MIST: PHYSICS-CONSTRAINED SOURCE–TRANSPORT DECOMPOSITION FOR CRYO-EM HETEROGENEITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning interpretable latent structure from indirect observations is difficult when multiple mechanisms can explain the same measurements. Cryo-electron microscopy (cryo-EM) exhibits this ambiguity because an apparent structural change may arise from either motion of existing molecular density or local density gain and loss. Reconstruction error alone cannot determine which explanation is more appropriate. We introduce MIST (Mechanism-Informed Source–Transport), a Gaussian reconstruction model that structures the representation and training of the two mechanisms without requiring an accurate precomputed consensus reconstruction. Motion transports existing density with approximately mass-preserving deformation, while Source models local density changes and activates dormant anchors when variation extends beyond the canonical support. During joint training, a residual guided joint teacher evaluates the local image responses of Motion and Source together, including their overlap, and proposes coordinated coefficient updates that are accepted only when they improve the training objective. After a reliable reconstruction has been established, a shared-density auxiliary objective encourages variation explainable by moving common density to be represented by Motion, while preserving the full reconstruction and the capacity of Source. Across synthetic and experimental cryo-EM datasets, MIST substantially reduces cross-mechanism leakage while maintaining high-fidelity heterogeneous reconstruction and producing clearer, more coherent structural details. These results demonstrate that mechanism-aware source–transport modeling improves mechanism attribution without compromising reconstruction fidelity.

1 INTRODUCTION

Cryo electron microscopy reveals structural variation from noisy particle images. This variation may reflect motion of existing material or changes in local composition. The two mechanisms can produce similar observations. Moving a component may resemble density loss at one location and density gain at another, while an absent component may be approximated by moving its density away. Accurate reconstruction alone therefore does not determine how variation should be attributed.

Existing methods provide expressive models of heterogeneous structures, including neural density fields, deformation models, and Gaussian representations. However, assigning motion and density variation to separate parameters does not fully resolve their ambiguity. Density changes can approximate relocation, and flexible deformation can absorb occupancy changes. The branches may also compete during training because both reduce the same reconstruction residual.

We introduce MIST, a Gaussian model that gives the two mechanisms different transformation capabilities. Motion deforms existing density while approximately preserving mass. Source changes local density and can activate dormant anchors when particle dependent structure appears outside the canonical support. This asymmetry encourages transport and density change to be represented by the corresponding mechanisms.

MIST also addresses their interaction during optimization. A residual guided joint teacher estimates how small changes in Motion and Source affect the rendered image and proposes a coordinated update in their coefficient space. The proposal is used only when evaluation through the original forward model improves the training objective. After joint reconstruction, a shared density auxiliary objective provides an additional signal to Motion for variation explainable by transporting

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

common density. Full reconstruction remains active, preserving the ability of Source to represent compositional variation.

We evaluate reconstruction quality and mechanism behavior on synthetic and experimental datasets with conformational and compositional heterogeneity. Our contributions are threefold.

- We introduce a Gaussian representation that combines approximately mass preserving Motion with local Source density variation and dormant anchors.
- We introduce a residual guided joint teacher that accounts for the local image responses of both branches during joint training.
- We introduce a shared density auxiliary objective that encourages Motion to explain variation in common density while retaining full reconstruction.

2 Related Work

Heterogeneous cryo-EM reconstruction aims to recover structural variability from large collections of noisy particle projections Frank & Ourmazd (2016); Katsevich et al. (2015); Moscovich et al. (2020). Early approaches such as 3D variability analysis Punjani & Fleet (2021) characterize heterogeneity through low-dimensional modes around a consensus reconstruction. RECOVAR Gilles & Singer (2025) develops a regularized covariance-based formulation and combines it with kernel regression to reconstruct heterogeneous states from a learned low-dimensional representation. Neural methods substantially increase the flexibility of the structural model Zhong et al. (2021b); Levy et al. (2022). CryoDRGN Zhong et al. (2021a) introduces a variational latent representation together with a neural implicit decoder, enabling continuous distributions of three-dimensional density to be learned directly from particle images. OPUS-DSD Luo et al. (2023) further incorporates structural modeling into the latent representation to recover complex conformational landscapes, while Het-SIREN Herreros et al. (2025) studies real-space neural representations together with heterogeneous reconstruction and pose estimation. CryoBench Jeon et al. (2024) provides standardized synthetic datasets, including both conformational and compositional variability, and establishes common evaluation protocols for comparing heterogeneous reconstruction methods.

