BOOK OF ABSTRACTS

of the

Begegnungszone

Statistical Physics and Machine Learning

SPAML23

Seminaris Hotel Leipzig

Organisers:

Johannes Zierenberg
MPI-DS Göttingen

Martin Weigel
TU Chemnitz

September 18–21, 2023

https://www.spaml23.tu-chemnitz.de/

Monday, September 18, 2023

12:30 – 14:00		**– Lunch at Seminaris Hotel –**
13:50 – 14:00		**– Opening –**

Session 1: Session I: Hidden Structures (Chair: Johannes Zierenberg)

14:00 – 14:30	**Jeff Byers**	Statistical Mechanics and the Geometry of Machine Learning
14:45 – 15:00	**Andrea Perin**	Quantifying the Improvements in Learning Efficiency Obtained by Leveraging Symmetries in the Data
15:15 – 15:30	**NN**	Brainstorming session
15:45 – 16:15		**– Coffee Break –**

Session 2: Session II: Interpretable Machine Learning (Chair: Jeremi Ochab)

16:15 – 16:45	**Marvin Wright**	Interpretable Machine Learning
17:00 – 17:15	**Konstantin Nikolaou**	Towards a Phenomenological Understanding of Neural Networks: Data
17:30 – 17:45	**NN**	Brainstorming session
18:30 – 20:00		**– Dinner at Seminaris Hotel –**
20:00 – 22:00		**– Posters and Welcome Reception –**

Tuesday, September 19, 2023

Session 3: Session III: Deep Networks (Chair: Jonathan Kadmon)

08:30 – 09:00	**Riccardo Zecchina**	Liquid States in Neural Networks
09:15 – 09:30	**Francesco Cagnetta**	Learning Hierarchical Compositionality With Deep Convolutional Networks: Insights From a Random Hierarchy Model
09:45 – 10:00	**Sebastiano Ariosto**	A Statistical Mechanics Framework for Deep Neural Networks Beyond the Infinite-Width Limit
10:00 – 10:30		**– Coffee Break –**

Session 4: Session IV: Training I (Chair: Sebastian Goldt)

10:30 – 11:00	**Beatriz Seoane**	Statistical Physics of Energy-Based Generative Models
11:15 – 11:30	**Frederieke Richert**	The Role of the Activation Function in Feedforward Learning Systems
11:45 – 12:00	**NN**	Brainstorming session
12:20 – 12:30		**– Conference Photo –**
12:30 – 14:00		**– Lunch at Seminaris Hotel –**

Session 5: Session V: Training II (Chair: Beatriz Seoane)

14:00 – 14:30	**Sebastian Goldt**	The Gaussian World is Not Enough – How Training Data Shapes Neural Representations
14:45 – 15:00	**Christian Keup**	Learning Dynamics in Deep Teacher-Student Models
15:15 – 15:30	**NN**	Brainstorming session
15:45 – 16:15		**– Coffee Break –**
16:15 – 18:30		**– Conference Excursion –**
19:00 – 21:00		**– Conference Dinner –**

Wednesday, September 20, 2023

Session 6: Session VI: Quantum Mechanics (Chair: Wolfhard Janke)

08:30 – 09:00	**Lei Wang**	Unlocking the Power of the Variational Free-Energy Principle with Deep Generative Models
09:15 – 09:30	**Luciano Vitteriti**	Transformer Variational Wave Functions for Frustrated Quantum Spin Systems
09:45 – 10:00	**NN**	Brainstorming session
10:00 – 10:30		**– Coffee Break –**

Session 7: Session VII: Recurrent Networks (Chair: Marylou Gabrié)

10:30 – 11:00	**Jonathan Kadmon**	Reliable Coding With Chaotic Neural Networks
11:15 – 11:30	**Freya Behrens**	The Backtracking Dynamical Cavity Method and Cellular Automata
11:45 – 12:00	**NN**	Brainstorming session
12:30 – 14:00		**– Lunch at Seminaris Hotel –**

Session 8: Session V: Sampling (Chair: Jeff Byers)

14:00 – 14:30	**Marylou Gabrié**	Assisting Sampling (of Physical States) With Generative Models
14:45 – 15:00	**Henrik Christiansen**	Learning How to Integrate: Accelerating Hamiltonian Monte Carlo with Machine Learning
15:15 – 15:30	**NN**	Brainstorming session
15:45 – 16:15		**– Coffee Break –**
16:15 – 18:30		**– Group Projects –**
18:30 – 20:00		**– Dinner at Seminaris Hotel –**

Thursday, September 21, 2023

Session 9: Session IX: Physical Systems (Chair: Martin Weigel)

