

Fangshuo (Jasper) Liao

Computer Science Department, Rice University
3108 Duncan Hall, 6100 Main Street, Houston, TX 77005
+1 281-745-3702
Fangshuo.Liao@rice.edu

RESEARCH INTEREST	Convergence theory for optimization algorithms in deep learning, neural network pruning, nonconvex optimization.	
ACADEMIC BACKGROUND	Ph.D. Computer Science	2021-now
	George R. Brown School of Engineering, Rice University Advisor: Prof. Anastasios Kyrillidis [website]	GPA: 4.00
	B.S. Computer Science	2016-2020
	George R. Brown School of Engineering, Rice University	GPA: 3.92
	B.A. Mathematics	2016-2020
	Wiess School of Engineering, Rice University	GPA: 3.92
JOURNAL PAPER	Fangshuo Liao and Anastasios Kyrillidis, “On the Convergence of Shallow Neural Network Training with Randomly Masked Neurons”, Transactions on Machine Learning Research (TMLR), 2022. [Link]	
PREPRINT	Cameron R Wolfe*, Jingkan Yang*, Fangshuo Liao* , Arindam Chowdhury, Chen Dun, Artun Bayer, Santiago Segarra, Anastasios Kyrillidis, “GIST: Distributed Training for Large-Scale Graph Convolutional Networks”, arXiv preprint arXiv:2102.10424, 2021. [Link]	
RESEARCH PROJECTS	<i>On the Convergence of Shallow Neural Network Training with Randomly Masked Neurons</i> <ul style="list-style-type: none">– Theoretically characterize the dynamic of the Neural Tangent Kernel of random subnetworks for one-hidden-layer multi-layer perceptron (MLP).– Establish the theoretical framework for the convergence of gradient descent in the local training of random sampled subnetworks. Based on this framework, I give the first convergence proof for Dropout, Multi-sample Dropout, and Independent Subnet Training on over-parameterized one-hidden-layer MLP in regression setting.– My theory demonstrate how the convergence behavior is affected by the size of the subnetworks, the subnetwork sampling method, the number of local gradient descent steps, and the number of the subnetworks. <i>GIST: Distributed Training for Large-Scale Graph Convolutional Networks</i> <ul style="list-style-type: none">– Define and characterize the dynamic of the Graph Independent Subnet Training Kernel (GIST-K). Based on this result, we prove the convergence of Graph Independent Subnet Training for a one-hidden-layer Graph Convolutional Neural Network in semi-supervised regression. <i>LoFT: Finding Lottery Tickets through Filter-wise Training</i> <ul style="list-style-type: none">– Prove the convergence for Lottery Ticket through Filter-wise Training (LoFT) algorithm for a one-hidden-layer CNN.	

*Equal Contribution

- Characterize the convergence path of LoFT, and shows that the weight that LoFT converges to stays near to the weight of gradient descent.

Strong Lottery Ticket Hypothesis with ϵ -Perturbation

- Extend the theoretical guarantee of the Random Subset Sum problem to a setting where each candidate in the summation is allowed and ϵ -perturbation. Show that under the goal of reaching a fixed accuracy, as the perturbation scale ϵ grows, the required number of candidates decreases.
- Collaborate with Zheyang Xiong on applying the theory of ϵ -perturbed random subset sum to neural networks, and show that, when allowed an ϵ -perturbation on the weight entries, the strong Lottery Ticket Hypothesis on deep neural networks requires a smaller over-parameterization.
- Propose the ϵ -bounded projected gradient descent for neural networks to verify that gradient descent is able to find a good set of such ϵ -perturbation.

Learning Compression Matrix for Structured Sparse Recovery with Unrolled LSTM Autoencoder

- Extend the ℓ_1 -AE algorithm for learning the compression matrix for sparse data to a different scheme: we introduced the LSTM unit in the unrolled decoder network to incorporate the sparsity structure in the decoding process.
- Test this method on a variety of structured sparse data, including block-sparse data and tree-sparse data. Demonstrate that the LSTM based algorithm shows better performance in a limited number of scenarios.

**TEACHING
ASSISTANT**

COMP 540 – Statistical Machine Learning

- *Spring 2022, 2021, 2020*
- Designing course project, improving and grading homework, giving multiple recitation lectures, holding office hours.

COMP 440/557 – Artificial Intelligence

- *Fall 2021, 2019*
- Improving and grading homeworks, giving recitation lectures, holding office hours.

MENTORSHIP

Co-advised with Prof. Anastasios Kyrillidis

- **Zheyang (Eddie) Xiong** (Rice University) Aug.2021-now
Strong Lottery Ticket Hypothesis with ϵ -Perturbation.
- **Yuan Gao** (Purdue University) May.2022-Aug.2022
Federated Learning using Graph Independent Subnet Training.
- **Kaichun Luo** (Rice University) May.2020-Aug.2021
Sparse Simplex Projection for Multi-label Classification and Neural Architecture Search.

**INVITED TALKS
& WORKSHOPS**

On the Convergence of Shallow Neural Network Training with Randomly Masked Neurons. Google's Federated Learning and Analytics Workshop. November, 2021.

Provable distributed Learning of Deep Neural Networks using Independent Subnet Training. Intel's MLWiNS Mid Year Workshop. April, 2021.

LoFT: Finding Lottery Tickets through Filter-wise Training. Intel's MLWiNS Annual Workshop. October, 2021.