Slow collective dynamics via data-driven approximation of the Koopman generator

Feliks Nüske[1]

[1] *Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany*

The determination of low-dimensional collective variables (CVs) is a critical problem in various application areas, including molecular dynamics simulations. However, finding good CVs, for example in the sense of retaining slow dynamical time scales of the full system, is highly non-trivial and requires the use of state-of-the art data analysis techniques.
In this talk, I will present recent theoretical and algorithmic progress on the determination of CVs and their associated effective dynamics using the projection formalism for stochastic dynamics as developed by Legoll and Lelièvre in [2]. In [6], it was shown that the projection formalism amounts to applying an orthogonal projection to the Koopman generator of the full system. The generator is projected onto the (still infinite-dimensional) space of functions which only depend on the CV space. The first result I will show is an error estimate comparing the slow time scales of the projected system to those of the full one. We arrive at a Galerkin-type estimate, bounding the time scale error in terms of the projection error for dominant eigenfunctions [4].
Then, I will move on to show how a data-driven matrix approximation of the projected generator can be obtained by a technique called generator extended dynamic mode decomposition (gEDMD) [1]. This method can also be formulated if a tensor product basis is used as Galerkin subspace [3], enabling the use of rich approximation spaces and potentially high-dimensional CVs. Finally, I will discuss bounds for the estimation error of the gEDMD method in terms of the amount of simulation data used to learn the Koopman generator [5].

References

[1] S. Klus, F. Nüske, S. Peitz, J. H. Niemann, C. Clementi, and C. Schütte. Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. *Physica D*, 406:132416, 2020.

[2] F. Legoll and T. Lelièvre. *Nonlinearity*, 23(9):2131, 2010.

[3] M. Lücke and F. Nüske. tgEDMD: Approximation of the Kolmogorov Operator in Tensor Train Format. *arXiv:2111.09606*, 2021.

[4] F. Nüske, P. Koltai, L. Boninsegna, and C. Clementi. Spectral properties of effective dynamics from conditional expectations. *Entropy*, 23(2):134, 2021.

[5] F. Nüske, S. Peitz, F. Philipp, M. Schaller, and K. Worthmann. Finite-data error bounds for koopman-based prediction and control. *arXiv:2108.07102*, 2021.

[6] W. Zhang, C. Hartmann, and C. Schütte. *Faraday Discuss.*, 195:365–394, 2016.