SORBONNE UNIVERSITÉ

Ecole doctorale SMAER

Sciences **M**écaniques, **A**coustique, **E**lectronique, **R**obotique

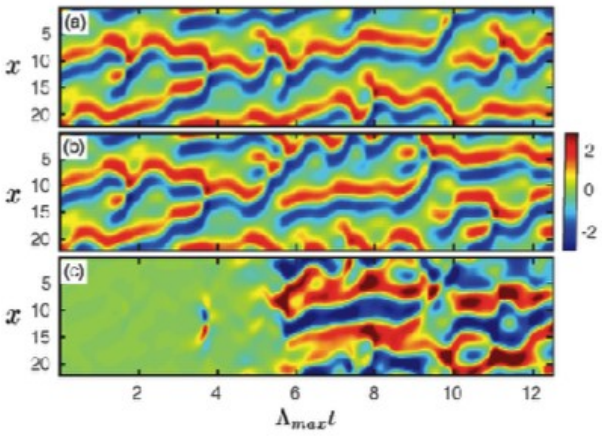
Goal-oriented physics-constrained machine learning for chaotic systems

The wide availability of data from everyone's numeric life (private photos, posts on social media, travel logs, etc.) and ubiquitous inexpensive sensors have triggered a surge of interest in data-driven techniques for exploiting this massive amount of data. Computational fluid dynamics relies on massive numerical simulations and generates a considerable amount of data. The rise of machine learning techniques has led to major efforts to enhance, and sometimes substitute, physics solvers. Moreover, new strategies in optimization and control could be developed along this promising way. Of particular interest for scientific purposes is the modeling of the behavior of a system from limited observations. In particular, one often needs to determine faithful and reliable governing equations accurately describing the time-evolution of the system under consideration.

Learning from data relies on the following steps:
- Fit each model from a given class of models on the available training data,
- Assess the quality of each model on a separate set of test data obtained from the same system as the training data,
- Retain the model leading to the lowest test error.

The class of models generically described as "neural networks" has seen a spectacular renewed interest in the last decade and has now become the gold standard in many applications. In particular, deep networks, involving anywhere from 5 to 1000+ hidden layers of neurons, have demonstrated impressive performance.

A similar framework has recently been used with a deep neural network architecture (Reservoir Computing) to learn a model for the chaotic Kuramoto-Sivashinsky system, [2, 3]. The resulting model is able to accurately predict the future state of the system up to about 6 Lyapunov exponent times. Similar results were obtained by [4] and our team on a similar case [1], as shown in Figure 1.



ED SMAER (ED391)
Tour 45-46 Bureau 205- case courrier 270- 4, place Jussieu - 75252 PARIS Cedex 05
: 01 44 27 40 71
charlotte.vallin@sorbonne-universite..fr

Figure 1: Prediction of KS. Top: direct numerical simulation. Center: Simulation obtained from the machine learning. Bottom: Difference of the two fields.

This remarkable result rises several largely open questions, also prototypical of a large class of learning situations:

- What are the reasons why this deep neural computing allows such a performance, well beyond the usual Lyapunov time limit ? How specific is the performance to the particular architecture of this network ?
- In a broad sense, what is the best quantity to learn for achieving good general performance? More precisely, what is the good metric to use as the training criterion to improve generalizability of the model learned ? In a dual view, what is a relevant nonlinear transform to pre-process the data so that it achieves a good validation error ?
- In many situations, expertise knowledge is available on the system under consideration. How to make use of this expertise and reduce the dimension of the learning space by introducing relevant constraints such as symmetries, invariants, causality, etc.?
- In situations of active learning when the samples can be chosen, how to benefit from past knowledge to improve the sampling scheme? How to adapt the design of experiment and explore the system in an efficient way, given the objective function?

The objectives of this thesis will be to improve the learning process of a data-driven approach by addressing the questions and limitations above, with particular attention to possible developments for engineering applications.

Specifically, different points will be addressed, including but not limited to:

1. Choice of the best cost function to get optimal results. Cost functions are typically defined in terms of integral (L2) norm of the misfit residual. This choice is not always necessary and one may want to employ other metrics. We would like to use a physics-based approach to decide which feature deserves particular attention. For instance, one may need to ensure a good learning in terms of Fourier spectrum, Lyapunov exponents, wavelet structure, optimal transport-related norms (e.g., Wasserstein distance), etc. We will first analyze how a standard L2 learning is relevant for other metrics. In a second step, these alternative metrics will be directly employed as objective functions in a goal-oriented approach.

2. Many key questions will be at the center of the investigation:
   How to sample the system under consideration in the most efficient way? How to most quickly discover its natural measure and adapt the sampling scheme accordingly? How to explore around the neutral manifold (when it exists) in useful directions? What is the minimal number of samples one needs to learn a model of an attractor of a given dimension?

From a technical point of view, the work will be first conducted on relatively low-dimensional dynamical systems such as Kuramoto-Sivashinky to develop the methodology. The simulation codes are already available and parallelized on GPUs.
In a second step, more complex systems such as fluid flows in both laminar and turbulent regimes will be considered.

This project is strongly multidisciplinary (applied maths, statistics, physics, fluid dynamics, data analysis, high- performance computing applied to engineering), at the edge of fundamental and applied research.
The successful candidate has a strong background in at least one of the following fields: statistics, applied mathematics, machine learning. He/She is also an interdisciplinary, team-oriented, person with good communication and writing skills.

References

[1] Bucci, M. A., Semeraro, O., Allauzen, A., Cordier, L., Wisniewski, G., & Mathelin, L. (2018). Control-oriented model learning with a recurrent neural network. *Bulletin of the American Physical Society*.

[2] Pathak, J., Lu, Z., Hunt, B. R., Girvan, M., & Ott, E. (2017). Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *27*(12), 121102.

[3] Pathak, J., Hunt, B., Girvan, M., Lu, Z., & Ott, E. (2018). Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Physical review letters*, *120*(2), 024102.

[4] Vlachas, P. R., Byeon, W., Wan, Z. Y., Sapsis, T. P., & Koumoutsakos, P. (2018). Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *474*(2213), 20170844.