DY 33: Machine Learning in Dynamics and Statistical Physics I

Time: Thursday 9:30–13:00 Location: H47

DY 33.1 Thu 9:30 H47

Learning Mechanisms of Neural Scaling Laws — •Konstantin Nikolaou[1], Samuel Tovey[1], Sven Krippendorf[2], and Christian Holm[1] — [1]Institute for Computational Physics, University of Stuttgart, Germany — [2]Cavendish Laboratory and DAMTP University of Cambridge, United Kingdom, CB3 0WA

Recent works have identified neural scaling laws, which describe the trade-off between neural network performance and computation cost. Understanding the underlying mechanisms leading to scaling behavior might be one of the most important questions in current machine-learning research.

We compare the behavior of neural networks for data and model scaling by analyzing the learning dynamics through the lens of the neural tangent kernel. We find similar performance scaling in both regimes but uncover fundamentally distinct internal model mechanisms underlying the scaling. Additionally, we investigate scaling towards the infinite-width limit of neural networks and identify a transition, we coin the Feature-Kernel Transition, separating two regimes: Below, a model refines features to resolve a task, while above the transition the refinement declines and the initial state becomes the dominant factor. We argue that the transition marks the trade-off between model size and maximum feature learning.

DY 33.2 Thu 9:45 H47

Finite integration time drives optimal dynamic range into subcritical regime — Sahel Azizpour[1,2], Viola Priesemann[3,4], •Johannes Zierenberg[3,4], and Anna Levina[1,2] — [1]Eberhard Karls University of Tübingen, Germany — [2]Max Planck Institute for Biological Cybernetics, Tübingen, Germany — [3]Max Planck Institute for Dynamics and Self Organisation, Göttingen, Germany — [4]Institute for the Dynamics of Complex Systems, University of Göttingen, Germany

Sensitivity to small changes in the environment is crucial for many real-world tasks, enabling living and artificial systems to make correct behavioral decisions. It has been shown that such sensitivity is maximized when a system operates near the critical point of a second-order phase transition. However, proximity to criticality introduces large fluctuations and diverging timescales. Hence, it would require impractically long integration periods to leverage the maximal sensitivity. Here, we analytically and computationally demonstrate how the optimal tuning of a recurrent neural network is determined given a finite integration time. Rather than maximizing the theoretically available sensitivity, we find networks to attain different sensitivity depending on the time available. Consequently, the optimal dynamic regime shifts from critical to subcritical when integration times are finite, highlighting the necessity of incorporating finite-time considerations into studies of information processing.

DY 33.3 Thu 10:00 H47

Self-Organizing Global Computation from Local Objective Functions Based on Partial Information Decomposition — Andreas C. Schneider[1,2], •Valentin Neuhaus[2,1], David A. Ehrlich[3], Abdullah Makkeh[3], Alexander S. Ecker[4,2], Viola Priesemann[2,1], and Michael Wibral[3] — [1]Institute for the Dynamics of Complex Systems, University of Göttingen, Germany — [2]Max Planck Institute for Dynamics and Self Organisation, Göttingen, German — [3]Campus Institute for Dynamics of Biological Networks, University of Göttingen — [4]Institute of Computer Science and Campus Institute Data Science, University of Göttingen

In modern deep neural networks, individual neuron learning dynamics are often obscure due to global optimization. In contrast, biological systems use self-organized, local learning to achieve robustness and efficiency with limited global information. We propose a method for achieving self-organization in artificial neurons by defining local learning goals based on information theory. These goals leverage Partial Information Decomposition (PID), which breaks down information from sources into unique, redundant, and synergistic contributions. Our framework enables neurons to locally determine how input classes contribute to the output, expressed as a weighted sum of PID terms derived from intuition or numerical optimization. This approach enhances task-relevant local information processing and neuron-level interpretability while maintaining strong performance, providing a principled foundation for local learning strategies.