A complementary direction introduces stronger structural assumptions on molecular motion. 3DFlex Punjani & Fleet (2023) models heterogeneous structures as deformations of a canonical density and exploits the assumption that molecular motion transports density while largely preserving local geometry. DynaMight Schwab et al. (2024) represents a molecule with Gaussian pseudo-atoms and learns particle-specific deformation fields, allowing continuous motions to be analyzed and subsequently used for reconstruction. CryoSTAR Li et al. (2024) incorporates structural information derived from atomic models to regularize heterogeneous reconstruction, while cryoSPHERE Ducrocq et al. (2025) learns molecular segmentation and approximately rigid motions from a nominal atomic structure. These methods demonstrate the benefit of replacing unrestricted density variation with representations that reflect molecular structure and motion.

Explicit Gaussian representations provide another important foundation for recent cryo-EM reconstruction methods. e2gmm Chen & Ludtke (2021) represents molecular density using Gaussian mixtures and describes continuous structural variability through particle-dependent changes of Gaussian parameters. More recent approaches build on differentiable Gaussian rendering Kerbl et al. (2023). GEM Qu et al. (2025) represents cryo-EM density with anisotropic Gaussians and uses an efficient real-space splatting formulation for reconstruction, while CryoSplat Chen & Ling (2026) develops a Gaussian splatting formulation adapted to the cryo-EM image formation process. These methods mainly establish Gaussians as compact and differentiable representations for molecular density reconstruction, providing a useful basis for extending explicit representations to heterogeneous structures.

Several recent methods use Gaussian representations specifically to model both conformational and compositional heterogeneity. CryoSPIRE Shekarforoush et al. (2025) introduces a hierarchical Gaussian mixture representation together with molecular part discovery, allowing groups of Gaussians with correlated motion or occupancy changes to be organized into higher-level structural components. CryoDCOV Huang & Liu (2026) explicitly separates conformational and compositional variables by associating Gaussian center displacement with structural deformation and Gaussian amplitude variation with occupancy change, and optimizes the two representations with separate latent

variables and alternating updates. GaussianEM He et al. (2025) also models heterogeneous structures through particle-dependent Gaussian geometry and density, while imposing local structural constraints to obtain coherent conformational trajectories. These works show that explicit Gaussian parameters provide a natural basis for representing mixed heterogeneity. MIST builds on this line of work but distinguishes physical mechanisms rather than only assigning different phenomena to different parameters or structural groups. Motion is constrained as approximately mass-preserving transport in which Gaussian position, covariance, and density change are coupled by the same deformation, while Source represents non-conservative density variation and can introduce density outside the canonical support. MIST further considers how these mechanisms compete during joint reconstruction and refines their attribution after a reliable reconstruction has been obtained.

3 METHOD

3.1 PROBLEM SETTING

We observe noisy cryo-EM projections I with particle poses and contrast transfer functions (CTFs) supplied by preprocessing Scheres (2012); Punjani et al. (2017). A particle-specific structure consists of anisotropic Gaussians $G_i = (\mu_i, \Sigma_i, \rho_i)$ with center, covariance, and density weight. We omit particle indices except when averaging across particles. The image formation operator $\mathcal{P}$ gives

$$\hat{I} = \mathcal{P}(\mathcal{G}), \qquad \mathcal{L}_{\text{rec}} = \|I - \hat{I}\|_2^2. \tag{1}$$

Because transport and local density gain or loss can yield similar projections, reconstruction alone cannot assign mechanisms. MIST uses a Motion latent z_m for transport and a Source latent z_s for non-conservative density variation, without assuming unique identifiability.

3.2 MIST OVERVIEW

Figure 1 summarizes MIST. A shared Gaussian *Core* represents canonical density, while initially dormant *anchors* allow Source to add particle-dependent support beyond the Core. Motion deforms existing density; Source modulates Core density and activates anchors. Both branches jointly reconstruct images in Stage I, aided periodically by a residual-guided teacher. Stage II adds a shared-density objective that updates Motion while full reconstruction continues for both branches.

3.3 MECHANISM-ASYMMETRIC GAUSSIAN REPRESENTATION

Approximately mass-preserving Motion. Motion predicts $\phi(x) = x + u(x, z_m)$ with local Jacobian $F_i = \nabla\phi(\mu_i)$. The same deformation transforms Gaussian center, covariance, and density:

$$\boxed{\mu_i' = \phi(\mu_i), \qquad \Sigma_i' = F_i \Sigma_i F_i^\top, \qquad \rho_i' = \frac{\rho_i}{\max(|\det F_i|, \epsilon)}} \tag{2}$$

Determinant compensation links density to local volume change and biases Motion toward mass-preserving transport. We regularize local rigidity, orientation, volume change, and neighboring deformation to discourage folding and fragmented motion. Global invertibility is not required.

