Lorenzo Noci

Ph.D candidate in Deep Learning at ETH Zurich

lorenzo.noci@inf.ethz.ch +39 3920855212 Twitter LinkedIn Website Google Scholar

Education

Sep. 2021 - June/Sep 2025 **Ph.D. in Deep Learning** ETH Zurich, Switzerland

Supervised by Prof. Thomas Hofmann at the Data Analytics Lab. Currently working on scaling limits of neural networks at large width and depth. Applications: architecture design and hyperparameter transfer.

Sep. 2018 - June 2021 **M.Sc. in Data Science** ETH Zurich, Switzerland

Thesis title: "Understanding Bayesian Neural Networks through the Cold Posterior Effect, the Prior Predictive Distribution and PAC-Bayes Bounds". Average Grade: 5.5/6.0

Oct. 2014 - July 2017 **Bachelor in Computer Science and Engineering** Politecnico di Milano, Italy

Final mark: 110/110 with honors. Average Grade: 29.22/30

Sep. 2015 - June 2018 **Bachelor in Electronic Information Engineering** Tongji University, Shanghai, China

Participated to the double degree project 'Politong' between Politecnico di Milano and Tongji University with fully founded scholarship.

Featured Publications

Authors who equally contributed to a publication are marked with a †.

1. Blake Bordelon†, **Lorenzo Noci**†, Mufan Bill Li, Boris Hanin, Cengiz Pehlevan "Depthwise Hyperparameter Transfer in Residual Networks: Dynamics and Scaling Limit." In *International Conference on Learning Representations (ICLR), 2024*.
2. **Lorenzo Noci**†, Chuning Li†, Mufan Bill Li†, Bobby He, Thomas Hofmann, Chris Maddison, Daniel M. Roy. "The Shaped Transformer: Attention Models in the Infinite Depth-and-Width Limit" In *Advances in Neural Information Processing Systems (NeurIPS), 2023*
3. **Lorenzo Noci**†, Sotiris Anagnostidis†, Luca Biggio†, Antonio Orvieto†, Sidak Pal Singh†, and Aurelien Lucchi. "Signal propagation in transformers: Theoretical perspectives and the role of rank collapse." In *Advances in Neural Information Processing Systems (NeurIPS), 2022*
4. **(Oral)** Gregor Bachmann†, **Lorenzo Noci**†, and Thomas Hofmann. "How Tempering Fixes Data Augmentation in Bayesian Neural Networks." In *International Conference on Machine Learning, pages 1244–1260. PMLR, 2022*
5. **(Spotlight) Lorenzo Noci**†, Gregor Bachmann†, Kevin Roth†, Sebastian Nowozin, and Thomas Hofmann. "Precise Characterization of the Prior Predictive Distribution of deep relu networks." In *Advances in Neural Information Processing Systems, 34, 2021*

Working Experience

Mar. 2020 - Sep. 2020 **Research Scientist (Data Science) Intern** Amazon, Luxembourg

I worked in the Operation Research Science Team. My project was **productionalized and adopted across Europe and US**, and focused on adaptive importance sampling for combinatorial optimization in high dimension.

Apr. 2019 - Mar. 2020 **Research Assistant - Data Engineer** ETH Zurich

Worked part time with the Biomedical Informatics (BMI) group at ETH on building efficient data processing pipelines.

Sep. 2017 - Sep. 2018 **Data Scientist** BUYINT, Shanghai, China

Worked on extracting features from financial documents, recommending companies for potential acquisition.

Awards and other Academic Activity

Sept. 2024 **Google PhD Fellowship 2024**. Area: Machine Intelligence

Mar. 2025 - Jun. 2025 **(Incoming) Visiting PhD** Supervisor: Prof. Boris Hanin Princeton University, NJ, United States

Aug. 2024 **Visiting Research Collaborator** Supervisor: Prof. Daniel Roy Vector Institute, Toronto, Canada

Apr. 2024 - Sep. 2025 **Visiting Research Collaborator**. Host: Antonio Orvieto Max Planck Inst. for Intelligent Systems

Talks

Dec. 2024 **EPFL Pre-NeurIPS Event**
Super Consistency of Neural Network Landscapes and Learning Rate Transfer

Aug. 2024 **Vector Institute**
Hyperparameter Transfer in Wide-and-Deep Neural Networks with Insights from the Loss Landscape

Oct. 2024 **Max Planck Institute for Intelligent Systems**
The Infinite Depth-and-Width Limit(s) of Neural Networks

Nov. 2023 **Institute of Machine Learning, ETH Zurich**
The Infinite Depth-and-Width Limit(s) of Neural Networks

Jan. 2023 **Google Zurich**.
Signal Propagation in Transformers

Publications & Preprints

Authors who equally contributed to a publication are marked with a †.

