

CryoFlow: Multi-modal amortized pose inference for Cryo-EM using normalizing flow

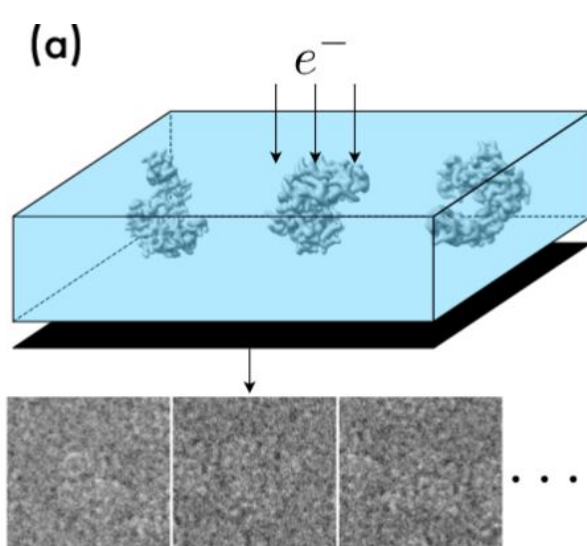
Andy Cai

Department of Computer Science, University of Toronto

Introduction

Background: Cryo-electron microscopy (Cryo-EM)

- Reconstruct 3D potential of a molecule from a series of very noisy 2D projections at random orientations
- Classical techniques used EM; fit orientations to volume estimate, then v.v
- Becomes too expensive for large datasets
- Modern amortized inference methods (e.g. CryoAI) treat problem like training an autoencoder, learn pose estimation (encoder) + volume estimation (decoder) simultaneously [1]
- However, sometimes unable to escape symmetric local minima

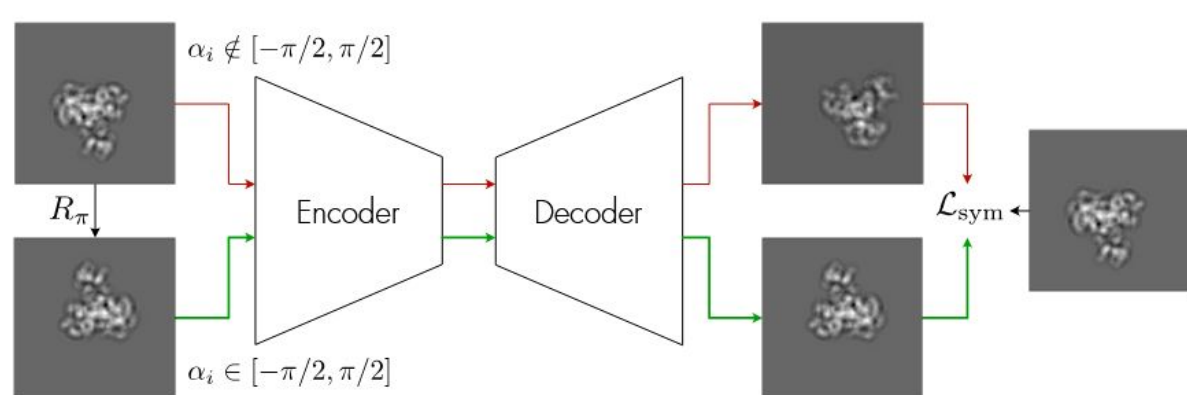


Motivation:

- Potential reason for local minima is that pose space is topologically complex (group of 3D rotations is not simply connected)
- Prior work by Esmaeili et al. [2] demonstrates this effect in SO(2), and proposes a solution by using a normalizing flow based encoder

Related Work

- CryoAI [1] tried to solve symmetric local minima using a 180° invariant loss; however, some structures were still unable to converge
- de Ruijter et al. [3] proposed using a group equivariant pose predictor (90° rotations and reflections), but still sometimes found lack of convergence