08:30 – 09:00	**Miriam Klopotek**	From Model Systems of Matter to Physical Computing and Physics-Explainable Machine Learning
09:15 – 09:30	**Riccardo Rende**	Optimal Inference of a Generalised Potts Model by Single-Layer Transformers with Factored Attention
09:45 – 10:00	**NN**	Brainstorming session
10:00 – 10:30		**– Coffee Break –**
10:30 – 12:00		**– Panel Discussion: Open Problems and Challenges –**
12:00 – 12:15		**– Summary –**
12:15 – 12:30		**– Conference Closing –**
12:30 – 14:00		**– Lunch at Seminaris Hotel –**

A1 – A Statistical Mechanics Framework for Deep Neural Networks Beyond the Infinite-Width Limit

Sebastiano Ariosto
University of Insubria, Italy
sebastiano.ariosto@gmail.com

Decades-long literature testifies to the success of statistical mechanics at clarifying fundamental aspects of deep learning. Yet the ultimate goal remains elusive: we lack a complete theoretical framework to predict practically relevant scores, such as the train and test accuracy, from knowledge of the training data. Huge simplifications arise in the infinite-width limit, where the number of units N_l in each hidden layer ($l = 1, \dots, L$, being L the finite depth of the network) far exceeds the number P of training examples. This idealisation, however, blatantly departs from the reality of deep learning practice, where training sets are larger than the widths of the networks. Here, we show one way to overcome these limitations. The partition function for fully-connected architectures, which encodes information about the trained models, can be evaluated analytically with the toolset of statistical mechanics. The computation holds in the "thermodynamic limit" where both N_l and P are large and their ratio $\alpha_l = P/N_l$, which vanishes in the infinite-width limit, is now finite and generic. This advance allows us to obtain (i) a closed formula for the generalisation error associated to a regression task in a one-hidden layer network with finite α_1; (ii) an approximate expression of the partition function for deep architectures (technically, via an "effective action" that depends on a finite number of "order parameters"); (iii) a link between deep neural networks in the proportional asymptotic limit and Student's t processes; (iv) a simple criterion to predict whether finite-width networks (with ReLU activation) achieve better test accuracy than infinite-width ones. As exemplified by these results, our theory provides a starting point to tackle the problem of generalisation in realistic regimes of deep learning.

A2 – The Backtracking Dynamical Cavity Method and Cellular Automata

Freya Behrens
École polytechnique fédérale de Lausanne, Switzerland
freya.behrens@epfl.ch

Despite their simple description as a local update rules, Cellular Automata (CA) may exhibit very complex dynamic behaviours. This phase of complex behaviour is sometimes referred to as "the edge of chaos" and understanding which properties of a rule lead to this non-trivial dynamics has long been of interest to researchers. Similar phase transitions can nicely be analysed in random boolean networks, a more general formula- tion of CA on graphs with independent rules per node. However, this breaks two properties inherent to traditional CA, i.e. the regularity of the grid structure and the uniformity of the rule over the complete grid. To bridge this gap between the amenable and interesting, we analyse CA's on random sparse graphs. We particularly focus on rules with short attractors to analyse the zoology of phase transitions that occur for these rules depending on the initial state of the dynamics. This has implications for CA-amenable phenomena like majority dynamics, opinion formation and even neural network like models of computation. From a methodological point of view, our work is interesting as we use the dynamic cavity method in a to analyse dynamics the backwards in time from the attractors, novel approach which could be interesting to other problems.

A3 – Statistical Mechanics and the Geometry of Machine Learning

Jeff Byers
U.S. Naval Research Laboratory, Washington, DC, U.S.A.
jeff.byers@nrl.navy.mil

What intellectual threads connect Statistical Mechanics and Machine Learning?

First consider that unlike most of physics, statistical mechanics is rooted in the core issues of how do we represent the uncertainty of our knowledge about objects and processes, making the discipline more epistemological in its foundations. Foremost among epistemic concepts at the crossroads of machine learning and statistical mechanics is entropy. Entropy connects our epistemic uncertainty with constraints on probabilistic representations and a fundamental measure of information content. Second, the representations used in machine learning typically embed data in a high-dimensional feature space that posits Euclidean distance as a surrogate for entropy. Similarly, statistical mechanics confronts the issue of how to distribute probability in high-dimensional phase spaces subject to observational constraints. Finally, the underlying models in statistical mechanics can be viewed as a kind of log-prior in Bayesian inference. Observations act as constraints on the prior induced by the model. In this way, statistical mechanics proposes symmetries at a global scale while machine learning often uses kernels as a form of similarity measure on the local scale. The two approaches can be seen as different perspectives that formally link Hamiltonians to their corresponding propagators or Green's functions.

