Statistics and classification of spatial rotations and projection directions considering molecular symmetry in 3D electron cryo-microscopy

Qi Zhang[1, 2, 3], Hai Lin[†4, 5], and Mingxu Hu[†1, 2, 3]

[1]Key Laboratory of Protein Sciences (Tsinghua University), Ministry of Education, Beijing, China

[2]School of Life Science, Tsinghua University, Beijing, China

[3]Beijing Advanced Innovation Center for Structural Biology

[4]Department of Mathematical Sciences, Tsinghua University, Beijing, China

[5]Yau Mathematical Sciences Center, Tsinghua University, Beijing, China

†Correspondence should be addressed to M.H. (humingxu@mail.tsinghua.edu.cn) and H.L. (hailin@mail.tsinghua.edu.cn).

Abstract

Electron cryo-microscopy (cryoEM) three-dimensional (3D) reconstruction is based on estimations of orientations of projection images or 3D volumes. It is common that the macromolecules studied by cryoEM have molecular symmetry. However, for neither spatial rotations nor projection directions, currently there is no statistics and classification method which takes molecular symmetry into consideration. This limits the applications of advanced statistical methods in cryoEM. In this research, based on the concept of non-unique-games (NUG), we propose statistics and classification methods of both spatial rotations and projection directions when molecular symmetry is considered. Such methods are realized in the open-source python package pySymStat.

1 Introduction

In the three-dimensional (3D) reconstruction of electron cryo-microscopy (cryoEM), particle images are projections of biological samples at random orientations. Therefore, the analysis of such orientations is the key step in the reconstruction of 3D density maps[1, 2, 3, 4, 5, 6, 7, 8]. Accordingly, the analysis of both spatial rotations and projection directions is fundamentally important for cryoEM. Algorithms are being developed for such analyses; there are growing demands for adopting advanced statistical methods in their development. It is common that the macromolecules studied by cryoEM have molecular symmetry[9, 10]. However, the true topology induced by the molecular symmetry is currently neglected as there are no statistics and classification methods considering molecular symmetry. This limits the applications of advanced statistical methods in cryoEM.

In this study, we propose statistics and classification methods of both spatial rotations and projection directions which take molecular symmetry into consideration. Such methods are realized in the open-source python package pySymStat. The unit quaternion[11, 12, 13] is widely used for describing rotations; however, recognizing the majority of the readers might be unfamiliar with it, Section 2 gives a brief overview as well as clarifies the notations used in this work. In Section 3, distance, variance, and mean are defined in the context of considering molecular symmetry. In Section 4, such variance and mean problems are solved by a method developed based on non-unique-games (NUG)[14, 15]. NUG is a mathematical framework for solving a type of optimization problem defined over a locally compact group. The first step in computing the variance and mean is converting it into a NUG problem. Then, NUG produces a convex combination[16] of possible solutions. Finally, a leaderboard algorithm is proposed for recovering the true solution from the possible solutions. In Section 5, the performance of the proposed methods is analyzed. After determining the distance, variance, and mean, classification is introduced in Section 6. The potential applications of the proposed methods are discussed in Section 7.

2 Background of the unit quaternion description

The unit quaternion description of rotation has been well studied and applied in many fields, including robotics[17, 18, 19, 20], aviation[21, 22, 23, 24], computational vision[25, 26, 27], as well as cryoEM[28, 29, 30, 31, 32, 7, 33, 5, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43]. In the present work, we give a very brief introduction of the unit quaternion description of spatial rotations. For general introductions of quaternion, we refer the reader to [12, 13, 44, 45].

Notations in this article are specified according to the following rules. The unit quaternion in vector form, or saying a unit four-dimensional (4D) vector, are denoted by a letter with a bold font, for example, $\mathbf{q}$. The molecular symmetry group is denoted by G. An element in the molecular symmetry group G is denoted by $\mathbf{g}$. Such an element is also in the unit quaternion description. Meanwhile, the number of elements in G is denoted by $|G|$. The set containing all 3D spatial rotations is denoted by $SO(3)$.

A quaternion is an extension of a complex number, composed of a real part and three imaginary parts. The quaternion is commonly viewed as a 4D vector. If its norm equals 1, i.e., $|\mathbf{q}| = 1$, the quaternion is called a unit quaternion. Quaternion multiplication follows special rules, the definition of which can be

found in many references[12, 13]. The quaternion multiplication operator is denoted by $\otimes$ in the present work.

The rotation of a 3D vector $\vec{\mathbf{v}}$ to another vector $\vec{\mathbf{v}}'$ can be achieved by quaternion multiplication. By viewing $\vec{\mathbf{v}}$ as $\mathbf{v} = \begin{pmatrix} 0 \\ \vec{\mathbf{v}} \end{pmatrix}$ in quaternion form, the rotation induced by a unit quaternion is defined by[12]

$$\mathbf{v}' = \begin{pmatrix} 0 \\ \vec{\mathbf{v}}' \end{pmatrix} = \mathbf{q} \otimes \begin{pmatrix} 0 \\ \vec{\mathbf{v}} \end{pmatrix} \otimes \mathbf{q}^*, \tag{1}$$

where $\mathbf{q}^*$ is the conjugate[45, 12, 13] of $\mathbf{q}$. Therefore, $\mathbf{q}$ is a description of such spatial rotation. Such rotation is also denoted by

$$\vec{\mathbf{v}}' = R_{\mathbf{q}}(\vec{\mathbf{v}}). \tag{2}$$

For sequential rotations, such as a spatial rotation $\mathbf{q}_1$ followed by a spatial rotation $\mathbf{q}_2$, the overall rotation is calculated as $\mathbf{q}_2 \otimes \mathbf{q}_1$[12]. When molecular symmetry is not considered, the distance between the two spatial rotations $\mathbf{q}_1$ and $\mathbf{q}_2$ can be defined by[45].

$$d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_2) := \max_{\vec{\mathbf{v}} \in S^2} d_{S^2}(R_{\mathbf{q}_1}(\vec{\mathbf{v}}), R_{\mathbf{q}_2}(\vec{\mathbf{v}})), \tag{3}$$

where d_{S^2} is the natural distance in the unit sphere S^2.

Meanwhile, there are several different definitions of the projection direction in the cryoEM field. Some of them, however, are ambiguous. Hence, we explicitly define the projection direction as follows, in the present work. We denote the operator of projection for a spatial rotation $\mathbf{q}$ as $\mathcal{P}^{\mathbf{q}}$. In other words, $\mathcal{P}^{\mathbf{q}}(V)$ stands for the image obtained as follows. The map V is rotated by spatial rotation $\mathbf{q}$; then, the rotated map is projected along the Z axis. The defining of projection direction $\vec{\mathbf{v}}_p$ can be achieved by defining the set of projected images along $\vec{\mathbf{v}}_p$, which is denoted by $\mathbf{P}_{\vec{\mathbf{v}}_p}$. This set is defined by

$$\mathbf{P}_{\vec{\mathbf{v}}_p} := \{\mathcal{P}^{\mathbf{q}}(V)\}_{\left\{\mathbf{q} \middle| R_{\mathbf{q}}(\vec{\mathbf{v}}_p) = \vec{\mathbf{v}}_n\right\}}, \tag{4}$$

where $\vec{\mathbf{v}}_n = (0, 0, 1)^T$ is the unit vector along the Z axis. In other words, every projection image in this set is projected along the Z axis from a rotated volume. Such a rotation transforms $\vec{\mathbf{v}}_p$ to $\vec{\mathbf{v}}_n$. For two projection directions, $\vec{\mathbf{v}}_{p_1}$ and $\vec{\mathbf{v}}_{p_2}$, they are equal if and only if their corresponding sets of projected images are equal, i.e., $\mathbf{P}_{\vec{\mathbf{v}}_{p_1}} = \mathbf{P}_{\vec{\mathbf{v}}_{p_2}}$. The definition of the distance between the two projection directions $\vec{\mathbf{v}}_{p_1}$ and $\vec{\mathbf{v}}_{p_2}$ is assigned by the natural distance in the unit sphere S^2, when molecular symmetry is not taken into consideration.

3 Definitions of distance, variance, and mean considering molecular symmetry

The variance and the mean, two of the most basic statistics, need to be defined in the context of molecular symmetry. The distance definition serves as the prerequisite of the variance and mean definitions.

For considering molecular symmetry, the set of equivalent spatial rotations and the set of equivalent projection directions are defined first (Section 3.1). In the present work, the distances between two spatial rotations and two projection directions are defined by the distances between the corresponding two sets of equivalent spatial rotations and two sets of equivalent projection directions, respectively (Section 3.2). The variance and mean are defined based on theses distance definitions (Section 3.3).

3.1 The set of equivalent spatial rotations and the set of equivalent projection directions considering molecular symmetry

The symmetry property of a molecule can be represented by a molecular symmetry group. For a molecule with symmetry G, the spatial rotation $\mathbf{q}$ is equivalent to $\mathbf{q} \otimes \mathbf{g}_m$ for any $\mathbf{g}_m$ in a molecular symmetry group G. This is proven in Appendix I. We denote all of these equivalent spatial rotations of $\mathbf{q}$ as $\mathbf{q} \otimes G$, defined by

$$\mathbf{q} \otimes G := \{\mathbf{q} \otimes \mathbf{g}_m\}_{1 \leq m \leq |G|} . \tag{5}$$

In mathematical terms, $\mathbf{q} \otimes G$ is a left coset[46, 47] of G in $SO(3)$. An element in the set of equivalent spatial rotations $\mathbf{q} \otimes G$, i.e., $\mathbf{q} \otimes \mathbf{g}_m$, is referred to as a representative[46, 47] of this set.

Meanwhile, when the molecular symmetry G is taken into consideration, the projection direction $\vec{\mathbf{v}}_p$ is equivalent to $R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)$ for any $\mathbf{g}_m$ in G; this is proven in Appendix II. Here, an inverse operation of $\mathbf{g}_m$ exists for the completeness of the concept. $R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)$ is the equivalent projection direction generated by rotating the map V by the symmetry element $\mathbf{g}_m$ (see Appendix II).

We denote all of these equivalent projection directions of $\vec{\mathbf{v}}_p$ as $R_{G^{-1}}(\vec{\mathbf{v}}_p)$, defined by

$$R_{G^{-1}}(\vec{\mathbf{v}}_p) := \left\{R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)\right\}_{1 \leq m \leq |G|} . \tag{6}$$

An element in the set of equivalent projection direction $R_{G^{-1}}(\vec{\mathbf{v}}_p)$, i.e., $R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)$, is referred to as a representative of this set.

For a set of spatial rotations $\{\mathbf{q}_i\}_{1 \leq i \leq N}$ and a set of projection directions $\{\vec{\mathbf{v}}_{p_i}\}_{1 \leq i \leq N}$, their statistical analyses are actually performed on $\{\mathbf{q}_i \otimes G\}_{1 \leq i \leq N}$ and $\{R_{G^{-1}}(\vec{\mathbf{v}}_{p_i})\}_{1 \leq i \leq N}$ respectively, when the molecular symmetry G is taken into consideration.

3.2 Definitions of the distance between two spatial rotations and the distance between two projection directions considering molecular symmetry

Definitions of the distance between two sets of equivalent spatial rotations and the distance between two sets of equivalent projection directions are the prerequisites of considering the statistics and classification of spatial rotations and projection directions, respectively, when molecular symmetry is taken into consideration.