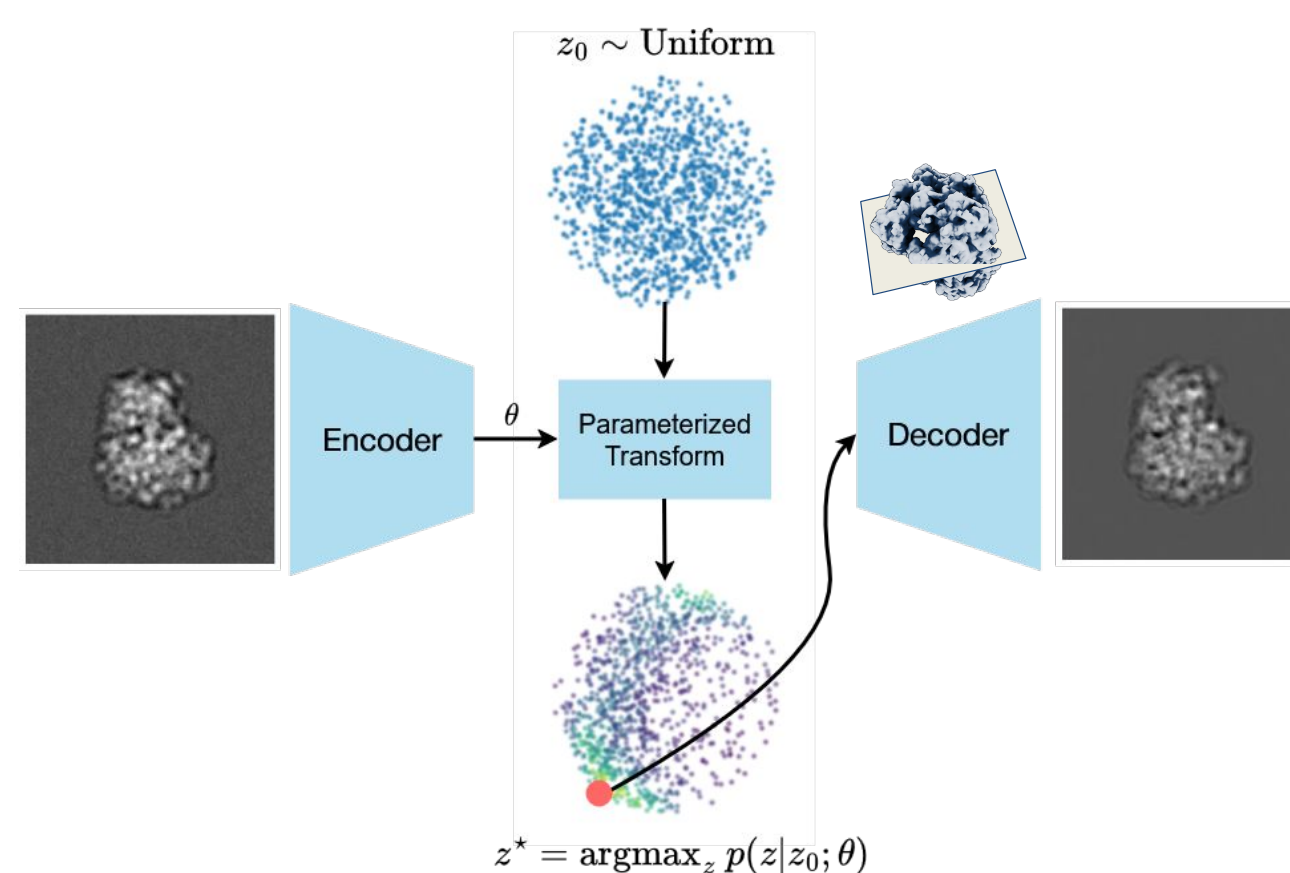
- CryoSPIN [4] utilizes multiple heads as an extension of the 2 heads approach of CryoAI, as well as using a semi-amortized approach; this mostly eliminates symmetric minima, but have to decide when to switch to auto-decoding

References

- [1] B. Esmaeili, R. Walters, H. Zimmermann, and J.-W. van de Meent, "Topological Obstructions and How to Avoid Them," NeurIPS, 2023
- [2] A. Levy et al., "CryoAI: Amortized Inference of Poses for Ab Initio Reconstruction of 3D Molecular Volumes from Real Cryo-EM Images," ECCV, 2022
- [3] L. de Ruijter and G. Cesa, "Equivariant amortized inference of poses for cryo-EM," ICLR, 2024
- [4] S. Shekarforoush, D. B. Lindell, M. A. Brubaker, and D. J. Fleet, "CryoSPIN: Improving Ab-Initio Cryo-EM Reconstruction with Semi-Amortized Pose Inference", NeurIPS, 2024
- [5] J. Köhler, M. Invernizzi, P. de Haan, and F. Noé, "Rigid Body Flows for Sampling Molecular Crystal Structures", ICML, 2023.

