SOE 8: Machine Learning in Dynamics and Statistical Physics II (joint session DY/SOE)

Time: Tuesday 9:30–13:00 Location: BH-N 243

SOE 8.1 Tue 9:30 BH-N 243

Pareto-Based Selection of Data-Driven Ordinary Differential Equations — •Gianmarco Ducci, Karsten Reuter, and Christoph Scheurer — Fritz-Haber-Institut der MPG, Berlin

Data-driven approaches enable the approximation of governing laws of physical processes with parsimonious equations. However, they face challenges due to inherent noise in data, which impacts the sparsity of the result. While a great effort over the last decade has been made in this field, data-driven approaches generally rely on the paradigm of imposing a fixed base of library functions. In order to promote sparsity, finding the optimal set of basis functions is a necessary condition but a challenging task to guess in advance.

In this work, we propose an alternative approach which consists of optimizing the very library of functions while imposing sparsity. The robustness of our results is not only evaluated by the quality of the fit of the discovered model, but also by the statistical distribution of the residuals with respect to the original noise in the data. The model selection is then chosen from a subset of optimal models obtained in a Pareto fashion. We illustrate how this method can be used as a tool to derive microkinetic equations from experimental data.

SOE 8.2 Tue 9:45 BH-N 243

Accurate Memory Kernel Extraction from Discretized Time-Series Data — •Lucas Tepper — Department of Physics, Freie Universität Berlin

Memory effects emerge whenever the dynamics of complex many-body systems are projected onto low-dimensional observables. Accounting for memory effects using the framework of the generalized Langevin equation (GLE) has proven efficient, accurate and insightful, particularly when working with high-resolution time series data. However, in experimental systems, high-resolution data is often unavailable, raising questions about the effect of the data resolution on the estimated GLE parameters. Using molecular dynamics (MD) data of a small, alpha-helix-forming peptide, I demonstrate that the direct memory extraction remains accurate when the discretization time is below the memory time. For discretization times exceeding the memory time, I show that a Gaussian Process Optimization (GPO) scheme estimates accurate memory kernels by minimizing the deviation of discretized two-point correlation functions between MD and GLE simulations. The GPO scheme stays accurate for discretization times below the longest time scale in the data, typically the barrier crossing time.

SOE 8.3 Tue 10:00 BH-N 243

Coarse-graining non-equilibrium systems with machine learning: from conceptual challenges to new approaches — •Patrick Egenlauf[1,2] and Miriam Klopotek[2] — [1]University of Stuttgart, Interchange Forum for Reflecting on Intelligent Systems, IRIS3D project, Stuttgart, Germany — [2]University of Stuttgart, Stuttgart Center for Simulation Science, SimTech Cluster of Excellence EXC 2075, Stuttgart, Germany

Machine learning (ML) was previously shown to effectively coarse-grain configurations of many-body systems. We want to investigate ML applications to address the dynamic coarse-graining of non-equilibrium many-body systems. Our research aims to advance ML methods while avoiding conventional assumptions. The focus is on time-dependent datasets and their broader implications for understanding causality. We introduce innovative techniques by incorporating general theory, including the time-dependent generalized Langevin equation [1], for building and interpreting time-dependent learning techniques [2]. This provides a distinctive ML perspective that extends to various applications for dynamical systems beyond equilibrium states. This study offers new ways to improve our understanding and manipulation of complex non-equilibrium many-body dynamics using ML.

[1] Schilling, T. (2022). Coarse-grained modelling out of equilibrium. *Physics Reports*, 972, 1-45.
[2] Nakajima, K., and Fischer, I. (2021). *Reservoir Computing*. Springer Singapore.