DY 33.4 Thu 10:15 H47

Explaining Near-Zero Hessian Eigenvalues Through Approximate Symmetries in Neural Networks — •Marcel Kühn and Bernd Rosenow — Institute for Theoretical Physics, University of Leipzig, 04103 Leipzig, Germany

The Hessian matrix, representing the second derivative of the loss function, offers crucial insights into the loss landscape of neural networks and significantly influences optimization algorithms, model design, and generalization in deep learning. A common characteristic of the Hessian eigenspectrum is the presence of a few large eigenvalues alongside a bulk of near-zero eigenvalues. We propose that this bulk structure arises from approximate symmetries inherent in network architectures – an often overlooked aspect. First, we demonstrate that in deep, fully connected linear networks, exact continuous symmetries that leave the loss invariant lead to zero eigenvalues in the Hessian. These zero eigenvalues and their corresponding eigenvectors can be attributed to symmetries such as rotations between weight layers. Extending this concept, we suggest that in networks with nonlinear activation functions, approximate symmetries introduce a large number of small but finite eigenvalues, viewed as perturbations of the linear case. We illustrate this phenomenon in a two-layer ReLU student-teacher setup and in a multi-layer network trained on CIFAR-10, showing that eigenvectors with small eigenvalues predominantly align with symmetry directions. Finally, we apply our symmetry-based analysis to convolutional networks, demonstrating the generality of our approach in understanding the Hessian eigenspectrum across different architectures.

DY 33.5 Thu 10:30 H47

Efficient mapping of phase diagrams with conditional Boltzmann Generators — •Maximilian Schebek[1], Michele Invernizzi[2], Frank Noé[1,2,3,4], and Jutta Rogal[1,5] — [1]Fachbereich Physik, Freie Universität Berlin, 14195 Berlin — [2]Fachbereich Mathematik und Informatik, Freie Universität Berlin, 14195 Berlin — [3]Department of Chemistry, Rice University, Houston, 77005, Texas, USA — [4]AI4Science, Microsoft Research, 10178 Berlin — [5]Department of Chemistry, New York University, New York, NY 10003, USA

The accurate prediction of phase diagrams is of central importance for both fundamental and applied material sciences. However, the computational prediction of the relative stability between phases based on their free energy is a daunting task, as traditional free energy estimators require a large amount of uncorrelated equilibrium samples over a grid of thermodynamic states. In this work, we develop deep generative machine learning models based on the Boltzmann Generator approach for entire phase diagrams, employing normalizing flows conditioned on the thermodynamic states that they map to. By training a single model to transform the equilibrium distribution sampled at only one reference thermodynamic state to a wide range of target temperatures and pressures, we can efficiently generate equilibrium samples across the entire phase diagram. We demonstrate our approach by predicting the solid-liquid coexistence line for a Lennard-Jones system in excellent agreement with state-of-the-art free energy methods while significantly reducing the number of energy evaluations needed.

DY 33.6 Thu 10:45 H47

Sampling rare events with neural networks: Machine learning the density of states — •Moritz Riedel[1], Johannes Zierenberg[2], and Martin Weigel[1] — [1]Institute of Physics, Technische Universität Chemnitz, 09107 Chemnitz, Germany — [2]Max Planck Institute for Dynamics and Self-Organization, 37077 Göttingen, Germany

Neural networks can be trained to generate samples from the Boltzmann distribution of many-particle systems. If suitable architectures such as normalizing flows or variational autoregressive networks are chosen, exact generation weights are known and hence present biases can be corrected for. Still, such networks typically struggle to learn and reproduce configurations from the full range of configuration space since effects such as mode collapse occur. For the simulation of rare events and suppressed states accessible in generalized frameworks such as the multicanonical ensemble such broad exploration is crucial. Here, we show how a combination of variational autoregressive network and autoencoder allows for a systematic exploration of configuration space

in spin models, during which the network is able to learn the density of states. We demonstrate the efficacy of the approach in the Potts system in the strong first-order regime.