It is a common approach to define the distance between two sets by the distance between their

representatives[48]. For spatial rotations, we define the distance between $\mathbf{q}_1 \otimes G$ and $\mathbf{q}_2 \otimes G$ by

$$d_{SO(3)/G}\left(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G\right) := \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2} \in G} d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}), \tag{7}$$

which is the shortest distance in $SO(3)$ by selecting representatives $\mathbf{q}_1 \otimes \mathbf{g}_{t_1}$ and $\mathbf{q}_2 \otimes \mathbf{g}_{t_2}$ from $\mathbf{q}_1 \otimes G$ and $\mathbf{q}_2 \otimes G$, respectively.

Similar to the distance definition of two spatial rotations, we define the distance between two sets of equivalent projection directions $R_G\left(\vec{\mathbf{v}}_{p_1}\right)$ and $R_G\left(\vec{\mathbf{v}}_{p_2}\right)$ by

$$d_{S^2/G}\left(R_G\left(\vec{\mathbf{v}}_{p_1}\right), R_G\left(\vec{\mathbf{v}}_{p_2}\right)\right) := \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2} \in G} d_{S^2}\left(R_{\mathbf{g}_{t_1}^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right), R_{\mathbf{g}_{t_2}^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)\right), \tag{8}$$

which is the shortest distance in S^2 by selecting representatives $R_{\mathbf{g}_{t_1}^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right)$ and $R_{\mathbf{g}_{t_2}^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$ from $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right)$ and $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$, respectively. d_{S^2} is the natural distance in the unit sphere S^2.

These definitions, i.e., Equation (7) and Equation (8), both satisfy all of the requirements of the distance definition in mathematics (see Appendix III and Appendix IV, respectively), namely, the symmetry property, the positive definiteness property, and the triangle inequality property.

3.3 Definition of variance and mean of spatial rotations and projection directions considering molecular symmetry

In the same spirit of the distance definition, we define the variance of spatial projections $\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}$ and projection directions $\{R_G\left(\vec{\mathbf{v}}_{p_i}\right)\}_{1 \le i \le N}$, respectively as

$$\mathrm{Var}\left(\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}\right) := \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2}, \cdots, \mathbf{g}_{t_N} \in G} \left(\sum_{1 \le i,j \le N} d^2_{SO(3)}\left(\mathbf{q}_i \otimes \mathbf{g}_{t_i}, \mathbf{q}_j \otimes \mathbf{g}_{t_j}\right)\right) \tag{9}$$

and

$$\mathrm{Var}\left(\{R_G\left(\vec{\mathbf{v}}_{p_i}\right)\}_{1 \le i \le N}\right) := \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2}, \cdots, \mathbf{g}_{t_N} \in G} \left(\sum_{1 \le i,j \le N} d^2_{S^2}\left(R_{\mathbf{g}_{t_i}^{-1}}\left(\vec{\mathbf{v}}_{p_i}\right), R_{\mathbf{g}_{t_j}^{-1}}\left(\vec{\mathbf{v}}_{p_j}\right)\right)\right). \tag{10}$$

In other words, for spatial rotations, a representative $\mathbf{q}_i \otimes \mathbf{g}_{t_i}$ is chosen from each set of equivalent spatial rotations $\mathbf{q}_i \otimes G$, such that all representatives are as close as possible overall. The projection direction follows the same principle. Furthermore, we define the mean of spatial rotations $\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}$ and projection directions $\{R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_i}\right)\}_{1 \le i \le N}$, respectively as

$$\mathrm{Mean}\left(\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}\right) := \left(\operatorname*{argmin}_{\mathbf{q}} \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2}, \cdots, \mathbf{g}_{t_N} \in G} \sum_{i=1}^{N} d^2_{SO(3)}\left(\mathbf{q}, \mathbf{q}_i \otimes \mathbf{g}_{t_i}\right)\right) \otimes G \tag{11}$$

and

$$\mathrm{Mean}\left(\{R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_i}\right)\}_{1 \le i \le N}\right) := R_{G^{-1}}\left(\operatorname*{argmin}_{\vec{\mathbf{v}}_p} \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2}, \cdots, \mathbf{g}_{t_N} \in G} \sum_{i=1}^{N} d^2_{S^2}\left(\vec{\mathbf{v}}_p, R_{\mathbf{g}_{t_i}^{-1}}\left(\vec{\mathbf{v}}_{p_i}\right)\right)\right). \tag{12}$$

In other words, for spatial rotations, the mean of $\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}$ is defined by the spatial rotation closest to $\{\mathbf{q}_i \otimes G\}_{1 \le i \le N}$. The mean of the projection direction follows the same principle.

The problems of determining the variance and the mean are in some sense equivalent. For spatial rotations, the mean definition by Equation (11) can be approximated by (see Appendix V)

$$\mathrm{Mean}(\{\mathbf{q}_i \otimes G\}_{1\le i\le N}) \approx \left(\underset{\mathbf{q}}{\mathrm{argmin}} \sum_{i=1}^{N} d^2_{SO(3)}(\mathbf{q}, \mathbf{q}_i \otimes \tilde{\mathbf{g}}_{t_i})\right) \otimes G, \tag{13}$$

where

$$\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N} = \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\mathrm{argmin}} \left(\sum_{1\le i,j\le N} d^2_{SO(3)}(\mathbf{q}_i \otimes \mathbf{g}_{t_i}, \mathbf{q}_j \otimes \mathbf{g}_{t_j})\right) \tag{14}$$

and $\mathrm{argmin}_{\mathbf{q}} \sum_{i=1}^{N} d^2_{SO(3)}(\mathbf{q}, \mathbf{q}_i \otimes \tilde{\mathbf{g}}_{t_i})$ is solvable by an established method[28] once $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$ are determined. Meanwhile, for projection directions, the mean definition by Equation (12) can be approximated by (see Appendix VI)

$$\mathrm{Mean}(\{R_{G^{-1}}(\vec{\mathbf{v}}_{p_i})\}_{1\le i\le N}) \approx R_{G^{-1}}\left(\underset{\vec{\mathbf{v}}_p}{\mathrm{argmin}} \sum_{i=1}^{N} d^2_{S^2}\left(\vec{\mathbf{v}}_p, R_{\tilde{\mathbf{g}}_{t_i}^{-1}}(\vec{\mathbf{v}}_{p_i})\right)\right), \tag{15}$$

where

$$\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N} = \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\mathrm{argmin}} \left(\sum_{1\le i,j\le N} d^2_{S^2}\left(R_{\mathbf{g}_{t_i}^{-1}}(\vec{\mathbf{v}}_{p_i}), R_{\mathbf{g}_{t_j}^{-1}}(\vec{\mathbf{v}}_{p_j})\right)\right) \tag{16}$$

and $\mathrm{argmin}_{\vec{\mathbf{v}}_p} \sum_{i=1}^{N} d^2_{S^2}\left(\vec{\mathbf{v}}_p, R_{\tilde{\mathbf{g}}_{t_i}^{-1}}(\vec{\mathbf{v}}_{p_i})\right)$ is also solvable by an established method[49] once $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$ are determined.

In other words, the representatives chosen in the determination of the variances, i.e., Equation (9) and (10), determine $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$. The means of spatial rotations and projection directions can be approximated by Equation (13) and Equation (15), by using $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$.

In conclusion, choosing the overall-closest representatives, i.e., $\mathbf{q}_1 \otimes \tilde{\mathbf{g}}_{t_1}, \mathbf{q}_2 \otimes \tilde{\mathbf{g}}_{t_2}, \cdots, \mathbf{q}_N \otimes \tilde{\mathbf{g}}_{t_N}$ for spatial rotations, and $R_{\tilde{\mathbf{g}}_{t_1}^{-1}}(\vec{\mathbf{v}}_{p_1}), R_{\tilde{\mathbf{g}}_{t_2}^{-1}}(\vec{\mathbf{v}}_{p_2}), \cdots, R_{\tilde{\mathbf{g}}_{t_N}^{-1}}(\vec{\mathbf{v}}_{p_N})$ for projection directions, where $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$ are selected by Equation (14) and Equation (16), is the key step in computing the statistics of spatial rotations and projection directions, respectively, when molecular symmetry is taken into consideration. In other words, once Equation (14) and Equation (16) are solved, the variance and the mean of $\{\mathbf{q}_i \otimes G\}_{1\le i\le N}$ and $\{R_{G^{-1}}(\vec{\mathbf{v}}_{p_i})\}_{1\le i\le N}$ can be calculated, respectively.

4 NUG and the determination of representatives

In this section, we present an approximation algorithm for determining $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$ in both Equation (14) and Equation (16), which is equivalent to the variance and the mean determination, as discussed above. This approximation algorithm is based on the framework of non-unique games (NUG), developed by Lederman and Singer[14, 15]. Therefore, we refer to this algorithm as the NUG approach. The $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$ determination in both Equation (14) and Equation (16) is first translated into a NUG problem (see Section 4.1). Then, by making use of representation theory[50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 14, 15, 61] (see Section 4.2 for a brief overview), the NUG problem is relaxed into a semi-positive definite programming (SDP) problem[62, 63] (see Section 4.3 and Section 4.4), which can be solved by a convex optimization

technique[62, 16]. However, the solution of the SDP problem is a "relaxation" of the NUG problem. In other words, the constraint in the SDP problem is looser than that of the NUG problem. Therefore, in the present work, we propose a leaderboard algorithm for "re-constraining" the SDP solution into a proper solution for the NUG problem (see Section 4.5). Currently, such algorithms are mainly adopted as leaderboard peptide sequencing in bioinformatics[64, 65].

4.1 Determination of representatives is a NUG problem

The original NUG problem is defined by the following optimization problem

$$\underset{x_1,x_2,\ldots,x_N\in\mathcal{X}}{\operatorname{argmin}} \sum_{1\leq i,j\leq N} f_{ij}(x_i x_j^{-1}), \tag{17}$$

where $\mathcal{X}$ is a given locally compact group, and $\{f_{ij}\}_{1\leq i,j\leq N}$ are N^2 given real valued functions on $\mathcal{X}$.

The $\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N}$ determination for both spatial rotation statistics, i.e., Equation (14), and projection direction statistics, i.e., Equation (16), can be derived in the same form (see Appendix VII and Appendix VIII)

$$\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N} = \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\leq i,j\leq N} f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}) \right). \tag{18}$$

In the above equation, $f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})$ is assigned as (see Appendix VII),

$$f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}) = d^2_{SO(3)}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}, \mathbf{q}_i^{-1}\otimes\mathbf{q}_j) \tag{19}$$

for the representative determination of spatial rotations. Meanwhile, for the representative determination of projection directions, $f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})$ is assigned as (see Appendix VIII),

$$f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}) = d^2_{S^2}\left(\vec{\mathbf{v}}_{p_i}, R_{\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}}\left(\vec{\mathbf{v}}_{p_j}\right)\right). \tag{20}$$

Moreover, the molecular symmetry group G is a finite group. Hence, the determination of representatives meets all of the requirements of being a NUG problem. In the following, we solve this NUG problem for determining the variance and the mean for both spatial rotations and projection directions.

4.2 Recall of representation theory

As in the NUG problem, the variables in the optimization are elements of the molecular symmetry group G. However, it is often difficult to optimize over a finite group. Proposed by Singer and Lederman[14, 15], such problems can be converted into optimizations over matrices using representation theory[50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61]. Representation theory provides a general Fourier transform between a group and matrices, making such conversions possible.