SOE 8.4 Tue 10:15 BH-N 243

Statistical criteria for the prediction of dynamical clustering in granular gases — •Sai Preetham Sata[1,2], Dmitry Puzyrev[2], and Ralf Stannarius[1,2] — [1]Institute of Physics, Otto-von-Guericke University, Magdeburg, Germany — [2]Department of Microgravity and Translational Regenerative Medicine and MARS, Otto von Guericke University, Magdeburg, Germany

Granular gases excited by external forces can undergo transitions from the homogeneous to a dynamical cluster state [1, 2], depending on filling fraction, excitation parameters and container geometry. We compare two statistical criteria for the clustering transition, viz. the Kolmogorov-Smirnov Test (KS-Test) on the particle number density profile and the so-called caging-effect based on the local packing fraction [2]. Both criteria are evaluated for various combinations of system parameters in the VIP-Gran experiment [3] and combined into one dataset. This allows us to compare existing clustering criteria and tune them to provide matching clustering thresholds. The aim is to develop improved threshold criteria. Machine learning models are trained with this dataset to predict whether particular parameters lead to homogeneous or dynamical cluster states.

This study is supported by DLR projects VICKI and EVA II(50WM2252 and 50WK2348)

References: [1] É. Falcon et al., Phys. Rev. Lett., 83:440-443, 1999 [2] E. Opsomer et al., Europhys. Lett., 99:40001, 2012 [3] S. Aumaître et al., Rev. Sci. Instr., 89, 2018.

SOE 8.5 Tue 10:30 BH-N 243

Excitability and Memory in a Time-Delayed Optoelectronic Neuron — •Jonas Mayer Martins[1], Svetlana V. Gurevich[1], and Julien Javaloyes[2] — [1]Institute for Theoretical Physics, University of Münster, Wilhelm-Klemm-Str. 9 and Center for Nonlinear Science (CeNoS), University of Münster, Corrensstrasse 2, 48149 Münster, Germany — [2]Departament de Física and IAC-3, Universitat de les Illes Balears, C/ Valldemossa km 7.5, 07122 Mallorca

We study the dynamics of an optoelectronic circuit composed of a nanoscale resonant-tunneling diode (RTD) in the excitable regime driving a nanolaser diode (LD) coupled via time-delayed feedback. Using a combination of numerical path-continuation methods and time simulations, we demonstrate that the RTD-LD system can serve as an artificial neuron, generating pulses in the form of temporal localized states (TLSs) that can be employed as memory for neuromorphic computing. In particular, our findings reveal that the prototypical delayed FitzHugh–Nagumo model previously employed to model the RTD-LD resembles our more realistic model qualitatively only in the limit of a slow RTD. We show that the RTD time scale plays a critical role in how the RTD-LD can be used as memory because it governs a shift in pulse interaction forces from repulsive to attractive, leading to a transition from stable to unstable multi-pulse TLSs. Our theoretical analysis uncovers novel features and challenges, including the multistability of TLSs and attractive interaction forces, stemming from the previously neglected intrinsic dynamics of the laser. These dynamics are crucial to consider for the memory properties of the RTD-LD.

SOE 8.6 Tue 10:45 BH-N 243

Anisotropic diffusion analysis in confined geometries — •Kevin Höllring[1], Andreas Baer[1], Nataša Vučemilović-Alagić[2], David M. Smith[2], and Ana-Sunčana Smith[1,2] — [1]PULS Group, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), 91058 Erlangen, Germany — [2]Group of Computational Life Sciences, Ruđer Bošković Institute, 10000 Zagreb, Croatia

In various systems, liquid and particle transport are of major importance to the viability of chemical like catalysis or adsorption. Most of these systems involve interfaces and confined geometries, where the prerequesites for the application of classical analysis techniques like the Einstein/MSD or the Green-Kubo/ACF approach are not fulfilled. To facilitate the resolution of diffusion coefficients in such system, we propose a novel approach built around the analysis of time statistics of particles in subspaces of the system using the Smoluchowski equation. For simple point-like particles, we propose a explicit analytic formula to link mean lifetimes and diffusivity, with an extension to account for the impact of locally induced drift as a consequence of complex effective interaction potentials. For more complex particles like in ionic liquids, we provide an extended technique able to resolve the impact of internal degrees of freedom, through which we can not only analyze the evolution of transport but also characteristic changes

in the conformational behavior of particles close to interfaces. Overall, this approach can be used to demonstrate a characteristic oscillatory behavior of particle diffusivity in confinement close to interfaces so far not reported in literature.

