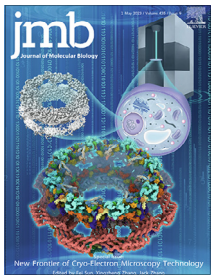
Review Article



Methods for Cryo-EM Single Particle Reconstruction of Macromolecules Having Continuous Heterogeneity

Bogdan Toader [1*], **Fred J. Sigworth** [2] **and Roy R. Lederman** [1]

*1 - **Department of Statistics and Data Science,** Yale University, United States*
*2 - **Department of Cellular and Molecular Physiology,** Yale University, United States*

Correspondence to Bogdan Toader: *bogdan.toader@yale.edu (B. Toader) @BogdanToader0 (B. Toader), @roylederman (R.R. Lederman)*
https://doi.org/10.1016/j.jmb.2023.168020
Edited by Jack Zhang

Abstract

Macromolecules change their shape (conformation) in the process of carrying out their functions. The imaging by cryo-electron microscopy of rapidly-frozen, individual copies of macromolecules (single particles) is a powerful and general approach to understanding the motions and energy landscapes of macromolecules. Widely-used computational methods already allow the recovery of a few distinct conformations from heterogeneous single-particle samples, but the treatment of complex forms of heterogeneity such as the continuum of possible transitory states and flexible regions remains largely an open problem. In recent years there has been a surge of new approaches for treating the more general problem of continuous heterogeneity. This paper surveys the current state of the art in this area.

Introduction

Over the past few years, the combination of cryo-electron microscopy (cryo-EM) imaging and single-particle analysis has been established as the method of choice for studying the structure of large protein complexes at atomic or near-atomic resolution.[1] Its recent success has been enabled by advances in detector technology, sample preparation techniques and the availability of advanced image processing software packages. The two other major structure-determination methods are X-ray crystallography, which requires the sample to be crystallized, and NMR, which is useful only with relatively small proteins.

Cryo-EM single-particle analysis (we will denote this simply as cryo-EM) involves the imaging of individual copies (called particles) of a macromolecular structure. Through the computational processing of 10^4 to 10^6 of such particle images, 3D density maps can be obtained through *single-particle reconstruction*. To density maps of sufficient resolution, atomic structures can then be fitted, with ~ 4 Å being the worst resolution for successful ab initio fitting. Because the macromolecules are suspended in solution before they are rapidly frozen, particle images are likely to reflect a more native conformation, but also may contain frozen instances of flexibility or conformational variation. Therefore, one of the promises of cryo-EM is that researchers will be able to construct a complete picture of all the possible conformations of the imaged structures.

Conformational changes are key to the function of many macromolecular machines. The molecular motors dynein and kinesin undergo cyclical changes as a chemical reaction (hydrolysis of ATP) drives a mechanical stepping motion that moves cargo along a microtubule filament. Glucose transporters allow cells to take up this nutrient through a conformational cycle that enforces the transport of one or two Na^+ ions with

each glucose molecule. DNA replication is carried out by the replisome, a large combination of molecular machines that, in stepwise fashions, unwind the double-stranded DNA and synthesize new complementary DNA strands.

While existing methods excel at reconstructing clearly-defined discrete conformations from cryo-EM data, the problem of reconstructing continuously-varying conformations of a macromolecule is where most state of the art methods fall short. This is the *continuous heterogeneity reconstruction* problem. Fortunately, there has been a surge in the effort devoted to developing computational tools for continuous heterogeneity reconstruction, and this is the focus of this survey article. Our aim is to highlight the defining characteristics of each method and the conceptual overlaps and differences between them.

We note that there are many ideas and approaches for continuous heterogeneity and many papers introduce multiple new ideas and combine multiple approaches. For brevity and clarity, we cluster together different works and omit important implementation details. This paper covers the conceptual families of ideas and does not make specific recommendations about software to use. Some of the work surveyed is theoretical or less accessible to the user. Where available, we included links to some of the software that may be more accessible.

Complementary surveys

There are a number of recent surveys covering other aspects of the cryo-EM pipeline.

A general review of the computational challenges and the main components in the analysis pipeline are available in.[2] A comprehensive description of the mathematical aspects of the problem, focusing on homogeneous reconstruction and validation, is available in.[3]

Discrete heterogeneity is discussed in.[4] A survey of earlier work on continuous heterogeneity, with a comprehensive survey of normal mode analysis (the section ''Normal Modes and Predefined Spaces of Conformations'') is available in.[5] In,[6] the focus is on the interpretation of the energy landscape resulting from the heterogeneity analysis using likelihood-based methods. An up-to-date overview of the full cryo-EM pipeline, from sample preparation to the latest reconstruction methods, including time-resolved cryo-EM is available in.[7]

The recent reviews[8,9] focus on machine learning approaches to cryo-EM. Specifically, the former gives an overview of machine learning algorithms used in each step in the cryo-EM pipe-line, from pre-processing and particle picking to 3D reconstruction and post-processing, while the latter is a thorough survey of deep generative modeling techniques for 3D reconstruction.

In this review, we summarize the state of the art methods for analyzing continuous heterogeneity in cryo-EM. Our aim is to sort the main families of ideas in the area and convey some of the main ideas of each technique.

Outline of the paper

We first discuss a simplified image formation model that most cryo-EM reconstruction methods assume, as well as discrete heterogeneous reconstruction and multi-body refinement, in the ''Preliminaries'' section.

In the ''Manifold Learning on Particle Images'' and ''Manifolds of Volumes'' sections, we describe manifold learning approaches to continuous heterogeneity reconstruction, specifically applied to particle images and reconstructed volumes respectively. In the ''Principal Volumes and Linear Models for Volumes'' and ''Normal Modes and Predefined Spaces of Conformations'' sections, we discuss linear models based on covariance estimation and normal mode analysis respectively.

In the ''Nonlinear Models: Hyper-Molecules'' section, we discuss nonlinear models for continuous heterogeneous reconstruction which conceptually fit into the ''hyper-molecules'' framework, including traditional and deep learning algorithms. In the ''Bypassing the Estimation of Latent Variables'' section, we discuss inference methods based on distribution matching.

Finally, we conclude the article with a discussion in the ''Discussion and Perspectives'' section.

Preliminaries

Image formation model and homogeneous reconstruction

In this section, we discuss a simplified model for image formation that is the basis of all the approaches in this survey. For simplicity, we begin this discussion with a model for the tomographic projections of a single conformation in the *homogeneous* case. While each specific method may be a variation of this model and contain additional details, this model provides a useful baseline.

Let $V(\mathbf{r})$ represent the electrostatic potential of the molecule of interest in a specific conformation at location $\mathbf{r} \in \mathbb{R}^3$. Throughout this survey, we will use the terms *volume* and *density* interchangeably to refer to V.

The *particle images* I_i ($i = 1, \ldots, M$) are given by the following linear model of the image formation process:

$$I_i = (C_i \circ T_i \circ P \circ R_i)V + \eta_i, \tag{1}$$

for $i = 1, \ldots, M$, where R_i is a 3D rotation operator corresponding to the orientation of the volume V, P is the 2D projection operator, T_i is the 2D translation operator corresponding to the offset of the projected volume with respect to the center of the image, C_i is the *contrast transfer function* (CTF) operator applied to

the projected image and $\circ$ denotes the composition of these operators. Both R_i and T_i are latent variables specific to each particle image I_i and are not known in advance. The CTF operator C_i can be different for each particle image, and in our simplified discussion we assume that it is known. For simplicity, we assume that the particle images are cropped in advance from the larger micrograph produced by the microscope.

Lastly, Gaussian noise η_i is applied to each image. In reality, the noise introduced in cryo-EM images is shot noise,[10] but the Gaussian assumption is commonly used in the reconstruction literature. While the details of the noise model and its estimation are specific to each reconstruction method, in our discussion we assume, for exposition purposes, that the noise parameters have been estimated in advance. For example, RELION computes the radial power spectrum of the noise during the homogeneous reconstruction.[11]

Remark 1 Cryo-EM algorithms make extensive use of the Fourier transform of images and density functions for efficient representation and computation. While the distinction between representations and operations in the Fourier domain and in the spatial domain is very important in implementations, it is not specific to the study of heterogeneity. For brevity, we omit the detailed discussion of Fourier vs. spatial domain implementations in different algorithms and focus on the main ideas that are specific to the study of heterogeneity.

In the case of homogeneous reconstruction, the usual approach is to solve the maximum-a posteriori (MAP) problem given the set of particle images $\{I_i\}$:

$$\operatorname{argmax}_V \ln p(V|\{I_i\}), \tag{2}$$

where the log-posterior distribution is given by:

$$\ln p(V|\{I_i\}) = \sum_{i=1}^{M} \ln \int_{\phi_i} p(I_i|V,\phi_i)p(\phi_i)\mathrm{d}\phi_i + \ln p(V). \tag{3}$$

Here, the likelihood function $p(I_i|V,\phi_i)$ is Gaussian and determined by the image formation model (1), and the rotation and the translation R_i, T_i of each particle image are paired into one pose variable ϕ_i whose assumed joint distribution is denoted by $p(\phi_i)$. From a Bayesian point of view, $p(\phi_i)$ and $p(V)$ play the role of priors (of the pose and the volume respectively), and in some methods (e.g., RELION), they are adjusted iteratively during reconstruction.

Discrete heterogeneity

The traditional approach to the heterogeneity problem is to perform 3D classification, where each particle is assigned to one of a small number of different, optimized reference volumes (in some algorithms, a probability of assignment to each references volume is considered). This works best when the underlying macromolecule has a small number of discrete states, and these are distinguishable in the set of single-particle images.

