

# Revealing biomolecular structure and motion with neural ab initio cryo-EM reconstruction

Junwen Liao

Qiuqin College, Tsinghua university

Sept 26th, 2024

# Contents

Introduction

DRGN-AI

Results

# Background

## Forward Model

$$I_i = C_i * P_{\phi_i} V_i + \eta_i, \quad (1)$$

where  $I_i$  is the measured data,  $C_i$  denotes the point spread function,  $P$  represents the projection under rotation  $R$  and translation  $t$  ( $\phi_i = (R, t)$ ),  $V_i$  is the object to be reconstructed and  $\eta_i$  is the noise.

# Reconstruction neural network

## Cryo-DRGN

- ▶ Architecture: VAE
- ▶ Input:  $I_i, C_i, \phi_i$
- ▶ Output:  $z_i \rightarrow V_i$
- ▶ Loss: just like standard VAE (reconstruction + KL)

## 3DFlex

- ▶ Architecture: Auto-decoder
- ▶ Input:  $I_i, C_i, \phi_i$
- ▶ Output:  $V, z_i \rightarrow f_i, V_i = f_i(V)$
- ▶ Loss: reconstruction term and non-rigidity penalty.

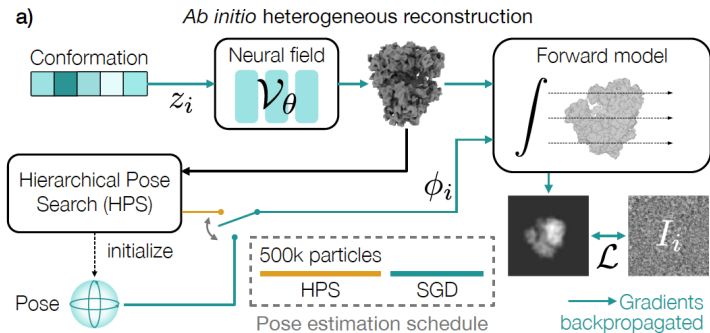
# Contents

Introduction

**DRGN-AI**

Results

# Architecture overview

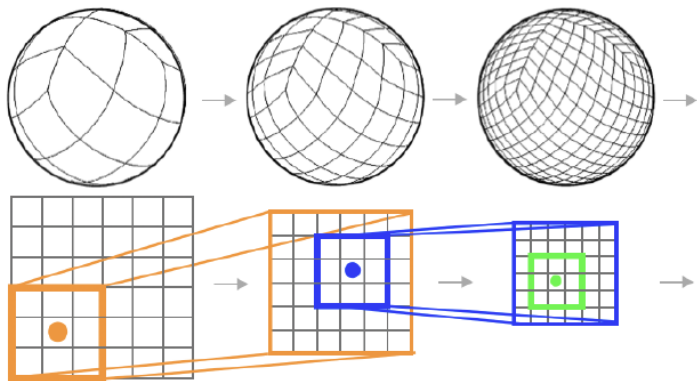


DRGN-AI Architecture [1]

$$I_i = C_i * P_{\phi_i} \mathcal{V}_\theta(z_i) + \eta_i$$

# Grid search

## (b) 5-D Pose search



HPS [2]

## DRGN-AI

- ▶ Architecture: Auto-decoder
- ▶ Input:  $I_i, C_i$
- ▶ Output:  $\phi_i, z_i \rightarrow V_i$
- ▶ Loss: reconstruction term only.



# Contents

Introduction

DRGN-AI

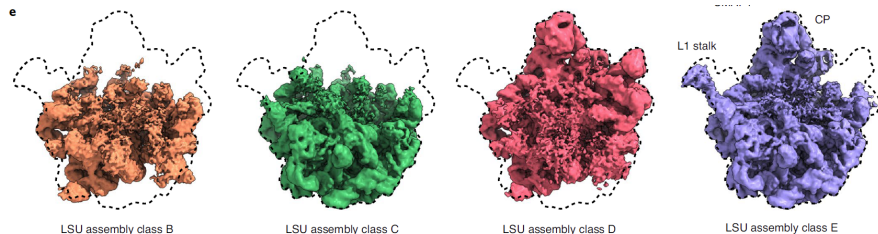
Results

# Pose estimate

<i>Mean Error</i>	cryoSPARC	cryoSPARC (refined)	DRGN-AI (HPS only)	DRGN-AI
Out-of-plane (deg.)	3.81	<u>0.64</u>	0.92	<b>0.62</b>
In-plane (deg.)	3.67	<u>0.57</u>	0.81	<b>0.56</b>
Translation (pix.)	0.32	<b>0.13</b>	0.17	<u>0.14</u>
Resolution (pix.)	2.4	<b>2.1</b>	2.4	<u>2.2</u>

mean error [1]

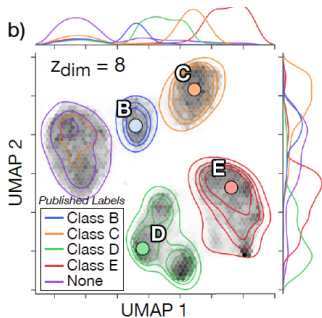
# Clustering for EMPIAR-10076



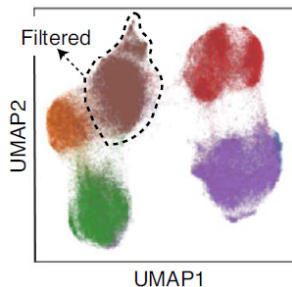
LSU dataset [3]

Four major states and one unassigned state for this dataset.