DY 33.7 Thu 11:00 H47

stable diffusion for microstructure: from microstructural properties to 2D-to-3D reconstruction — •Yixuan Zhang[1], Teng Long[2], Mian Dai[1], and Hongbin Zhang[1] — [1]TU Darmstadt, Darmstadt, Germany — [2]Shandong University, Jinan, China

We propose a novel framework that combines Stable Diffusion and ControlNet to generate microstructures tailored to specific properties, such as coercivity. By leveraging latent alignment techniques, our method enables direct reconstruction of 3D microstructures from 2D inputs, ensuring geometric and property consistency across dimensions. This approach not only facilitates accurate 2D-to-3D reconstruction but also opens possibilities for studying and predicting microstructural transformations during various manufacturing processes. By integrating generative AI with material design, this work provides a robust foundation for property-driven microstructure generation, offering a potential pathway to optimize materials for targeted applications.

15 min. break

DY 33.8 Thu 11:30 H47

Machine learning for prediction of dynamical clustering in granular gases — •Sai Preetham Sata[1], Dmitry Puzyrev[2,1], and Ralf Stannarius[3,2] — [1]AMS, Otto von Guericke University, Germany — [2]MTRM and MARS, Otto von Guericke University, Germany — [3]Department of Engineering, Brandenburg University of Applied Sciences, Germany

Granular gases are sparse ensembles of free-moving macroscopic particles that interact via inelastic collisions. One peculiar property of granular gas is dynamical clustering, i.e. spontaneous increase of local number density. To quantify this effect, microgravity experiments and simulations were performed [1-3] and two gas-cluster transition criteria were established:Kolmogorov-Smirnov test, and caging effect criterion [2]. We perform simulations based on the VIP-GRAN experiment [3] and test these criteria for various combinations of system parameters, revealing their advantages and drawbacks. In addition, we investigate additional criteria that can help to understand the dynamical properties of gas-cluster transition. Based on the simulation data, machine learning can be used to detect dynamical clusters and predict the state of the system for a given set of system parameters. This study is funded by the German Aerospace Center (DLR) within projects VICKI (50WM2252) and EVA II (50WK2348). References: [1] É. Falcon et al., Phys. Rev. Lett., 83:440, 1999. [2] E. Opsomer et al., Europhys. Lett., 99:40001, 2012. [3] S. Aumaître et al., Rev. Sci. Instr., 89:075103, 2018.

DY 33.9 Thu 11:45 H47

Automated construction of complex reaction networks — •Weiqi Wang[1], Xiangyue Liu[1], and Jesús Pérez Ríos[2] — [1]Fritz-Haber-Institut, Berlin — [2]Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA

Kinetic models are essential for understanding chemical reaction mechanisms and estimating reaction rates. Typically, kinetic models are constructed based on transition state theory, using stable and intermediate species with zero-Kelvin energy calculations. However, they often fail to account for temperature effects and anharmonic influences, limiting their accuracy for real-world reactions.

This talk will discuss our method for automatically constructing reaction networks at finite temperatures using *ab initio* molecular dynamics simulations. Based on extensive sampling of configurational space, temperature-dependent free energies, and transition probabilities can be derived, enabling the construction of reaction networks to analyze temperature effects.

DY 33.10 Thu 12:00 H47

Data-Driven Sparse Identification with Adaptive Function Bases — •Gianmarco Ducci, Maryke Kouyate, Karsten Reuter, and Christoph Scheurer — Fritz-Haber-Institut der MPG, Berlin

Interpretable data-driven methods have proven viable for deriving kinetic equations directly from experimental data. However, such numerical methods are inherently susceptible to noise, which affects the sparsity in the resulting models. In order to promote such a sparsity condition, finding the optimal set of basis functions is a necessary prerequisite, but yet a challenging task to determine in advance.

We here present our in-house developed `ddmo` (Data-Driven Model Optimizer) software, which allows precise control over the space of candidate constituent terms. Such a complete framework comprises two main novel features. The first feature permits to include parametric functions in the library. The second feature is an adaptive library sizing routine that progressively adds or removes elements based on the learning from the dataset. We show a practical application of our algorithm tailored at identifying Langmuir-Hinshelwood mechanisms from experimental data.