Source modulation and dormant anchors. For each Core Gaussian, Source changes canonical density ρ_i^0 through a particle-independent spatial gate g_i and particle-specific amplitude $r_i(z_s)$:

$$\rho_i' = \rho_i^0 \exp\big[\text{clip}\big(g_i r_i(z_s)\big)\big]. \tag{3}$$

Core modulation acts only where canonical Gaussians provide support. Initially near-zero-density anchors allow Source to represent density elsewhere; spatial sparsity and anchor birth regularization constrain their use. Thus Motion relocates existing density, whereas Source changes local density and can create new support.

3.4 RESIDUAL-GUIDED JOINT TEACHER

Both branches can reduce the same residual $r = I - \hat{I}$. During Stage I, a periodic teacher coordinates their updates in the compact coefficient space $c = [c_m, c_s]$. For sampled directions d_k, central

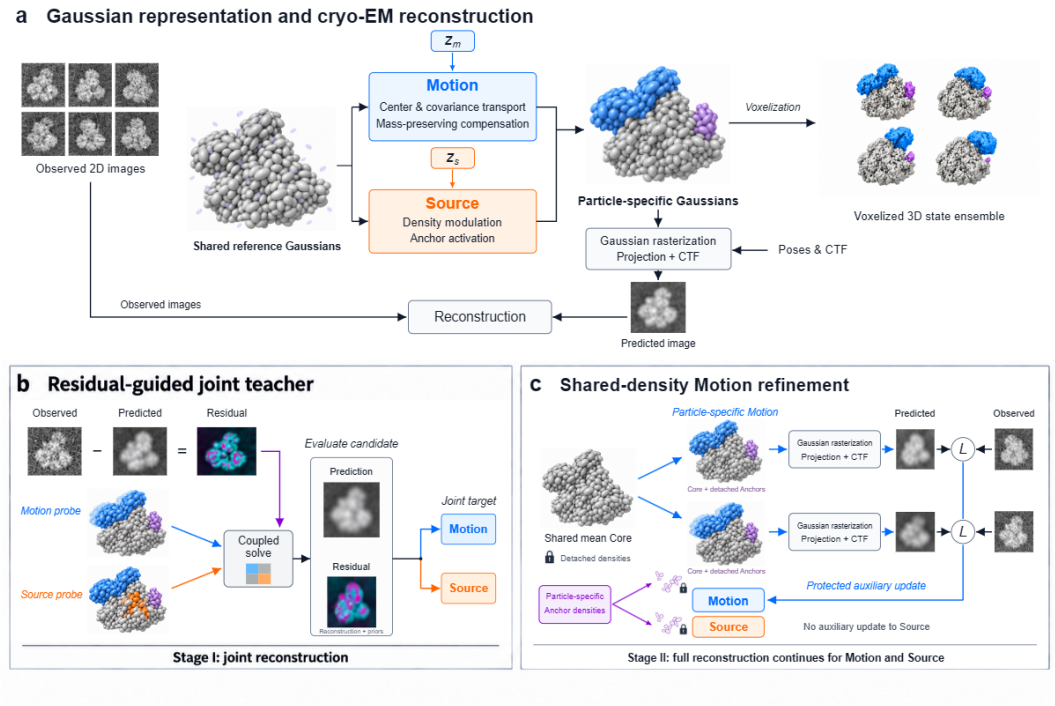

Figure 1: **Overview of MIST.** (a) Motion transports shared Gaussian density, whereas Source modulates Core density and activates dormant anchors; the resulting state is projected using particle pose and CTF. (b) A residual-guided joint teacher proposes coordinated branch updates during Stage I. (c) Shared-density refinement adds a protected Motion update during Stage II while full reconstruction continues for both branches.

differences estimate image responses $J_k \approx [\hat{I}(c+\varepsilon d_k) - \hat{I}(c-\varepsilon d_k)]/(2\varepsilon)$. Combining Motion and Source responses as $J = [J_m, J_s]$, the teacher proposes

$$\boxed{\Delta c = \arg\min_{\|\Delta c\| \le \tau} \|r - J\Delta c\|_2^2 + \lambda \|\Delta c\|_2^2} \tag{4}$$

where τ limits the step and λ damps it. The joint solve captures overlap through $J_m^\top J_s$. We render each proposed state with the original nonlinear model and retain it only if the training objective improves. Accepted coefficients serve as temporary targets for branch predictors and particle latents; ordinary joint reconstruction remains active.

