



Einladung zum Oberseminar Mathematik des Maschinellen Lernens und Angewandte Analysis

Julius-Maximilians-Universität Würzburg
Professur für Mathematik des Maschinellen Lernens

M.Sc. Anna Shalova

Eindhoven University of Technology

Stationary Measures of Noisy Transformers

Transformers are, arguably, the state-of-the-art models in natural language processing and various other applications. To build a theoretical understanding of transformers, a toy transformer model, allowing for a wide range of analytical tools, was recently introduced by Geshkovski et al. (2024). In this talk I will discuss the aforementioned toy transformer model and its noisy counterpart, McKean-Vlasov PDE with exponential kernel. The talk is mostly based on our recent work with André Schlichting (arXiv:2412.14813) in which we study stationary solutions of McKean-Vlasov equation on Riemannian manifolds. Our analysis extends the equivalence of the energetic problem formulation to the manifold setting and allows us to characterize critical points of the corresponding free energy functional. On a sphere, we employ the properties of spherical convolution to study the bifurcation branches around the uniform state and characterize the phase transition of the model. I will give an overview of the results and show how our findings apply to noisy transformers.

Ort: Mathematik Ost, Seminarraum 40.01.003

Zeit: Mittwoch, 28.05.2025 14:15

Zu diesem Vortrag laden wir Sie herzlich ein.

gez. Leon Bungert