At the interface of physics, mathematics and artificial intelligence

Report of Contributions

Contribution ID: **2** Type: **not specified**

High dimensional Stochastic Gradient Descent, summary statistics and effective dynamics

I will survey recent progress about the behavior of SGD in high dimensional settings.
(in joint works with Aukosh Jagannath and Reza Gheissari, and upcoming work with them and Jiaoyang Huang)

Presenter: BEN AROUS, Gérard (New York U.)

Contribution ID: **3** Type: **not specified**

Recent developments in (hyperbolic) string vertices

String vertices are the geometric ingredient underlying string field theory. In this talk, I will outline the bootstrap formalism for constructing hyperbolic string vertices. The emphasis will be on how machine learning may provide a natural numerical framework to explicitly realize this construction. Based on 2211.09129, 2302.12843.

Presenter: FIRAT, Atakan Hilmi (MIT; IAIFI)

Contribution ID: **4** Type: **not specified**

Statistical Physics of Energy-Based generative models

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.

Presenter: Prof. SEOANE, Beatriz (Paris-Saclay University)

Contribution ID: **5** Type: **not specified**

Towards a phenomenological understanding of neural networks

A theory of neural networks (NNs) built upon collective variables would provide scientists with the tools to better understand the learning process at every stage. I argue that a fruitful path for this endeavour of understanding non-linear neural network dynamics is to consider the analogy with physical systems. As an example, I demonstrate that the dynamics of neural networks trained with gradient descent and the dynamics of scalar fields in a flat, vacuum energy dominated Universe are structurally profoundly related. This duality provides the framework for synergies between these systems, to understand and explain neural network dynamics and new ways of simulating and describing early Universe models. In an attempt to capture the dynamics in the non-linear regime effectively, I then introduce two such variables, the entropy and the trace of the empirical neural tangent kernel (NTK) built on the training data passed to the model. We empirically analyze the NN performance in the context of these variables and find that there exists a correlation between the starting entropy, the trace of the NTK, and the generalization of the model computed after training is complete. This framework is then applied to the problem of optimal data selection for the training of NNs.

Presenter: KRIPPENDORF, Sven (LMU)

Contribution ID: **6** Type: **not specified**

GeN S: the new generation of sampling algorithms

The task of learning non-trivial probability densities is a crucial problem in machine learning with an uncountable number of applications in, among others, computer vision, sound synthesis, text generation, and natural sciences. The subfield of machine learning that leverages deep learning to learn complicated probability distributions and samples from them is known as Generative AI. Recently, the relevance of this problem has been extensively studied in several domains where deep generative models have been proposed to sample non-trivial Boltzmann-like densities in, for instance, lattice quantum field theory and statistical mechanics. Learning the underlying density, namely a normalized Boltzmann distribution, greatly improves the sampling task and opens the possibility of estimating physical observables, such as the partition function and related thermodynamic observables, which are notoriously hard to be estimated using standard sampling methods. I refer to this novel approaches as GEnerative Neural Samplers, or, "Gen S" for short: the new generation of sampling algorithms.

In my talk, I'll start by describing some key general concepts such as the path integral formalism of quantum field theory and so-called Generative Neural Samplers with Exact Probability (GNSEP), namely, generative models having the desirable feature of allowing for an exact form of the probability density, thereby enabling exact likelihood inference. I will then discuss in more detail two examples of GNSEP, e.g., normalizing flows and variational autoregressive models, and present two Gen S algorithms, namely neural-enhanced sampling approaches, known as NeuralMCMC (NMCMC) and Neural Importance Sampling (NIS). Application of the NIS approach for estimating thermodynamic observables in the context of statistical mechanics and scalar lattice field theory will be presented. I'll conclude the talk by summarizing the results and briefly discussing the challenges and the current limitations of these methods.

Presenter: NICOLI, Kim (Bonn U.)

Contribution ID: **7** Type: **not specified**

Machine Learning and Flows for Lattice QCD

Recently, there have been some very impressive advances in generative models for sound, text and images. In this talk, I will look into applications of generative models to Lattice QCD. The models I will consider are flows, which are families of diffeomorphisms transforming simple base distributions into complicated target distributions. I will explain why we believe that flows are suitable for Lattice QCD. I will give details on how we built these flows, and explain how known symmetries of LQCD can be incorporated into them.

Presenter: RACANIÈRE, Sébastien (DeepMind)

Contribution ID: **8** Type: **not specified**

An Equivariant Neural Decoder for Quantum Error Correction

Quantum error correction is a critical component for scaling up quantum computing. Given a quantum code, an optimal decoder maps the measured code violations to the most likely error that occurred, but its cost scales exponentially with the system size. Neural network decoders are an appealing solution since they can learn from data an efficient approximation to such a mapping and can automatically adapt to the noise distribution. In this work, we introduce a data efficient neural decoder that exploits the symmetries of the problem. We characterize the symmetries of the optimal decoder for the toric code and propose a novel equivariant architecture that achieves state of the art accuracy compared to previous neural decoders.

Presenter: Prof. BONDESAN, Roberto (Imperial College London)

Contribution ID: **9** Type: **not specified**

Quantum error mitigation for Physics and vice versa

First, I will describe a specific quantum error mitigation scheme designed for parametric circuits accessible by classical computations in some range of their parameters. I will demonstrate the work of the scheme on the example of the 4-spin anti-ferromagnetic Ising model in the transverse field, and discuss possible applications to the sign problem in Monte Carlo simulations.
Then, I will indicate relations between quantum error mitigation and perturbative computations and will introduce a concept of error mitigation based on the variational perturbation theory.

Presenter: SAZONOV, Vasily (CEA-LIST)

Contribution ID: **10** Type: **not specified**

The math of training large neural networks, with some analogy to physics

Recently, the theory of infinite-width neural networks led to the first technology, muTransfer, for tuning enormous neural networks that are too expensive to train more than once. For example, this allowed us to tune the 6.7 billion parameter version of GPT-3 using only 7% of its pretraining compute budget, and with some asterisks, we get a performance comparable to the original GPT-3 model with twice the parameter count. In this talk, I will explain the core insight behind this theory. In fact, this is an instance of what I call the *Optimal Scaling Thesis*, which connects infinite-size limits for general notions of “size”to the optimal design of large models in practice. I'll end with several concrete key mathematical research questions whose resolutions will have incredible impact on the future of AI.

Presenter: YANG, Greg (Microsoft Research)