The goal of the talk is to suggest that the future of machine learning and artificial intelligence can be guided by developing algorithms that explicitly learn local symmetry operations on constraint manifolds from data in high-dimensional spaces. Whereas, the existing deep neural network algorithms can only stumble into the appropriate representations with enormous amounts of data and fortuitous choices of network architecture.

A4 – Learning Hierarchical Compositionality With Deep Convolutional Networks: Insights from a Random Hierarchy Model

Francesco Cagnetta
École polytechnique fédérale de Lausanne, Switzerland
francesco.cagnetta@epfl.ch

Understanding the inner workings of deep learning methods remains a significant challenge. What properties of real data make these methods so successful and how are they learned? Lacking an answer to these questions we cannot even estimate the order of magnitude of the number of training examples necessary to learn a given task. In this presentation I will explore our approach to solving this challenge, which relies on studying a synthetic classification task — the Random Hierarchy Model — built to mimic the hierarchical and compositional structure of natural data. As the subject of an image (e.g. a dog) consists of features (head, body, limbs), themselves consisting of sub-features (eyes, nose and mouth for the head), each of the classes in our model corresponds to a number of synonymic compositions of high-level features, in turn composed of lower-level features. Thanks to our specific model choice, we are able to predict how many training points are required to learn the task as a function of the number of classes, number of synonyms and depth of the hierarchy. Interestingly, this number is strictly related to the detectability of the correlations between the low-level features of the data and their class.

A5 – Learning How to Integrate: Accelerating Hamiltonian Monte Carlo With Machine Learning

Henrik Christiansen
NEC Research Laboratories Europe, Heidelberg, Germany
henrik.christiansen@neclab.eu

The performance of Hamiltonian Monte Carlo crucially depends on its parameters, i.e., the time step and the number of integration steps. We propose a fully-differentiable framework to optimize these parameters using gradient information. It is demonstrated for the one dimensional harmonic oscillator and alanine dipeptide, where we find that our approach finds good parameters, eliminating the need for hand-tuning these. The set-up can also be extended to learn more parameters, such as atom-dependent timesteps. There, we find roughly a 25% improvement in autocorrelation times.

A6 – Exploring the Role of Sparsity in Soft Committee Machines From a Statistical Mechanics Framework

Otávio Citton
University of Groningen, Netherlands
o.c.citton@rug.nl

Sparsity has proven to be a valuable technique in training deep neural networks, effectively mitigating overfitting and enhancing generalization. However, the precise extent of its influence on the learning process remains unclear. Motivated by a potential theoretical understanding of the performance of such models, we employ the framework of statistical mechanics to investigate the learning of artificial neural networks in the presence of sparsity-promoting terms. Specifically, we focus our analysis on soft committee machines and utilize the replica formalism to evaluate its free energy. By incorporating sparsity into the Hamiltonian, we aim to gain insights into its impact on the network's capacity, generalization ability, and overall performance. Through systematic statistical mechanical analysis, we explore how these new terms influence networks with different activation functions.

A7 – Assisting Sampling (of Physical States) With Generative Models

Marylou Gabrié
École Polytechnique, Paris, France
marylou.gabrie@polytechnique.edu

Monte Carlo methods share a long history of interaction with statistical mechanics. In these lectures, we will discuss further opportunities in enhancing sampling with machine learning. These methods rely largely on deep generative models which are very flexible probabilistic models and provide independent samples from complex high-dimensional distributions at negligible costs. On the other hand, sampling exactly a target distribution, such a Bayesian posterior or the Boltzmann distribution of a physical system, is typically challenging: either because of dimensionality, multi-modality, ill-conditioning or a combination of the previous. I will present different approaches recently proposed, the challenges they tackle and the remaining limitations.

A8 – Predicting Chaotic Systems With Active Matter Reservoir Computing

Mario Gaimann
University of Stuttgart, Germany
mario.gaimann@simtech.uni-stuttgart.de

Spatio-temporal prediction of chaotic systems is a challenging problem that is relevant for many fields (weather, finance, energy, and other dynamic systems). Recurrent neural networks and specifically neuron-based reservoir computing were previously used to approach this problem [1,2]. However, these learning systems are typically treated as black boxes, and do not incorporate reasoning or analysis in terms of physical laws and dynamics. Here we study the non-equilibrium dynamics of simple active matter models serving as reservoir computing substrates [3]. This allows us to determine and interpret the state of our reservoir and relate the learning problem to other generic phenomena in statistical physics. With this knowledge we aim to understand optimal conditions for learning in terms of physics.

[1] Tanaka, G. et al. (2019), Neural Networks 115, 100-123.
[2] Nakajima, K. and Fischer, I. (2021). Reservoir Computing. Springer Singapore.
[3] Lymburn, T. et al. (2021), Chaos 31(3), 033121.

