

Statistical estimation of spatially-resolved heterogeneity from cryo EM images

Peter Doerschuk¹, Yunye Gong², Nan Xu³, John Johnson⁴

¹Cornell University ²Cornell University, ³Cornell University, ⁴The Scripps Research Institute
pd83@cornell.edu

There are many methods for characterizing the heterogeneity of an ensemble of particles from single-particle cryo EM images. This talk concerns a method [1-3] based on describing the electron scattering intensity of the particle as a Fourier series and describing the coefficients of the Fourier series as random variables that are independent and identically distributed from instance to instance of the particle. The heterogeneity is characterized by estimating the mean and variance of the coefficients from the image data by a maximum likelihood estimator. The mean results give a reconstruction and the variance results give a spatially-resolved characterization of the heterogeneity of the ensemble of particles. When symmetry is present, the method can allow each instance of the particle to lack symmetry while imposing the symmetry on the statistics of the particle. This avoids anomalous peaks in the variance map located on and near symmetry axes of the particle [4]. Imposing symmetry on the statistics also allows the computation of ensemble averages of the product of the electron scattering intensity at two different locations which can be used to detect allosteric interactions between the different locations. We demonstrate the method on the bacteriophage HK97 where we show that binding of the maturation protease on the inner surface of the capsid has wide-ranging effects on the heterogeneity of the outer surface of the capsid.

[1] Y. Gong et. al. *J. Structural Biology*, 193(3):188-195, March 2016.

[2] N. Xu et. al. *J. Structural Biology*, 202(2):129-141, May 2018.

[3] N. Xu et. al. *IEEE Trans. Image Processing*, 28(11):5479-5494, 2019.

[4] S. J. Ludtke, "Methods in Enzymology", 579:159-189, 2016.