

# MICHAEL E. SANDER

Personal webpage: <https://michaelsdr.github.io/>

## EDUCATION

---

- Ecole Normale Supérieure de Paris, France** *2020 – 2024 (expected)*  
Ph.D. candidate in Machine Learning  
**Dissertation:** "Differentiable Efficient Learning through Discretized Dynamics"  
Advisors: Prof. Gabriel Peyré and Dr. Mathieu Blondel
- Sorbonne Université** *2019 - 2020*  
M.S. in Mathematics: "Mathematics of the modeling"
- Ecole Normale Supérieure Paris-Saclay** *2019 - 2020*  
M.S. in Machine Learning: Mathematics, Vision and Learning (MVA)
- Ecole polytechnique** *2016 - 2020*  
B.S. in Mathematics: "Cursus Ingénieur Polytechnicien"  
Mathematics, Applied Mathematics, Computer Science

## RESEARCH EXPERIENCE

---

- Visiting student, Tokyo University and RIKEN AIP, Tokyo** *May - August 2023*  
In-Context-Learning in Transformers, with Prof. Taiji Suzuki
- Student Researcher, Google DeepMind, Paris** *September 2022 - March 2023*  
Differentiable programming with Dr. Mathieu Blondel
- M.S. internship. DMA, Ecole Normale Supérieure, Paris** *April 2020 - September 2020*  
Deep invertible neural networks, with Prof. Gabriel Peyré and Dr. Mathieu Blondel
- M.S. internship. Behnia Lab, Colombia University, New York** *April 2019 - August 2019*  
Processing properties of color pathways in the eye and statistics of natural visual scenes, with Prof. Rudy Behnia
- B.S. internship. SOPHiA Genetics, Geneva** *June 2018 - September 2018*  
Statistical modeling of DNA sequencing, with Dr. Christian Pozzorini

## PUBLICATIONS

---

- [8] Michael E. Sander, Raja Giryes, Taiji Suzuki, Mathieu Blondel, Gabriel Peyré. How do Transformers perform In-Context Autoregressive Learning? *ICML*, 2024.
- [7] Pierre Marion\*, Yu-Han Wu\*, Michael E. Sander, Gérard Biau. Implicit regularization of deep residual networks towards neural ODEs. *ICLR*, 2024. **Spotlight**
- [6] Michael E. Sander, Tom Sander, Maxime Sylvestre. Unveiling the secrets of paintings: deep neural networks trained on high-resolution multispectral images for accurate attribution and authentication. *QCAV*, 2023.
- [5] Michael E. Sander, Joan Puigcerver, Josip Djolonga, Gabriel Peyré, Mathieu Blondel. Fast, Differentiable and Sparse Top-k: A convex analysis perspective. *ICML*, 2023.
- [4] Michael E. Sander, Pierre Ablin, Gabriel Peyré. Do Residual Neural Networks discretize Neural Ordinary Differential Equations? *NeurIPS*, 2022.
- [3] Samy Jelassi, Michael E. Sander, Yuanzhi Li. Vision Transformers provably learn spatial structure. *NeurIPS*, 2022.

[2] Michael E. Sander, Pierre Ablin, Mathieu Blondel, Gabriel Peyré. Sinkformers: Transformers with Doubly Stochastic Attention. *AISTATS*, 2022.

[1] Michael E. Sander, Pierre Ablin, Mathieu Blondel, Gabriel Peyré. Momentum Residual Neural Networks. *ICML*, 2021.

## OPEN SOURCE PYTHON SOFTWARE

---

Summary on my GitHub page: <https://github.com/michaelsdr>

-momentumnet, sinkformers, resnets-nodes, sparse-soft-topk: lead developer

-jaxopt: contributor

## PROFICIENCY IN CODING LANGUAGES

---

Python, Pytorch, JAX

## TALKS

---

CMStatistics, London, 2021. *Deep Invertible Residual Neural Networks*

SODA seminar, Saclay, 2021. *Momentum Residual Neural Networks*

Curves and Surfaces, Arcachon, 2022. *Momentum Residual Neural Networks*

MIND seminar, Saclay, 2022. *Sinkformers: Transformers with Doubly Stochastic Attention*

ICSIDS, Florence, 2022. *Do Residual Neural Networks discretize Neural Ordinary Differential Equations?*

Google, Slope Team, Paris, 2023. *Transformers and Neural Ordinary Differential Equations*

Cornell University, Peter McMahon's group, Online, 2023. *Second order Ordinary Differential Equations for Physical Neural Networks*

Tokyo University, 2023 *Do Residual Neural Networks discretize Neural Ordinary Differential Equations?*

MIND Seminar, Saclay, 2023 *On the relationship between Residual Neural Networks and Neural Ordinary Differential Equations*

Center for data science, Ecole Normale Supérieure, Paris, 2023 *On the relationship between Residual Neural Networks and Neural Ordinary Differential Equations*

PGMODAYS, Saclay, 2023 *Smooth and Sparse top-k operator*

## COMMUNITY SERVICE

---

Reviewer at NeurIPS 2023 and AISTATS 2024

## TEACHING

---

**Teacher in Statistics and Probability theory, Ecole Normale Supérieure, Paris** *2020 - 2024*  
In full charge of a mathematical course for students at Ecole Normale Supérieure

**Teaching assistant in Machine Learning, Ecole Normale Supérieure, Paris** *2023-2024*  
Optimization for Machine Learning

**Mathematical examiner for B.S. students at Lycée Henri IV, Paris** *2017 - 2020*  
In charge of evaluating students on theoretical mathematical concepts during oral sessions

## LANGUAGES

---

**French** (mother tongue), **English** (fluent), **Hebrew** (fluent), **Chinese** (good speaking)