
Learning to recover orientations from projections in single-particle cryo-EM

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 A major challenge in single-particle cryo-electron microscopy (cryo-EM) is that
2 the orientations adopted by the 3D particles prior to imaging are unknown; yet, this
3 knowledge is essential for high-resolution reconstruction. We present a method
4 to recover these orientations directly from the acquired set of 2D projections.
5 Our approach consists of two steps: (i) the estimation of distances between pairs
6 of projections, and (ii) the recovery of the orientation of each projection from
7 these distances. In step (i), pairwise distances are estimated by a Siamese neural
8 network trained on synthetic cryo-EM projections from resolved bio-structures.
9 In step (ii), orientations are recovered by minimizing the difference between
10 the distances estimated from the projections and the distances induced by the
11 recovered orientations. We evaluated the method on synthetic cryo-EM datasets.
12 Current results demonstrate that orientations can be accurately recovered from
13 projections that are shifted and corrupted with a high level of noise. The accuracy
14 of the recovery depends on the accuracy of the distance estimator. While not
15 yet deployed in a real experimental setup, the proposed method offers a novel
16 learning-based take on orientation recovery in SPA. Our code is available at <https://github.com/anonymous/protein-reconstruction>.
17

18 1 Introduction

19 Single-particle cryo-electron microscopy (cryo-EM) has revolutionized the field of structural biology
20 over the last decades [1, 2, 3]. The use of electron beams to image ice-embedded samples has
21 permitted the recovery of 3D bio-structures at unprecedented resolution. This “resolution revolution”
22 has had a tremendous impact in biomedical research, providing invaluable insights into the biological
23 processes that underlie many current diseases.

24 In single-particle cryo-EM, every 3D particle adopts a random orientation θ_i in the ice layer before
25 being imaged. Hence, the projection geometry associated to each acquired 2D projection (Figure 1)
26 is unknown. Yet, this knowledge is essential for the tomographic reconstruction of bio-structures [4].
27 We consider that a cryo-EM measurement (*i.e.*, a projection) $\mathbf{p}_i \in \mathbb{R}^{n_p}$ is acquired through

$$\mathbf{p}_i = \mathbf{C}_\varphi \mathbf{S}_{\mathbf{t}_i} \mathbf{P}_{\theta_i} \mathbf{x} + \mathbf{n}, \quad (1)$$

28 where $\mathbf{x} \in \mathbb{R}^{n_x}$ is the unknown 3D density map [5] (Coulomb potential). The operator $\mathbf{P}_{\theta_i} : \mathbb{R}^{n_x} \rightarrow$
29 \mathbb{R}^{n_p} is the projection along the orientation θ_i (*i.e.*, the x-ray transform). The operator $\mathbf{S}_{\mathbf{t}_i} : \mathbb{R}^{n_p} \rightarrow$
30 \mathbb{R}^{n_p} is a shift of the projection by $\mathbf{t}_i = (t_{i_1}, t_{i_2})$. The convolution operator $\mathbf{C}_\varphi : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_p}$ models
31 the microscope point-spread function (PSF) with parameters $\varphi = (d_1, d_2, \alpha_{ast})$, where d_1 is the
32 defocus-major, d_2 is the defocus-minor, and α_{ast} is the angle of astigmatism [6, 7]. Finally, $\mathbf{n} \in \mathbb{R}^{n_p}$
33 represents additive noise. Figure 11 illustrates the effect of projection, shift, and noise. The challenge
34 is then to reconstruct \mathbf{x} from a set of projections $\{\mathbf{p}_i\}_{i=1}^P$ acquired along unknown orientations.

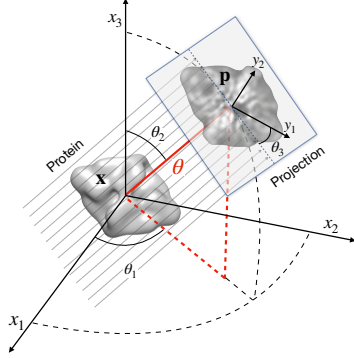


Figure 1: Geometry of the imaging model defined in (1). The 3D density \mathbf{x} in the coordinate system (x_1, x_2, x_3) is imaged along the orientation θ to produce the 2D projection \mathbf{p} in the coordinate system (y_1, y_2) of the microscope’s detector plane. The orientation $\theta = (\theta_3, \theta_2, \theta_1)$ is decomposed as the direction $(\theta_2, \theta_1) \in [0, \pi] \times [0, 2\pi[$ (parameterizing the sphere \mathbb{S}^2) and the in-plane rotation $\theta_3 \in [0, 2\pi[$ (parameterizing the circle \mathbb{S}^1). In our work, we represent the orientation θ as a unit quaternion q .

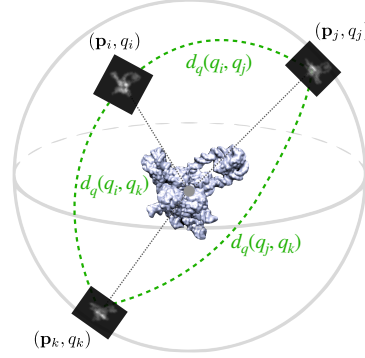


Figure 2: Single-particle cryo-EM produces P projections (with P in the order of 10^5) from unknown orientations: $\{(\mathbf{p}_i, q_i)\}_{i=1}^P$. Observing that distances between orientations constrain the latter, we aim to *recover the orientations* $\{q_i\}$ from $\{d_q(q_i, q_j)\}$, where $d_q(q_i, q_j)$ is the distance (angle) between orientations q_i and q_j . Observing that the similarity between projections depends on their relative orientation, we aim to *estimate the distance* $d_q(q_i, q_j)$ from the projections $(\mathbf{p}_i, \mathbf{p}_j)$.

35 A popular approach is to alternatively refine the 3D structure and estimated orientations [8, 9, 10, 11,
 36 12, 13]. Yet, the outcome of these iterative-refinement procedures is often predicated on the quality
 37 of the initial reconstruction, or, equivalently, on the initial estimation of the orientations [14, 15].

38 Several methods have been designed to produce a first rough *ab initio* structure for the refinement
 39 procedure [16]. *Moment-matching* techniques [17, 18, 19, 20] reconstruct an initial structure such
 40 that the first few moments of the distribution of its theoretical measurements match the ones of
 41 its experimental projections; however, they typically remain sensitive to error in data and can
 42 require relatively high computational complexity. Based on the central-slice theorem, *common-lines*
 43 methods [21, 8, 22, 23, 24, 25, 26] aim at uniquely determining the orientations of each projection by
 44 identifying the common-lines between triplets of projections—a real challenge given the massive
 45 amount of noise. Alternatively, the marginalized maximum likelihood (ML) formulation of the
 46 reconstruction problem [11]—classically used for the iterative-refinement procedures themselves—
 47 can be minimized using stochastic gradient descent [27]. This permits to avoid the need for an initial
 48 volume estimate, at the possible cost of greater convergence instability.

49 More recently, the recovery of geometrical information from unknown view tomography of 2D point
 50 sources has been proposed [28], but the extension to 3D cryo-EM tomography is not straightforward.
 51 Finally, [29] proposed to recover the in-plane rotations by learning to embed projections in an
 52 appropriate latent space, but only after directions had been estimated through three rounds of 2D
 53 classification in RELION.