As some readers might not be familiar with representation theory, in this section, some useful concepts and methods of representation theory are briefly presented. As the molecular symmetry group G is considered in the present work and it is a finite group, this overview is limited to representation theory of the finite group[57, 58, 59, 60, 61].

For a finite group G, it corresponds to several intrinsically different irreducible representations. The number of irreducible representations of G is denoted by K. The irreducible representations of G are denoted by $\{\boldsymbol{\rho}_k\}_{1\leq k\leq K}$. The irreducible representation $\boldsymbol{\rho}_k$ is a function, which maps every element in a group G to a real or complex $d_k \times d_k$ unitary matrix, where d_k is the dimensionality of the k-th irreducible representation. When G is given, its irreducible representations, i.e., $\{\boldsymbol{\rho}_k(\mathbf{g}_m)\}_{1\leq k\leq K}$ are known. The irreducible representations have the property described below.

By the Peter-Weyl theorem[66], a function $f : G \to \mathbb{R}$ has a general Fourier transformation, as follows. The k-th Fourier coefficient is defined by

$$\hat{\mathbf{f}}^{(k)} := \frac{1}{|G|}\sum_{m=1}^{|G|} f(\mathbf{g}_m)\boldsymbol{\rho}_k(\mathbf{g}_m)^*, \tag{21}$$

where $\boldsymbol{\rho}_k(\mathbf{g}_m)^*$ is the conjugate transpose of $\boldsymbol{\rho}_k(\mathbf{g}_m)$. Then, the general Fourier expansion of f is

$$f(\mathbf{g}_m) = \sum_{k=0}^{K-1} d_k \, tr\left(\hat{\mathbf{f}}^{(k)}\boldsymbol{\rho}_k(\mathbf{g}_m)\right), \tag{22}$$

where $tr\left(\hat{\mathbf{f}}^{(k)}\boldsymbol{\rho}_k(\mathbf{g}_m)\right)$ is the trace of the matrix $\hat{\mathbf{f}}^{(k)}\boldsymbol{\rho}_k(\mathbf{g}_m)$.

4.3 Matrix form of a NUG by representation theory

By a general Fourier transformation using representation theory, introduced as Equation (21) and Equation (22), the NUG problem, i.e., Equation (14), is converted into the matrix form as (see Appendix IX)

$$\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N} = \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \sum_{k=0}^{K-1} tr\left(\hat{\mathbf{F}}^{(k)}\mathbf{X}^{(k)}\right) \tag{23}$$

where

$$\hat{\mathbf{F}}^{(k)} = d_k \begin{pmatrix} \hat{\mathbf{f}}_{11}^{(k)} & \hat{\mathbf{f}}_{21}^{(k)} & \cdots & \hat{\mathbf{f}}_{N1}^{(k)} \\ \hat{\mathbf{f}}_{12}^{(k)} & \hat{\mathbf{f}}_{22}^{(k)} & \cdots & \hat{\mathbf{f}}_{N2}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{f}}_{1N}^{(k)} & \hat{\mathbf{f}}_{2N}^{(k)} & \cdots & \hat{\mathbf{f}}_{NN}^{(k)} \end{pmatrix}, \tag{24}$$

in which

$$\hat{\mathbf{f}}_{ij}^{(k)} = \sum_{m=1}^{|G|} f_{ij}(\mathbf{g}_m)\boldsymbol{\rho}_k^*(\mathbf{g}_m), \tag{25}$$

and

$$\mathbf{X}^{(k)} = \begin{pmatrix} \boldsymbol{\rho}_k(\mathbf{g}_{t_1}\otimes\mathbf{g}_{t_1}^{-1}) & \boldsymbol{\rho}_k(\mathbf{g}_{t_1}\otimes\mathbf{g}_{t_2}^{-1}) & \cdots & \boldsymbol{\rho}_k(\mathbf{g}_{t_1}\otimes\mathbf{g}_{t_N}^{-1}) \\ \boldsymbol{\rho}_k(\mathbf{g}_{t_2}\otimes\mathbf{g}_{t_1}^{-1}) & \boldsymbol{\rho}_k(\mathbf{g}_{t_2}\otimes\mathbf{g}_{t_2}^{-1}) & \cdots & \boldsymbol{\rho}_k(\mathbf{g}_{t_2}\otimes\mathbf{g}_{t_N}^{-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\rho}_k(\mathbf{g}_{t_N}\otimes\mathbf{g}_{t_1}^{-1}) & \boldsymbol{\rho}_k(\mathbf{g}_{t_N}\otimes\mathbf{g}_{t_2}^{-1}) & \cdots & \boldsymbol{\rho}_k(\mathbf{g}_{t_N}\otimes\mathbf{g}_{t_N}^{-1}) \end{pmatrix}. \tag{26}$$

4.4 Convex relaxation of NUG to SDP

In the matrix form of a NUG problem, i.e., Equation (23), $\hat{\mathbf{F}}^{(k)}$ is known. As shown by Equation (25), the element of $\hat{\mathbf{F}}^{(k)}$, i.e., $\hat{\mathbf{f}}_{ij}^{(k)}$, is determined by the given symmetry group G and the values of function f_{ij}. In

both cases (spatial rotations and projection directions), the values of the function f_{ij} are pre-determined by $\{\mathbf{q}_i\}_{1\le i\le N}$ and $\{\vec{\mathbf{v}}_{p_i}\}_{1\le i\le N}$ by Equation (19) and Equation (20), respectively. Meanwhile, $\mathbf{X}^{(k)}$ is composed of the variables, i.e., $\{\mathbf{g}_{t_i}\}_{1\le i\le N}$, in the optimization problem. Thus, Equation (23) is an optimization problem over the variables $\left\{\mathbf{X}^{(k)}\right\}_{0\le k\le K-1}$, subject to Equation (26) as constraints.

However, these constraints make the defined domain in this optimization problem in general non-convex. Thus, convex optimization techniques are not suitable, leading to the difficulty of determining a global optimal solution. However, such matrices $\left\{\mathbf{X}^{(k)}\right\}_{1\le k\le K}$ can be relaxed into their convex hull[15, 67, 68] (see Appendix X), converting this NUG problem to an SDP problem. It is a classic problem to determine the solution of an SDP problem. In convex optimization, the global optimal solution is guaranteed. One can obtain the solution using open-source packages[69, 70, 71, 72]. Thus, instead of solving $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$, the solution matrices $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$ are obtained. Such solution matrices $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$ are in the convex hull of the matrices defined in Equation (26).

4.5 Recovery of representatives from the solution of SDP

As the constraint in the SDP problem is looser than that in the NUG problem, the solution matrices of the SDP problem may no longer comply with the constraints of the NUG problem, i.e., Equation (26). In other words, $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$, i.e., the solution matrices of the SDP problem, may violate Equation (26). However, determination of representatives, i.e, determination of $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$, is needed for calculating mean and variance of rotations considering molecular symmetry. Thus, a recovery algorithm is needed for "re-constraining" $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$ to obtain $\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N}$.

We propose this recovery algorithm based on the following. Each block of the matrices $\left\{\mathbf{X}^{(k)}\right\}_{0\le k\le K-1}$, i.e., $\mathbf{X}_{ij}^{(k)}$, is constrained as

$$\mathbf{X}_{ij}^{(k)} = \boldsymbol{\rho}_k\left(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}\right), \tag{27}$$

shown in Equation (26). Meanwhile, such a NUG problem is relaxed into an SDP problem by relaxing the above constraints into their convex hull. Based on the nature of a convex hull, the solution matrices of an SDP problem are the convex combinations of its constraints. In other words, each block $\tilde{\mathbf{X}}_{ij}^{(k)}$ of $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$ is the convex combination of $\boldsymbol{\rho}_k(\mathbf{g}_m)$ as

$$\begin{aligned}
\tilde{\mathbf{X}}_{ij}^{(k)} &= \sum_{m=1}^{|G|}\lambda_{ij}^{\mathbf{g}_m}\boldsymbol{\rho}_k(\mathbf{g}_m)\\
\lambda_{ij}^{\mathbf{g}_m} &\ge 0, \qquad 1\le i,j\le N, \qquad 1\le m\le |G|\\
\sum_{m=1}^{|G|}\lambda_{ij}^{\mathbf{g}_m} &= 1, \qquad 1\le i,j\le N,
\end{aligned} \tag{28}$$

where $\lambda_{ij}^{\mathbf{g}_m}$ is the coefficient of the convex combination. $\lambda_{ij}^{\mathbf{g}_m}$ is determined as

$$\lambda_{ij}^{\mathbf{g}_m} = \frac{1}{|G|}\left(\sum_{k=0}^{K-1} d_k\, tr\left(\boldsymbol{\rho}_k(\mathbf{g}_m)^*\tilde{\mathbf{X}}_{ij}^{(k)}\right)\right) \tag{29}$$

by the Peter-Weyl theorem[66], i.e., Equation (21) and Equation (22).

Comparing the constraints of the block matrices of the NUG problem, i.e., Equation (27), and the solution block matrices of the SDP problem as the convex combinations of $\boldsymbol{\rho}_k(\mathbf{g}_m)$, i.e., Equation (28), the convex combination coefficient $\lambda_{ij}^{\mathbf{g}_m}$ can be considered as the "probability" of $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ being $\mathbf{g}_m$.

For determining $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$ in the original NUG problem, $N-1$ pairs of relations between two of them are required. In other words, once proper $N-1$ relations, i.e., $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$, are determined, $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$ can also be determined. Meanwhile, as previously stated, $\lambda_{ij}^{\mathbf{g}_m}$ indicates the "probability" of $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ being $\mathbf{g}_m$. There are in total $N \times N \times |G|$ guesses of every $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ being every $\mathbf{g}_m$ in G; each has the "probability" $\lambda_{ij}^{\mathbf{g}_m}$. Hence, these guesses are redundant for solving $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$.

In the present work, we propose a leaderboard algorithm for recovering $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$ from such redundant guesses. This algorithm follows the following three principles.

The first principle of this algorithm is sequentially determining each $\tilde{\mathbf{g}}_{t_i}$. As G is a group, the following property holds. For all $1 \le t_i, t_j, t_k \le N$, the relations between $\tilde{\mathbf{g}}_{t_i}$ and $\tilde{\mathbf{g}}_{t_j}$, i.e., $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$, and the relation between $\tilde{\mathbf{g}}_{t_j}$ and $\tilde{\mathbf{g}}_{t_k}$, i.e., $\tilde{\mathbf{g}}_{t_j} \otimes \tilde{\mathbf{g}}_{t_k}^{-1}$, are recovered. In addition, the relation between $\tilde{\mathbf{g}}_{t_i}$ and $\tilde{\mathbf{g}}_{t_k}$, i.e., $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_k}^{-1}$, is also recovered. Therefore, this sequential determination of $\tilde{\mathbf{g}}_{t_i}$ is a classic union-find[73]. In this union-find, if the relation between $\tilde{\mathbf{g}}_{t_i}$ and $\tilde{\mathbf{g}}_{t_j}$ is decided, $\tilde{\mathbf{g}}_{t_i}$ is connected to $\tilde{\mathbf{g}}_{t_j}$. Once every $\tilde{\mathbf{g}}_{t_i}$ is connected to any other $\tilde{\mathbf{g}}_{t_j}$, then $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$ are determined.

