# Tetrahedral and Octahedral reconstruction

Yuval Meylakh

April 12, 2025

**Abstract**

This report examines Ab initio reconstruction of 3-dimensional volumes of bio-molecules with tetrahedral (T) and octahedral (O) symmetries according to the method described by [1] . Using the fourier projection slice theorem and the common lines approach one can estimate the relative rotations of each projection. To solve the handedness ambiguity between each relative rotation, the output is run through the J synchronization algorithm [2]. The method then leverages a property unique to $T$ and $O$ symmetry groups in SO(3) to estimate the rotations of each projection. Finally, the 3-dimensional volume of the bio-molecule is restored using the Fourier based Iterative Reconstruction Method [3].

## 1 Introduction

The paper describes in detail an algorithm for estimating the 3-dimensional volume of the bio-molecule $V$, Given a set of $n$ 2-dimensional projections of a bio-molecule with $T$ or $O$ symmetry $\{P_{R_i}\}_{i\in[n]}$ taken at unknown rotations $\{R_i\}_{i\in[n]}$. The algorithm builds upon the Fourier projection slice theorem that states that

$$\hat{P}_{R_i}(\omega_x,\omega_y)=\hat{\psi}(\omega_x R_i^1+\omega_y R_i^2),\quad (\omega_x,\omega_y)\in R^2 \tag{1}$$

the Fourier transform of the projection image $\hat{P}_{R_i}(\omega_x,\omega_y)$ at the cartesian coordinates $(\omega_x,\omega_y)$ is equal to the Fourier transform $\hat{\psi}$ of the electrostatic potential function of the molecule $\psi$ at the coordinates $(\omega_x R_i^1+\omega_y R_i^2)\in R^3$. Where $R_i^1$ and $R_i^2$ are the first and second column of the rotation matrix $R_i$. The set of coordinates $(\omega_x R_i^1+\{\omega_y R_i^2)|(\omega_x,\omega_y)\in R^2\}$ describes a 2-dimensional 'slice' that goes through the center of the 3-dimensional volume rotated by $R_i$. The bio-molecule in question has tetrahedral or octahedral symmetry, and therefore suffers from symmetry ambiguity. Meaning that $R_i$ will produce the same common line with $g_kR_j$ for every $g_k\in G$ where $G$ is the symmetry group in question, either $G=T\subset SO(3)$ or $G=O\subset SO(3)$. Each two slices $P_{R_i}$ and $P_{g_kR_j}$ with different rotations $R_i$, $g_kR_j$ intersect at a line that passes through the center of the volume as well as the center of the projection images, referred to as the "common line" and defined by

$$q^k_{R_i,R_j}=\frac{R_i^3\times g^{(k)}R_j^3}{\left\|R_i^3\times g^{(k)}R_j^3\right\|} \tag{2}$$

The common line in each projection is located at different angles in polar coordinates, denoted $\alpha_{R_i,R_j}$ and $\beta_{R_i,R_j}$. Namely,

$$\hat{P}_{R_i}(\xi\cos(\alpha_{R_i,R_j}),\xi\sin(\alpha_{R_i,R_j}))=\hat{P}_{R_j}(\xi\cos(\beta_{R_i,R_j}),\xi\sin(\beta_{R_i,R_j}))\quad \xi\in R \tag{3}$$

And the Fourier representation of common lines can be defined with

$$\hat{q}^k_{R_j,R_i}=\cos(\alpha^k_{R_i,R_j})\,R_i^1+\sin(\alpha^k_{R_i,R_j})\,R_i^2=\cos(\beta^k_{R_i,R_j})\,g_kR_j^1+\sin(\beta^k_{R_i,R_j})\,g_kR_j^2. \tag{4}$$

From this equation the angles can be reproduced from $R_i^Tg_kR_j$ with The algorithm also makes use of relative rotations noted as $R_{ij}$ and defined as satisfying

$$\left\{R_{ij}^Tg_kR_{ji}\right\}_{k=1}^n=\left\{R_i^Tg_kR_j\right\}_{k=1}^n,\quad g_k\in G. \tag{5}$$

This report begins by describing all stages of the algorithm to reconstruct the bio-molecule. Followed by results showing the accuracy of the reconstruction with regards to a reference volume both for simulated and experimental data.

