

Single particle reconstruction (SPR) in cryo-electron microscopy (cryo-EM) has recently emerged as the method of choice to determine the structure of biological macromolecules to near atomic resolution. The typical procedure for obtaining the final high resolution 3D structure is by starting with an initial guess and iteratively refining it using the acquired dataset of the molecule’s 2D projection images. The final estimate from the refinement procedure is known to often depend heavily on the initial model used as the starting point, thereby making a good initial estimate crucial for success.

In this thesis, we propose and test two novel approaches, which we call Orthogonal Extension and Orthogonal Replacement, for 3D ab-initio and homology modeling in SPR using cryo-EM and X-ray free electron lasers (XFEL). Our approach is inspired by the molecular replacement technique used in X-ray crystallography. We first test both approaches on noisy synthetic datasets.

Motivated by the need for a reliable estimator of the covariance matrix, we develop a new image restoration method to perform CTF correction and denoising in a single step. Through results on several experimental datasets, we demonstrate the efficacy of our method as a single, preliminary step to inspect particle images, detect outliers, and estimate the covariance matrix of the underlying clean images. Our covariance matrix estimator is asymptotically consistent and successfully corrects for the CTF.

An immediate application of improved covariance estimation is an improvement in the 2D classification or class averaging procedure in the cryo-EM pipeline. We digress from 3D homology/ab-initio modeling to focus on this application. Since different cryo-EM images are affected by noise as well as different CTF’s or point spread functions from the microscope, the Euclidean distance between two images is not an optimal metric for their affinity. We derive and test a new affinity measure akin to the Mahalanobis distance to compare cryo-EM images belonging to different defocus groups. We demonstrate that the new metric leads to an improvement in nearest neighbor detection and therefore the obtained class averages.

Finally, we revisit the homology modeling procedure of Orthogonal Extension. We incorporate our improved covariance matrix estimator into the Orthogonal Extension algorithm and propose a family of asymptotically unbiased estimators to recover the 3D structure. We demonstrate the advantage of our estimator through numerical experiments on synthetic and experimental datasets. We foresee this method as a good way to provide models to initialize refinement, directly from experimental images without performing class averaging and orientation estimation in cryo-EM and XFEL. Our second algorithm for ab-initio modeling, Orthogonal Replacement, is tested on synthetic datasets. In future work, Orthogonal Replacement would require designing an appropriate experiment to collect datasets that would facilitate its usage.