ML meets LFT July 24-26 2024, Swansea University: Abstracts

Lingxiao Wang

Learning hadron interactions from lattice QCD

In this study, we develop a deep learning method to learn hadron interactions from Lattice QCD calculated correlations unsupervisedly. We present our approach of using neural networks to model potential functions that are learned from Nambu-Bethe-Salpeter (NBS) wave functions. This allows most general forms of interaction potentials to be incorporated into a Schrödinger-like equation for detailed hadron interaction analysis.

Simran Singh

Testing machine learning against finite size scaling for the chiral phase transition

We extend an existing ML analysis using masked autoregressive flows to estimate the conditional probability density of the chiral condensate conditioned on the number of spatial lattice sites, gauge coupling and quark masses for Nf=5, degenerate light quarks. This was previously done for HISQ fermions by Neumann et.al to determine the Z2 phase boundary in this theory. In this contribution we extend this analysis to previously published data (by F. Cuteri et.al., JHEP 2021) to determine whether the ML analysis can recover the critical mass as determined by the finite size scaling analysis and hence offer an alternative.

Elia Cellini

Stochastic normalizing flows for new theories and observables

Normalizing Flows (NFs) are a class of deep generative models proposed as a promising alternative to traditional Markov Chain Monte Carlo methods in lattice field theory calculations. In this talk, we explore Stochastic Normalizing Flows (SNFs), a combination of NF layers and out-of-equilibrium stochastic updates. We outline the relationship of this extended class of deep generative algorithms with Crooks' theorem and Jarzynski's equality, two fundamental results in non-equilibrium statistical mechanics. We then present numerical results for several observables in lattice-regularized Effective String Theory, a powerful non-perturbative framework used to study confinement in pure gauge theory, and for the calculation of the entanglement entropy in scalar field theory.

Alessandro Nada

Sampling SU(3) pure gauge theory with out-of-equilibrium evolutions and stochastic normalizing flows

Non-equilibrium Monte Carlo simulations based on Jarzynski's equality are a well-understood method to compute differences in free energy and to sample from a target probability distribution that suffers from long autocorrelation times. Out-of-equilibrium evolutions are conceptually similar to Normalizing Flows and they can be combined into a recently-developed architecture called Stochastic Normalizing Flows (SNF). We first outline two computational strategies to mitigate critical slowing down in SU(3) pure gauge theory, either by switching from open to periodic boundary conditions or by changing the gauge coupling over each out-of-equilibrium evolution, with a focus on the promising scaling with the volume. Then, we introduce a SNF with gauge-equivariant layers between the out-of-equilibrium Monte Carlo updates, we analyse the improvements over the purely stochastic approach, and we conclude with a discussion on future prospects.

Ankur Singha

Multilevel sampling of lattice theories using RG-inspired autoregressive models

We introduce a method for multilevel sampling of lattice theories, using insights from Renormalization Group (RG) analysis to design model architectures from coarse to fine levels. The interaction range of spins at each level, derived from the RG transformation of the lattice Hamiltonian, determines the CNN kernel size in the model. By training the model at a coarse level first and then using it as the initial point for the next level, we enhance training efficiency. This multilevel approach addresses the limitations of existing models, offering better scalability and accuracy for large-scale lattice sampling problems.

Tej Kanwar

Neural-network contour deformations for the signal-to-noise problem

Path integral contour deformations provide a systematically exact method for redefining lattice field theory observables to minimize their statistical noise. This talk details recent developments based on applying a hierarchical U-net neural network architecture to define families of contour deformations for SU(N) variables. Numerical optimization within this family enables exponential improvements in the signal-to-noise ratio of Wilson loops in proof-of-principle applications to 2+1D SU(N) lattice gauge theories. The choice of gauge-fixing is a key ingredient in the success of this method, and a potential method to optimize over gauge-fixing schemes will be discussed.

Alexander Rothkopf

Learning optimal kernels for real-time complex Langevin

In this talk I will present recent progress in extending the range of validity of complex Langevin real-time simulations for 1+1d scalar field theory using a simple reinforcement learning approach. Our strategy relies on supplying complex Langevin with prior information of the simulated system to learn an optimal non-neutral kernel that modifies the convergence properties of the underlying stochastic process. Ongoing work towards retaining continuum symmetries in the discrete setting, crucial to exploit as prior information, are discussed.

Biagio Lucini

Topological data analysis for lattice gauge theories

It is well known that in non-Abelian gauge theories topologically non-trivial configurations play a crucial role in determining non-perturbative behaviour. Topological data analysis is an investigation framework that enables robust definitions of topology for discrete sets of points. Therefore, lattice studies of gauge theories with non-trivial topological content provide ideal application scenarios for topological data analysis. In this talk, after a general introduction to the subject, I will show how topological data analysis can be used to characterise configurations with non-trivial topological features. Building on those results, I will provide constructions of (pseudo-)order parameters that enable to study quantitatively the deconfining phase transition in U(1), SU(2) and SU(3) lattice gauge theories. The role of machine learning in this process will be highlighted.

Ryan Abbott

Progress in normalizing flows for 4d gauge theories

Normalizing flows have recently arisen as a potential tool for aiding in sampling lattice field theories. In this talk I will give an overview of our groups' recent progress in applying normalizing flows to 4-dimensional nonabelian gauge theories, as well as current efforts to scale normalizing flows towards modern lattice field theory calculations.

