# Partial differential equation discovery for spatio-temporal simulations in cells

Supervision
- Hugues Berry, http://www.inrialpes.fr/Berry/
- Audrey Denizot, https://adenizot.github.io/
- Thomas Guyet, http://thomas.guyet.info/

Research Team: AIstroSight / Inria Lyon

Location: Inria, Centre de Lyon, 56 Boulevard Niels Bohr, 69100 Villeurbanne

Starting: November, 2023

Application deadline: by June 26th, 2023 (*application details below*)

Keywords: Biochemical Processes, Computational Biology, Dynamical systems, Data-driven modeling, Equation Discovery, Nonlinear Partial Differential Equations, Reaction-Diffusion Simulations, Stochastic processes

## Background

Brain diseases emerge from perturbations of biochemical processes, such as local variations in the concentration of some molecules. Identifying potential treatments (*i.e.* active drugs) thus requires to understand the sources of dysfunctions at the microscopic scale, which is challenging. First, biological systems are highly complex: they are nonlinear, include unobserved variables, their dynamics occurs on multiple interdependent spatial and temporal scales, and the physical principles that govern their dynamics are partly unknown. Second, data are scarce and some parameter spaces, notably for processes occurring in nanoscopic sub-cellular domains, cannot be reached experimentally. In this context, mechanistic spatially-extended stochastic models are essential to capture the complex chemical interactions that underlie the functioning of brain cells. Notably, they are useful to simulate the mechanism of action of a candidate drug, *i.e.* its interactions with its cellular targets and its impact on cell dynamics. As this approach is resource and time demanding, it cannot be used to simulate brain dynamics at larger spatio-temporal scales, thus hindering its ability to predict whether the drug of interest is likely to treat the disease. A tool that links such fine-grained microscopic models of drug-cell interactions with brain dynamics at higher temporal and spatial scales is thus of high interest in the search for therapies for brain disorders, yet is currently lacking.

Recently, data-driven machine learning approaches have emerged as tools of interest to address this challenge. These include physics-informed neural networks [1,2], observation-based model parametrization methods [3] or data-driven equation discovery [4-9]. The latter aims at establishing a mathematical model, expressed in the form of partial/ordinary differential equations, that describes an observed system that is constrained by hidden physical laws. By identifying both the structure of the equations and parameter values, this approach has the advantage of being explanatory, providing a physical interpretation of the mechanisms that govern the dynamics of the modeled system. This is one of the major assets of this technique compared to approaches that provide descriptive black-box models. Representative equation discovery algorithms include DeepMoD [4], D-CODE [5], D-CIPHER [6], SINDy [7], PDE-Find [8] or WSINDy [9].

The main limitation of the state-of-the-art equation discovery methods is the strong assumption of a uniform field on the whole domain (spatially and temporally). However, a cell is a biological system

with localized features, where molecular processes are affected by the topology of the cell itself. For example, interactions at the center of the cell differ from those that occur close to the cell membrane because of distinct diffusion constraints. Local sites of inter-cellular communication can also result in heterogeneous distributions of molecules within the cell (*e.g.* “injection” sites). These spatial inhomogeneties (*e.g.* variability of cell shape, border conditions and punctual sources) are thus important to take into account to model the internal functioning of cells. There is thus a need for equation discovery methods that capture the complex spatial heterogeneities inherent to biological systems.

Objectives, challenges & novelty

The goal of this PhD project is to develop a data-driven partial differential equation (PDE) discovery method for complex dynamical systems such as brain cell. The algorithm will be evaluated on its ability to robustly and accurately learn cell function at the macroscopic scale from data simulated at the nanoscopic level. The simulated data will be obtained from agent-based models of biochemical reaction kinetics developed in the team (such as [10]), that account for the intracellular movements of individual molecules and ions, their interactions, as well as their modulation by cellular spatial heterogeneity. These simulators are partly stochastic (initial conditions, Brownian motion, probabilistic reactions upon encounter) such that the same parameter set can be used to generate large collections of data.

As those models are not based on differential equations, a major challenge of this project will be to evaluate the discovered equations, as the ground-truth equations are only known for a few simple biochemical reactions. For more complex biological processes, the discovered equations will be evaluated based on their ability to predict the dynamics of the system accurately.

The novelty of this project compared to the state-of-the-art methods is that the algorithm will learn from *in silico* data. This is a strong asset as it provides more flexibility to generate complex datasets with various spatio-temporal properties and lifts constraints on data availability.

This project, by providing innovative numerical approaches to learn explanatory mathematical models of brain cell dynamics, is an important milestone in the global effort to build digital twins for neuroscience.

Candidate profile

We are looking for a student with a Master degree who has experience in at least one of the following areas: data science/machine learning/mathematical modeling as well as an interest in cell biology/neuroscience.

Proficiency in written and oral English is required. No knowledge of French is needed.

Most importantly, we are looking for future colleagues who are eager to learn and grow, and who are driven by scientific curiosity.

How to apply

Applicants should submit the following documents to the team of supervisors by June 26th: :

- A CV
- A one-page statement presenting your interests and goals
- The contact information of at least 2 referees

Candidates are welcome to contact the project supervisors if they have questions regarding this position.

Contact information of the supervisors :
Hugues Berry: hugues.berry@inria.fr
Audrey Denizot: audrey.denizot@inria.fr

Thomas Guyet: thomas.guyet@inria.fr

References

[1] Yin, Y., Le Guen, V., Dona, J., de Bézenac, E., Ayed, I., Thome, N., & Gallinari, P. (2021). Augmenting physical models with deep networks for complex dynamics forecasting. *Journal of Statistical Mechanics: Theory and Experiment*, *2021*(12), 124012.
[2] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, *378*, 686-707.
[3] Peng, G. C., Alber, M., Buganza Tepole, A., Cannon, W. R., De, S., Dura-Bernal, S., ... & Kuhl, E. (2021). Multiscale modeling meets machine learning: What can we learn?. Archives of Computational Methods in Engineering, 28, 1017-1037.
[4] Both, G. J., Choudhury, S., Sens, P., & Kusters, R. (2021). DeepMoD: Deep learning for model discovery in noisy data. *Journal of Computational Physics*, *428*, 109985.
[5] Qian, Z., Kacprzyk, K., & van der Schaar, M. (2021, September). D-CODE: Discovering Closed-form ODEs from Observed Trajectories. In *International Conference on Learning Representations*.
[6] Kacprzyk, K., Qian, Z., & van der Schaar, M. (2022). D-CIPHER: Discovery of Closed-form PDEs. *arXiv preprint arXiv:2206.10586*.
[7] Champion, K., Lusch, B., Kutz, J. N., & Brunton, S. L. (2019). Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, *116*(45), 22445-22451.
[8] Rudy, Samuel H., et al. (2017). Data-driven discovery of partial differential equations. Science Advances 3.4: e1602614
[9] Messenger, D. A., & Bortz, D. M. (2021). Weak SINDy for partial differential equations. Journal of Computational Physics, 443, 110525.
[10] Denizot, A., Arizono, M., Nagerl, V. U., Soula, H. and Berry, H., (2019). Simulation of calcium signaling in fine astrocytic processes: effect of spatial properties on spontaneous activity. PLoS Computational Biology,15(8):e1006795