

# Kaifeng LYU

## Personal Information

---

**Name:** Kaifeng Lyu (or Kaifeng Lv)      **Chinese Name:** 吕凯风  
**E-mail:** vfleaking@gmail.com  
klyu@cs.princeton.edu

## Education

---

### Aug 2024 (Expected): Ph.D. in Computer Science, Princeton University

**Note:** I could not go to the U.S. immediately after my bachelor's graduation since my visa application underwent an extremely long administrative processing for two years. I studied at Tsinghua University from Oct 2019 to Jul 2021 until the visa was finally approved.

**Advisor:** Prof. Sanjeev Arora.

### Jul 2019: B.Eng. in Computer Science and Technology, Tsinghua University

Studied in Yao Class (a special Computer Science class directed by Prof. Andrew C. Yao).

Yao Award, Silver Medal, 2018 (Top 3 out of 32 in Yao Class).

## Experiences

---

**Starting from Sep 2025 (Expected):**      *Assistant Professor at IIIS, Tsinghua University*

**Jun 2023 — Sep 2023:**      *Student Researcher at Google (Mentor: Sashank J. Reddi)*

**Jul 2018 — Sep 2018:**      *Visiting Student at Princeton University (Advisor: Prof. Sanjeev Arora)*

**Feb 2018 — May 2018:**      *Visiting Student at MIT (Advisor: Prof. Ryan Williams; Mentor: Lijie Chen)*

**Mar 2016 — Aug 2016:**      *Research Intern at Megvii (Face++) Inc. (Mentor: Dr. Gang Yu)*

## Selected Honors

---

- NeurIPS Top Reviewers Award, 2022.
- Gordon Y.S. Wu Fellowships in Engineering, 2021.
- NeurIPS Outstanding Reviewer Award, 2021 (Top 8%).
- 1st Place, ACM-ICPC Asia EC-Final Contest, 2015 (with Yuhao Du, Lijie Chen).
- 1st Place, ACM-ICPC Asia Hefei Regional Contest, 2015 (with Yuhao Du, Lijie Chen).
- Freshman Scholarship, Tsinghua University, 2015.
- Gold Medal (2nd Place), Chinese National Olympiad in Informatics, 2014.
- Gold Medal (5th Place), Asia-Pacific Informatics Olympiad, 2013.

## Teaching Experiences

---

- Fall 2022. Teaching Assistant for COS521: Advanced Algorithm Design (by Prof. Matt Weinberg & Prof. Huacheng Yu), Princeton University.
- Spring 2021. Teaching Assistant for COS598B: Advanced Topics in Computer Science: Mathematical Understanding of Deep Learning (by Prof. Sanjeev Arora), Princeton University.

- Spring 2020. Teaching Assistant for Mathematics for Computer Science (by Prof. Andrew Chi-Chih Yao), Tsinghua University.
- Spring 2019. Teaching Assistant for Distributed Computing (by Prof. Wei Chen), Tsinghua University.

## Professional Services

---

- Organizer, NeurIPS 2023 Workshop on Mathematics of Modern Machine Learning.
- Conference Reviewer: ICML (2020-2023), NeurIPS (2020-2023), ICLR (2022-2024), COLT (2020), AAAI (2020), KDD (2022).
- Journal Reviewer: TMLR, JMLR, TPAMI, AIJ.
- Organizer, Yao Class Seminar, Tsinghua (Fall 2019, Fall 2020, Spring 2021).

## Selected Papers

---

(Contribution order by default; Asterisk \* stands for equal contribution.)

- **A Quadratic Synchronization Rule for Distributed Deep Learning**  
Xinran Gu\*, **Kaifeng Lyu\***, Sanjeev Arora, Jingzhao Zhang, Longbo Huang.  
*The 12th International Conference on Learning Representations (ICLR 2024).*
- **Why (and When) does Local SGD Generalize Better than SGD?**  
Xinran Gu\*, **Kaifeng Lyu\***, Longbo Huang, Sanjeev Arora.  
*The 11th International Conference on Learning Representations (ICLR 2023).*
- **Dichotomy of Early and Late Phase Implicit Biases Can Provably Induce Grokking**  
**Kaifeng Lyu\***, Jikai Jin\*, Zhiyuan Li, Simon S. Du, Jason D. Lee, Wei Hu.  
*The 12th International Conference on Learning Representations (ICLR 2024).*
- **On the SDEs and Scaling Rules for Adaptive Gradient Algorithms**  
Sadhika Malladi\*, **Kaifeng Lyu\***, Abhishek Panigrahi, Sanjeev Arora.  
*The 36th Conference on Neural Information Processing Systems (NeurIPS 2022).*
- **Understanding the Generalization Benefit of Normalization Layers: Sharpness Reduction**  
**Kaifeng Lyu**, Zhiyuan Li, Sanjeev Arora.  
*The 36th Conference on Neural Information Processing Systems (NeurIPS 2022).*
- **Gradient Descent Maximizes the Margin of Homogeneous Neural Networks**  
**Kaifeng Lyu**, Jian Li.  
*The 8th International Conference on Learning Representations (ICLR 2020)*  
**Oral Presentation (Top 1.9%).**