The main assumption here is that the reference volumes obtained are a good representation of the different conformations of the macromolecule, which requires the removal of the outliers before reconstruction and that each class contains enough particle images for the reconstruction to be possible. Usually the number of reference volumes, or classes, is specified in advance, but it can change during optimization if some classes are not very populated.[12]

First introduced in,[13] this method consists of maximum likelihood or maximum-a posteriori optimization for K different volumes, or K *classes*, $\{V_1(\mathbf{r}),\ldots,V_K(\mathbf{r})\}$. Then, the log posterior distribution (3) becomes:

$$\ln p(\{V_k\}|\{I_i\}) = \sum_{i=1}^{M} \ln\left(\sum_{k=1}^{K}\int_{\phi_i} p(I_i|k,\{V_k\},\phi_i)p(k,\phi_i|\{V_k\})\mathrm{d}\phi_i\right) + \sum_{k=1}^{K}\ln p(V_k). \tag{4}$$

The K volumes are usually initialized as low-resolution reconstructions from K randomly drawn subsets of the particle images, and the optimization can be performed, for example, with the expectation–maximization algorithm[13,11,14] or with stochastic gradient descent or its variants[A].[15,12]

Multi-body extension of traditional analysis

A step up in complexity is the multi-body refinement[B].[16–18] Here, the volume is assumed to consist of a small number of rigid components that move relative to each other and are identical across the particle images.

This method starts with applying standard homogeneous reconstruction techniques to obtain a consensus reconstruction of the bulk of the volume. Then, the user specifies a set of 3D masks that define regions of the volume containing individual components that move with respect to the bulk. Each of these components is treated like a rigid body. Lastly, independent homogeneous reconstruction is performed for each component using the particle images where, for each component reconstruction, the CTF-affected 2D projections of the other components have been subtracted from the particle images. The subtraction is done either before the separate components reconstruction step[16,17] or during the reconstruction process in an iterative fashion.[18] In the latter approach, the relative orientation of each

[A] For software, see for example the RELION https://relion.readthedocs.io or CryoSPARC https://cryosparc.com packages.

[B] For software that performs multi-body analysis, see RELION in footNote A.

component is refined for each particle image and updated at every iteration before subtraction, which leads to improved results.

One obvious disadvantage of this approach is that it relies on the user's previous knowledge of the structural domains of the molecule and, therefore, it is susceptible to human bias. In addition, multi-body refinement is limited to rigid variability and has difficulty at the interface between the moving components.

Conformation space

Given a set of particle images, the objective of traditional homogeneous reconstruction is to recover a single volume which, according to the forward model (1) (or a variation of it), generated the particle images. When there is conformational heterogeneity that a single homogeneous reconstruction model cannot capture, homogeneous reconstruction algorithms recover some average structure (or even worse, corrupted structures). Similarly, the main underlying assumption in discrete heterogeneous reconstruction is that the particles cluster around a small number of distinct dominating structures, which the reconstruction process aims to recover.

In both cases, an important step of the data preprocessing is the removal of outliers that do not fit with any of the structures that one wants to recover. Moreover, information from particles on the boundary between different "discrete" structures, which may represent rare or less likely conformations, is lost when they are assigned to one of the volumes or classified as an outlier.

The strength of cryo-EM for structure determination of biological samples is that it is able to capture the sample and its conformations without being constrained to a crystal lattice. Cryo-EM observes individual molecules rather than crystal ensemble averages. Therefore, avoiding averaging of these conformations through reconstruction of one or several discrete volumes is a key motivation for continuous heterogeneity reconstruction.

While the output of homogeneous or discrete heterogeneous reconstruction consists of a small number of discrete volumes, the output of continuous heterogeneity analysis should capture the full space of conformations of the molecule. This often takes the form of a *low-dimensional manifold*. The specifics vary significantly between different techniques. Throughout this survey, we will refer to this manifold as the *conformation space* or *latent space* representation. Given such a manifold and a coordinate vector on it, one should be able to recover a volume corresponding to the conformation at those coordinates. In Figure 1, we illustrate examples of volumes along a trajectory in the conformation space for the calmodulin protein.

Manifold Learning on Particle Images

One of the first ideas for characterizing continuous heterogeneity in cryo-EM was based on the observation that a continuous space of volumes could be represented mathematically as a manifold of volumes; similarly, the space of particle images, which are the projections of different molecules in different viewing directions, form a manifold of images.[19] Manifold learning techniques are usually based on organizing observations (particle images) based on a norm of the difference between every two observations. The study of cryo-EM data as a manifold of particle images presents several challenges. For example, in the simplest form of manifold learning, each variable such as the viewing direction, in-plane rotations, in-plane translations, and CTF, introduces additional dimensions to the problem, making the analysis impractical even before considering the heterogeneity. Furthermore, the technical properties of the difference between particle images lead to mathematical difficulty in identifying the observed manifold with the product $SO(3) \times \tau$ of viewing directions and heterogeneity space. Finally, the high level of noise in cryo-EM makes the difference between particle images noisy. A sequence of followup work introduced additional ideas to make the analysis practical[20–25] by assuming known viewing directions and analyzing the manifold of conformations as viewed from each direction using the diffusion map algorithm [26]; the maps are then aligned across all different directions either by using nonlinear Laplacian spectral analysis (NLSA)[24] or a novel algorithm ESPER[C,25].

Experiments performed using simulated data in[25] show that the conformation manifold is not constructed well for certain projection directions where the range of the conformation variability is small. The manifold embeddings are also affected negatively by insufficient samples for each state in the range and low SNR, effects that are equally visible for both PCA and Diffusion Maps. To circumvent these issues, ESPER[25] rotates the manifolds' coordinates (*eigenfunction realignment*) so that each dimension of the heterogeneity variable is captured in a separate eigenbasis of the embedding. Then, a subspace partitioning procedure is performed to assign particles to each state along the trajectory representing the continuous heterogeneity and produce a 2D movie for each viewing direction. The direction of each movie is determined using optical flow and belief propagation using the method in,[23] after which the individual states along the trajectory can be reconstructed using standard homogeneous reconstruction methods.

[C] For the ManifoldEM/ESPER software, see the code repositories: https://github.com/GMashayekhi/ManifoldEM_Matlab, https://github.com/evanseitz/ManifoldEM_Python, https://github.com/evanseitz/cryoEM_ESPER.

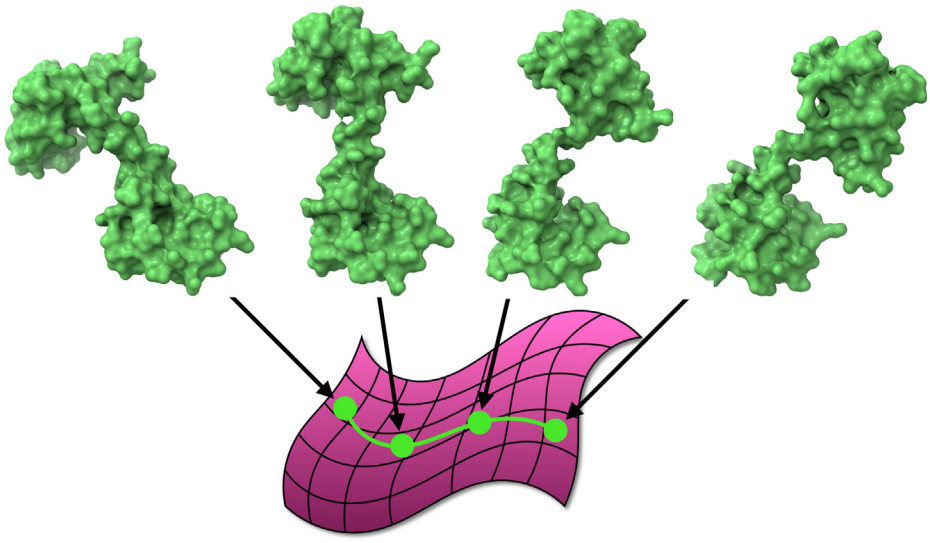


Figure 1. Examples of conformations of the calmodulin protein along a trajectory on its conformation manifold (illustration). Volumes obtained with UCSF ChimeraX version 1.4 (2022–06-03) using its *morph* function and the structures from the Protein Data Bank with entryIDs 1CFD and 3CLN.

The manifold approach in this section is characterized by sidestepping the modeling of the volumes themselves during the analysis. Instead, this approach aims to build a map of conformations for each viewing direction. Once the map is available, one selects particle images from many viewing directions of roughly the same conformation and reconstructs that conformation from the particle images using traditional reconstruction algorithms. In the following sections, we will discuss approaches that rely on explicit modeling of the volumes.

Manifolds of Volumes

Many of the challenges presented in the previous section can be avoided by applying manifold learning techniques directly to volumes, which is the approach we discuss in this section. Clearly, cryo-EM experiments produce tomographic projections, not volumes. Therefore, the approach in this section of not self-contained in most cases. One work on manifolds of volumes in[27] assumes that the volumes are given. A later method in[28] generates the volumes by performing 3D classification of reconstructed volumes from a number of bootstrapped subsets of the data.

Once the volumes are available, both methods propose a way of obtaining the low dimensional embedding corresponding to the manifold of conformations. In StructMap[D],[27] the volumes are aligned using rigid body alignment as well as flexible alignment using their normal mode representations (which will be discussed in the "Normal Modes and Predefined Spaces of Conformations" section), and define a dissimilarity matrix which is then used to perform multi-dimensional scaling. For the specific details, see.[27] In AlphaCryo4D[E],[28] the low dimensional manifold embedding is obtained by first extracting features from the volumes using an autoencoder neural network and then applying t-SNE[29] to the volumes and their features.

Once the low dimensional manifold of conformations is obtained, one can analyze the conformational landscape, for example by performing clustering on the embedding, or high-resolution homogeneous refinement based on specific regions of the manifold.