SOE 8.7 Tue 11:00 BH-N 243

Data assimilation of cardiac dynamics by means of adjoint optimization — •Inga Kottlarz[1,2,3,4], Sebastian Herzog[2,4,5], Patrick Vogt[2,3], Stefan Luther[1,2,4], and Ulrich Parlitz[2,3,4] — [1]Institute for Pharmacology and Toxicology, UMG Göttingen, Germany — [2]MPI for Dynamics and Self-Organization, Göttingen, Germany — [3]Institute for the Dynamics of Complex Systems, University of Göttingen, Germany — [4]German Center for Cardiovascular Research, Partner Site Niedersachsen, Göttingen, Germany — [5]III. Institute of Physics, University of Göttingen, Germany

Cardiac muscle tissue is an excitable medium that can exhibit a range of dynamics of different complexity, from planar waves to spiral waves to spatiotemporal chaos, the latter being associated with (fatal) cardiac arrhythmia.

Both the prediction of such high dimensional chaotic time series, as well as the reconstruction of their (not yet fully observable) complete dynamical state are ongoing challenges. In recent years, machine learning approaches have gained popularity for solving these problems, which can be advantageous if we do not have much knowledge about the dynamical system in question, but are limited by the large amounts of training data that is needed and often not available for biological systems. We present adoptODE, an adjoint optimization framework for estimating model parameters and unobserved variables. We showcase the adjoint method*s effectiveness in optimizing high-dimensional problems with thousands of unknowns, serving as a valuable tool for bridging the gap between empirical data and theoretical models.

15 min. break

SOE 8.8 Tue 11:30 BH-N 243

Collective Variables for Neural Networks — •Konstantin Nikolaou[1], Samuel Tovey[1], Sven Krippendorf[2], and Christian Holm[1] — [1]Institute for Computational Physics, University of Stuttgart, Germany — [2]Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität, Germany

Neural Networks have witnessed extensive integration across diverse domains within physics. However, our focus shifts towards the inverse problem: How can neural networks benefit from physics? Learning with a neural network involves algorithmically assimilating information into a model. Nevertheless, the process of neural learning remains largely elusive, given the challenge of understanding how to extract information from its dynamics. Analogous to dynamical systems in statistical physics, describing neural network training involves an extensive number of degrees of freedom, which appears to benefit from a description through macroscopic quantities. To that end, we introduce Collective Variables for neural networks, tracing out microscopic degrees of freedom to describe and analyze the learning process at every stage. We investigate the initial state as well as the learning dynamics of the network in the context of the Collective Variables. We find a correlation between the initial network state and the generalization of the model computed after training. Moreover, we use the collective variables to identify and analyze stages arising in the dynamics of the learning process.

SOE 8.9 Tue 11:45 BH-N 243

Fluctuating weight dynamics and loss landscape in deep linear networks — •Markus Gross — DLR, Institute for AI Safety and Security, Germany

Understanding how weights fluctuate during training of neural networks and how this impacts the loss landscape is key to optimizing training processes and performance. We investigate this dynamics in (deep) linear networks within the continuum limit of stochastic gradient descent. For a two-layer network, we highlight the role of the inter-layer coupling and analytically derive from first principles a recently discovered key relation between weight fluctuations and loss landscape. The uncovered behaviors are rooted in general statistical properties of the network architecture und training data.

Reference: https://arxiv.org/abs/2311.14120

SOE 8.10 Tue 12:00 BH-N 243

Loss is More: Exploring the weight space of a perceptron via enhanced sampling techniques — •Margherita Mele[1], Roberto Menichetti[1], Alessandro Ingrosso[2], and Raffaello Potestio[1] — [1]Physics Department, University of Trento, via Sommarive, 14 I-38123 Trento, Italy — [2]The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy

Understanding how input data properties influence the learning process in artificial networks is crucial. The assumption of Gaussian i.i.d. inputs has long been foundational, yet questioning its constraints is now essential. Our approach utilises enhanced sampling methods from soft matter physics to exhaustively explore the loss profile and reconstruct the density of states of networks with discrete weights, addressing optimization in highly rugged landscapes even in simple architectures. These methods, effective in real datasets, enable exploration of data dimensionality and structure impact.

