

To be published in Journal of the Optical Society of America A:

Title: Plankton Reconstruction through Robust Statistical Optical Tomography

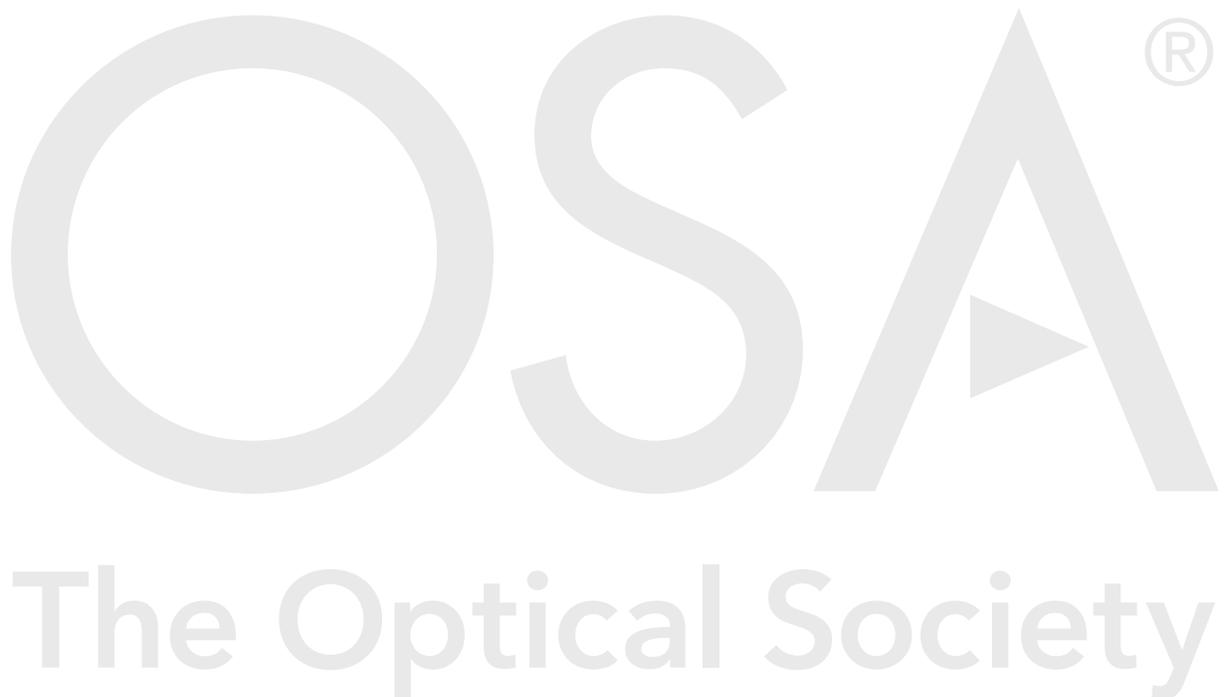
Authors: roi ronon, Yacov Attias, Yoav Schechner, Jules Jaffe, Eric Orenstein

Accepted: 07 July 21

Posted 12 July 21

DOI: <https://doi.org/10.1364/JOSAA.423037>

© 2021 Optical Society of America



Plankton Reconstruction through Robust Statistical Optical Tomography

ROI RONEN^{1,*}, YACOV ATTIAS¹, YOAV Y. SCHECHNER¹, JULES S. JAFFE², AND ERIC ORENSTEIN²

¹Viterbi Faculty of Electrical Eng., Technion, Haifa, Israel

²Scripps Institution of Oceanography, University of California, San Diego

*roironen@campus.technion.ac.il

Compiled July 7, 2021

Plankton interact with the environment according to their size and 3D structure. To study them outdoors, these translucent specimens are imaged *in situ*. Light projects through a specimen in each image. The specimen has a random scale, drawn from the population's size distribution and random unknown pose. The specimen appears only once before drifting away. We achieve 3D tomography using such a random ensemble, to statistically estimate an average volumetric distribution of the plankton type and specimen size. To counter errors due to non-rigid deformations, we weight the data, drawing from advanced models developed for Cryo-Electron Microscopy. The weights convey the confidence in the quality of each datum. This confidence relies on a statistical error model. We demonstrate the approach on live plankton using an underwater field microscope. © 2021 Optical Society of America

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

1. INTRODUCTION

It is important to study plankton for their role in both local and global ecosystems. These organisms are vital to the ecology of the planet [1], as they play a critical role in producing breathable oxygen (50% due to marine phytoplankton) and serve as the base of the marine food web. More recently, an increase in harmful algal blooms has been devastating to marine and fresh-water supplies [2, 3], while having a negative impact on coastal ecosystems.

Over decades, the basic taxonomy of plankton has been established both by morphological observations and by DNA studies that allow unique attribution. Modern oceanographers increasingly rely on *in situ* imaging systems to study populations in their natural habitat. However, almost all available inferences about plankton morphology derived from these instruments rely on two-dimensional (2D) projections [4–8] with the third dimension being conjectured. The determination of three-dimensional (3D) structure is important as it defines the *surface-area-to-volume ratio*, which is fundamental to the absorption and diffusion of nutrients and waste products. Also, flow is linked to this ratio, through the Reynolds number [9]. In addition to the 3D structure, it is important to assess the size distribution of the population (Fig. 1).

Recent advances in 3D imaging focus on lab systems [10–12]. However, many plankton are easily deformed by collection and are thus difficult to study in the lab. There are therefore a host of ocean and lake systems [13–15], to observe these organisms *in situ*. These systems look outward from their housing into

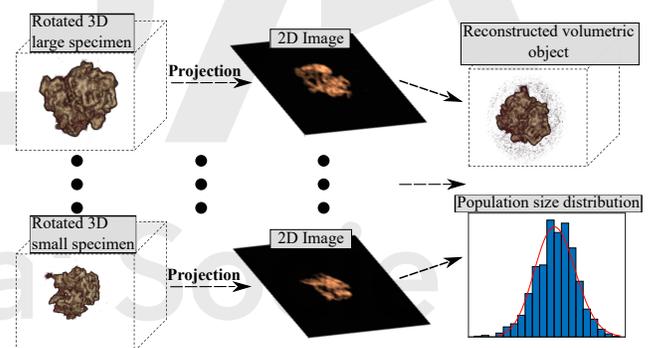


Fig. 1. Specimens viewed in nature have random sizes and poses. Light passes through, yielding a single 2D image per specimen. The goal is to tomographically recover the 3D volumetric distribution of the plankton (on average), and estimate the population size distribution.

the surrounding water. In the majority of systems, imaging is accomplished via the presence of organisms that simply pass through their field of view.

In order to image in relatively high resolution small plankton (<100 micro-meters), there are four kinds of common *in situ* imaging systems that can be submerged and operated [16, 17]. (i) The Imaging FlowCytobot (IFCB) [15] uses either hydrodynamic or acoustic particle focusing. This results in particles

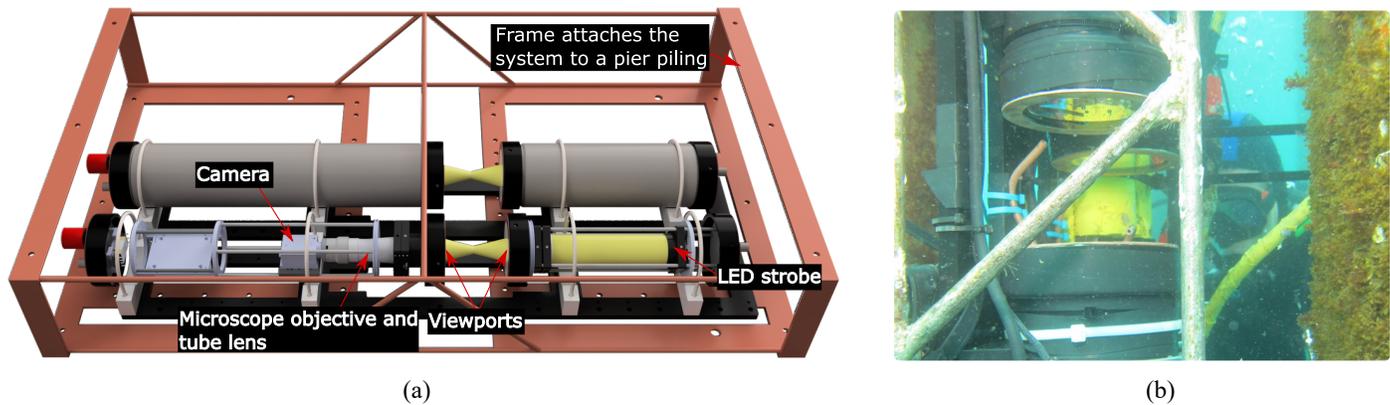


Fig. 2. (a) An illustration of the Scripps Plankton Camera (SPC). (b) An image of the SPC *in situ*. The darkfield microscope's sample volume is between the ports.

being positioned in the volume in focus, flowing through a chamber. In cases of more spherical particles, IFCB captures a single shot of a randomly sized organism [18] at a random pose. Other systems of relevance to our study are based either on (ii) holography [19–21], (iii) shadowgraph illumination [7, 22] or (iv) darkfield microscopy.

Holography and shadowgraph imaging are viable modalities. However, we found several advantages to darkfield microscopy [23]: (a) Darkfield imaging is common, low cost and robust; (b) Our experience with on-axis holography indicates that the point spread function is range-dependent (a quadratic phase effect) that generally yields lower contrast than darkfield. We also found, in the case of holography, difficulty in yielding viable depth information at the thicknesses of the organisms that we study. A darkfield image is a simple projection rather than a phase measurement. Projected images can provide information that can yield 3D structure using tomography, as we describe in this paper; (c) Small scattering particles in turbid waters, as are commonly found in coastal regions, yield random localized specks in darkfield images. This randomness can be filtered-out by tomography (which is based on multiple images). On-axis holography (the practical basis for digital holography) responds to scatterers by non-local speckled patterns, which reduce overall contrast; (d) Darkfield illumination enhances edges of objects whose refractive index is close to that of the surrounding medium. This leads to better resolved appendages and internal structures of plankton. We thus chose to work with darkfield optical microscopes to maximize taxonomic resolution, enhance edge contrast and ensure high performance in scattering-rich environments.

Underwater computer vision has benefited from a dramatic increase in computing capabilities, addressing image quality [24–28] and underwater robotics [29–34]. It is fitting that the benefits of computational imaging should be applied to *in situ* study of plankton. Tomography is a computational method that we use. Usually, tomography is performed around a single, temporally fixed object [35–39]. Here we are interested in drawing structure and statistics about a *population*. The data is an ensemble of random views of random objects within a population, each having a different random and unknown scale, orientation and location (See Fig. 1). A basis to our solution is *Single Particle Reconstruction* (SPR) [40], which is used in Cryo-Electron Microscopy (Cryo-EM) [41, 42]. SPR handles random rotations but not scale variation in a population. An attempt to generalize SPR

to handle variable scales was done in [43], where the approach termed 3D-POP is defined. It is based on pairwise constraints on rotations and scales. However, the work in [43] is prone to significant estimation errors. Hence, our paper employs a more robust SPR approach [44] and generalizes it to analyze data subject to scale variations. The proposed method uses a set of *triplets* of images to assess how indicative are constraints derived from image pairs. Then, this assessment is used to weight each constraint. The constraints are then efficiently and optimally fused using linear algebra.

