

# Jichu Li

Tel: +86-181-5508-2105 | Email: lijichu52@gmail.com  
Google Scholar | Github | Homepage | Last update: May, 2026

## EDUCATION

### Renmin University of China

B.S. in Statistics, School of Statistics

Beijing, China

Sep 2023 - Jun 2027 (expected)

**Overall GPA:** 3.98/4.00    **Ranked:** 2/168

### Relevant Coursework:

Mathematical Analysis I(93) II(94) III(93) | Advanced Algebra I(98) II(97) | Probability Theory(98) | Real Analysis(96)  
Mathematical Statistics(94) | Python Programming for Artificial Intelligence(96) | Data Structure and Algorithm(94)  
Machine Learning(98) | Stochastic Process(97) | Optimization(97) | Nonparametric Statistics(96)

### Renmin University of China

B.Econ.(Minor) in Finance, Gaoli Academy, Gaoling School of Artificial Intelligence

Beijing, China

Sep 2024 - Jun 2027 (expected)

### Toyota Technological Institute at Chicago

Research Visiting Student

Chicago, IL, USA

Jun 2026 – Sep 2026

**Host Faculty:** Prof. Zhiyuan Li

## PUBLICATIONS & PREPRINTS (\*equal contribution)

### [P1] Learning with Synthetic Data via SGD in High-Dimensional Linear Regression

Jichu Li, Difan Zou.

*Preprint submitted to NeurIPS 2026.*

*ICML 2026 Workshop on High-dimensional Learning Dynamics (HiLD)*

### [C4] The Implicit Bias of Steepest Descent with Mini-batch Stochastic Gradient

Jichu Li, Xuan Tang, Difan Zou.

*International Conference on Machine Learning (ICML) 2026.*

arXiv:2602.11557

### [C3] A Convergence Analysis of Adaptive Optimizers under Floating-point Quantization

Xuan Tang\*, Jichu Li\*, Difan Zou.

*International Conference on Learning Representations (ICLR) 2026.*

arXiv:2510.21314

### [C2] Long-range Modeling and Processing of Multimodal Event Sequences

Jichu Li\*, Yilun Zhong\*, Zhiting Li, Feng Zhou, Quyu Kong.

*International Conference on Learning Representations (ICLR) 2026. (Meta Score: 7; Top 1.28%)*

arXiv:2602.01125

### [C1] DanmakuTPPBench: A Multi-modal Benchmark for Temporal Point Process Modeling and Understanding

Yue Jiang, Jichu Li, Yang Liu, Dingkan Yang, Feng Zhou, Quyu Kong.

*Neural Information Processing Systems (NeurIPS) 2025.*

arXiv:2505.18411

## RESEARCH INTEREST

My research interests lie at the intersection of machine learning theory and optimization. I am dedicated to building a rigorous theoretical understanding of modern optimization algorithms in deep learning, analyzing their convergence, implicit bias, and generalization [P1][C3][C4]. At the same time, I am interested in applying these theoretical insights to the design and analysis of large language models (LLMs/MLLMs), including LLM-based approaches for modeling complex sequential and event-based data [C1][C2].

## RESEARCH EXPERIENCE

### The University of Hong Kong    Research Intern

Advisor: Prof. Difan Zou

May 2025 – Present

Hong Kong SAR, China

### Project (i): Convergence analysis of optimizers in mixed-precision training [C3]

Developed a rigorous theoretical framework to analyze the convergence properties of the Quantized Adam and Muon algorithm, a crucial component in mixed-precision training for deep learning models. Proved the convergence of Quantized Adam and Muon by establishing a bounded error for the expected squared gradient norm, demonstrating that the algorithm reliably converges under

specific conditions. Analyzed the combined impact of weight and gradient quantization on the training process, providing a theoretical basis for understanding and optimizing mixed-precision training.

**Project (ii): The Implicit bias of stochastic steepest descent [C4]**

Analyzed the implicit bias of mini-batch stochastic steepest descent (e.g., SignSGD, Muon) in multi-class classification, rigorously characterizing how stochasticity and batch size affect the convergence direction under general norm geometries. Proved that momentum ensures convergence in small-batch regimes, while variance reduction removes batch-dependent margin gaps, enabling stochastic algorithms to recover the full-batch max-margin solution, with fully explicit, dimension-free convergence rates improving prior bounds. Investigated the vanilla batch-size-one steepest descent without momentum or variance reduction and revealed its convergence to a fundamentally different bias.

**Project (iii): Learning with synthetic data via SGD [P1]**

Analyzed the effect of synthetic data on one-pass SGD in high-dimensional linear regression under model shift. Proved that mixed real-synthetic training causes strong model collapse through a non-vanishing excess-risk floor, whereas two-stage training provably eliminates this floor by training first on synthetic data and then on real data. Further Derived scaling laws under a random sketch model showing that larger models amplify synthetic-induced degradation under mixing, while high-quality synthetic data may improve two-stage training by reducing bias.

**Renmin University of China**   Research Intern  
Advisor: **Prof. Feng Zhou**

Jan. 2025 – Sep. 2025  
Beijing, China

**Project: LLM-based multimodal event sequence modeling [C1][C2]**

Designed and constructed a novel multimodal dataset benchmark incorporating diverse data types (text, timestamps, figures) to establish a standard for model evaluation. Integrated Large Language Models (LLMs) with Temporal Point Processes (TPPs) to significantly enhance inference and predictive capabilities for complex event sequences.

HONORS & AWARDS

---

|                                                                                                       |            |
|-------------------------------------------------------------------------------------------------------|------------|
| <b>National Scholarship (twice).</b> (Top scholarship, Ministry of Education, China. CNY 20000)       | 2024, 2025 |
| <b>HKU Summer Research Stipend.</b> (The University of Hong Kong. HK\$ 19600)                         | 2025       |
| <b>Honorable Mention.</b> Mathematical Contest in Modeling and Interdisciplinary Contest in Modeling. | 2025       |
| <b>Provincial First Prize.</b> Mathematics competition of Chinese College Students, CMC               | 2024       |
| <b>Provincial First Prize.</b> Contemporary Undergraduate Mathematical Contest in Modeling.           | 2024       |

PROFESSIONAL SERVICE

---

**Conference Reviewer**   ICML 2026, NeurIPS 2026  
**Instructor** - Statistical Investigation Association of Renmin University of China (2024-now):  
Taught Python programming and wrote tutorial articles for student groups at Renmin University of China.

SKILLS

---

**Programming:** Python(Proficient), C/C++, R.  
**Typesetting:** L<sup>A</sup>T<sub>E</sub>X, Markdown.  
**Language:** English (TOEFL: 110), Mandarin (Native)