# 2 The Alogrithm

## 2.1 Estimate relative rotations

To find the relative rotations between two projections $P_{R_i}$ and $P_{R_j}$ the method constructs a score function $\pi_{ij}(R_q, R_s)$ for any two rotations $R_q, R_s \in SO(3)$ that estimates how close they are to the relative rotations $\{R_{ij}, R_{ji}\}$. This is done by taking the polar Fourier transform of each image, denoted as $\hat{P}_i$ and $\hat{P}_j$. Each column in the transformed projection $\hat{P}_i$ is the Fourier transform of a ray starting from the center of the image at a certain angle $\varphi$, denoted $\hat{P}_{i,\varphi}$. Cross correlation between two rays of projections $\hat{P}_i$ and $\hat{P}_j$ at angles $\varphi$ and $\theta$ is denoted as

$$\rho_{ij}(\theta, \phi) = \Re \int_0^\infty \frac{(\hat{P}_{i,\varphi}(\xi))^* \hat{P}_{j,\theta}(\xi) d\xi}{\|\hat{P}_{i,\varphi}(\xi)\|_{L_2} \|\hat{P}_{j,\theta}(\xi)\|_{L_2}}. \tag{6}$$

The method uses the common line angles $\alpha^k_{R_q,R_s}, \beta^k_{R_q,R_s}$ to find the correlation between $R_q$ and $R_s$. For any two rotations the correlation should be maximal both at the common lines and their respective self common lines. The score function is therefore defined by

$$\pi_{ij}(R_q, R_s) = \prod_{k\in[n]} \rho(\alpha^k_{R_q,R_s}, \beta^k_{R_q,R_s}) \prod_{k\in\{2,..,n\}} \rho(\alpha^k_{R_s,R_s}, \beta^k_{R_s,R_s}) \rho(\alpha^k_{R_q,R_q}, \beta^k_{R_q,R_q}). \tag{7}$$

The relative rotations are hence estimated by

$$\{\tilde{R}_{ij}, \tilde{R}_{ji}\} = \arg\max_{R_q,R_s\in SO(3)} \pi_{ij}(R_q, R_s) \tag{8}$$

Finding the correct rotations requires running through all pairs in $SO(3)$. A subset of rotations $SO_G(3) \subset SO(3)$ with a modifiable resolution to use as candidates for the algorithm.

## 2.2 Handedness synchronization

The values along a common line of the projection stay the same even if the projection is inverted along this line. Hence the previous algorithm can output a "differently handed" relative rotations for each projection pair. Each pair relative rotations $\{\tilde{R}_{ij}, \tilde{R}_{ij}\}$ can estimate either $\{R_{ij}, R_{ij}\}$ or $\{JR_{ij}J, JR_{ij}J\}$ with $J$ denoted as the reflection matrix $diag(1, 1, -1)$. To solve this ambiguity, the algorithm applies the J synchronization method in which will result in either all the rotations estimating the set $\{R_{ij}, R_{ij}\}_{i>j\in[n]}$ or they all estimate the set $\{JR_{ij}J, JR_{ij}J\}_{i>j\in[n]}$. The synchronization method is done by finding the handedness between triplets of rotations $(\tilde{R}_{ij}, \tilde{R}_{jk}, \tilde{R}_{ki})_{i<j<k}$ by minimizing

$$C(\mu_{ij}, \mu_{jk}, \mu_{ki}) = ||J^{\mu_{ij}} \tilde{R}_{ij} J^{\mu_{ij}} J^{\mu_{jk}} \tilde{R}_{jk} J^{\mu_{jk}} J^{\mu_{ki}} \tilde{R}_{ki} J^{\mu_{ki}} - I|| \tag{9}$$

where $\mu_{ij}, \mu_{jk}, \mu_{ki} \in \{+1, -1\}$ and the results are denoted as

$$\mu^*_{ij}, \mu^*_{jk}, \mu^*_{ki} = \arg\min_{\mu_{ij},\mu_{jk},\mu_{ki}} C(\mu_{ij}, \mu_{jk}, \mu_{ki}) \tag{10}$$

The method then constructs the adjacency matrix $\Sigma^*$ of the graph $\Sigma = (V, E)$ where vertices $V$ are the rotations $\{\tilde{R}_{ij}\}_{i>j\in[n]}$ and the edges $E$ where each edge between $R_{ij}$ and $R_{kl}$ are denoted by

$$\Sigma_{(i,j),(k,l)} = \begin{cases} 1, & \text{if } |\{i,j\} \cap \{k,l\}| = 1, \text{ and } \mu_{ij} = \mu_{kl}, \\ -1, & \text{if } |\{i,j\} \cap \{k,l\}| = 1, \text{ and } \mu_{ij} \neq \mu_{kl}, \\ 0, & \text{if } |\{i,j\} \cap \{k,l\}| = 0. \end{cases}$$

The method then retrieves the eigenvector of the leading eigenvalue. Finding the eigenvalue and eigenvector is preformed by the signs power method in order to reduce computation time complexity. The value of the eigenvector at index $(i, j)$ corresponds to the handedness of the rotation $\tilde{R}_{ij}$. The synchronization algorithm is described in detail in [3].