The second principle of this algorithm is as follows. During sequentially deciding $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$, the relations between $\tilde{\mathbf{g}}_{t_i}$ and $\tilde{\mathbf{g}}_{t_j}$ are sequentially determined in the descending order of their certainty. For example, $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ and $\tilde{\mathbf{g}}_{t_{i'}} \otimes \tilde{\mathbf{g}}_{t_{j'}}^{-1}$ are two relations. As previously stated, the "probability" of $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ being $\mathbf{g}_m$ is $\lambda_{ij}^{\mathbf{g}_m}$. If the greatest value in $\left\{\lambda_{ij}^{\mathbf{g}_m}\right\}_{1 \le m \le |G|}$ is greater than that in $\left\{\lambda_{i'j'}^{\mathbf{g}_m}\right\}_{1 \le m \le |G|}$, then the relation between $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$ is more certain than that between $\tilde{\mathbf{g}}_{t_{i'}} \otimes \tilde{\mathbf{g}}_{t_{j'}}^{-1}$. Therefore, the relation between $\tilde{\mathbf{g}}_{t_i}$ and $\tilde{\mathbf{g}}_{t_j}$ must be determined before the relation between $\tilde{\mathbf{g}}_{t_{i'}}$ and $\tilde{\mathbf{g}}_{t_{j'}}$. The reason of adopting this principle is as follows. As previously stated, it is redundant for solving $\{\tilde{\mathbf{g}}_{t_i}\}_{1 \le i \le N}$ by $N \times N \times |G|$ guesses of all $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$. By deciding relations with high certainty first, high quality information is used for solving the problem which contributes to the high confidence of the recovery.

The third principle of this algorithm is as follows. For the relation $\tilde{\mathbf{g}}_{t_i} \otimes \tilde{\mathbf{g}}_{t_j}^{-1}$, instead of determining it solely by $\mathbf{g}_m$ of the single greatest value of $\left\{\lambda_{ij}^{\mathbf{g}_m}\right\}_{1 \le m \le |G|}$, several trials are examined. These trials are based on multiple greatest values of $\left\{\lambda_{ij}^{\mathbf{g}_m}\right\}_{1 \le m \le |G|}$. During the sequential determination of each $\tilde{\mathbf{g}}_{t_i}$, multiple sets of trials are tracked; these are selected by partial cost (see Appendix XI). Ultimately, the most qualified set of trials is selected as the final solution. This third principle makes the proposed algorithm in this work tally with the leaderboard algorithm.

5 Performance of the NUG approach

As the NUG approach is an approximation algorithm, its performance, i.e., the capability of it reaching global optimal resolution, requires examination. The capability of solving Equation (14), i.e., selection of representatives for solving the variance and the mean of spatial rotations, is presented below (see Section 5.1). Moreover, currently, there is no accepted approximation algorithm for solving this problem. Therefore, we designed several naive approaches. However, even the best one of them is rather poor in

performance (see Section 5.2). For solving variance and mean of projection directions, the performance is similar, as Equation (16) has a similar construction to Equation (14).

5.1 Capability of reaching global optimal solution

Currently, there is no algorithm for obtaining a global optimal solution, other than the brute force approach. In the brute force approach, for each $\tilde{\mathbf{g}}_i$, there are $|G|$ potential $\mathbf{g}_m$ to be considered. Therefore, there are $|G|^N$ trials to be examined. In other words, the computation cost grows exponentially as the number of spatial rotations grows. Hence, the global optimal is only practically solvable by the brute force approach when N is limited.

Several molecular symmetries are tested. For 3-fold cyclic (C_3), 7-fold dihedral (D_7), and tetrahedral (T) symmetries, $N = 5$ is used for performance tests due to the computational limitations of a modern workstation. Meanwhile, for octahedral (O) and icosahedral (I) symmetries, N is further limited to 3.

Such tests demonstrate that, in almost all cases, the NUG approach reaches a global optimal solution (see Table 1). Moreover, the capability of reaching a global optimal solution increases as the spatial rotations concentrate. Spatial rotations with four angular precision values[45] are tested. For the angular precision values higher than 7.9°, all cases tested reach a global optimal (see Table 1). Moreover, for cases in which global optimal solutions are not reached, the cost gaps between the obtained solution and the global optimal are small. For all cases tested, the maximum gap is approximately 5% (see Table 1).

5.2 Comparison with a naive approach

The NUG approach is compared with a naive approach we design, as currently there is no accepted approximation algorithm available for solving Equation (14). We design several naive approaches, and one naive approach performs better than the others. The superior naive approach is described in Appendix XII.

However, compared to the NUG approach, the naive approach performs rather poorly (see Table 1). For spatial rotations in a uniform distribution, i.e., angular precision equals to 90°, $177 \sim 533$ out of 1000 tested cases do not reach a global optimal, under conditions of various molecular symmetries. Moreover, the cost gaps between the obtained solution and the global optimal solution varies from 23.0% to 70.0% with molecular symmetries. Such incapability to reach a global optimal is alleviated as the spatial rotations concentrate. However, even when angular precision is 1.6°, there are cases which fail to reach global optimal solutions. Especially, the maximum cost gaps between the obtained solutions and the global optimal remain large even though the spatial rotations concentrate.

6 Classification considering molecular symmetry

As the distance is defined and the mean can be computed, the classification of both spatial rotations and projection directions can be achieved by several algorithms, e.g., classic k-means clustering[74]. Due to the similar construction of the distance definition and the mean calculation between the spatial rotations

and the project directions, we can restrict our exposition to projection directions; equivalent results hold for spatial rotations.

As an example, we classified projection directions into five classes with several molecular symmetries (Figure 1). One may notice that, for some classes, the regions seem to be not connected. For example, consider the classes indicated by blue and turquoise in C_3, and the classes indicated by turquoise and red in D_7. However, when molecular symmetry is considered, these regions are in fact connected. Taking the red class with D_7 symmetry as an example, although the right small region seems apart from the left large region, it can be rotated by a symmetry operation and then be attached to the left large region.

In contrast, the traditional method of projection directions classification with molecular symmetry is as follows. A fundamental domain of S^2/G is selected, and the projection directions are rotated by a symmetry operation into this domain. Classification without considering molecular symmetries is then performed on these rotated projection directions. In contrast to the classification method we propose, the true topology induced by the molecular symmetry is neglected. The closer the projection direction is to the edge of the fundamental domain, the more severe the consequences (i.e., error) of neglecting the molecular symmetry.

7 Discussion about potential applications

In the present work, we have enabled the statistics and classification method for both spatial rotations and projection directions considering molecular symmetry, by developing an approximation algorithm for calculating the mean and the variance. The proposed statistics and classification method has the potential to be adopted in a wide range of applications.

A quaternion-assisted angular reconstruction algorithm[30, 31] contributes to solving a series of structures at 10-30Å resolution in 90s[32]. The key step in this algorithm is solving the absolute orientation problem[28, 29], which is related to the determination of the mean of a series of spatial rotations. With this present work, it becomes feasible to solve such problems in the context of considering molecular symmetry. This might lead to rethinking the possibility of adopting a quaternion-assisted angular reconstruction algorithm in state-of-the-art cryoEM reconstruction. Furthermore, the statistics of spatial rotations may contribute to the adoption of advanced statistical method, such as particle filter[75, 76, 8] or Kalman filter algorithm[77, 78, 79], in cryoEM image processing.

For the statistics of spatial rotations, they may be adopted to resolve continuous structural changes by manifold embedding[80, 81]. This is because in such methods, images are grouped by projection directions for the purpose of mapping the continuum of the states of a molecule in each projection direction. Moreover, they can be used for analyzing molecule heterogeneity and misalignment. As an example, 124,810 single particle images of a Tc complex are classified by projection directions when C_5 symmetry is considered (Figure 2). The orientation heterogeneity of the tail exists, and approximately 12.5% images are misaligned. However, a 3.7Å resolution structure can be obtained from this dataset.

In summary, the statistics tools developed in this work form a system for analyzing spatial rotations and projection directions in cryoEM when molecular symmetry is considered. They provide more freedom

and efficiency when processing and understanding images in cryoEM. We have realized the proposed methods in the open-source python package pySymStat, and it can be easily adopted or integrated in other algorithms or software.

Acknowledgements

This work was supported by fund from Advanced Innovation Center for Structural Biology (to M.H.), Yau Mathematical Sciences Center (to H.L.), and grant TH-533310008 of Tsinghua University (to H.L.). The authors are grateful to Dr. Xueming Li for providing the Tc complex dataset.

Contributions

M.H., Q.Z. and H.L. initialized the project and designed the methods. Q.Z. and H.L. clarified mathematical issues and derived formulas. M.H. and Q.Z. wrote pySymStat and validated its performance. M.H. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

[1] Sjors H. W. Scheres. Relion: Implementation of a bayesian approach to cryo-em structure determination. *Journal of Structural Biology*, 180(3):519–530, 2012.

[2] Ali Punjani, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker. cryosparc: algorithms for rapid unsupervised cryo-em structure determination.(cryo-electron microsopy single-particle ab initio reconstruction and classification)(report). *Nature Methods*, 14(3):290, 2017.

[3] Nikolaus Grigorieff. Frealign: High-resolution refinement of single particle structures. 157(1):0–125.

[4] Timothy Grant, Alexis Rohou, and Nikolaus Grigorieff. cistem: User-friendly software for single-particle image processing. 7(a2):C1368–C1368, 2018.

[5] Guang Tang, Liwei Peng, Philip R. Baldwin, Deepinder S. Mann, Wen Jiang, Ian Rees, and Steven J. Ludtke. Eman2: An extensible image processing suite for electron microscopy. 157(1):0–46.

[6] Tanvir R Shaikh, Haixiao Gao, William T Baxter, Francisco J Asturias, Nicolas Boisset, Ardean Leith, and Joachim Frank. Spider image processing for single-particle reconstruction of biological macromolecules from electron micrographs. 3(12):1941–1974.

[7] Joachim Frank, Brian Shimkin, and Helen Dowse. Spider—a modular software system for electron image processing. *Ultramicroscopy*, 6(1):343–357.

[8] Mingxu Hu, Hongkun Yu, Kai Gu, Zhao Wang, Huabin Ruan, Kunpeng Wang, Siyuan Ren, Bing Li, Lin Gan, Shizhen Xu, Guangwen Yang, Yuan Shen, and Xueming Li. A particle-filter framework for robust cryo-em 3d reconstruction. *Nature Methods*, 15(12):1083–1089, 2018.

[9] David S Goodsell and Arthur J Olson. Structural symmetry and protein function. *Annual review of biophysics and biomolecular structure*, 29(1):105–153, 2000.

[10] Lucy R Forrest. Structural symmetry in membrane proteins. *Annual review of biophysics*, 44:311–337, 2015.

[11] William Rowan Hamilton. On a new species of imaginary quantities connected with a theory of quaternions. In *Proceedings of the Royal Irish Academy*, volume 2, pages 4–1, 1844.

[12] John Vince. *Quaternions for Computer Graphics*. London: Springer London, London, 2011.

[13] John H. Conway and Derek A. Smith. *On Quaternions and Octonions*. 1 edition, 2003.