54 Despite the aforementioned advances, providing a robust initial volume remains a challenge due to
 55 the high-dimensionality and ill-posedness of the underlying optimization problem. On the other hand,
 56 the remarkable ability of convolutional neural networks to capture relevant representations of images
 57 has had a profound influence in imaging [30]. In this work, we present a learning-based approach to
 58 recover the unknown orientations directly from the acquired set of projections—without the need for
 59 an intermediate reconstruction procedure or an initial volume estimate.

60 2 Method

61 Our approach relies on two observations (Figure 2), yielding two steps (Figure 3). First, the more
 62 similar two projections $(\mathbf{p}_i, \mathbf{p}_j)$, the more likely they originated from two particles that adopted close

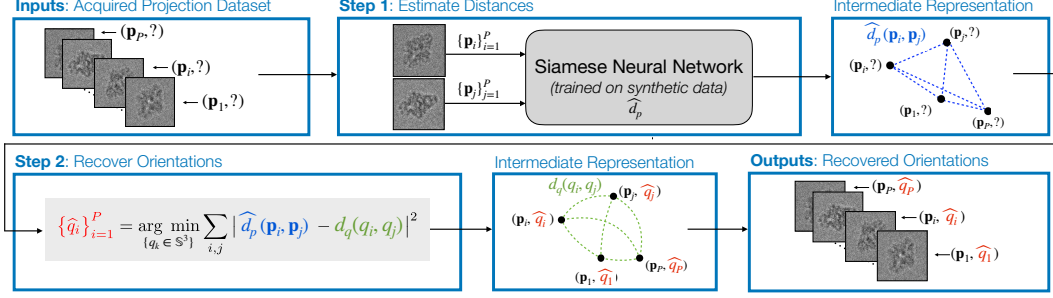


Figure 3: Our method consists of two steps. First, we estimate distances between pairs of projections. Second, we recover the orientation of each projection from these distances.

63 orientations (q_i, q_j) in the ice prior to imaging;¹ this observation guides a number of applications in
 64 the field [2]. Hence, we aim to *estimate distances* between orientations $d_q(q_i, q_j)$ from the projections
 65 as $\widehat{d}_p(\mathbf{p}_i, \mathbf{p}_j)$, which we discuss in §2.2. Second, an orientation q is constrained by the distances
 66 between itself and the other orientations $\{d(q, q_j)\}$. Hence, we aim to *recover orientations* $\{\widehat{q}_k\}$ such
 67 that the induced distances $\{d_q(\widehat{q}_i, \widehat{q}_j)\}$ are close to the estimated distances $\{\widehat{d}_p(\mathbf{p}_i, \mathbf{p}_j)\}$, which we
 68 discuss in §2.3. All in all, from a set of projections $\{\mathbf{p}_k\}$, we aim to recover their orientations $\{\widehat{q}_k\}$
 69 such that $d_q(\widehat{q}_i, \widehat{q}_j) \approx \widehat{d}_p(\mathbf{p}_i, \mathbf{p}_j) \approx d_q(q_i, q_j)$, with equality if \widehat{d}_p and $\{\widehat{q}_k\}$ are perfectly estimated.
 70 Our approach is similar to [31]. While the authors reconstruct 2D images from 1D projections, they
 71 rely on the same two-step approach: they (i) estimate distances as $\widehat{d}_p(\mathbf{p}_i, \mathbf{p}_j) = \|\mathbf{p}_i - \mathbf{p}_j\|_2$ then
 72 (ii) recover the orientations by spectrally embedding that distance graph. The Euclidean distance is
 73 however not robust to perturbations: for example, two projections that only differ by a shift \mathbf{S}_t of one
 74 pixel would be considered far apart while their orientations are the same. They noted that issue and
 75 we observed it too (Appendix E). To circumvent this, we propose to *learn* \widehat{d}_p from examples (§2.2).

76 2.1 Representation of orientations with quaternions

77 The orientation of a 3D particle with respect to the microscope’s detector plane is a rotation relative to
 78 a reference orientation (Figure 1). The group of all 3D rotations under composition is identified with
 79 $\mathbf{SO}(3)$, the group of 3×3 orthogonal matrices with determinant 1 under matrix multiplication. A
 80 rotation matrix $\mathbf{R}_\theta \in \mathbf{SO}(3)$ can be decomposed as a product of $\binom{3}{2} = 3$ independent rotations, for
 81 example as $\mathbf{R}_\theta = \mathbf{R}_{\theta_3} \mathbf{R}_{\theta_2} \mathbf{R}_{\theta_1}$, where $\theta = (\theta_3, \theta_2, \theta_1) \in [0, 2\pi[\times [0, \pi] \times [0, 2\pi[$ are the (extrinsic
 82 and proper) Euler angles in the ZYZ convention (a common parameterization in cryo-EM) [32].

83 While Euler angles are a concise representation of orientation (3 numbers for 3 degrees of freedom),
 84 they suffer from a topological constraint—there is no covering map from the 3-torus to $\mathbf{SO}(3)$ —
 85 which manifests itself in the *gimbal lock*, the loss of one degree of freedom when $\theta_2 = 0$. This makes
 86 their optimization by gradient descent (§2.3) problematic. On the other hand, optimizing rotation
 87 matrices (made of 9 numbers) would require computationally costly constraints (orthogonality and
 88 determinant 1) to reduce the degrees of freedom to 3. Moreover, the distance between orientations
 89 cannot be directly computed from Euler angles and is costly (30 multiplications) to compute from
 90 rotation matrices [33]. We solve both problems by representing orientations with unit quaternions.

91 Quaternions $q \in \mathbb{H}$ are an extension of complex numbers² of the form $q = a + bi + cj + dk$ where
 92 $a, b, c, d \in \mathbb{R}$. Unit quaternions $q \in \mathbb{S}^3$, where $\mathbb{S}^3 = \{q \in \mathbb{H} : |q| = 1\}$ is the 3-sphere (with
 93 the additional group structure inherited from quaternion multiplication), concisely and elegantly
 94 represent a rotation of angle θ about axis (x_1, x_2, x_3) as $q = \cos(\theta/2) + x_1 \sin(\theta/2)\mathbf{i} + x_2 \sin(\theta/2)\mathbf{j} +$
 95 $x_3 \sin(\theta/2)\mathbf{k}$. They parameterize rotation matrices as

$$\mathbf{R}_q = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2bd + 2ac \\ 2bc + 2ad & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2cd + 2ab & a^2 - b^2 - c^2 + d^2 \end{pmatrix}.$$

¹Up to protein symmetries, which we discuss later.

²The algebra \mathbb{H} is similar to the algebra of complex numbers \mathbb{C} , with the exception of multiplication being non-commutative.

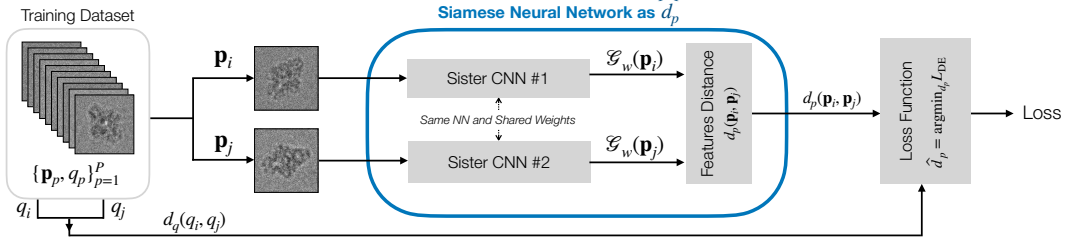


Figure 4: Distance learning. We are looking for a distance \hat{d}_p between projections that is an accurate estimator of the distance d_q between their orientations. We propose to parameterize \hat{d}_p as a Siamese neural network (SNN), trained on a synthetic dataset of projections with associated orientation.

