Zhongjie Shi — Curriculum Vitae

Georgia Institute of Technology, Atlanta – United States
+1 4043240577 • zshi332@gatech.edu •

Research Interests

Deep learning theory, Neural network approximation theory, Operator learning, Optimization, Distributed learning, Trustworthy AI.

Professional Experiences

Georgia Institute of Technology (GaTech) **United States**
Visiting Assistant Professor in Mathematics *2025-Present*
Mentor: Prof. Wenjing Liao

The University of Hong Kong (HKU) **Hong Kong**
Postdoctoral Fellow in Computing and Data Science *2023-2025*
Mentor: Prof. Yuan Cao

Katholieke Universiteit Leuven (KU Leuven) **Belgium**
Postdoctoral Fellow in Electrical Engineering *2022-2023*
Mentor: Prof. Johan A.K. Suykens

Education

City University of Hong Kong (CityU) **Hong Kong**
Ph.D. in Data Science *2018-2022*
Supervisor: Prof. Ding-Xuan Zhou

City University of Hong Kong (CityU) **Hong Kong**
B.Sc. in Computing Mathematics; Minor in Computer Science *2014-2018*
Advisor: Prof. Xiang Zhou

Honors and Awards

Outstanding Research Thesis Award (8 awardees at CityU in 2022) (10/2022) *CityU*
HKSAR Government Scholarship (Top 2% university-wide) (08/2017) *CityU*
HKSAR Government Scholarship Fund - Talent Development Scholarship (08/2015) *CityU*

Publications

* indicates corresponding author. ____ indicates alphabetical order.

1. **Zhongjie Shi***, and Wenjing Liao. Learning Theory of Transformers: Local-to-Global Approximation via Softmax Partition of Unity. *The 40th Annual Conference on Neural Information Processing Systems*, 2026.

2. **Zhongjie Shi***, Fanghui Liu, Yuan Cao, and Johan A.K. Suykens. Can overfitted deep neural networks in adversarial training generalize? An approximation viewpoint. *SIAM Journal on Mathematics of Data Science*, 8(2): 225-256, 2026.

3. **Zhongjie Shi**, Puyu Wang, Chenyang Zhang, and Yuan Cao. Towards Understanding Generalization in DP-GD: A Case Study in Training Two-Layer CNNs. *The 40th Annual AAAI Conference on Artificial Intelligence*, 2026.
4. Zhan Yu, **Zhongjie Shi**, and Deming Yuan. High-Probability Convergence Theory for Distributed Composite Optimization with Sub-Weibull Noises. *IEEE Transactions on Automatic Control*, doi: 10.1109/TAC.2026.3687492, 2026.
5. Zhan Yu, **Zhongjie Shi**, Deming Yuan, and Daniel W. C. Ho. Generic Frameworks for Distributed Functional Optimization and Learning Over Time-Varying Networks. *IEEE Transactions on Control of Network Systems*, doi: 10.1109/TCNS.2026.3701536, 2026.
6. Zhan Yu, Daniel W. C. Ho, Deming Yuan, and **Zhongjie Shi**. Distributed composite optimization with distributed composite block mirror descent. *IEEE Transactions on Automatic Control*, doi: 10.1109/TAC.2026.3652901, 2026.
7. **Zhongjie Shi***, Jun Fan, Linhao Song, Ding-Xuan Zhou, and Johan A.K. Suykens. Nonlinear functional regression by functional deep neural network with kernel embedding. *Journal of Machine Learning Research*, 26(284): 1-49, 2025.
8. **Zhongjie Shi**, Zhiying Fang, and Yuan Cao. Approximation and estimation capability of Vision Transformers for hierarchical compositional models. *Applied and Computational Harmonic Analysis*, 82: 101849, 2026.
9. **Zhongjie Shi**, Zhan Yu, and Ding-Xuan Zhou. Learning theory of distribution regression with neural networks. *Constructive Approximation*, 62(1): 61-104, 2025.
10. Zhan Yu, Jun Fan, **Zhongjie Shi**, and Ding-Xuan Zhou. Distributed gradient descent for functional learning. *IEEE Transactions on Information Theory*, 70(9): 6547-6571, 2024.
11. Tong Mao, **Zhongjie Shi***, and Ding-Xuan Zhou. Approximating functions with multi-features by deep convolutional neural networks. *Analysis and Applications*, 21(1): 93-125, 2023.
12. Tong Mao, **Zhongjie Shi**, and Ding-Xuan Zhou. Theory of deep convolutional neural networks III: Approximating radial functions. *Neural Networks*, 144: 778-790, 2021.
13. Zhan Yu, Daniel W. C. Ho, **Zhongjie Shi**, and Ding-Xuan Zhou. Robust kernel-based distribution regression. *Inverse Problems*, 37(10): 105014, 2021.