[14] Roy R. Lederman and Amit Singer. A representation theory perspective on simultaneous alignment and classification. *Applied and Computational Harmonic Analysis*, 2016.

[15] Afonso S Bandeira, Yutong Chen, Roy R Lederman, and Amit Singer. Non-unique games over compact groups and orientation estimation in cryo-em. *Inverse Problems*, 36(6):064002, 2020.

[16] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. 2004.

[17] Serdar Kucuk and Zafer Bingul. *Robot kinematics: Forward and inverse kinematics*. INTECH Open Access Publisher, 2006.

[18] Eduardo Bayro-Corrochano. *Geometric computing: for wavelet transforms, robot vision, learning, control and action*. Springer Publishing Company, Incorporated, 2010.

[19] JR Dooley and J Michael McCarthy. Spatial rigid body dynamics using dual quaternion components. In *Proceedings. 1991 IEEE International Conference on Robotics and Automation*, pages 90–95. IEEE, 1991.

[20] Erol Ozgur and Youcef Mezouar. Kinematic modeling and control of a robot arm using unit dual quaternions. *Robotics and Autonomous Systems*, 77:66–73, 2016.

[21] Bong Wie and Peter M Barba. Quaternion feedback for spacecraft large angle maneuvers. *Journal of Guidance, Control, and Dynamics*, 8(3):360–365, 1985.

[22] Lawrence E Barker Jr, Roland L Bowles, and Louise H Williams. Development and application of a local linearization algorithm for the integration of quaternion rate equations in real-time flight simulation problems. 1973.

[23] An-Min Zou and Krishna Dev Kumar. Quaternion-based distributed output feedback attitude coordination control for spacecraft formation flying. *Journal of Guidance, Control, and Dynamics*, 36(2):548–556, 2013.

[24] Nuno Filipe and Panagiotis Tsiotras. Adaptive position and attitude-tracking controller for satellite proximity operations using dual quaternions. *Journal of Guidance, Control, and Dynamics*, 38(4):566–577, 2015.

[25] John Vince. *Quaternions for computer graphics*. Springer Science & Business Media, 2011.

[26] Ramakrishnan Mukundan. Quaternions: From classical mechanics to computer graphics, and beyond. In *Proceedings of the 7th Asian Technology conference in Mathematics*, pages 97–105, 2002.

[27] David Eberly. *3D game engine design: a practical approach to real-time computer graphics*. CRC Press, 2006.

[28] Berthold Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4:629–642, 1987.

[29] George Harauz. Representation of rotations by unit quaternions. *Ultramicroscopy*, 33(3):209–213.

[30] Neil A Farrow and E Peter Ottensmeyer. A posteriori determination of relative projection directions of arbitrarily oriented macromolecules. *JOSA A*, 9(10):1749–1760, 1992.

[31] Neil A Farrow and F Peter Ottensmeyer. Automatic 3d alignment of projection images of randomly oriented objects. *Ultramicroscopy*, 52(2):141–156, 1993.

[32] Michael Schatz, EV Orlova, P Dube, H Stark, F Zemlin, and M Van Heel. Angular reconstitution in three-dimensional electron microscopy: Practical and technical aspects. *Scanning Microscopy*, 11:179–193, 1997.

[33] Joachim Frank, Michael Radermacher, Pawel Penczek, Jun Zhu, Yanhong Li, Mahieddine Ladjadj, and Ardean Leith. Spider and web: processing and visualization of images in 3d electron microscopy and related fields. *Journal of structural biology*, 116(1):190–199, 1996.

[34] PR Baldwin and Pawel A Penczek. The transform class in sparx and eman2. *Journal of structural biology*, 157(1):250–261, 2007.

[35] John Kuo. *Electron microscopy: methods and protocols*, volume 369. Springer Science & Business Media, 2007.

[36] Ingolf Sommer and Richard Brimacombe. Methods for refining interactively established models of ribosomal rna towards a physico-chemically plausible structure. *Journal of Computational Chemistry*, 22(4):407–417, 2001.

[37] JR Plaisier, L Jiang, and JP Abrahams. Cyclops: New modular software suite for cryo-em. *Journal of structural biology*, 157(1):19–27, 2007.

[38] Xiangan Liu, Wen Jiang, Joanita Jakana, and Wah Chiu. Averaging tens to hundreds of icosahedral particle images to resolve protein secondary structure elements using a multi-path simulated annealing optimization algorithm. *Journal of structural biology*, 160(1):11–27, 2007.

[39] Maya Topf, Keren Lasker, Ben Webb, Haim Wolfson, Wah Chiu, and Andrej Sali. Protein structure fitting and refinement guided by cryo-em density. *Structure*, 16(2):295–307, 2008.

[40] Sebastian Wasilewski and Peter B Rosenthal. Web server for tilt-pair validation of single particle maps from electron cryomicroscopy. *Journal of structural biology*, 186(1):122–131, 2014.

[41] Arun Prasad Pandurangan, Daven Vasishtan, Frank Alber, and Maya Topf. r-tempy: simultaneous fitting of components in 3d-em maps of their assembly using a genetic algorithm. *Structure*, 23(12):2365–2376, 2015.

[42] Thomas Hrabe. Localize.pytom: a modern webserver for cryo-electron tomography. *Nucleic acids research*, 43(W1):W231–W236, 2015.

[43] Gydo Cp van Zundert and Alexandre Mjj Bonvin. Fast and sensitive rigid-body fitting into cryo-em density maps with powerfit. *AIMS Biophysics*, 2(2):73–87, 2015.

[44] John C Hart, George K Francis, and Louis H Kauffman. Visualizing quaternion rotation. *ACM Transactions on Graphics (TOG)*, 13(3):256–276, 1994.

[45] Mingxu Hu, Qi Zhang, Jing Yang, and Xueming Li. Unit quaternion description of spatial rotation in 3d electron cryo-microscopy. *bioRxiv*, 2019.

[46] Nathan Jacobson. *Lectures in abstract algebra: Basic concepts*. Springer, 1958.

[47] David Steven Dummit and Richard M Foote. *Abstract algebra*, volume 3. Wiley Hoboken, 2004.

[48] Anuj Srivastava and Eric P. Klassen. *Functional and Shape Data Analysis*. New York, NY: Springer New York, New York, NY, 2016.

[49] K. V. Mardia. *Directional statistics*. Chichester ; New York : Wiley, Chichester ; New York, [2nd ed.].. edition, 2000.

[50] Charles W Curtis and Irving Reiner. *Representation theory of finite groups and associative algebras*, volume 356. American Mathematical Soc., 1966.

[51] Jean-Pierre Serre. *Linear representations of finite groups*, volume 42. Springer, 1977.

[52] Roe Goodman and Nolan R Wallach. *Representations and invariants of the classical groups*. Cambridge University Press, 2000.

[53] Pavel I Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to representation theory*, volume 59. American Mathematical Soc., 2011.

[54] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.

[55] Hans Kurzweil and Bernd Stellmacher. The theory of finite groups, 2004.

[56] I. Martin Isaacs. *Finite group theory*, volume 92. American Mathematical Soc., 2008.

[57] Michael J Collins. *Representations and characters of finite groups*, volume 22. Cambridge University Press, 1990.

[58] Roger William Carter. Finite groups of lie type: Conjugacy classes and complex characters. *Pure Appl. Math., Wiley*, 44, 1985.

[59] Ronald Mirman. *Point groups, space groups, crystals, molecules*. World Scientific, 1999.

[60] Christopher Bradley and Arthur Cracknell. *The mathematical theory of symmetry in solids: representation theory for point groups and space groups*. Oxford University Press, 2009.

[61] Walter Feit. *The representation theory of finite groups*. Elsevier, 1982.

[62] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *Siam Review*, 38(1):49–95.

[63] H. Wolkowicz. *Semidefinite Programming*. Springer Berlin Heidelberg, 2007.

[64] Andreas D Baxevanis, Gary D Bader, and David S Wishart. *Bioinformatics*. John Wiley & Sons, 2020.

[65] David W Mount and David W Mount. *Bioinformatics: sequence and genome analysis*, volume 1. Cold spring harbor laboratory press New York:, 2001.

[66] H. Weyl and P. Peter. Die vollstandigkeit der primitiven darstellungen einer geschlossenen kontinuierlichen gruppe. 97(1):737–755, 1927.

[67] Afonso S Bandeira, Moses Charikar, Amit Singer, and Andy Zhu. Multireference alignment using semidefinite programming. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 459–470, 2014.

[68] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 205–214, 2006.

[69] Michael Grant, Stephen Boyd, and Yinyu Ye. Cvx: Matlab software for disciplined convex programming, 2009.

[70] Michael Grant and Stephen Boyd. Cvx: Matlab software for disciplined convex programming, version 2.1, 2014.

[71] Guillaume Sagnol et al. Picos, a python interface to conic optimization solvers. In *Proceedings of the in 21st International Symposium on Mathematical Programming*, 2012.

[72] Martin S Andersen, Joachim Dahl, and Lieven Vandenberghe. Cvxopt: A python package for convex optimization, version 1.1. 6. *Available at cvxopt. org*, 54, 2013.

[73] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[74] Robert Sedgewick and Kevin Wayne. *Algorithms*. Addison-wesley professional, 2011.

[75] M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002.

[76] Fredrik Gustafsson, Fredrik Gunnarsson, Niclas Bergman, Urban Forssell, Jonas Jansson, Rickard Karlsson, and P-J Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on signal processing*, 50(2):425–437, 2002.

[77] Daniel Choukroun, Itzhack Y Bar-Itzhack, and Yaakov Oshman. Novel quaternion kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, 42(1):174–190, 2006.

[78] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. 1995.

[79] Andrew C Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

[80] Joachim Frank and Abbas Ourmazd. Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-em. *Methods*, 100:61–67, 2016.

[81] Suvrajit Maji, Hstau Liao, Ali Dashti, Ghoncheh Mashayekhi, Abbas Ourmazd, and Joachim Frank. Propagation of conformational coordinates across angular space in mapping the continuum of states from cryo-em data by manifold embedding. *Journal of Chemical Information and Modeling*, 2020.

Table 1: Performance of the NUG approach, compared with a naive approach

Method	Symmetry	$\epsilon = 90°$*		$\epsilon \approx 36°$		$\epsilon \approx 7.9°$		$\epsilon \approx 1.6°$	
		nW/nT**	δ***	nW/nT	δ	nW/nT	δ	nW/nT	δ
NUG approach	C_3	1/1000	1.78%	1/1000	1.15%	0/1000	0	0/1000	0
	D_7	3/1000	5.07%	1/1000	0.005%	0/1000	0	0/1000	0
	T	0/1000	0	0/1000	0	0/1000	0	0/1000	0
	O	1/1000	0.002%	0/1000	0	0/1000	0	0/1000	0
	I	0/1000	0	1/1000	1.74%	0/1000	0	0/1000	0
naive approach	C_3	533/1000	23.0%	260/1000	33.7%	27/1000	8.6%	1/1000	18.6%
	D_7	511/1000	33.7%	441/1000	39.2%	90/1000	44.3%	5/1000	10.6%
	T	444/1000	44.8%	367/1000	51.2%	57/1000	43.9%	2/1000	3.27%
	O	485/1000	59.0%	429/1000	55.2%	78/1000	49.6%	6/1000	36.6%
	I	177/1000	70.0%	177/1000	75.6%	47/1000	60.5%	3/1000	30.1%

* ϵ stands for the angular precision[45] of the testing spatial rotations.
** nW stands for the number of cases that fail to reach a global optimum. nT stands for the number of cases tested.
*** δ stands for the maximum relative cost gap between the obtained solutions and the global optimal.

