CFEL SCIENCE

SEMINAR

13th April 2022 - 2:00 p.m.

Virtual meeting room in ZOOM (ID: 895 1951 2902 / PW: 925239)

Cecilia Casadei

Department of Biology, ETH Zürich and Paul Scherrer Institut, Switzerland

Randomness in time-resolved serial crystallography data: alternatives to the binning and averaging approach

Time binning approaches have proved successful in dealing with the issues of data incompleteness and partiality in time-resolved serial crystallography and produced molecular movies that offer unprecedented insight into the function of biomolecules in action. Yet the averaging procedure that underlies such approaches can imply a loss of timing information.

While the data manifold of the time-evolving system is intrinsically one-dimensional—and its trajectory likely explores only a low-dimensional subspace of the high-dimensional data space, the physical specificities of the diffraction experiment introduce data incompleteness and partiality, artificially increasing the apparent dimension of the subspace in which the data points lie. Time-lagged embedding mitigates such issues and allows—at least in favourable situations—to reconstruct the underlying system dynamics by singular value decomposition in the supervector space (singular spectrum analysis). Nonetheless this procedure is impractical in the case of large data sets, and in some circumstances may not be able to provide a low-rank approximation of the system dynamics [1]. Time-lagged embedding can be combined with data filtering in supervector space to ease these issues. This can be done—for instance, by using a set of orthonormal trigonometric functions as subspace basis, or alternatively a data-driven set of basis functions. The latter approach is called the nonlinear Laplacian spectral analysis (NLSA) [1], which was first applied to serial crystallography in [2].

These concepts are exemplified using synthetic models and preliminary serial crystallography results from the membrane protein bacteriorhodopsin in the first ps after photoactivation.

[1] D. Giannakis and A. J. Majda, *PNAS* **109**, 2222 (2012), doi: 10.1073/pnas.1118984109
[2] A. Hosseinizadeh *et al.*, *Nature* **599**, 697 (2021), doi: 10.1038/s41586-021-04050-9

Host: Robin Santra – CFEL-DESY Theory Division