Since this approach relies on volumes that are produced by some other methods, in inherits some difficulties in generating such volumes from other methods. That being said, many of the methods in the remaining sections involve some components of dimensionality reduction or manifold learning on the latent conformation space, which are conceptually related to the methods in this section and the previous section. In particular, a different take on manifold learning of volumes, presented in[30] combines manifold learning techniques with covariance estimations that we discuss in the next section.

Principal Volumes and Linear Models for Volumes

The approaches in this section and the remaining sections incorporate generative models that describe the space of conformations directly. This

[D] StructMap is distributed as part of the ContinuousFlex plugin for Scipion (see footNote 7).

[E] The AlphaCryo4D code can be found at https://github.com/alphacryo4d/alphacryo4d/ and

is in contrast to the ''Manifold Learning on Particle Images'' section, which maps the conformation space without creating an internal representation of the space of possible volumes, and in contrast to the ''Manifolds of Volumes'' section, where the manifold is not inherently a generative model (with some exceptions such as[30] which integrates ideas from the ''Manifolds of Volumes'' and ''Principal Volumes and Linear Models for Volumes'' sections).

One approach to representing the densities is based on linear combinations of a number of given *principal volumes*. These principal volumes are a subset of the principal components of the dataset consisting of all the volumes represented by the particle images (see more details after Eq. (6)). In this representation, we have a reference volume, which we denote by X_0, and K principal volumes, which we denote by $\{X_k\}_{k=1}^{K}$. The density $V_m := V(m, \cdot)$ of a particular conformation is provided as a linear combination of the principal volumes,

$$V_m \approx X_0 + \sum_{k=1}^{K} b_k^m X_k,$$

where the coefficient b_k^m is the weight that the k-th principal volume receives in the m-th conformation; there is one conformation (and one set of coefficients) corresponding to each particle image. The coefficients b_k^m are different for each conformation and can be positive or negative, and the principal volumes are often chosen to be orthogonal to each other, such that $\int X_k(x) X_n(x) dx = 0$ if $k \neq n$ (excluding X_0). Once the principal volumes are obtained, the standard practice is to visualize each principal mode separately. The expression for the movie of the k-th principal mode as a function of "time" τ (negative or positive) is

$$V^k(\tau) = X_0 + \tau X_k. \tag{6}$$

We call τ "time" here because the common way to visualize the principal volume is a movie that shows V^k at a sequence of values of τ, with consecutive frames representing small increments in τ. The small increments in τ imply small differences in the volumes $V^k(\tau)$ rendered in the movie, yielding a smooth movie of transitions between states. Importantly, τ does not imply any relation to the temporal behavior of the actual molecule. More specifically, τ can be seen as a conformation variable in a subset of the conformation space determined by the principal volume X_k.

If the volumes in different conformations were given (and aligned), a natural optimal choice of principal volumes X_k is obtained by a standard Principal Component Analysis (PCA) of the densities. This idea was first introduced in[31] and further developed in,[32] where the volumes are obtained by homogeneous reconstruction using resampled subsets of the particle images.

Remarkably, the authors of[33,34] demonstrate that it is possible to estimate the 3D covariance matrix and principal volumes directly from the 2D particle images under mild assumptions. This surprising fact is evident from the expressions for the covariance in Fourier domain, extending the Fourier slice theorem. Specifically, in,[34] appropriate estimators $\widetilde{X}^M$ and Σ^M for X_0 and the covariance matrix Σ_0 are defined using least-squares optimization problems. These estimators are shown to be consistent, i.e. they converge to μ_0 and Σ_0 as the number of particle images M tends to infinity. Finding μ_M and Σ_M then involves solving two linear systems. The work in[35–37] improves the solvers, both in terms of scalability and generality. The authors of[33] propose a probabilistic PCA approach, where the principal volumes are estimated directly from the data without first computing the covariance matrix Σ_0. This method is improved in[38] and in[39] (implemented in CryoSPARC), which allow reconstruction of volumes of higher resolution[F].

The interpretation of heterogeneity based on the modes in (6) is widely used in practice and has been demonstrated in several examples.[35,38,39] However, as discussed in more detail in,[40] the covariance method has limited applicability to high-resolution approximation of large continuous conformational variability. The problem, in a nutshell, is that high-resolution large continuous conformational variability does not behave like Eq. (6). Instead, one would require a larger number of high-resolution principal volumes and to identify specific combinations of coefficients b_k^m in Eq. (5) that represent valid conformations.

A refined and potentially more interpretable analysis is proposed in.[30] Each particle image can be best explained as a tomographic projection of a particular volume that is a specific linear combination (Eq. (5)) with particular coefficients b_k^m. This translates each particle image to an approximate volume, which can then be used for manifold learning on volumes (the ''Manifolds of Volumes'' section). The algorithm in[30] further extends the manifold-learning approach to produce refined basis volumes, called *spectral volumes*, that are compatible with the recovered manifold.

A key limitation of this approach to the analysis of heterogeneity is the prerequisite of known viewing directions for each particle image. Large conformational heterogeneity makes it challenging to define and compute consistent viewing directions.

Normal Modes and Predefined Spaces of Conformations

In this section, we discuss a direct representation of the positions of atoms (or pseudoatoms) in each conformation. Let $\mathbf{r}_l = (x_l, y_l, z_l)^T$ be the position of

F Code for the principal volumes method is available as part of the ASPIRE software package http://spr.math.princeton.edu, as well as in CryoSPARC with its 3D variability analysis functionality (link given in footNote 1).

the l-th atom (or a pseudoatom) in a molecule, and let $\mathbf{r} = (\mathbf{r}_1^\intercal, \mathbf{r}_2^\intercal, \ldots)^\intercal \in \mathbb{R}^{3N}$ be the concatenation of the coordinates of N atoms.

The structural variability of molecules can be simulated and studied computationally by Normal Mode Analysis (NMA).[41] NMA produces a linearized model for small perturbations around a reference location of the atoms. The best reference structure is an (already determined) atomic model of the molecule, where the way in which the protein is folded will distinguish rigid (e.g. hydrogen-bonded) from non-rigid regions. An alternative but much less powerful approach treats a low-resolution envelope of the molecule (say, as determined by cryo-EM) as an elastic structure. In this model-free NMA, larger pseudo-atoms are used for a coarse representation of the molecule.

Analogously to the "Principal Volumes and Linear Models for Volumes" section, the positions of atoms $\mathbf{r}^n$ in any conformation can be approximated by a linear combination of the normal modes $\{\Delta_k\}_{k=1}^K$ added to the reference positions $\mathbf{r}^0$:

$$\mathbf{r}^m \approx \mathbf{r}^0 + \sum_{k=1}^{K} b_k^m \Delta_k \tag{7}$$

where the coefficients b_k^m determine the amplitude of the perturbation in the m-th conformation and each set of coefficients $\{b_k^m\}_{k=1}^K$ corresponds to a particle image. The coefficients b_k^m are different for each conformation, and they can be positive or negative. While the normal modes $\{\Delta_k\}_{k=1}^K$ are determined by the physics of the molecule, the coefficients $\{b_k^m\}_{k=1}^K$ are chosen so that the conformations fit the cryo-EM particle images.

The starting point of NMA is defining a potential function of the energy of the reference reconstruction, which is usually considered to be a low energy conformation.

There are two distinct approaches to obtaining the reference structure: a model-based approach, where the reference structure is obtained from the atomic model of the molecule of interest, and a model-free NMA, where larger pseudo-atoms are used for a coarser representation of the molecule (obtained, for example, from a volume determined using cryo-EM). The quality of the normal modes varies with the quality of the model.

Then, the normal modes are the eigenvectors of the Hessian of the potential function, with the normal modes corresponding to the large eigenvalues describing large, collective motions of atoms and the modes corresponding to small eigenvalues describing more localized motions. Once the normal modes are approximated from an approximate model of the molecule, cryo-EM comes into the heterogeneity analysis to find the that best combination of normal modes to explain each particle image.

In,[42] a simple potential function that models the distances between pairs of close atoms as harmonic oscillators is introduced and is shown to accurately describe collective motions of atoms in a molecule as well as more complex and computationally expensive potential that explicitly model bond lengths, bond angles and dihedral angles. A web server allowing a user to compute the normal modes of a protein based on this idea is described in,[43] and the same idea enables recent approaches like HEMNMA[44] to scale to full cryo-EM datasets[G].

In early work in the context of cryo-EM, NMA has been used for flexible fitting of high resolution structures from X-ray crystallography to low-resolution maps from cryo-EM[45] and for finding new conformations of a previously determined structure.[46] These ideas are taken further in HEMNMA,[44] a method for computing the conformational space of a reference structure by fitting the normal mode coefficients b_k^m that best explain each observed particle image in a cryo-EM dataset. This works by iteratively refining the normal mode coefficients and pose coordinates (rotation angles and translations) for each particle image.

In,[47,48] global collective motions described by NMA are combined with local atomic displacements given by Hamiltonian Monte Carlo sampling or molecular dynamics simulations. In DeepHEMNMA,[49] a deep neural network is trained on particle images, normal mode coefficients and pose parameters obtained from HEMNMA for a subset of the data, which then outputs the corresponding normal mode coefficients and pose parameters for the remaining data. In,[50] an unsupervised learning approach is used to estimate normal mode coefficients and in-plane rotations, while in,[51] the coefficient of one normal mode is sampled using a deep encoder neural network together with the CTF defocus and in-plane orientation of the particle.

Subsets of particle images that have similar values of coefficients are assumed to have the same conformations; such subsets can be used to reconstruct the volume directly from the images, or the conformations can be better visualized in a lower dimensional space by performing PCA on the normal mode coefficients. An important prerequisite for the algorithms is the reference X_0 and its atomic model, which is used to compute the normal modes $\{\Delta_k\}_{k=1}^K$.

WarpCraft[52] models continuous heterogeneity by using normal mode analysis to combine different components of the volume, which avoids the downsides of multi-body refinement of rigid components.

