

## PATT Position in APPLIED MATHEMATICS

Talks via zoom link:

**Thursday, 28 January 2021**

Sujet : PATT position in APPLIED MATH - presentation

Heure : 28 janv. 2021 09:45 AM Zurich

Participer à la réunion Zoom

<https://epfl.zoom.us/j/86059374376>

**10:00 – 11:00      Dr Lénaïc CHIZAT – Université Paris-Sarclay, France**

Title: Analysis of Gradient Descent on Wide Two-Layer ReLU Neural Networks

Abstract:

Artificial neural networks are a family of parametric models which, given a training set of sample/label pairs, can be trained to predict the labels of new samples. To do so, the training algorithm updates the parameters using variants of the gradient descent method on a well-chosen objective function (the empirical risk, with potentially a regularization term). In this talk, we propose an analysis of gradient descent on wide two-layer ReLU neural networks (networks with many parameters but only one simple "positive part" non-linearity) that leads to sharp characterizations of the learned predictor. The main idea is to study the dynamics when the width of the neural network goes to infinity, which can be written as a Wasserstein gradient flow. While this dynamics evolves on a non-convex landscape, we show that when the parameters are initialized properly, its limit is a global minimizer. We also study the "implicit bias" of this algorithm in various situations: among all the minimizers, we show that it selects a specific one which depends, among other things, on the initialization and the choice of objective function. Along the way, we discuss what these results tell us about the statistical performance of these models.

This is based on joint work with Francis Bach.