Zhi-Qin John Xu

Tenure-track Associate Professor

Artificial Neural Science

Shanghai Jiao Tong University

Email: xuzhiqin@sjtu.edu.cn

🌐 https://ins.sjtu.edu.cn/people/xuzhiqin

Born: Oct, 1990

Research Interests

Deep learning theory, AI for Science, Computational neuroscience

Research Experience

2019 – Now	🔖	**Tenure-track Associate Professor.** Shanghai Jiao Tong University, Shanghai, China
2016 – 2019	🔖	**Post-Doctoral Associate.** New York University Abu Dhabi; Courant Institute, New York University

Education

2012 – 2016	🔖	Ph.D. in Mathematics, School of Mathematical Sciences, Shanghai Jiao Tong University, China. Thesis title: *Statistical Methods in Neuronal Dynamical Data.*
2008 – 2012	🔖	B.S. in Physics (major) and Mathematics (minor), Zhiyuan College, Shanghai Jiao Tong University, China.

Teaching

Open course in Bilibili	🔖	Please search for the uploader: 许志钦 (followers > 31,800, view counts > 820,000)
Main course	🔖	Data Science (undergraduate), The mathematical foundation of artificial intelligence (graduate), Optimization (graduate), Statistical computing and machine learning (undergraduate)

Community Service

Managing editor	🔖	for *Journal of Machine Learning.*
One of organizers	🔖	for "the second China Conference of Scientific Machine Learning", 2024.
	🔖	for "科技论剑：深度学习的前沿基础", 2024.
	🔖	for "AI for Science Summer School", 2023.
	🔖	for "Theory and Application of Deep Learning Summer School", 2023.
	🔖	for "the first China Conference of Scientific Machine Learning", 2022.
	🔖	for "Summer school on deep learning theory", 2020.
One of general chairs	🔖	for the conference of "Mathematical and Scientific Machine Learning", 2022.

Selected Honors

Shanghai Jiao Tong University Outstanding Bachelor's Thesis Advisor (上海交通大学优异学士学位论文指导教师), 2023

The Second Prize of the National Teaching Achievement Award in Higher Education (Ranked 11th)(高等教育（本科）国家级教学成果奖二等奖（排名第十一）), 2022

the Special Prize of the Shanghai Municipal Teaching Achievement Award in Higher Education (Ranked 8th) (上海市高等教育优秀教学成果特等奖（排名第八）), 2022

Outstanding Class Advisor at Zhiyuan College, Shanghai Jiao Tong University (上海交通大学致远学院优秀班主任), 2022

World artificial intelligence conference youth outstanding paper nomination (世界人工智能大会青年优秀论文提名奖), 2021

Shanghai Overseas High-level Talent Program (上海海外高层次人才引进计划), 2020.

Research Publications

* indicates the corresponding author

# Equal contribution

Representative works

Language Model: We design anchor functions to study language models, especially for understanding the reasoning ability.

- Zhongwang Zhang, Pengxiao Lin, Zhiwei Wang, Yaoyu Zhang, **Zhi-Qin John Xu***, Initialization is Critical to Whether Transformers Fit Composite Functions by reasoning or Memorizing, NeurIPS (2024).

Frequency Principle: We found DNNs bias towards low frequency during training. This F-Principle provides much understanding of the strength and limit of deep learning and inspires a series of algorithms to overcome the difficulty of learning high frequency.

- **Zhi-Qin John Xu***, Yaoyu Zhang, Tao Luo, Yanyang Xiao, Zheng Ma, Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks, arXiv preprint: 1901.06523 (2019), Communications in Computational Physics (2020), 2021 world artificial intelligence conference youth outstanding paper nomination.

  link: `https://www.global-sci.org/intro/article_detail/cicp/18395.html`

- **Zhi-Qin John Xu***, Yaoyu Zhang, and Yanyang Xiao, Training behavior of deep neural network in frequency domain, arXiv preprint: 1807.01251, (2018), 26th International Conference on Neural Information Processing (ICONIP 2019).

  link: `https://link.springer.com/chapter/10.1007/978-3-030-36708-4_22`

- Ziqi Liu, Wei Cai, **Zhi-Qin John Xu***, Multi-scale Deep Neural Network (MscaleDNN) for Solving Poisson-Boltzmann Equation in Complex Domains, arxiv 2007.11207, Communications in Computational Physics (2020).

  link: `https://www.global-sci.org/intro/article_detail/cicp/18402.html`

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

  link: `https://link.springer.com/article/10.1007/s42967-024-00398-7`

Condensation: We found that neurons in the same layer tends to get clustered into more and more groups, where neurons in each group are similar. This condensation shows the DNN progressively increases its complexity during the training from parameter perspective. We have a series of works studying condensation from its dynamics, loss landscape and generalization.

