Mechanistic inference in complex systems

Internship offer, w/ Serge Dmitrieff, Institut Jacques Monod

In complex systems, comparing experiments to theory often implying aligning experimental results along a *master curve*. This is achieved by performing various linear or non-linear operations on the experimental parameters in order to plot the experimental measurement as a function of a single, complex effective parameter. This implies that there exists a well defined function mapping experimental parameters to a an observable.

In practice, plotting these master curves simply having a theoretical expectation. Sometimes, the parameter combination is simple enough that a master curve can be guessed or built by the physicist intuition. However, there are cases where the number of parameters is too large, or the function to complicated to be properly intuited. In this case, machine learning tools can be used to infer a mapping function from the experimental data. However this does not provide us with additional understanding of the physical system.

Project

The goal of this project is to obtain mechanistic understanding from inferred master curves. For this we will construct an architecture similar to an auto-encoder to identify the latent space, i.e. the minimal-dimension space required to account for experimental observations. Each dimension should correspond to an effective parameter combination.

We will then employ sensitivity analysis and clustering techniques to dissect the parameter combination giving rise to this latent space, i.e. find how various parameters take part in the effective combinations. The goal it to come up with a first proof-of-concept tool to show that complex parameter combinations can thus be inferred. We will start with a known system to compare our finding to a ground truth.

This project would be extendable to a PhD project by applying it biophysical simulations. In addition, extension to probabilistic formalism is possible, using variational auto-encoders. In this case, an analytical approach using information theory could also be considered.

Methods

The student should be interested in (physics-informed) machine learning and in the mathematics of dimensionality reduction. The student should be proficient in Python to build a machine learning pipeline and analysis scripts.

Mentoring

The student will be mentored by Serge Dmitrieff (*Cellular Spatial Organization*, Institut Jacques Monod, Paris), a theoretical physicist developing analytical and numerical tools applied to cell mechanics and data analysis : `www.biophys.info`. The team "Cellular Spatial Organization" is an interdisciplinary team that hosts both theoreticians and experimentalists.

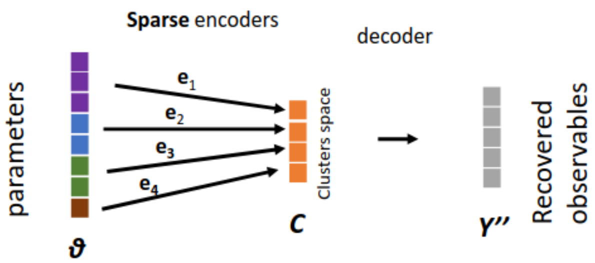

Figure 1: Schematic of an encoder/decoder pair to recover effective parameter combinations.