CHAOYUE LIU

(+1) 517 348 5311 ⋄ chl212@ucsd.edu

BRIEF INTRODUCTION

I am currently a post-doctoral researcher at Halicioglu Data Science Institute in UC San Diego. My research focuses on theoretical and mathematical foundations of deep learning optimization and algorithmic acceleration. I am enthusiastic in theoretically understanding the dynamics of neural network training, and the mechanisms behind. My recent research effort has been devoted to finding fundamental properties of neural networks and algorithms that improve training of neural networks.

EDUCATION

The Ohio State University *August 2015 - May 2021*
Degree: Ph.D., Computer Science
Advisor: Prof. Mikhail Belkin

Michigan State University *August 2012 - May 2015*
Ph.D. student in Physics

Tsinghua University *August 2010 - June 2012*
Degree: Master of Science, Physics

Tsinghua University *August 2006 - June 2010*
Degree: Bachelor of Science, Physics

WORKING EXPERIENCES

Halicioglu Data Science Institute, UC San Diego *Sep 2022 - Present*
Postdoctoral Researcher

Facebook Inc. *May 2021 - Sep 2022*
Research Scientist (Machine Learning)

Facebook Inc. *May 2019 - August 2019*
Machine Learning Intern (Ph.D.)

PUBLICATION

Chaoyue Liu, Like Hui. "ReLU soothes the NTK condition number and accelerates optimization for wide neural networks". Arxiv:2305.08813. (under review)

Chaoyue Liu, Amirhesam Abedsoltan, Mikahil Belkin. "On Emergence of Clean-Priority Learning in Early Stopped Neural Networks". arXiv:2306.02533. (under review)

Libin Zhu, **Chaoyue Liu**, Adityanarayanan Radhakrishnan, Mikhail Belkin. "Catapults in SGD: spikes in the training loss and their impact on generalization through feature learning". arXiv:2306.04815 (under review)

Libin Zhu, **Chaoyue Liu**, and Mikhail Belkin. "Quadratic models for understanding neural network dynamics".*International Conference on Learning Representations (ICLR)*, 2024.

Chaoyue Liu, Dmitriy Drusvyatskiy, Mikhail Belkin, Damek Davis, Yi-An Ma. "Aiming towards the minimizers: fast convergence of SGD for overparametrized problems". *Neural Information Processing Systems (NeurIPS)*, 2023.

Chaoyue Liu, Dmitriy Drusvyatskiy, Mikhail Belkin, Damek Davis, Yi-An Ma. "SGD batch saturation for training wide neural networks". *Neural Information Processing Systems (NeurIPS) Optimization Workshop*, 2023.

Chaoyue Liu, Libin Zhu, and Mikhail Belkin. "Transition to Linearity of Wide Neural Networks is an Emerging Property of Assembling Weak Models". *International Conference on Learning Representations (ICLR)*, 2022. (**spotlight paper**, 5.2% of all submissions)

Chaoyue Liu, Libin Zhu, and Mikhail Belkin. "Loss landscape and optimization framework for systems of non-linear equations: the lessons of deep learning". *Applied and Computational Harmonic Analysis*, 2022.

Libin Zhu, **Chaoyue Liu**, and Mikhail Belkin. "Transition to Linearity of General Neural Networks with Directed Acyclic Graph Architecture". *Neural Information Processing Systems (NeurIPS)*, 2022.

Kun Jin, **Chaoyue Liu**, Cathy Xia. "Two-sided Wasserstein Procrustes Analysis", *International Joint Conference on Artificial Intelligence (IJCAI)*, 2021.

Chaoyue Liu. "Understanding and Accelerating the Optimization of Modern Machine Learning". Ph.D. dissertation, The Ohio State University. 2021.

Chaoyue Liu, Libin Zhu, and Mikhail Belkin. "On the Linearity of Large Non-linear Models: When and Why Tangent Kernel Is Constant". *Neural Information Processing Systems (NeurIPS)*, 2020. (**spotlight paper**, 3.0% of all submissions)

Chaoyue Liu, Mikhail Belkin. "Accelerating SGD with Momentum for Over-parameterized learning". *International Conference on Learning Representations (ICLR)*, 2020. (**spotlight paper**, 4.2% of all submissions)

Kun Jin, **Chaoyue Liu**, Cathy Xia. "OTDA: a Unsupervised Optimal Transport framework with Discriminant Analysis for Keystroke Inference", *IEEE Conference on Communications and Network Security (CNS)*, 2020.

Chaoyue Liu, Mikhail Belkin. "Parametrized Accelerated Methods Free of Condition Number". arXiv: 1802.10235.

Chaoyue Liu, Mikhail Belkin. "Clustering with Bregman Divergences: an Asymptotic Analysis". *Neural Information Processing Systems (NIPS)*, 2016.

TALKS

Transition to Linearity and Optimization Theories for Wide Neural Networks, Control and Pizza (Co-PI) seminar, UC San Diego, Nov 2023.

Transition to Linearity as An Emerging Property of Assembly Model, ICLR conference spotlight talk, April 2022.

Transition to Linearity of Wide Neural Networks, Math Machine Learning Seminar, Max Planck Institute & UCLA, April 2022.

Large Non-linear Models: Transition to Linearity & An Optimization Theory, NSF-Simons Journal Club, January 2021.

Transition to Linearity of Wide Neural Networks. NeurIPS conference spotlight talk, December 2020.

Accelerating SGD with Momentum for Over-parameterized Learning, ICLR conference spotlight talk, April 2020.

Clustering with Bregman Divergence, Artificial Intelligence Seminar, Ohio State University, November 2016.

POSITION OF RESPONSIBILITY

Publication Reviewer

NeurIPS: 2018, 2019, 2020, 2021, 2022, 2023
ICLR: 2021, 2022, 2023, 2024
ICML: 2019, 2020,2021, 2022, 2023
AAAI: 2020, 2021, 2022
UAI: 2019,2020
JASA: 2021
ALT: 2021 (External reviewer)

Grant Reviewer

Swiss National Science Foundation (SNSF), 2022.

Teaching Assistant of Department of Physics & Astronomy at MSU

· PHY 820 (Classical Mechanics),	Fall 2012
· PHY 851 (Quantum Mechanics I),	Fall 2012
· PHY 231 (Introductory Physics I),	Fall 2014
· PHY 803 (Elementary Particle Physics),	Spring 2015

Teaching Assistant of Department of Computer Science & Engineering at OSU

· CSE 2321 (Discrete Structure),	Fall 2015
· CSE 2111 (Modeling and Problem Solving with Spreadsheets and Databases),	Spring 2016
· CSE 5523 (Machine Learning),	Spring 2017, Spring 2018, Spring 2019
· CSE 3421 (Intro. to Computer Architecture),	Autumn 2018

HONORS AND AWARDS

NIPS travel award	2016
Guanghua Scholarship for Excellent students	2011