## Other Papers

---

(Contribution order by default; Asterisk \* stands for equal contribution.)

- **Efficient Stagewise Pretraining via Progressive Subnetworks**  
Abhishek Panigrahi, Nikunj Saunshi, **Kaifeng Lyu**, Sobhan Miryoosefi, Sashank J. Reddi, Satyen Kale, Sanjiv Kumar.  
Preprint.
- **DistillSpec: Improving Speculative Decoding via Knowledge Distillation**  
Yongchao Zhou, **Kaifeng Lyu**, Ankit Singh Rawat, Aditya Krishna Menon, Afshin Rostamizadeh, Sanjiv Kumar, Jean-François Kagy, Rishabh Agarwal.  
*The 12th International Conference on Learning Representations (ICLR 2024).*
- **The Marginal Value of Momentum for Small Learning Rate SGD**  
Runzhe Wang, Sadhika Malladi, Tianhao Wang, **Kaifeng Lyu**, Zhiyuan Li.  
*The 12th International Conference on Learning Representations (ICLR 2024).*

- **Understanding Incremental Learning of Gradient Descent: A Fine-grained Analysis of Matrix Sensing**  
Jikai Jin, Zhiyuan Li, **Kaifeng Lyu**, Simon S. Du, Jason D. Lee.  
*The 40th International Conference on Machine Learning (ICML 2023).*
- **New Definitions and Evaluations for Saliency Methods: Staying Intrinsic, Complete and Sound**  
Arushi Gupta\*, Nikunj Saunshi\*, Dingli Yu\*, **Kaifeng Lyu**, Sanjeev Arora.  
*The 36th Conference on Neural Information Processing Systems (NeurIPS 2022).*  
**Oral Presentation (Top 1.9%).**
- **Towards Resolving the Implicit Bias of Gradient Descent for Matrix Factorization: Greedy Low-Rank Learning**  
Zhiyuan Li, Yuping Luo, **Kaifeng Lyu**. (Alphabetical order)  
*The 9th International Conference on Learning Representations (ICLR 2021).*
- **Gradient Descent on Two-layer Nets: Margin Maximization and Simplicity Bias**  
**Kaifeng Lyu\***, Zhiyuan Li\*, Runzhe Wang\*, Sanjeev Arora.  
*The 35th Conference on Neural Information Processing Systems (NeurIPS 2021).*
- **Reconciling Modern Deep Learning with Traditional Optimization Analyses: The Intrinsic Learning Rate**  
Zhiyuan Li\*, **Kaifeng Lyu\***, Sanjeev Arora.  
*The 34th Conference on Neural Information Processing Systems (NeurIPS 2020).*
- **Theoretical Analysis of Auto Rate-Tuning by Batch Normalization**  
Sanjeev Arora, Zhiyuan Li, **Kaifeng Lyu**. (Alphabetical order)  
*The 7th International Conference on Learning Representations (ICLR 2019).*
- **Fine-grained Complexity Meets  $IP = PSPACE$**   
Lijie Chen, Shafi Goldwasser, **Kaifeng Lyu**, Guy N. Rothblum, Aviad Rubinfeld. (Alphabetical order)  
*The 2019 ACM-SIAM Symposium on Discrete Algorithms (SODA 2019).*
- **Single-Source Bottleneck Path Algorithm Faster than Sorting for Sparse Graphs**  
Ran Duan, **Kaifeng Lyu** and Yuanhang Xie. (Alphabetical order)  
*The 45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).*
- **Learning Gradient Descent: Better Generalization and Longer Horizons**  
**Kaifeng Lv\***, Shunhua Jiang\*, Jian Li.  
*The 34th International Conference on Machine Learning (ICML 2017).*

## Open Source Project

Aug 2014 — now: Universal Online Judge (UOJ)

<https://uoj.ac>

A popular online judge system for OI (Olympiad in Informatics) in China.  
Open-sourced since 2016 (<https://github.com/vfleaking/uoj>).