Employing benchmarks (e.g. MNIST, FashionMNIST) and in silico datasets, our study investigates the role of various input-data properties, including class imbalance, separation, item mislabelled, and input-output correlation. Our findings bridge theoretical and applied aspects, shedding light on the limitations and extensions of Gaussian i.i.d. assumptions. This work provides pivotal insights into the interplay between input data properties and network learning, advancing our understanding of how artificial networks adapt to different information contexts.

SOE 8.11 Tue 12:15 BH-N 243

Emergent oscillating dimensionality transformations in deep learning — •Pascal de Jong, Felix J. Meigel, and Steffen Rulands — Arnold Sommerfeld Center for Theoretical Physics, Department of Physics, Ludwig- Maximilians-Universität, München, Germany

Artificial intelligence relies on deep neural networks (DNNs), which comprise a large set of nonlinear nodes connected by weights. The functioning of DNNs and their ability to generalize to unseen data are examples of complex behavior, which due to their highly nonlinear nature is poorly understood. Here, we show that training DNNs universally leads to oscillating weight topologies that alter the embedding dimensions of hidden data representations in different layers. Specifically, using a path representation of DNNs, we derive equations for the time evolution of the weights. We show that training leads to a structure, in which weights are focused on a subset of nodes, and the degree of focusing oscillates across layers. We empirically confirm these findings by studying the training dynamics of large DNNs on different data sets. Finally, we show that these structures imply a repeated dimensional decrease and increase of the hidden data representations. Our results highlight that emergent dynamics during training can lead to universal network topologies with implications for their function.

SOE 8.12 Tue 12:30 BH-N 243

Investigating the Evolution of Fisher Information for Neural Network Dynamics — •Marc Sauter, Samuel Tovey, Konstantin Nikolaou, and Christian Holm — Institute for computational physics, Stuttgart, Germany

Machine learning has proven to be a powerful tool with remarkable effectiveness for various physical applications. Even though Neural Networks offer great potential for solving complex tasks, their black-box nature limits the information obtainable about the underlying mechanisms they employ. Especially in physics, where the exact methodologies of solving tasks are just as important as finding solutions itself, advances in interpretable machine learning promise to aid the applicability of Neural Networks task-solving capabilities greatly. The Fisher Information Matrix (FIM) is a statistical measure known for its ability to identify second order phase transitions in physical systems. It can also be used for analyzing learning dynamics, where it expresses correlations between influences of parameters in Neural Networks. However, because of its size, computation of the FIM is currently intractable for Neural Networks used in common problem settings. As a way of obtaining parts of the information contained in the FIM, we introduce a novel mathematical relationship between the trace of the FIM and the Neural Tangent Kernel, a smaller observable of neural network training. We apply this approach for simple test models and discuss arising research topics.

SOE 8.13 Tue 12:45 BH-N 243

Near-zero-cost post-training uncertainties for deep learning architectures — •Filippo Bigi, Sanggyu Chong, Michele Ceriotti, and Federico Grasselli — Laboratory of Computational Science and Modeling (COSMO), IMX, École Polytechnique Fédérale

de Lausanne, Switzerland

Over the last decade, deep learning models have shown impressive performance and versatility on an extremely wide range of tasks. However, their probability estimates are unreliable, especially outside of the training distribution, with neural networks often returning overconfident results when queried on unfamiliar data. Although several uncertainty quantification schemes are available, their practical downsides hinder their widespread adoption. We propose a novel method for estimating the predictive uncertainties of deep learning architectures based on the interpretation of the last layer of neural networks as a linear Gaussian process. Contrary to previous methods, the proposed approach is simple, scalable, does not involve modification of the architecture or the training procedure, can be applied to trained models *a posteriori*, and generates uncertainty estimates with a single forward pass at negligible additional cost. We demonstrate the accuracy and practicality of our scheme on a wide range of machine learning datasets.