## 2.3 Estimate the rotations for volumes with T and O symmetry

Due to the symmetry ambiguity of the molecule, it is sufficient to estimate the set $\{g_\tau g_i R_i\}_{i\in[n]}$ where $g_\tau, g_i$ are arbitrary rotations in the symmetry group $O$ or $T$. $g_\tau$ is applied on all rotations, while $g_i$ is applied on rotation $R_i$. Denote $v_i^{(\sigma_i(\tau(m)))}$ as the $\tau(m)$'th row of $g_i R_i$ (or the $\sigma_i(\tau(m))$'th row of $R_i$) where $\tau(m)$ and $\sigma_i(m)$ are permutations on $m \in \{1,2,3\}$ corresponding to $g_\tau$ and $g_i$ respectively. Further, denote $v_{\sigma,\tau(m)}$ as the concatenation of all vectors in the $\tau(m)$ 'th row of each rotation $\{g_i R_i\}_{i\in[n]}$

$$v_{\sigma,\tau(m)} = (v_1^{(\sigma_i(\tau(m))}, ..., v_n^{(\sigma_i(\tau(m))}) \tag{11}$$

Finally denote the matrix $H_{\sigma,\tau(m)}$ by

$$H_{\sigma,\tau(m)} = v_{\sigma,\tau(m)}^T v_{\sigma,\tau(m)} \tag{12}$$

$H_{\sigma,\tau(m)}$ is a rank-1 $3N \times 3N$ block matrix whose $(i,j)$'th $3\times 3$ block is given by the rank-1 matrix $v_i^{(\sigma_i(\tau(m)))^T} v_j^{(\sigma_j(\tau(m)))}$. The Algorithm begins by constructing the $\tilde{H}_{\sigma,\tau(m)}$ matrices using the estimated relative rotations. Then the estimated $\tilde{v}_{\sigma,\tau(m)}$ vectors are produced by finding the eigenvector of the leading eigenvalue of $\tilde{H}_{\sigma,\tau(m)}$ with SVD. The values of $\hat{v}_{\sigma,\tau(m)}$ are used to construct the estimations $\{g_i \tilde{R}_i\}_{i\in[n]}$ of the $\{g_i R_i\}_{i\in[n]}$ rotations. The $\hat{H}_{\sigma,\tau(m)}$ matrices are constructed block by block. The paper shows that each block $(i,j)$ is equal to $R_{i,j} e_{kk} R_{j,i}$ for some unknown $k,l \in \{1,...,9\}$ where $e_{kl} \in R^{3\times 3}$ are single entry matrices with 1 at index $(k,l)$. First, the $3\times 3$ block in the (1,2)'th index, denoted as $\hat{H}^{(1,2)}_{\sigma,\tau(m)}$, is computed by

$$\hat{H}^{(1,2)}_{\sigma,\tau(m)} = \hat{R}_{12} e_{mm} \hat{R}_{21} \tag{13}$$

This equation is proved in proposition 8 in the paper by leveraging the fact that $N_{SO(3)}(O) = O$ and $N_{SO(3)}(T) = T$ where $N_{SO(3)}(O)$ is the normalizer of $SO(3)$ on $O$ defined by $N_{\tilde{H}}(\tilde{G}) = \left\{\tilde{h} \in \tilde{H} : \tilde{h}^T \tilde{G} \tilde{h} = \tilde{G}\right\}$. Next the rest of the first row of blocks $(1,i)_{i\in\{3,..,n\}}$ are computed by

$$\tilde{H}^{(1,i)}_{\sigma,m}|_{i\in 3,...,n} = \underset{M\in\{\tilde{R}_{1,i}^T e_{rr} \tilde{R}_{i,1}\}_{r=1,2,3}}{\arg\max} \left\| \tilde{H}^{(1,2)^T}_{\sigma,m} M \right\|_F. \tag{14}$$

