

Hongyi (Henry) Jin

📞 6153972330 | ✉️ hjenryin@gmail.com, hongyi@cs.ucla.edu | 🌐 @hjenryin

Looking for research interns; available dates are flexible.

PROFILE

- I'm a first-year PhD student at UCLA CS department under the supervision of Prof. Baharan Mirzasoleiman, working on **efficient and trustworthy machine learning algorithms**, especially in the context of reasoning LLMs and video diffusion models.
- My past experience include **adversarial attacks** in image classifiers, LLM watermarks and multi-agent LLM systems, **interpretable classifier** of medical images from **weak labels**, guaranteed **fairness** for targeted populations, and **efficient and preference-aligned** video diffusion models.
- Proficient in Python, PyTorch and C++.

EDUCATION

Computer Science Department, UCLA

PH.D. IN COMPUTER SCIENCE

Los Angeles, CA

Sep. 2025 – (Expected) Jun. 2030

- Advised by Prof. Baharan Mirzasoleiman. GPA: 4.0/4.0

Yao Class, Tsinghua University

B.S. IN COMPUTER SCIENCE AND TECHNOLOGY

Beijing, China

Aug. 2021 – Jun. 2025

- Special CS Pilot Class (established by Turing Award Laureate, Andrew Chi-Chih Yao). **GPA: 3.97/4.0 (Rank: 2/89)**

EXPERIENCES

Tencent

STUDENT RESEARCH INTERN (TENCENT QINGYUN PLAN)

Shanghai, China

Jul. 2025 – Aug. 2025

- Research Initiative: Researched methods for **step distillation** and **DPO** in video diffusers, developing a warmup phase to enhance stability during adversarial distillation, enabling a 100x reduced generation latency.
- Product development: Resolved **distribution-shift** issue causing artifacts in portrait animation from camera-motion videos, enabling realistic dance animation for **Wechat Channels**.

Software and Societal System Department, Carnegie Mellon University

RESEARCH INTERN

Pittsburgh, PA

Feb. 2024 – Aug. 2024

- Advised by Prof. Steven Wu
- Proposed a discretization-free algorithm that improves multicalibration, a key metric in **algorithm fairness**, removing the long-standing need of discretizing the outcome space.

ONGOING PROJECTS

Which Reasoning Traces Are Worth Generating Further? Data Curation for Training Reasoning Models

IN SUBMISSION TO *ICML2026*, FIRST AUTHOR

Oct. 2025 – Now

- Addressed the high cost of teacher reasoning generation **LLM reasoning distillation** by identifying diverse and difficult reasoning examples early during teacher generation, pruning the examples that aren't worth continuing.
- Used **loss** signals on **partial reasoning traces** evaluated by **randomly perturbed checkpoints** to remove easy and redundant examples, eliminating the inefficiency in previous work of handling gradients, long reasoning traces, and finetuning on the full dataset.
- Outperformed existing baselines by up to **2%** while generating only as few as **9%** tokens for reasoning traces.

PUBLICATIONS

- [1] J. Cheng, B. Ma, X. Ren, **H. H. Jin**, Y. Kai, P. Zhang, W. Li, Y. Zhou, T. Zheng, and Q. Lu, "POSE: Phased one-step adversarial equilibrium for video diffusion models," in *The Fortieth AAAI Conference on Artificial Intelligence*, 2026 (to appear). [Online]. Available: <https://arxiv.org/abs/2508.21019>
- [2] **H. H. Jin**, Z. Ding, D. D. Ngo, and S. Wu, "Discretization-free multicalibration through loss minimization over tree ensembles," in *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. [Online]. Available: <https://openreview.net/forum?id=56Y2HRjPlp>