

Nikolaos Tsilivis

ntsilivis@ttic.edu, +1-347-721-1666
6045 S Kenwood Ave, Chicago, IL 60637

Research Interests	Foundations of Machine Learning, Robust Machine Learning, Applications of Machine Learning in Natural Sciences.	
Employment	Toyota Technological Institute at Chicago Research Assistant Professor	Chicago, IL, USA Sep 2026 -
	Fundamental AI Research (FAIR), Meta Visiting Researcher	NYC, NY, USA Sep 2024 - Sep 2025
Education	New York University Ph.D. in Data Science Ph.D. Advisor: Julia Kempe	NYC, NY, USA Sep 2021 - Aug 2026
	École Normale Supérieure de Paris Visiting Ph.D. student Hosts: Julia Kempe and Bruno Loureiro	Paris, France Apr - Jun 2024, Oct - Dec 2025
	Toyota Technological Institute at Chicago Visiting Ph.D. student Host: Nathan Srebro	Chicago, IL, USA Jan - Mar 2024
	Kempner Institute, Harvard University Visiting Ph.D. student Hosts: Boaz Barak and Cengiz Pehlevan	Cambridge, MA, USA Sep - Dec 2023
	National Technical University of Athens Diploma in Electrical and Computer Engineering (B.Sc. & M.Sc.) Thesis Advisor: Petros Maragos	Athens, Greece Sep 2014 - Jul 2021
	KTH Royal Institute of Technology Exchange Studies	Stockholm, Sweden Jan - Jun 2018
Publications	<i>Conferences</i> 20. How reinforcement learning after next-token prediction facilitates learning N. Tsilivis , E. Malach, K. Ullrich, J. Kempe ICLR 2026 EurIPS 2025 Workshop on Principles of Generative Modeling <i>★Contributed talk★</i> 19. OpenApps: Simulating Environment Variations to Measure UI Agent Reliability K. Ullrich, J. Su, C. Shi, A. Subramonian, A. Bar, I. Evtimov, N. Tsilivis , R. Balestrieri, J. Kempe, M. Ibrahim ICLR 2026 <i>★Oral★</i>	

18. Flavors of Margin: Implicit Bias of Steepest Descent in Homogeneous Neural Networks
N. Tsilivis, G. Vardi, J. Kempe
 ICLR 2025
17. The Evolution of Statistical Induction Heads: In-Context Learning Markov Chains
 E. Edelman*, **N. Tsilivis***, B. L. Edelman, E. Malach, S. Goel
 NeurIPS 2024
16. The Price of Implicit Bias in Adversarially Robust Generalization
N. Tsilivis, N. Frank, N. Srebro, J. Kempe
 NeurIPS 2024
15. What Can the Neural Tangent Kernel Tell Us About Adversarial Robustness?
N. Tsilivis, J. Kempe
 NeurIPS 2022
14. Sparsity in Max-Plus Algebra and Applications in Multivariate Convex Regression
N. Tsilivis, A. Tsiamis, P. Maragos
 ICASSP 2021
13. Sparse Approximate Solutions to Max-Plus Equations
N. Tsilivis, A. Tsiamis, P. Maragos
 Discrete Geometry and Mathematical Morphology 2021 *★Invited to the special issue★*

Journals

12. Flavors of Margin: Implicit Bias of Steepest Descent in Homogeneous Neural Networks (extended version)
N. Tsilivis, E. Gronich, J. Kempe, G. Vardi
 Journal of Machine Learning Research 27
11. On the Geometry of Regularization in Adversarial Training: High-Dimensional Asymptotics and Generalization Bounds
 M. Vilucchio, **N. Tsilivis**, B. Loureiro, J. Kempe
 TMLR 2026
10. Kernels, data & physics
 F. Cagnetta, D. Oliveira, M. Sabanayagam, **N. Tsilivis**, J. Kempe (α - β order)
 Journal of Statistical Mechanics: Theory and Experiment 2024
9. On the Robustness of Neural Collapse and the Neural Collapse of Robustness
 J. Su, Y. S. Zhang, **N. Tsilivis**, J. Kempe
 TMLR 2024
8. Attacking Bayes: On the Adversarial Robustness of Bayesian Neural Networks
 Y. Feng, T. Rudner, **N. Tsilivis**, J. Kempe
 TMLR 2023 *★Reproducibility certification★*
7. Toward a Sparsity Theory on Weighted Lattices
N. Tsilivis, A. Tsiamis, P. Maragos
 Journal of Mathematical Imaging and Vision 2022

Workshops

6. A Tale of Two Circuits: Grokking as Competition of Sparse and Dense Subnetworks
W. Merrill*, **N. Tsilivis***, A. Shukla
ICLR 2023 Workshop on Mathematical and Empirical Understanding of Foundation Models
5. Can we achieve robustness from data alone?
N. Tsilivis, J. Su, J. Kempe
ICML 2022 Workshop on New Frontiers in Adversarial Machine Learning
4. Adversarial Noise Injection for Learned Turbulence Simulations
J. Su, J. Kempe, D. Fielding, **N. Tsilivis**, M. Cranmer, S. Ho
NeurIPS 2022 Workshop on Machine Learning and the Physical Sciences

Preprints

3. Learning through Internalization
N. Tsilivis, N. Joshi, M. Medvedev, J. Kempe, N. Srebro, arXiv 2026
2. Learning to Think from Multiple Thinkers
N. Joshi, R. Magen, N. Srebro, **N. Tsilivis**, G. Vardi (α - β order), arXiv 2026
1. Extracting Finite Automata from RNNs Using State Merging
W. Merrill*, **N. Tsilivis***, arXiv 2022

Awards

Coefficient Giving Grant - 88\$k (2026-2027): Academic grant for project: “Theoretical Foundations of Robust Autoregressive Learning as a Means of Designing Safe and Robust Models”.