A comprehensive discussion of NMA is available in[41] and in the context of cryo-EM in.[5]

There are other approaches based on the physics of the molecules to provide conceptually related, but very different conformation spaces. Specifically, in,[53] molecular dynamics simulations are biased

[G] HEMNMA and DeepHEMNMA are distributed as part of the ContinuousFlex plugin for the Scipion software package https://github.com/scipion-em/scipion-em-continuousflex.

with cryo-EM data by sampling a posterior determined by an energy function that combines a standard molecular dynamics energy term and one that takes into account the cryo-EM data, and in,[54] CryoBIFE uses a Bayesian approach to extract the free energy profile of a molecule directly from cryo-EM images together with its uncertainty.

Remark 2 We note that in most normal mode analysis work described in the current section, the space of conformations is predetermined by the fixed normal modes expressed in Eq. (7). These normal modes are not inherently based on cryo-EM data, but rather on some external predefined model of the molecule (although that model might be derived from some reference homogeneous model reconstructed from cryo-EM data).

Nonlinear Models: Hyper-Molecules

In the "Principal Volumes and Linear Models for Volumes" and "Normal Modes and Predefined Spaces of Conformations" sections, we discussed linear models, and in the "Manifold Learning on Particle Images" section, we discussed a nonlinear approach based on manifolds of two-dimensional tomographic projections. In this section, we discuss nonlinear models for the volume V. The general formulation of nonlinear models in cryo-EM, laid out in,[55,56] models all the conformations of the molecule in one *hyper-molecule* function $V(\boldsymbol{\tau}, \mathbf{r})$, where $\boldsymbol{\tau}$ is the conformation variable (position in the conformation space) and $\mathbf{r}$ is a point in space (or a frequency). In the simplest case, $\boldsymbol{\tau} := \tau$ is a scalar and the function $V(\tau, \mathbf{r})$ is analogous to a movie: if we fix a point τ in "time" we obtain a single conformation $V(\tau, \cdot)$ analogous to a single still frame in a video. In other words, $V(\tau, \mathbf{r})$ is a complete description of a continuum of conformations. More generally, $\boldsymbol{\tau}$ can be a high dimensional vector capturing complex continuous heterogeneity. For example, when $\boldsymbol{\tau}$ is two dimensional, one can imagine a planar map where each point represents a conformation.

The hyper-molecule function $V(\boldsymbol{\tau}, \mathbf{r})$ has been implemented in many different ways in different works.

It can be argued that a continuous model can be approximated with a sufficiently large number of discrete samples, or a sufficiently large number of classes in 3D classification in cryo-EM. This would be analogous to the sequence of individual frames that represents a movie. However, the discrete 3D classes are analogous to an unordered collection of frames, which is not as interpretable as a movie. More importantly, the relation between adjacent frames in a video reduces the amount of information required to represent the movie (indeed, compressed videos are more efficient than a collection of images representing the same frames); this is loosely translates to fewer particle images one would need in order to recover $V(\tau, \cdot)$ compared to the number of particle images required to recover a very large number of discretized classes. We revisit the motivation for continuous functions more formally in the "Continuous vs. discrete models" section.

Orthogonal basis functions

The hyper-molecule $V(\boldsymbol{\tau}, \cdot)$ can be represented in different ways. The first classic harmonic analysis approach proposed in[55,56] uses linear combinations of orthogonal basis functions P_k:

$$V(\boldsymbol{\tau}, \mathbf{r}) = \sum_k a_k P_k(\boldsymbol{\tau}, \mathbf{r}). \tag{8}$$

One practical implementation of (8) is a product of three dimensional basis functions (such as prolate spheroidal functions[57,58,56]) and one dimensional basis functions.

The principle was demonstrated to work using a stochastic gradient decent like algorithm on synthetic data in,[55] and using a Markov Chain Monte Carlo (MCMC) algorithm and synthetic and real data in.[56] In both cases, the viewing direction and conformational state are not assumed to be known a priori. While[55,56] argue the hyper-molecules are the natural generalization of traditional 3D volumes to the case of continuous heterogeneity, it argued that it would become increasingly difficult to rigorously generalize the traditional expectation–maximization and branch-and-bound algorithms to high-dimensional latent spaces that involve viewing direction, translation and complex conformation variables. Therefore, they propose formal MCMC algorithms and speculate about possible use of variational approximations.

Remark 3 To contrast between Eqs. (8) and (5), we note that the coefficients a_k in (8) are a shared part of the model, whereas the coefficients b_k^m in (5) are different for each individual particle image.

Remark 4 In fact, the linear density model in covariance approach in Eq. (5) is a special case of nonlinear models. For example, one can define $\boldsymbol{\tau}$ to simply be the vector of coefficients $\boldsymbol{\tau} := (b_1, b_2, \ldots . b_K)$, so that $V_m(\boldsymbol{r}) = V(\boldsymbol{\tau}^m, \boldsymbol{r}) = X_0 + \sum_{k=1}^K b_k^m X_k$.

CryoDRGN

CryoDRGN[59,60] introduces new ideas inspired by the success of Variational AutoEncoders (VAEs)[61] in other applications[H]. VAEs have two components, an *encoder* which is optimized to provide approximations of the distribution of latent variables, and a *decoder*, which is optimized to reproduce particle images given the latent variables as inputs. Figure 2 is a schematic illustration of the cryoDRGN architecture. Many of the neural network based approaches discussed in the remainder of this section have an analogous architecture.

[H] The CryoDRGN software is available at https://cb.csail.mit.edu/cb/cryodrgn.

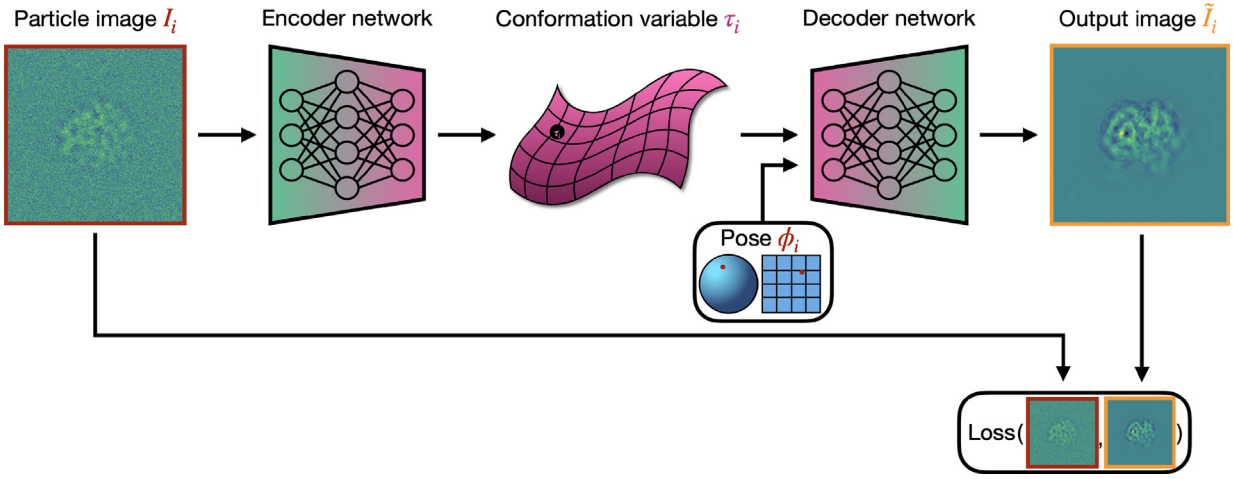


Figure 2. CryoDRGN architecture. In order to train the encoder and the decoder networks, a particle image I_i is given as input to the encoder, which outputs a latent space representation (or conformation variable) $\boldsymbol{\tau}_i$ of the particle image. Then, the decoder network takes $\boldsymbol{\tau}_i$ and the pose variable ϕ_i as input and outputs a projected image $\tilde{I}_i$ of the estimated volume. The loss $\mathscr{L}\left(I_i, \tilde{I}_i\right)$ between the input particle image and the output image is computed and the weights of the networks are adjusted accordingly. The GMM and GMM with folding constraints models described in the "GMM model" and "GMM with folding constraints" sections follow a similar workflow, with the most significant difference being that the decoder outputs the atom positions in the conformation corresponding to $\boldsymbol{\tau}_i$, which are then used to generate the output particle image $\tilde{I}_i$.

Remark 5 The use of a deep neural network in CryoDRGN and other methods below might create a perception that CryoDRGN is somehow pretrained on some dataset with many known structures; in fact, rather than being pre-trained, CryoDRGN fits the weights of the neural networks separately for each dataset to which it is applied.

CryoDRGN's decoder is a hyper-molecule function $V(\boldsymbol{\tau}, \mathbf{r})$, implemented as a standard multilayer perceptron (MLP), with an interesting twist. The conformation latent variable $\boldsymbol{\tau}$ (typically 8 dimensional) feeds into the MLP in the traditional way. The coordinates $\mathbf{r} = (r_1, r_2, r_3)$ (in the Fourier-Hartley domain) first go through positional encoding where each of the three coordinates is augmented; for example, the coordinate r_1 is augmented by the vector $\tilde{\mathbf{r}}_1$ with:

$$\tilde{\mathbf{r}}_1[2i] = \sin\left(r_1 N\pi(2/N)^{2i/N}\right), \tag{9}$$

$$\tilde{\mathbf{r}}_1[2i+1] = \cos\left(r_1 N\pi(2/N)^{2i/N}\right), \tag{10}$$

where $i = 1, \ldots, N/2$, and N is the size of the box containing the volume of interest. The positional encoding modification to the MLP by augmenting the input vectors has been used in transformers and vision applications; it plays an important role in fitting high-frequency components of functions.[62]

In order to produce a tomographic projection from a particular viewing direction and conformation $\boldsymbol{\tau}$, CryoDRGN computes a grid of points $r_{1,1}, r_{1,2}, \ldots$ in the Fourier-Hartley domain and evaluates $V(\boldsymbol{\tau}, r_{1,1}), V(\boldsymbol{\tau}, r_{1,2}), \ldots$ at these points. The grid of values is the Hartley transform of the particle image.