A9 – Policy Guided Monte Carlo in Supercooled Liquids

Leonardo Galliano
University of Trieste, Italy
leonardo.galliano@phd.units.it

We use and extend a recently proposed policy-guided Monte Carlo (PGMC) algorithm, which identifies efficient MC moves as part of an optimisation process inspired by reinforcement learning. First, we generalise the method to the case of continuous degrees of freedom. We focus on binary mixture models for which simple swap MC is inefficient and which are sufficiently robust against crystallisation, at least when simulated with conventional molecular dynamics. We show that for one of these mixtures PGMC identifies a biased swap move that is significantly more efficient than simple swaps.

A10 – Self-Learning Population Simulations for a Frustrated Ising Model

Denis Gessert
Universität Leipzig, Germany
gessert@itp.uni-leipzig.de

Population annealing (PA) is a Monte Carlo method well suited for problems with a rough free energy landscape such as glassy systems. PA is similar to repeated simulated annealing, with the addition of a resampling step at each temperature. To achieve best performance, the PA parameters have to be chosen carefully. Both the set of temperatures, as well as the number of sweeps carried out at each temperature, form a several hundred-dimensional parameter space. In practice, these are often chosen a priori following very simple patterns, e.g., a temperature set given by linearly increasing inverse temperature with a constant number of updates at each temperature. While it is now common to choose the temperatures adaptively, the number of updates still is usually either constant or piecewise constant. We propose a fully adaptive variant of the PA algorithm in which both the temperature step and the number of updates are chosen adaptively. Besides few hyper-parameters, this method requires no previous knowledge on the underlying model and works in an online self-learning manner. To test our method we study the Ising model with

competing ferromagnetic ($J_1 > 0$) nearest and antiferromagnetic ($J_2 < 0$) next-to-nearest neighbor interactions on the honeycomb lattice.

A11 – The Gaussian World is Not Enough – How Training Data Shapes Neural Representations

Sebastian Goldt
International School for Advanced Studies, Trieste, Italy
goldt.sebastian@gmail.com

What do neural networks learn from their data? We discuss this question in two learning paradigms: supervised classification with feed-forward networks, and masked language modelling with transformers. First, we give analytical and experimental evidence for a "distributional simplicity bias", whereby neural networks learn increasingly complex distributions of their inputs. We then show that neural networks learn from the higher-order cumulants (HOCs) more efficiently than lazy methods, and show how HOCs shape the learnt features. We finally characterise the distributions that are learnt by single- and multi-layer transformers, and discuss implications for learning ground states of quantum many-body problems.

A12 – Applying the Rayleigh-Ritz Procedure to the Nelson Picture of Quantum Mechanics using Neural Networks and Genetic Algorithms

Kai-Hendrik Henk
Martin-Luther-Universität Halle-Wittenberg, Germany
kai-hendrik.henk@physik.uni-halle.de

The stochastic formulation of Quantum mechanics, also known as the Nelson picture, is lacking a good method for calculating the osmotic velocity $u(x)$, which contains all the information of a system in this picture. obtaining it from the Schrödinger wave function is conceptually unsatisfying due to it's complete dependence on another formulation of Quantum mechanics. The other method, stochastic optimal control, is independent from other formulations but suffers from inefficiencies and difficulties in programmability. Here, I introduce a genetic algorithm with neural networks which uses the Rayleigh-Ritz variational principle to obtain the osmotic velocity for ground states. As a proof of concept, we calculated $u(x)$ for one dimensional systems, the harmonic oscillator, the double well and the Pöschl-Teller potential. To obtain exited states, we calculate ground states of super symmetrical partner Hamiltonians for each of these potentials.

A13 – A Fast and Flexible, Multilayer Ising Simulation Platform for Quick Experimentation and Analysis

Fabian IJpelaar
University of Groningen, Netherlands
f.i.ijpelaar@rug.nl

We present an interactive simulation platform for Ising- and Ising-like models, leveraging the Julia programming language for enhanced speed, flexibility, and extensibility. Our software offers unique features that enable highly flexible and fast experimentation with Ising models, including the ability to construct weight matrices based on arbitrary spin distance functions, easy creation of neural architectures, and a suite of fast, real-time analysis tools.