1. Bobby He†, **Lorenzo Noci**, Daniele Paliotta, Imanol Schlag, Thomas Hofmann. "Understanding and Minimising Outlier Features in Neural Network Training." In *Advances in Neural Information Processing Systems (NeurIPS), 2024*.
2. **Lorenzo Noci**†, Alex Meterez†, Thomas Hofmann, Antonio Orvieto. "Why do Learning Rates Transfer? Reconciling Optimization and Scaling Limits for Deep Learning." In *Advances in Neural Information Processing Systems (NeurIPS), 2024*.
3. Jacopo Graldi†, Giulia Lanzillotta, **Lorenzo Noci**, Benjamin Grewe, Thomas Hofmann. "Exploring the Limits of Feature Learning in Continual Learning" In *Continual FoMo Workshop at NeurIPS, 2024*.
4. Blake Bordelon†, **Lorenzo Noci**†, Mufan Bill Li, Boris Hanin, Cengiz Pehlevan. "Depthwise Hyperparameter Transfer in Residual Networks: Dynamics and Scaling Limit." In *International Conference on Learning Representations (ICLR), 2024*.
5. Kai Lion†, **Lorenzo Noci**, Thomas Hofmann, Gregor Bachmann. "How Good is a Single Basin?" In *Artificial Intelligence and Statistics (AISTATS), 2024*
6. **(Spotlight)** Sotiris Anagnostidis†, Dario Pavllo, Luca Biggio, **Lorenzo Noci**, Aurelien Lucchi, Thomas Hofmann "Dynamic Context Pruning for Efficient and Interpretable Autoregressive Transformers". In *Advances in Neural Information Processing Systems (NeurIPS), 2023*
7. **Lorenzo Noci**†, Chuning Li†, Mufan Bill Li†, Bobby He, Thomas Hofmann, Chris Maddison, Daniel M. Roy. "The Shaped Transformer: Attention Models in the Infinite Depth-and-Width Limit" In *Advances in Neural Information Processing Systems (NeurIPS), 2023*
8. Gül Sena Altıntaş, Gregor Bachmann, **Lorenzo Noci**, Thomas Hofmann "Disentangling Linear Mode-Connectivity" In *UniReps Workshop at NeurIPS, 2023*
9. Sanghwan Kim†, **Lorenzo Noci**, Antonio Orvieto, and Thomas Hofmann. "Achieving a Better Stability-Plasticity Trade-off via Auxiliary Networks in Continual Learning." In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*
10. Sotiris Anagnostidis†, Gregor Bachmann†, **Lorenzo Noci**†, and Thomas Hofmann. "The curious case of benign memorization." In *International Conference on Learning Representations (ICLR), 2023.*

11. **Lorenzo Noci**[†], Sotiris Anagnostidis[†], Luca Biggio[†], Antonio Orvieto[†], Sidak Pal Singh[†], and Aurelien Lucchi. "Signal propagation in transformers: Theoretical perspectives and the role of rank collapse." In *Advances in Neural Information Processing Systems (NeurIPS), 2022*

12. **(Oral)** Gregor Bachmann[†], **Lorenzo Noci**[†], and Thomas Hofmann. "How Tempering Fixes Data Augmentation in Bayesian Neural Networks." In *International Conference on Machine Learning, pages 1244–1260. PMLR, 2022*

13. **(Spotlight) Lorenzo Noci**[†], Gregor Bachmann[†], Kevin Roth[†], Sebastian Nowozin, and Thomas Hofmann. "Precise Characterization of the Prior Predictive Distribution of deep relu networks." In *Advances in Neural Information Processing Systems, 34, 2021*

14. **Lorenzo Noci**[†], Kevin Roth[†], Gregor Bachmann[†], Sebastian Nowozin, and Thomas Hofmann. "Disentangling the roles of curation, data-augmentation and the prior in the cold posterior effect." In *Advances in Neural Information Processing Systems (NeurIPS), 34, 2021.*

15. Mario Arduini[†], **Lorenzo Noci**[†], Federico Pirovano[†], Ce Zhang, Yash Raj Shrestha, and Bibek Paudel. "Adversarial learning for debiasing knowledge graph embeddings." In *Workshop on Mining and Learning with Graphs at KDD.*