Previous algorithms store or compute for each particle image I_i some form of explicit latent variables such as the viewing direction R_i and conformation $\boldsymbol{\tau}_i$. For example, MCMC software stores these latent variables and software like RELION compares each particle image to projections of different volumes in different viewing directions. Instead of storing the explicit $\boldsymbol{\tau}_i$ for each particle image I_i, CryoDRGN's encoder is optimized to take the image I_i itself as input and compute a sample from $p(\boldsymbol{\tau}_i|I_i)$. Anecdotal evidence points to similarities between the encoder implementation and explicit latent variables.[63]

The optimization objective for the encoder and decoder networks is the standard VAE variational lower bound of the model evidence, which includes the reconstruction error as the mean squared error between the reconstructed image and the input image. We note that the encoder and decoder networks are optimized together for each dataset; neither is pretrained on other datasets.

The original CryoDRGN implementation was demonstrated with known precomputed viewing directions. Followup work[64] demonstrated a hybrid use of explicit latent variables for viewing directions and translations, and an encoder for the conformation variables. Related work on CryoFIRE[65] uses an encoder that outputs both conformation and pose variables using amortized inference, while in,[66] score-based generative diffusion models[67,68] are used to sample from the latent space of CryoDRGN, showing significant improvements.

Adding model constraints

The functions used to express hyper-molecules in[55,56] and the deep learning version in[59] are very generic in their ability to express a wide variety of changing volumes; including ones that are not physically plausible. In,[56] it is pointed out that complex heterogeneity presents an information problem. This can be illustrated by looking at a simplified discretized case. Suppose that a good reconstruction of a volume requires 10^5 particle images. Now suppose that the molecule has a flexible region and that the flexibility can be adequately represented using 10 states, each requiring 10^5 particle images for a good reconstruction (we will see in the discussion that continuous models may require fewer particle images, but the principle is similar). For convenience, let us assume that the particle images are conveniently equally distributed among the different states and that we are able to associate each image to the correct state. Now suppose that there are 2 such flexible regions; this means that there are 100 combinations of states of the two regions, each requiring 10^5 particle images. The number of particle images grows exponentially fast with the complexity and number of regions (if not mitigated somehow).[56] proposes to enforce structure in the models that capture any of the physical properties that can be assumed about the molecule; the example presented there is explicit decomposition of the volume into regions that are allowed move independently, and it is argued that the number of coefficients describing the problem (loosely reflecting on the number of particle images) grows linearly (or even remains constant) as the number of regions grows. We note that in contrast to the multi-body approach to decomposition of the volume (the ''Multi-body extension of traditional analysis'' section), the components in the example in[56] are not rigid.

The following sections discuss several different approaches that enforce models on hyper-molecules.

GMM Model

A Gaussian mixture model (GMM) was introduced in[69] as a different model for expressing $V(\boldsymbol{\tau}, \mathbf{r})$[I]. In GMM models, the volume is represented as a sum of N Gaussians centered at spatial coordinates $\{\mathbf{r}_j(\boldsymbol{\tau})\}_{j=1}^N \subset \mathbb{R}^3$, which vary continuously with the conformation variable $\boldsymbol{\tau}$:

$$V(\boldsymbol{\tau}, \mathbf{r}) = \sum_{j=1}^{N} a_j(\boldsymbol{\tau}) e^{-\frac{|\mathbf{r}-\mathbf{r}_j(\boldsymbol{\tau})|^2}{2\sigma_j(\boldsymbol{\tau})^2}}, \tag{11}$$

where $\{a_j(\boldsymbol{\tau})\}_{j=1}^N, \{\sigma_j(\boldsymbol{\tau})\}_{j=1}^N \subset \mathbb{R}_+$ are the amplitudes and widths of the Gaussians respectively and are also dependent on the conformation variable $\boldsymbol{\tau}$.

[I] The code is distributed as part of the EMAN2 software at https://blake.bcm.edu/emanwiki/EMAN2/e2gmm.

The GMM model in[69] is based an autoencoder architecture like CryoDRGN. The decoder in this model takes the latent conformation variable $\boldsymbol{\tau}$ (with a default dimension four) and produces a list of centers $\{\mathbf{r}_j(\boldsymbol{\tau})\}_{j=1}^N$ that can then be used to evaluate the tomographic projection for any viewing direction. The encoder computes the latent conformation variable $\boldsymbol{\tau}$ for any given images (with notable differences that we omit here from the CryoDRGN encoder).

The training of the neural network is performed in two steps: the encoder is first trained to match a pre-computed neutral structure corresponding to $\boldsymbol{\tau} = \boldsymbol{0}$, and then both the encoder and the decoder are trained to match the variability in the data. The second step of the training is similar to that of CryoDRGN variable ϕ_i.

Remark 6 The NMA representation using pseudoatoms with locations given in (7) is a special case of the GMM representation in (11) (but not the specific algorithm discussed in this section). We can see this by letting $a_j(\boldsymbol{\tau}) = a_0$ and $\sigma_j(\boldsymbol{\tau}) = \sigma_0$ for all $j = 1, \ldots, N$, where a_0 and σ_0 are the fixed values of the amplitudes and widths of the pseudoatoms in (7) and by letting

$$\mathbf{r}_j(\boldsymbol{\tau}_m) = \mathbf{r}_j^0 + \sum_{k=1}^{K} b_k(\boldsymbol{\tau}_m)\Delta_{k_j}, \tag{12}$$

where $\mathbf{r}_j^0 \in \mathbb{R}^3$ is the vector of coordinates corresponding to the j-th pseudoatom in the reference structure, $\Delta_{k_j} \in \mathbb{R}^3$ contains the entries of the normal mode Δ_k corresponding to the j-th pseudoatom, and the normal mode coefficients $b_k^m := b_k(\boldsymbol{\tau}_m)$ are functions of the conformation variable $\boldsymbol{\tau}_m$.

GMM with folding constraints

The methods described thus far make limited use of what we know about the molecules imaged in the experiment. Biomolecules are a folded chain of amino acids, and in many cases one has information about the chain and even some approximation of the folded structure. This information can be incorporated into a more nuanced model.

AtomVAE[70] [J] models the molecule at the atomic level, expressing $f(\boldsymbol{\tau}, \mathbf{r})$ as a GMM model where each atom (other than hydrogen) is represented by a Gaussian. The architecture of AtomVAE is also based on a VAE approach, where the encoder consists of three networks: one that takes a particle image and encodes it into a general latent vector, one that takes the general latent vector and encodes the conformational landscape, and one that take the general latent vector and encodes the pose. The decoder takes the sampled latent conformation vector and the pose as inputs and outputs the translations with respect to a neutral structure, previously

[J] The survey in[9] refers to this method as AtomVAE.

obtained using homogeneous reconstruction (or purely computational methods).

CryoFold[71] models the molecule at the residue level, expressing $f(\boldsymbol{\tau}, \mathbf{r})$ as a GMM model where each amino acid is represented by two Gaussians, one for the backbone and one for the side chain. It assumes known, fixed, viewing directions and uses a VAE architecture similar to that of CryoDRGN and outputs the translations of the Gaussians with respect to a neutral structure, as done by AtomVAE.

In both methods, structural constraints related to the positions of the atoms in the backbone and the side chains are specified in the loss function, in addition to the error in the reconstructed images. They have been demonstrated on synthetic data and they require a good initialization with a reasonable reference structure.

Deformation based models

It is appropriate to model many cases of continuous heterogeneity as deformations of a reference volume; this may exclude the binding and unbinding of subunits or ligands, such as elongation factors on ribosomes. Given a reference volume $V_0(\mathbf{r})$, a hyper-molecule description of the deformed volume can loosely be formulated as

$$V(\boldsymbol{\tau}, \mathbf{r}) = V_0\Big(g[\boldsymbol{\tau}]^{-1}(\mathbf{r})\Big),$$

where $g[\boldsymbol{\tau}]$ expresses the deformation of the reference to the conformation $\boldsymbol{\tau}$ (a more nuanced technical description may also account for the Jacobian of the deformation). Deformation based models attempt to fit both the reference volume $V_0(\mathbf{r})$ and the deformations $g(\boldsymbol{\tau})$.

Deformation functions based on Zernike polynomials have been used in,[72–74] which combine ideas from manifolds of volumes described in the "Manifolds of Volumes" section. 3DFlex[75] uses a deep neural network that produces a deformation field generator g as a function of the latent conformation variable $\boldsymbol{\tau}$. 3DFlex optimizes the parameters of the deformation generator network, the canonical volume, and the latent conformation variable of each particle image in a maximum-likelihood framework. The viewing directions and the CTF parameters are assumed to be known, but can be further refined by 3DFlex[K].

Follow up work on the GMM method in[69] is presented in,[76] which combines physical constraints and deformation models in a more scalable version of the original GMM algorithm.

Bypassing the Estimation of Latent Variables

Let us assume that the particle images are generated according to a classic generative model, where latent variables such as conformation, viewing direction, translation (and possibly CTF parameters) are selected at random and used to generate a particle image along the lines of Eq. (1). Let us further assume that the distributions of these latent variables and the noise are known. In theory, this would give us a distribution of particle images that is conditional on the hyper-molecule.

It was shown in[77–80] that a given distribution of particle images can be theoretically traced back to a unique volume (up to trivial symmetries). For simplicity, we limit the description in this section to the problem of homogeneous reconstruction when there is no conformational heterogeneity in the data and we will point to the way it generalizes to the heterogeneous case without delving into details.