A14 – Reliable Coding With Chaotic Neural Networks

Jonathan Kadmon
Hebrew University, Jerusalem, Israel
jonathan.kadmon@mail.huji.ac.il

Artificial neural networks frequently serve as frameworks for exploring neural computation in the brain; however, significant discrepancies exist between these artificial constructs and their biological counterparts. Specifically, neurons in the brain demonstrate strong recurrent connectivity and exhibit highly irregular activity, phenomena often attributed to chaotic dynamics. Despite the erratic behavior of individual neurons, cortical circuits achieve remarkable computational reliability. In this talk, I will show how large, chaotic networks can be reliable. Specifically, I will address two key questions: (1) How do high-dimensional networks approximate arbitrary dynamical systems? and (2) How do large chaotic networks achieve robust generalization?

A15 – Understanding Neural Network Models for Phase Recognition

Shashank Suresha Kallappara
Technische Universität Chemnitz, Germany
shashank-suresha.kallappara@physik.tu-chemnitz.de

Neural Networks have been used successfully to detect the different phases of various spin systems. We investigate and compare the performances of a few different models and use the smallest networks to make a landscape which can be used to visualised training of the network.

A16 – Learning Dynamics in Deep Teacher-Student Models

Christian Keup
École polytechnique fédérale de Lausanne, Switzerland
christian.keup@epfl.ch

While 2-layer networks are increasingly well understood theoretically, there is still no analytically tractable toy-model of deep networks beyond deep linear and lazy models. I give an overview of recent statphys approaches to predict learning curves and generalization errors in deep feed-forward networks, including our own work in this direction.

A17 – From Model Systems of Matter to Physical Computing and Physics-Explainable Machine Learning

Miriam Klopotek
Universität Stuttgart, Germany
miriam.klopotek@simtech.uni-stuttgart.de

Emergent phenomena are central to the physics of soft and hard condensed matter: Out of equilibrium, the occurrence of phase transitions as well as collective behavior – especially under external driving – manifest the 'surprise' inherent to these systems as they self-assemble into many different meso- and macroscopic forms and dynamical states. Arguably, matter can be 'creative' in the rich, complex response to changing external conditions: In this talk, we argue this can be used for learning problems directly, as well as for interpreting machine learning algorithms and how they are trained.

To the latter, we first show results for generative models trained on data from statistical model systems of hard rods, enabling a deepened physics-interpretation and -explanation of β-variational autoencoders [1,2]. To the former, we discuss how active-matter model systems can perform computations in unconventional ways within the reservoir computing paradigm [3]. In turn, such physical computing can be interpreted and eventually customized based on physical arguments, viz. dynamical heterogeneity and collective properties.

[1] P. Quiring, M. Klopotek and M. Oettel, Phys. Rev. E 100, 012707 (2019). M. D. Klopotek, PhD Dissertation, Tübingen (2021).

[2] D. P. Kingma and M. Welling, ICLR (2014). D. J. Rezende, S. Mohamed, and D. Wierstra, ICML (2014), p. 1278–1286.

[3] T. Lymburn et al., Chaos 31, 033121 (2021).

A18 – Fast, Hierarchical, and Adaptive Algorithm for Metropolis Monte Carlo Simulations of Long-Range Interacting Systems

Fabio Müller

Universität Leipzig, Germany

fmueller@itp.uni-leipzig.de

We present a fast, hierarchical, and adaptive algorithm for Metropolis Monte Carlo simulations of systems with long-range interactions that reproduces the dynamics of a standard implementation exactly, ie, the generated configurations and consequently all measured observables are identical, allowing in particular for nonequilibrium studies. The method is demonstrated for the power-law interacting long-range Ising and XY spin models with nonconserved order parameter and a Lennard-Jones particle system, all in two dimensions. The measured run times support an average complexity $\mathrm{O}(N \log N)$, where N is the number of spins or particles. Importantly, prefactors of this scaling behavior are small, which in practice manifests in speedup factors larger than 10^4. The method is general and will allow the treatment of large systems that were out of reach before, likely enabling a more detailed understanding of physical phenomena rooted in long-range interactions.

A19 – Towards a Phenomenological Understanding of Neural Networks: Data

Konstantin Nikolaou

University of Stuttgart, Germany

knikolaou@icp.uni-stuttgart.de

Currently we witness quick-growing research efforts in materials science and nanoelectronics, where the goal is to leverage unconventional physical effects for extremely low-energy and high-bandwidth computing. In this field, Ising-like physical dynamics are a promising candidate. Our engine enables theorists to explore the many ways how Ising-like dynamical systems can be exploited for practical computations.

A20 – Comparing Cross-Correlation Estimators in fMRI Data

Jeremi Ochab
Jagiellonian University Krakow, Poland
jeremi.ochab@uj.edu.pl

Functional magnetic resonance imaging (fMRI) signals are notoriously challenging to analyse due to their very low temporal resolution (seconds), leading to a considerably larger number of time series than time points ($N/T \gg 1$), and a non-trivial auto-correlation and cross-correlation (CC) structure. In this contribution, we present an analysis of a memory distortion experiment [1] based on sample Pearson correlation, detrended CC [2] and non-linear asymmetric CC based on filtering high-amplitude events, rBeta [3], and their shrinkage estimators [3].