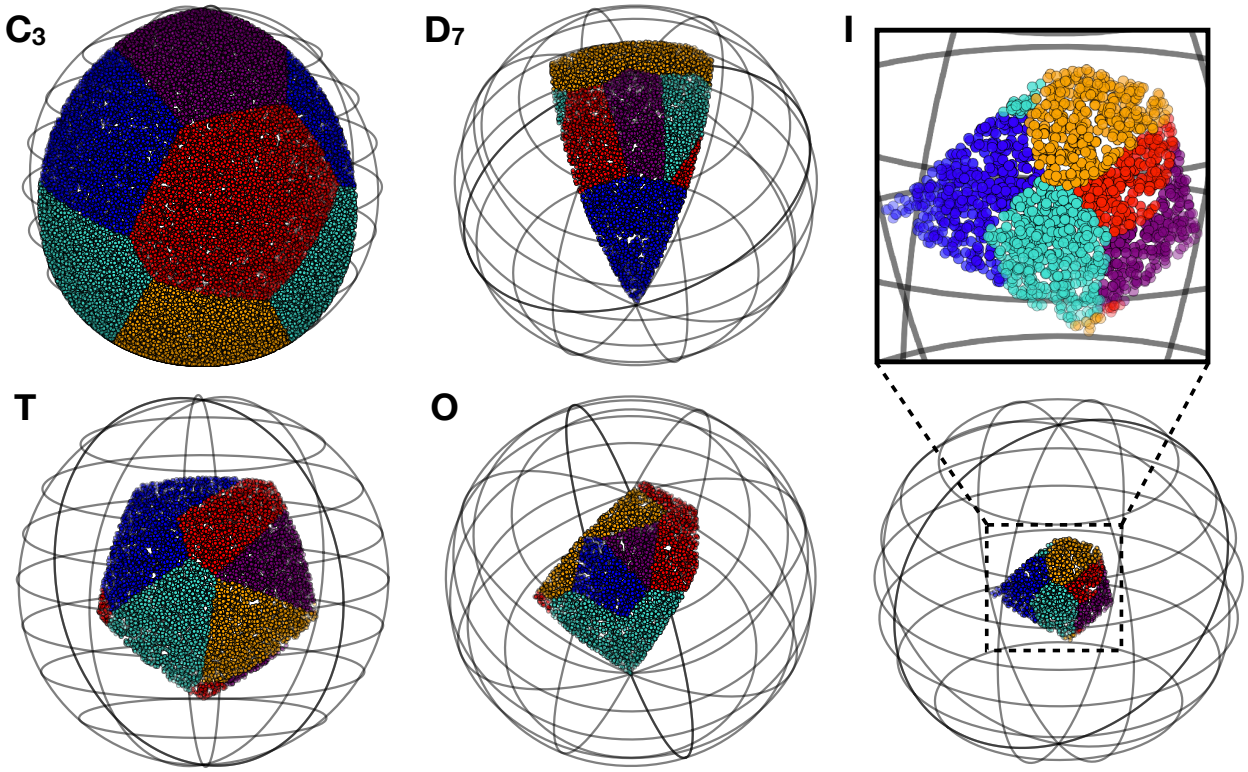

Figure 1: Classification of projection directions considering molecular symmetries. With several molecular symmetries, the projection directions are classified into five classes; these are distinguished by color. For presentation purposes, in each set of equivalent projection directions $R_{G^{-1}}(\vec{\mathbf{v}}_p)$, an element within a given fundamental domain is selected for display. The spheres are rotated in order to place projection directions towards the reader.

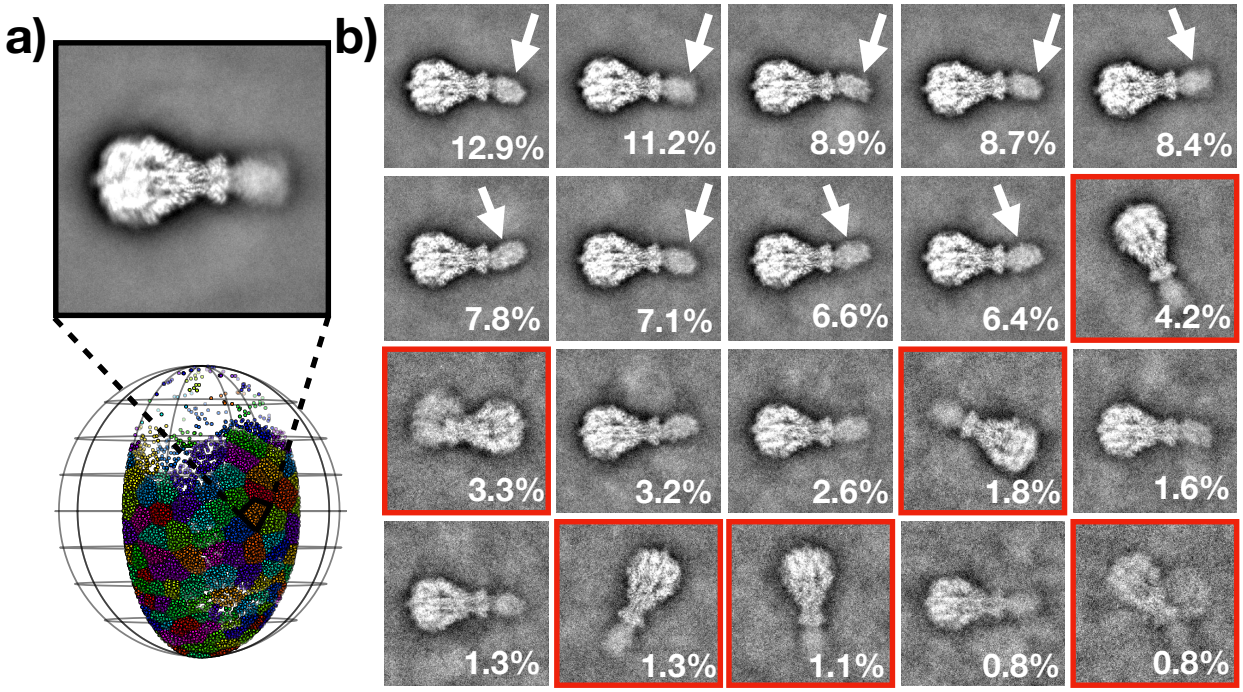

Figure 2: Projection direction classification of a Tc complex. a, 124,810 single particle images of a Tc complex are classified by their projection directions into 100 classes, when C_5 molecular symmetry is taken into consideration. Classes are distinguished by colors. The images from a selected class (orange) are averaged and then displayed. **b**, The images from the selected class are further classified into 20 classes. The orientation heterogeneity of the tail is indicated by an arrow. The misaligned images are boxed by red lines.

Appendix I The sufficient condition of two spatial rotations being equivalent considering molecular symmetry

In this section, we prove the following statement. For a molecule with G symmetry, the spatial rotation $\mathbf{q}$ is equivalent to $\mathbf{q} \otimes \mathbf{g}_m$ for any $\mathbf{g}_m$ in the molecular symmetry group G. This statement can be paraphrased as follows. For two spatial rotations $\mathbf{q}_1$ and $\mathbf{q}_2$, the sufficient condition of $\mathbf{q}_1$ and $\mathbf{q}_2$ being equivalent considering molecular symmetry G is that there exists an element $\mathbf{g}_m$ in G such that $\mathbf{q}_1 = \mathbf{q}_2 \otimes \mathbf{g}_m$.

A coordinate in the map is denoted by $\vec{\mathbf{v}}$. The value at coordinate $\vec{\mathbf{v}}$ in this map is denoted by $h(\vec{\mathbf{v}})$. As this map has molecular symmetry G,

$$h(R_{\mathbf{g}_m}(\vec{\mathbf{v}})) = h(\vec{\mathbf{v}}), \tag{30}$$

for all $\mathbf{g}_m$ in G.

The value at the coordinate $\vec{\mathbf{v}}$ in the rotated map with a spatial rotation $\mathbf{q}$ is denoted by $h_{\mathbf{q}}$. By the definition of the spatial rotation,

$$h_{\mathbf{q}}(\vec{\mathbf{v}}) = h(R_{\mathbf{q}^{-1}}(\vec{\mathbf{v}})). \tag{31}$$

The sufficiency is proved as follows. As there exists an element $\mathbf{g}_m$ in G such that $\mathbf{q}_1 = \mathbf{q}_2 \otimes \mathbf{g}_m$, it can be concluded that

$$\begin{aligned} h_{\mathbf{q}_1}(\vec{\mathbf{v}}) =& h(R_{\mathbf{q}_1^{-1}}(\vec{\mathbf{v}})) \\ =& h(R_{\mathbf{g}_m^{-1} \otimes \mathbf{q}_2^{-1}}(\vec{\mathbf{v}})) \\ =& h(R_{\mathbf{g}_m^{-1}}(R_{\mathbf{q}_2^{-1}}(\vec{\mathbf{v}}))) \end{aligned} \tag{32}$$

By Equation (30), it can be further concluded that

$$\begin{aligned} h_{\mathbf{q}_1}(\vec{\mathbf{v}}) =& h(R_{\mathbf{q}_2^{-1}}(\vec{\mathbf{v}})) \\ =& h_{\mathbf{q}_2}(\vec{\mathbf{v}}). \end{aligned} \tag{33}$$

As stated in Equation (33), the $\mathbf{q}_1$ rotated map is identical to the $\mathbf{q}_2$ rotated map. Hence, $\mathbf{q}_1$ is equivalent to $\mathbf{q}_2$. Therefore, we have proven the sufficiency of the condition.

Appendix II The sufficient condition of two projection directions being equivalent considering molecular symmetry

In this section, we continue to use the notations and conclusions presented in Appendix I. Next, we prove the following statement. For a molecule with G symmetry, the projection direction $\vec{\mathbf{v}}_p$ is equivalent to $R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)$ for any $\mathbf{g}_m$ in a molecular symmetry group G.

By the concept of the projection image set defined by Equation (4), the equivalence between two projec-

tion directions can be defined by the equivalence between the two sets of projected images. As the projection operator $\mathcal{P}^{\mathbf{q}}$ is defined by the operation of projecting along Z axis after after rotating the map V by spatial rotation $\mathbf{q}$ and the map V has molecular symmetry G, it stands that

$$\mathcal{P}^{\mathbf{q}}(V) = \mathcal{P}^{\mathbf{q}\otimes\mathbf{g}_m}(V), \tag{34}$$

for any spatial rotation $\mathbf{q}$ and any element $\mathbf{g}_m \in G$. Therefore, by definition of the projection image set, i.e., Equation (4), it can be concluded that

$$\begin{aligned}\mathbf{P}_{\vec{\mathbf{v}}_p} &= \{\mathcal{P}^{\mathbf{q}}(V)\}_{\{\mathbf{q}|R_{\mathbf{q}}(\vec{\mathbf{v}}_p)=\vec{\mathbf{v}}_n\}} \\ &= \left\{\mathcal{P}^{\mathbf{q}\otimes\mathbf{g}_m}(V)\right\}_{\{\mathbf{q}|R_{\mathbf{q}}(\vec{\mathbf{v}}_p)=\vec{\mathbf{v}}_n\}} \\ &= \{\mathcal{P}^{\mathbf{q}}(V)\}_{\left\{\mathbf{q}\middle|R_{\mathbf{q}\otimes\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)=\vec{\mathbf{v}}_n\right\}} \\ &= \{\mathcal{P}^{\mathbf{q}}(V)\}_{\left\{\mathbf{q}\middle|R_{\mathbf{q}}\left(R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)\right)=\vec{\mathbf{v}}_n\right\}} \\ &= \mathbf{P}_{R_{\mathbf{g}_m^{-1}}(\vec{\mathbf{v}}_p)}\end{aligned} \tag{35}$$

Therefore, we have proven the sufficiency of the condition.