Explained in page 17 in the paper that only when the row indices are aligned, $\sigma_{12}(\tau_{12}(m)) = \pm\sigma_{1i}(\tau_{1i}(r))$, the value in the maximizing function is non zero. Then we can compute the rest of the upper triangle by

$$\tilde{H}^{(i,j)}_{\sigma,m}|_{i<j} = \underset{M\in\{\pm\tilde{R}_{ij}^T e_{kl} \tilde{R}_{ji}\}_{k,l=1,2,3}}{\arg\min} \left\| M - \tilde{H}^{(1,i)^T}_{\sigma,m} \tilde{H}^{(1,j)}_{m} \right\|_F. \tag{15}$$

Explained in proposition 9 in the paper. $\hat{H}_{\sigma,\tau(m)}$ is hermitian since $(v^T_{\sigma,\tau(m)} v_{\sigma,\tau(m)}) = v^T_{\sigma,\tau(m)} v_{\sigma,\tau(m)}$. Hence the lower triangle is given by the transposed matrix at the corresponding index of the upper triangle.

$$\tilde{H}^{(i,j)}_{\sigma,m}|_{i>j} = \tilde{H}^{(j,i)^T}_{\sigma,m} \tag{16}$$

Finally, the diagonal is computed by

$$\tilde{H}^{(i,i)}_{\sigma,m} = \frac{1}{N-1} \sum_{\substack{j=1 \\ j\neq i}}^{N} \tilde{H}^{(j,i)^T}_{\sigma,m} \tilde{H}^{(j,i)}_{\sigma,m} \tag{17}$$

Explained in page 19 of the paper.

## 2.4 Fourier based Iterative Reconstruction Method

The Fourier based Iterative Reconstruction Method (FIRM) is used to reconstruct the 3-dimensional volume from the 2-dimensional projections given their estimated rotations. The method is presented in the paper [**?**]. The paper begins by describing the forward-projector function that obtains the Fourier transform of the 2-dimensional projection defined as

$$(\mathcal{A}(V))(k_1,k_2,m) = \sum_{\mathbf{j}} V_{\mathbf{j}} \exp\left(-i \cdot \langle \mathbf{j}, R_m^{-1}(\omega_{k_1}, \omega_{k_2}, 0)\rangle\right) \cdot h_m\left(\|\omega_{k_1,k_2}\|\right) \tag{18}$$

Where $V$ is the given volume, $\omega_{k_1}, \omega_{k_2} = 2\pi/n$ are cartesian coordinates of $k_1, k_2 \in Z$, $m$ is the index of the projection with rotation matrix $R_m$. $V_{\mathbf{j}}$ is the voxel in index $\mathbf{j} \in Z^3$ of the volume $V$. $h_m$ is the contrast transfer function (CTF) of the electron microscope for projection $m$. For this algorithm the CTF is set to $h_m \equiv 1$. From the forward-projector function a back-projector function is derived as

$$(\mathcal{A}^* g)(\mathbf{j}) = \sum_{m=1}^{M} \sum_{\mathbf{k}} \exp\left(i \cdot \langle \mathbf{j}, R_m^{-1}(\omega_{k_1}, \omega_{k_2}, 0)\rangle\right) \cdot h_m\left(\|\omega_{k_1,k_2}\|\right) \cdot g_{k_1,k_2,m}. \tag{19}$$

where $g$ is the set of projections and $g_{k_1,k_2,m}$ is the pixel at the index $(k_1, k_2)$ of projection $m$. The task at hand is finding the volume $V$ that when projected produces the projections $g$ :

$$\mathcal{A}(V) = g \tag{20}$$

applying the back-projector to each side results in

$$\mathcal{A}^*\mathcal{A}(V) = \mathcal{A}^* g \tag{21}$$

The paper then derives the convolution kernel **Ker** that satisfies

$$\mathcal{A}^*\mathcal{A}(V) = \mathbf{Ker} \star V \tag{22}$$

denoted by

$$\mathbf{Ker}(\mathbf{n}) = \sum_{m=1}^{M} \sum_{\mathbf{k}} \exp\left(i \cdot \langle \mathbf{n}, R_m^{-1}(\omega_{k_1}, \omega_{k_2}, 0)\rangle\right) \cdot h_m\left(\|\omega_{k_1,k_2}\|\right)^2 \tag{23}$$

Now the task can be described as

$$\mathbf{Ker} \star V = \mathcal{A}^* g \tag{24}$$

The algorithm first back-projects the images at the estimated rotations to compute $\mathcal{A}^* g$. Then computes the kernel given the rotations. Finally, since the constructed **Ker** is positive semi-definite, the Conjugate gradient method is then used to estimate the volume $V$.