- Zhongwang Zhang, **Zhi-Qin John Xu***, Implicit regularization of dropout. Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2024).

  链接: `https://ieeexplore.ieee.org/document/10412142`

- Yaoyu Zhang*, Zhongwang Zhang, Tao Luo, **Zhi-Qin John Xu***, Embedding Principle of Loss Landscape of Deep Neural Networks. NeurIPS Spotlight (2021).

link: `https://papers.nips.cc/paper/2021/hash/7cc532d783a7461f227a5da8ea80bfe1-Abstract.html`

- Hanxu Zhou, Qixuan Zhou, Tao Luo, Yaoyu Zhang*, **Zhi-Qin John Xu***, Towards Understanding the Condensation of Neural Networks at Initial Training. arxiv 2105.11686, NeurIPS, 2022.

  link: `https://papers.nips.cc/paper_files/paper/2022/hash/0f4d1fc085b7504c140e66bb26ed8842-Abstract-Conference.html`

- Tao Luo#, **Zhi-Qin John Xu**#, Zheng Ma, Yaoyu Zhang*, Phase diagram for two-layer ReLU neural networks at infinite-width limit, arxiv 2007.07497, Journal of Machine Learning Research (2021).

  link: `https://jmlr.csail.mit.edu/papers/v22/20-1123.html`

AI for Science: We focus on solving high-dimensional ODEs, especially for combustion application. We have developed a DNN-assisted sampling method for model reduction and a multiscale sampling method for surrogate model.

- Zhiwei Wang, Yaoyu Zhang, Pengxiao Lin, Enhan Zhao, Weinan E, Tianhan Zhang*, **Zhi-Qin John Xu***, Deep mechanism reduction (DeePMR) method for fuel chemical kinetics, Combustion and Flame, 2024.

  链接: `https://www.sciencedirect.com/science/article/pii/S0010218023006600`

- 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, 2022.

  link: `https://www.sciencedirect.com/science/article/pii/S0010218022003340`

Full list

Preprint

- Zhangchen Zhou, Yaoyu Zhang, **Zhi-Qin John Xu***, A rationale from frequency perspective for grokking in training neural network, arxiv 2405.17479 (2024)

- Zhiwei Wang, Yunji Wang, Zhongwang Zhang, Zhangchen Zhou, Hui Jin, Tianyang Hu, Jiacheng Sun, Zhenguo Li, Yaoyu Zhang, **Zhi-Qin John Xu***, Towards Understanding How Transformer Perform Multi-step Reasoning with Matching Operation, arxiv 2405.15302 (2024)

- (Alphabetic order) Liangkai Hang, Dan Hu, **Zhi-Qin John Xu***, Input gradient annealing neural network for solving low-temperature Fokker-Planck equations, arxiv 2405.00317 (2024)

- Zhongwang Zhang#, Zhiwei Wang#, Junjie Yao, Zhangchen Zhou, Xiaolong Li, Weinan E, **Zhi-Qin John Xu***, Anchor function: a type of benchmark functions for studying language models, arxiv 2401.08309 (2024)

- Xiaolong Li, **Zhi-Qin John Xu***, Zhongwang Zhang, Loss Spike in Training Neural Networks. arxiv 2305.12133 (2023)

- 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.

- Zhangchen Zhou, Hanxu Zhou, Yuqing Li*, **Zhi-Qin John Xu***, Understanding the Initial Condensation of Convolutional Neural Networks, arxiv 2305.09947.

Deep Learning

- Zhongwang Zhang, Pengxiao Lin, Zhiwei Wang, Yaoyu Zhang, **Zhi-Qin John Xu***, Initialization is Critical to Whether Transformers Fit Composite Functions by Inference or Memorizing, NeurIPS (2024).