3.5 Shared-Density Motion Refinement

Stage II encourages variation explainable by transport of common density to remain in Motion. We average reconstructed Core density across N particles, $\bar{\rho}_i = N^{-1} \sum_{n=1}^{N} \rho_{i,n}^{\text{core}}$, and pair it with each particle's Motion latent and retained anchor densities. Core and anchor densities are detached in this auxiliary path. Its prediction $\hat{I}_{\text{shared}}$ defines

$$\mathcal{L}_{\text{shared}} = \|I - \hat{I}_{\text{shared}}\|_2^2. \tag{5}$$

Only Motion receives this auxiliary gradient; full heterogeneous reconstruction still updates both branches. To protect it, let g_{rec} and g_{aux} be their respective Motion gradients. When they conflict, we project away the opposing auxiliary component Yu et al. (2020):

$$g_{\text{aux}} \leftarrow g_{\text{aux}} - \frac{g_{\text{aux}}^\top g_{\text{rec}}}{\|g_{\text{rec}}\|_2^2 + \epsilon} g_{\text{rec}}, \quad \text{if } g_{\text{aux}}^\top g_{\text{rec}} < 0. \tag{6}$$

We also ramp the auxiliary weight and cap its gradient relative to reconstruction. Source remains available for variation that shared-density transport cannot explain.

Table 1: Per-conformation FSC-AUC on CryoBench. Published baselines use the fixed-pose setting. Higher is better.

Method	IgG-1D ↑	Ribosembly ↑
CryoDRGN	0.366 ± 0.003	0.415 ± 0.019
CryoDRGN-AI-fixed	0.366 ± 0.001	0.372 ± 0.032
OPUS-DSD	0.340 ± 0.006	0.372 ± 0.046
3DFlex	0.336 ± 0.002	–
MIST	0.355 ± 0.006	$0.403 \pm 0.016^{\dagger}$

†Current MIST Ribosembly statistics use 100 matched reconstructions.

Table 2: Component ablation on IgG-1D and Ribosembly. Values are mean ± standard deviation.

	IgG-1D		Ribosembly	
Variant	Full ↑	Fixed ↑	Full ↑	Fixed ↑
MIST	0.3546 ± 0.0058	0.3520 ± 0.0056	0.4028 ± 0.0160	0.3821 ± 0.0184
w/o Auxiliary	0.3550 ± 0.0060	0.3485 ± 0.0053	0.4022 ± 0.0171	0.3731 ± 0.0234
w/o Teacher	0.3547 ± 0.0071	0.3522 ± 0.0059	0.4001 ± 0.0166	0.3795 ± 0.0202
w/o Anchors	0.3318 ± 0.0050	0.3351 ± 0.0049	0.3919 ± 0.0154	0.3737 ± 0.0153
w/o MassComp	0.3585 ± 0.0060	0.3517 ± 0.0057	0.4007 ± 0.0153	0.3953 ± 0.0171

4 Experiments

4.1 Experimental Setup

We evaluate MIST on IgG-1D and Ribosembly from CryoBench (Jeon et al., 2024), and on the experimental datasets EMPIAR-10076 (Davis et al., 2016) and EMPIAR-10028 (Wong et al., 2014). Since MIST uses particle poses and CTF parameters from preprocessing, we compare with the fixed-pose CryoBench settings of CryoDRGN, CryoDRGN-AI, OPUS-DSD, and 3DFlex.

For synthetic datasets, we use per-conformation FSC-AUC. For MIST ablations, *Full* denotes the complete Motion–Source reconstruction, while *Fixed-Motion* fixes Core density to the shared estimate and retains state-specific Motion.

To evaluate determinant compensation, we further measure the relative integrated-mass error

$$E_{\text{mass}} = \frac{\sum_g |M_{g,\text{after}} - M_{g,\text{before}}|}{\sum_g M_{g,\text{before}}}, \tag{7}$$

where lower values indicate better mass consistency. A same-geometry test evaluates an identical learned deformation with and without determinant compensation. UMAP (McInnes et al., 2018) is additionally used to visualize the effect of the joint teacher on latent organization.

4.2 Reconstruction and Ablation

Table 1 compares MIST with representative CryoBench methods. MIST achieves competitive reconstruction quality while introducing explicit constraints on how heterogeneity is represented. On IgG-1D, MIST reaches 0.355, above OPUS-DSD and 3DFlex and close to CryoDRGN and CryoDRGN-AI-fixed. On Ribosembly, MIST reaches 0.403, above CryoDRGN-AI-fixed and OPUS-DSD and close to CryoDRGN.

We next remove each main component individually. Table 2 shows that the components affect different properties of the model rather than simply changing overall FSC.

4.3 Mechanism Analysis

Mass-preserving Motion. Determinant compensation is evaluated directly through mass consistency. As shown in Table 3, MIST reduces the relative mass error to 0.034% on IgG-1D and 0.012%

Table 3: Mass-consistency evaluation. Lower is better.

