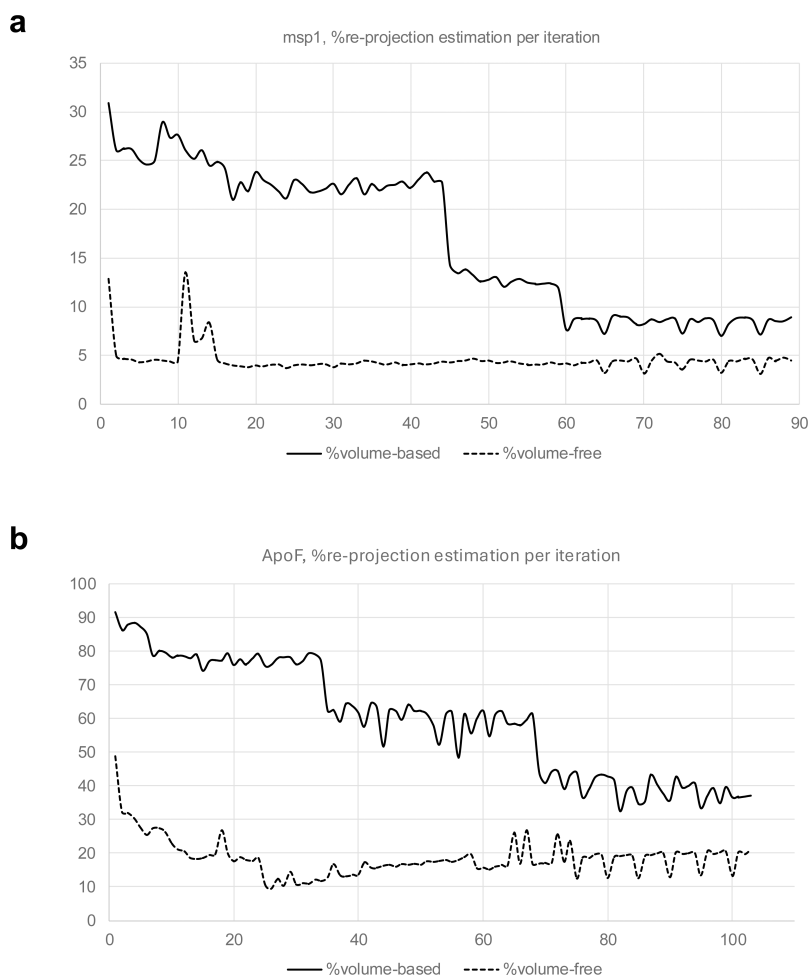
$N_{sample}$  is the number of particles subjected to update in each iteration according to our previously described importance sampling scheme (Van *et al.*, 2025).

Dataset	$N_{particles}$	$N_{sample}$	$V_{based}, N_{iters}$	$V_{free}, N_{iters}$	$V_{based}, total\ time\ (h, m)$	$V_{free}, total\ time\ (h, m)$
Streptavidin	8431	All	$111 \pm 6$	$107 \pm 8$	$0h26m \pm 2m$	$0h25m \pm 4m$
Msp1	112349	10000	$89 \pm 1$	$91 \pm 3$	$3h31m \pm 1m$	$2h58m \pm 59m$
Export gate	128994	5000	$100 \pm 2$	$113 \pm 0$	$1h44m \pm 1m$	$1h57m \pm 1m$
ApoF	200000	5000	$113 \pm 0$	$104 \pm 0$	$4h15m \pm 1m$	$3h33m \pm 10m$
CNGA1	225285	5000	$94 \pm 3$	$93 \pm 1$	$3h6m \pm 3m$	$2h51 \pm 1m$

free approach only uses 5%. Overall execution times and number of iterations needed for convergence are presented in Table 2, indicating a best-case speedup of  $\sim 1.8$  and a worst-case speedup of  $\sim 0.9$  (a slow-down) for msp1. In the ApoF benchmark [Fig. 5(b)], reprojection estimation in the volume-based approach corresponds to 58% of the total per-iteration wall-clock time on average, whereas the volume-free approach only uses 19%. The 24 symmetry operations of the octahedral group increase the cost of reprojection estimation in ApoF relative to asymmetric datasets such as msp1. The overall speedup measured for the ApoF benchmark is  $1.2 \pm 0.06$ .

Overall performance gains correlate with the fraction of runtime spent on reprojection estimation in the volume-based approach; phases where projection matching dominates show limited speedup. Total execution times for representative datasets are summarized in Table 2.