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

- Zhiwei Wang, Lulu Zhang, Zhongwang Zhang, **Zhi-Qin John Xu***, Loss Jump During Loss Switch in Solving PDEs with Neural Networks, arxiv 2405.03095 (2024), Communications in Computational Physics (2024).

- Tianyi Chen, **Zhi-Qin John Xu***, Efficient and Flexible Method for Reducing Moderate-size Deep Neural Networks with Condensation, arxiv 2405.01041 (2024), Entropy, 2024.

- Zheng-An Chen, Yuqing Li*, Tao Luo, Zhangchen Zhou, **Zhi-Qin John Xu***, Phase Diagram of Initial Condensation for Two-layer Neural Networks, CSIAM Transactions on Applied Mathematics, 2024.

- Zhongwang Zhang, Yuqing Li*, Tao Luo*, **Zhi-Qin John Xu***, Stochastic Modified Equations and Dynamics of Dropout Algorithm. ICLR 2024.

- Xiong-bin Yan, **Zhi-Qin John Xu**, Zheng Ma*, Laplace-fPINNs: Laplace-based fractional physics-informed neural networks for solving forward and inverse problems of subdiffusion, East Asian Journal on Applied Mathematics (2024)

- Zhongwang Zhang, **Zhi-Qin John Xu***, Implicit regularization of dropout. Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (2024).

- Zhiwei Wang, Yaoyu Zhang, Pengxiao Lin, Enhan Zhao, Weinan E, Tianhan Zhang*, **Zhi-Qin John Xu***, Deep mechanism reduction (DeePMR) method for fuel chemical kinetics, Combustion and Flame, 2024.

- Zhiwei Bai, Tao Luo, **Zhi-Qin John Xu***, Yaoyu Zhang*, Embedding Principle in Depth for the Loss Landscape Analysis of Deep Neural Networks. CSIAM Transactions on Applied Mathematics, 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

- Lulu Zhang, Wei Cai, **Zhi-Qin John Xu*** A Correction and Comments on "Multi-Scale Deep Neural Network (MscaleDNN) for Solving Poisson-Boltzmann Equation in Complex Domains. CiCP, 28 (5): 1970–2001, 2020". Communications in Computational Physics, 2023.

- Runze Mao, Minqi Lin, Yan Zhang, Tianhan Zhang, **Zhi-Qin John Xu**, Zhi X Chen*, DeepFlame: A deep learning empowered open-source platform for reacting flow simulations. Computer Physics Communications, 2023.

- Yifan Peng, Dan Hu*, **Zhi-Qin John Xu**, A Non-Gradient Method for Solving Elliptic Partial Differential Equations with Deep Neural Networks. Journal of Computational Physics, 2023.

- (Alphabetic order) Xi-An Li, **Zhi-Qin John Xu**, Lei Zhang*, Subspace Decomposition based DNN algorithm for elliptic type multi-scale PDEs. Journal of Computational Physics, 2023.

- Zhemin Li, **Zhi-Qin John Xu**, Tao Luo, Hongxia Wang*, A regularized deep matrix factorized model of matrix completion for image restoration, IET Image Processing (2022).

- 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, NeurIPS2022.

- 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, 2022.

- 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, 2022.

- Lulu Zhang, **Zhi-Qin John Xu***, Yaoyu Zhang*, Data-informed Deep Optimization. PLoS One (2022).

- Lulu Zhang, Tao Luo, Yaoyu Zhang, **Zhi-Qin John Xu***, Zheng Ma*, MOD-Net: A Machine Learning Approach via Model-Operator-Data Network for Solving PDEs. arxiv 2107.03673, Communications in Computational Physics (CiCP), 2022.

- Hanxu Zhou, Qixuan Zhou, Tao Luo, Yaoyu Zhang*, **Zhi-Qin John Xu***, Towards Understanding the Condensation of Two-layer Neural Networks at Initial Training. arxiv 2105.11686, NeurIPS, 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 Transactions on Applied Mathematics, 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 , arxiv 2105.11675, Mathematical and Scientific Machine Learning (MSML22), 2022.

- (Alphabetic order) Tao Luo, Zheng Ma, **Zhi-Qin John Xu**, Yaoyu Zhang, On the exact computation of linear frequency principle dynamics and its generalization, arxiv 2010.08153, SIAM Journal on Mathematics of Data Science, 2022.

