

# Hanning Zhang

Homepage: [hanningzhang.github.io](https://hanningzhang.github.io)  $\diamond$  [Google Scholar](#)  
Phone: (+1) 217-979-3032  $\diamond$  Email: [hanningzhang@ucsb.edu](mailto:hanningzhang@ucsb.edu)

## EDUCATION

|                                                                                                            |                                     |
|------------------------------------------------------------------------------------------------------------|-------------------------------------|
| University of California, Santa Barbara<br><i>PhD in Computer Science</i>                                  | Sep. 2026 -<br><b>CPT Available</b> |
| University of Illinois, Urbana-Champaign (UIUC)<br><i>MS in Computer Science (thesis-track)</i>            | Aug. 2024 – May. 2026               |
| The Hong Kong University of Science and Technology<br><i>B.S. in Computer Science (First Class Honors)</i> | Sep. 2020 – Jun. 2024               |

## INDUSTRIAL EXPERIENCE

|                                                                                                                               |                                      |
|-------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|
| Amazon<br><i>Applied Scientist Intern, Supervisor: Dr. Lin Zhong</i><br>Topic: Multi-task and Multi-domain learning for Ads   | May 2026 - Aug 2026<br>Palo Alto, CA |
| Amazon<br><i>Applied Scientist Intern, Supervisor: Dr. Liang Qiu</i><br>Topic: Reward Models for Machine Learning Engineering | May 2025 - Aug 2025<br>Palo Alto, CA |

## RESEARCH SUMMARY

My current research interests are **Selective Tool Usage**, **Efficient Agent Systems**, and **Agentic RL**. Previously, I worked on **LLM Reasoning and Alignment**.

## PUBLICATIONS (\* Equal Contribution)

- [1] H. Mao\*, **Hanning Zhang\***, S. Diao, J. Jiang, H. Zhang, Z. Dong, S. Han, and P. Molchanov, “Hybrid Intelligence: An On-Device Orchestrator That Delegates to Cloud Experts,” Under Review. (**Recent Research**)
- [2] **Hanning Zhang\***, R. Wang\*, R. Pan\*, W. Wang\*, B. Meng, and T. Zhang, “PhysProver: Advancing Automatic Theorem Proving for Physics,” at *Findings of EMNLP 2026*.
- [3] **Hanning Zhang**, J. Song, J. Zhu, Y. Wu, T. Zhang, and C. Niu, “OpenGenAlign: A Preference Dataset for Trustworthy Reward Modeling in Open-Ended, Long-Context Generation,” at *Findings of ACL 2026*.
- [4] W. Xiong\*, **Hanning Zhang\***, C. Ye\*, L. Chen, N. Jiang, and T. Zhang, “Self-rewarding Correction for Mathematical Reasoning,” in *Arxiv*.
- [5] J. Yao\*, Y. Hao\*, **Hanning Zhang**, H. Dong, W. Xiong, N. Jiang, and T. Zhang, “Optimizing Chain-of-Thought Reasoners via Gradient Variance Minimization in Rejection Sampling and RL,” in *NeurIPS 2025*.
- [6] **Hanning Zhang\***, S. Diao\*, Y. Lin\*, Y. Fung, Q. Lian, X. Wang, Y. Chen, H. Ji, and T. Zhang, “R-Tuning: Instructing Large Language Models to Say ‘I Don’t Know’,” in *NAACL 2024*. (**Outstanding Paper Award, 6/2434 = 0.25%**)
- [7] **Hanning Zhang**, J. Yao, C. Ye, W. Xiong, and T. Zhang, “Online-DPO-R1: Unlocking Effective Reasoning Without the PPO Overhead,” in *Notion Blog*.
- [8] **Hanning Zhang**, P. Wang, S. Diao, Y. Lin, R. Pan, H. Dong, D. Zhang, P. Molchanov, and T. Zhang, “Entropy-Regularized Process Reward Model,” in *TMLR 2025*.
- [9] R. Pan\*, D. Zhang\*, **Hanning Zhang\***, X. Pan\*, M. Xu, J. Zhang, R. Pi, X. Wang, and T. Zhang, “ScaleBiO: Scalable Bilevel Optimization for LLM Data Reweighting,” in *ACL 2025*.

- [10] H. Hao\*, X. Pan\*, **Hanning Zhang\***, C. Ye, R. Pan, and T. Zhang, “**Understanding Overadaptation in Supervised Fine-Tuning: The Role of Ensemble Methods**,” in *ICML 2025*.
- [11] Y. Lin, H. Lin, W. Xiong, S. Diao, J. Liu, J. Zhang, R. Pan, H. Wang, W. Hu, **Hanning Zhang**, H. Dong, R. Pi, H. Zhao, N. Jiang, H. Ji, Y. Yao, and T. Zhang, “**Mitigating the Alignment Tax of RLHF**,” in *EMNLP 2024*.
- [12] Y. Hao\*, Y. Lu\*, **Hanning Zhang**, X. Shen, and T. Zhang, “**Towards Better Generalization via Distributional Input Projection Network**,” Under Review at *TMLR*.
- [13] Y. Wu, J. Song, **Hanning Zhang**, T. Zhang, and C. Niu, “**DuaShepherd: Integrating Stepwise Correctness and Potential Rewards for Mathematical Reasoning**,” in *Arxiv*.

## RESEARCH EXPERIENCE

---

### Research Assistant at ZenAI Lab

University of California, Santa Barbara

*Supervisor: Professor Zhen Dong*

*Jun 2026 - Now*

- Built Hybrid Intelligence, an on-device orchestrator that delegates to cloud experts, with NVIDIA collaborators (Under Review).

### Research Assistant at ScaleML Lab

University of Illinois Urbana-Champaign

*Supervisor: Professor Tong Zhang*

*Mar 2024 - May 2026*

- Research on LLM alignment, generalization, and reasoning: process reward modeling, ensemble methods against catastrophic forgetting, and data/training-efficient reasoning.
- Published in TMLR, ACL, EMNLP, ICML, and NeurIPS; open-sourced multiple training pipelines and datasets.

### Research Intern at Blender Lab

HKUST and UIUC

*Supervisors: Prof. Heng Ji and Professor Tong Zhang*

*May 2023 - Mar 2024*

- Proposed Refusal-Aware Fine-tuning to teach LLMs to say “I don’t know,” improving factuality and calibration (NAACL 2024, Outstanding Paper Award).

## SELECTED AWARDS

---

**Siebel Scholar of Class 2026 (\$35,000)**

**Outstanding Paper Award at NAACL 2024**