# 3 Results

## 3.1 Simulated data

The reconstruction is preformed on projections of the reference molecule EMD-4179 [4]. Projections are taken by preforming a Non-uniform Fast Fourier transform on the reference volume along random rotation grids. For variety, and in order to show a more gradual decline in accuracy, a more granular selection of the number of projections and Signal to Noise Ratio (SNR) is presented.

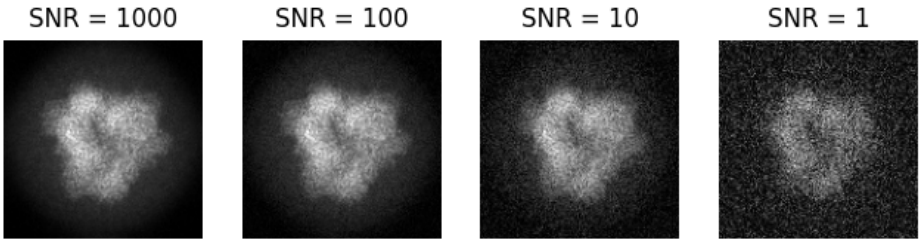


Figure 1: Simulated projection images of EMD-4179 ($T$ symmetry) with different values for Signal to Noise Ratio (SNR)

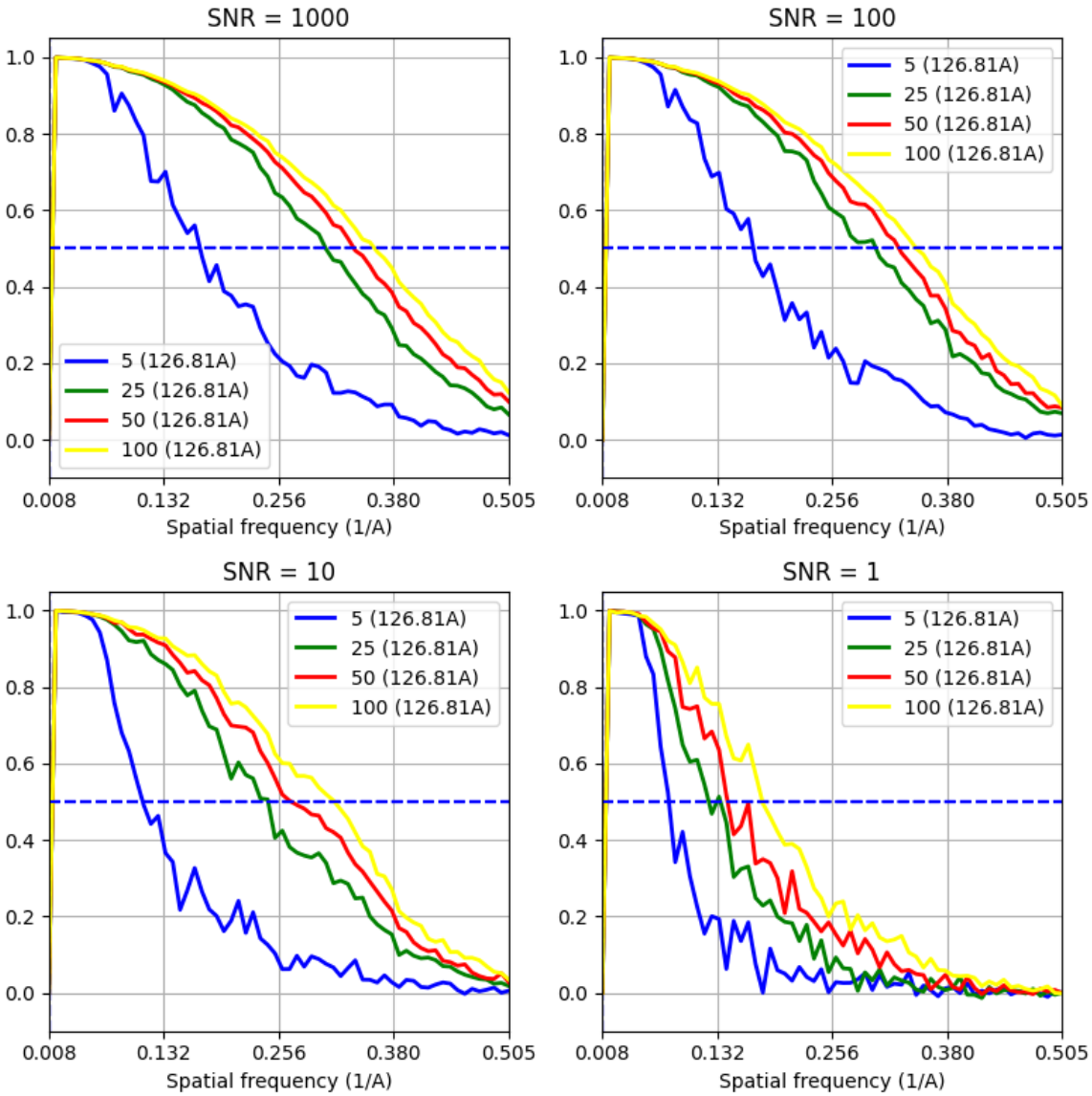


Figure 2: Fourier shell correlation curves for volumes reconstructed from simulated projections of EMD-4179 ($T$ symmetry). Each plot represents a different signal to noise ratio (SNR), while each color represents a different number of projections used.

((a)) ((b))

Figure 3: Comparison between the reference volume EMD-4179 ($T$ symmetry) (a) and the reconstructed volume (b) with 100 projection slices and 1000 SNR.

### 3.2 Experimental data

An important distinction between the algorithm used in the paper and the algorithm reproduced for this report is the lack of regard for shifts. For simulated data this has no effect since the reference volume is aligned and centered. For experimental data this can sway the results dramatically. The reconstruction is preformed on the raw data projections from EMPIAR-10026 [5] which corresponds to the volume EMD-2788 [6] reconstructed using RELION. The data set provided is made up of 483 projections with $132 \times 132$ pixels each. In order to save computation time, a 100 out of 483 images were selected randomly for reconstruction. The projection images seem to have a very high SNR, predicting a grim outcome for the accuracy of the reconstruction.