We show the effectiveness of the method on simulated data and publically available brightfield images from the Imaging FlowCytobot (IFCB). We also show it using darkfield images of plankton acquired by our *in situ* system: the Scripps Plankton Camera (SPC). SPC is an *in situ* darkfield microscope developed to continuously monitor plankton populations [45] (Fig. 2). It is a free-space imaging system, with the light and camera focused on a plane in an open volume of fluid. The SPC uses a white LED illumination source with the central rays blocked by an opaque disk. Objects in the focal plane thus scatter light toward the image plane while the direct light continues past the lens, thus creating the darkfield effect. The paper uses images from two test SPC datasets. Both were taken by a version of the SPC outfitted with a 0.5x microscope objective, yielding an object sampling interval of $7.4 \mu\text{m}/\text{pixel}$ with adequate contrast at $30 \mu\text{m}$. All presented SPC plankton samples are about 0.5–2.5mm long. The object sampling interval of the IFCB is $0.29 \mu\text{m}/\text{pixel}$, enabling imaging of objects in an order of magnitude smaller than the SPC. All plankton samples were manually classified by an expert taxonomist.

2. BACKGROUND ON SINGLE PARTICLE RECONSTRUCTION

A. Tomography Using Random Projections

We denote 3D spatial coordinates by $\mathbf{x}=(x_1, x_2, x_3)^T$, where $(\cdot)^T$ denotes transposition. A specimen has a volumetric optical extinction distribution $\beta(\mathbf{x})$. In darkfield imaging, $\beta(\mathbf{x})$ expresses the density of scattering particles in the object. In brightfield imaging, $\beta(\mathbf{x})$ represents the optical density of attenuating media.

Computational tomography (CT) estimates $\beta(\mathbf{x})$ using a set of N multi-view orthographic projections. Relative to the camera, the sample orientation in the n th image \mathbf{y}_n is expressed by a

3D rotation matrix \mathbf{R}_n . Orthographic projection through the translucent sample is expressed by linear integration along the x_3 axis

$$\mathbf{y}_n = \mathcal{P}_{x_1, x_2}^{(n)} \{\beta\} = \int \beta(\mathbf{R}_n^{-1} \mathbf{x}) dx_3, \quad (1)$$

where x_1, x_2 are image axes. Organisms have a range of scattering characteristics. For the sake of our analysis, let the objects be weak scatterers, as commonly assumed in microscopy restoration problems [46, 47]. Images are thus approximately compliant with the linear model of Eq. (1).

There are strong and widely used methods for estimating a volumetric distribution $\beta(\mathbf{x})$ using tomography based on such a set of projections $\{\mathbf{y}_n\}_{n=1}^N$. This paper thus does not focus on estimating $\beta(\mathbf{x})$. We note, however, that common tomographic methods [48–55] assume that the set of N orientation matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_N\}$ is known. Moreover, common tomography assumes that the specimen itself does not change between projections.

A random biological microscopic specimen collected *in situ* does not comply with these assumptions. Its orientation is random and unknown. Furthermore, each projection image corresponds to a *different specimen*, hence having random, unknown scale and lateral displacement. This paper focuses on recovering the unknown scale, displacement and rotation of each specimen, in order to facilitate tomographic recovery of the 3D structure.

B. 3D-POP

There is a prior method for estimating the unknown scale, displacement and rotation of each specimen: 3D-POP [43]. Our paper presents methods that yield better efficiency, accuracy and robustness. We briefly describe the main elements of 3D-POP, drawing significantly on its notations.

Analysis is done in the 3D spatial frequency domain $\mathbf{k}=(k_1, k_2, k_3)$. The 3D Fourier transform of the volumetric object is $\mathcal{F}_{\mathbf{k}}\{\beta\}$. A *central slice* is a frequency-plane in the 3D Fourier space that passes through the origin, $\mathbf{k}=0$. Any central slice is a 2D domain, which can be written in terms of a *rotated* horizontal slice,

$$\mathbf{R}_n^{-1}[k_1, k_2, 0]^T, \quad (2)$$

where \mathbf{R}_n is a 3D rotation matrix indexed n , as in Eq. (1). Figure 3 illustrates 3D Fourier space (3D k-space) and two central slices n, m in this space. These slices respectively represent the planar domains $\mathbf{R}_n^{-1}[k_1, k_2, 0]^T$ and $\mathbf{R}_m^{-1}[k_1, k_2, 0]^T$ in 3D k-space.

Now let us consider the 2D Fourier transform of projection image \mathbf{y}_n . Denote the resulting transformed representation by $\tilde{\mathbf{y}}_n(\rho, \psi)$, where ρ, ψ are polar coordinates in the 2D spatial frequency domain. According to the *Central Slice Theorem*, the 2D function $\tilde{\mathbf{y}}_n$ is equivalent to the central slice $\mathbf{R}_n^{-1}[k_1, k_2, 0]^T$ in 3D k-space (Eq. 2).

Any two central slices n, m intersect at a *common line* in 3D k-space (Fig. 3). Correspondingly, this line in 3D k-space is equivalent to a radial line in the 2D function $\tilde{\mathbf{y}}_n$ at an angle denoted by $\psi=\psi_{n \rightarrow m}$. It is also equivalent to a radial line in the 2D function $\tilde{\mathbf{y}}_m$, at an angle denoted by $\psi'=\psi_{m \rightarrow n}$.

Within these 2D functions, finding the angle-pair $\psi_{n \rightarrow m}, \psi_{m \rightarrow n}$ which correspond to a common line yields constraints [42] about the 3D rotations $\mathbf{R}_n, \mathbf{R}_m$. The constraints

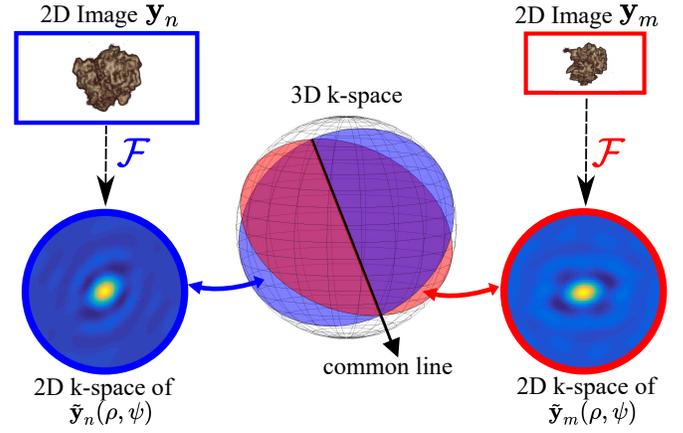


Fig. 3. The 3D Fourier domain, and two central slices $\tilde{\mathbf{y}}_n, \tilde{\mathbf{y}}_m$. Each central slice is perpendicular to the direction-vector of the respective projection in the microscope. The two slices intersect at a *common tilt axis* (common line) in the 3D Fourier domain.

are:

$$\begin{bmatrix} \cos \psi_{n \rightarrow m} \\ \sin \psi_{n \rightarrow m} \\ 0 \end{bmatrix} = \mathbf{R}_n^{-1} \mathbf{R}_m \begin{bmatrix} \cos \psi_{m \rightarrow n} \\ \sin \psi_{m \rightarrow n} \\ 0 \end{bmatrix}. \quad (3)$$

Then there are two challenges:

- Estimating all projection rotations $\{\mathbf{R}_n\}_{n=1}^N$ using constraints as in Eq. (3). This is a key for subsequent tomographic recovery of $\beta(\mathbf{x})$.
- To yield constraint as Eq. (3), the angle-pair $\psi_{n \rightarrow m}, \psi_{m \rightarrow n}$ must be estimated reliably and efficiently for many real-data image-pairs. This challenge is exacerbated by the fact that $\beta(\mathbf{x})$ is randomly shifted and scaled between images of different organisms collected *in situ*.

Let us start with challenge b. In the 2D Fourier domain, the 1-D function on a radial line $\tilde{\mathbf{y}}_n(\rho, \psi_{n \rightarrow m})$ is expected to highly correlate with the radial function $\tilde{\mathbf{y}}_m(\rho, \psi_{m \rightarrow n})$ since both correspond to the same common line. The correlation should generalize to account for a relative scale $M_{n \rightarrow m}$ between the specimens viewed in images n and m . Relative spatial scale is expressed as an inverse-scale in the Fourier domain. The correlation should also generalize to account for a lateral shift $\tau_{n \rightarrow m}$ of the specimen along the common line direction. This shift yields a linear phase $2\pi\rho\tau_{n \rightarrow m}$ in the Fourier domain. Consequently, in Ref. [43] the Fourier-domain radial correlation is defined by an inner product of scale and phase-shifted 1D functions sampled along two directions, ψ, ψ' , using variable scale M and lateral shift τ :

$$\Sigma_{nm}(\psi, \psi', M, \tau) = \int_{-\rho_{\max}}^{\rho_{\max}} \tilde{\mathbf{y}}_n^* \left(\frac{\rho}{M}, \psi \right) \tilde{\mathbf{y}}_m(\rho, \psi') e^{2\pi i \rho \tau} d\rho. \quad (4)$$

Here $[-\rho_{\max}, \rho_{\max}]$ is the effective radial domain of the functions, and $i = \sqrt{-1}$. Ref. [43] estimates the angle-pair, relative scale and relative shift by

$$\{\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}, \hat{M}_{n \rightarrow m}, \hat{\tau}_{n \rightarrow m}\} = \arg \max_{\psi, \psi', M, \tau} \frac{\Sigma_{nm}(\psi, \psi', M, \tau)}{\sqrt{\Sigma_{nn}(\psi, \psi, M, \tau)} \sqrt{\Sigma_{mm}(\psi', \psi', M, \tau)}}. \quad (5)$$

We now describe challenge a. Eqs. (3,5) yield pairwise relations on relative rotations, scales and translations. Using a large number of images, these relations are then compounded to yield overall solutions to pose and scale of each specimen in a global coordinate system. Ref. [43] found $\hat{M}_{n \rightarrow m}$ empirically in the pairwise optimization (5). The unknown scale of specimen n is M_n . Denote $\bar{M}_n = \log(M_n)$. Then $\bar{M}_n - \bar{M}_m = \log(\hat{M}_{n \rightarrow m})$,

Consequently, a data term that penalizes scale inconsistencies is

$$\mathcal{D}^{\text{scale}} \left(\{\bar{M}_n\}_{n=1}^N \right) = \sum_{n \neq m} \left| \bar{M}_n - \bar{M}_m - \log \hat{M}_{n \rightarrow m} \right|^2. \quad (6)$$

Adding a simple regularization term $\mathcal{R}^{\text{scale}}$, Ref. [43] then solves linearly the optimization problem

$$\{\hat{M}_n\}_{n=1}^N = \arg \min_{\bar{M}_n} \left[\mathcal{D}^{\text{scale}} (\{\bar{M}_n\}) + \mathcal{R}^{\text{scale}} (\{\bar{M}_n\}) \right]. \quad (7)$$

The scale of specimen n is $\hat{M}_n = \exp \hat{M}_n$. A similar approach optimizes the translations of all specimens [43], using $\hat{t}_{n \rightarrow m}$ that were obtained empirically in the pairwise optimization (5).