Fernando Romero Lopez

Applications of flow models to the generation of correlated lattice QCD ensembles

Machine-learned normalizing flows can be used in the context of lattice quantum field theory to generate statistically correlated ensembles of lattice gauge fields at different action parameters. In this talk, we show examples on how these correlations can be exploited for variance reduction in the computation of observables. Three different proof-of-concept applications are presented: continuum limits of gauge theories, the mass dependence of QCD observables, and hadronic matrix elements based on the Feynman-Hellmann approach. In all three cases, statistical uncertainties are significantly reduced when machine-learned flows are incorporated as compared with the same calculations performed with uncorrelated ensembles or direct reweighting.

Mathis Gerdes

Exploring continuous normalizing flows for gauge theories

We explore equivariant architectures for continuous flows on gauge theories, taking inspiration from previous successes for scalar theories, and going beyond gradient flows. We focus on expressivity of the vector field architecture while maintaining computational efficiency, showing promising results for pure SU(2) Yang-Mills theory.

Akio Tomiya

MLPhys in Japan and developments of CASK: Gauge symmetric transformer

This presentation will cover three main topics related to the MLPhys project, which integrates machine learning and AI into physics to tackle complex computational challenges. First, I will introduce the MLPhys project and its goals. Funded by Japan's Grants-in-Aid for Scientific Research (KAKENHI) and supported by the Fugaku supercomputer, this project aims to combine machine learning with physics. Next, I will discuss our development of the Gauge Symmetric Transformer (CASK: Covariant Attention for SU(N) Kernel). CASK, applicable to fermions and extendable as stout smearing, is demonstrated within self-learning HMC (SLHMC) to improve lattice QCD simulations. Finally, I will briefly cover the application of sparse modeling and All-Mode Averaging (AMA) techniques for computing 1/D^n, enhancing the precision of physical calculations.

David Müller

Lattice simulations with machine-learned classically perfect fixed-point actions

Fixed-point actions are classically perfect lattice actions, i.e., they are free from classical lattice artifacts. Furthermore, they exhibit suppressed quantum artifacts. Monte Carlo simulations employing such actions may provide a way to efficiently approach the continuum limit on coarse lattices, thereby avoiding critical slowing down and topological freezing. Extending our previous work, we use lattice gauge equivariant convolutional neural networks (L-CNNs) to approximate a fixed-point action for SU(3) gauge theory in four dimensions to previously unseen accuracy. Using this new parametrization, we perform HMC simulations and classically perfect gradient flow. Our self-consistent approach aims to extract gradient flow observables on much coarser lattices compared to simulations using the Wilson action.

Chanju Park

Empirical phase diagram of neural network and spin glass theory

The stochastic gradient update of a neural network can be described by a Langevin equation with the strength of fluctuation proportional to $\alpha/|B|$, which can be interpreted as the temperature of the system. Here we show an empirical phase diagram of one hidden layer neural network with tangent hyperbolic activation function, where three distinctive phases can be classified. The control parameter of the phase transition is shown to be $\alpha/|B|$ and the initial width of the weight matrix σ_W. Then, we argue that the phase diagram can be understood in the context of the spin glass theory, where each phase corresponds to ferromagnetic, paramagnetic, and spin glass.

Tomasz Stebel

Entanglement entropy with generative neural networks

In this talk I will describe a method to estimate Rényi entanglement entropy, which is based on the replica trick and generative neural networks with explicit probability estimation. We demonstrate it on a one-dimensional quantum Ising spin chain. As the generative model, we use a hierarchy of autoregressive deep neural networks, allowing us to simulate up to 32 spins. We calculate the second Rényi entropy and its derivative. This method can be extended to any spin system or lattice field theory if the appropriate sampling algorithms are available.

Shiyang Chen

Exploring generative networks for manifolds with non-trivial topology

The expressive power of neural networks in modelling complex distributions is desirable to bypass topological freezing and critical slowing down in simulations of lattice field theory. Some approaches suffer from problems with topology, which may lead to some classes of configurations not being generated. In this talk, I will present a novel generative approach inspired by a model previously introduced in the ML community (GFlowNets) and simulate triple ring models and phi^4 lattice model to demonstrate the capabilities of the method to solve issues connected with ergodicity.

Gert Aarts

Weight matrix dynamics and Dyson Brownian motion

We apply concepts from random matrix theory to describe stochastic weight matrix dynamics, using the framework of Dyson Brownian motion. We derive the linear scaling rule between the learning rate of the optimisation and the mini-batch size, and identify universal and non-universal aspects of the weight matrix dynamics. We test our hypothesis in the (near-)solvable case of the Gaussian Restricted Boltzmann Machine, and explicitly identify the Wigner surmise and semi-circle, and the linear scaling rule.

Matteo Favoni

Towards the application of random matrix theory to neural networks

Random matrix theory was first examined in the context of nuclear physics to investigate properties of heavy atom nucleus spectra. This theory is suited for an application to machine learning algorithms, specifically to study the properties of their weight matrices. In this presentation, we study a teacher-student model and discuss the role of hidden layers, showing that the matrix eigenvalues characterizing well-trained models are distributed according to the Wigner's surmise and a generalized version of the Wigner's semicircle.