96 Note that $\mathbb{S}^3 \rightarrow \mathbf{SO}(3)$ is a two-to-one mapping (a double cover) as q and $-q$ represent the same
 97 orientation. Unlike Euler angles, \mathbb{S}^3 is isomorphic to the universal cover of $\mathbf{SO}(3)$. Hence, the
 98 distance between two orientations, *i.e.*, the length of the geodesic between them on $\mathbf{SO}(3)$, is

$$d_q : \mathbb{S}^3 \times \mathbb{S}^3 \rightarrow [0, \pi], \quad (2)$$

$$d_q(q_i, q_j) = 2 \arccos(|\langle q_i, q_j \rangle|),$$

99 where $\langle \cdot, \cdot \rangle$ is the inner product, and the absolute value $|\cdot|$ ensures that $d_q(q_i, q_j) = d_q(q_i, -q_j)$. The
 100 distance $d_q(q_i, q_j)$ corresponds to the magnitude of the rotation \mathbf{R}_* such that $\mathbf{R}_{q_i} = \mathbf{R}_* \mathbf{R}_{q_j}$ [33].

101 2.2 Distance learning

102 We aim to estimate a function \hat{d}_p such that $\hat{d}_p(\mathbf{p}_i, \mathbf{p}_j) \approx d_q(q_i, q_j)$. While we could in principle
 103 design \hat{d}_p , that would be intricate—if not impossible—partly because the invariants are difficult
 104 to specify. We instead opt to learn \hat{d}_p , capitalizing on (i) the powerful function approximation
 105 capabilities of neural networks, and (ii) the possibility to generate realistic datasets supported by the
 106 availability of numerous 3D atomic models³ and our ability to model the cryo-EM imaging procedure.

107 From a training dataset $\{\mathbf{p}_i, q_i\}_{i=1}^P$, we learn the projection distance

$$\hat{d}_p = \arg \min_{d_p} L_{DE}, \quad \text{where} \quad L_{DE} = \sum_{i,j} |d_p(\mathbf{p}_i, \mathbf{p}_j) - d_q(q_i, q_j)|^2 \quad (3)$$

108 is the loss and d_q is defined in (2). The d_p is parameterized as the Siamese neural network (SNN) [34]

$$d_p(\mathbf{p}_i, \mathbf{p}_j) = d_f(\mathcal{G}_w(\mathbf{p}_i), \mathcal{G}_w(\mathbf{p}_j)),$$

109 where \mathcal{G}_w is a convolutional neural network with weights w that is trained to extract the most relevant
 110 features $\mathbf{f}_i \in \mathbb{R}^{n_f}$ from a projection \mathbf{p}_i . SNNs, also termed “twin networks”, are commonly used in
 111 the field of deep metric learning to learn similarity functions [35]. We set the feature space distance
 112 d_f as the cosine distance to facilitate the learning of a \hat{d}_p that respects the elliptic geometry of \mathbb{S}^3
 113 (Appendix F). Figure 4 illustrates the proposed learning paradigm.

114 As evaluating a sum over P^2 pairs is computationally intractable for cryo-EM datasets with typically
 115 P in the order of 10^5 projections, we sample the sum and minimize (3) with stochastic gradient
 116 descent (SGD) over small batches of pairs. The weights w are updated by back-propagation.

117 The architecture of \mathcal{G}_w is described in Appendix G. When designing the architecture, we constrain
 118 the functional space from which the trained \mathcal{G}_w is drawn and express our prior expert knowledge. For
 119 example, we realize shift invariance, *i.e.*, a guarantee that a shift \mathbf{S}_t does not change our estimated
 120 distances and orientations, with a fully convolutional architecture. Size invariance, *i.e.*, taking
 121 projections \mathbf{p} of varying sizes n_p while yielding a representation \mathbf{f} of a fixed size n_f , is realized by a
 122 final average pooling layer. As we do not (yet) know how to realize an invariance to noise or PSF, we
 123 resort to data augmentation, *i.e.*, training on perturbed projections. In §3.4, we show that a built-in
 124 invariance (shift) is far preferable to one learned through augmentation (noise). Finally, as projections

³<https://www.ebi.ac.uk/pdbe/emdb>

125 are made by integrating through the 3D volume, projections from opposed directions are mirrors of
 126 each other.⁴ That is another kind of physical knowledge that should ideally be built into our method.

127 One could hope to train \mathcal{G}_w to directly map projections to orientations as $\hat{q}_i = \mathbf{f}_i = \mathcal{G}_w(\mathbf{p}_i)$. While
 128 that would avoid the orientation recovery step, a space of $n_f = 4$ dimensions does not have room for
 129 \mathcal{G}_w to represent the other factors of variation in \mathbf{p} , such as different noise levels, PSFs, or proteins.
 130 We tested that hypothesis in Appendix F.

131 2.3 Orientation recovery

132 The task of recovering points based on their relative distances has been extensively studied. Many
 133 methods aim at mapping high-dimensional data onto a lower-dimensional space while preserving
 134 distances, primarily for dimensionality reduction and data visualization. Well-known examples
 135 include MDS [36], Isomap [37], LLE [38], Laplacian eigenmaps [39], t-SNE [40], and UMAP [41].
 136 The embedding of distance matrices in Euclidean space (given by their eigenvectors) is especially
 137 well-described. In particular, the framework of Euclidean distance matrices (EDMs) [42] provides
 138 theoretical guarantees on the recovery of points from distances.

139 We however aim to embed the orientations q in \mathbb{S}^3 (§2.1), a setting for which we are unaware of any
 140 theoretical characterization (*e.g.*, on the shape of the loss function or its behavior when distances are
 141 missing or noisy). The fact that \mathbb{S}^3 is locally Euclidean does however offer some hope. Indeed, despite
 142 the non-convexity and the lack of theoretical guarantees, we are able to appropriately minimize our
 143 loss function, as we experimentally demonstrate in Appendix D.

144 We recover the orientations of a set of projections $\{\mathbf{p}_k\}_{k=1}^P$ through

$$\{\hat{q}_k\}_{k=1}^P = \arg \min_{\{q_k \in \mathbb{S}^3\}} L_{\text{OR}}, \quad \text{where} \quad L_{\text{OR}} = \sum_{i,j} \left| \hat{d}_p(\mathbf{p}_i, \mathbf{p}_j) - d_q(q_i, q_j) \right|^2 \quad (4)$$

145 is the loss and \hat{d}_p is the estimator trained in (3). Note that the sole difference with (3) is that the
 146 minimization is performed over the orientations q rather than the distance d_p . Here again, we sample
 147 the sum in practice and minimize (4) with mini-batch SGD. Sampling the sum amounts to building a
 148 sparse (instead of complete) distance graph before embedding, a common strategy.