Submissions

14. Zhan Yu, **Zhongjie Shi**, and Ding-Xuan Zhou. Learning Theory of Decentralized Robust Kernel-Based Learning Algorithm. Under review at *Journal of Machine Learning Research*.
15. **Zhongjie Shi**, Rongjie Lai, Alexander Cloninger, and Wenjing Liao. Transformers as Cross-Task Learners: Shared Structure Drives Sample Efficiency in In-Context Learning. arXiv preprint arXiv:2609.29060, 2026. Submitted.
16. Haoran Yan, **Zhongjie Shi**, Yuanze Xi, Peng Chen, and Wenjing Liao. Neural Scaling Laws of Transformer Operator Network, arXiv preprint arXiv:2609.33533, 2026. Submitted.
17. Zhan Yu, **Zhongjie Shi**, and Daniel W. C. Ho. Non-convex Distributed Stochastic Optimization under Heavy-Tailed Noise: High-Probability Convergence Approach. To Submit.

Teaching Services

- **Lecturer**
  - MATH 2551 (BSc): Multivariable Calculus, Georgia Institute of Technology (Fall 2026)
  - MATH 2550 (BSc): Introduction to Multivariable Calculus, Georgia Institute of Technology (Spring 2026)

- MATH 1553 (BSc): Introduction to Linear Algebra, Georgia Institute of Technology (Fall 2025)

◦ **Teaching Assistant**
- SDSC 8003 (PhD): Machine Learning, City University of Hong Kong (Fall 2021)
- SDSC 1001 (BSc): Introduction to Data Science, City University of Hong Kong (Spring 2021)
- SDSC 6011 (MSc): Optimization for Data Science, City University of Hong Kong (Fall 2019)
- MA 1201 (BSc): Calculus and Basic Linear Algebra II, City University of Hong Kong (Spring 2019)
- MA 2503 (BSc): Linear Algebra, City University of Hong Kong (Fall 2018)

Presentations and Posters

Scientific Machine Learning: Theory, Algorithms, Applications, and Agentic AI Workflows **2026**
University of Notre Dame
Lightning Talk: Approximation and Generalization of Transformers for In-Context Learning with Cross-Task Structure.

2026 Southeast Applied & Computational Math (ACM) Student Workshop **2026**
Emory University
Poster: Towards Understanding Generalization in DP-GD: A Case Study in Training Two-Layer CNNs.

2026 Georgia Scientific Computing Symposium **2026**
University of Georgia
Poster: Towards Understanding Generalization in DP-GD: A Case Study in Training Two-Layer CNNs.

Research at the Interface of Applied Mathematics and Machine Learning **2025**
University of Houston
Poster: Towards Understanding Generalization in DP-GD: A Case Study in Training Two-Layer CNNs.

MATH 5470: Statistical Machine Learning **2025**
Hong Kong University of Science and Technology
Talk: Can we avoid robust overfitting in adversarial training? - An approximation viewpoint.

SIAM Seminar **2024**
Georgia Institute of Technology
Talk: Nonlinear functional regression by functional deep neural network with kernel embedding.

7th International Conference on Econometrics and Statistics (EcoSta 2024) **2024**
Beijing Normal University
Talk: Nonlinear functional regression by functional deep neural network with kernel embedding.

SDSC Seminar **2024**
City University of Hong Kong
Talk: Nonlinear functional regression by functional deep neural network with kernel embedding.

Mentoring Experience

Ph.D. Research Mentoring **2026-Present**
Georgia Institute of Technology
Student: Haoran Yan, a second-year Ph.D. student supervised by Prof. Wenjing Liao at Georgia Tech.
Project: Neural Scaling Laws of Transformer Operator Network.
Role: Supporting the student's research under Prof. Wenjing Liao's supervision, with regular meetings and guidance on mathematical formulation, proof of learning theory, and writing.

Master's Thesis Day-to-day Advisor (Master Artificial Intelligence) **2022–2023**
KU Leuven
Student: Xinjie Zeng, later a Ph.D. student supervised by Prof. Johan Suykens at KU Leuven.
Project: Utilization of convolutional kernels in deep restricted kernel machines for image tasks.
Role: Proposed the research problem, guided the student through the relevant literature, held regular meetings to discuss the mathematical ideas and implementation details, and provided feedback on experiments and thesis writing.

Professional Services

- **Academic Event Organization and Outreach**
  - Organizer, 2nd Atlanta Mathematics of Data Science Bootcamp, Georgia Institute of Technology, Atlanta, GA (September 19–20, 2026)
- **Journal Reviewing Activities**
  - IEEE Transactions on Neural Networks and Learning Systems
  - Neural Networks
  - Applied and Computational Harmonic Analysis
  - IEEE Transactions on Automatic Control
  - Applied Numerical Mathematics
  - Analysis and Applications
  - Signal Processing
  - IEEE Transactions on Pattern Analysis and Machine Intelligence
  - Journal of Approximation Theory
- **Conference Reviewing Activities**
  - The 38th and 40th Conferences on Neural Information Processing Systems (NeurIPS 2024, 2026)
  - The Thirteenth International Conference on Learning Representations (ICLR 2025)
  - The 40th Annual AAAI Conference on Artificial Intelligence (AAAI 2026)