# Yaoyu Zhang

#### zhyy.sjtu@sjtu.edu.cn

324 Build. No.5, Science Buildings, Shanghai Jiao Tong University

#### Tenure-track Associate Professor, 2020. 9 -

Institute of Natural Sciences and School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai

### Member, 2019. 9 – 2020. 7

School of Mathematics, Institute for Advanced Study, Princeton, New Jersey

#### Post-doctoral Associate, 2016-2019

NYUAD Institute, New York University Abu Dhabi and Courant Institute of Mathematical Sciences, New York University Supervisors: David Cai, David W. McLaughlin

## Education

### Ph.D in Applied Mathematics, 2012-2016

School of Mathematical Sciences and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China Advisors: Douglas Zhou, David Cai PhD thesis (in Chinese): Sampling Artifacts of Granger Causality and the Reliable Reconstruction of Neuronal Network Connectivity

## B.S. in Applied Physics, minor in Applied Mathematics, 2008-2012

Zhiyuan College, Shanghai Jiao Tong University, Shanghai, China

## **Research Interests**

#### **Deep Learning Theory**

- Condensation phenomenon
- Loss landscape structure
- Implicit regularization

## **Publications**

#### **Highlights in Deep Learning theory**

• Optimistic Estimate

[1] **Yaoyu Zhang\***, Zhongwang Zhang, Leyang Zhang, Zhiwei Bai, Tao Luo, Zhi-Qin John Xu\*, Linear Stability Hypothesis and Rank Stratification for Nonlinear Models, arxiv 2211.11623 (2022)

[2] **Yaoyu Zhang\***, Zhongwang Zhang, Leyang Zhang, Zhiwei Bai, Tao Luo, Zhi-Qin John Xu\*, Optimistic Estimate Uncovers the Potential of Nonlinear Models. arXiv:2307.08921, (2023).

[3] **Yaoyu Zhang\***, Leyang Zhang, Zhongwang Zhang and Zhiwei Bai, Local Linear Recovery Guarantee of Deep Neural Networks at Overparameterization. arXiv:2406.18035, (2024).

[4] (Alphabetic order) Tao Luo\*, Leyang Zhang, **Yaoyu Zhang**, Structure and Gradient Dynamics Near Global Minima of Two-layer Neural Networks. arXiv:2309.00508, (2023).

• Embedding Principle

[1] **Yaoyu Zhang**\*, Zhongwang Zhang, Tao Luo, Zhi-Qin John Xu\*, Embedding Principle of Loss Landscape of Deep Neural Networks. NeurIPS 2021 spotlight.

[2] **Yaoyu Zhang\***, Yuqing Li, Zhongwang Zhang, Tao Luo, Zhi-Qin John Xu\*, Embedding Principle: a hierarchical structure of loss landscape of deep neural networks. Journal of Machine Learning, 1(1), pp. 60-113, 2022.

[3] Zhiwei Bai, Tao Luo, Zhi-Qin John Xu\*, **Yaoyu Zhang**\*, Embedding Principle in Depth for the Loss Landscape Analysis of Deep Neural Networks. CSIAM Trans. Appl. Math., 5 (2024), pp. 350-389.

• Phase diagram

[1] Tao Luo#, Zhi-Qin John Xu#, Zheng Ma, **Yaoyu Zhang**\*, Phase diagram for two-layer ReLU neural networks at infinite-width limit, Journal of Machine Learning Research (JMLR) 22(71):1–47, (2021)

[2] Hanxu Zhou, Qixuan Zhou, Zhenyuan Jin, Tao Luo, **Yaoyu Zhang**, Zhi-Qin John Xu\*, Empirical Phase Diagram for Three-layer Neural Networks with Infinite Width. arxiv 2205.12101, (2022)

## • Frequency Principle

[First paper] Zhiqin Xu, Yaoyu Zhang, Yanyang Xiao, Training Behavior of Deep Neural Network in Frequency Domain, International Conference on Neural Information Processing (ICONIP), pp. 264-274, 2019. (arXiv:1807.01251, Jul 2018)

[2021 World Artificial Intelligence Conference Youth Outstanding Paper Nomination Award] Zhi-Qin John Xu\*, Yaoyu Zhang, Tao Luo, Yanyang Xiao, Zheng Ma, Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks, Communications in Computational Physics (CiCP) 28(5). 1746-1767, 2020.