	Actual error (%) ↓		Same geometry (%) ↓	
Dataset	MIST	w/o MassComp	Comp.	Uncomp.
IgG-1D	**0.034**	26.403	**0.034**	20.809
Ribosembly	**0.012**	33.036	**0.012**	33.453

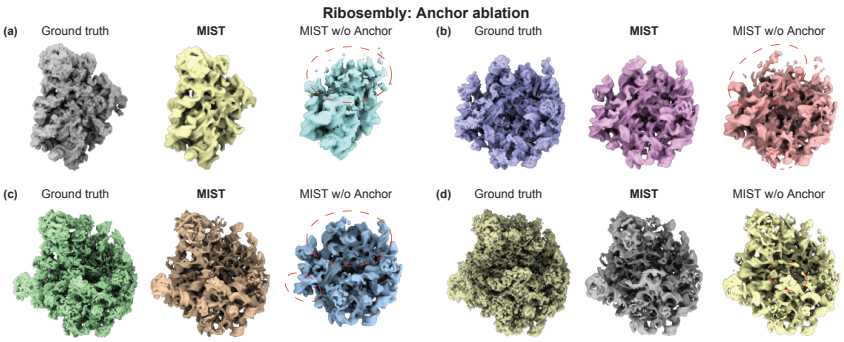

Figure 2: **Anchor ablation on Ribosembly.** Ground truth, MIST, and MIST without anchors are shown for four representative states. Dashed circles mark local structural differences.

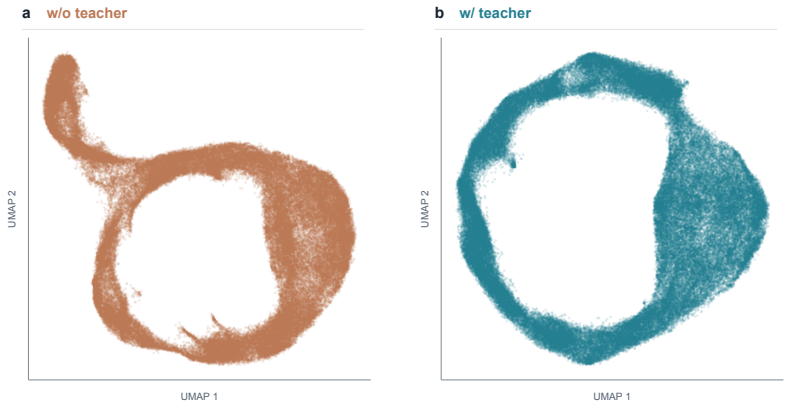

Figure 3: **Effect of the joint teacher.** UMAP visualization without the teacher (left) and with the teacher (right).

on Ribosembly, compared with 26.40% and 33.04% without compensation. The same-geometry test gives the same trend, directly confirming the effect of the compensation rule.

Dormant Source anchors. Removing anchors produces the largest consistent drop in Full FSC-AUC, from 0.3546 to 0.3318 on IgG-1D and from 0.4028 to 0.3919 on Ribosembly. Figure 2 shows corresponding local structural degradation, supporting the role of anchors in extending Source beyond the canonical Core support.

Residual-guided joint teacher. Removing the teacher produces only small changes in FSC, but changes the organization of the learned latent representation. Figure 3 shows that the complete model forms a more coherent continuous embedding, whereas the no-teacher model develops a pronounced additional branch. This is consistent with the teacher coordinating Motion and Source during joint optimization.

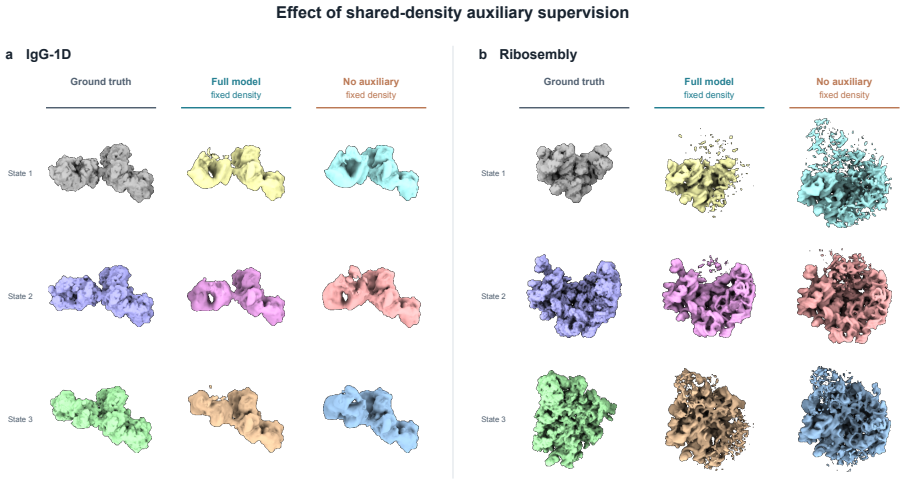

Figure 4: **Effect of shared-density refinement.** Ground truth, full-model fixed-density reconstruction, and no-auxiliary fixed-density reconstruction are shown for IgG-1D and Ribosembly.

