Francesca Mignacco (IPHT Saclay, France)

Title: Understanding the dynamics of multi-pass SGD in high dimensional non-convex problems

Abstract: Artificial neural networks trained with gradient-based algorithms have achieved impressive performances in a variety of applications. In particular, the stochastic gradient descent algorithm (SGD) proved to be surprisingly efficient in navigating high-dimensional complex loss landscapes. However, the theory behind this practical success remains largely unexplained. In this talk, I will consider phase retrieval as a prototype high-dimensional non-convex problem and show how to apply dynamical mean-field theory from statistical physics to track analytically the learning dynamics of multi-pass SGD. I will further discuss some intriguing properties of the landscape and the role played by different types of noise of stochastic gradient-based algorithms.