Problems that persist

There are several problems with Eqs. (4,5) and their implementation. In [43], optimization (5) was done by exhaustive search over a 4-dimensional (4D) domain of variables. To make the search space manageable, Ref. [43] used a coarse-to-fine approach starting from a sparse sampling of the domain of variables. Maximization (5) is performed on this small sparse set. This results in a narrower 4D domain, to be searched in finer resolution. However, noise and other factors may lead to errors in (5). An error occurring in a coarse-sampling stage can lead to intolerable errors which cannot be solved in later, finer-scale sampling. In effect, robustness and accuracy are sacrificed for numerical tractability.

The model in Eqs. (4,5) assumes a similarity relation between specimens. It does not account for non-rigid deformations, which are common in natural non-rigid objects in a population. A non-rigid deformation is a perturbation that may lead to deviation from the estimated pose variables. It may also lead to an inaccurate solution of Eqs. (5): by chance, correlation of an arbitrary orientation scale and translation can yield better correlation (4) than the correct pose due to the deformation. Hence, a major error in pose propagates to major errors in tomographic recovery. In this paper, we use a more robust approach.

The biggest challenge is that of recovering the orientation. Based on the pairwise relations as in (3,5), all rotations were estimated in [43] using a computationally complex, non-convex, constrained optimization as follows:

$$\begin{aligned} \{\hat{\mathbf{R}}_n\}_{n=1}^N = \\ \arg \min_{\mathbf{R}_1, \dots, \mathbf{R}_N} \sum_{n \neq m} \left\| \mathbf{R}_n \begin{bmatrix} \cos \psi_{n \rightarrow m} \\ \sin \psi_{n \rightarrow m} \\ 0 \end{bmatrix} - \mathbf{R}_m \begin{bmatrix} \cos \psi_{m \rightarrow n} \\ \sin \psi_{m \rightarrow n} \\ 0 \end{bmatrix} \right\|_2^2 \\ \text{s.t. } \mathbf{R}_n \mathbf{R}_n^T = \mathbf{I}, \det(\mathbf{R}_n) = 1, n = 1, \dots, N. \end{aligned} \quad (8)$$

Here \mathbf{I} is a 3×3 identity matrix [42].

3. STATISTICAL MODEL FOR ROBUST PLANKTON RECONSTRUCTION

A. Estimating Rotations

To mitigate this complexity, we describe a numerically efficient method that solves $\{\hat{\mathbf{R}}_n\}_{n=1}^N$ given pairwise relations. Contrary to Eq. (8), the method here is not formulated as a non-convex problem or using approximate relaxation. Instead, it uses a linear algebra approach [44] to achieve accurate results quickly. Define

$$\mathbf{R}_{m \rightarrow n} \equiv \mathbf{R}_n \mathbf{R}_m^{-1}. \quad (9)$$

Use the angle pair $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ and a matrix $\mathbf{R}_{m \rightarrow n}$ which satisfies Eqs. (3,9) as found as in [56]. Then, define matrices

$$\mathbf{S} = \begin{bmatrix} \mathbf{I} & \mathbf{R}_{1 \rightarrow 2} & \dots & \mathbf{R}_{1 \rightarrow n} & \dots & \mathbf{R}_{1 \rightarrow N} \\ \mathbf{R}_{2 \rightarrow 1} & \mathbf{I} & \dots & \mathbf{R}_{2 \rightarrow n} & \dots & \mathbf{R}_{2 \rightarrow N} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ \mathbf{R}_{m \rightarrow 1} & \mathbf{R}_{m \rightarrow 2} & \dots & \mathbf{R}_{m \rightarrow n} & \dots & \mathbf{R}_{m \rightarrow N} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ \mathbf{R}_{N \rightarrow 1} & \mathbf{R}_{N \rightarrow 2} & \dots & \mathbf{R}_{N \rightarrow n} & \dots & \mathbf{I} \end{bmatrix} \quad (10)$$

and

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_N \end{bmatrix} \quad (11)$$

Denote the n -th block of a $3N \times 3$ matrix by $(\cdot)_n$ (i.e. rows numbered $3n-2, 3n-1, 3n$ in the matrix). Then from (9),

$$(\mathbf{S}\mathbf{H})_n = \sum_{m=1}^N \mathbf{R}_{m \rightarrow n} \mathbf{R}_m = \mathbf{N}\mathbf{R}_n = \mathbf{N}(\mathbf{H})_n \quad (12)$$

leading to

$$\mathbf{S}\mathbf{H} = \mathbf{N}\mathbf{H}. \quad (13)$$

Therefore, the three columns of \mathbf{H} are eigenvectors of \mathbf{S} . For each of these three eigenvectors, the corresponding eigenvalue is N . Recall that $\mathbf{R}_m^{-1} = \mathbf{R}_m^T$ and thus that $\mathbf{S} = \mathbf{H}\mathbf{H}^T$, which means that \mathbf{S} has rank 3. Therefore, N is the only non-zero eigenvalue of \mathbf{S} .

Using this formulation, we propose an algorithm for empirically estimating the rotations matrices $\{\hat{\mathbf{R}}_n\}_{n=1}^N$.

1. Construct the empirical matrix \mathbf{S} , using the estimated angle pair $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ and Eqs. (3,9,10).
2. Perform singular value decomposition (SVD) on \mathbf{S} . Select the three largest singular values and their corresponding eigenvectors. Denote these column vectors as $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$.
3. Construct the estimated rotations by concatenating these column vectors into a matrix:

$$\begin{bmatrix} \hat{\mathbf{R}}_1 \\ \vdots \\ \hat{\mathbf{R}}_N \end{bmatrix} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]. \quad (14)$$

This approach generalizes to yield more robust results in the presence of errors in the estimated angle pair $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$.

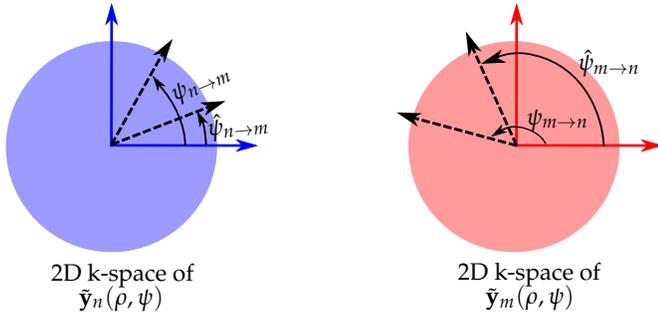


Fig. 4. A central slice in 3D Fourier domain (seen in Fig. 3) is equivalent to a 2D Fourier transform of an acquired image. A common line (also seen in Fig. 3) is at respective corresponding angles $\psi_{n \rightarrow m}, \psi_{m \rightarrow n}$ in the 2D Fourier domains. Noisy data (e.g. due to non-rigid deformation of specimens) leads to erroneous estimation of these angles, $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$.

We describe in Secs. B, C an approach to yield a *confidence weight* $w_{n,m}$ to each $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ pair. A low value of $w_{n,m}$ indicates a low confidence in the correctness of this pair of angles. The weights are normalized so as to satisfy

$$\sum_m w_{n,m} = N, \quad w_{n,m} \geq 0, \quad \forall n, m. \quad (15)$$

These weights propagate to the empirical matrix \mathbf{S} : each block $\mathbf{R}_{m \rightarrow n}$ in \mathbf{S} is replaced by $w_{n,m} \mathbf{R}_{m \rightarrow n}$. This yields an improved matrix $\tilde{\mathbf{S}}$. Similarly to Eq. (12),

$$(\tilde{\mathbf{S}}\mathbf{H})_n = \sum_{m=1}^N w_{n,m} \mathbf{R}_{m \rightarrow n} \mathbf{R}_m = \mathbf{R}_n \sum_{m=1}^N w_{n,m} = N (\mathbf{H})_n. \quad (16)$$

Analogously to Eq. (13), the three columns of \mathbf{H} are eigenvectors of $\tilde{\mathbf{S}}$ with the corresponding eigenvalue N . To estimate the rotations matrices $\{\hat{\mathbf{R}}_n\}_{n=1}^N$, the algorithm above is run using $\tilde{\mathbf{S}}$ instead of \mathbf{S} . The advantage of this method is that common lines having a lower confidence have a smaller effect on the estimation of $\{\hat{\mathbf{R}}_n\}_{n=1}^N$.

B. A Statistical Model of Common Lines

To derive weights $w_{n,m}$, we need to have a measure of confidence or uncertainty for each common line. The uncertainty depends on a probabilistic model of common lines, the noise sources which affect this model and finite sampling effects. From our experience, errors in $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ cause problems that are more severe than problems stemming from errors in scale or shift. Hence we focus our attention on angular errors.

Consider the 2D Fourier representation \tilde{y}_n (Fig. 4). The representation of the true common line with \tilde{y}_m is a line at true angle $\psi_{n \rightarrow m}$. Its empirical estimation is $\hat{\psi}_{n \rightarrow m}$. The corresponding angular error is

$$\Delta\psi_{nm} = \psi_{n \rightarrow m} - \hat{\psi}_{n \rightarrow m}. \quad (17)$$

The direction of any common line in each 2D Fourier domain is estimated only modulo 180° . The reason is that the correlation (Eq. 4) between any two such Fourier-domain lines is invariant to reflection around the zero-frequency. Thus $\Delta\psi_{nm} \in [-90^\circ, 90^\circ]$.

In Cryo-EM literature [44], a parametric model had been suggested to express the distribution of $\Delta\psi_{nm}$. The model dictates that:

- There is probability P that a pair $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ is *indicative*: the

angles are random though roughly correct. The probability is assumed to be Gaussian, in the form $\Delta\psi_{nm}, \Delta\psi_{mn} \sim \mathcal{N}(0, \sigma^2)$. Here σ is the standard deviation of angular errors in indicative common lines.

- There is probability $1 - P$ that a pair $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$ is *arbitrary*: each of these angles is random, distributed uniformly around a circle.