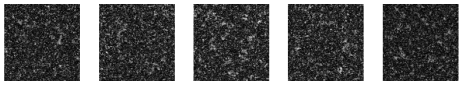

Figure 4: Randomly selected images from the raw data provided for EMPIAR-10026 ($O$ symmetry)

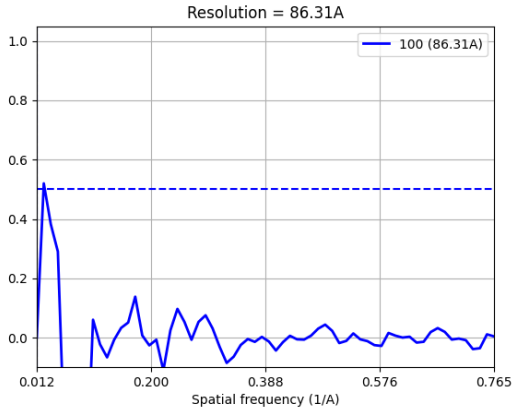


Figure 5: Fourier shell correlation curve for volume reconstructed from raw data of EMPIAR-10026 ($O$ symmetry) with the reconstructed volume by RELION.

((a)) ((b))

Figure 6: Comparison between the volume reconstructed by RELION - ($O$ symmetry) (a) and the volume reconstructed by the proposed algorithm (b) with 100 projection slices from the dataset EMPIAR-10026.

## References


[1] Adi Shasha Geva. A common lines approach for ab-initio modeling of molecules with tetrahedral and octahedral symmetry. *SIAM Journal on Imaging Sciences*, 16(4):1978–2014, 2022.

[2] Yoel Shkolnisky. Synchronization. *Methods in bioimaging 0372-4034*, 2024.

[3] Yoel Shkolnisky. A fourier-based approach for iterative 3d reconstruction from cryo-em images. *http://spr.math.princeton.edu/sites/spr.math.princeton.edu/files/reconstruction$_v$4.pdf*, 2024.

[4] Schwieters CD Gauto DF, Estrozi LF. Combined solid-state nmr, solution-state nmr and em data for structure determination of the tetrahedral aminopeptidase tet2 from p. horikoshii. *https://www.ebi.ac.uk/emdb/EMD-4179*, 2017.

[5] Russo CJ. Raw data for the 3d structure of horse spleen apoferritin determined by electron cryomicroscopy. *https://www.ebi.ac.uk/empiar/EMPIAR-10026/*, 2014.

[6] Russo CJ. 3d structure of horse spleen apoferritin determined by electron cryomicroscopy. *https://www.ebi.ac.uk/emdb/EMD-2788?tab=overview*, 2014.