The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

Speaker: Bruno LOUREIRO (ENS Paris, France)

Overlaps are all you need

In this tutorial I will discuss some recent progress on our understanding of the one-pass stochastic gradient dynamics for two-layer neural networks based on the tools from statistical physics. The discussion will be two-fold:

First, I will derive a unified picture of the dynamics is terms of order parameters known as "overlaps" in the statistical physics literature. This will bridge the seminal work of David Saad and Sara Solla [1] for high-dimensional narrow networks with the more recent work in the so-called "mean-field" or wide regime [2,3].

Second, I will discuss how this same idea can be adapted to characterise the class of functions that can be learned with a single gradient step [4].

[1] https://journals.aps.org/pre/abstract/10.1103/PhysRevE.52.4225
[2] https://proceedings.neurips.cc/paper_files/paper/2022/hash/939bb847ebfd14c6e4d3b5705e562054-Abstract-Conference.html
[3] https://arxiv.org/abs/2302.05882
[4] https://arxiv.org/abs/2305.18270