- (Alphabetic order) Tao Luo, Zheng Ma, **Zhi-Qin John Xu**, Yaoyu Zhang, Theory of the frequency principle for general deep neural networks, arXiv preprint, 1906.09235 (2019), CSIAM Transactions on Applied Mathematics, 2022.

- Yaoyu Zhang*, Zhongwang Zhang, Tao Luo, **Zhi-Qin John Xu***, Embedding Principle of Loss Landscape of Deep Neural Networks. NeurIPS Spotlight (2021).

- 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, 2021.

- **Zhi-Qin John Xu***, Hanxu Zhou, Deep frequency principle towards understanding why deeper learning is faster, arxiv 2007.14313, The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21).

- Tao Luo#, **Zhi-Qin John Xu**#, Zheng Ma, Yaoyu Zhang*, Phase diagram for two-layer ReLU neural networks at infinite-width limit, arxiv 2007.07497, Journal of Machine Learning Research (2021).

- (Alphabetic order) Xi-An Li, **Zhi-Qin John Xu***, Lei Zhang, A multi-scale DNN algorithm for nonlinear elliptic equations with multiple scales, arxiv 2009.14597, Communications in Computational Physics (2020).

- Ziqi Liu, Wei Cai, **Zhi-Qin John Xu***, Multi-scale Deep Neural Network (MscaleDNN) for Solving Poisson-Boltzmann Equation in Complex Domains, arxiv 2007.11207, Communications in Computational Physics (2020).

- Yaoyu Zhang, **Zhi-Qin John Xu***, Tao Luo, Zheng Ma, A type of generalization error induced by initialization in deep neural networks. arXiv preprint: 1905.07777 (2019). 1st Mathematical and Scientific Machine Learning Conference (MSML2020).

- **Zhi-Qin John Xu***, Yaoyu Zhang, Tao Luo, Yanyang Xiao, Zheng Ma, Frequency Principle: Fourier Analysis Sheds Light on Deep Neural Networks, arXiv preprint: 1901.06523

(2019), Communications in Computational Physics (2020), 2021 world artificial intelligence conference youth outstanding paper nomination.

- **Zhi-Qin John Xu***, Yaoyu Zhang, and Yanyang Xiao, Training behavior of deep neural network in frequency domain, arXiv preprint: 1807.01251, (2018), 26th International Conference on Neural Information Processing (ICONIP 2019).

Computational Neuroscience

- **Zhi-Qin John Xu**, Xiaowei Gu, Chengyu Li, David Cai, Douglas Zhou, David W. McLaughlin. Neural networks of different species, brain areas and states can be characterized by the probability polling state, European Journal of Neuroscience (2020).

- **Zhi-Qin John Xu**, Douglas Zhou, David Cai, Swift Two-sample Test on High-dimensional Neural Spiking Data, arxiv preprint 1811.12314, (2018).

- **Zhi-Qin John Xu**, Fang Xu, Guoqiang Bi, Douglas Zhou, David Cai, A Cautionary Tale of Entropic Criteria in Assessing the Validity of Maximum Entropy Principle, (2018). (Europhysics Letters)

- **Zhi-Qin John Xu**, Jennifer Crodelle, Douglas Zhou, David Cai, Maximum Entropy Principle Analysis in Network Systems with Short-time Recordings, Physical Review E, DOI: 10.1103/PhysRevE.99.022409, (2019).

- **Zhi-Qin John Xu**, Douglas Zhou, David Cai, Dynamical and Coupling Structure of Pulse-Coupled Networks in Maximum Entropy Analysis, Entropy 2019, 21(1).

- **Zhi-Qin John Xu**, Guoqiang Bi, Douglas Zhou, and David Cai, A dynamical state underlying the second order maximum entropy principle in neuronal networks, Communications in Mathematical Sciences, 15 (2017), pp. 665–692.

- Douglas Zhou, Yanyang Xiao, Yaoyu Zhang, **Zhiqin Xu**, and David Cai, Granger causality network reconstruction of conductance-based integrate-and-fire neuronal systems, PloS one, 9 (2014).

- Douglas Zhou, Yanyang Xiao, Yaoyu Zhang, **Zhiqin Xu**, and David Cai, Causal and structural connectivity of pulse-coupled nonlinear networks, Physical review letters, 111 (2013), p. 054102.