Friedrich-Alexander-Universität
Research Center for
Mathematics of Data | MoD

FAU MoD Lecture Series

#FAUMoDLecture

WWW.MOD.FAU.EU

Holger Rauhut

UNIVERSITÄT MÜNCHEN

The implicit bias phenomenon in deep learning

Deep neural networks are usually trained by minimizing a non-convex loss functional via (stochastic) gradient descent methods. Unfortunately, the convergence properties are not very well-understood. Moreover, a puzzling empirical observation is that learning neural networks with a number of parameters exceeding the number of training examples often leads to zero loss, i.e., the network exactly interpolates the data. Nevertheless, it generalizes very well to unseen data, which is in stark contrast to intuition from classical statistics which would predict a scenario of overfitting.

A current working hypothesis is that the chosen optimization algorithm has a significant influence on the selection of the learned network. In fact, in this overparameterized context there are many global minimizers so that the optimization method induces an implicit bias on the computed solution. It seems that gradient descent methods and their stochastic variants favor networks of low complexity (in a suitable sense to be understood), and, hence, appear to be very well suited for large classes of real data. Initial attempts in understanding the implicit bias phenomen considers the simplified setting of linear networks, i.e., (deep) factorizations of matrices. This has revealed a surprising relation to the field of low rank matrix recovery (a variant of compressive sensing) in the sense that gradient descent favors low rank matrices in certain situations. Moreover, restricting further to diagonal matrices, or equivalently factorizing the entries of a vector to be recovered, shows connection to compressive sensing and l1-minimization.

WHEN

Tue. **December 3, 2024**
14:30H (Berlin time)

WHERE

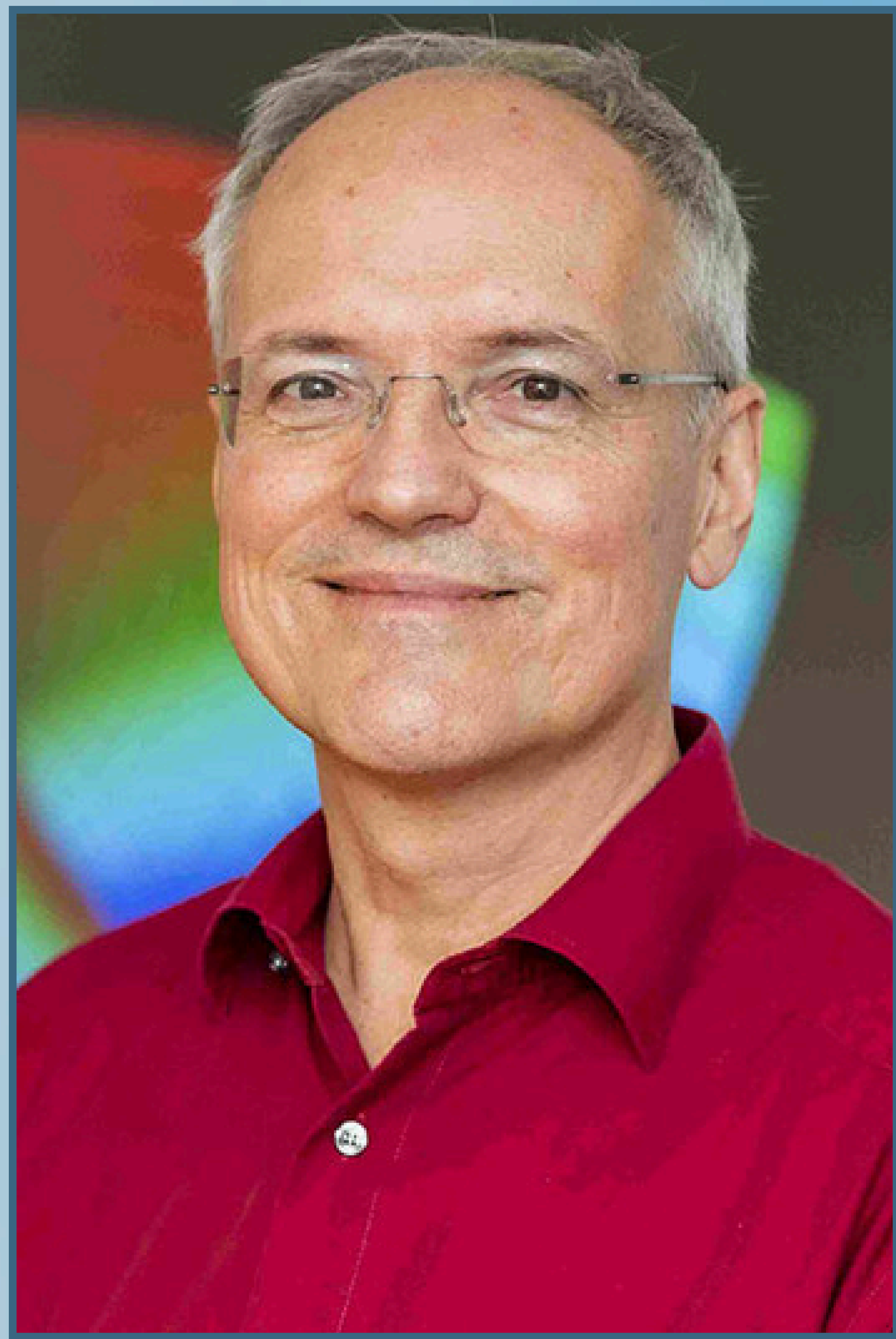
On-site / Online

FAU. Friedrich-Alexander-Universität Erlangen-Nürnberg
Room H13
Johann-Radon-Hörsaal
Cauerstraße 11
91058 Erlangen
Bavaria, Germany

FAU Zoom link:
https://go.fau.de/1bcfg

Meeting ID:
667 9081 1368

PIN code:
716845



Christian Bär

UNIVERSITÄT POTSDAM

Counterintuitive approximations

One of the oldest questions in differential geometry is whether every curved space arises as a subspace of Euclidean space $\mathbb{R}^n$. More precisely, does every Riemannian manifold admit an isometric (i.e. length preserving) smooth embedding into $\mathbb{R}^n$? Nash showed that the answer is yes but the dimension n of the Euclidean space may be much larger than that of the manifold. Often there are non-isometric embeddings of much lower codimension.
The Nash-Kuiper embedding theorem is a prototypical example of a counterintuitive approximation result in geometry: any short (but highly non-isometric) embedding of a Riemannian manifold into Euclidean space can be approximated by isometric embeddings. They are generally not smooth but of regularity C^1. This implies that any surface with a given geometry can be isometrically C^1-embedded into an arbitrarily small ball in $\mathbb{R}^3$. For C^2-embeddings this completely false due to curvature restrictions.

After explaining this historical and conceptual context, I will present a general result which ensures approximations by maps satisfying strongly overdetermined equations (such as being isometric) on open dense subsets. This will be illustrated by three examples: Lipschitz functions with surprising derivative, surfaces in 3-space with unexpected curvature properties, and a similar statement for abstract Riemannian metrics on manifolds. Our method is based on “cut-off homotopy”, a concept introduced by Gromov in 1986.
This is based on joint work with Bernhard Hanke.