

Tom Sander

+33 6 59 19 58 02 / tom.sander@polytechnique.edu / <https://github.com/tomsander1998>

EDUCATION

École polytechnique (IP Paris) - Palaiseau 2023 – 2026

Ph.D. in Mathématiques for Privacy-Preserving Machine Learning

- Supervisor: Prof. Alain Durmus, CMAP

École Normale Supérieure Paris-Saclay (Paris-Saclay University) - Gif-sur-Yvette 2021 – 2022

MVA Master 2 (MS in Mathematics for Machine Learning, “Mathématiques, Vision, Apprentissage”)

- Relevant courses: *Computational Optimal Transport, Computer Vision, Computational Statistics, Convex Optimization, Reinforcement Learning, Statistical Learning, Advanced Deep Learning, Time series, Kernel methods*
- Grade: 18.3/20, Highest honors with congratulations from the jury/*Mention Très Bien avec felicitations du jury*

École Polytechnique – Palaiseau, France 2018 - 2022

France’s top university for science and engineering, Cycle Ingénieur Polytechnicien (X2018)

- Master level program in Applied Math and Computer Science, specializing in Machine Learning and Optimization
- Coursework includes *Machine Learning (1&2), Operations Research, Optimization, Stochastic Calculus, Big Data Processing with C++, Statistics, Algorithmics, Reinforcement Learning, Modeling Random Phenomena*
- GPA: Masters (3A): 3.96/4.0, Bachelor (prépa/1A/2A): 3.93/4.0

Lycée Henri-IV Preparatory Program - Paris, France 2016 - 2018

Program in advanced pure math, computer science and physics leading to nationwide entrance examinations to the Grandes Écoles for scientific studies (“MP option informatique”)*

- Coursework in Mathematics includes *Differential equations, Vector Calculus, Analysis, Linear and Abstract Algebra*
- Coursework in Physics includes *Newtonian Mechanics, Electrokinetic, Thermodynamics, Quantum Physics*
- Coursework in Computer Science includes *Data structures and algorithms, Programming in Ocaml and Python*

(SELECTED) RESEARCH WORK

Meta AI Research / CMAP Polytechnique - Paris, France March 2023 – March 2026

CIFRE Ph.D. in Privacy-preserving Machine Learning

- “Implicit Bias in Noisy-SGD: With Applications to Differentially Private training” (submitted to AISTATS2024)
- “Differentially Private Vision-Language Foundation Model Via Image Captioning”, (submitted to ICLR 2024)

Meta AI Research - Paris, France June 2022 – October 2022

Research Internship in Privacy-preserving Machine Learning, <https://github.com/facebookresearch/tan>

- “TAN without a Burn: Scaling laws of DP-SGD”, arxiv: <https://arxiv.org/abs/2210.03403>
- We introduced TAN, a simulation framework for large scale Differentially Private (DP) learning that helps reduce the computational cost of hyper parameter search with DP-SGD by a factor of x100, pushing private state of the art performance of private training on ImageNet by +10%.
- Subject of MVA masters’ thesis, grade: 19/20 (jury: Julien Mairal), **Accepted at ICML2023**

Polytechnique Montréal - Montréal, Canada March 2021 – August 2021

Research Internship in Operations Research and Reinforcement Learning (<https://github.com/corail-research/SeaPearl.jl>)

- Building a hybrid solver in Julia based on Constraint programming and Deep Reinforcement Learning (SeaPearl). Supervised by Quentin Cappart and Louis-Martin Rousseau from CORAIL (Combinatorial Optimization Reasoning in Artificial Intelligence Laboratory).
- ‘Best scientific internship award’ from Ecole Polytechnique (Paris), grade: 19/20 (jury: Eric Moulines)

LIMSI-CNRS - Massy, France September 2020 – February 2021

Research assistant in NLP, under the supervision of Prof. François Yvon

- “Unsupervised learning for multilingual embeddings: theoretical aspects and computer simulation”
- Using Optimal Transport between monolingual embeddings can be efficient when coupled with adequate optimization methods to obtain multilingual embeddings in an unsupervised manner. We analyzed the influence of the frequencies

initially given to the words on the result of Sinkhorn's algorithm. We used word frequencies of appearance in large corpuses, and the results show that it gives better accuracy for uncommon words. Git repos available