Therefore, the probability density function (PDF) of $\Delta\psi_{nm}$ is modeled by

$$f(\Delta\psi_{nm}) = \frac{P}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\Delta\psi_{nm})^2}{2\sigma^2}\right] + \frac{(1-P)}{180^\circ}. \quad (18)$$

The values of P and σ depend on noise sources in the data. In Cryo-EM, the radiometric signal-to-noise ratio (SNR) is very low and there are no concerns of scale variations or non-rigid deformations. In our case, however, the opposite holds. Hence, we tested in simulations whether this model can be relevant.

B.1. The Effect of Radiometric Noise

We simulated a synthetic volume and projected it using $N = 100$ random orientations. The projections included photon (Poisson) noise, typical to optical cameras. The best quality had photon noise which corresponds to a full-well depth of 10000 photoelectrons, i.e. SNR=100 near camera saturation. The lower quality corresponded to a full-well depth of 100 photoelectrons, i.e. SNR=10 near camera saturation. Eqs. (4,5,17) were then run, after which the model of Eq. (18) was fitted to the histogram of errors. The fit was good yet *insensitive to the radiometric SNR*, yielding $P = 0.897 \pm 0.007$ and $\sigma = 1.25^\circ \pm 0.01^\circ$ across the range. Similar results were obtained when introducing random lateral displacements and scales to the imaged samples. For this reason, we do not see radiometric noise as a major source of uncertainty in common line estimation, in our type of imaging.

B.2. Specimen Deformation as Noise

The objects we consider always have variance in sub-structures, i.e., non-rigid deformations. Such random morphological noise disrupts proper detection and matching of common line pairs. To induce such noise in a simulation, we followed a recipe by Ref. [43]. Let the longest dimension of a volumetric distribution $\beta(\mathbf{x})$ be bounded by G voxels. The domain size of $\beta(\mathbf{x})$ is $G \times G \times G$ voxels. This distribution is deformed to $\beta(\mathbf{x} + \mathbf{v}^{(n)}[\mathbf{x}])$. Here $\mathbf{v}^{(n)} = [v_1, v_2, v_3]$ represents a random vector (deformation) field of the specimen corresponding to image n . There is a 3D vector of deformation at every point in the object.

The deformation direction $\hat{\mathbf{v}}^{(n)} = [\hat{v}_1^{(n)}, \hat{v}_2^{(n)}, \hat{v}_3^{(n)}]$ is randomly sampled on the unit sphere. The deformation magnitude varies spatially as a sinusoid having a random amplitude $V^{(n)} \sim \mathcal{U}(0, V_{\max}G)$. Each spatial component $c \in \{1, 2, 3\}$ has random phase $\phi_c^{(n)} \sim \mathcal{U}(-\pi, \pi)$. The distortion of component c is

$$v_c^{(n)} = \hat{v}_c^{(n)} V^{(n)} \sin\left[2\pi x_c / (G - 1) + \phi_c^{(n)}\right]. \quad (19)$$

This deformation field model represents a basic Fourier component of distortion. We set $G = 100$. For each value of V_{\max} , we computed $N = 100$ random projections and deformations. Eqs. (4,5,17) were run, and then the model of Eq. (18) was fitted to the histogram of angular errors (Fig. 5). Across a range of V_{\max} values, the fit is good. The values of P and σ are sensitive to the magnitude of distortion amplitude V_{\max} (Fig. 6).

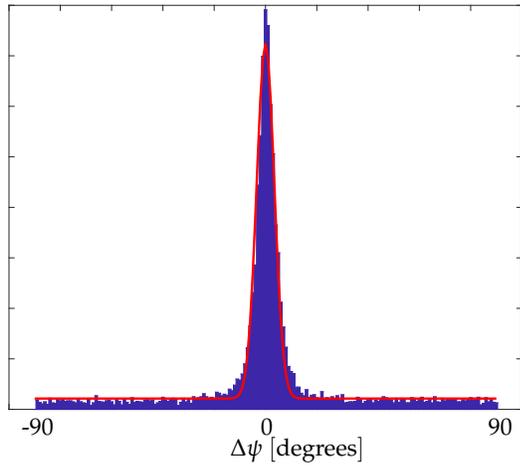


Fig. 5. A histogram of the angular error $\Delta\psi_{nm}$ (in degrees) with $V_{\max} = 9\%$, when data is affected by simulated random non-rigid deformations. The red curve is a fit of the histogram to Eq. (18). The fit has a coefficient of correlation 0.98.

C. Common Line Triplets

To establish the weights $w_{n,m}$, we need a statistical model. Section B presented a generic model parameterized by P and σ . When using real-data, we do not know the values of these statistical parameters. Thus, we derive these parameters based on the actual real-data, which includes many (noisy) estimates of angles $\{\hat{\psi}_{n \rightarrow m}\}_{n,m}$.

We denote the n, m pair of images as $\overline{n, m}$. For any $\overline{n, m}$, we should assess the likelihood that the estimated angle-pair is indicative or arbitrary. This can be done by relying on *common line triplets*. Denote a different triplet of projection images (indexed n, m, l) of the same object as n, m, l . It corresponds to true rotations $\mathbf{R}_n, \mathbf{R}_m, \mathbf{R}_l$, and consequently to true three matrices $\mathbf{R}_{n \rightarrow m}, \mathbf{R}_{m \rightarrow l}, \mathbf{R}_{l \rightarrow n}$, following Eq. (9). To be correct rotation matrices, they must satisfy

$$\mathbf{R}_{n \rightarrow m} \mathbf{R}_{m \rightarrow l} \mathbf{R}_{l \rightarrow n} - \mathbf{I} = 0. \quad (20)$$

When processing real-data, we do not know the true rotations. Instead, we only have noisy estimates $\hat{\mathbf{R}}_{n \rightarrow m}, \hat{\mathbf{R}}_{m \rightarrow l}, \hat{\mathbf{R}}_{l \rightarrow n}$, due to the noisy estimated set $\{\hat{\psi}_{n \rightarrow m}\}_{\forall m \neq n}$. Generally Eq. (20) is not satisfied exactly when based on estimated angles and matrices.

C.1. Triplet Score and Statistics

Ref. [44] proposed a *score* to quantify how closely Eq. (20) is approximated using estimated matrices

$$s_{n,m,l} \equiv s(\hat{\mathbf{R}}_{n \rightarrow m} \hat{\mathbf{R}}_{m \rightarrow l} \hat{\mathbf{R}}_{l \rightarrow n}), \quad (21)$$

whose range is $s_{n,m,l} \in [0, 1]$. The score decreases the further Eq. (20) is from being satisfied.

Because the empirical values of $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow l}, \hat{\psi}_{l \rightarrow n}$ are random, then $s_{n,m,l}$ is random. If all three angles $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow l}, \hat{\psi}_{l \rightarrow n}$ are indicative, then $\overline{n, m, l}$ is said to be an *indicative triplet*, and $s_{n,m,l}$ is likely to be high. However, if any of these angles is arbitrary, then $\overline{n, m, l}$ is said to be an *arbitrary triplet*, and $s_{n,m,l}$ is likely to be low.

These likelihoods are expressed by PDFs [44]. For *arbitrary triplets*, the PDF of s is denoted by $f_{\text{triplet}}^{\text{arbitrary}}(s)$, which can be pre-calculated using a simulation. This simulation does not rely

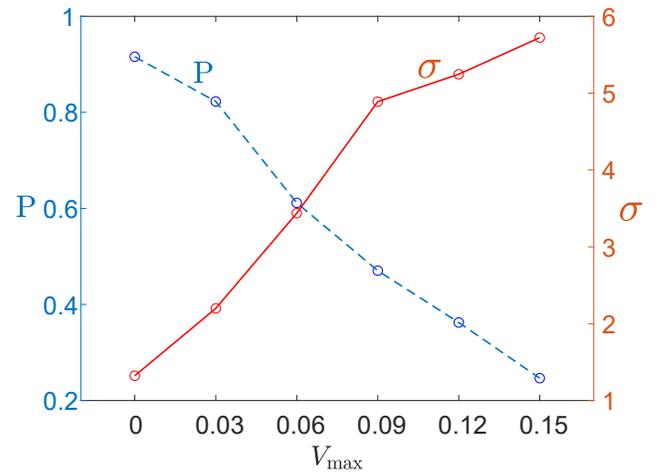


Fig. 6. The influence of the deformation coefficient V_{\max} on the parameters of the statistical model (18). The stronger the deformations, the lower is the probability that correlation-based angle-pairs are indicative. Pairs that are indicative have a larger uncertainty as V_{\max} increases.

on any image data. Rather, it tests relations between random rotation matrices:

- Draw a set of a large number of random uniformly distributed common line angles $\{\hat{\psi}_{n \rightarrow m}^{\text{simul}}\}_{m,n}$.
- Calculate $s \in [0, 1]$ for triplets drawn from this set using (21).
- Calculate a normalized histogram of s and fit a simple continuous curve [44] to it, in the domain $s \in [0, 1]$.

This yields $f_{\text{triplet}}^{\text{arbitrary}}(s)$, and it is independent of the particular mechanism underlying any arbitrary orientation errors. The score is maximal for $s = 0$ and decays to zero at $s = 1$ (Fig. 7).

For *indicative triplets*, the PDF of s is analogously denoted by $f_{\text{triplet}}^{\text{indicative}}(s|\sigma)$. Here σ is the standard deviation of angular errors in indicative common lines, as used in Sec. B.

The PDF $f_{\text{triplet}}^{\text{indicative}}(s|\sigma)$ is likewise pre-calculated in simulations, irrespective of the underlying mechanism causing orientation errors:

- Draw a set of a large number of arbitrary orientations $\{\mathbf{R}_n\}_n$, randomly uniformly distributed.
- Calculate the set of corresponding correct common line angles $\{\psi_{n \rightarrow m}^{\text{simul}}\}_{m,n}$.
- Perturb these angles by random Gaussian errors having standard deviation σ (as in Eq. 18) for $P = 1$. This yields a set of indicative $\{\hat{\psi}_{n \rightarrow m}^{\text{simul}}\}_{m,n}$.
- Calculate $s \in [0, 1]$ for triplets drawn from this set, using (21).
- Calculate a normalized histogram of s and fit a simple continuous curve [44] to it, in the domain $s \in [0, 1]$. This fitted curve is parameterized by σ .

For the special case $\sigma = 6$, $f_{\text{triplet}}^{\text{indicative}}(s|\sigma)$ is indeed maximal close to $s = 1$ and decays to zero at $s = 0$ (Fig. 7).

