Hanmin Li 李 瀚民

About Me

Place and Date of Birth:	Anhui, China	31 July 1999
Address:	Exploration Avenue, KAUST, 23955, Thuwal, Saudi Arabia
email:	hanmin.li AT kaust.edu.sa

Education

Present	Ph.D. student in Computer Science
King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
Supervisor: Prof. Peter Richtárik
December 2022	Master of Engineering in Computer Science
King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
Major: Computer Science
Department: Computer, Electrical and Mathematical Sciences and Engineering
Gpa: 3.86/4.00
July 2021	Bachelor of Engineering in Computer Science and Technology
University of Science and Technology of China (USTC), Hefei, Anhui, China
Major: Computer Science and Technology
Department: School of Gifted Young
Gpa: 3.64/4.30
August 2018	Summer School Student
University of Texas at Austin, Austin, Texas
Major: Computer Science (Software Engineering)

Publications and preprints

Preprint	"The Ball-Proximal (="Broximal") Point Method: a New Algorithm, Convergence Theory, and Applications"
Kaja Gruntkowska, **Hanmin Li**, Aadi Rane, Peter Richtárik
arXiv preprint arXiv:2502.02002
Workshop	"On the Convergence of FedProx with Extrapolation and Inexact Prox"
Hanmin Li and Peter Richtárik.
arXiv preprint arXiv:2410.01410 Oct, 2024; OPT2024-NeurIPS Poster
Conference	"The Power of Extrapolation in Federated Learning"
Hanmin Li, Kirill Acharya, and Peter Richtárik.
Conference on Neural Information Processing Systems, 2024
Workshop	"Variance Reduced Distributed Non-Convex Optimization Using Matrix Stepsizes"
Hanmin Li, Avetik Karagulyan, and Peter Richtárik.
arXiv preprint arXiv:2310.04614 Oct, 2023; FL@FM-NeurIPS'23 Poster
Conference	"Det-CGD: Compressed Gradient Descent with Matrix Stepsizes for Non-Convex Optimization"
Hanmin Li, Avetik Karagulyan, and Peter Richtárik.
International Conference on Learning Representations 2024.

JOURNAL "SD^2: spatially resolved transcriptomics deconvolution through integration of dropout and spatial information" Haoyang Li, **Hanmin Li**, Juexiao Zhou, Xin Gao. Bioinformatics, 38(21), pp.4878-4884. SEPTEMBER, 2022.

INVITED TALKS

7 May 2024
Vienna, Austria

International Conference on Learning Representations,
ICLR Poster
"Det-CGD: Compressed Gradient Descent with Matrix Stepsizes for Non-Convex Optimization".

26 June 2024
Lund, Sweden

EURO working group on Continuous Optimization
Invited Talk
"Compressed Gradient Descent with Matrix Stepsizes for Non-Convex Optimization".

12 Dec. 2024
Vancouver, Canada

Conference on Neural Information Processing Systems,
NeurIPS Poster
"The Power of Extrapolation in Federated Learning".

15 Dec. 2024
Vancouver, Canada

Annual Workshop on Optimization for Machine Learning,
NeurIPS Workshop Poster
"On the Convergence of FedProx with Extrapolation and Inexact Prox".

REVIEW SERVICES

NeurIPS 24, 25', NeurIPS OPT-ML 24', ICLR 25', ICML 25', JMLR, IEEE TNNLS, IEEE TSP, Optimization Methods and Software.

WORK EXPERIENCE

JAN 2021 - JUNE 2021

Research Intern at QULAB/USTC
Supervised by Prof. Kun Qu

SEPT 2019 - MARCH 2019

Research Intern at USTC
Supervised by Prof. Yongkun Li

SCHOLARSHIPS AND CERTIFICATES

SEPT. 2024

Marked as "outstanding" in the Ph.D. student yearly evaluation by my advisor Prof. Peter Richtárik. KAUST

SEPT. 2023

Marked as "outstanding" in the Ph.D. student yearly evaluation by my advisor Prof. Peter Richtárik. KAUST

SEPT. 2019

Scholarship for outstanding students (top % 20 students)
School of Gifted Young, USTC

SEPT. 2018

Scholarship for outstanding students (top % 20 students)
School of Gifted Young, USTC

SEPT. 2017

Scholarship for Shitsan Pai class of talented students, (top % 10 students)
University of Science and Technology of China

SEPTEMBER 2019

TOEFL®: 110 (READING:30; LISTENING:30; SPEAKING:23; WRITING:27)

FEBRUARY 2018

GRE®: 333 (VERBAL:163; QUANTITATIVE:170; AW:3,5)

Research Interests

My research focus lies at the intersection of optimization and large language models (LLMs), with a focus on training efficiency and scalability. I am also interested in distributed training and the theoretical foundations of learning from decentralized data.

More broadly, my interests also extends to the theory of modern machine learning, including first-order methods, convex and non-convex optimization, and operator theory, as well as applied areas like deep learning and language modeling.

Currently, I am working on:

- Distributed training of large language models (LLMs), including experience with large-scale GPU clusters and training using PyTorch Distributed Data Parallel (DDP).

- Efficient optimizer design for large-scale training, with a focus on advancing the Muon optimizer and its variants to achieve faster convergence and improved scalability.

- Designing efficient algorithms for large language models (LLMs), with a focus on both theoretical analysis and empirical validation.

Technical Skills

Python	PyTorch, with hands-on experience in PyTorch DistributedDataParallel (DDP), including customization of optimizers and operators, as well as development and communication optimization through gradient bucketing. Other packages such as NumPy, Pandas, Matplotlib ...
Others	C++, C, R
Math	In depth skills such as advanced linear algebra, convex & non-convex optimization, first order methods, operator theory, probabilistic ...

Languages

CHINESE:	Mothertongue
ENGLISH:	Fluent
FRENCH:	Intermediate