Nicolas Zucchet

PhD Student · ETH Zürich

nzucchet@ethz.ch | nicolaszucchet.github.io | NicolasZucchet | NicolasZucchet | Google Scholar

Education

PhD student *Zürich, Switzerland*

ETH Zürich *09/2021 - now*

- Doctoral student in artifical intelligence and computational neuroscience in the lab of Prof. Dr. Angelika Steger, jointly supervised by Dr. João Sacramento.

Computer science MSc *Zürich, Switzerland*

ETH Zürich *09/2019 - 04/2021*

- Theoretical foundations of artificial intelligence (optimization, neuroscience, machine learning).
- Master thesis: "Equilibrium propagation for bilevel optimization", supervised by Dr. João Sacramento.

Ingénieur polytechnicien (MSc) *Palaiseau, France*

École polytechnique *09/2016 - 08/2019*

- Multidisciplinary education (applied mathematics, computer science, economics), specialization in deep learning and computer vision during the last two semesters.

Classe préparatoire aux grandes écoles (MPI*) *Versailles, France*

Lycée Hoche *09/2014 - 08/2016*

- Intensive trainining for competitive entrance exams to french Grandes Écoles (fundamental mathematics, computer science, physics).

Professional experience

Student researcher *London, United Kingdom*

Google DeepMind *09/2024-02/2025*

- Research internship in the team of Dr. Soham De, studying the learning dynamics of language models.

Visiting student *Bern, Switzerland*

University of Bern *05/2021 - 08/2021*

- Probabilistic methods for continual learning, hosted by Prof. Dr. Jean-Pascal Pfister.

Research internship *Paris, France*

Prophesee *05/2019 - 08/2019*

- Deep learning algorithms for event-based cameras in autonomous cars.

Software developer intern *Sophia-Antipolis, France*

Amadeus *06/2018 - 08/2018*

- Design and development of continuous integration tools.

Publications

Preprints

[3] How do language models learn facts? Dynamics, curricula and hallucinations, **N. Zucchet**, J. Bornschein, S. Chan, A. Lampinen, R. Pascanu, S. De

[2] Uncovering mesa-optimization algorithms in Transformers, J. von Oswald*, E. Niklasson*, M. Schlegel*, S. Kobayashi, **N. Zucchet**, N. Scherrer, N. Miller, M. Sandler, B. Agüera y Arcas, M. Vladymyrov, R. Pascanu and J. Sacramento.

[1] Gated RNNs discover attention, **N. Zucchet***, S. Kobayashi*, Y. Akram*, J. von Oswald, M. Larcher, A. Steger†, J. Sacramento†.

Conference and journal papers

[6] Recurrent neural networks: vanishing and exploding gradients are not the end of the story, **N. Zucchet**, A. Orvieto. 38th Conference on Neural Information Processing Systems (NeurIPS), 2024.

[5] Online learning of long-range dependencies. **N. Zucchet***, R. Meier*, S. Schug*, A. Mujika and J. Sacramento. 37th Conference on Neural Information Processing Systems (NeurIPS), 2023.

[4] The least-control principle for local learning at equilibrium. A. Meulemans*, **N. Zucchet***, S. Kobayashi*, J. von Oswald and J. Sacramento. 36th Conference on Neural Information Processing Systems (NeurIPS), 2022.

[3] A contrastive rule for meta-learning. **N. Zucchet***, S. Schug*, J. von Oswald*, D. Zhao and J. Sacramento. 36th Conference on Neural Information Processing Systems (NeurIPS), 2022.

[2] Beyond backpropagation: Bilevel optimization through implicit differentiation and equilibrium propagation. **N. Zucchet** and J. Sacramento. Neural Computation 34 (12), 2022.

[1] Learning where to learn: Gradient sparsity in meta and continual learning. J. von Oswald*, D. Zhao*, S. Kobayashi, S. Schug, M. Caccia, **N. Zucchet** and J. Sacramento. 35th Conference on Neural Information Processing Systems (NeurIPS), 2021.

Workshop papers

[1] Random initialisations performing above chance and how to find them. F. Benzing, S. Schug, R. Meier, J. von Oswald, Y. Akram, **N. Zucchet**, L. Aitchison†, A. Steger†. OPT2022: 14th Annual Workshop on Optimization for Machine Learning (NeurIPS), 2022.

Talks

Dec. 2024 **Learning in the brain: what can we learn from deep state-space models?**, Linderman's lab, Stanford.
Jan. 2024 **Online learning of long-range dependencies**, Swiss Computational Neuroscience Retreat.
Feb. 2023 **The least-control principle for local learning at equilibrium**, Swiss Computational Neuroscience Retreat.
Nov. 2022 **The least-control principle for local learning at equilibrium**, Jean-Rémi King's Brain & AI group at Meta AI.
July 2022 **Biologically plausible bilevel optimization**, Rafal Bogacz's group at Oxford.
June 2022 **Bilevel optimization in neural networks**, MLSS^N 2022 lecture with J. Sacramento, A. Meulemans and S. Schug.
Feb. 2022 **A contrastive rule for meta-learning**, Swiss Computational Neuroscience Retreat.

Teaching experience

2022-24 **Algorithms lab**, Teaching assistant (Master's course).
2022-24 **Algorithms and probability**, Teaching assistant (Bachelor's course).
2021 **Randomized algorithms**, Teaching assistant (Master's course).
2016-17 **High-school mathematics**, Full-time tutoring of high school students from deprived neighborhoods.

Mentoring

2023 **Kevin Lopez**, Master thesis, ETH Zürich (co-supervised with J. Sacramento).
2023-24 **Yanick Schimpf**, Bachelor thesis, ETH Zürich.
2022-23 **Qianqian Feng**, Research project, ETH Zürich (co-supervised with J. Sacramento).
2022 **Anja Surina**, Master thesis, ETH Zürich (co-supervised with J. Sacramento and S. Schug).

Community service

2023-25 **ICML**, Reviewer.
2023-25 **ICLR**, Reviewer.
2023-24 **NeurIPS**, Reviewer (Best reviewer award in 2024).
2025 **TMLR**, Reviewer.
2025 **COLM**, Reviewer.

Awards

2022 **Scholar award**, NeurIPS.