149 2.4 Evaluation

150 While not a part of the method *per se*, we must evaluate the quality of the recovered orientations.
 151 Unfortunately, we cannot directly take the difference between the recovered orientations $\{\hat{q}_k\}_{k=1}^P$
 152 and the true orientations $\{q_k\}_{k=1}^P$ as orientations are rotations up to an arbitrary reference orientation.
 153 Any global rotation or reflection of the recovered orientations is as valid as any other, *i.e.*, $d_q(q_i, q_j) =$
 154 $d_q(\mathbf{T}q_i, \mathbf{T}q_j) \forall \mathbf{T} \in \mathbf{O}(4)$, where $\mathbf{O}(4)$ is the group of 4×4 orthogonal matrices. Hence, we align
 155 the sets of orientations and compute the *mean orientation recovery error* as

$$E_{\text{OR}} = \min_{\mathbf{T} \in \mathbf{O}(4)} \frac{1}{P} \sum_{i=1}^P |d_q(q_i, \mathbf{T}\hat{q}_i)|. \quad (5)$$

156 We implement \mathbf{T} as a product of $\binom{4}{2} = 6$ independent rotations and an optional reflection:

$$\mathbf{T} = \begin{bmatrix} m & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \prod_{1 \leq i < j \leq 4} \mathbf{T}_{\theta_{ij}}, \quad m \in \{-1, 1\}, \theta_{ij} \in [0, 2\pi[, \quad (6)$$

157 where $\mathbf{T}_{\theta_{ij}} \in \mathbf{SO}(4)$ is a rotation by angle θ_{ij} on the (x_i, x_j) plane.

158 In practice, we again minimize (5) with mini-batch SGD. Because $\mathbf{O}(4)$ is disconnected, we optimize
 159 the 6 angles separately for $m = 1$ (proper rotations) and $m = -1$ (improper rotations). Figure 15
 160 shows an alignment to $E_{\text{OR}} = 0$ after a perfect recovery.

⁴That fact prevents the resolution of chirality, *i.e.*, we cannot distinguish a protein from its mirrored form.

161 3 Experiments

162 We first evaluated whether orientation recovery through (4) was feasible assuming perfect distances,
163 and how it was affected by errors in the distances (§3.2). We then learned to estimate the distances
164 through (3), and evaluated the accuracy of this procedure (§3.3) and its robustness to perturbations of
165 the projections (§3.4). Finally, we ran the whole machinery on a synthetic dataset to assess how well
166 orientations could be recovered from estimated distances (§3.5).

167 3.1 Experimental conditions

168 **Density maps.** We considered two proteins (Figure 10): the β -galactosidase, a protein with a
169 dihedral (D2) symmetry, and the lambda excision HJ intermediate (HJI), an asymmetric protein
170 with local cyclic (C1) symmetry. Their deposited PDB atomic models are 5a1a [43] and 5j0n [44],
171 respectively. From these atomic models, we generated the density maps in Chimera [45] by fitting the
172 models with a 1Å map for 5a1a and a 3.67Å map for 5j0n; this gave us a volume of $110 \times 155 \times 199$
173 voxels for 5a1a and one of $69 \times 57 \times 75$ voxels for 5j0n.

174 **Protein symmetries.** Symmetries are problematic when learning distances: two projections can
175 be identical while not originating from the same orientation, which breaks an axiom of distance
176 functions (identity of indiscernibles). Figure 16b illustrates this problem. To capture only one of four
177 identical projections of 5a1a, we restricted directions to $(\theta_2, \theta_1) \in [0, \pi[\times [0, \frac{\pi}{2}[$ (a quarter of the
178 sphere, illustrated in Figure 12a) for that protein. This treatment of symmetries is incomplete⁵ but
179 sufficient for a proof-of-concept.

180 **Projections.** Using the ASTRA projector [46], we generated $P = 5,000$ synthetic projections
181 of 275×275 pixels (downsampled to 116×116) for 5a1a and 116×116 pixels for 5j0n, taken
182 from uniformly sampled orientations.⁶ We then perturbed the measurements with different levels of
183 additive Gaussian noise [47, 48] and off-centering shifts. Figure 11 displays some samples.

184 **Datasets.** For each protein, we split the projections into training, validation, and test subsets, and
185 created *disjoint* pairs of projections from each (Table 1). The training and validation sets were used to
186 train and evaluate the SNN, while the test set was used to evaluate orientation recovery given a trained
187 SNN. Sampling orientations (mostly) uniformly induces a distribution of distances that is skewed
188 towards larger distances (shown in Figure 12b). As this would skew L_{DE} and bias \hat{d}_p , we further
189 sampled 1% of the training and validation pairs to make the distribution of distances uniform—for
190 \hat{d}_p to be uniformly accurate over the whole $[0, \pi]$ range of distances (see Appendix B for further
191 illustrations). While 1,650 projections were enough to perfectly reconstruct the density maps (as
192 shown in Figures 9e and 9j), our method is not limited by the number of projections as optimization
193 is done per batch. Optimization settings are described in Appendix C.

194 3.2 Sensitivity of orientation recovery to errors in distance estimation

195 We first evaluated the feasibility of orientation recovery assuming that the exact distances were known.
196 The method successfully recovers the orientation of every projection in this case (see Appendix D).

197 To evaluate the robustness of (4), we perturbed the distances prior to recovery with an error sampled
198 from a Gaussian distribution with mean 0 and variances $\sigma^2 \in [0.0, 0.8]$. Figure 5 shows that the
199 recovery error E_{OR} is a monotonic function of the error in distances: from $E_{OR} = 0$ with exact
200 distances to $E_{OR} \approx 0.2$ radians ($\approx 11.5^\circ$) for $\sigma^2 = 0.8$.

201 These results demonstrate that the performance of orientation recovery (4) depends on the quality of
202 the estimated distances, which advocates for a proper and extensive training of the SNN. Moreover,
203 we observe that L_{OR} is a reliable proxy for E_{OR} , allowing us to assess recovery performance in the
204 absence of ground-truth orientations (*i.e.*, when recovering the orientations of real projections).

⁵The remaining issue is that one of four distances is arbitrarily chosen per pair of projections.

⁶Orientations used in §3.2 (Figure 5) and §3.4 (Figure 7) were actually obtained by uniformly sampling the Euler angles θ , constrained to $(\theta_3, \theta_2, \theta_1) \in [0, 2\pi[\times [0, \frac{\pi}{2}[\times [0, 2\pi[$ for 5j0n. Our conclusions would be identical if orientations were uniformly sampled from $\text{SO}(3)$ instead.

Table 1: Split of $P = 5,000$ projections in training, validation, and test subsets.

Dataset	P	P^2	Used pairs
Training	2,512 (50%)	6,310,144	63,101
Validation	838 (17%)	702,244	7,022
Test	1,650 (33%)	2,722,500	2,722,500

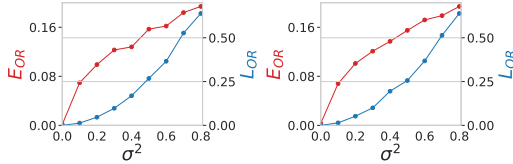
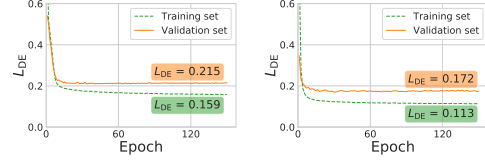
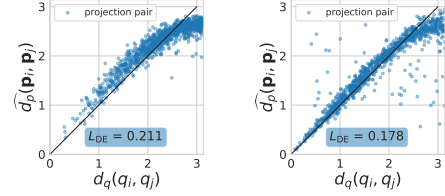


Figure 5: Orientation recovery from perturbed distances on 5j0n (left) and 5a1a (right).



