

# Curriculum vitæ: Felix Dangel



Felix Dangel is a Postdoctoral researcher at the [Vector Institute](#) in Toronto. He finished his PhD at [Philipp Hennig's](#) lab at the University of Tübingen and the Max Planck Institute for Intelligent Systems. His PhD focused on extending gradient backpropagation to efficiently extract higher-order geometrical and statistical information about the loss landscape of neural networks to improve their training and inspire novel algorithmic research. Before, he studied Physics at the University of Stuttgart with a focus on the simulation of quantum many-body systems with tensor networks. He is passionate about

- developing automatic differentiation tricks to tackle efficient extraction of richer deep learning quantities, like second-order and per-sample information, and integrating that functionality into machine learning libraries,
- using these quantities to build better algorithms or gain insights into deep learning phenomena, and
- releasing code that empowers the community (see for example [cockpit](#), [backpack](#), and [vivit](#)).

## Education

**now** Postdoctoral researcher, Vector Institute, Toronto

–

**2023** With: Prof. Dr. Yaoliang Yu

**2023** **PhD in Computer Science**, Max Planck Institute for Intelligent Systems & University of Tübingen

– Thesis: *Backpropagation beyond the Gradient*

**2018** Advisor: Prof. Dr. Philipp Hennig

**2018** **Researcher**, University of Stuttgart

– Paper: *Topological invariants in dissipative extensions of the Su-Schrieffer-Heeger model*

**2017** Host: Institute for Theoretical Physics I

**2017** **MSc in Physics**, University of Stuttgart

– Thesis: *Bosonic many-body systems with topologically nontrivial phases subject to gain and loss*

**2015** Advisor: P. D. Holger Cartarius

**2015** **BSc in Physics**, University of Stuttgart

– Thesis: *Microscopic description of a coupling process for  $\mathcal{PT}$ -symmetric Bose-Einstein condensates*

**2012** Advisor: Prof. Dr. Günter Wunner

## Publications

- *Structured Inverse-Free Natural Gradient: Memory-Efficient & Numerically-Stable KFAC for Large Neural Nets*, pre-print 2023  
W. Lin, **F. Dangel**, R. Eschenhagen, K. Neklyudov, A. Kristiadi, R. Turner, A. Makhzani ([code](#))

- *On the Disconnect Between Theory and Practice of Overparametrized Neural Networks*, pre-print 2023  
J. Wenger, **F. Dangel**, A. Kristiadi ([pdf](#) | [arXiv](#))
- *Convolutions Through the Lens of Tensor Networks*, pre-print 2023  
**F. Dangel** ([pdf](#) | [arXiv](#) | [code](#))
- *The Geometry of Neural Nets' Parameter Spaces Under Reparametrization*, NeurIPS 2023  
A. Kristiadi, **F. Dangel**, P. Hennig ([pdf](#) | [arXiv](#))
- *ViViT: Curvature access through the generalized Gauss-Newton's low-rank structure*, TMLR 2022  
**F. Dangel**, L. Tatzel, P. Hennig ([pdf](#) | [journal](#) | [arXiv](#) | [code](#) | [www](#))  
We present novel ways to compute with curvature that allow quantifying noise in both gradient and curvature.
- *Cockpit: A Practical Debugging Tool for Training Deep Neural Networks*, NeurIPS 2021 (poster)  
F. Schneider, **F. Dangel**, P. Hennig ([pdf](#) | [conference](#) | [arXiv](#) | [code](#) | [www](#) | [video](#))  
Just like classic debuggers help figuring out bugs in code, our neural debugger assists deep learning engineers in troubleshooting training. It is used to debug [non-standard deep learning tasks](#) that do not work out of the box in contrast to standard problems, and may soon be introduced as a tool to ML students in the classroom.
- *BackPACK: Packing more into backprop*, ICLR 2020 (spotlight)  
**F. Dangel**, F. Kunstner, P. Hennig ([pdf](#) | [conference](#) | [arXiv](#) | [code](#) | [www](#) | [video](#))  
The paper received [perfect scores](#) before the rebuttal. BackPACK has raised awareness in the community on issues with autodiff in ML frameworks. Since its introduction, major engines like TensorFlow (Google) and PyTorch (Meta) have provided some of the functionalities as first-party solutions ([TF](#), [PT](#)) with our approach.
- *Modular Block-diagonal Curvature Approximations for Feedforward Architectures*, AISTATS 2020 (poster)  
**F. Dangel**, S. Harmeling, P. Hennig ([pdf](#) | [conference](#) | [arXiv](#) | [code](#) | [video](#))  
A take on the Hessian chain rule for neural networks: Layer-wise Hessians can be computed with a procedure similar to gradient backpropagation, where the propagated quantities are the Hessians.
- *Topological invariants in dissipative extensions of the Su-Schrieffer-Heeger model*, Phys. Rev. A 2018  
**F. Dangel**, M. Wagner, H. Cartarius, J. Main, G. Wunner ([pdf](#) | [journal](#) | [arXiv](#))
- *Numerical calculation of the complex Berry phase in non-Hermitian systems*, Acta Polytechnica 2018  
M. Wagner, **F. Dangel**, H. Cartarius, J. Main, G. Wunner ([pdf](#) | [journal](#) | [arXiv](#))

