

ALESSANDRO FAVERO

RESEARCH INTERESTS	Physics-AI Fellow E-mail: af940@cam.ac.uk Website: alesfav.github.io LinkedIn: linkedin.com/in/alesfav X: x.com/alesfav	Centre for Mathematical Sciences University of Cambridge Cambridge, UK Infosys-Cambridge AI Centre London, UK
CURRENT AND PAST POSITIONS	Physics of learning/AI: generalization, data structure, compositionality, geometric priors, scaling laws, probabilistic graphical models, formal grammars, statistical mechanics. Foundation models: diffusion models, world models, vision-language models, multimodal models, post-training, model editing, model merging, AI for scientific discovery.	
	University of Cambridge , Cambridge, UK Research Associate	October 2025 to Present
	Inaugural <i>Physics-AI Fellow (independent postdoc)</i> in the Department of Applied Mathematics and Theoretical Physics (DAMTP). Member of the <i>Infosys-Cambridge AI Centre</i> , aimed at driving scientific discoveries through automation and understanding machine learning using physics methods.	
	Amazon Web Services Artificial Intelligence (AWS AI) , Santa Clara, California Applied Scientist	July to October 2023
	Internship at AWS AI Labs on vision-language alignment of multimodal LLMs with the fundamental research team led by Prof. Stefano Soatto.	
	EPFL , Lausanne, Switzerland Predoctoral Research Scholar	November 2020 to April 2021
	<i>Master's valorization</i> research scholarship on the statistical physics of AI systems in the Institute of Physics.	
	INRIM – Italian National Metrology Research Institute , Torino, Italy Research Intern	October 2017 to January 2018
	Undergraduate internship on space-time quantum correlations in the Quantum Optics Laboratory led by Prof. Marco Genovese.	
EDUCATION	EPFL , Lausanne, Switzerland Ph.D., Physics, AI	2025
	Advisors: Prof. Matthieu Wyart, Prof. Pascal Frossard. Dissertation: "The physics of data and tasks: Theories of locality and compositionality in deep learning". G-Research EPFL PhD award in maths and data science (first prize).	
	Sorbonne Université , Paris, France M.S., Fundamental Physics, Specialization in Complex Systems	2020
	<i>Mention très bien</i> (highest honors).	
	SISSA, ICTP, Politecnico di Torino , Trieste-Torino, Italy M.S., Physics of Complex Systems, Mathematical Modeling	2020
	<i>110/110 cum laude</i> (highest honors). International Honors Track (competitive admission, 20 students per cohort). Thesis at EPFL: "Spectral analysis of infinitely-wide convolutional neural networks".	
	Politecnico di Torino , Torino, Italy B.S., Engineering Physics, Information Engineering	2018
	Thesis: "Topological quantum computation with non-abelian anyons".	

AWARDS

- G-Research EPFL PhD prize in maths and data science (5,000 CHF), 2025.
- Recipient of *Physics-AI Fellowship*, DAMTP & Infosys-Cambridge AI Centre, 2025.
- Notable reviewer, ICLR, 2025.
- Dean's award for teaching excellence (1,000 CHF), EPFL, 2024.
- Top reviewer award, NeurIPS, 2024.
- Six-months *Master's valorization* research scholarship on statistical mechanics of deep learning (15,000 CHF), EPFL Institute of Physics, 2020.
- Merit-based scholarship for thesis abroad, Politecnico di Torino (2,500 EUR), 2020.
- Erasmus+ scholarship (3,000 EUR), 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 (1,800 EUR), 2018.
- Top 200 engineering admission tests (8,000 applicants), Politecnico di Torino, 2014.

**REFEREED
PUBLICATIONS**

Automatically-updated list on [Google Scholar](#).

Bibliographic metrics: 500+ citations, h-index 9 (Google Scholar, as of Oct 2025).

* denotes co-first authorship.

- [1] Cagnetta, F., Favero, A., Sclocchi, A. and Wyart, M., 2025. Scaling laws and representation learning in simple hierarchical languages: Transformers versus convolutional architectures. *Physical Review E (PRE)*.
- [2] Wang, K., Qin, Y., Dimitriadis, N., Favero, A. and Frossard, P., 2025. MEMOIR: Lifelong Model Editing with Minimal Overwrite and Informed Retention for LLMs. To appear in *Advances in Neural Information Processing Systems (NeurIPS)*, 38.
- [3] Favero*, A., Sclocchi*, A., Cagnetta, F., Frossard, P. and Wyart, M., 2025. How compositional generalization and creativity improve as diffusion models are trained. *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, PMLR 267, pp.16286-16306.
Workshop version presented at the ICLR 2025 Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy.
- [4] Sclocchi*, A., Favero*, A., Levi*, N. I. and Wyart, M., 2025. Probing the Latent Hierarchical Structure of Data via Diffusion Models. *International Conference on Learning Representations (ICLR)*.
Workshop version presented at the NeurIPS 2024 Workshop on Scientific Methods for Understanding Deep Learning. **Oral**.
Included in the 2025 special issue on the Statistical Physics aspects of Machine Learning, *Journal of Statistical Mechanics: Theory and Experiment*, 2025(8), p.084005.
- [5] 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. *International Conference on Learning Representations (ICLR)*.
- [6] Sclocchi, A., Favero, A. and Wyart, M., 2025. A Phase Transition in Diffusion Models Reveals the Hierarchical Nature of Data. *Proceedings of the National Academy of Sciences (PNAS)*, 122 (1), e2408799121.
- [7] Hazimeh*, A., Favero*, A. and Frossard, P., 2024. Task Addition and Weight Disentanglement in Closed-Vocabulary Models. *ICML 2024 Efficient Systems for Foundation Models Workshop*.
- [8] Cagnetta, F., Petrini, L., Tomasini, U.M., Favero, A. and Wyart, M., 2024. How Deep Neural Networks Learn Compositional Data: The Random Hierarchy Model. *Physical Review X*, 14(3), p.031001.
- [9] 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. *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.