All performance benchmarks were performed with a minimum of five repeats on a dual-socket Intel Xeon Gold 6242R workstation ( $2 \times 20$  cores, 2 hardware threads per core; 80 logical CPUs total; 3.10 GHz base frequency, 4.10 GHz Turbo; x86\_64 architecture) with 252 GB RAM. The system was configured with two NUMA nodes. Hyper-threading was



**Figure 5**

Fraction of per-iteration wall-clock time spent on reprojection estimation. Percentage of per-iteration wall-clock time attributed to reprojection estimation for the volume-based (solid) and volume-free (dashed) approaches. Reprojection estimation corresponds to 3D Fourier gridding and volume reprojection in the volume-based approach, and to direct estimation in the polar Fourier domain in the volume-free approach. (a) Msp1. (b) ApoF.

enabled and all 80 logical CPUs were utilized. The code was compiled using *GCC 15.2.0* with optimization flags `-O3 -march = native -funroll-loops -fPIC`. Ten distributed-memory partitions were used, each running eight OpenMP shared-memory threads in parallel. At most, 36 GB of RAM was utilized at any one time across all runs.

Although reprojection estimation with common lines is substantially faster than 3D Fourier gridding, acceleration of the full application is not guaranteed. For example, in the export gate and CNGA1 benchmarks the time spent on matching projections dominates and no speedup is observed. Our results demonstrate that while reprojection estimation can be significantly accelerated, the overall performance gain is limited by the projection matching, which dominates total runtime. Future optimization efforts will focus on accelerating the projection matching and exploring hybrid strategies that combine volume-based and volume-free updates to enhance convergence robustness.

#### 4. Discussion

Efforts to detect common lines in large collections of noisy projection images typically lead to challenging optimization problems. Prior work approached this through combinatorial search, eigenvector methods and semidefinite programming, among other techniques (Singer & Shkolnisky, 2011; Singer *et al.*, 2010; Elmlund *et al.*, 2010; Elmlund & Elmlund, 2012; Muller *et al.*, 2024; Liu *et al.*, 2024; Herman & Kalinowski, 2008; Wang *et al.*, 2024). The approach we introduce here instead directly couples the manifold structure of the rotation group  $SO(3)$  (Agricola *et al.*, 2011) with the projection-slice theorem (Bracewell, 1956) through the polar Fourier transform (Baddour, 2019).

Rather than identifying common lines explicitly, we discretize  $SO(3)$  and represent each particle in a form that maps naturally onto this discretized rotational manifold, requiring no voting, ranking, dimensionality reduction, algebraic constraints or spherical embeddings. The resulting compact mathematical model provides a direct correspondence between particle orientations and the discretized rotation group, enabling the generation of reprojection estimates for arbitrary orientation configurations by averaging along lines in the discretized polar Fourier domain. A consequence of operating exclusively in reciprocal space is that real-space priors – such as solvent masking (Chen *et al.*, 2013), neural-network-based denoising (Kimanius *et al.*, 2024) or non-uniform regularization (Punjani *et al.*, 2020) – cannot be applied directly without additional transformation steps.

The ability of our method to produce high-quality *ab initio* 3D maps without any volume masking – beyond the spherical particle mask applied prior to polar Fourier transformation – is important for future integration into automated cryo-EM structure-determination pipelines. The resulting *ab initio* map quality distinguishes our approach from previous common-lines-based methods, which typically yield only low-resolution initial models. Moreover, the robustness of the optimization procedure is demonstrated by a single failed run on the

52 kDa streptavidin dataset out of 103 total runs across all datasets. This suggests that the optimization strategy deployed is well suited to the problem, although future work could explore whether slightly weaker but computationally faster optimization schemes might offer favorable trade-offs.

Both the present study and our recent work on probabilistic projection matching (Van *et al.*, 2025) build on the Bayesian framework for single-particle reconstruction established in *RELION* (Scheres, 2012a,b). That said, our implementations differ from a strict Bayesian derivation. We employ a renormalized likelihood objective to improve numerical conditioning, we replace weighted orientation marginalization with probabilistic hard-assignment optimization and we deploy importance sampling guided by the preceding 2D analysis step for accelerated 3D orientation search (Van *et al.*, 2025). These are deliberate design choices aimed at improving numerical stability, efficiency, convergence behavior and computational performance within the *SIMPLE* software environment.

Our overarching objective is to enable real-time cryo-EM image processing on commodity CPU hardware, providing statistical feedback about data quality during acquisition. This imposes practical constraints that can favor computational efficiency over strict Bayesian formalism. Our methods are therefore best understood as pragmatic adaptations that retain the core statistical principles of the Bayesian framework while restructuring the geometry and computational organization of the orientation-recovery problem to meet our engineering and performance objectives. The key distinction of our common-lines approach lies not in the statistical objective but in the computational representation of the projection and reconstruction operators.

