

# Bolian Li

PhD Student

✉ li4468 [AT] purdue [DOT] edu

🐙 lblaoke

🌐 lblaoke.github.io

## Education

### PhD, Computer Science

Purdue University

Advisor: Ruqi Zhang

2022 – Present

West Lafayette, IN, USA

### BE, Computer Science and Technology

Tianjin University

Advisor: Changqing Zhang

GPA: 3.8/4.0

2018 – 2022

Tianjin, China

## Experiences

### Amazon.com, Inc.

Applied Scientist Intern, working with Rui Song, Sheng Wang, Hengrui Cai, and Qingkai Zeng  
Reward Modeling Generalization

May 2025 – Aug. 2025

### University of Illinois Urbana-Champaign

Research Intern, working with Hanghang Tong  
Graph Representation Learning

May 2021 – Nov. 2021

### Institute of Computing Technology, Chinese Academy of Sciences

Research Intern, working with Shuhui Wang  
Video Object Detection

Jan. 2021 – Feb. 2021

## Research Interests

I am interested in the post-training of large language models (LLMs). Recently, I focus on the sampling efficiency of post-training algorithms. I am also widely interested in safety alignment, reward generalization, Bayesian deep learning, and data imbalance.

## Publications

### Reward-Shifted Speculative Sampling Is An Efficient Test-Time Weak-to-Strong Aligner

Conference on Empirical Methods in Natural Language Processing (EMNLP 2025)

Bolian Li, Yanran Wu, Xinyu Luo, Ruqi Zhang

### More is Less: The Pitfalls of Multi-Model Synthetic Preference Data in DPO Safety Alignment

Conference on Language Modeling (COLM 2025)

Yifan Wang, Runjin Chen, Bolian Li, David Cho, Yihe Deng, Ruqi Zhang, Tianlong Chen, Zhangyang Wang, Ananth Grama, Junyuan Hong

### Cascade Reward Sampling for Efficient Decoding-Time Alignment

Conference on Language Modeling (COLM 2025)

Bolian Li\*, Yifan Wang\*, Anamika Lochab\*, Ananth Grama, Ruqi Zhang

### Bayesian Computation in Deep Learning

MCMC Handbook

Wenlong Chen\*, Bolian Li\*, Ruqi Zhang, Yingzhen Li

### Stacey: Promoting Stochastic Steepest Descent via Accelerated $\ell_p$ -Smooth Nonconvex Optimization

International Conference on Machine Learning (ICML 2025)

Xinyu Luo\*, Site Bai\*, Bolian Li\*, Petros Drineas, Ruqi Zhang, Brian Bullins

**Making Reliable and Flexible Decisions in Long-tailed Classification**

Transactions on Machine Learning Research (TMLR)  
*Bolian Li, Ruqi Zhang*

**ETA: Evaluating Then Aligning Safety of Vision Language Models at Inference Time**

International Conference on Learning Representations (ICLR 2025)  
*Yi Ding, Bolian Li, Ruqi Zhang*

**Entropy-MCMC: Sampling from Flat Basins with Ease**

International Conference on Learning Representations (ICLR 2024)  
*Bolian Li, Ruqi Zhang*

**Long-tailed Classification from a Bayesian-decision-theory Perspective**

Symposium on Advances in Approximate Bayesian Inference (AABI 2023)  
*Bolian Li, Ruqi Zhang*

**Trustworthy Long-tailed Classifications**

The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022)  
*Bolian Li, Zongbo Han, Haining Li, Huazhu Fu, Changqing Zhang*

**Graph Communal Contrastive Learning**

The Web Conference (WWW 2022), Oral Presentation  
*Bolian Li, Baoyu Jing, Hanghang Tong*

*Services*

---

<b>Reviewer</b>	ICML 2025, NeurIPS 2024; ICLR 2025; AISTATS 2024, 2025; UAI 2024, 2025; SPIGM 2023, 2024; NextGenAISafety 2024
<b>Teaching Assistant</b>	CS 57800; CS 38003; CS 17600

*Awards*

---

Tan Xuguang Alumni Scholarship, <i>Tianjin University</i>	2020
Tianjin University Vision Education Fund, <i>Tianjin University</i>	2019