



Institut für Mathematik

Seminar zur Stochastik

Donnerstag, 29. Juni 2023
15 Uhr s.t.
SR 025 August-Bebel-Str. 4

Herr Dr. Sebastian Kassing
(Universität Bielefeld)

“Stochastic Modified Flows, Mean-Field Limits and Dynamics of Stochastic Gradient Descent”

Abstract: Stochastic gradient descent algorithms (SGD) are the most common way to train neural networks. Due to the non-convexity of the corresponding loss landscape, the analysis of the optimization dynamics is highly challenging. A popular technique for analyzing SGD is to approximate the discrete dynamics by suitable stochastic differential equations (SDEs) called stochastic modified equations. In this talk, we propose new SDEs that appear as limiting dynamics for SGD in the small learning rate regime. These SDEs are driven by a cylindrical Brownian motion and improve the stochastic modified equations by having regular diffusion coefficients and by matching the multi-point statistics. As a second contribution, we introduce distribution dependent stochastic modified flows which we prove to describe the fluctuating limiting dynamics of SGD in the small learning rate - infinite width scaling regime.

Alle Interessierte sind herzlich eingeladen