DY 33.11 Thu 12:15 H47

Kalman filter enhanced adversarial Bayesian optimization for active sampling in inelastic neutron scattering — Yixuan Zhang[1], •Nihad Abuawwad[2], Samir Lounis[2], and Hongbin Zhang[1] — [1]TU Darmstadt, Darmstadt, Germany — [2]Peter Grünberg Institute (PGI), Jülich, Germany

Spin waves, or magnons, are fundamental excitations in magnetic materials that provide insights into their dynamic properties and interactions. Magnons are the building blocks of magnonics, which offer promising perspectives for data storage, quantum computing, and communication technologies. These excitations are typically measured through inelastic neutron or x-ray scattering techniques, which involve heavy and time-consuming measurements, data processing, and analysis based on various theoretical models. Here, we introduce a machine learning algorithm that integrates adaptive noise reduction and active learning sampling, which enables the restoration from minimal inelastic neutron scattering point data of spin wave information and the accurate extraction of magnetic parameters, including hidden interactions. Our findings, benchmarked against the magnon spectra of CrSBr, significantly enhance the efficiency and accuracy in addressing complex and noisy experimental measurements. This advancement offers a powerful machine learning tool for research in magnonics and spintronics, which can also be extended to other characterization techniques at large facilities.

DY 33.12 Thu 12:30 H47

Accelerating the Training and Improving the Reliability of Machine-Learned Interatomic Potentials for Strongly Anharmonic Materials through Active Learning — •Kisung Kang, Thomas A. R. Purcell, Christian Carbogno, and Matthias Scheffler — The NOMAD Laboratory at the FHI of the Max Planck Society

Machine-learned interatomic potentials (MLIP) can efficiently implement molecular dynamics (MD) simulations with large spatial and long time scales. However, immature training for rare dynamical events, such as defect creation, may happen due to their absence or insufficiency in training data or their fadeout during regularization, leading to the critical deterioration of MLIP predictions regarding dynamical properties like transport phenomena. To improve the MLIP's reliability and accelerate the whole training process, we adopt a sequential active learning ($\mathcal{AL}$) scheme via MD employing MLIP (MLIP-MD) and uncertainty estimates [1]. In each iterative step, MLIP-MD serves as an efficient exploration tool for configurational space to generate training data, while uncertainty estimates identify unfamiliar data to be sampled for subsequential MLIP models. The representative examples of CuI and $AgGaSe_2$ among 112 materials display erroneous MLIP predictions of missing and fictitious rare events. We demonstrate how $\mathcal{AL}$ addresses these issues, specifically correcting unfamiliar regions for the MLIP potential energy surface. At last, the over(under)estimation of their phonon lifetimes is rectified after the $\mathcal{AL}$ steps.
[1] K. Kang, T. A. R. Purcell, *et al.*, arXiv:2409.11808 (2024).

DY 33.13 Thu 12:45 H47

Molecular Dynamics of Endohedral CaX@C60 Fullerenes: Reproducing Correlated Movement Features Using Machine Learning Applications — •Mihaela Cosinschi[1,3], Amanda Teodora Preda[1,3], Calin Andrei Pantis Simut[1,3], Nicolae Filipoiu[1,3], Ioan Ghitiu[4], Mihnea Alexandru Dulea[3], Andrei Manolescu[5], and George Alexandru Nemnes[1,2,3] — [1]University of Bucharest, Faculty of Physics, Magurele, Romania — [2]Research Institute of the University of Bucharest, Bucharest, Romania — [3]Horia Hulubei National Institute for Physics and Nuclear Engineering, Magurele, Romania — [4]National Institute for Laser, Plasma and Radiation Physics, Magurele, Romania — [5]Department of Engineering, School of Technology, Reykjavik University, Reykjavik, Iceland

Fullerenes are allotropes of carbon with remarkable properties due to their high degree of symmetry, cage-like structures and ability to support addition of internal or external atoms. In the present work, we have conducted a molecular dynamics (MD) study on the C60 fullerene containing one to four encapsulated calcium atoms. All-atomistic ab initio DFT methods were employed to perform calculations through the Orca MD Module, albeit at a high computational cost. Results proved that the internal atoms adopt minimal-energy configurations and exhibit coupled motion, maintaining constant characteristics after a period of equilibration. Furthermore, we have built an artificial neural network (ANN) that can pick up the dynamics patterns and recreate trajectories to reasonable accuracy, allowing for MD calculations in significantly shorter times, even under small perturbations.