Lindsay M. Smith

lindsay.smith@princeton.edu | LS12005CS@gmail.com | lindsaymsmith3 | smithlindsay.github.io

Research Interest: My research focuses on the **science of AI**: I apply techniques and ideas from physics to artificial neural networks to understand how they learn. My current research in machine learning/AI includes projects in **interpretability, AI for scientific discovery, in-context learning, LLM multi-agent interactions, and AI safety**.

Education

M.A., Ph.D. — Princeton University — *Physics* 2022 - Present (expected 2027)
NSF GRFP; Research Advisors: Profs. William Bialek (Princeton) and David Schwab (CUNY)

B.A. — University of Pennsylvania — *Physics (Honors)* 2018 - 2022
Research Advisor: Prof. Dani Bassett
Cum Laude, Minors in Mathematics and French and Francophone Studies

Publications

* indicates equal contribution.

7. **Lindsay M. Smith***, Ananya Malik*, Edward James Young, Puria Radmard, Cameron Tice, and Hannes Whittingham (2026). Chain-of-Thought Injection as an Inference-Time Safety Intervention. *Accepted at **ICLR 2026** Workshop on Logical Reasoning of LLMs*, https://openreview.net/forum?id=v0XkjgeD6U.

6. Colin Scheibner*, **Lindsay M. Smith***, and William Bialek (2025). Large language models and the entropy of English. *Under review*, https://arxiv.org/abs/2512.249692.

5. Jeff Shen* & **Lindsay M. Smith*** (2025). Alice: An Interpretable Neural Architecture for Generalization in Substitution Ciphers. *Under review*, https://arxiv.org/abs/2509.07282.

4. Chase Goddard, **Lindsay M. Smith**, Vudtiwat Ngampruetikorn*, David J. Schwab* (2025). When can in-context learning generalize out of task distribution? ***ICML 2025***, https://arxiv.org/abs/2506.05574.

3. **Lindsay M. Smith**, Chase Goddard, Vudtiwat Ngampruetikorn*, David J. Schwab* (2024). Model Recycling: Model component reuse to promote in-context learning. ***NeurIPS 2024** Workshop on Scientific Methods for Understanding Deep Learning,* https://openreview.net/forum?id=vWSu8nEURM.

2. Chase Goddard, **Lindsay M. Smith**, Vudtiwat Ngampruetikorn*, David J. Schwab* (2024). Specialization-generalization transition in exemplar-based in-context learning. ***NeurIPS 2024** Workshop on Scientific Methods for Understanding Deep Learning,* https://openreview.net/forum?id=D1ui5QwHqF.

1. **Lindsay M. Smith**, Jason Z. Kim, Zhixin Lu, and Dani S. Bassett (2022). Learning continuous chaotic attractors with a reservoir computer, ***Chaos*** 32, 011101, https://doi.org/10.1063/5.0075572. *Selected as an Editor's Pick and publicized with a Scilight summary: https://doi.org/10.1063/10.0009079.*

Experience

Junior Research Scientist – Polymathic AI, New York, NY Mar 2026 - Aug 2026
Internship. Project: AI for Scientific Discovery

Mentee – MARS (Mentorship for Alignment Research Students), Cambridge, UK July 2025 - Feb 2026
MARS 3.0 participant at the Cambridge AI Safety Hub. Co-first author on project investigating AI control and steering via Chain-of-Thought (CoT) injections.

Honors and Awards

Citadel GQS Physics PhD Fellowship Finalist (top 3; 1 awarded) 2026

National Science Foundation (NSF) ACCESS Computing Award – 200,000 credits ($\sim$ $5000) 2025

I applied for and am the Principal Investigator for this award.

NSF AI Institutes Virtual Organization (AIVO) Travel Grant – $2000 2025

American Physical Society (APS) GSNP Student Speaker Award Finalist – $500 2025

NSF Graduate Research Fellowship Program (GRFP) – $138,000 + $12,000 bonus from Princeton 2022 - 2027

Charlotte and Morris Tanenbaum *52 Graduate Fellowship in Physical or Life Sciences – $40,000 2022 - 2023

Joseph Henry Merit Award – $5,000 2022

University Scholars Program – $7,000 2020 - 2022

Applied for and awarded summer research funding in 2020 and 2021.

National French Honor Society – Pi Delta Phi 2020 - Present

Sister Loretta Thome Scholarship – $10,000 2018 - 2023

Presentations

APS Global Summit, Denver, CO 2026

Talk: "Universality of LLM mechanisms across scale and diversity"

Princeton Center for Theoretical Science Workshop: The Physics of John Hopfield, Princeton, NJ 2025

Poster: "Alice: An Interpretable Neural Architecture for Generalization in Substitution Ciphers"

Meta 2025 PhD Forum, Menlo Park, CA 2025

Poster: "Alice: An Interpretable Neural Architecture for Generalization in Substitution Ciphers"

APS Global Summit, Anaheim, CA 2025

Talk: "Multi-Agent Debate: Analyzing Consensus in Networks of LLM Agents"

(GSNP Student Speaker Award Finalist)

APS March Meeting, Chicago, IL 2022

Talk: "Learning Continuous Chaotic Attractors with a Reservoir Computer"

Conference for Undergraduate Women in Physics (CUWiP), Virtual 2022

Poster: "Learning Continuous Chaotic Attractors with a Reservoir Computer"

Penn Research Expo, Phila., PA 2020, 2021, 2022

Posters: "Development of control in brain networks over temporal and spatial scales using graph models", "Learning Continuous Chaotic Attractors with a Reservoir Computer"

CUWiP, Virtual 2021

Lightning Talk: "Development of control in brain networks over temporal and spatial scales using graph models"

University Scholars Lunch Talk, Phila., PA 2020, 2022

Talks: "Development of control in brain networks over temporal and spatial scales using graph models", "Learning Continuous Chaotic Attractors with a Reservoir Computer"

APS March Meeting, Virtual 2020

Poster: "Development of control in brain networks over temporal and spatial scales using graph models"

Mentoring, Outreach, Teaching

PHY 103: General Physics I, **Teaching Assistant** 2025

Princeton Towards an Inclusive Physics cOmmunity in graduaTe Students (TIPOTS), **Mentor** 2024 - Present

Princeton Women in Physics, **Executive Board** 2023 - Present

Princeton Physics EDI, **Events Committee** 2022 - 2024

Princeton Physics Mentorship, **Mentor** 2022 - 2024

Mentored one to two undergraduate physics students each semester. Met at least once a semester to give career and academic advice.

CIS 110, **Tutor** 2021 - 2022

Tutored two to three students weekly in CIS 110: Introduction to Computer Programming.

Side By Side Agency, **Mentor** 2021

Mentored a student on her research project exploring astrophysics, advising her how to create a poster and conduct independent research.

Skills

Python, PyTorch, Jupyter, Git, MATLAB, Java, C++, ROOT, Mathematica, LaTeX, Linux