Uncovering regionally coordinated changes in brain activity between experimental conditions involved comparing distributions of correlation matrices' eigenvalues. For that purpose, we additionally performed agglomerative hierarchical clustering of eigenvectors based on their similarity. The rBeta correlation matrix was symmetrised before shrinking. The signals considered in this work were averages over all voxels belonging to an anatomical atlas region.

The CC between brain areas were found indicative of differences between spontaneous brain activity (resting state) and other tasks, as well as between different tasks (verbal vs non-verbal) and experimental stages (memorisation and recollection). Other statistically significant effects associated with smaller eigenvalues are yet to be interpreted psychologically. Detailed statistical analyses were performed for the rBeta method, providing additional results on the different types of memorised stimuli (matching, non-matching, intentionally confusing). In terms of methodology, the detrended correlations turned out to be more sensitive (leading to stronger statistical effects) than Pearson correlations, and clustered eigenvalues to unclustered ones. The rBeta yielded qualitatively different results than other methods. The stability of the eigenvector clustering is still to be scrutinised together with the optimisation of the detrending parameters in detrended CC. The shrinkage makes the results more consistent across correlation matrix estimation methods.

[1] Ochab JK, Wątorek M, Ceglarek A, et al., *Sci. Rep.* **12** (2022) 17866.

[2] Kwapień J, Oświęcimka P, Drożdż S, *Phys. Rev. E* **92** (2015) 052815.

[3] Cifre I, Miller Flores MT, Penalba L, Ochab JK, Chialvo DR, *Front. Neurosci.* **15** (2021) 1194.

[4] Ceglarek A, Ochab JK, Cifre I, et al., *Front. Neurosci.* **15** (2021) 1611.

[5] Burda Z, Jarosz A, *Physical Review E* **105** (2022) 034136.

A21 – Quantifying the Improvements in Learning Efficiency Obtained by Leveraging Symmetries in the Data

Andrea Perin
Aalto University Helsinki, Finland
andrea.perin@aalto.fi

The problem of linear separability has been recently studied in the context of concept manifolds (Chung et al., 2018) and group structured data (Farrell et al., 2022). A more realistic approach to qualifying the goodness of data representations in machine learning, however, is instead few shot classification (i.e., generalising classification from a few examples of each class). While previous work has studied few shot learning for general concept manifolds (Sorscher et al., 2021), the case of explicit group structures in the manifolds has not been studied. We present our current framework, specifically studying the effect of considering symmetries in few shot classification of group-structured concept manifolds, present some considerations that arise from it on the specific case study of MNIST digits and rotations, and discuss possible extensions to non linear learning methods.

A22 – Optimal Inference of a Generalised Potts Model by Single-Layer Transformers With Factored Attention

Riccardo Rende
International School for Advanced Studies, Trieste, Italy
rrende@sissa.it

Transformers are the type of neural networks that has revolutionised natural language processing and protein science. Their key building block is a mechanism called self-attention which is trained to predict missing words in sentences. Despite the practical success of transformers in applications it remains unclear what self-attention learns from data, and how. Here, we give a precise analytical and numerical characterisation of transformers trained on data drawn from a generalised Potts model with interactions between sites and Potts colours. While an off-the-shelf transformer requires several layers to learn this distribution, we show analytically that a single layer of self-attention with a small modification can learn the Potts model exactly in the limit of infinite sampling. We show that this modified self-attention, that we call "factored", has the same functional form as the conditional probability of a Potts spin given the other spins, compute its generalisation error using the replica method from statistical physics, and derive an exact mapping to pseudo-likelihood methods for solving the inverse Ising and Potts problem.

A23 – The Role of the Activation Function in Feedforward Learning Systems

Frederieke Richert
University of Groningen, Netherlands
f.richert@rug.nl

Statistical physics has played an important role in the last decades in increasing the theoretical knowledge about the typical behaviour of neural networks. One phenomenon that has surfaced recently is different phase transitions in off-line learning for two-layer feedforward networks depending on the activation function used in the network. It was shown [Oostwal, Straat, Biehl. "Hidden unit specialization in layered neural networks: ReLU vs. sigmoidal activation", 2021] that while the sigmoidally shaped erf results in a first-order phase transition when varying the training set size, the now commonly used ReLU activation function displays a second-order specialisation of hidden units in shallow networks. This can be of practical relevance, because in the case of a first-order phase transition a suboptimal solution remains stable after the transition point and training might get stuck in this solution instead of converging to the optimal solution. Inspired by this finding, we investigate the influence of the activation function on the generalisation behaviour of layered feedforward neural networks. A first goal is to identify features of the activation functions that lead to the respective phase transition. In search for these features, we apply methods from statistical physics of learning, such as high temperature and annealed approximation as well as the replica formalism. We obtained first results for several relevant activation functions. This includes the GELU activation function and a convex combination of the GELU and the erf, which prove that convexity is not indicative of the type of phase transition.