One of the methods to compute parameters from a distribution in many statistical problems is the method of moments, where moments of a distribution are computed and used to estimate the parameters of the distribution. Indeed, back in the 1970s, Zvi Kam found that one can compute moments of the distribution of observed images and use these to reconstruct the volume.[81] The method has been further developed in recent years, for example in.[79,80,82] Remarkably, this can be applied directly to micrographs, without particle picking.[83] This approach, as well as the broader area of multi-reference alignment, has been shown to extend to the case of heterogeneity (e.g.,[84,85]).

Recently, machine learning approaches have emerged for approximating distributions. CryoGAN[78] is a machine learning approach that directly finds a volume that generates the distribution of observed particle images. We note that, in practice, the data consists of a large but finite number of samples from the distribution, not the distribution itself. Considering the many latent variables (viewing direction, shifts, CTF), it might be surprising that this is enough for CryoGAN to recover a volume, but the idea has been demonstrated in.[78] The method has been extended to the case of heterogeneity in.[86]

Discussion and Perspectives

Having surveyed the main approaches to continuous heterogeneity analysis, in this section we discuss some of the broader questions that these approaches pose.

Reading the output: Explicit models vs. reconstruction from images

Given the output of a continuous heterogeneity reconstruction algorithm, there are generally two approaches to obtaining the volume corresponding to a conformation at a specific value of the conformation variable $\boldsymbol{\tau} = \boldsymbol{\tau}^*$.

K 3DFlex is part of the CryoSPARC software package (link in footNote 1).

One approach involves selecting all the particle images $\{I_i\}$ that have been assigned values $\{\boldsymbol{\tau}_i\}$ that are within a certain distance from $\boldsymbol{\tau}^*$ and feed them to a homogeneous reconstruction algorithm. This produces one volume that represents that specific region of the conformation space. Alternatively, one can directly use the explicit model generated by the algorithm. For example, the hyper-molecule models described in the "Nonlinear Models: Hyper-Molecules" section have an internal representation of each state, which allows the user to generate the volume $V(\boldsymbol{\tau}^*, \cdot)$. Similarly, Eq. (6) gives a way to visualize principal volumes.

Some of the methods presented in this survey rely on homogeneous reconstruction to generate volumes with different conformations, as they lack an explicit representation of the volumes (the manifold of images methods in the "Manifold Learning on Particle Images" section are an example). However, most of the other methods lend themselves to both approaches to generating heterogeneous volumes. In this case, the homogeneous reconstructions from subsets of the particle images can be used as a form of validation, as the final volumes are reconstructed directly from the data: if the volumes appear to be high-resolution and biologically plausible, users tend to accept them as a successful run of the algorithm. Such a result is especially convincing when generating volumes directly from the model would result in volumes that are relatively low-resolution due to model constraints such as a limited number of principal volumes computed in the principal volume approach or a small number of Gaussians in the GMM models. Furthermore, as models get more elaborate (e.g., atomic models), there are concerns that the results present artifacts that reflect bias and error in the model, which are somewhat mitigated if one obtains homogeneous reconstructions from subsets of the data.

That being said, visualization based on models can provide higher-resolution volumes, which reflects the advantages of continuous models over discrete reconstructions. As an illustrative example, consider a molecule with one flexible region that requires 10 separate discrete models to be fully captured in the reconstruction, for which we are given 10^6 particle images. If we restrict our attention to the rigid part of the volume, applying one of the hyper-molecule algorithms would use all 10^6 particle images for the reconstruction of this region, while selecting subsets of 10^5 particle images for homogeneous reconstruction of one of the conformations leads to only this subset of the data being used to reconstruct the rigid part, which would otherwise benefit from the full dataset. Modeling considerations and the ability to refine the viewing directions taking into account flexible regions extends this argument to other less obvious cases.

Continuous vs. discrete models

Continuous functions can be approximated by a sufficiently large number of discrete samples, and therefore it is compelling to deduce that traditional discrete 3D classification is a good discretization of continuous heterogeneity. In this section, we point to several advantageous features of continuous models.

First, in contrast to discrete heterogeneity analysis, continuous heterogeneity analysis does not require the user to specify the number of classes to be reconstructed. This choice, which is not trivial when the underlying heterogeneity is concentrated in discrete states and influences the resolution of the inferred model, is particularly difficult when the heterogeneity is inherently continuous. Instead, approaches for continuous heterogeneity analysis explicitly model the continuum of states, where states that are sufficiently distinct in the particle images tend to naturally appear as dense "islands" on the conformation manifold.

Second, methods for continuous heterogeneity are able to use all the available data to produce higher resolution volumes in any conformation, while discrete heterogeneity approaches only use a subset of the data for a particular conformation. The example given in the "Reading the output: Explicit models vs. reconstruction from images" section illustrates this aspect: while the rigid part of a volume can benefit from the full dataset, 3D classification methods only use the particle images corresponding to one specific conformation to reconstruct the volume, including the rigid part, leading to lower resolution than possible. Here, one could correctly argue that the rigid region can be reconstructed using homogeneous reconstruction separately from the heterogeneous region, as it is done in multi-body refinement (the "Multi-body extension of traditional analysis" section). However, continuous heterogeneity analysis may still perform better, as it can update the viewing directions estimations (potentially yielding better estimation of the conformation variable) and it models the continuum of states more accurately than discrete models, which do not have a notion of "neighboring states". For example, the deformation model described in the "Deformation based models" section severely restricts the space of possible conformations to a (limited) deformation of a reference volume, and therefore all the particle images from all conformations contribute high-resolution information to the reference volume.

Third, an implication of the previous argument is that continuous heterogeneity analysis has an advantage in analyzing rare conformations using the continuity between conformations. Using the deformation model as an example, when the deformation assumption is applicable, this model constructs a high resolution reference volume based on the highly populated conformations and uses the particle images from a rare conformation effectively to determine the path of deformation of the reference volume. This leads to a high resolution approximation of the rare conformation as the correct deformation of the reference volume.

One last argument regards the advantage of continuous heterogeneity analysis over multi-body analysis (the ''Multi-body extension of traditional analysis'' section). While it has been implemented as an extension of traditional reconstruction algorithms, multi-body analysis can also be seen as a special case of the hyper-molecule model. Specifically, is has proven to be very effective in the analysis of rigid components moving with respect to each other, but it falls short when such structural assumptions do not hold and it fails to resolve the interface regions between the rigid components, issues which are addressed in the hyper-molecule models for examples.

Finally, it is worth noting that high resolution homogeneous reconstruction and traditional 3D classification as well as multi-body analysis are available in mature software and are well understood by practitioners. In contrast, practitioners still face the dilemma of which continuous heterogeneity software, model assumptions and parameters to use in continuous heterogeneity analysis, so it is not the aim of this article to argue that the problem has been solved or that higher resolution is always easily attainable, but to highlight conceptual advantages on which these algorithms increasingly capitalize.

Outliers, junk, and rare conformations

Practitioners have long used 2D and 3D classification not only to identify actual conformations, but also to identify "junk"; particles that produce low quality volumes and objects that were incorrectly picked as particles are discarded. One of the caveats in this procedure is that it is difficult to determine whether high quality particle images are also discarded. Perhaps most importantly, it is likely that (valid) particle images from rare conformations are discarded, too. This is an undesirable outcome of the reconstruction algorithm not being able to create high quality coherent classes using them.

Preliminary informal evidence suggests that an analogous procedure is applicable to different algorithms for continuous heterogeneity analysis: junk particles appear to be "pushed" out to separate regions of the conformation space, reducing their influence on the "valid" regions, which in this case correctly include particles corresponding to the rare conformations in addition to the ones that would be used by homogeneous and discrete heterogeneous algorithms. It is therefore conceivable that some of the continuous heterogeneity analysis algorithms strike a better balance between using rare-conformation particle images and discarding junk.

Interpretation of output

The output of continuous heterogeneity analysis varies considerably between different methods. However, there are two main families of outputs. One is a model of the space of molecular structures (e.g., the principal volume expansion of V_m in (5), the hyper-molecule function $V(\boldsymbol{\tau}, \mathbf{r})$ in (8)), which does not exist in the image manifold approach in the ''Manifold Learning on Particle Images'' section. The other family of outputs consists of the estimates of the conformation variable $\boldsymbol{\tau}$ for each particle image, which does not exist in the methods in the ''Bypassing the Estimation of Latent Variables'' section.

From both types of outputs, one can obtain a low dimensional representation of the latent conformation space of $\boldsymbol{\tau}$, where it is common to try to identify physically meaningful clusters or manifolds. For example, in,[20] the density of the particles in the latent space is related to the energy landscape of the ribosome. In other work, various dimensionality reduction and clustering algorithms are applied to the latent space to identify dominating conformations. We note, however, that the latent conformation variable does not immediately correspond to a physically meaningful and interpretable quantity, it is not necessarily unique even up to trivial symmetries and can in principle be arbitrarily deformed (for some motivating examples, see[87]). Different algorithms and even the same algorithm with a different random seed can produce different latent space representations. Similarly, different post-processing dimensionality reduction and clustering algorithms (and again, even different random seeds) can produce different results. In cases of relatively simple heterogeneity, it is plausible that an expert user would be able to probe structures in the latent space together with the models or reconstructions associated with them in order to understand the mechanisms that it captures. However, complex heterogeneity and more automated tools may require more work on the interpretation of the latent space.

Mathematically, this is a problem related to the manifold metric as transformed by the measurement and reconstruction process, and while it has been studied in other fields,[88–91] it still requires further investigation in a cryo-EM context. Moreover, establishing an appropriate metric for the conformation space that is physically meaningful is a problem that has not been explored yet.

