MINGZE WANG

210, Jingyuan Building #6, Peking University, Beijing, China, 100084
mingzewang@stu.pku.edu.cn

SHORT BIO

I am a third-year Ph.D student in Computational Mathematics, Peking University. I am very fortunate to be advised by Prof. Weinan E. Prior to that, I received my B.S. degree in Pure and Applied Mathematics (ranking 1/111 for the first three years during my undergraduate study) from Zhejiang University in 2021. My homepage is https://wmz9.github.io/.

EDUCATION

Peking University — Beijing, China
Ph.D Candidate, *Computational Mathematics* — *2021.09 - Present*
School of Mathematical Sciences
Advisor: Prof. Weinan E.

Zhejiang University — Hangzhou, China
Bachelor of Science, *Pure and Applied Mathematics* — *2017.09 - 2021.06*
School of Mathematical Sciences
Academic ranking: 1/111, Comprehensive ranking: 1/111, Major GPA: 4.84/5 (95.5/100).

EXPERIENCE

Peking University — Beijing, China
Teaching assistant: Deep Learning Theory, taught by Prof. Zhiyuan Li — *Summer School 2023.*
Teaching assistant: Calculus (A) — *Fall 2021*
Teaching assistant: Calculus (B) — *Fall 2022, 2023; Spring 2022, 2023*

Moqi Technology — Beijing, China
Algorithm Intern — *2021.09 - 2022.06*
Work on image processing and privacy protection for biometric technology.

RESEARCH INTERESTS

I am broadly interested in theory, algorithm and application of machine learning. I am also interested in non-convex and convex optimization. Specifically, my recent research topics are

- **Deep learning theory**: optimization, generalization, and implicit bias.
  - **Optimization**: When training neural networks, why can optimization algorithms converge to global minima? [1][4]
  - **Implicit Bias**: When training neural networks, why can optimization algorithms converge to global minima with favorable generalization ability (even without any explicit regularization)? Such as flat-minima-bias [2][5] and max-margin-bias aspects [4][6].
  - **Algorithm Design**: For machine learning problems, design new optimization algorithms which can converge to global minima with better generalization ability. [6]
  - **Generalization**: How to measure the generalization ability of neural networks. [3]
- **Foundation Model and Transformer**: theory and algorithm.
  - **Expressive Power**: The expressive power of Transformer (On the preparation).
  - **Algorithm Design**: (On the preparation).
- **Non-convex and Convex Optimization**: theory and algorithm.
  - **Convex Optimization in ML**. [6]

  - **Non-convex Optimization in ML**. [1][4]

- **CV and NLP**: algorithm and application.
- **AI for Compositional Optimization**: theory and algorithm.

PUBLICATIONS & PREPRINTS

[1] **Mingze Wang**, Chao Ma. **Early Stage Convergence and Global Convergence of Training Mildly Parameterized Neural Networks.** *Thirty-sixth Conference on Neural Information Processing Systems* (**NeurIPS 2022**)(73 pages). *2022.*

[2] Lei Wu, **Mingze Wang**, Weijie J. Su. **The alignment property of SGD noise and how it helps select flat minima: A stability analysis.** *Thirty-sixth Conference on Neural Information Processing Systems* (**NeurIPS 2022**). *2022.*

[3] **Mingze Wang**, Chao Ma. **Generalization Error Bounds for Deep Neural Networks Trained by SGD.** Under review. *arXiv preprint: 2206.03299, 2022.*

[4] **Mingze Wang**, Chao Ma. **Understanding Multi-phase Optimization Dynamics and Rich Nonlinear Behaviors of ReLU Networks.** *Thirty-seventh Conference on Neural Information Processing Systems* (**NeurIPS 2023**, **Spotlight (Top 3.5%)**)(94 pages). *2023.*

[5] **Mingze Wang**, Lei Wu. **The Noise Geometry of Stochastic Gradient Descent: A Quantitative and Analytical Characterization.** *NeurIPS 2023 Workshop on Mathematics of Modern Machine Learning* (**NeurIPS 2023 Workshop M3L**). *arXiv preprint: 2310.00692, 2023.*

[6] **Mingze Wang**, Zeping Min, Lei Wu. **Achieving Margin Maximization Exponentially Fast via Progressive Norm Rescaling.** *arXiv preprint: 2311.14387, 2023.*

[7] **Mingze Wang**, Weinan E. **Understanding the Expressive Power of Transformer for Next Token Prediction.** On the preparation. *2023.*

SELECTED TALKS & PRESENTATIONS

Some mathematical modeling problems and machine learning theory, *Schlumberger (Beijing).* *2023.07*

SERVICE

Conference: Conference on Neural Information Processing Systems (NeurIPS); International Conference on Learning Representations (ICLR).

Journal: Journal of Machine Learning Research (JMLR); Journal of Machine Learning (JML).

SELECTED AWARDS & HONOURS

BICMR Mathematical Award for Graduate Students (top 1%, 110,000 RMB) *2023.11*
Schlumberge Scholarship (30,000 RMB) *2022.10*
PKU Academic Innovation Award (top 1%) *2022.10*
Outstanding Graduate of Zhejiang Province (top 5%) *2021.05*
Outstanding Graduate of ZJU *2021.05*
Chinese National Scholarship (top 1%) *2019.10*
First Class Scholarship of ZJU (top 3%) *2019, 2020.10*
Zhejiang Provincial Government Scholarship *2018.10*
First Prize of Mathematical Contest in Modeling of ZJU (top 1%) *2020.06*
Meritourious Award in The Mathematical Contest in Modeling *2020.02*
National Second Prize of Chinese Undergraduate Mathematical Contest in Modeling (top 2.5%) *2019.10*

SELECTED UNDERGRADUATE TRANSCRIPT

Real Analysis	100	Functional Analysis	100	Partial Differential Equation	100
Scientific Computing	100	Mathematical Analysis (II)	99	Differential Geometry	99
Point Topology	99	Mathematical Physics	97	Complex Analysis	97
Calculus (I)	97	Stochastic Process	96	Foundation of Analysis	96