There are several promising directions for further development of this methodology. One important avenue is to evaluate how the wide range of previously developed optimization strategies for 2D multi-reference alignment perform when augmented with common-lines constraints. Such investigations may give rise to a new class of methods capable of rapidly extracting 3D information at very early stages of data processing. The framework should also be extended to multi-volume *ab initio* 3D reconstruction, in analogy with the recent generalization of probabilistic projection matching (Van *et al.*, 2025). Finally, it remains to be determined whether jointly leveraging common-lines constraints and volume-based projection matching can improve the robustness of 3D orientation refinement for challenging datasets.

#### 5. Conclusions

Our discrete polar Fourier formulation for single-particle 3D reconstruction is designed to meet the computational requirements of on-the-fly image processing during microscope data acquisition. In future work, we will implement and benchmark real-time versions of this algorithm to further enhance the live cryo-EM structure-determination capabilities of our *SIMPLE* software suite (Caesar *et al.*, 2020). While implemented within the *SIMPLE* framework, the formulation

is general and applicable to any projection-matching or probabilistic orientation-recovery scheme.

*SIMPLE* is available for download at <https://github.com/hael/SIMPLE>.

## Acknowledgements

The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services. We thank Dr Susan M. Lea for assistance preparing the manuscript. Author contributions are as follows: conception/design of the work, CTSV, HE; software design, CTSV, CFR, JJEC, RM-P, HE. All authors contributed to analysis/interpretation of data/results and writing of the manuscript.

## Funding information

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH).

## References

Agricola, I., Becker-Bender, J. & Friedrich, T. (2011). *Ann. Glob. Anal. Geom.* **40**, 67–84.

Baddour, N. (2019). *Mathematics* **7**, 698.

Baldwin, P. R. & Penczek, P. A. (2007). *J. Struct. Biol.* **157**, 250–261.

Beatty, P. J., Nishimura, D. G. & Pauly, J. M. (2005). *IEEE Trans. Med. Imaging* **24**, 799–808.

Bracewell, R. N. (1956). *Aust. J. Phys.* **9**, 198–217.

Caesar, J., Reboul, C. F., Machello, C., Kiesewetter, S., Tang, M. L., Deme, J. C., Johnson, S., Elmlund, D., Lea, S. M. & Elmlund, H. (2020). *J. Struct. Biol. X* **4**, 100040.

Chen, S., McMullan, G., Faruqi, A. R., Murshudov, G. N., Short, J. M., Scheres, S. H. & Henderson, R. (2013). *Ultramicroscopy* **135**, 24–35.

Crowther, R. A., Amos, L. A., Finch, J. T., De Rosier, D. J. & Klug, A. (1970). *Nature* **226**, 421–425.

De Rosier, D. J. & Klug, A. (1968). *Nature* **217**, 130–134.

Elmlund, D., Davis, R. & Elmlund, H. (2010). *Structure* **18**, 777–786.

Elmlund, D. & Elmlund, H. (2009). *J. Struct. Biol.* **167**, 83–94.

Elmlund, D. & Elmlund, H. (2012). *J. Struct. Biol.* **180**, 420–427.

Elmlund, H., Lundqvist, J., Al-Karadaghi, S., Hansson, M., Hebert, H. & Lindahl, M. (2008). *J. Mol. Biol.* **375**, 934–947.

Fuller, S. D., Butcher, S. J., Cheng, R. H. & Baker, T. S. (1996). *J. Struct. Biol.* **116**, 48–55.

Grigorieff, N. (2007). *J. Struct. Biol.* **157**, 117–125.

Hall, R. J., Siridechadilok, B. & Nogales, E. (2007). *J. Struct. Biol.* **159**, 474–482.

Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. & Downing, K. H. (1990). *Biochem. Soc. Trans.* **18**, 844.

Herman, G. T. & Kalinowski, M. (2008). *Ultramicroscopy* **108**, 327–338.

Jackson, J. I., Meyer, C. H., Nishimura, D. G. & Macovski, A. (1991). *IEEE Trans. Med. Imaging* **10**, 473–478.

Jaitly, N., Brubaker, M. A., Rubinstein, J. L. & Lilien, R. H. (2010). *Bioinformatics* **26**, 2406–2415.

Kimanius, D., Jamali, K., Wilkinson, M. E., Lövestam, S., Velazhahan, V., Nakane, T. & Scheres, S. H. W. (2024). *Nat. Methods* **21**, 1216–1221.

Liu, J. X., Lu, Y. G. & Zhu, L. (2024). *Brief. Bioinform.* **25**, bbad473.

Lyumkis, D., Brilot, A. F., Theobald, D. L. & Grigorieff, N. (2013). *J. Struct. Biol.* **183**, 377–388.

Muller, T., Duncan, A. L., Verbeke, E. J. & Kileel, J. (2024). *Biol. Imaging* **4**, e9.

