Understanding Turbulence via Machine Learning

Sergio Chibbaro[1], Cyril Furtlehner[1,2], Christophe Josserand[3],
[1] LISN Université Paris-Saclay [2] INRIA, TAU team, Université Paris-Saclay [3] LadHyX, IPP
*

Context In the last decade, machine learning (ML), and more specifically deep neural networks (DNN), have thoroughly renewed the research perspectives in many fields. In this work we would like to assess the capability of ML to give insights on the physical structure of the cascade process in turbulence.

Project The most notable feature of turbulence is the presence of a cascade, as a prototypical of a strong out-of-equilibrium steady statistical process, as firstly analysed by Kolmogorov, Onsager and Heisenberg. The problem is known to be described by hydrodynamic Navier-Stokes equations [1], which are however extremely complex and have not allowed any successful first-principle approach. We consider here a dynamical system model of turbulence, namely the shell model of turbulence. The problem is described by the following set of equations:

$$\frac{du_n}{dt} = i(ak_{n+1}u_{n+2}u^*_{n+1} + bk_nu_{n+1}u^*_{n-1} - ck_{n-1}u_{n-1}u_{n-2}) - \nu k_n^2 u_n + f_n \qquad (1)$$

for $n = 1, 2, 3, ...$, u_n being a Fourier component of the velocity fied, associated to the wave numbers k_n. These are taken to be $k_n = k_0\lambda^n$, with $\lambda > 1$ being the shell spacing parameter, and $k_0 > 0$. In order for the shell model to be a system of the hydrodynamic type we require that in the inviscid ($\nu = 0$) and unforced ($f_n = 0$) case the model will have at least one quadratic invariant, that is the energy. Despite the simplicity of the model with respect to the original Navier-Stokes equations, Shell models reproduce many of the intrinsic feature of turbulent cascade, notably the energy spectrum and intermittency, that is extreme event statistics.

Learning an effective shell model: The question then is whether it is possible to learn an effective shell model

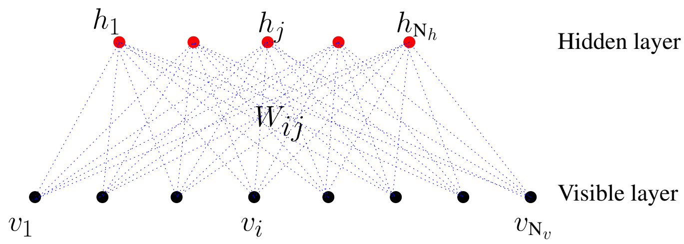

Figure 1: bipartite structure of the RBM.

from Navier-Stokes turbulent flow. The dependence between the various modes can be recovered in principle by the generalized linear repsonse theory once we are given a "good" joint-probability stationary distribution of the modes. The first simple model to check would be to learn a multi-variate Gaussian distribution with sparse inverse covariance matrix, which would correspond to use the linear response theory. To go beyond the linear approximation, we could consider as a possible candidate the restricted Boltzmann machine (see e.g. [3]), which is a probabilistic model of the form

$$P(\mathbf{v}, \mathbf{h}) = P_{\text{prior}}(\mathbf{v})P_{\text{prior}}(\mathbf{h}) \exp\Big(\sum_{ij} v_i W_{ij} h_j\Big)$$

defined as a two layers neural network, the first one corresponding to the visible variables we want to model – here the u_n at different time-steps – the second layer consisting of hidden variables (never observed) being there to build the correct dependencies between the visible variables via the weight matrix W. Once the model is learned, thanks to some systematic expansion of the energy function w.r.t. entries of W it is then in principle possible to extract the dominant couplings, and thus an effective model. Another possibility is to directly compute the generalize response function of the model. If successful we could then see whether an effective shell model can be extracted from complete Navier-Stokes turbulent flows. The project is interdisciplinary with an interplay of applied mathematics, statistical physics, fluid mechanics and informatics. Depending on the skills of the candidate, different tracks can be explored.

*mail: sergio.chibbaro@universite-paris-saclay.fr, cyril.furtlehner@inria.fr, christophe.josserand@ladhyx.polytechnique.fr