

Samy Jelassi

sjelassi@fas.harvard.edu | 609-933-5773 | sjelassi.github.io

Employment

Harvard University , Postdoctoral Fellow, School of Engineering and Applied Sciences (SEAS) Hosts: Boaz Barak and Sham Kakade Research topics: LLM architectures, Optimization, Long-context, RL with LLMs.	2025 - Present
Harvard University , Research fellow, Center of Mathematical Sciences and Applications	2023 - 2025

Education

Princeton University , PhD, Operations Research Department. Advised by Boris Hanin Thesis: Algorithmic and architectural implicit biases in deep learning	2017 – 2023
ENS Cachan , Master of Arts in Applied Mathematics with distinction. Advised by Francis Bach Thesis: Variance-Reduced Gradient Descent Methods	2015 – 2017
ENS Lyon , Bachelor in Computer Science with distinction.	2014 – 2015
Lycée Louis-le-Grand , Classes Préparatoires aux Grandes Écoles. University-level preparation for the competitive entrance to French Engineering Schools	2011 – 2014

Internships

Google Research (NYC) , hosted by Srinadh Bhojanapalli and Sashank Reddi	2022
Google Deepmind (London) , hosted by Bernardo Avila Pires and Rémi Munos	2021
Facebook AI Research (NYC) , hosted by Aaron Defazio	2020

Selected works

Matching Features, Not Tokens: Energy-Based Fine-Tuning of Language Models S. Jelassi* , M. Kwun*, R. Zhao*, Y. Li, N. Fusi, Y. Du, S. Kakade, C. Domingo-Enrich* submitted	2026
Let's (not) just put things in Context: Test-time Training for Long-context LLMs R. Bansal, A. Zhang, R. Tiwari, L. Madaan, S. Duvvuri, F. Devvrit, D. Brandfonbrener, D. Alvarez-Melis, P. Bhargava, M. Kale, S. Jelassi International Conference on Learning Representations (ICLR) 2026, https://arxiv.org/abs/2512.13898	2026
Echo chamber: RL post-training amplifies behaviors learned in pretraining R. Zhao*, A. Meterez*, S. Kakade, C. Pehlevan, S. Jelassi [†] , E. Malach [†] Conference on Language Modeling (COLM) 2025, https://arxiv.org/abs/2504.07912	2025
Mixture of Parrots: Experts improve memorization more than reasoning S. Jelassi , C. Mohri, D. Brandfonbrener, A. Gu, N. Vyas, N. Anand, D. Alvarez-Melis, Y. Li, S. Kakade, E. Malach International Conference on Learning Representations (ICLR) 2025, oral presentation (top 10%) at the "Mathematics of modern machine learning" workshop, NeurIPS 2024, https://arxiv.org/abs/2410.19034	2025
Repeat after me: Transformers are better than state space models at copying S. Jelassi , D. Brandfonbrener, S. Kakade, E. Malach	2024

International Conference on Machine Learning (ICML) 2024,
<https://arxiv.org/abs/2402.01032>

Conference papers

- Parameter-Efficient Reinforcement Learning using Prefix Optimization** 2026
S. Jelassi*, I. Rocha Filho*, R. Zhao, S. Kakade, E. Malach
International Conference on Learning Representations (ICLR) 2026,
<https://openreview.net/forum?id=SLhLUdlaqc>
- Let Me Think! A Long Chain-of-Thought Can Be Worth Exponentially Many Short Ones** 2025
P. Mirtaheri*, E. Edelman*, S. Jelassi, E. Malach, E. Boix-Adsera
Conference on Neural Information Processing Systems (NeurIPS) 2025,
<https://arxiv.org/abs/2505.21825>
- To backtrack or not to backtrack: When sequential search limits model reasoning** 2025
T. Qin, D. Alvarez-Melis, S. Jelassi*, E. Malach*
Conference on Language Modeling (COLM) 2025,
<https://arxiv.org/abs/2504.07052>
- Universal length generalization with turing programs** 2025
K. Hou, D. Brandfonbrener, S. Kakade, S. Jelassi*, E. Malach*
International Conference on Machine Learning (ICML) 2025,
<https://arxiv.org/abs/2407.03310>
- The Role of Sparsity for Length Generalization in Transformers** 2025
N. Golowich, S. Jelassi, D. Brandfonbrener, S. Kakade, E. Malach
International Conference on Machine Learning (ICML) 2025,
<https://arxiv.org/abs/2502.16792>
- LoRA Soups: Merging LoRAs for Practical Skill Composition Tasks** 2025
A. Prabhakar, Y. Li, K. Narasimhan, S. Kakade, E. Malach, S. Jelassi
International Conference on Computational Linguistics (COLING) 2025, Industry track.
<https://arxiv.org/abs/2410.13025>
- Q-Probe: A Light Approach to Reward Maximization for Language Models** 2024
K. Li, S. Jelassi, H. Zhang, S. Kakade, M. Wattenberg, D. Brandfonbrener
International Conference on Machine Learning (ICML) 2024,
<https://arxiv.org/abs/2402.14688>
- Vision transformers provably learn spatial structure** 2022
S. Jelassi, M. Sander, Y. Li
Conference on Neural Information Processing Systems (NeurIPS) 2022,
<https://arxiv.org/abs/2210.09221>
- Towards understanding how momentum improves generalization in deep learning** 2022
S. Jelassi, Y. Li
International Conference on Machine Learning (ICML) 2022,
<https://arxiv.org/abs/2207.05931>
Oral presentation (top 5%) at "Overparameterization: Pitfalls & Opportunities" workshop, ICML 2021.
- Auction learning as a two-player game** 2021
J. Rahme, S. Jelassi, S. M. Weinberg
International Conference on Learning Representations (ICLR) 2021,
<https://arxiv.org/abs/2006.05684>