Meta AI Research Grant (2024-2025): Covers full tuition and stipend as a part of the Meta AI Mentorship Program for one academic year.

Gerondelis Foundation Grant (2024): Awarded to Greek students pursuing graduate studies in the United States.

Center for Data Science Fellowship (2021-2026): Covers tuition and living expenses for 5 years.

Thomaideio Award (Publications) (2021): Awarded to undergraduate students of the National Technical University of Athens who published a research paper before their graduation.

Teaching

Teaching Assistant for DS-GA 1005, NYU: *Inference and Representation* (2025)
Instructor: Joan Bruna

Teaching Assistant for DS-GA 2003, NYU: *Introduction to Data Science for Ph.D. Students* (2022)
Instructors: Kyunghyun Cho, Cristina Savin, Julia Kempe

Co-authored lecture notes for Julia Kempe’s lectures at the 2022 Les Houches Summer School on Statistical Physics and Machine Learning

Invited talks

ICTP, Youth in High-Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference (2026)
Title: *Learning through Internalization* ([video link](#))

EPFL, Lausanne Event on Machine Learning and Neural Network Theory (2026)

Title: *How reinforcement learning after next-token prediction facilitates learning*

Toyota Technological Institute at Chicago, Young Researcher Seminar Series (2026)

Title: *How reinforcement learning after next-token prediction facilitates learning*

Center for Computational Mathematics, Flatiron Institute (2026)

Title: *How reinforcement learning after next-token prediction facilitates learning*

EurIPS 2025, Workshop on Principles of Generative Modeling (2025)

Title: *How reinforcement learning after next-token prediction facilitates learning*

Stanford University, Simons Collaboration on the Physics of Learning and Neural Computation Kickoff Workshop (2025)

Title: *How reinforcement learning after next-token prediction facilitates learning*

Microsoft Research NYC (2025)

Title: *How reinforcement learning after next-token prediction facilitates learning*

École Normale Supérieure de Paris, CSD Seminar (2025)

Title: *How reinforcement learning after next-token prediction facilitates learning*

Toyota Technological Institute at Chicago, NSF TRIPODS Workshop (2024)

Title: *The Price of Implicit Bias in Robust ML*

Flatiron Institute, Center for Computational Mathematics (2024)

Title: *The Price of Implicit Bias in Robust ML*

University of California, Los Angeles, IPAM, Workshop II: Theory and Practice of Deep Learning (2024)

Title: *The Price of Implicit Bias in Robust ML* ([video link](#))

Harvard University, ML Foundations group (2023)

Title: *The best algorithm for adversarial training*

New York University, J. Bruna's group (2023)

Title: *Discontinuous Deep Learning, Grokking & More*

University of California, Irvine - GoalLab (2022)

Title: *Lazy Optimization Regimes in Deep Learning*

New York University, CDS PhD seminar (2022)

Title: *What Can The Neural Tangent Kernel Tell Us About Adversarial Robustness?*

Service

Mentoring & Community Service:

- CDS & NYU GSAS Ph.D. Mentoring Programs: Paired with incoming students to provide guidance and support (2024-2025)
- Supervision of undergraduate & master students:
 - Aman Shukla (CDS, NYU), 2022-2023: optimization dynamics in neural networks
 - Ya Shi Zhang (Courant, NYU), 2022-2023: adversarial robustness of neural networks
 - Sergey Sedov (CDS, NYU), 2024-2025: efficient methods for reasoning in language models
 - Iason Kalogiannis (ECE, NTUA), 2025-2026: mean-field analysis of neural networks with ranking layers
- Member of the Student Inclusion and Belonging Advisory Board at CDS, NYU (Spring 2025)

- NeurIPS 2024 Volunteer.

Reviewing: NeurIPS 2024 (**Top reviewer**), 2026, ICLR 2023, 2025 (**Notable reviewer**), AISTATS 2025, Physical Review Letters, TMLR, UniReps Workshop (NeurIPS 2023), Re-Align Workshop (ICLR 2024), HiLD Workshop (ICML 2024), SciFor4DL Workshop (NeurIPS 2024), M3L Workshop (NeurIPS 2024), MOSS Workshop (ICML 2025), PriGM Workshop (EurIPS 2025).
Program Committee: ALT 2027.

**Programming
Skills**

Languages: Python, C/C++, Java, SML, Prolog
Other: PyTorch, JAX

Languages

Greek (native), English (fluent)