Carrefour Data-lab & Polytechnique - Massy, France

September 2019 – April 2020

One-year research project under the supervision of Prof. Eric Moulines (PSC)

- “Optimization of short-term statistical prevision models for sales using AI methods”
- One-year team research project under the guidance of Prof. Eric Moulines from Polytechnique's Applied Math Laboratory (CMAP) and with the Carrefour Data lab of Massy. We worked on the short-term forecast of Carrefour's sales (French supermarket chain). After a preprocessing of all their sales data and feature engineering, we delivered a pipeline that gave better accuracy than Carrefour's model. The difference was made by factoring in external data, such as sport events dates, protest movements and studying the theory of the tools in depth.

École polytechnique - Palaiseau, France

April 2020 - June 2020

Research project under the supervision of Prof. Emmanuel Gobet (Modal)

- “Simulation of rare events for a complete study of the Galton-Watson process”
- Events that occur with a probability lower than $10^{-4}/10^{-5}$ are called “rare”. Obtaining information about such events is computationally challenging, and classic Monte Carlo method is not sufficient to analyze them. I developed a pipeline that mainly uses selection/mutation algorithms and Importance Sampling for a complete study of the Galton Watson Process in the subcritical case (when the population almost surely disappears).

LEADERSHIP AND INTERNSHIP EXPERIENCE

BNP-PARIBAS, PACE - Paris, France

June 2020 – September 2020

Summer internship, Data science

- The team works on AI research and application for all marketing or prices related topics at BNP (French leading bank). I worked on a Private Financial Management project, using time series Machine Learning algorithms to predict user's transactions patterns in order to help them manage their spendings. It implied working with Big Data (millions of transactions per day), processed and trained in Python. Algorithms are currently running on the bank's mobile application

University of Science and Technology of Hanoi (USTH) - Hanoi, Vietnam

October 2018 – April 2019

Civil Volunteer Service, first-year internship part of Polytechnique's curriculum

- I assisted the University of Science and Technology of Hanoi (USTH), a French Vietnamese university, in its development. I worked directly with the vice-principal with the additional mission to give math classes. I also contributed to the international opening of the university and built a marketing strategy to attract students.

Lycée Henri-IV - Paris, France

September 2019 – Current

Oral examiner in math for undergraduates preparing for Grandes Écoles (colleur)

- In charge of evaluating undergraduate students on theoretical and mathematical concepts during weekly oral sessions

École polytechnique - Palaiseau, France

April 2019 – Septembre 2020

Vice-President of the entrepreneurship association (CSU)

- Organization of the “Tech Forum”, 10k budget, 1,000 participants, 30 startups

French Army - France

August 2018 – October 2018

- Intensive military training part of Polytechnique's curriculum

Additional Courses

Coursera, DeepLearning.AI – Online Courses

August 2021 – October 2021

- Deep Learning Specialization: *Neural Networks, Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization, Structuring Machine Learning Projects, Convolutional Neural Networks*
- Natural Language Processing Specialization: *Classification and Vector Spaces, Probabilistic Models, Sequence Models, Attention Models*

AWARDS

- *Best Research Internship (2021)*, Ecole Polytechnique

- *Outstanding Investment (OI) award (2021)*, Ecole Polytechnique, attributed to a student who contributed greatly to the community of the school by its extra-curricular involvement
- *First price in Mathematics of the school (2016)*, Ecole Alsacienne

SKILLS AND INTERESTS

Language Skills: Fluent English (Toefl: 110/120), Native French, Intermediate Chinese and Hebrew

Linguistic exchanges: Sydney (Australia), Beijing (China), San Francisco (US)

Computer Skills: Python (Pytorch, TensorFlow, Trax), R, Julia, C++, Java, Caml, Lisp

Interests: Entrepreneurship, Backgammon, Traveling, Cinema, Ski, Basketball