C.2. Estimating The Statistical Model Parameters

To estimate P and σ , we use the pre-calculated PDFs $f_{\text{triplet}}^{\text{arbitrary}}(s)$, $f_{\text{triplet}}^{\text{indicative}}(s|\sigma)$ together with the set $\{\hat{\psi}_{n \rightarrow m}\}_{m,n}$ empirically derived from the real-data. Consequently the weights $w_{n,m}$ are derived as in Sec. C.3. Recall that for each m, n , the probability that $\{\hat{\psi}_{n \rightarrow m}\}$ is indicative is P . Assume that the esti-

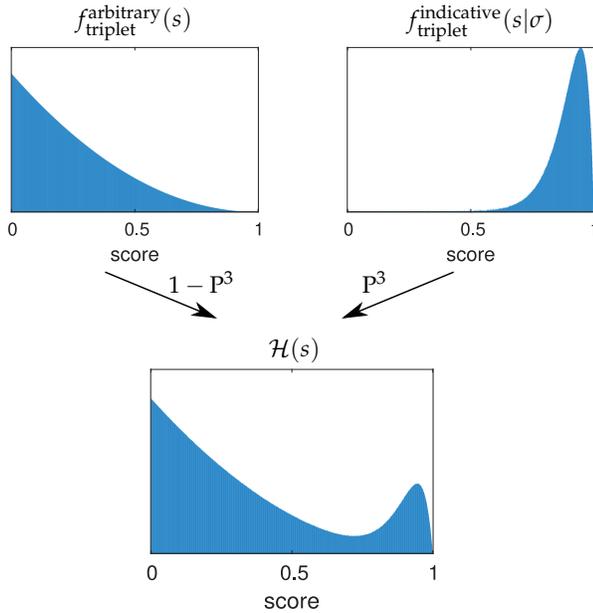


Fig. 7. [Top-Left] The pre-computed PDF of a triplet score s , if the triplet of estimated relative rotations is arbitrary. [Top-Right] The pre-computed PDF of a triplet score s , if the triplet of estimated relative rotations is indicative. Here estimated relative rotations have errors whose standard deviation is $\sigma = 6^\circ$. [Bottom] Empirical histogram $\mathcal{H}(s)$ of triplet scores. It is modelled as a convex sum of the above PDFs, controlled by P . Fitting this parametric model to $\mathcal{H}(s)$ yields an estimate of $\hat{P}, \hat{\sigma}$.

mation of each common line is approximately independent of the estimation of the other lines. Hence, the probability that a triplet is indicative is $\approx P^3$, and the probability that a triplet is arbitrary is approximately $1 - P^3$.

From the data-based estimation of $\{\hat{\psi}_{n \rightarrow m}\}_{m,n}$, the matrices $\{\hat{\mathbf{R}}_{n \rightarrow m}\}_{m,n}$ are derived. Then Eq. (21) yields a set of scores $\{s_{n,m,l}\}_{n \neq m \neq l=1}^N$. The values in the set of scores yield a normalized histogram $\mathcal{H}(s)$. According to the model,

$$\mathcal{H}(s) \approx P^3 f_{\text{triplet}}^{\text{indicative}}(s|\sigma) + (1 - P^3) f_{\text{triplet}}^{\text{arbitrary}}(s). \quad (22)$$

The values of P and σ are then searched such that $\mathcal{H}(s)$ fits the model: a convex sum of pre-computed PDFs. This yields estimation of the statistical model parameters, $\hat{P}, \hat{\sigma}$.

C.3. Confidence Weights

To derive the weights $w_{n,m}$ we can now rely on the estimated $\hat{P}, \hat{\sigma}$ and the set of empirical scores $\{s_{n,m,l}\}_{n \neq m \neq l=1}^N$. The confidence in the angles of \bar{n}, \bar{m} , $\hat{\psi}_{n \rightarrow m}, \hat{\psi}_{m \rightarrow n}$, is assessed by the set of triplet scores associated with \bar{n}, \bar{m} :

$$S_{n,m} \equiv \{s_{n,m,l} \mid \forall l \notin \{n,m\}\}. \quad (23)$$

The likelihoods that this set is drawn from indicative or arbitrary common lines are $\ell_{n,m}^{\text{indicative}}(\hat{\sigma})$ and $\ell_{n,m}^{\text{arbitrary}}$, respectively. Details of estimates of $\ell_{n,m}^{\text{indicative}}(\hat{\sigma})$ and $\ell_{n,m}^{\text{arbitrary}}$ are given in the Appendix. Given $\ell_{n,m}^{\text{indicative}}(\hat{\sigma})$ and $\ell_{n,m}^{\text{arbitrary}}$, the likelihood of the scores $S_{n,m}$ is $\hat{P} \ell_{n,m}^{\text{indicative}}(\hat{\sigma}) + (1 - \hat{P}) \ell_{n,m}^{\text{arbitrary}}$. The probability density of n, m being indicative, given the set $S_{n,m}$, is thus [44]

$$p_{n,m} = \frac{\hat{P} \ell_{n,m}^{\text{indicative}}(\hat{\sigma})}{\hat{P} \ell_{n,m}^{\text{indicative}}(\hat{\sigma}) + (1 - \hat{P}) \ell_{n,m}^{\text{arbitrary}}}. \quad (24)$$

The confidence weights are set to be

$$w_{n,m} = \gamma p_{n,m}, \quad (25)$$

where γ is a normalization constant to ensure that Eq. (15) is satisfied.

From this result, a higher weight is given to a common line pair which yields a higher likelihood of indicative triplet scores. This means that the particular common line being weighted has high support across the entire set of common lines. A common line pair, which has little support from the other common lines for being indicative, has a low probability for being indicative and thus has a lower weight. The weight (Eq. 25) is used for constructing the matrix $\tilde{\mathbf{S}}$ (Sec. A). In turn, this enhances the estimation of the rotation matrices $\{\hat{\mathbf{R}}_n\}_{n=1}^N$.

D. A Constraint for Scale Estimation

Recall from Eq. (5) that pairwise rotation angles, scales and translations are searched in a 4D domain. This is a computationally expensive process that is prone to errors in estimated pairwise rotation angles, which can in turn induce errors in scale. Our procedure (Secs. A-C) reduces these problems, yielding a robust estimation of the rotations, despite uncertainties and significant outliers. The estimated rotations $\{\hat{\mathbf{R}}_n\}_{n=1}^N$ consequently *constrain the pairwise angles*. The constrained angles then lead to a simpler search, whose results proved to be more accurate. We follow these steps:

(a) Run Eq. (5) for initial values of $\{\hat{\psi}_{n \rightarrow m}\}_{m,n}$, as well as initial estimates of pairwise scales and shifts. The search is executed on a coarse 4D grid for the sake of speed. The results can have many outliers with large uncertainties. Discard the initial estimates of pairwise scales and shifts.

(b) Run the process described in Secs. A-C to estimate the rotations $\{\hat{\mathbf{R}}_n\}_{n=1}^N$.

(c) Use $\{\hat{\mathbf{R}}_n\}_{n=1}^N$ to derive the common line between any pair of central slices \bar{n}, \bar{m} . This is done in closed-form and yields a new set of common line pairs, denoted by $\{\tilde{\psi}_{n \rightarrow m}\}_{m,n}$. The new set of common lines $\{\tilde{\psi}_{n \rightarrow m}\}_{m,n}$ is thus a result of robust estimation, which accounts for rotations of all projections in the image set. This is contrary to the set $\{\hat{\psi}_{n \rightarrow m}\}_{m,n}$ which was crudely sampled based on pairs of images.

(d) Constrain Eq. (5) to run on the common line pair $\tilde{\psi}_{n \rightarrow m}, \tilde{\psi}_{m \rightarrow n}$:

$$\{\hat{M}_{n \rightarrow m}, \hat{\tau}_{n \rightarrow m}\} = \underset{M, \tau}{\operatorname{argmax}} \frac{\Sigma_{nm}(M, \tau)}{\sqrt{\Sigma_{nm}(M, \tau)} \sqrt{\Sigma_{mm}(M, \tau)}} \quad (26)$$

where $\Sigma_{nm}(M, \tau) \equiv \Sigma_{nm}(\tilde{\psi}_{n \rightarrow m}, \tilde{\psi}_{m \rightarrow n}, M, \tau)$. This is a 2D search, contrary to Eq. (5).

The resulting pairwise scales are then used in Eqs. (6,7), to yield all scales, globally. Similarly, pairwise translations are fed into a linear optimization process. This yields all translations in a global frame, akin to [43]. Once rotations, scales, and translations are estimated, the pairwise search is repeated. In the new iteration, the search domain is narrowed and sampled more finely. This step is integrated into a broader coarse-to-fine approach that follows [43].

Eventually, an estimate of all orientation, scales and shifts is obtained. Consequently, tomographic 3D recovery of the objects is performed, yielding $\hat{\beta}(x)$. We used an Algebraic Reconstruction Technique [57, 58] for this purpose.

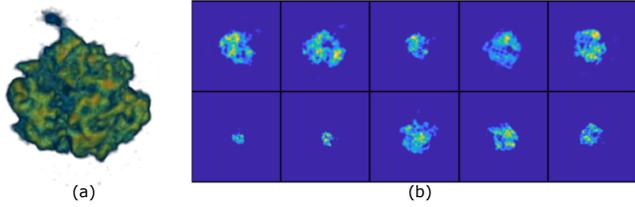


Fig. 8. (a) A 3D volumetric object is randomly *scaled, deformed* shifted and projected onto 100 random orientations. (b) Here we show several image samples from this simulation.

4. RESULTS

A. Quantitative Criteria

We performed simulations as well as experiments using real-data. For quantitative criteria, we follow [43] by noting that all rotations can only be estimated up to an unknown global rotation \mathbf{O} . Similarly, all scales are defined up to a global unknown M_{rel} . Hence [43], when ground-truth is available during simulations, error measures for rotations and scales are respectively

$$\epsilon_{\text{rot}} = \frac{1}{N} \min_{\mathbf{O}} \sum_{n=1}^N \|\mathbf{R}_n - \mathbf{O}\hat{\mathbf{R}}_n\|_2^2, \quad (27)$$

$$\epsilon_{\text{scale}} = \frac{1}{N} \min_{M_{\text{rel}}} \sum_{n=1}^N |M_n - M_{\text{rel}}\hat{M}_n|^2. \quad (28)$$

The relative error of the estimated object density $\hat{\beta}(\mathbf{x})$ is

$$\epsilon_{\text{dens}} = \frac{\int |\hat{\beta}(\mathbf{x}) - \beta(\mathbf{x})| d\mathbf{x}}{\int \beta(\mathbf{x}) d\mathbf{x}}. \quad (29)$$

Since our experiments on real *in situ* plankton images have no ground-truth, quality is assessed via cross-validation. Rotations, scales and translations are estimated using the entire ensemble. Then a single image \tilde{n} is excluded from the N -sample ensemble. The 3D estimation $\hat{\beta}$ thus uses $N - 1$ images. Then, $\hat{\beta}$ is re-projected using $\hat{\mathbf{R}}_{\tilde{n}}, \hat{M}_{\tilde{n}}, \hat{\tau}_{\tilde{n}}$, yielding a synthetic image $\hat{\mathbf{y}}_{\tilde{n}}$. It is compared to the true image $\mathbf{y}_{\tilde{n}}$ using relative root-mean-squared error. This is repeated $\forall \tilde{n} \in [1, N]$. There are N_{pixel} pixels. The re-projection error is

$$\epsilon_{\text{re-proj}} = \frac{1}{N\sqrt{N_{\text{pixel}}}} \sum_{\tilde{n}=1}^N \frac{\|\mathbf{y}_{\tilde{n}} - \hat{\mathbf{y}}_{\tilde{n}}\|_2}{\max(\mathbf{y}_{\tilde{n}})}. \quad (30)$$

Plankton ecologists usually model population size distribution as a log-normal PDF [59]. Let an object be of size (e.g., length) ϱ . Then the PDF model is

$$f^{\text{size}}(\varrho) = \frac{1}{\varrho\sqrt{2\pi\nu^2}} \exp\left\{-\frac{[\ln(\varrho) - \zeta]^2}{2\nu^2}\right\}. \quad (31)$$

Here ζ, ν are the mean and standard deviation in the logarithm domain. We can then assess the quality of the estimated scales by fitting Eq. (31) to the histogram of $\hat{M}_{\tilde{n}}$.