A24 – Statistical Physics of Energy-Based Generative Models

Beatriz Seoane
Université Paris-Saclay, France
beseoane@ucm.es

Energy-based models (EBMs) are powerful generative machine learning models that are able to encode the complex distribution of a dataset in the Gibbs-Boltzmann distribution of a model energy function. This means that, if properly trained, they can be used to synthesize new samples that resemble those of the dataset as closely as possible, but also

that this energy function can be used to "learn" something about the building mechanisms of the dataset under study. Indeed, EBMs can be considered a powerful modeling tool for arbitrary data if one were able to map complex energy functions defined in a neural network into spin-interaction Hamiltonians that can be explored using standard statistical physics tools. Such an approach has long been used in physics for inverse Ising problems. The goal now is to extend this approach to more complex energy functions that can encode all higher order correlations in complex data. While this program is very encouraging, training good EBMs is particularly challenging, mainly because they rely on long Monte Carlo sampling processes to estimate the log-likelihood gradient. In my talk, I will present some results on the interpretability of shallow EBMs and discuss how computational statistical physics is a valuable tool for understanding and improving and controlling the training of EBMs.

A25 – Transformer Variational Wave Functions for Frustrated Quantum Spin Systems

Luciano Lori Viteritti
University of Trieste, Italy
lucianoloris.viteritti@phd.units.it

The Transformer architecture has become the state-of-art model for natural language processing tasks and, more recently, also for computer vision tasks, thus defining the Vision Transformer (ViT) architecture. The key feature is the ability to describe long-range correlations among the elements of the input sequences, through the so-called self-attention mechanism. Here, we propose an adaptation of the ViT architecture with complex parameters to define a new class of variational neural-network states for quantum many-body systems, the ViT wave function. We apply this idea to the one dimensional $J_1 - J_2$ Heisenberg model, demonstrating that a relatively simple parametrization gets excellent results for both gapped and gapless phases. In this case, excellent accuracies are obtained by a relatively shallow architecture, with a single layer of self-attention, thus largely simplifying the original architecture. Still, the optimization of a deeper structure is possible and can be used for more challenging models, most notably highly-frustrated systems in two dimensions. The success of the ViT wave function relies on mixing both local and global operations, thus enabling the study of large systems with high accuracy.

A26 – Unlocking the Power of the Variational Free-Energy Principle with Deep Generative Models

Lei Wang
Chinese Academy of Sciences, Beijing, China
wanglei@iphy.ac.cn

At finite temperatures, Nature tries to balance energy and entropy to achieve minimal free energy. Ironically, it was rather difficult to turn such a free energy minimization principle into a practical algorithm, especially for quantum matter consisting of fermions. Such a difficulty is mainly due to the prohibitive entropy calculation in the variational free energy, also known as the "intractable" partition functions problem.

We harness the power of the least free energy principle by employing tractable deep generative models including the normalizing flow and autoregressive model as the variational ansatz. For classical statistical physics problems, this is a new development that "inherits the glorious tradition"; while for the quantum many-body problem, it is expected to become a unique technique that is "one of a kind". I will illustrate these developments with concrete examples such as spin glasses and homogeneous electron gases. In the latter case, the calculation has provided new insight into quasiparticle effective mass in the prototypical fermi liquid.

A27 – Online Learning Dynamics of the Random Feature Model in the Student-Teacher Framework

Roman Worschech
Universität Leipzig, Germany
roman.worschech@mis.mpg.de

Deep neural networks are widely used forecasting algorithms whose predictive power often improves as the number of weights increases, leading to over-parametrization. We consider a two-layered neural network whose first layer is frozen while the last layer is trainable, known as the random feature model. We express over-parametrization in the context of a student-teacher framework and show how it improves the generalizability of the network by deriving a set of differential equations for weight dynamics. We find that the asymptotic generalization error is bounded from below for weak correlations of student and teacher weights. In the case of high correlations, the student can perfectly predict the teacher's output if its hidden layer's size increases exponentially faster than the size of the input layer.