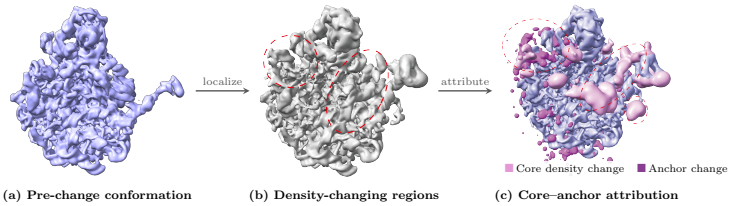

Figure 5: **Source interpretation on EMPIAR-10076.** (a) Pre-change conformation. (b) Localized density-changing regions. (c) Core–Anchor attribution of Source variation.

Shared-density Motion refinement. Removing the auxiliary objective leaves Full FSC-AUC nearly unchanged, but reduces Fixed-Motion FSC-AUC from 0.3520 to 0.3485 on IgG-1D and from 0.3821 to 0.3731 on Ribosembly. Figure 4 shows the same trend qualitatively, with the largest degradation appearing on Ribosembly. This indicates that the auxiliary objective improves the conformational information retained by Motion without changing the overall reconstruction capacity.

4.4 Experimental Cryo-EM Results

We finally examine whether the learned decomposition remains interpretable on experimental cryo-EM data.

EMPIAR-10076. Figure 5 shows a representative density-changing process. MIST localizes the changing regions and further attributes the recovered Source variation to Core density modulation and Anchor activation. This illustrates how Source represents density changes on existing support while using anchors when additional support is required.

EMPIAR-10028. EMPIAR-10028 contains both positional motion and density-changing variation. Figure 6 shows representative states together with their mechanism decomposition. The displaced peripheral region is captured by Motion, while Source variation concentrates on the chain-like density change. This example shows that both mechanisms remain active and spatially distinct on experimental data.

Together, the controlled and experimental results show that MIST maintains competitive reconstruction quality while producing a structured Motion–Source decomposition: Motion obeys the intended

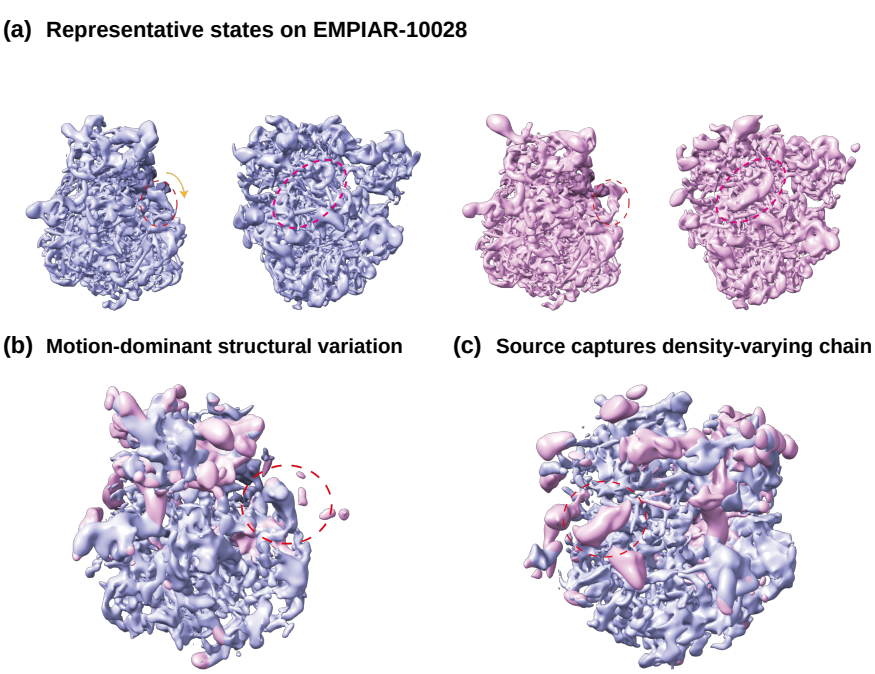

Figure 6: **Motion–Source decomposition on EMPIAR-10028.** (a) Representative heterogeneous states. (b) Motion-dominant structural variation. (c) Source variation localized on the density-varying chain.

mass-consistency constraint, anchors extend Source support, the teacher improves latent organization, and shared-density refinement strengthens Motion attribution.