- [10] Ortiz-Jimenez*, G., Favero*, A. and Frossard, P., 2023. Task Arithmetic in the Tangent Space: Improved Editing of Pre-Trained Models. Advances in Neural Information Processing Systems (NeurIPS), 36, pp.66727-66754.
Oral (top 0.54% of 12,000+ submissions).
- [11] Barak, B., Carrell, A., Favero, A., Li, W., Stephan, L. and Zlokapa, A., 2024. Computational complexity of deep learning: Fundamental limitations and empirical phenomena. Journal of Statistical Mechanics: Theory and Experiment, 2024(10), p.104008.
- [12] Cagnetta*, F., Favero*, A. and Wyart, M., 2023. What Can Be Learnt With Wide Convolutional Neural Networks?. 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.
- [13] Favero*, A., Cagnetta*, F. and Wyart, M., 2021. Locality defeats the curse of dimensionality in convolutional teacher-student scenarios. 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.
- [14] Petrini, L., Favero, A., Geiger, M. and Wyart, M., 2021. Relative stability toward diffeomorphisms indicates performance in deep nets. 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

- [15] Favero*, A., Sclocchi*, A. and Wyart, M., 2025. Bigger Isn't Always Memorizing: Early Stopping Overparameterized Diffusion Models. arXiv preprint arXiv: 2505.16959.
Workshop version presented at the ICML 2025 Workshop on the Impact of Memorization on Trustworthy Foundation Models.
- [16] Abdelraheem*, A., Favero*, A., Bovet, G. and Frossard P., 2025. Backdoor Unlearning Through Linear Task Decomposition in Multimodal Models. arXiv preprint arXiv:2510.14845.
Workshop version presented at the ICML 2025 Workshop on Machine Unlearning for Generative AI.

CONFERENCE ABSTRACTS

- [17] 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.
- [18] Favero, A., Cagnetta, F. and Wyart, M., 2023. Statistical Mechanics of Infinitely-Wide Convolutional Networks. Bulletin of the American Physical Society.
- [19] 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

- **Bernoulli Center for Fundamental Studies**, Fundamental principles of Learning and Representation Workshop: from Brains to LLMs, Lausanne, 2026.
- **University of Cambridge**, DAMTP Seminar, Cambridge, 2025. *Creativity in diffusion models: Insights from statistical physics and compositional grammars.*
- **Perimeter Institute**, Theory + AI: Theoretical Physics for AI, Waterloo, 2025. *Creativity by compositionality in generative diffusion models.* [\[Recording\]](#)

- **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*.
- **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.
- Youth in high-dimensions: machine learning, high-dimensional statistics, and inference for the new generation, ICTP, 2021.
- The hitchhiker’s guide: machine learning for condensed matter (4 days), 2021, ICTP.
- Spring college in physics of complex systems (1 week, interrupted due to the pandemic), 2020, ICTP, Trieste.

TEACHING

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
	<ul style="list-style-type: none"> • L. B., 2025, M.S. Physics, EPFL. • C. A. B., 2024, M.S. Cyber Security, EPFL–ETH Zurich. • T. H., 2023, M.S. Physics, EPFL.
	Semester projects (Ph.D.)
	<ul style="list-style-type: none"> • A. A., 2024, Ph.D. Computer Science, EPFL. • A. H., 2023, Ph.D. Computer Science, EPFL.
REVIEWER	Reviewer/referee for 20+ papers per year in machine learning and physics.
	<ul style="list-style-type: none"> • Advances in Neural Information Processing Systems (NeurIPS). <i>2024 Top Reviewer</i>. • International Conference on Learning Representations (ICLR). <i>2025 Notable Reviewer</i>. • 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	<ul style="list-style-type: none"> • ELLIS (European Lab for Learning & Intelligent Systems) PhD Recruiting Committee, Evaluator (a.y. 2024-25).
MEDIA & PRESS	<ul style="list-style-type: none"> • EPFL News, “A hidden hierarchy in AI image generation” Mar 2025. (Regarding PNAS 2025 publication). [Article]
SOFTWARE	<ul style="list-style-type: none"> • Programming. Python, C, C++, UNIX shell scripting. • Scientific and ML Libraries. NumPy, Matplotlib, scikit-learn, PyTorch, JAX, HF Transformers. • HPC. SLURM, Docker, K8s, Amazon Elastic Compute Cloud (EC2).
CERTIFICATES	IELTS Academic 8.5/9 (CEFR C2).