O’Sullivan, J. D. (1985). *IEEE Trans. Med. Imaging* **4**, 200–207.

Penczek, P. A. (2010). *Methods Enzymol.* **482**, 1–33.

Penczek, P. A., Renka, R. & Schomberg, H. (2004). *J. Opt. Soc. Am. A* **21**, 499–509.

Penczek, P., Radermacher, M. & Frank, J. (1992). *Ultramicroscopy* **40**, 33–53.

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. (2017). *Nat. Methods* **14**, 290–296.

Punjani, A., Zhang, H. & Fleet, D. J. (2020). *Nat. Methods* **17**, 1214–1221.

Reboul, C. F., Bonnet, F., Elmlund, D. & Elmlund, H. (2016). *Structure* **24**, 988–996.

Reuter, M. & Fischl, B. (2011). *NeuroImage* **57**, 19–21.

Romanov, E., Bendory, T. & Ordentlich, O. (2021). *SIAM J. Math. Data Sci.* **3**, 494–523.

Rosenthal, P. B. & Henderson, R. (2003). *J. Mol. Biol.* **333**, 721–745.

Saxton, W. O. & Baumeister, W. (1982). *J. Microsc.* **127**, 127–138.

Schatz, M., Orlova, E. V., Dube, P., Jäger, J. & van Heel, M. (1995). *J. Struct. Biol.* **114**, 28–40.

Scheres, S. H. (2012a). *J. Struct. Biol.* **180**, 519–530.

Scheres, S. H. & Chen, S. (2012). *Nat. Methods* **9**, 853–854.

Scheres, S. H., Valle, M. & Carazo, J. M. (2005). *Bioinformatics* **21**, ii243–ii244.

Scheres, S. H. W. (2012b). *J. Mol. Biol.* **415**, 406–418.

Scheres, S. H. W., Valle, M., Nuñez, R., Sorzano, C. O. S., Marabini, R., Herman, G. T. & Carazo, J. M. (2005). *J. Mol. Biol.* **348**, 139–149.

Sigworth, F. J. (1998). *J. Struct. Biol.* **122**, 328–339.

Singer, A., Coifman, R. R., Sigworth, F. J., Chester, D. W. & Shkolnisky, Y. (2010). *J. Struct. Biol.* **169**, 312–322.

Singer, A. & Shkolnisky, Y. (2011). *SIAM J. Imaging Sci.* **4**, 543–572.

Sorzano, C. O. S., Bilbao-Castro, J. R., Shkolnisky, Y., Alcorlo, M., Melero, R., Caffarena-Fernández, G., Li, M., Xu, G., Marabini, R. & Carazo, J. M. (2010). *J. Struct. Biol.* **171**, 197–206.

Sorzano, C. O. S., Jiménez-Moreno, A., Maluenda, D., Martínez, M., Ramírez-Aportela, E., Krieger, J., Melero, R., Cuervo, A., Conesa, J., Filipovic, J., Conesa, P., del Caño, L., Fonseca, Y. C., Jiménez-de la Morena, J., Losana, P., Sánchez-García, R., Strelak, D., Fernández-Giménez, E., de Isidro-Gómez, F. P., Herreros, D., Vilas, J. L., Marabini, R. & Carazo, J. M. (2022). *Acta Cryst. D* **78**, 410–423.

Stark, H., Mueller, F., Orlova, E. V., Schatz, M., Dube, P., Erdemir, T., Zemlin, F., Brimacombe, R. & van Heel, M. (1995). *Structure* **3**, 815–821.

Unser, M., Sorzano, C. O. S., Thévenaz, P., Jonić, S., El-Bez, C., De Carlo, S., Conway, J. F. & Trus, B. L. (2005). *J. Struct. Biol.* **149**, 243–255.

Van, C. T. S., Reboul, C. F., Caesar, J. J. E., Meana-Pañeda, R., Lountos, G. T., Deme, J. C., Bryant, O. J., Johnson, S., Piczak, C. T., Valkov, E., Lea, S. M. & Elmlund, H. (2025). *Acta Cryst. D* **81**, 396–409.

Van Heel, M. (1987). *Ultramicroscopy* **21**, 111–123.

Wang, X., Jin, Q., Zou, L., Lin, X. & Lu, Y. (2024). *IEEE/ACM Trans. Comput. Biol. Bioinf.* **21**, 2496–2509.

Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. (2012). *Structure* **20**, 237–247.

Yang, Z. & Penczek, P. A. (2008). *Ultramicroscopy* **108**, 959–969.

Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. (2021). *Nat. Methods* **18**, 176–185.

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. & Scheres, S. H. (2018). *eLife* **7**, e42166.