Validation

Validation has always been a nuanced problem in cryo-EM. A common sanity check for homogeneous reconstruction is the Fourier shell correlation (FSC) procedure[92,93]: the data is split into two subsets and two distinct reconstructions are obtained from them. The correlations between the two volumes are computed for each frequency radius and the resolution of the final reconstruction is given by the frequency radius where the FSC curve drops below a specific value (usually 0.143). Above this resolution, the reconstruction is considered to overfit the noise. In addition, a qualitative procedure is for an expert to evaluate and confirm the plausibility of the reconstructed molecule.

There is no clear solution for validating results of continuous heterogeneity analysis. To illustrate the difficulty, consider two hyper-molecules that are generated by one method from two separate subsets of the data. Even if the solutions are equivalent, they can assign different values of τ to the same conformation, in which case it is not clear how to compare the two outputs. One idea proposed in[75] for deformation models (the "Deformation based models" section) is to compute the FSC between two reference models that the algorithm computes for two different subsets of the data. However, this idea only applies to deformation models, since other methods do not necessarily compute a reference volume.

Given the wide variety of families of ideas for analysis of continuous heterogeneity and the different types of output they produce, it is even more challenging to find metrics that are applicable to all methods. Beyond the evaluation of outputs for individual problems, there is a desire to determine which tools perform best, which reinforces the need for evaluation metrics and benchmark datasets, something that is not currently available. Having said this, given the many forms of heterogeneity in applications and the fundamental differences between the algorithms for heterogeneity analysis, one should be cautious that the metrics of choice and the benchmark datasets used do not exclude some of the existing and new ideas.

Conclusions

In this survey, we covered a broad range of methods for analyzing continuous heterogeneity in cryo-EM. In addition, we highlighted the advantages of continuous heterogeneity analysis and the way they circumvent some of the tradeoffs that practitioners encounter in traditional methods involving homogeneous reconstruction, discrete heterogeneous reconstruction and multibody analysis.

As evident in this survey, many different approaches have been proposed in recent years. Most of these ideas are still in active development and it is likely that new ideas will emerge in the coming years. Some of the algorithms, like the various members of the non-linear models in the "Nonlinear Models: Hyper-Molecules" section, share conceptual similarities that make it plausible that we will see convergent software packages that offer multiple and customizable models, as envisioned in.[56] At the same time, there are significant differences between the methods, making different approaches well-suited for different problems.

One of the open problems in this area is the validation of the results of the algorithms. A related difficulty is in comparing different algorithms in the absence of good metrics and benchmarks; indeed, the different approaches differ even in the form of output they produce. A silver lining in the large number of fundamentally different algorithms that are becoming available is that one could conceivably apply several different algorithms to the same dataset and examine how well their conclusions agree qualitatively, which could serve as a temporary form of validation while the area matures.

Finally, we have we witnessed in recent years the immense success of protein structure prediction algorithms such as AlphaFold[94] and RoseTTAfold.[95] While these algorithms are not currently applicable to continuous heterogeneity (and, arguably, neither to discrete heterogeneity), it is plausible that they will evolve in this direction as a next step in complexity. Conversely, as cryo-EM datasets become more massive and broader in scope, incorporating physical models as priors or constraints obtained from structure prediction software will become a necessity. Therefore, there is an opportunity for continuous heterogeneity analysis methods to bridge the gap between such software and the massive experimental datasets. Preliminary examples of this general direction can be found in,[70] which integrates insights from AlphaFold, and in.[54]

While we do not discuss cryo-electron tomography (cryo-ET) explicitly in this survey, many of the approaches to continuous heterogeneity in cryo-EM can be generalized to cryo-ET. One of the complications in cryo-ET is that the targets of the analysis are typically measured in situ, and do not float freely as in typical cryo-EM. The continuous heterogeneity model may capture heterogeneity in the environment instead of heterogeneity in the target (which may or may not be a desirable result, depending on the application).

CRediT authorship contribution statement

Bogdan Toader: Conceptualization, Writing – original draft, Writing – review & editing. **Fred J. Sigworth:** Conceptualization, Writing – original

draft, Writing – review & editing, Funding acquisition. **Roy R. Lederman:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank David Silva Sanchez for his help.
The work is supported in part by NIH grants R01GM136780, R01NS021501 and AFOSR FA9550-21-1-0317.

Received 16 November 2022;
Accepted 16 February 2023;
Available online 28 February 2023

Keywords:
cryo-EM;
single particle reconstruction;
continuous heterogeneity;
conformation manifold

[A] For software, see for example the RELION https://relion.readthedocs.io or CryoSPARC https://cryosparc.com packages.

References

1. (2016). Method of the year 2015. *Nat. Methods* **13**(1), 1–1.
2. Bendory, T., Bartesaghi, A., Singer, A., (2020). Single-particle cryo-electron microscopy: mathematical theory, computational challenges, and opportunities. *IEEE Signal Process. Mag.* **37** (2), 58–76.
3. Singer, A., Sigworth, F.J., (2020). Computational methods for single-particle cryo-EM. *Ann. Rev. Biomed. Data Sci.* **3**, 163–190.
4. Jonić, S., (2016). Cryo-electron microscopy analysis of structurally heterogeneous macromolecular complexes. *Comput. Struct. Biotechnol. J.* **14**, 385–390.
5. Sorzano, C.O.S. et al, (2019). Survey of the analysis of continuous conformational variability of biological macromolecules by electron microscopy. *Acta Crystallogr. Section F Struct. Biol. Commun.* **75** (1), 19–32.
6. Cossio, P., Hummer, G., (2018). Likelihood-based structural analysis of electron microscopy images. *Curr. Opin. Struct. Biol.* **49**, 162–168.
7. DeVore, K., Chiu, P.-L., (2022). Probing structural perturbation of biomolecules by extracting Cryo-EM data heterogeneity. *Biomolecules* **12** (5), 628.
8. Wu, J.-G. et al, (2022). Machine learning for structure determination in single-particle cryo-electron microscopy: A systematic review. *IEEE Trans. Neural Networks Learn. Syst.* **33** (2), 452–472.
9. Donnat, C. et al, (2022). Deep generative modeling for volume reconstruction in cryo-electron microscopy. *J. Struct. Biol.* **214** (4)
10. Glaeser, R.M., Nogales, E., Chiu, W., (2021). Single-particle Cryo-EM of Biological Macromolecules. IOP Publishing.
11. Scheres, S.H., (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180** (3), 519–530.
12. Kimanius, D. et al, (2021). New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478** (24), 4169–4185.
13. Scheres, S.H.W. et al, (2007). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods* **4** (1), 27–29.
14. Lyumkis, D. et al, (2013). Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* **183** (3), 377–388.
15. Punjani, A. et al, (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14** (3), 290–296.
16. Bai, X.-C. et al, (2015). Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife* **4**, e11182.
17. Ilca, S.L. et al, (2015). Localized reconstruction of sub-units from electron cryomicroscopy images of macromolecular complexes. *Nat. Commun.* **6**
18. Nakane, T. et al, (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**
19. Schwander, P. et al, (2010). Mapping the conformations of biological assemblies. *New J. Phys.* **12**, 035007.
20. Dashti, A. et al, (2014). Trajectories of the ribosome as a Brownian nanomachine. *Proc. Nat. Acad. Sci.* **111** (49), 17492–17497.
21. Schwander, P., Fung, R., Ourmazd, A., (2014). Conformations of macromolecules and their complexes from heterogeneous datasets. *Philosoph. Trans. Roy. Soc. B: Biol. Sci.* **369** (1647), 20130567.
22. Frank, J., Ourmazd, A., (2016). Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods. Single Particle Cryo-EM. from sample to reconstruction* **100**, 61–67.
23. Maji, S. et al, (2020). Propagation of conformational coordinates across angular space in mapping the continuum of states from Cryo-EM data by manifold embedding. *J. Chem. Inf. Model.* **60** (5), 2484–2491.
24. Dashti, A. et al, (2020). Retrieving functional pathways of biomolecules from single-particle snap-shots. *Nat. Commun.* **11** (1), 4734.
25. Seitz, E. et al, (2022). Recovery of conformational continuum from single-particle Cryo-EM images: optimization of manifoldem informed by ground truth. *IEEE Trans. Comput. Imag.* **8**, 462–478.
26. Coifman, R.R., Lafon, S., (2006). Diffusion maps. *Appl. Comput. Harmonic Anal. Special Issue: Diffusion Maps Wavelets* **21** (1), 5–30.
27. Sanchez Sorzano, C.O. et al, (2016). StructMap: Elastic distance analysis of electron microscopy maps for studying conformational changes. *Biophys. J.* **110** (8), 1753–1765.
28. Wu, Z. et al, (2022). Visualizing conformational space of functional biomolecular complexes by deep manifold learning. *Int. J. Mol. Sci.* **23** (16), 8872.