New Technique

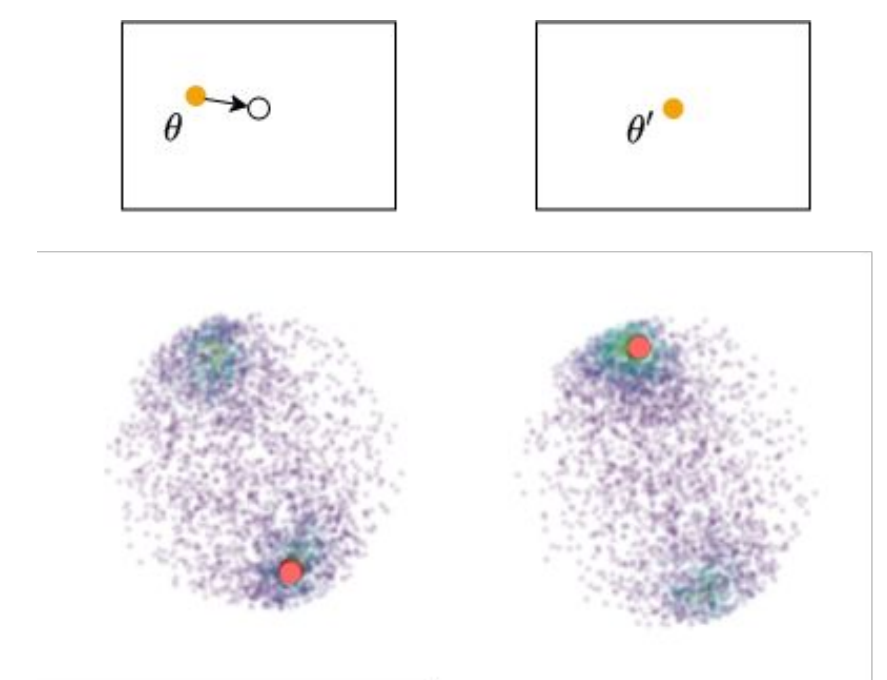
- Instead of predicting a single orientation, the pose encoder produces a latent vector that parametrizes a distribution over the space of orientations;
- Does this by applying a parameterized invertible transformation to the uniform distribution (normalizing flow)



Parameterized Transform
proposed by Köhler et al. [5]:

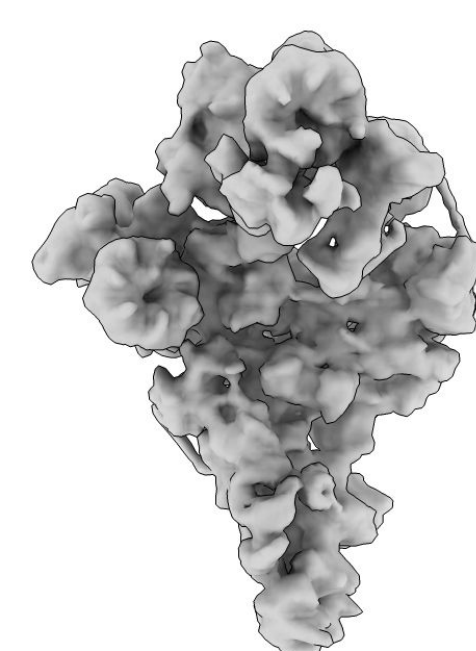
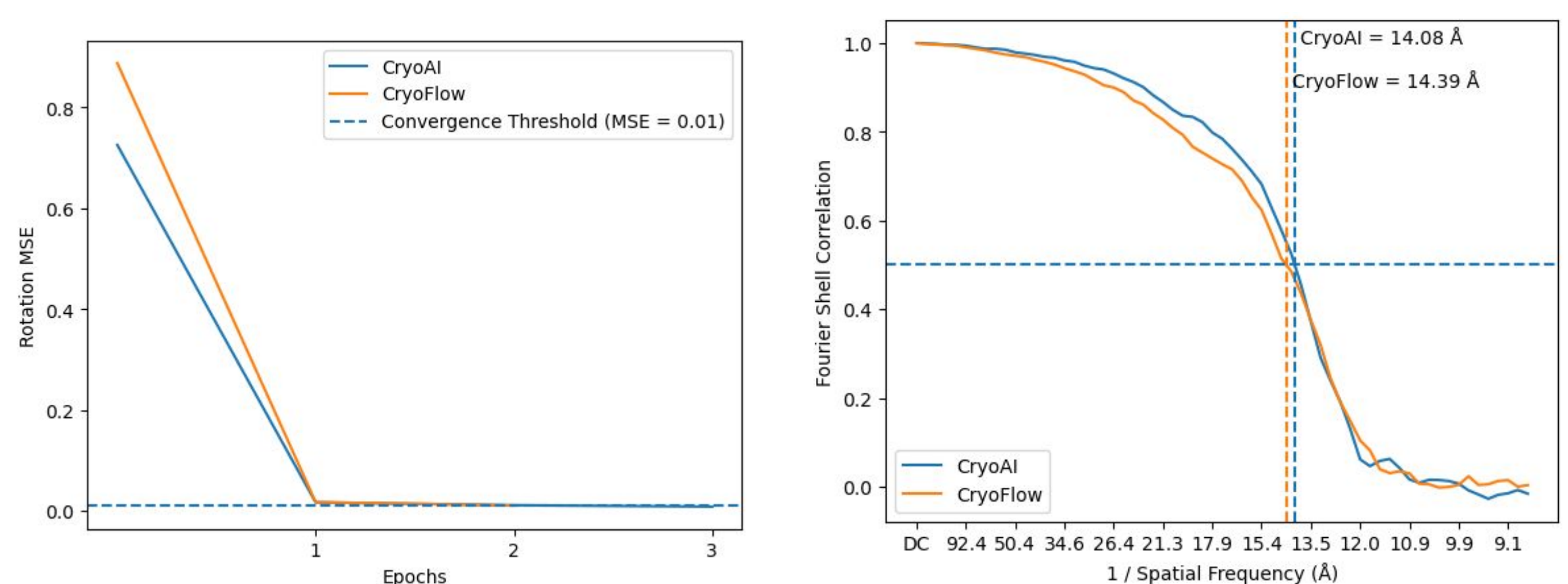
$$\Phi_{CG}(\mathbf{x}; \theta) = \frac{\nabla_{\mathbf{x}} \phi(\mathbf{x}; \theta)}{\|\nabla_{\mathbf{x}} \phi(\mathbf{x}; \theta)\|^2}$$
$$\phi(\mathbf{x}; \mathbf{W}, \mathbf{u}, \mathbf{b}, c) = \mathbf{u}^T (\log(\mathbf{b}) + \cosh(\mathbf{W}\mathbf{x})) + c\mathbf{x}^T \mathbf{x}$$