Appendix III Proof of well-definedness of the distance between two spatial rotations considering molecular symmetry

In this section, Equation (7) is proved to be a well-defined distance definition. The three requirements of distance definition in mathematics, i.e., the symmetry property, the positive definiteness property, and the triangle inequality property, are proven to be satisfied for Equation (7).

The symmetry property is such that

$$d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) = d_{SO(3)/G}(\mathbf{q}_2 \otimes G, \mathbf{q}_1 \otimes G) \tag{36}$$

holds for all $R_{\mathbf{q}_1}, R_{\mathbf{q}_2} \in SO(3)$. It is proven as follows.

$$\begin{aligned}d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) &= \min_{\mathbf{g}_{t_1},\mathbf{g}_{t_2}\in G} d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}) \\ &= \min_{\mathbf{g}_{t_1},\mathbf{g}_{t_2}\in G} d_{SO(3)}(\mathbf{q}_2 \otimes \mathbf{g}_{t_2}, \mathbf{q}_1 \otimes \mathbf{g}_{t_1}) \\ &= d_{SO(3)/G}(\mathbf{q}_2 \otimes G, \mathbf{q}_1 \otimes G)\end{aligned} \tag{37}$$

as $d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}) = d_{SO(3)}(\mathbf{q}_2 \otimes \mathbf{g}_{t_2}, \mathbf{q}_1 \otimes \mathbf{g}_{t_1})$ holds, because $d_{SO(3)}$ is also a distance, therefore

it has the symmetry property.

The positive definiteness property contains two parts. The first part is that

$$d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) \geq 0 \tag{38}$$

holds for all $R_{\mathbf{q}_1}, R_{\mathbf{q}_2} \in SO(3)$, which can be derived by $d_{SO(3)}(\mathbf{q}) \geq 0$ for all $R_{\mathbf{q}} \in SO(3)$. The second part is that

$$d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) = 0 \tag{39}$$

if and only if $\mathbf{q}_1 \otimes G = \mathbf{q}_2 \otimes G$. This statement is proven as follows. When $\mathbf{q}_1 \otimes G = \mathbf{q}_2 \otimes G$, there exist $\mathbf{g}_{t_1}$ and $\mathbf{g}_{t_2}$ such that $\mathbf{q}_1 \otimes \mathbf{g}_{t_1} = \mathbf{q}_2 \otimes \mathbf{g}_{t_2}$. Thus, $d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}) = 0$, as $d_{SO(3)}$ also has the positive definiteness property. Therefore, $d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) = 0$. Conversely, when $d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) = 0$, there exist $\mathbf{g}_{t_1}$ and $\mathbf{g}_{t_2}$ such that $\mathbf{q}_1 \otimes \mathbf{g}_{t_1} = \mathbf{q}_2 \otimes \mathbf{g}_{t_2}$. If $\mathbf{q}_1 \otimes G \neq \mathbf{q}_2 \otimes G$, without losing generality, there exists $\mathbf{g}'_{t_1}$ such that $\mathbf{q}_1 \otimes \mathbf{g}'_{t_1} \notin \mathbf{q}_2 \otimes G$. Therefore, $\mathbf{q}_2 = \mathbf{q}_1 \otimes \mathbf{g}_{t_1} \otimes \mathbf{g}_{t_2}^{-1} = \left(\mathbf{q}_1 \otimes \mathbf{g}'_{t_1}\right) \otimes \left(\mathbf{g}'^{-1}_{t_1} \otimes \mathbf{g}_{t_1} \otimes \mathbf{g}_{t_2}^{-1}\right)$ implies that $\mathbf{g}'^{-1}_{t_1} \otimes \mathbf{g}_{t_1} \otimes \mathbf{g}_{t_2}^{-1} \notin G$, which conflicts with the property of the group G. Thus, $\mathbf{q}_1 \otimes G = \mathbf{q}_2 \otimes G$.

The triangle inequality property is that

$$d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) \leq d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_3 \otimes G) + d_{SO(3)/G}(\mathbf{q}_3 \otimes G, \mathbf{q}_2 \otimes G) \tag{40}$$

holds for all $R_{\mathbf{q}_1}, R_{\mathbf{q}_2}, R_{\mathbf{q}_3}$ in $SO(3)$. It is proven as follows.

$$\begin{aligned} & d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) \\ = & \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2} \in G} d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}) \\ \leq & \min_{\mathbf{g}_{t_1}, \mathbf{g}_{t_2}, \mathbf{g}_{t_3} \in G} \left(d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_3 \otimes \mathbf{g}_{t_3}) + d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}_{t_3}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2})\right) \end{aligned} \tag{41}$$

as $d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}) \leq d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_3 \otimes \mathbf{g}_{t_3}) + d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}_{t_3}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2})$ holds for any $\mathbf{g}_{t_3} \in G$, because $d_{SO(3)}$ also has the triangle inequality property.

Moreover, distance in $SO(3)$, i.e., $d_{SO(3)}$, is invariant under left and right multiplication[1]. In other words, $d_{SO(3)}$ has the following properties,

$$d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_2) = d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{q}_3, \mathbf{q}_2 \otimes \mathbf{q}_3) \tag{42}$$

and

$$d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_2) = d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{q}_1, \mathbf{q}_3 \otimes \mathbf{q}_2) \tag{43}$$

for any $R_{\mathbf{q}_1}$, $R_{\mathbf{q}_2}$ and $R_{\mathbf{q}_3}$ in $SO(3)$. Therefore, Equation (41) can be further converted into

$$\begin{aligned}
&d_{SO(3)/G}(\mathbf{q}_1 \otimes G, \mathbf{q}_2 \otimes G) \\
\leq &\min_{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\mathbf{g}_{t_3}\in G} \left(d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_3 \otimes \mathbf{g}_{t_3} \otimes \mathbf{g}_{t_1}^{-1}) + d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}_{t_3} \otimes \mathbf{g}_{t_2}^{-1}, \mathbf{q}_2)\right) \\
= &\min_{\mathbf{g},\mathbf{g}'} \left(d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_3 \otimes \mathbf{g}) + d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}', \mathbf{q}_2)\right) \\
= &\min_{\mathbf{g}} d_{SO(3)}(\mathbf{q}_1, \mathbf{q}_3 \otimes \mathbf{g}) + \min_{\mathbf{g}} d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}, \mathbf{q}_2) \\
= &\min_{\mathbf{g}_{t_1},\mathbf{g}_{t_3}\in G} d_{SO(3)}(\mathbf{q}_1 \otimes \mathbf{g}_{t_1}, \mathbf{q}_3 \otimes \mathbf{g}_{t_3}) + \min_{\mathbf{g}_{t_3},\mathbf{g}_{t_2}\in G} d_{SO(3)}(\mathbf{q}_3 \otimes \mathbf{g}_{t_3}, \mathbf{q}_2 \otimes \mathbf{g}_{t_2}).
\end{aligned} \tag{44}$$

Therefore, we have proven the suitability of Equation (7) as a distance definition.

Appendix IV Proof of well-definedness of distance definition between two projection directions considering molecular symmetry

In this section, Equation (8) is proven to be a well-defined distance definition. Three requirements of distance definition in mathematics, i.e., the symmetry property, the positive definiteness property and the triangle inequality property, are proven to satisfy Equation (8).

As the distance definition between two projection directions considering molecular symmetry, i.e., Equation (8), follows the same structure of the definition between two spatial rotations considering molecular symmetry, i.e., Equation (7), the symmetry property, the first part of the positive definiteness, and the triangle inequality property can be proven with the same procedure in Appendix III.

The second part of the positive definiteness is that

$$d_{S^2/G}\left(R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right), R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)\right) = 0 \tag{45}$$

if and only if $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right) = R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$. This statement is proven as follows. When $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right) = R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$, there exist $\mathbf{g}_{t_1}$ and $\mathbf{g}_{t_2}$ such that $R_{\mathbf{g}_{t_1}^{-1}}(\vec{\mathbf{v}}_{p_1}) = R_{\mathbf{g}_{t_2}^{-1}}(\vec{\mathbf{v}}_{p_2})$. As d_{S^2} also has positive definiteness property, $d_{S^2}\left(R_{\mathbf{g}_{t_1}^{-1}}(\vec{\mathbf{v}}_{p_1}), R_{\mathbf{g}_{t_2}^{-1}}(\vec{\mathbf{v}}_{p_2})\right) = 0$. Therefore, $d_{S^2/G}\left(R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right), R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)\right) = 0$. Conversely, when $d_{S^2/G}\left(R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right), R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)\right) = 0$, there exist $\mathbf{g}_{t_1}$ and $\mathbf{g}_{t_2}$ such that $R_{\mathbf{g}_{t_1}^{-1}}(\vec{\mathbf{v}}_{p_1}) = R_{\mathbf{g}_{t_2}^{-1}}(\vec{\mathbf{v}}_{p_2})$. If $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right) \neq R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$, without losing generality, there exists $\mathbf{g}'_{t_1}$ such that $R_{\mathbf{g}'^{-1}_{t_1}}\left(\vec{\mathbf{v}}_{p_1}\right) \notin R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$. Therefore, $R_{\mathbf{g}'^{-1}_{t_1}}\left(\vec{\mathbf{v}}_{p_1}\right) = R_{\mathbf{g}'^{-1}_{t_1}\otimes\mathbf{g}_{t_1}}\left(R_{\mathbf{g}_{t_1}^{-1}}(\vec{\mathbf{v}}_{p_1})\right) = R_{\mathbf{g}'^{-1}_{t_1}\otimes\mathbf{g}_{t_1}}\left(R_{\mathbf{g}_{t_2}^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)\right) = R_{\mathbf{g}'^{-1}_{t_1}\otimes\mathbf{g}_{t_1}\otimes\mathbf{g}_{t_2}^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right) \notin R_G\left(\vec{\mathbf{v}}_{p_2}\right)$, which conflicts with the property of the group G. Thus, $R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_1}\right) = R_{G^{-1}}\left(\vec{\mathbf{v}}_{p_2}\right)$.

Therefore, we have proven the suitability of Equation (8) as a distance definition.

Appendix V Approximation of the mean of spatial rotations by representatives selected in variance determination

In this section, we state that it is rational to approximate the mean of spatial rotations, i.e, Equation (11), by Equation (13).