29. van der Maaten, L., Hinton, G., (2008). Visualizing Data using t-SNE. *J. Machine Learn. Res.* **9**, 2579–2605.
30. Moscovich, A. et al, (2020). Cryo-EM reconstruction of continuous heterogeneity by Laplacian spectral volumes. *Inverse Prob.* **36** (2), 024003.
31. Liu, W., Frank, J., (1995). Estimation of variance distribution in three-dimensional reconstruction. I. Theory. *J. Opt. Soc. Am.. A, Opt. Image Sci. Vision* **12** (12), 2615–2627.
32. Penczek, P.A., Kimmel, M., Spahn, C.M.T., (2011). Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. *Structure* **19** (11), 1582–1590.
33. Tagare, H.D. et al, (2015). Directly reconstructing principal components of heterogeneous particles from cryo-EM images. *J. Struct. Biol.* **191** (2), 245–262.
34. Katsevich, E., Katsevich, A., Singer, A., (2015). Covariance matrix estimation for the Cryo-EM heterogeneity problem. *SIAM J. Imag. Sci.* **8** (1), 126–185.
35. Liao, H.Y., Hashem, Y., Frank, J., (2015). Efficient estimation of three-dimensional covariance and its application in the analysis of heterogeneous samples in cryo-electron microscopy. *Structure* **23** (6), 1129–1137.
36. Andén, J., Katsevich, E., Singer, A., (2015). Co-variance estimation using conjugate gradient for 3D classification in CRYO-EM. *In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 200–204.
37. Andén, J., Singer, A., (2018). Structural variability from noisy tomographic projections. *SIAM J. Imaging Sci.* **11** (2), 1441–1492.
38. Melero, R. et al, (2020). Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures. *IUCrJ* **7** (6), 1059–1069.
39. Punjani, A., Fleet, D.J., (2021). 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryoEM. *J. Struct. Biol.* **214** (4), 107894.
40. Sorzano, C.O.S., Carazo, J.M., (2021). Principal component analysis is limited to low-resolution analysis in cryoEM. *Acta Crystallogr. Section D: Struct. Biol.* **77** (6), 835–839.
41. Cui, Q., Bahar, I., (2005). Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems. CRC Press.
42. Tirion, M.M., (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **77** (9), 1905–1908.
43. Suhre, K., Sanejouand, Y.-H., (2004). ElNémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res.* **32** (Web Server issue), 610–614.
44. Jin, Q. et al, (2014). Iterative Elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes. *Structure* **22** (3), 496–506.
45. Tama, F., Miyashita, O., Brooks III, C.L., (2004). Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. *J. Struct. Biol. Time-Resolved Imag. Macromol. Processes Interact.* **147** (3), 315–326.
46. Brink, J. et al, (2004). Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. *Structure* **12** (2), 185–191.
47. Vuillemot, R., Jonić, S., (2021). Combined Bayesian and normal mode flexible fitting with Hamiltonian Monte Carlo sampling for cryo electron microscopy. *In: 2021 29th European Signal Processing Conference (EUSIPCO)*, pp. 1211–1215.
48. Vuillemot, R. et al, (2022). NMMD: Efficient Cryo-EM flexible fitting based on simultaneous normal mode and molecular dynamics atomic displacements. *J. Mol. Biol.* **434** (7), 167483.
49. Hamitouche, I., Jonic, S., (2022). DeepHEMNMA: ResNet-based hybrid analysis of continuous conformational heterogeneity in cryo-EM single particle images. *Front. Mol. Bio-sci.*.
50. Nashed, Y. et al, (2022). Heterogeneous reconstruction of deformable atomic models in Cryo-EM. *In: NeurIPS Workshop on Machine Learning for Structural Biology*.
51. Woollard, G. et al, (2022). Physics aware inference for the cryo-EM inverse problem: anisotropic network model heterogeneity, global pose and microscope defocus. *In: NeurIPS Workshop on Machine Learning for Structural Biology*.
52. Schilbach, S. et al, (2017). Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature* **551** (7679), 204–209.
53. Bonomi, M., Pellarin, R., Vendruscolo, M., (2018). Simultaneous determination of protein structure and dynamics using cryo-electron microscopy. *Biophys. J.* **114** (7), 1604–1613.
54. Giraldo-Barreto, J. et al, (2021). A Bayesian approach to extracting free-energy profiles from cryoelectron microscopy experiments. *Sci. Reports* **11** (1), 13657.
55. Lederman, R.R., Singer, A., (2017). Continuously heterogeneous hyper-objects in cryo-EM and 3-D movies of many temporal dimensions. *arXiv:1704.02899*.
56. Lederman, R.R., Andén, J., Singer, A., (2020). Hyper-molecules: on the representation and recovery of dynamical structures for applications in flexible macro-molecules in cryo-EM. *Inverse Prob.* **36** (4), 044005.
57. Slepian, D., (1964). Prolate spheroidal wave functions, Fourier analysis and uncertainty—IV: extensions to many dimensions; generalized prolate spheroidal functions. *Bell Syst. Tech. J.* **43** (6), 3009–3057.
58. Lederman, R.R., (2017). Numerical algorithms for the computation of generalized prolate spheroidal functions. *arXiv:1710.02874*.
59. Zhong, E.D. et al, (2019). Reconstructing continuous distributions of 3D protein structure from cryo-EM images. *In: International Conference on Learning Representations*.
60. Zhong, E.D. et al, (2021). CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18** (2), 176–185.
61. Kingma, D.P., Welling, M., (2014). Auto-encoding variational bayes. *In: 2nd International Conference on Learning Representations ICLR2014*.
62. Tancik, M. et al, (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *NeurIPS*.
63. Edelberg, D.G., Lederman, R.R., (2022). Using VAEs to learn latent variables: Observations on applications in Cryo-EM. *In preparation*.
64. Zhong, E.D. et al, (2021). CryoDRGN2: Ab initio neural reconstruction of 3D protein structures from real cryo-EM images. *In: 2021 IEEE/CVF International Conference on*

Computer Vision (ICCV). IEEE, Montreal, QC, Canada, pp. 4046–4055.
65. Levy, A. et al, (2022). Amortized inference for heterogeneous reconstruction in Cryo-EM. *In: 36th Conference on Neural Information Processing Systems (NeurIPS 2022)*.
66. Kreis, K. et al, (2022). Latent space diffusion models of cryo-EM structures. *In: NeurIPS Work-shop on Machine Learning for Structural Biology*.
67. Sohl-Dickstein, J. et al, (2015). Deep unsupervised learning using nonequilibrium thermodynamics. *In: Proceedings of the 32nd International Conference on Machine Learning. PMLR*, pp. 2256–2265.
68. Song, Y. et al, (2021). Score-based generative modeling through stochastic differential equations. *In: International Conference on Learning Representations*.
69. Chen, M., Ludtke, S.J., (2021). Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. *Nat. Methods*, 1–7.
70. Rosenbaum, D. et al, (2021). Inferring a continuous distribution of atom coordinates from CryoEM images using VAEs. *arXiv:2106.14108*.
71. Zhong, E.D. et al, (2020). Exploring generative atomic models in cryo-EM reconstruction. *In: NeurIPS Workshop on Machine Learning for Structural Biology*.
72. Calero, D.H. et al, (2021). Continuous heterogeneity analysis of CryoEM images through Zernike polynomials and spherical harmonics. *Microsc. Microanal.* **27** (S1), 1680–1682.
73. Herreros, D. et al, (2021). Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. *IUCrJ* **8** (6), 992–1005.
74. Herreros, D. et al, (2023). Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials. *Nat. Commun.* **14** (1), 154.
75. Punjani, A., Fleet, D.J., (2022). 3D flexible refinement: structure and motion of flexible proteins from Cryo-EM. *Microsc. Microanal.* **28**
76. Chen, M., Toader, B., Lederman, R., (2023). Integrating molecular models into CryoEM heterogeneity analysis using scalable high-resolution deep Gaussian mixture models. *J. Mol. Biol.* **435** (9), 168014.
77. Panaretos, V.M., (2009). On random tomography with unobservable projection angles. *Annals Stat.* **37** (6A), 3272–3306.
78. Gupta, H. et al, (2021). CryoGAN: A new reconstruction paradigm for single-particle cryo-EM via deep adversarial learning. *IEEE Trans. Comput. Imag.* **7**, 759–774.
79. Bendory, T. et al, (2017). Bispectrum inversion with application to multireference alignment. *IEEE Trans. Signal Process.* **66** (4), 1037–1050.
80. Bandeira, A.S., Niles-Weed, J., Rigollet, P., (2020). Optimal rates of estimation for multireference alignment. *Mathe. Stat. Learn.* **2** (1), 25–75.
81. Kam, Z., (1980). The reconstruction of structure from electron micrographs of randomly oriented particles. *J. Theor. Biol.* **82** (1), 15–39.
82. Perry, A. et al, (2019). The sample complexity of multireference alignment. *SIAM J. Mathe. Data Sci.* **1** (3), 497–517.
83. Bendory, T. et al, (2019). Multi-target detection with application to cryo-electron microscopy. *Inverse Prob.* **35** (10), 104003.
84. Lederman, R.R., Singer, A., (2020). A representation theory perspective on simultaneous alignment and classification. *Appl. Comput. Harmonic Anal.* **49** (3), 1001–1024.
85. Boumal, N. et al, (2018). Heterogeneous multireference alignment: A single pass approach. *In: 2018 52nd Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6.
86. Gupta, H. et al, (2020). Multi-CryoGAN: Reconstruction of continuous conformations in Cryo-EM using generative adversarial networks. In: Bartoli, A., Fusiello, A. (Eds.), *Computer Vision – ECCV 2020 Workshops. Lecture Notes in Computer Science*. Springer International Publishing, Cham, pp. 429–444.
87. Toader, B., Lederman, R.R., (2022). Remarks on manifold learning and physical phenomena. *In preparation*.
88. Talmon, R., Coifman, R.R., (2013). Empirical intrinsic geometry for nonlinear modeling and time series filtering. *Proc. Nat. Acad. Sci.* **110** (31), 12535–12540.
89. Meila, M., Perrault-Joncas, D., (2012). Metric learning and manifolds: Preserving the intrinsic geometry.
90. Schwartz, A., Talmon, R., (2019). Intrinsic isometric manifold learning with application to localization. *SIAM J. Imag. Sci.* **12** (3), 1347–1391.
91. Bertalan, T., Dietrich, F., Kevrekidis, I.G., (2021). Transformations between deep neural networks. *arXiv:2007.05646*.
92. Henderson, R. et al, (2012). Outcome of the first electron microscopy validation task force meeting. *Structure* **20** (2), 205–214.
93. Scheres, S.H.W., Chen, S., (2012). Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9** (9), 853–854.
94. Jumper, J. et al, (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 1–11.
95. Baek, M. et al, (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373** (6557), 871–876.