# Clustering



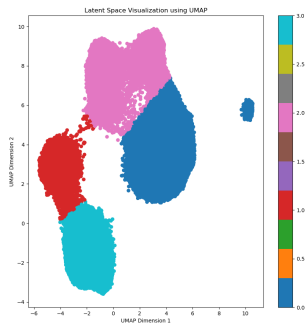
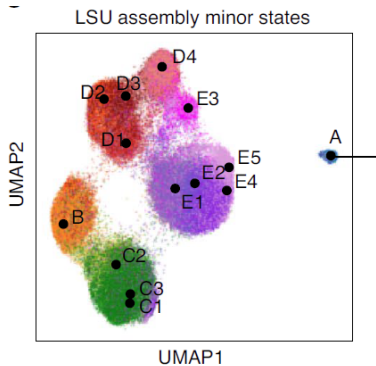
DRGN-AI [1]



CryoDRGN [3]

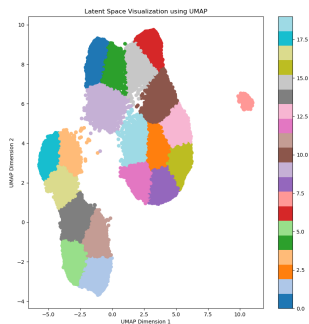
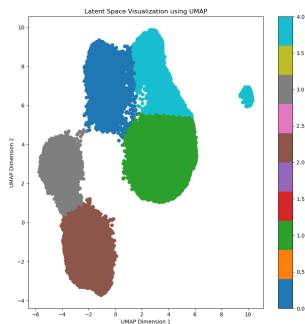
- ▶ CryoDRGN's latent variables are closer due to its KL loss term.
- ▶ The left one plot the KDE with 6 k-means clusters.

# Clustering



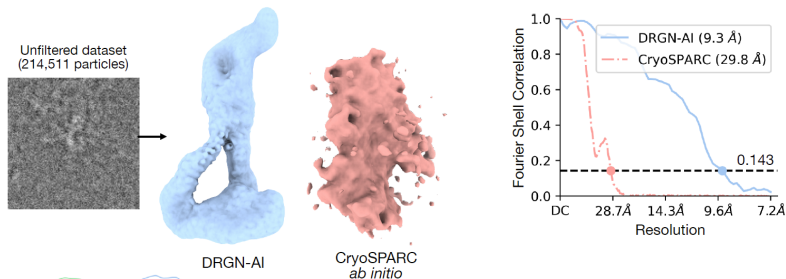
Real labels and  $K=4$

# Clustering



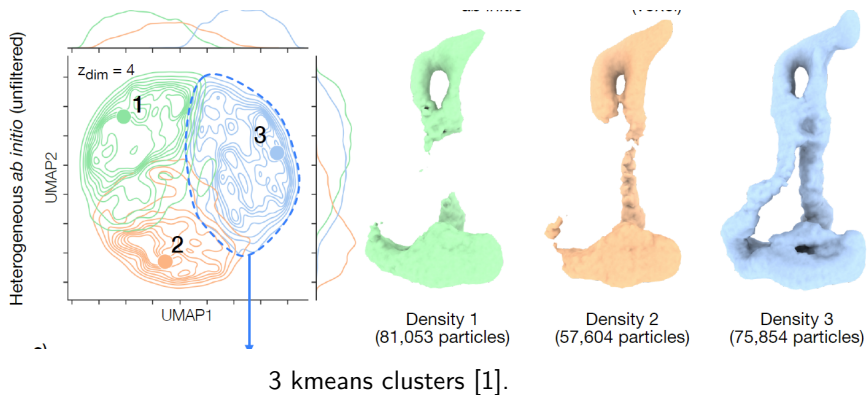
K=5 and K=20

# Ab initio



Ab initio with unfiltered datasets using DRGN-AI and CryoSPARC [1].

# Ab initio



Where are the junk images going?



# References

- [1]. Levy A, Grzadkowski M, Poitevin F, et al. Revealing biomolecular structure and motion with neural ab initio cryo-EM reconstruction[J]. bioRxiv, 2024: 2024.05. 30.596729.
- [2]. Zhong E D, Lerer A, Davis J H, et al. Cryodrgn2: Ab initio neural reconstruction of 3d protein structures from real cryo-em images[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 4066-4075.
- [3]. Zhong E D, Bepler T, Berger B, et al. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks[J]. Nature methods, 2021, 18(2): 176-185.