(a) Loss converged on 5j0n (left) and 5a1a (right).



(b) Relationship between \hat{d}_p and d_q on 1,000 pairs from the test sets of 5j0n (left) and 5a1a (right).

Figure 6: Distance learning.

205 3.3 Learning to estimate distances

206 We evaluated the ability of the SNN to learn to approximate the orientation distance d_q . For
 207 comparison, we evaluated a baseline, the Euclidean distance $\hat{d}_p(\mathbf{p}_i, \mathbf{p}_j) = \|\mathbf{p}_i, \mathbf{p}_j\|_2$, in Appendix E.

208 Figure 6a shows the convergence of L_{DE} , reached in about 50 epochs. Figure 6b shows the relationship
 209 between the distance \hat{d}_p estimated from projections and the true distance d_q . The outliers for 5a1a
 210 are explained by our incomplete treatment of its symmetry. While our learned distance function
 211 is a much better estimator than the Euclidean distance—compare Figure 6b with Figure 16—they
 212 share one characteristic: both plateau and underestimate the largest distances. We did attenuate
 213 the phenomenon by sampling training distances uniformly (see §3.1), and the issue is much less
 214 severe than with the Euclidean distance. An alternative could be to only rely on smaller distances for
 215 recovery. That would however require the addition of a spreading term in (4) to prevent the recovered
 216 orientations to collapse.

217 These results confirm that a SNN is able to estimate differences in orientations from projections alone,
 218 even though much has yet to be gained from improving upon the rather primitive SNN architecture
 219 we are currently using. The use of additional training data should help further diminish overfitting.

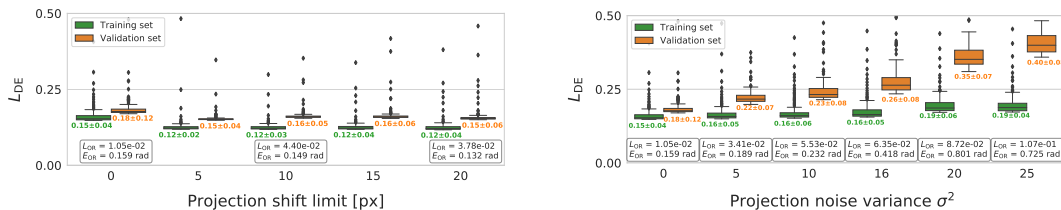
220 3.4 Sensitivity of distance learning to perturbations in the projections

221 We first demonstrated that the learning of distances is insensible to off-centering shifts (Figure 7a),
 222 which is expected given that shift invariance is built in our SNN (see §2.2).

223 As we cannot—or do not yet know how to—build noise invariance in the SNN architecture, we trained
 224 the SNN on noisy projections and evaluated whether it could learn to treat noise as an irrelevant
 225 information. Figure 7b shows $E_{OR} \approx 0.16$ radians ($\approx 9^\circ$) for noiseless projections and $E_{OR} \approx 0.42$
 226 radians ($\approx 24^\circ$) for a more realistic noise variance of $\sigma^2 = 16$ (with signal-to-noise ratio of -12 dB).
 227 Whereas a naive distance function (*e.g.*, an Euclidean distance) would be extremely sensitive to noise,
 228 the SNN mostly learned to discard it. Moreover, the observed overfitting indicates that more training
 229 data should further decrease the sensitivity of the SNN to noise.

230 Note that we did not evaluate sensitivity to the PSF at this stage but expect a similar behavior.

231 Here again (§3.2), we observed that (i) the estimation of more accurate distances (a smaller L_{DE})
 232 leads to the recovery of more accurate orientations (a smaller L_{OR} and E_{OR}), and that (ii) an higher
 233 recovery loss L_{OR} induces an higher error E_{OR} .



(a) Learning from shifted projections $\{\mathbf{S}_{t_i} \mathbf{P}_{\theta_i} \mathbf{x}\}$, with shifts t_{i_1} and t_{i_2} sampled from a triangular distribution with mean 0 and of increasing limits.

(b) Learning from noisy projections $\{\mathbf{P}_{\theta_i} \mathbf{x} + \mathbf{n}\}$, with white noise $\mathbf{n} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ of increasing variance σ^2 .

Figure 7: Sensitivity of distance learning to perturbations in the projections of 5j0n. The box plots show the distance learning loss L_{DE} (the distribution is taken over epochs). Boxes show the orientation recovery loss L_{OR} and error E_{OR} .

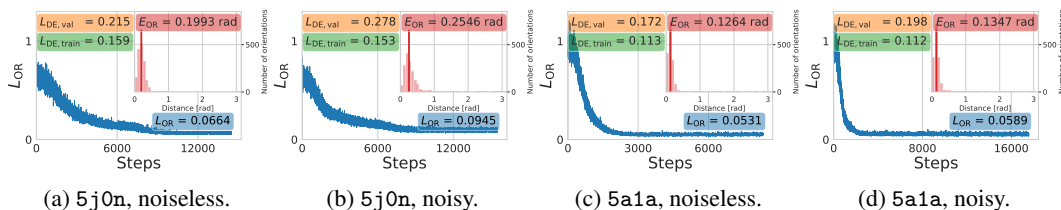


Figure 8: Distance learning and orientation recovery from estimated distances. The green and orange boxes show L_{DE} (3) on the training and validation sets. The blue curve shows the evolution of the recovery loss until convergence, with the minimum L_{OR} (4) highlighted. The red histogram shows the errors in the recovered orientations $\{d_q(q_i, \mathbf{T}\hat{q}_i)\}$, with the mean E_{OR} (5) highlighted.

234 3.5 Orientation recovery and reconstruction of density maps

235 As a proof-of-concept, we attempted to solve the full inverse problem posed by (1), *i.e.*, to reconstruct
 236 the density maps $\hat{\mathbf{x}}$ from sets of projections $\{\mathbf{p}_i\}$ and their orientations $\{\hat{q}_i\}$ recovered through the
 237 proposed method. It is worth noting that, at this stage of development, we only trained the SNN on
 238 projections originating from the protein we were attempting to reconstruct. In addition, reconstruction
 239 was performed with a direct reconstruction algorithm (ASTRA’s GPU implementation of the CGLS
 240 algorithm) rather than with a robust iterative method. This is a specific experimental case that only
 241 partially shines light on the applicability of the method in real situations; this is discussed in §4.

242 Figure 8a shows the recovery of orientations from distances that were estimated from noiseless
 243 projections of 5j0n. A mean error of $E_{OR} \approx 0.20$ radians ($\approx 11^\circ$) in the recovered orientations led
 244 to a reconstruction with a resolution of 12.2\AA at a Fourier shell coefficient (FSC) of 0.5, shown in
 245 Figure 9c. As predicted by our other experiments, corrupting the projections with noise ($\sigma^2 = 16$)
 246 negatively impacts the quality of the recovered orientations (Figure 8b); the obtained mean error
 247 is then $E_{OR} \approx 0.25$ radians ($\approx 14^\circ$). Unsurprisingly, this leads to a reconstruction with a lower
 248 resolution of 15.2\AA , shown in Figure 9d. (Note that reconstruction was here obtained from the
 249 noiseless projections, the goal being to evaluate only the impact of orientation mis-estimation.)

