Nikolaos Tsilivis

nt2231@nyu.edu, +1-347-721-1666
60 5th Ave, New York, NY 10011

Research Interests

Foundations of Machine Learning, Robust Machine Learning, Applications of Machine Learning in Physical Sciences.

Education

New York University — NYC, NY, USA
Ph.D. in Data Science, Center for Data Science — 2021-
Ph.D. Advisor: Julia Kempe

École Normale Supérieure de Paris — Paris, France
Visiting Ph.D. student at Center for Data Science — Apr - Jun 2024, Oct 2025 -
Hosts: Julia Kempe, Bruno Loureiro

Toyota Technological Institute at Chicago — Chicago, IL, USA
Visiting Ph.D. student — Jan - Mar 2024
Host: Nathan Srebro

Kempner Institute, Harvard University — Cambridge, MA, USA
Visiting Ph.D. student at Machine Learning Foundations group — Sep - Dec 2023
Hosts: Boaz Barak and Cengiz Pehlevan

National Technical University of Athens — Athens, Greece
Diploma in Electrical and Computer Engineering (B.Sc. & M.Sc.) — 2014-2021
Major: Computer Science, Signal Processing, Control Theory
Minor: Mathematics
Thesis: *Sparse Representations in Tropical Mathematics*
Advisor: Petros Maragos

KTH Royal Institute of Technology — Stockholm, Sweden
Exchange Studies — Jan - Jun 2018

Employment

Fundamental AI Research (FAIR), Meta — NYC, NY, USA
Visiting Researcher
Host: Karen Ullrich — Sep 2024 - Sep 2025

New York University — NYC, NY, USA
Visiting Researcher
Host: Julia Kempe — Mar 2021 - Jul 2021

CVSP Lab, National Technical University of Athens — Athens, Greece
Undergraduate Research Assistant
Host: Petros Maragos — Jan 2020 - Mar 2021

FS Unit, National Technical University of Athens — Athens, Greece
Undergraduate Research Assistant / Junior Developer — Nov 2015 - Jan 2016

Publications

Conferences

17. Flavors of Margin: Implicit Bias of Steepest Descent in Homogeneous Neural Networks
N. Tsilivis, G. Vardi, J. Kempe
ICLR 2025

16. The Evolution of Statistical Induction Heads: In-Context Learning Markov Chains
E. Edelman*, **N. Tsilivis***, B. L. Edelman, E. Malach, S. Goel
NeurIPS 2024

15. The Price of Implicit Bias in Adversarially Robust Generalization
N. Tsilivis, N. Frank, N. Srebro, J. Kempe
NeurIPS 2024

14. What Can the Neural Tangent Kernel Tell Us About Adversarial Robustness?
N. Tsilivis, J. Kempe
NeurIPS 2022

13. Sparsity in Max-Plus Algebra and Applications in Multivariate Convex Regression
N. Tsilivis, A. Tsiamis, P. Maragos
ICASSP 2021

12. 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

11. Kernels, data & physics
F. Cagnetta, D. Oliveira, M. Sabanayagam, **N. Tsilivis**, J. Kempe (α–β order)
Journal of Statistical Mechanics: Theory and Experiment 2024

10. On the Robustness of Neural Collapse and the Neural Collapse of Robustness
J. Su, Y. S. Zhang, **N. Tsilivis**, J. Kempe
TMLR 2024

9. Attacking Bayes: On the Adversarial Robustness of Bayesian Neural Networks
Y. Feng, T. Rudner, **N. Tsilivis**, J. Kempe
TMLR 2023, **Reproducibility Certification**

8. Toward a Sparsity Theory on Weighted Lattices
N. Tsilivis, A. Tsiamis, P. Maragos
Journal of Mathematical Imaging and Vision 2022

Workshops

7. 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

6. Can we achieve robustness from data alone?
N. Tsilivis, J. Su, J. Kempe
ICML 2022 Workshop on New Frontiers in Adversarial Machine Learning

5. 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

4. How reinforcement learning after next-token prediction facilitates learning
N. Tsilivis, E. Malach, K. Ullrich, J. Kempe, 2025

3. OpenApps: Simulating Environment Variations to Measure UI Agent Reliability
K. Ullrich, J. Su, C. Shi, A. Subramonian, A. Bar, I. Evtimov, **N. Tsilivis**, R. Balestriero, J. Kempe, M. Ibrahim, 2025

2. On the Geometry of Regularization in Adversarial Training: High-Dimensional Asymptotics and Generalization Bounds
M. Vilucchio, **N. Tsilivis**, B. Loureiro, J. Kempe, arXiv 2024

1. Extracting Finite Automata from RNNs Using State Merging
W. Merrill*, **N. Tsilivis***, arXiv 2022

Awards

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

École Normale Supérieure de Paris, CSD Seminar (2025)
Title: *How reinforcement learning after next-token prediction facilitates learning*

TTIC, 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 & Service:

- CDS & NYU GSAS Ph.D. Mentoring Programs: Paired with incoming students to provide guidance and support
- 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 optimization methods for reasoning in language models
  - Iason Kalogiannis (ECE, NTUA), 2025: dynamics analysis and regularization of tropical neural networks
- Member of the Student Inclusion and Belonging Advisory Board at CDS, NYU (Spring 2025)

Reviewing: ICLR 2023, 2025 (**Notable reviewer**), NeurIPS 2024 (**Top 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)

Programming Skills

Languages: Python, C/C++, Java, SML, Prolog

Other: PyTorch, JAX

Languages

Greek (native), English (fluent)