[Linear Frequency Principle] Yaoyu Zhang, Tao Luo, Zheng Ma, Zhi-Qin John Xu\*, Linear Frequency Principle Model to Understand the Absence of Overfitting in Neural Networks, Chinese Physics Letters (CPL) 38(3), 038701, 2021.

[Linear Frequency Principle] (Alphabetic order) Tao Luo\*, Zheng Ma, Zhi-Qin John Xu, Yaoyu Zhang, On the exact computation of linear frequency principle dynamics and its generalization, SIAM Journal on Mathematics of Data Science (SIMODS) to appear, arxiv 2010.08153 (2020).

[**Overview**] Zhi-Qin John Xu\*, **Yaoyu Zhang**, Tao Luo, Overview frequency principle/spectral bias in deep learning. Communications on Applied Mathematics and Computation (2024): 1-38.

### **Deep Learning theory**

- Leyang Zhang, **Yaoyu Zhang**, and Tao Luo. Geometry of Critical Sets and Existence of Saddle Branches for Two-layer Neural Networks. arXiv preprint arXiv:2405.17501 (2024).
- Zhongwang Zhang, Pengxiao Lin, Zhiwei Wang, **Yaoyu Zhang**, and Zhi-Qin John Xu. Initialization is Critical to Whether Transformers Fit Composite Functions by Inference or Memorizing. NeurIPS 2024
- Zhiwei Bai, Jiajie Zhao, and **Yaoyu Zhang\***. Connectivity Shapes Implicit Regularization in Matrix Factorization Models for Matrix Completion. NeurIPS 2024
- Zhiwei Wang, Yunji Wang, Zhongwang Zhang, Zhangchen Zhou, Hui Jin, Tianyang Hu, Jiacheng Sun, Zhenguo Li, **Yaoyu Zhang**, and Zhi-Qin John Xu. Towards understanding how transformer perform multi-step reasoning with matching operation. arXiv preprint arXiv:2405.15302 (2024).
- Leyang Zhang, Zhi-Qin John Xu, Tao Luo\*, **Yaoyu Zhang**\*, Limitation of characterizing implicit regularization by data-independent functions, Transactions on Machine Learning Research (2023).
- Hanxu Zhou, Qixuan Zhou, Tao Luo, **Yaoyu Zhang**\*, Zhi-Qin John Xu\*, Towards Understanding the Condensation of Neural Networks at Initial Training, NeurIPS 2022.
- (Alphabetic order) Tao Luo\*, Zheng Ma, Zhiwei Wang, Zhi-Qin John Xu, **Yaoyu Zhang**, An Upper Limit of Decaying Rate with Respect to Frequency in Deep Neural Network, MSML 2022
- (Alphabetic order) Jihong Wang, Zhi-Qin John Xu\*, Jiwei Zhang\*, **Yaoyu Zhang**, Implicit bias with Ritz-Galerkin method in understanding deep learning for solving PDEs, CSIAM Trans. Appl. Math. 3(2), pp. 299-317, 2022.
- (Alphabetic order) Tao Luo, Zheng Ma, Zhi-Qin John Xu, **Yaoyu Zhang**, Theory of the Frequency Principle for General Deep Neural Networks, CSIAM Trans. Appl. Math. 2 (2021), pp. 484-507
- Yaoyu Zhang, Zhi-Qin John Xu\*, Tao Luo, Zheng Ma, A type of generalization error induced by initialization in deep neural networks, MSML 2020.

## **Computational Neuroscience**

- Yaoyu Zhang, Lai-Sang Young, DNN-Assisted Statistical Analysis of a Model of Local Cortical Circuits, Scientific Reports 10, 20139, 2020.
- Yaoyu Zhang, Yanyang Xiao, Douglas Zhou, David Cai, Spike-Triggered Regression for Synaptic Connectivity Reconstruction in Neuronal Networks, Frontiers in Computational Neuroscience 11, 101, 2017.