250 Finally, Figures 8c,d show the recovery of orientations from noiseless and noisy projections of 5a1a.
 251 A mean error of $E_{OR} \approx 0.13$ radians ($\approx 7^\circ$) in both cases led to reconstructions with resolutions of
 252 8.0\AA and 9.6\AA , shown in Figures 9h,i. Distance estimation, orientation recovery, and reconstruction
 253 performed better on 5a1a than 5j0n because its ground-truth density is of higher resolution.

254 These results tend to indicate that a reasonable first structure can be reconstructed from projections
 255 whose orientations have been recovered through our method.

256 4 Discussion

257 In this work, we explored the use of distance learning between pairs of 2D cryo-EM projections
 258 from a 3D protein structure to infer the unknown orientation at which each projection was imaged

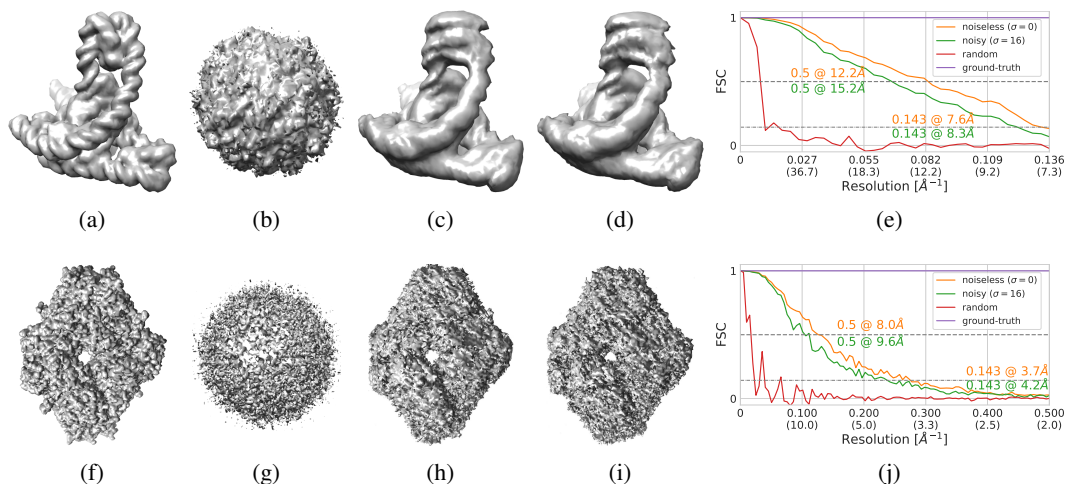


Figure 9: Density maps \hat{x} reconstructed from (a,f) ground-truth orientations, (b,g) random orientations, (c,h) orientations recovered from noiseless projections, and (d,i) orientations recovered from noisy projections. The Fourier shell correlation (FSC) curves in (e,j) indicate the resolutions of the densities (w.r.t. ground-truth densities, shown in Figures 10b,d).

259 from. Our two-step method relies on the estimation of pairwise distances between unseen projections,
 260 followed by the recovery of the orientations from these distances.

261 The method has been evaluated on synthetic datasets for two different proteins. The results provide
 262 key insights on the viability of the proposed scheme. First, they demonstrate that a SNN can learn
 263 a distance function between projections that estimates the difference in their orientation (§3.3) and
 264 that is invariant to shifts and robust to increasing levels of noise (§3.4)—an important condition in
 265 cryo-EM. Second, they demonstrate that an accurate estimation of distances leads to an accurate
 266 recovery of orientations (§3.2, §3.4). Finally, our method was able to recover orientations with
 267 an error of 0.12 to 0.25 radians (7 to 14°)—leading to an initial volume with a resolution of 8 to
 268 15Å (§3.5). In summary, the more accurate the estimated distances, the more precise the recovered
 269 orientations, and, ultimately, the higher-resolution the reconstructed volume.

270 While the method is not yet ready to be deployed in practice, we believe that a series of developments
 271 could make it relevant for single-particle cryo-EM reconstruction.⁷ As previously discussed, the
 272 results underline the importance of learning an accurate distance estimator. In this regard, the
 273 performance of the SNN could be improved. First, the architecture of the twin convolutional neural
 274 networks should be expanded and tuned. Second, training could be improved, perhaps by providing
 275 more supervision by separately predicting the differences in direction (θ_2, θ_1) and in-plane angle θ_3 .

276 Importantly, the SNN would be better trained on a more diverse cryo-EM dataset. Indeed, its success
 277 as a faithful estimator eventually relies on our capacity to generate a synthetic training dataset whose
 278 data distribution is diverse enough to cover that of unseen projection datasets. Such realistic cryo-EM
 279 projections could be generated by relying on a more expressive formulation of the cryo-EM physics
 280 and taking advantage of the thousands of atomic models available in the PDB. In particular, a
 281 necessary extension will be to include the effects of the PSF and to evaluate its impact.

282 A final phase of tests before deploying the method on real cryo-EM measurements will be to
 283 extensively test the method on “unseen proteins”, *i.e.*, proteins whose simulated projections have
 284 never been seen by the SNN. In this regard, an interesting aspect of our method is that the twin
 285 networks within the SNN intrinsically predict the *relationship* between projections, allowing the SNN
 286 as a whole to abstract the particular volume. Learning should benefit from the profound structural
 287 similarity shared by proteins—after all, they are all derived from the same 21 building blocks.

288 Training our 4.5M parameter model (see Appendices G and C) has the following negative environ-
 289 mental impact: it consumes 13 kWh of energy, which produces 6.36 lbs of CO₂ on average [49].

⁷Note that the present project will not be further continued by its authors due to other professional occupations. Hence, we strongly encourage anyone interested to build on these ideas and, hopefully, make it a practical tool.

