Alessandro Favero

Contact Information

EPFL
Institute of Physics
Institute of Electrical Engineering
CH-1015 Lausanne, Switzerland

Phone: +41-21-693-98-00
E-mail: alessandro.favero@epfl.ch
Website: alesfav.github.io
LinkedIn: linkedin.com/in/alesfav/

Research Interests

Theory & science of deep learning: generalization, task and data structure, compositionality, geometric priors, scaling laws, statistical physics of learning.
Foundation models: diffusion models, diffusion LLMs, vision-language models, multimodal LLMs, post-training, model editing, model merging.

Education

EPFL, Lausanne, Switzerland

Ph.D., Physics and AI — 2025
Advisors: Prof. Matthieu Wyart, Prof. Pascal Frossard.

Sorbonne Université, Paris, France

M.S., Fundamental Physics, Specialization in Complex Systems — 2020
Mention très bien (highest honors).

SISSA, ICTP, Politecnico di Torino, Trieste-Torino, Italy

M.S., Physics of Complex Systems, Mathematical Modeling for Engineering — 2020
110/110 cum laude (highest honors).

Politecnico di Torino, Torino, Italy

B.S., Engineering Physics, Information Engineering — 2018

Industry Experience

Amazon Web Services Artificial Intelligence (AWS AI), Santa Clara, California

Applied Scientist — July to October 2023

- Internship at AWS AI Labs working on the alignment and robustness of multimodal LLMs with the fundamental research team led by Prof. Stefano Soatto.

Academic Experience

EPFL, Lausanne, Switzerland

Predoctoral Research Scholar — November 2020 to April 2021

- *Master's valorization* research scholarship (18,000 USD) on the statistical physics of deep learning systems in the Institute of Physics.

Visiting Master's Thesis Student — April 2020 to October 2020

- Thesis project "*Spectral analysis of infinitely-wide convolutional neural networks*" in the Physics of Complex Systems Laboratory led by Prof. Matthieu Wyart.

INRiM – Italian National Metrology Research Institute, Torino, Italy

Undergraduate Research Intern — October 2017 to January 2018

- Internship on space-time quantum correlations in the Quantum Optics Laboratory led by Prof. Marco Genovese.

Refereed Publications

See also my Google Scholar and Semantic Scholar profiles.

* *denotes co-first authorship.*

[1] Favero*, A., Sclocchi*, A., Cagnetta, F., Frossard, P. and Wyart, M., 2025. How compositional generalization and creativity improve as diffusion models are trained. To appear in Proceedings of the 42nd International Conference on Machine Learning (ICML), PMLR 267.

Workshop version presented at the ICLR 2025 Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy.

[2] Sclocchi*, A., Favero*, A., Levi*, N. I. and Wyart, M., 2025. Probing the Latent Hierarchical Structure of Data via Diffusion Models. The Thirteenth International Conference on Learning Representations (ICLR).
Workshop version presented at the NeurIPS 2024 Workshop on Scientific Methods for Understanding Deep Learning. **Oral**.
To appear in the 2025 special issue on the Statistical Physics aspects of Machine Learning, Journal of Statistical Mechanics: Theory and Experiment, 2025.

[3] Wang, K., Dimitriadis, N., Favero, A., Ortiz-Jimenez, G., Fleuret, F. and Frossard, P., 2025. LiNeS: Post-training Layer Scaling Prevents Forgetting and Enhances Model Merging. The Thirteenth International Conference on Learning Representations (ICLR).

[4] Sclocchi, A., Favero, A. and Wyart, M., 2025. A Phase Transition in Diffusion Models Reveals the Hierarchical Nature of Data. In Proceedings of the National Academy of Sciences (PNAS), 122 (1), e2408799121.

[5] Hazimeh*, A., Favero*, A. and Frossard, P., 2024. Task Addition and Weight Disentanglement in Closed-Vocabulary Models. In ICML 2024 Efficient Systems for Foundation Models Workshop.

[6] Cagnetta, F., Petrini, L., Tomasini, U.M., Favero, A. and Wyart, M., 2024. How Deep Neural Networks Learn Compositional Data: The Random Hierarchy Model. In Physical Review X, 14(3), p.031001.