- **Yaoyu Zhang**, Yanyang Xiao, Douglas Zhou, David Cai, Granger Causality Analysis with Nonuniform Sampling and Its Application to Pulse-coupled Nonlinear Dynamics, Physical Review E 93, 042217, 2016.
- Douglas Zhou, **Yaoyu Zhang**, Yanyang Xiao, David Cai, Analysis of Sampling Artifacts on the Granger Causality Analysis for Topology Extraction of Neuronal Dynamics, Frontiers in Computational Neuroscience 8, 75, 2014.
- Douglas Zhou, **Yaoyu Zhang**, Yanyang Xiao, David Cai, Reliability of the Granger Causality Inference, New Journal of Physics 16 (4), 043016, 2014.
- Douglas Zhou, Yanyang Xiao, **Yaoyu Zhang**, Zhiqin Xu, David Cai, Granger Causality Network Reconstruction of Conductance-Based Integrate-and-Fire Neuronal Systems, PloS One 9 (2), e87636, 2014.
- Douglas Zhou, Yanyang Xiao, **Yaoyu Zhang**, Zhiqin Xu, David Cai, Causal and Structural Connectivity of Pulse-coupled Nonlinear Networks, Physical Review Letters 111 (5), 054102, 2013.

## **Deep Learning for Science**

- Zhiwei Wang, **Yaoyu Zhang**, Pengxiao Lin, Enhan Zhao, E. Weinan, Tianhan Zhang, and Zhi-Qin John Xu. Deep mechanism reduction (DeePMR) method for fuel chemical kinetics. Combustion and Flame 261 (2024): 113286.
- Tianhan Zhang\*, Yuxiao Yi, Yifan Xu, Zhi X. Chen, **Yaoyu Zhang**, Weinan E, Zhi-Qin John Xu\*, A multi-scale sampling method for accurate and robust deep neural network to predict combustion chemical kinetics. Combustion and Flame, 245, 112319, 2022.
- Lulu Zhang, Zhi-Qin John Xu\*, **Yaoyu Zhang**\*, Data-informed Deep Optimization. PLoS ONE 17 (6), e0270191, 2022
- Zhiwei Wang, **Yaoyu Zhang**, Yiguang Ju, Weinan E, Zhi-Qin John Xu\*, Tianhan Zhang\*, A deep learning-based model reduction (DeePMR) method for simplifying chemical kinetics. arxiv 2201.02025 (2022)
- Lulu Zhang, Tao Luo, Yaoyu Zhang, Weinan E, Zhi-Qin John Xu\*, Zheng Ma\*, MOD-Net: A Machine Learning Approach via Model-Operator-Data Network for Solving PDEs. Communications in Computational Physics (CiCP) (2022) to appear, arxiv 2107.03673 (2021)
- Tianhan Zhang, **Yaoyu Zhang**, Weinan E, Yiguang Ju, DLODE: a deep learning-based ODE solver for chemistry kinetics, AIAA Scitech 2021 Forum, 1139.

## **Select Presentations**

## Invited Talks

- **Optimistic Sample Size Estimate for Deep Neural Networks,** OneWorld Seminar, Oct. 9, 2024, Online
- Optimistic Estimate of Sample Efficiency for Neural Networks, Fudan University, Sep. 26, 2024, Shanghai

- **Optimistic Sample Size Estimate for Deep Neural Networks,** Numerical Analysis Seminar, The University of Hong Kong, Sep. 4, 2024, Online
- **Optimistic Estimate for Nonlinear Regression,** Tsinghua Statistics Seminar, Jun.28, 2023, Beijing
- Embedding Principle of Loss Landscape of Deep Neural Networks,海纳 百川上海青年学者讲座,华为上海研究所,2022.3.24, online
- Embedding Principle of Loss Landscape of Deep Neural Networks, AI TIME 青年科学家讲座, AI TIME, 2022.2.24, online
- Embedding Principle of Loss Landscape of Deep Neural Networks, Oct.16, 机器学习联合研讨计划, online
- Impact of initialization on generalization of deep neural networks, Oct.08, 2020, Math Machine Learning Seminar MPI MIS + UCLA, ZOOM.
- A Fourier Analysis Perspective of Training Dynamics of Deep Neural Networks, Dec.02, 2019, CMSA Special Seminar, Harvard University
- A Fourier Analysis Perspective of Training Dynamics of Deep Neural Networks, Nov.26, 2019, Seminar on Theoretical Machine Learning, Institute for Advanced Study, Princeton
- Quantitative Understanding of F-Principle, Jul.24, 2019, Beijing Computational Science Research Center
- Validity and reliability of Granger Causality analysis on neuronal network reconstruction, Nov.02, 2018, Purdue University
- Characterization of the training process of neural network in phase space, Oct.23, 2018, Rensselaer Polytechnic Institute
- Validity and Reliability of Granger Causality Analysis on Neuronal Network Reconstruction, Workshop on Data Assimilation and Information Theory, Jun.7, 2017, Fudan University
- Spike-Triggered Regression on Neuronal Network Reconstruction, Biomathematics/Computational Biology Colloquium Seminar, Nov.22, 2016, Courant Institute of Mathematical Sciences
- Reliability of Granger Causality Analysis and Spike-triggered Regression method on Neuronal Network Reconstruction, Dec.03, 2015, Rensselaer Polytechnic Institute
- Validity and Reliability of Granger Causality Analysis on Neuronal Network Reconstruction, Quantitative Biology Seminar, Nov.23, 2015, University of Michigan

