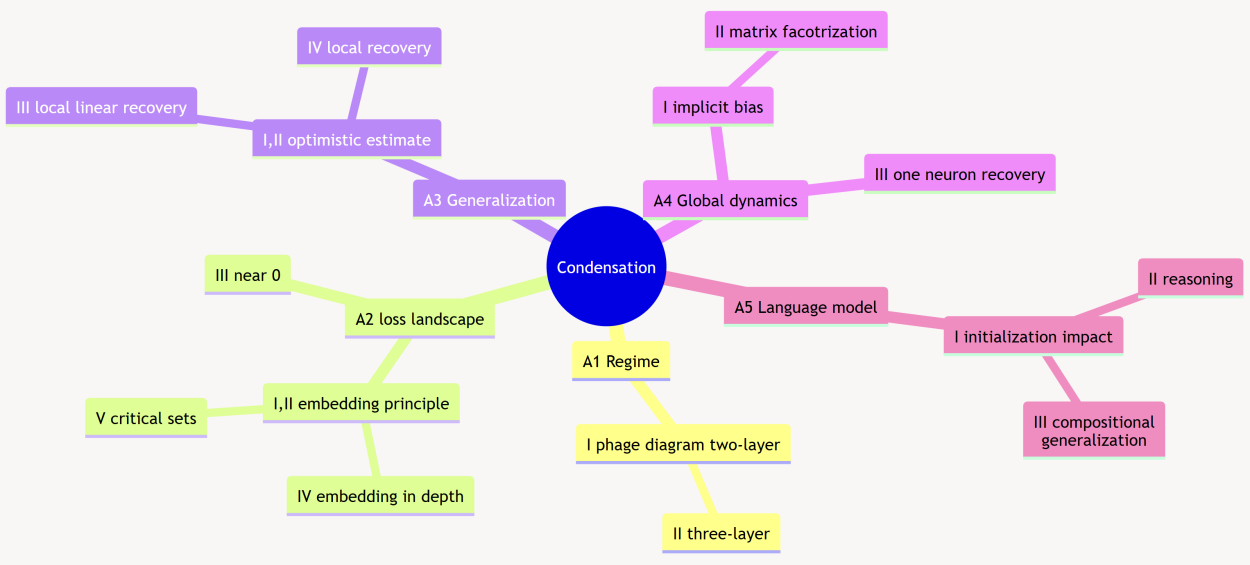
Publications

A. Condensation phenomenon of deep learning

Condensation phenomenon: Neurons in the same layer tends to align with one another during the training



A1. Regime of condensation—phase diagram series

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," NeurIPS 2022.

A2. Loss landscape structure—embedding principle series

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. Hanxu Zhou, Qixuan Zhou, Tao Luo, Yaoyu Zhang, Zhi-Qin John Xu, "Towards Understanding the Condensation of Neural Networks at Initial Training," NeurIPS 2022.
4. 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.
5. Leyang Zhang, Yaoyu Zhang, Tao Luo, "Geometry of Critical Sets and Existence of Saddle Branches for Two-layer Neural Networks," arXiv:2405.17501 (2024).

A3. Generalization advantage—optimistic estimate series

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, Zhiwei Bai, "Local Linear Recovery Guarantee of Deep Neural Networks at Overparameterization," arXiv:2406.18035 (2024).
4. Tao Luo, Leyang Zhang, Yaoyu Zhang, "Structure and Gradient Dynamics Near Global Minima of Two-layer Neural Networks," arXiv:2309.00508 (2023).

A4. Global dynamics and implicit bias

1. 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).
2. Zhiwei Bai, Jiajie Zhao, Yaoyu Zhang, "Connectivity Shapes Implicit Regularization in Matrix Factorization Models for Matrix Completion", NeurIPS 2024.
3. Jiajie Zhao, Zhiwei Bai, Yaoyu Zhang, "Disentangle Sample Size and Initialization Effect on Perfect Generalization for Single-Neuron Target," arXiv:2405.13787 (2024).

A5. Condensation in language models

1. 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.
2. Zhiwei Wang, Yunji Wang, Zhongwang Zhang, Zhangchen Zhou, Hui Jin, Tianyang Hu, Jiacheng Sun, Zhenguo Li, Yaoyu Zhang, Zhi-Qin John Xu, "The Buffer Mechanism for Multi-Step Information Reasoning in Language Models", arXiv:2405.15302 (2024).
3. Zhongwang Zhang, Pengxiao Lin, Zhiwei Wang, Yaoyu Zhang, Zhi-Qin John Xu, "Complexity Control Facilitates Reasoning-Based Compositional Generalization in Transformers", arXiv:2501.08537 (2025).

B. Frequency Principle of deep learning

Frequency Principle: neural networks tend to learn from low to high frequencies during the training.

1. **First Paper**: Zhiqin Xu, Yaoyu Zhang, Yanyang Xiao, "Training Behavior of Deep Neural Network in Frequency Domain," ICONIP, pp. 264-274, 2019. (arXiv:1807.01251, Jul 2018)

2. **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," CiCP 28(5). 1746-1767, 2020.
3. **Initialization effect**: Yaoyu Zhang, Zhi-Qin John Xu, Tao Luo, Zheng Ma, "A Type of Generalization Error Induced by Initialization in Deep Neural Networks," MSML 2020.
4. **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.
5. 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.
6. **Linear Frequency Principle**: 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 4 (4), 1272-1292, 2022.
7. **Minimal decay in frequency domain**: 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.
8. **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.
9. Zhangchen Zhou, Yaoyu Zhang, Zhi-Qin John Xu, "A rationale from frequency perspective for grokking in training neural network," arXiv:2405.17479 (2024).

C. Deep Learning for Science

- Zhiwei Wang, Yaoyu Zhang, Pengxiao Lin, Enhan Zhao, E. Weinan, Tianhan Zhang, 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.
- 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.
- 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 32(2) 299-335 2022.
- Tianhan Zhang, Yaoyu Zhang, Weinan E, Yiguang Ju, "DLODE: A Deep Learning-based ODE Solver for Chemistry Kinetics," AIAA Scitech 2021 Forum, 1139.

D. Computational Neuroscience

- Jing Yan, Yunxuan Feng, Wei Dai, Yaoyu Zhang, "State-dependent Filtering of the Ring Model," arXiv:2408.01817 (2024).
- 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.