290 **References**

- 291 [1] J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowell, and P. Schultz,
 292 “Cryo-electron microscopy of vitrified specimens,” *Quarterly Reviews of Biophysics*, vol. 21,
 293 no. 2, pp. 129–228, 1988.
- 294 [2] J. Frank, *Three-dimensional electron microscopy of macromolecular assemblies: Visualization*
 295 *of biological molecules in their native state*. Oxford University Press, 2006.
- 296 [3] Nature, “Method of the year 2015,” *Nature Methods*, vol. 13, no. 1, 2016.
- 297 [4] F. Natterer, *The mathematics of computerized tomography*. Society for Industrial and Applied
 298 Mathematics, jan 2001.
- 299 [5] F. DiMaio, D. A. Kondrashov, E. Bitto, A. Soni, C. A. Bingman, G. N. Phillips, and
 300 J. W. Shavlik, “Creating protein models from electron-density maps using particle-filtering
 301 methods,” *Bioinformatics*, vol. 23, no. 21, pp. 2851–2858, Nov. 2007. [Online]. Available:
 302 <https://academic.oup.com/bioinformatics/article/23/21/2851/374177>
- 303 [6] M. Vulović, R. B. Ravelli, L. J. van Vliet, A. J. Koster, I. Lazić, U. Lücken, H. Rullgård,
 304 O. Öktem, and B. Rieger, “Image formation modeling in cryo-electron microscopy,” *Journal of*
 305 *Structural Biology*, vol. 183, no. 1, pp. 19–32, Jul. 2013.
- 306 [7] H. Rullgård, L.-G. Öfverstedt, S. Masich, B. Daneholt, and O. Öktem, “Simulation of transmis-
 307 sion electron microscope images of biological specimens,” *Journal of Microscopy*, vol. 243,
 308 no. 3, pp. 234–256, 2011.
- 309 [8] P. A. Penczek, R. A. Grassucci, and J. Frank, “The ribosome at improved resolution: New
 310 techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological
 311 particles,” *Ultramicroscopy*, vol. 53, no. 3, pp. 251–270, 1994.
- 312 [9] T. Baker and R. Cheng, “A model-based approach for determining orientations of biological
 313 macromolecules imaged by cryoelectron microscopy,” *Journal of Structural Biology*, vol. 116,
 314 no. 1, pp. 120–130, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1047847796900209>
- 315 [10] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via
 316 the EM algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39,
 317 no. 1, pp. 1–38, 1977.
- 318 [11] F. J. Sigworth, “A maximum-likelihood approach to single-particle image refinement,” *Journal*
 319 *of structural biology*, vol. 122, no. 3, pp. 328–339, 1998.
- 320 [12] S. Scheres, “A bayesian view on cryo-EM structure determination,” *Journal of molecular*
 321 *biology*, vol. 415, no. 2, pp. 406–418, 2012.
- 322 [13] M. Zehni, L. Donati, E. Soubies, Z. J. Zhao, and M. Unser, “Joint Angular Refinement and
 323 Reconstruction for Single-Particle Cryo-EM,” *IEEE Transactions on Image Processing*, 2020.
- 324 [14] C. O. S. Sorzano, R. Marabini, A. Pascual-Montano, S. H. Scheres, and J. M. Carazo, “Opti-
 325 mization problems in electron microscopy of single particles,” *Annals of Operations Research*,
 326 vol. 148, no. 1, pp. 133–165, 2006.
- 327 [15] R. Henderson, A. Sali, M. L. Baker, B. Carragher, B. Devkota, K. H. Downing, E. H. Egelman,
 328 Z. Feng, J. Frank, N. Grigorieff, W. Jiang, S. J. Ludtke, O. Medalia, P. A. Penczek, P. B.
 329 Rosenthal, M. G. Rossmann, M. F. Schmid, G. F. Schröder, A. C. Steven, D. L. Stokes, J. D.
 330 Westbrook, W. Wriggers, H. Yang, J. Young, H. M. Berman, W. Chiu, G. J. Kleywegt, and C. L.
 331 Lawson, “Outcome of the first electron microscopy validation task force meeting,” *Structure*,
 332 vol. 20, no. 2, pp. 205–214, 2012.
- 333 [16] A. Singer and F. J. Sigworth, “Computational methods for single-particle electron cryomi-
 334 croscopy,” *Annual Review of Biomedical Data Science*, vol. 3, 2020.
- 335 [17] Z. Kam, “The reconstruction of structure from electron micrographs of randomly oriented
 336 particles,” in *Electron Microscopy at Molecular Dimensions*. Springer, 1980, pp. 270–277.
- 337 [18] D. B. Salzman, “A method of general moments for orienting 2d projections of unknown 3d
 338 objects,” *Computer vision, graphics, and image processing*, vol. 50, no. 2, pp. 129–156, 1990.
- 339 [19] A. Goncharov, “Integral geometry and three-dimensional reconstruction of randomly oriented
 340 identical particles from their electron microphotos,” *Acta Applicandae Mathematica*, vol. 11,
 341 no. 3, pp. 199–211, 1988.
- 342

- 343 [20] N. Sharon, J. Kileel, Y. Khoo, B. Landa, and A. Singer, “Method of moments for 3-d single
344 particle ab initio modeling with non-uniform distribution of viewing angles,” *Inverse Problems*,
345 2019.
- 346 [21] M. Van Heel, “Angular reconstitution: a posteriori assignment of projection directions for 3d
347 reconstruction,” *Ultramicroscopy*, vol. 21, no. 2, pp. 111–123, 1987.
- 348 [22] S. P. Mallick, S. Agarwal, D. J. Kriegman, S. J. Belongie, B. Carragher, and C. S. Potter,
349 “Structure and view estimation for tomographic reconstruction: A bayesian approach,” in *2006*
350 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*,
351 vol. 2. IEEE, 2006, pp. 2253–2260.
- 352 [23] A. Singer, R. R. Coifman, F. J. Sigworth, D. W. Chester, and Y. Shkolnisky, “Detecting consistent
353 common lines in cryo-EM by voting,” *Journal of structural biology*, vol. 169, no. 3, pp. 312–322,
354 2010.
- 355 [24] L. Wang, A. Singer, and Z. Wen, “Orientation determination of cryo-EM images using least
356 unsquared deviations,” *SIAM journal on imaging sciences*, vol. 6, no. 4, pp. 2450–2483, 2013.
- 357 [25] I. Greenberg and Y. Shkolnisky, “Common lines modeling for reference free ab-initio recon-
358 struction in cryo-EM,” *Journal of structural biology*, vol. 200, no. 2, pp. 106–117, 2017.
- 359 [26] G. Pragier and Y. Shkolnisky, “A common lines approach for ab initio modeling of cyclically
360 symmetric molecules,” *Inverse Problems*, vol. 35, no. 12, p. 124005, 2019.
- 361 [27] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, “cryoSPARC: Algorithms for
362 rapid unsupervised cryo-EM structure determination,” *Nature Methods*, vol. 14, no. 3, p. 290,
363 2017.
- 364 [28] M. Zehni, S. Huang, I. Dokmanić, and Z. Zhao, “Distance retrieval from unknown view
365 tomography of 2d point sources,” *Electronic Imaging*, vol. 2019, no. 13, pp. 134–1, 2019.
- 366 [29] N. Miolane, F. Poitevin, Y.-T. Li, and S. Holmes, “Estimation of orientation and camera
367 parameters from cryo-electron microscopy images with variational autoencoders and generative
368 adversarial networks,” 2019.
- 369 [30] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444,
370 2015.
- 371 [31] R. R. Coifman, Y. Shkolnisky, F. J. Sigworth, and A. Singer, “Graph laplacian tomography from
372 unknown random projections,” *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp.
373 1891–1899, 2008.
- 374 [32] C. Sorzano, R. Marabini, J. Vargas, J. Otón, J. Cuenca-Alba, A. Quintana, J. de la Rosa-Trevín,
375 and J. Carazo, “Interchanging geometry conventions in 3dem: mathematical context for the
376 development of standards,” in *Computational Methods for Three-Dimensional Microscopy*
377 *Reconstruction*. New York, NY: Springer New York, 2014, pp. 7–42. [Online]. Available:
378 https://doi.org/10.1007/978-1-4614-9521-5_2
- 379 [33] D. Q. Huynh, “Metrics for 3D rotations: Comparison and analysis,” *Journal of Mathematical*
380 *Imaging and Vision*, vol. 35, no. 2, pp. 155–164, 2009.
- 381 [34] S. Chopra, R. Hadsell, and Y. LeCun, “Learning a similarity metric discriminatively, with
382 application to face verification,” in *2005 IEEE Computer Society Conference on Computer*
383 *Vision and Pattern Recognition (CVPR’05)*, vol. 1. IEEE, 2005, pp. 539–546.
- 384 [35] D. Yi, Z. Lei, S. Liao, and S. Z. Li, “Deep metric learning for person re-identification,” in *2014*
385 *22nd International Conference on Pattern Recognition*. IEEE, 2014, pp. 34–39.
- 386 [36] M. A. Cox and T. F. Cox, “Multidimensional scaling,” in *Handbook of data visualization*.
387 Springer, 2008, pp. 315–347.
- 388 [37] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, “A global geometric framework for nonlinear
389 dimensionality reduction,” *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000. [Online].
390 Available: <https://science.sciencemag.org/content/290/5500/2319>
- 391 [38] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,”
392 *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- 393 [39] M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data represen-
394 tation,” *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