By observing the definition of mean and the definition of variance of spatial rotations, i.e., Equation (11) and (9), respectively, one can find that the spirit of the selection of $\mathbf{g}_{t_i}$ for the mean and the variance are the same. In other words, $\mathbf{g}_{t_i}$ is selected for each $\mathbf{q}_i$ to make $\mathbf{q}_i \otimes \mathbf{g}_{t_i}$ as close as possible overall. Therefore, it is natural to replace the $\mathbf{g}_{t_i}$ in the definition of the mean by the $\mathbf{g}_{t_i}$ in the definition of the variance. Thus, Equation (13) is a rational approximation of Equation (11).

Appendix VI Approximation of mean of projection directions by representatives selected in variance determination

In this section, we state that it is rational to approximate the mean of projection directions, i.e., Equation (12), by Equation (15).

As the relation between the mean and the variance of spatial rotations is identical to those of the projection directions, the rationality of approximating Equation (12) by Equation (15) can be explained by using the same procedure as in Appendix V.

Appendix VII Derivation of representative determination of spatial rotations into a NUG form

In this section, we prove that Equation (14) can be derived into a NUG form, i.e., Equation (18).

Distance in $SO(3)$, i.e., $d_{SO(3)}$, is invariant under left and right multiplications[1], i.e., Equation (42) and Equation (43). Thus, Equation (14) can be derived as

$$\begin{aligned}
\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N} &= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\leq i,j\leq N} d^2_{SO(3)}(\mathbf{q}_i \otimes \mathbf{g}_{t_i}, \mathbf{q}_j \otimes \mathbf{g}_{t_j}) \right) \\
&= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\leq i,j\leq N} d^2_{SO(3)}(\mathbf{q}_i \otimes \mathbf{g}_{t_i} \otimes \mathbf{g}_{t_j}^{-1}, \mathbf{q}_j) \right) \\
&= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\leq i,j\leq N} d^2_{SO(3)}(\mathbf{g}_{t_i} \otimes \mathbf{g}_{t_j}^{-1}, \mathbf{q}_i^{-1} \otimes \mathbf{q}_j) \right).
\end{aligned} \tag{46}$$

By assigning $f_{ij}(\mathbf{g}_{t_i} \otimes \mathbf{g}_{t_j}^{-1})$ as Equation (19), we have proven that Equation (14) can be derived into a NUG form, i.e., Equation (18).

Appendix VIII Derivation of representative determination of projection directions into a NUG form

In this section, we prove that Equation (16) can be derived into a NUG form, i.e., Equation (18).

The distance in S^2, i.e., d_{S^2} is invariant under spatial rotation. In other words, d_{S^2} has the following property,

$$d_{S^2}\left(\vec{\mathbf{v}}_{p_1}, \vec{\mathbf{v}}_{p_2}\right) = d_{S^2}\left(R_{\mathbf{q}}\left(\vec{\mathbf{v}}_{p_1}\right), R_{\mathbf{q}}\left(\vec{\mathbf{v}}_{p_2}\right)\right) \tag{47}$$

for any projection directions $\vec{\mathbf{v}}_{p_1}$ and $\vec{\mathbf{v}}_{p_2}$, and any spatial rotation $\mathbf{q}$.

Thus, Equation (16) can be derived into

$$\begin{aligned}
\{\tilde{\mathbf{g}}_{t_i}\}_{1\le i\le N} &= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\le i,j\le N} d^2_{S^2}\left(R_{\mathbf{g}_{t_i}^{-1}}\left(\vec{\mathbf{v}}_{p_i}\right), R_{\mathbf{g}_{t_j}^{-1}}\left(\vec{\mathbf{v}}_{p_j}\right)\right)\right) \\
&= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\le i,j\le N} d^2_{S^2}\left(\vec{\mathbf{v}}_{p_i}, R_{\mathbf{g}_{t_i}}\left(R_{\mathbf{g}_{t_j}^{-1}}\left(\vec{\mathbf{v}}_{p_j}\right)\right)\right)\right) \\
&= \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \left(\sum_{1\le i,j\le N} d^2_{S^2}\left(\vec{\mathbf{v}}_{p_i}, R_{\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}}\left(\vec{\mathbf{v}}_{p_j}\right)\right)\right)
\end{aligned} \tag{48}$$

By assigning $f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})$ as Equation (20), we have proven that Equation (16) can be derived into a NUG form, i.e., Equation (18).

Appendix IX Fourier expansion of a NUG, and a matrix form

In this section, we prove that Equation (18) is equivalent to Equation (23).

As G is a molecular symmetry group, G is compact. By the Peter-Weyl Theorem[2], the generalized Fourier expansion of $f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})$ is

$$f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}) = \sum_{k=0}^{K-1} d_k\, tr\left(\hat{\mathbf{f}}_{ij}^{(k)}\boldsymbol{\rho}_k(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})\right) \tag{49}$$

where $\hat{\mathbf{f}}_{ij}^{(k)}$ is determined by Equation (21).

Thus, Equation (18) can be derived into

$$
\begin{aligned}
& \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \sum_{1\le i,j\le N} f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}) \\
= & \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \sum_{1\le i,j\le N}\sum_{k=0}^{K-1} d_k\, tr\left(\hat{\mathbf{f}}_{ij}^{(k)}\boldsymbol{\rho}_k(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})\right) \\
= & \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \sum_{k=0}^{K-1}\sum_{1\le i,j\le N} tr\left(\left(d_k\hat{\mathbf{f}}_{ij}^{(k)}\right)\left(\boldsymbol{\rho}_k(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})\right)\right) \\
= & \underset{\mathbf{g}_{t_1},\mathbf{g}_{t_2},\cdots,\mathbf{g}_{t_N}\in G}{\operatorname{argmin}} \sum_{k=0}^{K-1} tr\left(\hat{\mathbf{F}}^{(k)}\mathbf{X}^{(k)}\right)
\end{aligned}
\tag{50}
$$

where $\hat{\mathbf{F}}^{(k)}$ is determined by Equation (24) and $\mathbf{X}^{(k)}$ is determined by Equation (26). Hence, we have proven that Equation (18) is equivalent to Equation (23).

Appendix X SDP relaxation of the matrix form of the NUG problem

The variable matrices $\left\{\mathbf{X}^{(k)}\right\}_{1\le k\le K}$ is defined by Equation (26), in the matrix form of the NUG problem. However, the constraints of such matrices are in general non-convex. For the purpose of convex optimization, the constraints of $\left\{\mathbf{X}^{(k)}\right\}_{1\le k\le K}$ are relaxed by the method proposed in[3]

$$
\begin{aligned}
& \mathbf{X}^{(k)} \succeq 0 && 0\le k\le K-1 \\
& \mathbf{X}_{ii}^{(k)} = \mathbf{I}_{d_k\times d_k} && 0\le k\le K-1, \quad 1\le i\le N \\
& \mathbf{X}_{ij}^{(0)} = \mathbf{I}_{d_0\times d_0} && 1\le i,j\le N \\
& \sum_{k=0}^{K-1} d_k\, tr\left(\boldsymbol{\rho}_k(\mathbf{g_m})^*\mathbf{X}_{ij}^{(k)}\right)\ge 0 && 1\le i,j\le N, \quad 1\le m\le |G|
\end{aligned}
$$

where $\mathbf{X}^{(k)}\succeq 0$ stands for that $\left\{\mathbf{X}^{(k)}\right\}_{1\le k\le K}$ are semi-definite positive (SDP) matrices.

Therefore, the NUG problem is relaxed into a SDP problem, as

$$
\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1} = \underset{\left\{\mathbf{X}^{(k)}\right\}_{0\le k\le K-1}}{\operatorname{argmin}} \sum_{k=0}^{K-1} tr\left(\hat{\mathbf{F}}^{(k)}\mathbf{X}^{(k)}\right), \tag{51}
$$

where $\left\{\mathbf{X}^{(k)}\right\}_{1\le k\le K}$ are subject to the above relaxed constraint. The solution matrices $\left\{\tilde{\mathbf{X}}^{(k)}\right\}_{0\le k\le K-1}$ are in the convex hull of the matrices defined in Equation (26).

Appendix XI Partial cost

Solving $\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N}$ minimizes the cost function Equation (18). Thus, during the sequential recovery of representatives from the solution of the SDP problem, the partial cost must be calculated at each recovery step for the purpose of determining the keeping trials. As in Equation (18), the cost $\sum_{1\leq i,j\leq N} f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1})$ is the summation of f_{ij} of all relations $\left\{\tilde{\mathbf{g}}_{t_i}\otimes\tilde{\mathbf{g}}_{t_j}^{-1}\right\}_{1\leq i,j,\leq N}$. Therefore, the partial cost can be calculated as

$$\sum_{\substack{1\leq i,j\leq N \\ relation\ \tilde{\mathbf{g}}_{t_i}\otimes\tilde{\mathbf{g}}_{t_j}^{-1} determined}} f_{ij}(\mathbf{g}_{t_i}\otimes\mathbf{g}_{t_j}^{-1}). \tag{52}$$

In other words, the partial cost is the summation of f_{ij} of all determined relations.

Appendix XII Design of a naive approach

In the section, a naive approach for solving Equation (14) is described.

The principle of obtaining $\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N}$ from Equation (14) is that a representative $\mathbf{q}_i\otimes\mathbf{g}_{t_i}$ is chosen from each set of equivalent spatial rotations $\mathbf{q}_i\otimes G$, such that all representatives are as close as possible overall. In this naive approach, this principle is realized as follows.

First, a spatial rotation $\tilde{\mathbf{q}}$ is selected as follows. For each $\mathbf{q}_i\otimes G$ from $\{\mathbf{q}_i\otimes G\}_{1\leq i\leq N}$, the overall distance between it and the other set of equivalent spatial rotations is calculated as

$$\sum_{1\leq j\leq N, j\neq i} d_{SO(3)/G}\left(\mathbf{q}_i\otimes G, \mathbf{q}_j\otimes G\right). \tag{53}$$

The set of equivalent spatial rotations with minimum such distance can be considered at the "center" of $\{\mathbf{q}_i\otimes G\}_{1\leq i\leq N}$. Therefore, the spatial rotation $\tilde{\mathbf{q}}$ is selected as an element of the "center" set of equivalent spatial rotations.

Then, from each set of equivalent spatial rotations $\mathbf{q}_i\otimes G$, the one closest to the "center" $\tilde{\mathbf{q}}$, i.e., $\mathbf{q}_i\otimes\tilde{\mathbf{g}}_i$, is selected as the representative. Therefore, $\tilde{\mathbf{g}}_i$ is determined for a spatial rotation set $\mathbf{q}_i\otimes G$.

By this means, a representative $\mathbf{q}_i\otimes\mathbf{g}_{t_i}$ is chosen from each set of equivalent spatial rotations $\mathbf{q}_i\otimes G$. Such a representative is closest to the "center" $\tilde{\mathbf{q}}$. The principle of obtaining $\{\tilde{\mathbf{g}}_{t_i}\}_{1\leq i\leq N}$ from Equation (14) is realized to some extent.

References

[1] Mingxu Hu, Qi Zhang, Jing Yang, and Xueming Li. Unit quaternion description of spatial rotation in 3d electron cryo-microscopy. *bioRxiv*, 2019.

[2] H. Weyl and P. Peter. Die vollstandigkeit der primitiven darstellungen einer geschlossenen kontinuierlichen gruppe. 97(1):737–755, 1927.

[3] Afonso S Bandeira, Yutong Chen, Roy R Lederman, and Amit Singer. Non-unique games over compact groups and orientation estimation in cryo-em. *Inverse Problems*, 36(6):064002, 2020.