[7] Favero, A., Zancato, L., Trager, M., Choudhary, S., Perera, P., Achille, A., Swaminathan, A. and Soatto, S., 2024. Multi-Modal Hallucination Control by Visual Information Grounding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.14303-14312.
Also presented at MMFM2: The 2nd Workshop on What is Next in Multimodal Foundation Models?, Seattle, WA, 2024.

[8] Ortiz-Jimenez*, G., Favero*, A. and Frossard, P., 2023. Task Arithmetic in the Tangent Space: Improved Editing of Pre-Trained Models. In Advances in Neural Information Processing Systems (NeurIPS), 36, pp.66727-66754.
Oral (top 0.54%).

[9] Barak, B., Carrell, A., Favero, A., Li, W., Stephan, L. and Zlokapa, A., 2024. Computational complexity of deep learning: Fundamental limitations and empirical phenomena. In Journal of Statistical Mechanics: Theory and Experiment, 2024(10), p.104008.

[10] Cagnetta*, F., Favero*, A. and Wyart, M., 2023. What Can Be Learnt With Wide Convolutional Neural Networks?. In Proceedings of the 40th International Conference on Machine Learning (ICML), PMLR 202, pp.3347-3379.
Included in the 2024 special issue on the Statistical Physics aspects of Machine Learning, Journal of Statistical Mechanics: Theory and Experiment, 2024(10), p.104020.

[11] Favero*, A., Cagnetta*, F. and Wyart, M., 2021. Locality defeats the curse of dimensionality in convolutional teacher-student scenarios. In Advances in Neural Information Processing Systems (NeurIPS), 34, pp.9456-9467.
Included in the 2022 special issue on the Statistical Physics aspects of Machine Learning, Journal of Statistical Mechanics: Theory and Experiment, 2022(11), p.114012.

[12] Petrini, L., Favero, A., Geiger, M. and Wyart, M., 2021. Relative stability toward diffeomorphisms indicates performance in deep nets. In Advances in Neural Information Processing Systems (NeurIPS), 34, pp.8727-8739.
Included in the 2022 special issue on the Statistical Physics aspects of Machine Learning, Journal of Statistical Mechanics: Theory and Experiment, 2022(11), p.114013.

Pre-prints

[13] Wang, K., Qin, Y., Dimitriadis, N., Favero, A. and Frossard, P., 2025. MEMOIR: Lifelong Model Editing with Minimal Overwrite and Informed Retention for LLMs. arXiv preprint arXiv: 2506.07899.

[14] Favero*, A., Sclocchi*, A. and Wyart, M., 2025. Bigger Isn't Always Memorizing: Early Stopping Overparameterized Diffusion Models. arXiv preprint arXiv:2505.16959.

[15] Cagnetta, F., Favero, A., Sclocchi, A. and Wyart, M., 2025. Scaling laws and representation learning in simple hierarchical languages: Transformers vs. convolutional architectures. arXiv preprint arXiv:2505.07070.

Conference Abstracts

[16] Favero, A., Sclocchi, A., Cagnetta, F., Frossard, P. and Wyart, M., 2025. Compositional Generalization and Creativity in Language Diffusion Models. ACL 2025 Workshop on Structure-aware Large Language Models.

[17] Favero, A., Cagnetta, F. and Wyart, M., 2023. Statistical Mechanics of Infinitely-Wide Convolutional Networks. Bulletin of the American Physical Society.

[18] Petrini, L., Favero, A., Geiger, M. and Wyart, M., 2023. Diffeomorphisms invariance is a proxy of performance in deep neural networks. Bulletin of the American Physical Society.

