

# YIBO JACKY ZHANG

yibo@stanford.edu

<https://yibo.me>

## EDUCATION

---

### Stanford University

*present*

PhD Student in Computer Science

Advisor: Sanmi Koyejo

Department of Computer Science

### University of Illinois at Urbana-Champaign

M.S. in Computer Science

*2022*

Department of Computer Science

### University of Science and Technology of China

B.E. in Computer Science

*2019*

Department of Computer Science and Technology

## RESEARCH INTERESTS

---

I'm interested in solving fundamental AI problems through theoretical research that leads to real-world solutions. Currently, I work on learning in systems composed of interacting components that evolve dynamically and stochastically. See my selected works for more details.

## SELECTED PUBLICATIONS & PREPRINTS

---

- [1] Yibo Jacky Zhang, Sanmi Koyejo. **Exploring Dynamical Stochastic Field Theory for General Autonomous Learning**  
Preprint, 2025.
- [2] Yibo Jacky Zhang, Sanmi Koyejo. **A Framework for Objective-Driven Dynamical Stochastic Fields**  
Preprint, 2025.
- [3] Xiangwen Wang\*, Yibo Jacky Zhang\*, Zhoujie Ding, Katherine Tsai, Sanmi Koyejo. **Aligning Compound AI Systems via System-level DPO**  
*Conference on Neural Information Processing Systems (NeurIPS)*, 2025.

## ALL PUBLICATIONS & PREPRINTS

---

\* indicates equal contribution.

- [1] Yibo Jacky Zhang, Sanmi Koyejo. **Exploring Dynamical Stochastic Field Theory for General Autonomous Learning**  
Preprint, 2025.
- [2] Xiaoyang Wang, Yibo Jacky Zhang, Olawale Elijah Salaudeen, Mingyuan Wu, Hongpeng Guo, Chaoyang He, Klara Nahrstedt, Sanmi Koyejo. **Improving Single-round Active Adaptation: A Prediction Variability Perspective**  
*Transactions on Machine Learning Research (TMLR)*, 2025.
- [3] Yibo Jacky Zhang, Sanmi Koyejo. **A Framework for Objective-Driven Dynamical Stochastic Fields**  
Preprint, 2025.

- [4] Xiangwen Wang\*, Yibo Jacky Zhang\*, Zhoujie Ding, Katherine Tsai, Sanmi Koyejo. **Aligning Compound AI Systems via System-level DPO**  
*Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- [5] Zhenan Shao, Linjian Ma, Yiqing Zhou, Yibo Jacky Zhang, Sanmi Koyejo, Bo Li, Diane M Beck. **Probing Human Visual Robustness with Neurally-Guided Deep Neural Networks**  
Preprint, 2024.
- [6] Boxin Wang, Yibo Jacky Zhang, Yuan Cao, Bo Li, H. Brendan McMahan, Sewoong Oh, Zheng Xu, Manzil Zaheer. **Can Public Large Language Models Help Private Cross-device Federated Learning?**  
*NAACL*, 2024.
- [7] Enyi Jiang\*, Yibo Jacky Zhang\*, Oluwasanmi Koyejo. **Principled Federated Domain Adaptation: Gradient Projection and Auto-Weighting.**  
*International Conference on Learning Representations (ICLR)*, 2024.
- [8] Maohao Shen\*, Bowen Jiang\*, Jacky Y. Zhang\*, Oluwasanmi Koyejo. **Batch Active Learning from the Perspective of Sparse Approximation.**  
*NeurIPS 2022 Workshop on Human in the Loop Learning*, 2022.
- [9] Xiaojun Xu\*, Jacky Y. Zhang\*, Evelyn Ma, Danny Son, Oluwasanmi Koyejo, and Bo Li. **Adversarially Robust Models may not Transfer Better: Sufficient Conditions for Domain Transferability from the View of Regularization.**  
*International Conference on Machine Learning (ICML)*, 2022.
- [10] Kaizhao Liang\*, Jacky Y. Zhang\*, Boxin Wang, Zhuolin Yang, Oluwasanmi Koyejo, and Bo Li. **Uncovering the Connections Between Adversarial Transferability and Knowledge Transferability.**  
*International Conference on Machine Learning (ICML)*, 2021.
- [11] Jacky Y. Zhang, Rajiv Khanna, Anastasios Kyrillidis, and Oluwasanmi Koyejo. **Bayesian Core-sets: Revisiting the Nonconvex Optimization Perspective.**  
*International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2021. (Oral)
- [12] Maohao Shen, Jacky Y. Zhang, Leihao Chen, Weiman Yan, Neel Jani, Brad Sutton, and Oluwasanmi Koyejo. **Labeling Cost-Sensitive Batch Active Learning for Brain Tumor Segmentation.**  
*International Symposium on Biomedical Imaging (ISBI)*, 2021.
- [13] Xiaoyang Wang, Bo Li, Yibo Zhang, Bhavya Kailkhura, and Klara Nahrstedt. **Robusta: Robust AutoML for Feature Selection via Reinforcement Learning.**  
*AAAI 2021 Workshop Towards Robust, Secure and Efficient Machine Learning*.
- [14] Jacky Y. Zhang, Rajiv Khanna, Anastasios Kyrillidis, and Oluwasanmi Koyejo. **Learning Sparse Distributions using Iterative Hard Thresholding.**  
*Conference on Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 2019.
- [15] Yibo Zhang, Chao Qian, and Ke Tang. **Maximizing Monotone DR-submodular Continuous Functions by Derivative-free Optimization.**  
*Preprint: arXiv 1810.06833*, 2018.
- [16] Chao Qian, Yibo Zhang, Ke Tang, and Xin Yao. **On Multiset Selection with Size Constraints.**  
*AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, LA, 2018.