# Conference Talks

- Optimistic Estimate of Sample Efficiency for Neural Networks, 数学与人 工智能科学大会 minisymposium, Aug. 29, 2024, Chongqing
- Optimistic Estimate of Sample Efficiency for Neural Networks, CSML2024, Aug. 10, 2024, Shanghai
- **Optimistic Estimate for Deep Neural Networks,** SIAM Annual Meeting minisymposium, Jul. 20, 2024, Online.

- Understanding Condensation via Embedding Principle and Optimistic Estimate of Deep Neural Networks, EASIAM minisymposium, Jun. 30, 2024, Macao, China
- Embedding Principle: A hierarchical structure of deep learning loss landscape, ICIAM minisymposium, Aug.24, 2023, Tokyo
- Learning behavior of artificial neural networks,中国神经科学年会 (CNS2022), Nov. 29, 2022. online
- Embedding Principle of Loss Landscape of Deep Neural Networks (Poster), DEEPMATH, Nov.4, 2021, online
- Embedding Principle of Loss Landscape of Deep Neural Networks, 人工 智能的数学基础研讨会, Oct.16, Tong Ji University, Shanghai
- Embedding Principle of Loss Landscape of Deep Neural Networks, 机器 学习与科学计算学术研讨会, Aug.9, 2021, Wuhan
- Impact of initialization on generalization of deep neural networks, 第二届 人工智能的数理基础青年论坛, Dec.05, Beijing
- **Phase diagram for two-layer ReLU neural networks at infinite-width limit.** Nov.01, CSIAM, Changsha
- **Dynamics of Deep Neural Networks A Fourier Analysis Perspective,** Oct.4, 2019, Institute for Advanced Study, Princeton
- Explicitizing an Implicit Bias of the Frequency Principle in Two-layer Neural Networks (Poster), DEEPMATH, Nov.11, 2019, New York City
- A Novel Method on Neuronal Network Reconstruction, 4<sup>th</sup> Statistical Physics and Complex System Conference, Jul.18, 2017, Xi'an, Shanxi
- Spike-triggered Regression on Neuronal Network Reconstruction, SIAM Conference on Dynamical Systems, May.25, 2017, Snowbird, Utah
- Nonuniform Sampling Granger Causality Analysis and Its Application to Neuronal Network Reconstruction, SIAM Conference on Nonlinear Waves and Coherent Structures, Aug.11, 2016, Philadelphia
- Reliability of Granger Causality Analysis on Neuronal Network Reconstruction, SIAM Conference on Life Sciences, Jul.11, 2016, Boston
- Spike-triggered Regression method on Neuronal Network Reconstruction, SIAM Conference on Life Sciences, Jul.11, 2016, Boston, Massachusetts

# **Event Organization**

- Theory and Application of Deep Learning Summer School, Coorganizer, SJTU, Jul.3-7, 2023. Shanghai
- Advances in Deep Learning Theory Minisymposium, Organizer, SIAM Annual Meeting, Jul. 18-20, 2024. Online
- The China Conference on Scientific Machine Learning 2024, Chair of Organization Committee, Aug. 9-11, 2024. Shanghai

# **Teaching Experiences**

#### Courses

- Optimization methods (Graduate, MATH6015, Fall 2021)
- Statistical learning (Undergraduate MA451, Spring 2021)
- Mathematical analysis (Undergraduate MATH1214-01, MATH1608H-06, 2022-2025)

Teaching assistant (with recitations):

- Biological physics (Undergraduate, Spring 2015)
- Methods of mathematical physics (Graduate, Fall 2013)
- Introduction to data science (Undergraduate, Summer 2013)
- Functional analysis (Undergraduate, Spring 2013)
- Complex analysis (Undergraduate, Fall 2012)

## Honors

- Shanghai oversea high level talents program (2020)
- Qiu Suo outstanding young scholar program (2020)