Selected Talks

Perimeter Institute, Theory + AI: Theoretical Physics for AI, Waterloo, 2025. *Creativity by compositionality in generative diffusion models.*

Johns Hopkins University Department of Physics & Astronomy, Baltimore, 2025. *Creativity by compositionality in generative diffusion models.*

IBM Research, IBM Accelerated Discovery Seminar, Zurich, 2024. *Task arithmetic in the tangent space of pre-trained models.*

EPFL Center for Intelligent Systems, Lausanne, 2023. *Task arithmetic in the tangent space: Improved editing of pre-trained models.*

37th Conference on Neural Information Processing Systems, New Orleans, 2023. *Task arithmetic in the tangent space: Improved editing of pre-trained models.*

Amazon AI Labs, 2023. *Task arithmetic in the tangent space of pre-trained models.*

MIT Center for Biological and Computational Learning, Boston, 2023. *Deep convolutional networks in kernel regimes: invariances, locality, and compositionality.*

NYU Center for Data Science, New York, 2023. *Generalization properties of deep convolutional networks in kernel regimes.*

American Physical Society March Meeting, Statistical Physics Meets Machine Learning, Las Vegas, 2023. *Statistical mechanics of infinitely wide convolutional networks.*

EPFL Institute of Physics, Seminars in Physics of Bio/Complex Systems, Lausanne, 2023. *Symmetry, locality, and hierarchy in artificial neural networks.*

Rice University, Workshop on the Theory of Overparameterized ML, 2022. *Locality defeats the curse of dimensionality in convolutional teacher-student scenarios.*

Selected Posters

Flatiron Institute, Center for Computational Neuroscience, New York, 2024. *Hierarchies and compositionality in diffusion models.*

Oxford Department of Statistics, Workshop on Robustness in LLMs, Oxford, 2024. *Multi-modal hallucination control by visual information grounding.*

Princeton University ORFE Department, Princeton, 2022. *How wide convolutional neural networks learn hierarchical tasks.*

Simons Foundation, Simons Collaboration on Cracking the Glass Problem Meeting, New York, 2022. *Spatial locality and translational invariance in machine learning.*

Meetings and Schools

- Mathematics of machine learning, Italian National Institute for Advanced Mathematics (INdAM – Istituto Nazionale di Alta Matematica), Cortona, 2024 (*invited*).
- Analytical connectionism summer school, Flatiron Center for Computational Neuroscience, New York, 2024.
- Machine learning theory summer school, Princeton University, Princeton, 2022.
- Statistical physics and machine learning summer school, Les Houches School of Physics, Les Houches, 2022.

Teaching Experience

EPFL, Lausanne, Switzerland

Teaching assistant (*2024 Dean's award for teaching excellence*) Fall 2021 to present

- PHYS-316 Statistical Physics II: Phase Transitions and Critical Phenomena (Spring 2023, Spring 2024).
- PHYS-467 Machine Learning for Physicists (Fall 2021, Fall 2022, Fall 2023).
- PHYS-421 Physics Projects I: Statistical Mechanics of Deep Learning (Fall 2021).

Guest lecturer at CS-625 Transfer Learning and Meta-Learning (Spring 2024).

Advising and Mentoring

Master's theses

- C. A. B., 2024, M.S. Cyber Security, EPFL–ETH Zurich.
- T. H., 2023, M.S. Physics, EPFL.

Semester projects (Ph.D.)

- A. A., 2024, Ph.D. Computer Science, EPFL.
- A. H., 2023, Ph.D. Computer Science, EPFL.

Referee

- Advances in Neural Information Processing Systems (NeurIPS). *2024 Top Reviewer.*
- International Conference on Learning Representations (ICLR). *2025 Notable Reviewer.*
- International Conference on Machine Learning (ICML).
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
- Transactions on Machine Learning Research (TMLR).
- Physical Review Journals.

Academic Service

- ELLIS (European Lab for Learning & Intelligent Systems) PhD Recruiting Committee, Evaluator (a.y. 2024-25)

Awards

- **Notable reviewer**, ICLR, 2025.
- **Dean's award for teaching excellence** (1,100 USD), EPFL, 2024.
- **Top reviewer award**, NeurIPS, 2024.
- Master's valorization research scholarship (18,000 USD), EPFL, 2020.
- Merit-based scholarship for thesis abroad, Politecnico di Torino (2,800 USD), 2020.
- Erasmus+ scholarship (3,400 USD), Sorbonne Université, 2019.
- Fee reduction for high academic performance, Politecnico di Torino, 2019.
- Physics of complex systems international track fellowship, Politecnico di Torino, SISSA, ICTP (2,000 USD), 2018.
- Top 200 engineering admission tests (among 8,000 applicants), Politecnico di Torino, 2014.