5 Conclusion

We introduced MIST to distinguish motion of existing density from local density gain and loss in heterogeneous cryo-EM reconstruction. Its Gaussian representation couples deformation with approximate mass preservation while allowing Source to change density and activate dormant anchors. A residual-guided joint teacher coordinates the two branches during training, and a shared-density objective encourages Motion to explain variation in common density. On controlled datasets, MIST achieves competitive reconstruction quality; ablations support the intended roles of mass compensation, anchors, and Motion refinement. Results on two experimental datasets further show that the model recovers coherent heterogeneous structures. Because cryo-EM projections do not uniquely determine the underlying mechanism, MIST provides a physically informed attribution rather than a guarantee of unique separation.

AI use statement

We used generative AI tools to assist with writing and editing research code, drafting the conclusion, and checking the manuscript's LaTeX formatting. The authors take responsibility for the final code, results, and manuscript.

References

Muyuan Chen and Steven J. Ludtke. Deep learning-based mixed-dimensional gaussian mixture model for characterizing variability in cryo-EM. *Nature Methods*, 18(8):930–936, 2021. doi: 10.1038/s41592-021-01220-5. URL https://doi.org/10.1038/s41592-021-01220-5.

Suyi Chen and Haibin Ling. CryoSplat: Gaussian splatting for cryo-EM homogeneous reconstruction. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/5f173562e7662b14fb5c5695f225ea46-Abstract-Conference.html.

Joseph H. Davis, Yong Zi Tan, Bridget Carragher, Clinton S. Potter, Dmitry Lyumkis, and James R. Williamson. Modular assembly of the bacterial large ribosomal subunit. *Cell*, 167(6):1610–1622.e15, 2016. doi: 10.1016/j.cell.2016.11.020. URL https://doi.org/10.1016/j.cell.2016.11.020.

Gabriel Claude Jean Ducrocq, Lukas Grunewald, Sebastian Westenhoff, and Fredrik Lindsten. cryoSPHERE: Single-particle heterogeneous reconstruction from cryo EM. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/2ac2eac5098dba08208807b65c5851cc-Abstract-Conference.html.

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. doi: 10.1016/j.ymeth.2016.02.007. URL https://doi.org/10.1016/j.ymeth.2016.02.007.

Marc Aurèle Gilles and Amit Singer. Cryo-EM heterogeneity analysis using regularized covariance estimation and kernel regression. *Proceedings of the National Academy of Sciences*, 122(9):e2419140122, 2025. doi: 10.1073/pnas.2419140122. URL https://doi.org/10.1073/pnas.2419140122.

Bintao He, Yiran Cheng, Hongjia Li, Xiang Gao, Xin Gao, Fa Zhang, and Renmin Han. GaussianEM: Model compositional and conformational heterogeneity using 3D gaussians, 2025. URL https://arxiv.org/abs/2512.21599.

David Herreros, Carlos Perez Mata, Chari Noddings, Deli Irene, James Krieger, David A. Agard, Ming-Daw Tsai, Carlos Oscar Sanchez Sorzano, and Jose Maria Carazo. Real-space heterogeneous reconstruction, refinement, and disentanglement of CryoEM conformational states with HetSIREN. *Nature Communications*, 16:3751, 2025. doi: 10.1038/s41467-025-59135-0. URL https://doi.org/10.1038/s41467-025-59135-0.

Yue Huang and Manhua Liu. Disentangling conformational and compositional heterogeneity for cryo-EM reconstruction via decoupled gaussian mixture models. *Research Square*, 2026. doi: 10.21203/rs.3.rs-9494005/v1. URL https://doi.org/10.21203/rs.3.rs-9494005/v1. Preprint.

Minkyu Jeon, Rishwanth Raghu, Miro Astore, Geoffrey Woollard, Ryan Feathers, Alkin Kaz, Sonya M. Hanson, Pilar Cossio, and Ellen D. Zhong. CryoBench: Diverse and challenging datasets for the heterogeneity problem in cryo-EM. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2840. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/a2ef5ba272df8f168dc38037cc946be0-Abstract-Datasets_and_Benchmarks_Track.html.

Eugene Katsevich, Alexander Katsevich, and Amit Singer. Covariance matrix estimation for the cryo-EM heterogeneity problem. *SIAM Journal on Imaging Sciences*, 8(1):126–185, 2015. doi: 10.1137/130935434. URL https://doi.org/10.1137/130935434.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):139:1–139:14, 2023. doi: 10.1145/3592433. URL https://doi.org/10.1145/3592433.