- Mode of distribution is predicted orientation
- For latent vectors which produce multimodal distribution, small changes can result in discontinuous jumps in orientation space
- If latent vector produces unimodal distribution, more stable to change
- Could serve as a mechanism for "automatic" switching to auto-decoding

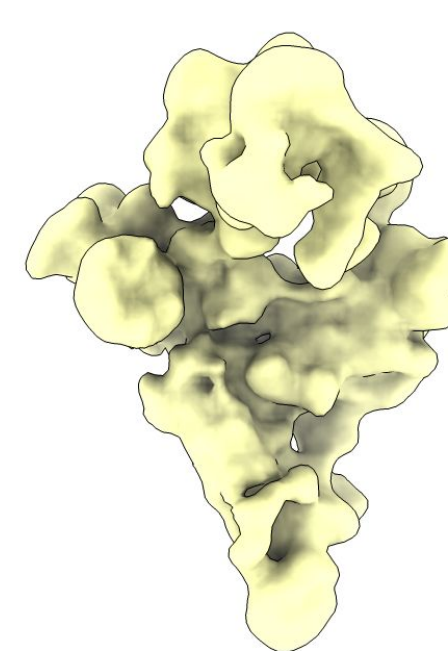


Experimental Results

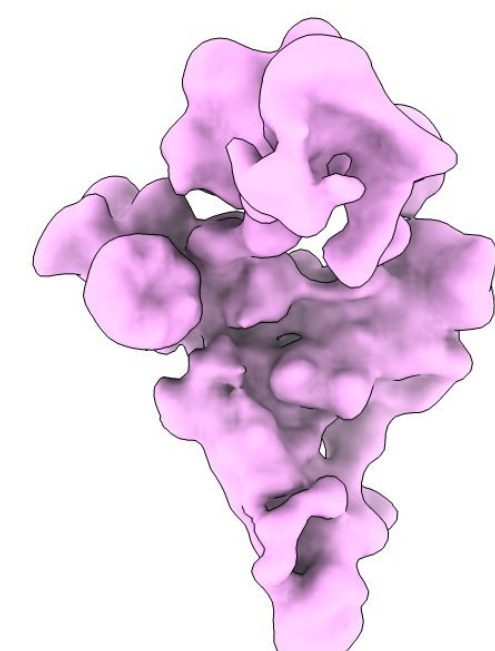
- Models trained on low noise (SNR = 10) synthetic dataset generated using structure of pre-catalytic spliceosome (PDB 5NRL)
- Models trained until MSE of aligned orientation matrices to ground truth was below 0.01



Ground Truth



CryoAI



CryoFlow

- For near noiseless dataset, CryoFlow performs the same as CryoAI, converging slightly faster but producing a slightly worse quality structure at convergence
- Convergence was always observed for both models
- Should extend tests to noisier datasets in future