B. Simulation based on Empirical Data

The simulation is based on a real object whose 3D volumetric distribution is empirically known [60]. In our simulations, this object underwent random deformations as described in Sec. B.2, with $V_{\text{max}} = 6\%$. It also underwent random scaling distributed as $\log_e(\text{scale}) \sim \mathcal{U}(-0.7, 0.7)$. This results in a scale factor of

	ϵ_{dens}	ϵ_{rot}	ϵ_{scale}
Ours	0.57 ± 0.05	$(3.9 \pm 0.5) \cdot 10^{-4}$	$(4.8 \pm 1) \cdot 10^{-5}$
Ref. [43]	0.88 ± 0.07	$(318 \pm 70) \cdot 10^{-4}$	$(5.4 \pm 2) \cdot 10^{-5}$

Table 1. Results corresponding to data described in Fig. 8. They result from six repeated simulations, each having different random samples. This yields the uncertainties quoted. The rotation errors for our method and [43] are $0.48^\circ \pm 0.27^\circ$ and $9.29^\circ \pm 6.84^\circ$, respectively.

$\times 4$ between the lengths of the largest and smallest specimen. The object was projected to $N = 100$ random directions, each yielding an image of size 126×126 pixels. The projection then underwent random lateral shifts in the image domain, distributed as $\text{shift} \sim \mathcal{U}(-10, 10)$ [pixels] in both image axes. The prototype object and sample projections are shown in Fig. 8.

The initial search domain in Eq. (5) used sampled angles with one-degree resolution. The domain had 10 relative-scale samples uniformly-spaced in the logarithm domain $[-\log(2), \log(2)]$ and 10 samples of relative offsets in a $[-15, 15]$ pixel domain. As in Sec. D, the domains become narrow through an iterative coarse-to-fine search. The initial pairwise matches were processed using the approach of this paper and compared to the result of [43]. Results are shown in Fig. 9 and Table 1.

C. Microscopy Experiments

C.1. Darkfield Data

The first dataset contains images of *Daphnia* sp., a species of freshwater zooplankton (Fig. 10). Outliers are first pruned out [43]: images whose estimated common lines pairs have low correlation values (under a manually chosen threshold) are considered outliers. For this fixed threshold, pruning resulted in a cluster of $N = 147$ samples for [43]. Correlations on the constrained lines described in Sec. D yielded $N = 145$ inlier samples for our approach. The initial search domain in Eq. (5) is as in the simulations. Comparing the resulting re-projections on sample frames from both approaches indicates that our approach is able to resolve finer details (Fig. 10). Quantitatively, our approach yields $\epsilon_{\text{re-proj}} = 1.38 \pm 0.002$. In comparison, the method of [43] yields $\epsilon_{\text{re-proj}} = 3.065 \pm 0.004$.

The second dataset consists of 1000 images of the cosmopolitan zooplankton *Oithona* sp. (Fig. 11). To keep runtime short in this large database, we used 50 equally spaced angular samples in the search domain of Eq. (5). We used the other domain sample parameters as those in the simulations. Pruning yielded $N = 921$ inlier samples for our approach and $N = 877$ for the method of [43]. Fig. 11 compares re-projections while Fig. 12 compares 3D recoveries. The proposed method yields $\epsilon_{\text{re-proj}} = 4.174 \pm 8 \cdot 10^{-4}$. In comparison, the method of [43] yields $\epsilon_{\text{re-proj}} = 6.559 \pm 0.002$.

Recall that our method yields an estimation of the size distribution. Figs. 13a,b show the estimated size histograms of *Oithona* sp. and *Daphnia* sp. and the corresponding PDFs fitted using Eq. (31). The estimated size histograms and fitted PDF by [43] are very similar to ours.

C.2. Brightfield Data

For direct comparison to the results of [43], we used the publicly available plankton image database provided by [61], specifically focusing on *Pyramimonas longicauda*. We note that this algal species is an order of magnitude smaller than the zooplankton

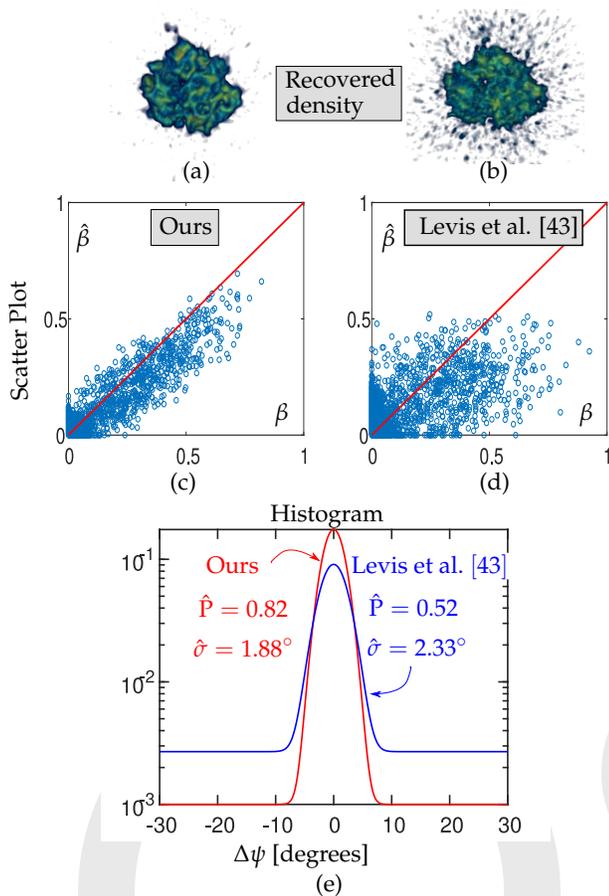


Fig. 9. Recovery using data described in Fig. 8. [Top] Visualization of the 3D object density estimated by (a) our method vs. (b) that of [43]. (c,d) The respective scatter plots of $\hat{\beta}$ vs. true β (values drawn from 1% random voxels). The red line represents a perfect fit. Our approach clearly yields a better fit to ground truth than [43]. (e) The statistical model (18) fitted to the empirical histogram of angular errors, similarly to Fig. 6. In the process described in Sec. D, estimated angle-pairs are better constrained. After a single iteration, our error model (red plot) has a narrower standard deviation and a higher indicative probability, compared to errors (blue plot) obtained in the method of [43].

considered in the other datasets. Pruning of poorly correlated data yielded a set of $N = 136$ images for our analysis and $N = 132$ for running [43]. Fig. 14 visualizes the respective 3D recoveries. Fig. 15 shows several re-projections and compares them to the real images. Both figures indicate that the proposed approach yields more complete results, of higher contrast. Quantitatively, our approach yielded $\epsilon_{\text{re-proj}} = 3.446 \pm 0.004$. In comparison, the method of [43] yielded $\epsilon_{\text{re-proj}} = 3.942 \pm 0.005$.

5. DISCUSSION

Using our method, massive 2D image databases collected globally by ocean and lake *in situ* observation systems can lead to 3D structure and population statistics estimates. Tackling this problem at scale necessitates more efficient processing. For example, using 3D volumetric models of plankton, learning-based methods can be trained to infer species and distributions, based

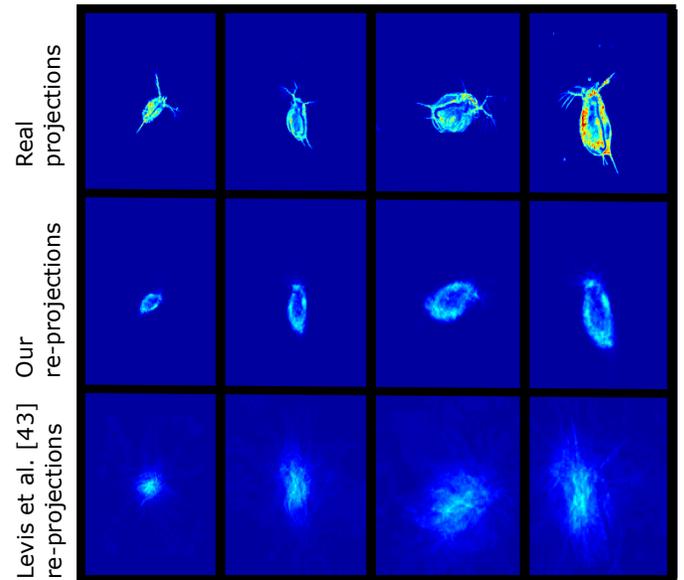


Fig. 10. Experiment on a population of *Daphnia* sp., imaged using our system. [Top] A few raw darkfield images. [Middle] Corresponding cross-validation re-projections, as described in Sec. A, using our tomographically estimated object. [Bottom] Corresponding cross-validation re-projections, using the estimation resulting from the method of [43].

on image projections taken at random poses.

Our experiments indicate that the method is robust to differences in organism size: the SPC images are of plankton in the 0.5-2.0 mm range, while those captured by the IFCB are 10s of microns across. The method seems to be robust to the body plan: *Daphnia* and *Oithona* are both zooplankton while *Pyramimonas* is a photosynthetic algae. We believe the approach will yield reasonable reconstructions for a variety of plankton, though domain experts should be consulted when approaching new taxa. Gelatinous organisms will be particularly challenging since much of their bodies are transparent in seawater.

The approach presented here is not limited to plankton. We believe it can be used in population analysis of a wide variety of small, translucent suspended objects. Beyond other organism classes, the method might also be applied to non-biological targets. These may include submerged suspended sediments, airborne aerosols and hydrosols, ice crystals in clouds and possibly small meteoroids.

6. BACKMATTER

Acknowledgments. We thank Aviad Levis and Tali Treibitz for their advice, and Johanan Erez, Ina Talmon and Daniel Yagodin for technical support. Yoav Schechner is the Mark and Diane Seiden Chair in Science at the Technion. He is a Landau Fellow - supported by the Taub Foundation. His work was conducted in the Ollendorff Minerva Center. Minvera is funded through the BMBF. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 810370-ERC-CloudCT.

