

Dimension Reduction Method for Noisy High-dimensional Images and Application to Cryogenic Electron Microscopy

Szu-Chi Chung

Institute of Statistical Science, Academia Sinica

Abstract

In contrast to X-ray crystallography, cryo-EM is amenable to the structural determination of proteins resistant to crystallization. Since molecules are captured in their native states, it can further analyze the conformational mixtures. With the enhancement of algorithms and GPU acceleration, cryo-EM has become the most efficient technique to solve structures of molecules at near-atomic resolution. Take a recent event, for example, after the outbreak of COVID-19 in January, the first structure of 2019-nCoV Spike trimer was published in March using cryo-EM, which has provided crucial medical insight for developing vaccines. However, the data characteristics include strong noise, huge dimension, large sample size and high heterogeneity with unknown orientations have made analysis very challenging. In the literature, dimension reduction plays an important role in overcoming the challenges above. The traditional methods utilized in the field, however, does not well suited for the scenario and they face bottleneck either in computation or performance. In this talk, I will first introduce the related background and principles of cryo-EM image processing. Second, I will discuss our proposed dimension reduction strategy called two-stage dimension reduction (2SDR) which alleviates the computation burden and improves performance over existing methods. Finally, I will demonstrate how the methods can be employed to improve the cryo-EM image processing tasks including denoising, 2D clustering and 3D volume classification.