Model inference in complex systems

Internship offer, w/ Serge Dmitrieff, Institut Jacques Monod

A striking feature of physical systems is the appearence of universal behaviours at large scales. For instance, the order parameter near a phase transition usually follows a power low of the reduced temperature defined by a critical exponent ; this critical exponent is usually shared by a large class of systems, independently of their microscopic details. As an other example, simple visco-elastic models are very useful in describing a wide variety of materials of vastly different compositions. This highlights that **the dimensionality of macroscopic observables is very small** compared to the microscopic dimensionality.

However, microscopic complexity cannot be bypassed that easily. For instance, experimental perturbations of complex systems are often microscopic, e.g changes in compositions. Numerical simulations also require microscopic parameters to predict the evolution of a system. In complex systems, there is usually no simple mapping between microscopic parameters and large-scale observables. Because of that, macroscopic models often cannot predict quantitatively the consequence of a microscopic change.

Understanding the **mapping between microscopic parameters and macroscopic behaviour** is thus a key challenge in complex systems. One of the main hurdles in the challenge is the high number of microscopic parameters. First, the cost of performing experiments or simulations for all values of each parameter becomes very high. Second, the visualization of results becomes increasingly hard as the number of parameters increases, and information contained in the data will be lost to us.

Project

The goal of this project is to develop a new approach to recover as much information as possible from experimental or simulated data sets with a large parameter space. For this, we need to :

- Identify the system's underlying dimensionality
- Identify groupings of highly correlated parameters

To achieve this, we will develop a specialized framework derived from *v*ariational auto-encoders, a machine-learning framework designed for dimensionality reduction and feature identification. A proof-of-concept program already exists, showing promising results on a minimal example. Strickingly, this tool is very frugal, costing less than a typical simulation run.

In this internship, we aim at going further than a proof-of-concept. After demonstrating its theoretical soundness,the student will validate this framework on toy models, before applying it to existing simulation results.

Methods

The student should be interested in dimensionality reduction, and its application to (bio-) physical systems. Python scripting is required to modify and generalize the existing Python proof-of-concept program. Some knowledge of Information Theory and/or Machine Learning methods is highly recommended.

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.