## Talks & Workshops

- Workshop poster presentation at NeurIPS [OPT23](#) on "Structured Inverse-Free Natural Gradient Memory-Efficient & Numerically-Stable KFAC for Large Neural Nets" ([poster](#))
- Invited talk at [Perimeter Institute](#) Machine Learning Initiative seminar titled "Deep Learning Convolutions Through the Lens of Tensor Networks" ([recording](#), [slides](#))
- Poster presentation at the [ELLIS Doctoral Symposium \(EDS\) 2022](#) in Alicante ([poster](#))
- Invited talk at the [ELISE Theory Workshop on ML Fundamentals 2022](#) at EURECOM in Sofia Antipolis
- Poster presentation at the [ELLIS Theory Workshop 2022](#) in Arenzano
- Session chair at the [Franco-German Research and Innovation Network on AI](#), June 2022
- Co-organization of the [ELLIS Doctoral Symposium \(EDS\) 2021](#) in Tübingen, to be held [2022 in Alicante](#)

- Invited overview talk about DL, seminar for [Integrated Engineering](#) students, DHBW CAS in Heilbronn
- Talk at the [DPG Spring Meeting](#) of the atomic Physics and quantum optics section 2017 in Mainz
- Participation at the "[Ferienakademie 2015](#)" summer school, organized by the TU Munich, the University of Stuttgart, and the FAU Erlangen in Sarntal (Northern Italy), talk about Lattice Boltzmann Methods.
- Participation at the "[Ferienakademie 2014](#)" summer school, organized by the TU Munich, the University of Stuttgart, and the FAU Erlangen in Sarntal (Northern Italy), talk about NMR & MRI

## Teaching & Reviewing

- Between fall 2018 and summer 2022, Felix taught seven (7) iterations of [software development practicals](#). In these courses, three PhD students supervise ~15 students whose task is to develop a machine learning prediction system for the German soccer league over the course of one term ([example](#)). The overall workload for a student is ~180 hours and the focus lies heavily on teaching good software development practices.
- Felix has worked with various students on different projects:
  - Elisabeth Knigge (high school student, summer internship) worked on making deep learning optimization methods more approachable to non-experts through visualization. By combining Tübingen's interesting topology with optimization methods, she created [intriguing wall art](#) for the Tübingen AI building.
  - [Jessica Bader](#) (research assistant) worked on broadening BackPACK's support for Kronecker-factorized curvature for Bayesian deep learning. She wrote the [interface for negative log-likelihood losses](#) to support KFAC and to enable applications with their Laplace approximation via the [laplace-torch](#) library.
  - [Tim Schäfer](#) ([Master thesis](#)), now a PhD student with [Anna Levina](#), added support for ResNets and recurrent architectures to BackPACK. The underlying converter that makes these architectures compatible can easily be enabled through an optional argument while extending the model.
  - [Shrisha Bharadwaj](#) (research assistant) improved BackPACK's code quality through additional tests, docstrings, and extended support for two-dimensional convolution to 1d and 3d.
  - [Paul Fischer](#) ([research project](#)), now a PhD student with [Christian Baumgartner](#), implemented and analyzed Hessian backpropagation for batch normalization to speed up its Hessian-vector product, [that can be slow \(page 7\)](#), through structural knowledge.
  - Christian Meier (Bachelor thesis): Activity prediction in smart home environments via Markov models.
- He reviewed for top-tier machine learning conferences and journals
  - [Advances in Neural Information Processing \(NeurIPS\)](#) (2020, 2021, 2022 ([HITY workshop](#)), 2023)
  - [International Conference for Machine Learning \(ICML\)](#) (2020, 2021, 2022)
  - [Journal of Machine Learning Research \(JMLR\)](#) (2021).
  - International Conference on Artificial Intelligence and Statistics ([AISTATS](#)) (2024)
- Served as reviewer for the [Vector Scholarship in Artificial Intelligence 2023-2024](#)