- 395 [40] L. Van der Maaten and G. Hinton, “Visualizing data using t-sne,” *Journal of machine learning*
396 *research*, vol. 9, no. 11, 2008.
- 397 [41] L. McInnes, J. Healy, and J. Melville, “Umap: Uniform manifold approximation and projection
398 for dimension reduction,” *arXiv preprint arXiv:1802.03426*, 2018.
- 399 [42] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, “Euclidean distance matrices: essential
400 theory, algorithms, and applications,” *IEEE Signal Processing Magazine*, vol. 32, no. 6, pp.
401 12–30, 2015.
- 402 [43] A. Bartesaghi, A. Merk, S. Banerjee, D. Matthies, X. Wu, J. L. Milne, and S. Subramaniam, “2.2
403 Å resolution cryo-EM structure of β -galactosidase in complex with a cell-permeant inhibitor,”
404 *Science*, vol. 348, no. 6239, pp. 1147–1151, 2015.
- 405 [44] G. Laxmikanthan, C. Xu, A. F. Brilot, D. Warren, L. Steele, N. Seah, W. Tong, N. Grigorieff,
406 A. Landy, and G. D. Van Duyne, “Structure of a holliday junction complex reveals mechanisms
407 governing a highly regulated dna transaction,” *Elife*, vol. 5, p. e14313, 2016.
- 408 [45] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and
409 T. E. Ferrin, “Ucsf chimera—a visualization system for exploratory research and analysis,”
410 *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.
- 411 [46] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and
412 J. Sijbers, “The ASTRA toolbox: A platform for advanced algorithm development in electron
413 tomography,” *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.
- 414 [47] C. Sorzano, L. De La Fraga, R. Clackdoyle, and J. Carazo, “Normalizing projection images:
415 A study of image normalizing procedures for single particle three-dimensional electron mi-
416 croscopy,” *Ultramicroscopy*, vol. 101, no. 2-4, pp. 129–138, 2004.
- 417 [48] H. Shigematsu and F. Sigworth, “Noise models and cryo-EM drift correction with a direct-
418 electron camera,” *Ultramicroscopy*, vol. 131, pp. 61–69, 2013.
- 419 [49] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for
420 modern deep learning research,” *Proceedings of the AAAI Conference on Artificial*
421 *Intelligence*, vol. 34, no. 09, pp. 13 693–13 696, Apr. 2020. [Online]. Available:
422 <https://ojs.aaai.org/index.php/AAAI/article/view/7123>
- 423 [50] T. Tieleman and G. Hinton, “Lecture 6.5-rmsprop: Divide the gradient by a running average of
424 its recent magnitude,” *COURSERA: Neural networks for machine learning*, vol. 4, no. 2, pp.
425 26–31, 2012.
- 426 [51] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint*
427 *arXiv:1412.6980*, 2014.
- 428 [52] H. B. McMahan, D. Golovin, S. Chikkerur, D. Liu, M. Wattenberg, A. M. Hrafinkelsson,
429 T. Boulos, J. Kubica, G. Holt, D. Sculley, M. Young, D. Ebner, J. Grady, L. Nie,
430 T. Phillips, and E. Davydov, “Ad click prediction: a view from the trenches,” in
431 *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery*
432 *and data mining - KDD '13*. ACM Press, 2013, p. 1222. [Online]. Available:
433 <http://dl.acm.org/citation.cfm?doid=2487575.2488200>

434 Checklist

- 435 1. For all authors...
- 436 (a) Do the main claims made in the abstract and introduction accurately reflect the paper’s
437 contributions and scope? [Yes]
- 438 (b) Did you describe the limitations of your work? [Yes] See §4.
- 439 (c) Did you discuss any potential negative societal impacts of your work? [Yes] We didn’t
440 identify any potential risk for improving protein imaging. Moreover, our work only
441 addresses a small step in a huge pipeline. We however mentioned the environmental
442 impact of training our model (see §4).
- 443 (d) Have you read the ethics review guidelines and ensured that your paper conforms to
444 them? [Yes]
- 445 2. If you are including theoretical results...

- 446 (a) Did you state the full set of assumptions of all theoretical results? [N/A]
447 (b) Did you include complete proofs of all theoretical results? [N/A]
- 448 3. If you ran experiments...
- 449 (a) Did you include the code, data, and instructions needed to reproduce the main exper-
450 imental results (either in the supplemental material or as a URL)? [Yes] We include
451 a URL in the Abstract to a git repository that includes code, data, and instructions to
452 reproduce our results. Moreover, notebooks and an interactive website are provided to
453 further play with the method.
- 454 (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they
455 were chosen)? [Yes] See §3.1 (including Table 1) for the preparation of data and how
456 they were split. See Appendix C for the hyperparameters.
- 457 (c) Did you report error bars (e.g., with respect to the random seed after running ex-
458 periments multiple times)? [Yes] When there was variance, e.g., on Figure 7 and
459 Figure 17b.
- 460 (d) Did you include the total amount of compute and the type of resources used (e.g., type
461 of GPUs, internal cluster, or cloud provider)? [Yes] See Appendix C.
- 462 4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
- 463 (a) If your work uses existing assets, did you cite the creators? [Yes] We used proteins
464 from the publicly available Protein Data Bank (PDB) and cited the ones we used, see
465 §3.1. We also used and cited the ASTRA toolbox in the same section.
- 466 (b) Did you mention the license of the assets? [Yes] The license of our code is mentioned
467 in the README.md and included in a LICENSE.txt file in our git repository. PDB
468 data are free of all copyright restrictions and made fully and freely available for both
469 non-commercial and commercial use.
- 470 (c) Did you include any new assets either in the supplemental material or as a URL? [Yes]
471 As a URL in the Abstract.
- 472 (d) Did you discuss whether and how consent was obtained from people whose data you're
473 using/curating? [N/A] Our data are proteins.
- 474 (e) Did you discuss whether the data you are using/curating contains personally identifiable
475 information or offensive content? [N/A] Our data are proteins.
- 476 5. If you used crowdsourcing or conducted research with human subjects...
- 477 (a) Did you include the full text of instructions given to participants and screenshots, if
478 applicable? [N/A]
- 479 (b) Did you describe any potential participant risks, with links to Institutional Review
480 Board (IRB) approvals, if applicable? [N/A]
- 481 (c) Did you include the estimated hourly wage paid to participants and the total amount
482 spent on participant compensation? [N/A]