Disclosures. The authors declare no conflicts of interest.

Data Availability Statement.

Data availability. Data underlying the results presented in this paper are available in Refs. [4, 61].

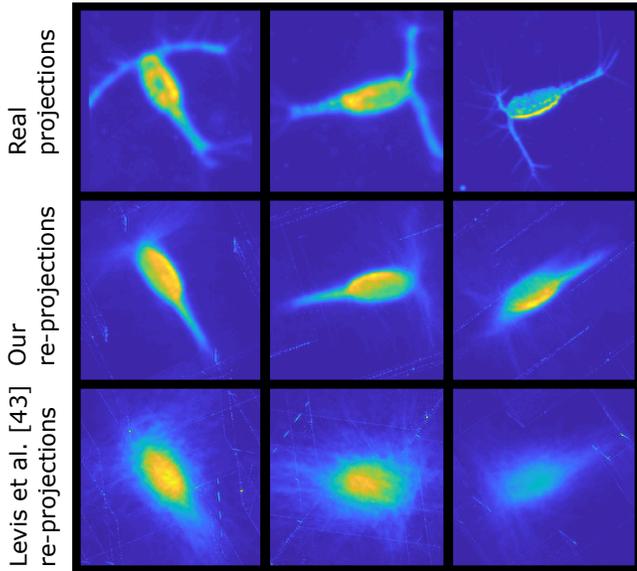


Fig. 11. Experiment on a population of *Oithona* sp., imaged using our system. [Top] A few raw darkfield images. [Middle] Corresponding cross-validation re-projections of our tomographically estimated object. [Bottom] Corresponding cross-validation re-projections of the object resulting from the method of [43].



Fig. 12. Volumetric density recovery: *Oithona* sp.

APPENDIX

We now describe how estimates of $\ell_{n,m}^{\text{indicative}}(\hat{\sigma})$ and $\ell_{n,m}^{\text{arbitrary}}$ are obtained, for used in Eq. (24).

- Assume for a moment that the n, m common line is arbitrary. Then by definition, all elements in $\mathcal{S}_{n,m}$ are scores of arbitrary triplets. Then, the likelihood of this set of scores being drawn from arbitrary common lines is

$$\ell_{n,m}^{\text{arbitrary}} \equiv \ell^{\text{arbitrary}}(\mathcal{S}_{n,m}) = \prod_{l \notin \{n,m\}} f_{\text{triplet}}^{\text{arbitrary}}(s_{n,m,l}). \quad (32)$$

Here, it is assumed that $s_{n,m,l}, s_{n,m,l'}$ are approximately statistically independent $\forall l \neq l'$, given that the common line of \bar{n}, \bar{m} is arbitrary.

- Assume for a moment that the common line of \bar{n}, \bar{m} is indicative. Then, $\bar{n}, \bar{m}, \bar{l}$ is indicative only if both the common lines of \bar{n}, \bar{l} and \bar{m}, \bar{l} are indicative as well. For such an event, the likelihood of the score $s_{n,m,l}$ is thus $\hat{P}^2 f_{\text{triplet}}^{\text{indicative}}(s_{n,m,l}|\hat{\sigma})$.

The complementing case is that either or both common lines of \bar{n}, \bar{l} and \bar{m}, \bar{l} are arbitrary. Then, the triplet $\bar{n}, \bar{m}, \bar{l}$ is arbitrary, despite the common line of \bar{n}, \bar{m} being indicative. In this case,

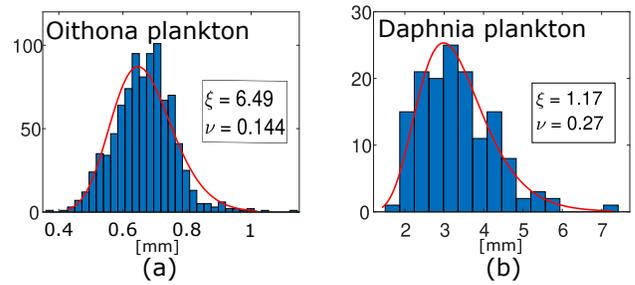


Fig. 13. Resulting size-distribution histograms (blue) and fitted (Eq. 31) density (red). (a) *Oithona* sp. The fit 95% confidence intervals are $\xi \in [6.48 \ 6.5]$, $\nu \in [0.138 \ 0.151]$. (b) *Daphnia* sp. The fit 95% confidence intervals are $\xi \in [1.12 \ 1.21]$, $\nu \in [0.244 \ 0.308]$.

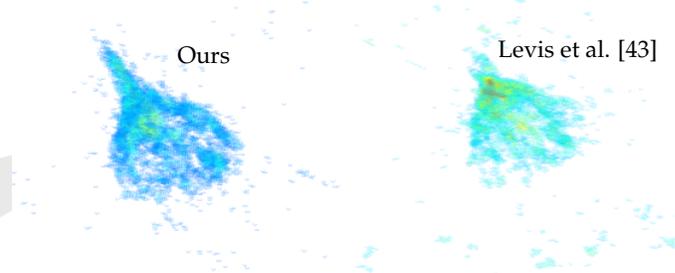


Fig. 14. Volumetric density recovery: *Pyramimonas longicauda*.

the likelihood of the score $s_{n,m,l}$ is thus $(1 - \hat{P}^2) f_{\text{triplet}}^{\text{arbitrary}}(s_{n,m,l})$. Consequently, in total, if the common line of \bar{n}, \bar{m} is indicative, the likelihood of the score $s_{n,m,l}$ is $\hat{P}^2 f_{\text{triplet}}^{\text{indicative}}(s_{n,m,l}|\hat{\sigma}) + (1 - \hat{P}^2) f_{\text{triplet}}^{\text{arbitrary}}(s_{n,m,l})$. Therefore, compounding all scores in $\mathcal{S}_{n,m}$, the likelihood of this set being drawn while the common line of \bar{n}, \bar{m} is indicative is

$$\ell_{n,m}^{\text{indicative}}(\hat{\sigma}) \equiv \ell^{\text{indicative}}(\mathcal{S}_{n,m}|\hat{\sigma}) = \prod_{l \notin \{n,m\}} \left[\hat{P}^2 f_{\text{triplet}}^{\text{indicative}}(s_{n,m,l}|\hat{\sigma}) + (1 - \hat{P}^2) f_{\text{triplet}}^{\text{arbitrary}}(s_{n,m,l}) \right]. \quad (33)$$

Aggregating both cases, recall that \bar{n}, \bar{m} has prior probabilities $\hat{P}, (1 - \hat{P})$ for being indicative or arbitrary, respectively.

REFERENCES

1. G. C. Hays, A. J. Richardson, and C. Robinson, "Climate change and marine plankton," *Trends ecology & evolution* **20**, 337–344 (2005).
2. J. Ryan, R. Kudela, J. Birch, M. Blum, H. Bowers, F. Chavez, G. Doucette, K. Hayashi, R. Marin, C. Mikulski, and J. Pennington, "Causality of an extreme harmful algal bloom in Monterey Bay, California, during the 2014–2016 northeast Pacific warm anomaly," *Geophys. Res. Lett.* **44**, 5571–5579 (2017).
3. C. Anderson, E. Berdalet, R. Kudela, C. Cusack, J. Silke, E. O'Rourke, D. Dugan, M. McCammon, J. Newton, S. Moore, and K. Paige, "Scaling up from regional case studies to a global harmful algal bloom observing system," *Front. Mar. Sci.* **6**, 250 (2019).
4. *Scripps plankton camera website*. <http://spc.ucsd.edu>.
5. G. Gorsky, M. Picheral, and L. Stemann, "Use of the underwater video profiler for the study of aggregate dynamics in the north mediterranean," *Estuarine, Coast. Shelf Sci.* **50**, 121–128 (2000).
6. C. S. Davis, S. M. Gallagher, M. Marra, and W. K. Stewart, "Rapid visualization of plankton abundance and taxonomic composition using

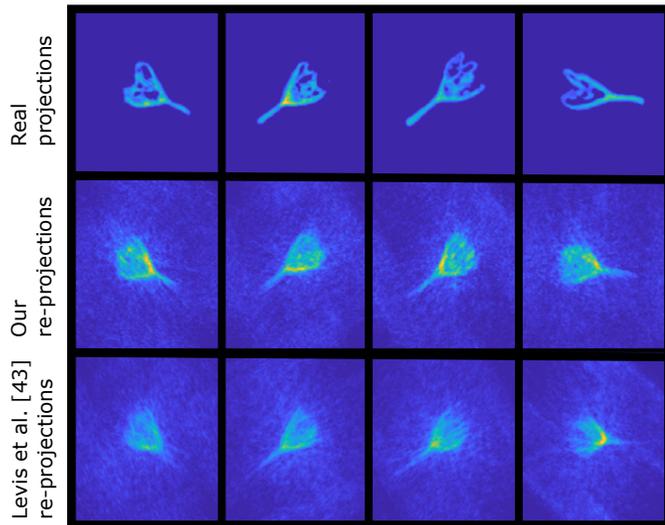


Fig. 15. Experiment on a population of *Pyramimonas longicauda*, from dataset [4]. [Top] A logarithm operation on raw bright-field images yields a linear model as in Eq. (1). A few samples are shown. [Middle] Corresponding cross-validation re-projections of our tomographically estimated object. [Bottom] Corresponding cross-validation re-projections of the object resulting from the method of [43].

- the Video Plankton Recorder," *Deep. Sea Res. Part II: Top. Stud. Oceanogr.* **43**, 1947–1970 (1996).
7. M. D. Ohman, R. E. Davis, J. T. Sherman, K. R. Grindley, B. M. Whitmore, C. F. Nickels, and J. S. Ellen, "Zooglider: an autonomous vehicle for optical and acoustic sensing of zooplankton," *Limnol. Oceanogr. Methods* **17**, 69–86 (2019).
 8. R. K. Cowen and C. M. Guigand, "In situ ichthyoplankton imaging system (ISIIS): system design and preliminary results," *Limnol. Oceanogr. Methods* **6**, 126–132 (2008).
 9. S. Vogel, *Life in Moving Fluids: The Physical Biology of Flow-Revised and Expanded Second Edition* (Princeton University Press, 2020).
 10. S. Lee, K. Kim, A. Mubarak, A. Panduwirawan, K. Lee, S. Lee, H. Park, and Y. Park, "High-resolution 3D refractive index tomography and 2D synthetic aperture imaging of live phytoplankton," *J. Opt. Soc. Korea* **18**, 691–697 (2014).
 11. S. Johnsen, *The optics of life: a biologist's guide to light in nature* (Princeton University Press, 2012).
 12. S. Colin, L. P. Coelho, S. Sunagawa, C. Bowler, E. Karsenti, P. Bork, R. Pepperkok, and C. De Vargas, "Quantitative 3D-imaging for cell biology and ecology of environmental microbial eukaryotes," *Elife* **6**, e26066 (2017).
 13. F. Lombard, E. Boss, A. M. Waite, J. Uitz, L. Stemann, H. M. Sosik, J. Schulz, J.-B. Romagnan, M. Picheral, and J. Pearlman, "Globally consistent quantitative observations of planktonic ecosystems," *Front. Mar. Sci.* **6**, 196 (2019).
 14. M. Picheral, L. Guidi, L. Stemann, D. M. Karl, G. Iddaoud, and G. Gorsky, "The underwater vision profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton," *Limnol. Oceanogr. Methods* **8**, 462–473 (2010).
 15. R. J. Olson and H. M. Sosik, "A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot," *Limnol. Oceanogr. Methods* **5**, 195–203 (2007).
 16. J. S. Jaffe, "Underwater optical imaging: the past, the present, and the prospects," *IEEE J. Ocean. Eng.* **40**, 683–700 (2014).
 17. A. D. Mullen, T. Treibitz, P. L. Roberts, and J. S. Jaffe, "An underwater microscope for in situ imaging of seafloor organism," in *Novel Techniques in Microscopy*, (OSA, 2017), pp. NTu1C–1.
 18. E. Acevedo-Trejos, G. Brandt, J. Bruggeman, and A. Merico, "Mechanisms shaping size structure and functional diversity of phytoplankton communities in the ocean," *Sci. reports* **5**, 8918 (2015).

