Weiyi Tian

weiyitian@uchicago.edu | weiyitian | weiyitian.com

EDUCATION

The University of Chicago Chicago, IL
Master of Science in Data Science *Mar 2026*

- GPA: 3.92/4.00
- Relevant Coursework: Mathematics of Generative Models, Mathematical Foundations of Machine Learning, Foundations of Machine Learning and AI, AI Agents

New York University New York City, NY
Bachelor of Arts in Computer Science and Mathematics *May 2024*

- GPA: 3.92/4.00
- Relevant Coursework: Linear Algebra, Multivariate Calculus, Real Analysis, Probability, Mathematical Statistics, Ordinary Differential Equations, Combinatorics, Abstract Algebra, Machine Learning, Artificial Intelligence, Data Structures, Algorithms

RESEARCH INTERESTS

Generative Modeling (LLMs, Diffusion), Reasoning and Post-Training, Mechanistic Interpretability

PUBLICATIONS

**Equal contribution*

Fine-Tuning Improves Information Conveyance in Language Models
Yuwei Cheng*, **Weiyi Tian***, Haifeng Xu
paper Under review at NeurIPS 2026

RESEARCH EXPERIENCE

Trajectory-Level Uncertainty in Autoregressive Generation Oct 2025 – Present
Advisor: Prof. Haifeng Xu *The University of Chicago*

- Implemented Monte Carlo estimators for Canopy Entropy, a tree-based measure of trajectory-level uncertainty in LLM generation, along with its normalized variants and a prompt-controlled correlation between length and uncertainty.
- Ran a sweep of 240K stochastic rollouts across base and instruct variants of Qwen3, Llama-3.1, and Gemma-3 on mathematical reasoning, coding, sentence completion, and story generation.
- Applied prompt-level cluster bootstrap to quantify standard errors; implemented linear mixed-effects models and ran quadratic robustness checks that revealed Gaussian misspecification, justifying the use of Beta regression.
- Showed that fine-tuning consistently shifts the correlation between generation length and entropy rate upward across tasks. Aligned models sustain information density over long generations rather than dilute it.
- Quantified that fine-tuned models convert token-level uncertainty into semantic variation more than twice as efficiently as base models, measured via embedding-based semantic diversity.

Manifold Geometry & Topology of LLM Knowledge Representations Feb 2026 – Present
Advisor: Prof. James Evans *The University of Chicago*

- Implemented a layer-wise geometric diagnostic on Llama-3.1 hidden states over cross-disciplinary arXiv abstracts, contrasting Euclidean (ambient) with kNN-graph geodesic (intrinsic) distances to localize where model representations depart from locally Euclidean geometry across depth.
- Built a persistent-homology analysis on the resulting hidden-state point clouds, tracking how topological features (connected components and loops) emerge and dissolve across transformer layers.

Activation Steering in Language and Vision-Language Models Mar 2025 – Sep 2025
Advisor: Prof. Ari Holtzman *The University of Chicago*

- Investigated why certain attention heads are more effective for activation steering by regressing per-head probe performance for political ideology on BOS attention-sink mass in Llama-2; found sink mass to

be a significant negative predictor after controlling for layer, suggesting heads less anchored to BOS encode political signal more linearly.

- Extended attention-head activation steering from text-only LLMs to the vision-language model Llava-v1.6 to reduce VQA hallucinations; evaluated hallucination rates via LLM-as-judge and model degradation via KL divergence.
- Observed overlap among the top steering attention heads for text and image interventions in Llava-v1.6, with text-derived steering vectors also reducing visual hallucinations, suggesting shared cross-modal representations in VLMs.

PROJECTS

Probability Flows in Transformer Representations Mar 2026

Advisor: Prof. Nisha Chandramoorthy *The University of Chicago*

- Framed the layer-wise evolution of Transformer representations as a discrete-time probability flow, with each layer inducing a pushforward between empirical measures of last-token hidden states.
- Quantified the flow via inter-layer 2-Wasserstein distances under entropic optimal transport, identifying a sharp surge in transport velocity at the terminal layer of GPT-2.

Speed–Accuracy Tradeoff in the Local Disinhibition Decision Model May 2022 – Sep 2022

Advisor: Prof. Paul Glimcher *New York University*

- Implemented simulations of an ODE-based neural circuit model to study the speed–accuracy tradeoff in animal decision-making, capturing flexible shifts in reaction-time distributions and accuracy rates.
- Reproduced empirical patterns from primate decision-making experiments at both behavioral (reaction-time distributions, accuracy rates) and neural (firing-rate trajectories, buildup rates, baseline shifts) levels.

AWARDS

Phi Beta Kappa *New York University*

Magna Cum Laude with High Honors in Computer Science *New York University*

University Honors Scholar *New York University*

Founder's Day Award for Academic Excellence *New York University*

Dean's List (2019–2022) *New York University*

Departmental Tuition Scholarship *The University of Chicago*

SKILLS

Programming: Python, Java, C, JavaScript, SQL

Machine Learning: PyTorch, HuggingFace, Weights & Biases

Infrastructure: Slurm, Hydra, Git, Bash, LaTeX