A Permutation-Equivariant Neural Network Architecture For Auction Design	2021
J. Rahme, S. Jelassi , J. Bruna, S. M. Weinberg AAAI Conference on Artificial Intelligence 2021, https://arxiv.org/abs/2003.01497	
Extragradient with player sampling for faster Nash equilibrium finding	2020
S. Jelassi , C. Domingo-Enrich, D. Scieur, A. Mensch, J. Bruna International Conference on Machine Learning (ICML) 2020, https://arxiv.org/abs/1905.12363	
A mean-field analysis of two-player zero-sum games	2019
C. Domingo-Enrich, S. Jelassi , A. Mensch, G. M. Rotskoff, J. Bruna Conference on Neural Information Processing Systems (NeurIPS) 2019, https://arxiv.org/abs/2002.06277	
Towards closing the gap between the theory and practice of SVRG	2019
O. Sebbouh, N. Gazagnadou, S. Jelassi , F. Bach, R. M. Gower Conference on Neural Information Processing Systems (NeurIPS) 2019, https://arxiv.org/abs/1908.02725	
Global convergence of neuron birth-death dynamics	2019
G. Rotskoff, S. Jelassi , J. Bruna, E. Vanden-Eijnden International Conference on Machine Learning (ICML) 2019, https://arxiv.org/abs/1902.01843	
Smoothed analysis of low-rank approach for smooth semidefinite programs	2019
T. Pumir*, S. Jelassi *, N. Boumal Oral presentation (top 3%) at the Conference on Neural Information Processing Systems (NeurIPS) 2018, https://arxiv.org/abs/1806.03763	
Journal papers	
Adaptivity without Compromise: A Momentumized, Adaptive, Dual Averaged Gradient Method for Stochastic Optimization	2022
A. Defazio, S. Jelassi Journal of Machine Learning Research 2022, https://arxiv.org/abs/2101.11075	
Depth separation beyond radial functions	2022
L. Venturi, S. Jelassi , T. Ozuch, J. Bruna Journal of Machine Learning Research 2022, https://arxiv.org/abs/2102.01621	
Preprints	
Depth Dependence of μP Learning Rates in ReLU MLPs	2023
S. Jelassi , B. Hanin, Z. Ji, S. Reddi, S. Bhojanapalli, S. Kumar https://arxiv.org/abs/2305.07810	
Length generalization in arithmetic transformers	2023
S. Jelassi , S. d'Ascoli, C. Domingo-Enrich, Y. Wu, Y. Li, F. Charton https://arxiv.org/abs/2306.15400	

Teaching

COS 485 Neural Networks: Theory and Applications, Teaching Assistant, Spring 2023.

ORF 350: Analysis of Big Data, Head Teaching Assistant, Spring 2019, 2021, 2022.

ECE 435/535, Machine Learning and Pattern Recognition, Teaching Assistant, Fall 2018, 2019, 2021.

ORF 409: Introduction to Monte Carlo Simulation, Teaching Assistant, Fall 2020.

Service

Reviewer, NeurIPS 2019-24, ICML 2020 & 2023, ICLR 2025, STOC 2025, JMLR.

Organizer, New Technologies in Mathematics Seminar at Harvard CMSA, Fall 2023 & Spring 2024.

Talks

Understanding RL with Verifiable Rewards through Distribution Sharpening 2025

Foundations of Post-training workshop, COLT 2025.

Mixture of Parrots: Experts improve memorization more than reasoning 2024

Mathematics of Modern Machine Learning workshop, NeurIPS 2024.

Algorithmic and architectural implicit biases in deep learning 2022

EPFL, Caltech, University of Toronto

Towards understanding how momentum improves generalization in deep learning 2022

International Conference on Machine Learning (ICML) 2021, 2022

Smoothed analysis of some machine learning problems 2019

Google Montreal

Smoothed analysis of the low-rank approach for smooth semidefinite program 2018

Plenary oral presentation at the Conference on Neural Information Processing Systems (NeurIPS) 2018.