19. H. Sun, P. W. Benzie, N. Burns, D. C. Hendry, M. A. Player, and J. Watson, "Underwater digital holography for studies of marine plankton," *Philos. Transactions Royal Soc. A: Math. Phys. Eng. Sci.* **366**, 1789–1806 (2008).
20. E. J. Davies and R. Nepstad, "In situ characterisation of complex suspended particulates surrounding an active submarine tailings placement site in a norwegian fjord," *Reg. Stud. Mar. Sci.* **16**, 198–207 (2017).
21. J. Sheng, E. Malkiel, and J. Katz, "Digital holographic microscope for measuring three-dimensional particle distributions and motions," *Appl. optics* **45**, 3893–3901 (2006).
22. S. Samson, T. Hopkins, A. Remsen, L. Langebrake, T. Sutton, and J. Patten, "A system for high-resolution zooplankton imaging," *IEEE J. oceanic Eng.* **26**, 671–676 (2001).
23. F. Lombard, E. Boss, A. M. Waite, M. Vogt, J. Uitz, L. Stemann, H. M. Sosik, J. Schulz, J.-B. Romagnan, M. Picheral, and J. Pearlman, "Globally consistent quantitative observations of planktonic ecosystems," *Front. Mar. Sci.* **6**, 196 (2019).
24. A. Zweifler, D. Akkaynak, T. Mass, and T. Treibitz, "In situ analysis of coral recruits using fluorescence imaging," *Front. Mar. Sci.* **4**, 273 (2017).
25. D. Akkaynak, T. Treibitz, B. Xiao, U. A. Gürkan, J. J. Allen, U. Demirci, and R. T. Hanlon, "Use of commercial off-the-shelf digital cameras for scientific data acquisition and scene-specific color calibration," *JOSA A* **31**, 312–321 (2014).
26. D. M. Kocak, F. R. Dalgleish, F. M. Caimi, and Y. Y. Schechner, "A focus on recent developments and trends in underwater imaging," *Mar. Technol. Soc. J.* **42**, 52–67 (2008).
27. A. Jordt-Sedlazeck and R. Koch, "Refractive structure-from-motion on underwater images," in *Proceedings of the IEEE International Conference on Computer Vision*, (2013), pp. 57–64.
28. C. Ancuti, C. O. Ancuti, C. De Vleeschouwer, R. Garcia, and A. C. Bovik, "Multi-scale underwater descattering," in *IEEE International Conference on Pattern Recognition*, (IEEE, 2016), pp. 4202–4207.
29. M. Sheinin and Y. Y. Schechner, "The next best underwater view," in *IEEE CVPR*, (2016), pp. 3764–3773.
30. H. Singh, J. G. Bellingham, F. Hover, S. Lemer, B. A. Moran, K. Von der Heydt, and D. Yoerger, "Docking for an autonomous ocean sampling network," *IEEE J. Ocean. Eng.* **26**, 498–514 (2001).
31. Y. Wang, S. Negahdaripour, and M. D. Aykin, "Calibration and 3d reconstruction of underwater objects with non-single-view projection model by structured light stereo imaging," *Appl. Opt.* **55**, 6564–6575 (2016).
32. T. Nakatani, S. Li, T. Ura, A. Bodenmann, and T. Sakamaki, "3d visual modeling of hydrothermal chimneys using a rotary laser scanning system," in *IEEE Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, (IEEE, 2011), pp. 1–5.
33. M. Leonardi, A. Stahl, M. Gazzea, M. Ludvigsen, I. Rist-Christensen, and S. M. Nornes, "Vision based obstacle avoidance and motion tracking for autonomous behaviors in underwater vehicles," in *IEEE OCEANS*, (IEEE, 2017), pp. 1–10.
34. E. Kelasidi, K. Y. Pettersen, J. T. Gravdahl, and P. Liljebäck, "Modeling of underwater snake robots," in *IEEE International Conference on Robotics and Automation*, (IEEE, 2014), pp. 4540–4547.
35. J. Gregson, M. Krimerman, M. B. Hullin, and W. Heidrich, "Stochastic tomography and its applications in 3d imaging of mixing fluids," *ACM TOG* **31**, 52–1 (2012).
36. A. Vainiger, Y. Y. Schechner, T. Treibitz, A. Avni, and D. S. Timor, "Optical wide-field tomography of sediment resuspension," *Opt. Express* **27**, A766–A778 (2019).
37. A. Aides, A. Levis, V. Holodovsky, Y. Y. Schechner, D. Althausen, and A. Vainiger, "Distributed sky imaging radiometry and tomography," in *2020 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2020), pp. 1–12.
38. A. Levis, Y. Y. Schechner, A. B. Davis, and J. Loveridge, "Multi-view

- polarimetric scattering cloud tomography and retrieval of droplet size," *Remote. Sens.* **12**, 2831 (2020).
39. A. Aides, Y. Y. Schechner, V. Holodovsky, M. J. Garay, and A. B. Davis, "Multi sky-view 3d aerosol distribution recovery," *Opt. express* **21**, 25820–25833 (2013).
40. J. Frank, "Single-particle reconstruction of biological macromolecules in electron microscopy—30 years," *Q. reviews biophysics* **42**, 139–158 (2009).
41. R. R. Lederman and A. Singer, "A representation theory perspective on simultaneous alignment and classification," *Appl. Comput. Harmon. Analysis* (2019).
42. L. Wang, A. Singer, and Z. Wen, "Orientation determination of Cryo-EM images using least unsquared deviations," *SIAM journal on imaging sciences* **6**, 2450–2483 (2013).
43. A. Levis, Y. Y. Schechner, and R. Talmon, "Statistical tomography of microscopic life," *IEEE CVPR* pp. 6411–6420 (2018).
44. I. Greenberg and Y. Shkolnisky, "Common lines modeling for reference free Ab-initio reconstruction in Cryo-EM," *J. structural biology* **200**, 106–117 (2017).
45. E. C. Orenstein, D. Ratelle, C. Briseño-Avena, M. L. Carter, P. J. Franks, J. S. Jaffe, and P. L. Roberts, "The scripps plankton camera system: A framework and platform for in situ microscopy," *Limnol. Oceanogr. Methods* (2020).
46. D. A. Agard, Y. Hiraoka, P. Shaw, and J. W. Sedat, "Fluorescence microscopy in three dimensions," *Methods cell biology* **30**, 353–377 (1989).
47. D. A. Agard, "Optical sectioning microscopy: cellular architecture in three dimensions," *Annu. review biophysics bioengineering* **13**, 191–219 (1984).
48. M. Alterman, Y. Y. Schechner, M. Vo, and S. G. Narasimhan, "Passive tomography of turbulence strength," in *ECCV*, (2014), pp. 47–60.
49. V. Holodovsky, Y. Y. Schechner, A. Levin, A. Levis, and A. Aides, "In-situ multi-view multi-scattering stochastic tomography," in *IEEE ICCP*, (2016), pp. 1–12.
50. D. Gürsoy, Y. P. Hong, K. He, K. Hujsak, S. Yoo, S. Chen, Y. Li, M. Ge, L. M. Miller, Y. S. Chu, and V. De Andrade, "Rapid alignment of nanotomography data using joint iterative reconstruction and reprojection," *Sci. Reports* **7**, 11818 (2017).
51. I. Ihrke and M. Magnor, "Image-based tomographic reconstruction of flames," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, (Eurographics Association, 2004), pp. 365–373.
52. A. Levis, Y. Y. Schechner, and A. B. Davis, "Multiple-scattering microphysics tomography," in *IEEE CVPR*, (2017), pp. 6740–6749.
53. D. Ren, C. Ophus, M. Chen, and L. Waller, "A multiple scattering algorithm for three dimensional phase contrast atomic electron tomography," *Ultramicroscopy* **208**, 112860 (2020).
54. B. D. Trifonov, "Tomographic reconstruction of transparent objects," Ph.D. thesis, University of British Columbia (2006).
55. A. Veeraraghavan, A. V. Genkin, S. Vitaladevuni, L. Scheffer, S. Xu, H. Hess, R. Fetter, M. Cantoni, G. Knott, and D. Chklovskii, "Increasing depth resolution of electron microscopy of neural circuits using sparse tomographic reconstruction," in *IEEE Conference on Computer Vision and Pattern Recognition*, (2010), pp. 1767–1774.
56. Y. Shkolnisky and A. Singer, "Viewing direction estimation in Cryo-EM using synchronization," *SIAM* **5**, 1088–1110 (2012).
57. R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-Ray photography," *J. theoretical Biol.* **29**, 471–481 (1970).
58. W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers, "Fast and flexible X-Ray tomography using the ASTRA toolbox," *Opt. express* **24**, 25129–25147 (2016).
59. Z. V. Finkel, J. Beardall, K. J. Flynn, A. Quigg, T. A. V. Rees, and J. A. Raven, "Phytoplankton in a changing world: cell size and elemental stoichiometry," *J. plankton research* **32**, 119–137 (2009).
60. C. L. Lawson, M. L. Baker, C. Best, C. Bi, M. Dougherty, P. Feng, G. Van Ginkel, B. Devkota, I. Lagerstedt, S. J. Ludtke, and R. Newman, "Emdatabank. org: unified data resource for cryoem," *Nucleic acids research* **39**, D456–D464 (2010).
61. H. Sosik, E. E. Peacock, and E. Brownlee, "Annotated plankton images data set for developing and evaluating classification methods," URL <http://darchive.mblwhoilibrary.org/handle/1912/7341> (2015).