Axel Levy, Gordon Wetzstein, Julien N. P. Martel, Frédéric Poitevin, and Ellen D. Zhong. Amortized inference for heterogeneous reconstruction in cryo-EM. In *Advances in Neural Information Processing Systems*, volume 35, 2022. doi: 10.52202/068431-0947. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/54b8b4e0b4ba4aad112e84f32e3b5dbb-Abstract-Conference.html.

Yilai Li, Yi Zhou, Jing Yuan, Fei Ye, and Quanquan Gu. CryoSTAR: Leveraging structural priors and constraints for cryo-EM heterogeneous reconstruction. *Nature Methods*, 21(12):2318–2326, 2024. doi: 10.1038/s41592-024-02486-1. URL https://doi.org/10.1038/s41592-024-02486-1.

Zhenwei Luo, Fengyun Ni, Qinghua Wang, and Jianpeng Ma. OPUS-DSD: Deep structural disentanglement for cryo-EM single-particle analysis. *Nature Methods*, 20(11):1729–1738, 2023. doi: 10.1038/s41592-023-02031-6. URL https://doi.org/10.1038/s41592-023-02031-6.

Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. UMAP: Uniform manifold approximation and projection. *Journal of Open Source Software*, 3(29):861, 2018. doi: 10.21105/joss.00861. URL https://doi.org/10.21105/joss.00861.

Amit Moscovich, Amit Halevi, Joakim Andén, and Amit Singer. Cryo-EM reconstruction of continuous heterogeneity by laplacian spectral volumes. *Inverse Problems*, 36(2):024003, 2020. doi: 10.1088/1361-6420/ab4f55. URL https://doi.org/10.1088/1361-6420/ab4f55.

Ali Punjani and David J. Fleet. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology*, 213(2):107702, 2021. doi: 10.1016/j.jsb.2021.107702. URL https://doi.org/10.1016/j.jsb.2021.107702.

Ali Punjani and David J. Fleet. 3DFlex: Determining structure and motion of flexible proteins from cryo-EM. *Nature Methods*, 20(6):860–870, 2023. doi: 10.1038/s41592-023-01853-8. URL https://doi.org/10.1038/s41592-023-01853-8.

Ali Punjani, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker. cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, 14(3):290–296, 2017. doi: 10.1038/nmeth.4169. URL https://doi.org/10.1038/nmeth.4169.

Huaizhi Qu, Xiao Wang, Gengwei Zhang, Jie Peng, and Tianlong Chen. GEM: 3D gaussian splatting for efficient and accurate cryo-EM reconstruction, 2025. URL https://arxiv.org/abs/2509.25075.

Sjors H. W. Scheres. RELION: Implementation of a bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, 180(3):519–530, 2012. doi: 10.1016/j.jsb.2012.09.006. URL https://doi.org/10.1016/j.jsb.2012.09.006.

Johannes Schwab, Dari Kimanius, Alister Burt, Tom Dendooven, and Sjors H. W. Scheres. DynaMight: Estimating molecular motions with improved reconstruction from cryo-EM images. *Nature Methods*, 21(10):1855–1862, 2024. doi: 10.1038/s41592-024-02377-5. URL https://doi.org/10.1038/s41592-024-02377-5.

Shayan Shekarforoush, David B. Lindell, Marcus A. Brubaker, and David J. Fleet. Reconstructing heterogeneous biomolecules via hierarchical gaussian mixtures and part discovery. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-2538. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/6d48a60086e5bfb5e3099f098bd4bf8e-Abstract-Conference.html.

Wilson Wong, Xiao-chen Bai, Alan Brown, Israel S. Fernandez, Eric Hanssen, Melanie Condron, Yan Hong Tan, Jake Baum, and Sjors H. W. Scheres. Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *eLife*, 3:e03080, 2014. doi: 10.7554/eLife.03080. URL https://doi.org/10.7554/eLife.03080.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 33, pp. 5824–5836, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/3fe78a8acf5fda99de95303940a2420c-Abstract.html.

Ellen D. Zhong, Tristan Bepler, Bonnie Berger, and Joseph H. Davis. CryoDRGN: Reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods*, 18(2):176–185, 2021a. doi: 10.1038/s41592-020-01049-4. URL https://doi.org/10.1038/s41592-020-01049-4.

Ellen D. Zhong, Adam Lerer, Joseph H. Davis, and Bonnie Berger. CryoDRGN2: Ab initio neural reconstruction of 3D protein structures from real cryo-EM images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4046–4055, 2021b. doi: 10.1109/ICCV48922.2021.00403. URL https://doi.org/10.1109/ICCV48922.2021.00403.