A28 – Interpretable Machine Learning

Marvin Wright
Universität Bremen, Germany
wright@leibniz-bips.de

Machine learning has become an indispensable tool in almost all scientific fields. With the rise of deep learning, machine learning revolutionized various applications, from image and speech recognition to natural language processing, but also the analysis of tabular and high-dimensional data has seen major improvements. However, in many applications, the translation of machine learning research into practice is hampered by the black box nature of machine learning models. Interpretable machine learning is an emerging field that aims to address this challenge, developing methods to increase the interpretability of black box models. In this talk, I will give an overview of state-of-the-art methods for interpretable machine learning, including partial dependence plots and Shapley values, but also point out current challenges and promising research directions.

A29 – Liquid States in Neural Networks

Riccardo Zecchina
Bocconi University, Milan, Italy
zecchina@gmail.com

This seminar will delve into the importance of liquid states in neural network architectures, specifically focusing on feedforward and attractor networks. For feedforward networks, we show that regions with 'liquid flatness' in the loss landscape are associated with minimizers that exhibit strong generalization in overparameterized, non-convex models. In the context of asymmetric attractor random networks, the discussion will highlight the coexistence of an exponential number of liquid attractors alongside chaotic fixed points. This interplay results in the existence of an exponentially large set of internal representations endowed with error-correcting capabilities.

These findings draw upon a large deviation technique and are supported by rigorous results (when possible) and numerical simulations.

Sebastiano Ariosto sebastiano.ariosto@gmail.com	Università dell'Insubria, Italy
Freya Behrens freya.behrens@epfl.ch	EPFL, Switzerland
Jeff Byers jeff.byers@nrl.navy.mil	Naval Research Laboratory, U.S.A.
Francesco Cagnetta francesco.cagnetta@epfl.ch	EPFL, Switzerland
Henrik Christiansen henrik.christiansen@neclab.eu	NEC Research Laboratories Europe, Heidelberg, Germany
Otávio Citton o.c.citton@rug.nl	University of Groningen, Netherlands
Sankar Prasad Das shankar@jnu.ac.in	Jawaharlal Nehru University, New Delhi, India
Marylou Gabrié marylou.gabrie@polytechnique.edu	École Polytechnique, Paris, France
Mario Gaimann mario.gaimann@simtech.uni-stuttgart.de	University of Stuttgart, Germany
Leonardo Galliano leonardo.galliano@phd.units.it	University of Trieste, Italy
Denis Gessert gessert@itp.uni-leipzig.de	Leipzig University, Germany
Sebastian Goldt goldt.sebastian@gmail.com	SISSA, Italy
Kai-Hendrik Henk kai-hendrik.henk@physik.uni-halle.de	Martin-Luther-University Halle-Wittenberg, Germany
Fabian IJpelaar f.i.ijpelaar@rug.nl	University of Groningen, Netherlands
Wolfhard Janke wolfhard.janke@itp.uni-leipzig.de	Leipzig University, Germany
Jonathan Kadmon jonathan.kadmon@mail.huji.ac.il	Hebrew University, Jerusalem, Israel
Shashank Suresha Kallappara shashank-suresha.kallappara@physik.tu-chemnitz.de	Chemnitz University of Technology, Germany
Christian Keup christian.keup@epfl.ch	EPFL, Switzerland

Miriam Klopotek
miriam.klopotek@simtech.uni-stuttgart.de
University of Stuttgart, Germany

Fabio Müller
fmueller@itp.uni-leipzig.de
Leipzig University, Germany

Konstantin Nikolaou
knikolaou@icp.uni-stuttgart.de
University of Stuttgart, Germany

Jeremi Ochab
jeremi.ochab@uj.edu.pl
Jagiellonian University, Cracow, Poland

Andrea Perin
andrea.perin@aalto.fi
Aalto University, Finland

Janett Prehl
janett.prehl@physik.tu-chemnitz.de
Chemnitz University of Technology, Germany

Riccardo Rende
rrende@sissa.it
SISSA, Italy

Frederieke Richert
f.richert@rug.nl
University of Groningen, Netherlands

Beatriz Seoane
beseoane@ucm.es
Université Paris-Saclay, France

Luciano Loris Viteritti
lucianoloris.viteritti@phd.units.it
University of Trieste, Italy

Lei Wang
wanglei@iphy.ac.cn
Chinese Academy of Sciences, China

Martin Weigel
martin.weigel@physik.tu-chemnitz.de
Chemnitz University of Technology, Germany

Roman Worschech
roman.worschech@mis.mpg.de
Leipzig University, Germany

Marvin Wright
wright@leibniz-bips.de
University of Bremen, Germany

Riccardo Zecchina
zecchina@gmail.com
Bocconi University, Milan, Italy

Johannes Zierenberg
johannes.zierenberg@ds.mpg.de
MPI-DS, Göttingen, Germany