Changyu Gao

Email:changyu.gao@wisc.edu
 cyugao in/changyu-gao

EDUCATION

University of Wisconsin–Madison — Madison, WI
Ph.D. Candidate in Industrial Engineering, Optimization Track — *Sep 2019 – Aug 2025 (Expected)*

University of Wisconsin–Madison — Madison, WI
M.S., Computer Science — *Feb 2021 – June 2024*

University of Science and Technology of China — Hefei, China
B.S., Mathematics and Applied Mathematics (Outstanding Graduate) — *Aug 2015 – June 2019*

RESEARCH EXPERIENCE

Advisor: **Prof. Stephen Wright** (All projects below are in collaboration with Prof. Wright)
Summary: design and analysis of optimization algorithms for machine learning applications.

Optimal Rates for Robust Stochastic Convex Optimization
Collaborator: **Andrew Lowy, Xingyu Zhou**

- Developed novel stochastic convex optimization algorithms with robustness guarantees.
- Achieved first optimal-rate results for robust stochastic convex optimization.
- Significantly improved sample complexity and relaxed strict requirements of existing algorithms, broadening their applicability.

Private Federated Learning
Collaborator: **Andrew Lowy, Xingyu Zhou**

- Designed novel federated learning algorithm with privacy guarantees for heterogeneous data
- Achieved optimal population excess risk bounds, surpassing previous state-of-the-art methods.
- Significantly improved the communication and gradient complexity over SOTA algorithms.

Differentially Private Optimization

- Innovated differentially private algorithms for finding approximate second-order stationary points
- Implemented adaptive line search and mini-batching strategies to enhance practical performance.
- Developed PyTorch implementation demonstrating empirical effectiveness through experiments.

Optimization Methods for Probabilistic Soft Logic (PSL)
Collaborator: **Charles Dickens, Connor Pryor, Lise Getoor**

- Implemented and tested HOGWILD! and Frank-Wolfe methods for PSL framework using Java.
- Executed inference experiments on real-world datasets, validating the practicality of these methods.

- Proved theoretical guarantees for the proposed bilevel formulation of PSL

Parameter Learning with Derivative-Free Optimization (DFO) Methods
- Implemented the parameter learning procedure for Lorenz96 model using DFO methods in Python
- Performed optimization with inexact function values using interpolation and Bayesian methods

WORK EXPERIENCE

Research Scientist Intern, Meta   Menlo Park, CA
Team: **Meta AI Research (FAIR) – Reality Labs**   *Sep 2022 – Jan 2023*
- **Adaptive Training for Transformer-based Models**:
  * Developed adaptive training algorithms and engineered gradient statistics analysis framework
  * Achieved baseline performance with reduced computation, improving training efficiency
  * Contributed to Meta's FairScale library, resolving critical gradient accumulation issues

Applied Scientist Intern, Amazon   Seattle, WA
Team: **Delivery Experience (DEX) – AI**   *May 2021 – Aug 2021*
- **Mining Inconsistency Issues using Semantic Search Model**:
  * Developed semantic search system for detecting customer experience inconsistencies
  * Enhanced search accuracy through fine-tuning approaches in TensorFlow
  * Identified and escalated critical inconsistency issues to the corresponding teams

PROGRAMMING SKILLS

Languages: Proficient: Python. Familiar: SQL, R, C, C++, Java

Frameworks: Pytorch, Tensorflow, JAX, Pandas, Numpy, Scipy

PUBLICATIONS

Changyu Gao, Andrew Lowy, Stephen J. Wright, Xingyu Zhou. Private Heterogeneous Federated Learning Without a Trusted Server Revisited: Error-Optimal and Communication-Efficient Algorithms for Convex Losses, Proceedings of the 41st International Conference on Machine Learning (**ICML 2024**), Vienna, Austria. PMLR 235, 2024. [**Poster Award**, Midwest Machine Learning Symposium 2024]

Changyu Gao and Stephen J. Wright. Differentially Private Optimization for Smooth Nonconvex ERM, arXiv preprint arXiv:2302.04972 (2023). [TPDP 2023 Poster]

Changyu Gao, Andrew Lowy, Stephen J. Wright, Xingyu Zhou. Optimal Rates for Robust Stochastic Convex Optimization, submitted.

Charles Andrew Dickens, **Changyu Gao**, Connor Pryor, Stephen J. Wright, Lise Getoor. Convex and Bilevel Optimization for Neuro-Symbolic Inference and Learning, Proceedings of the 41st International Conference on Machine Learning (**ICML 2024**), Vienna, Austria. PMLR 235, 2024.

Charles Andrew Dickens, Connor Pryor, **Changyu Gao**, Alon Albalak, Eriq Augustine, William Wang, Stephen J. Wright, and Lise Getoor. A mathematical framework, a taxonomy of modeling paradigms, and a suite of learning techniques for neural-symbolic systems, arXiv preprint arXiv:2407.09693 (2024).

Hobbies and interests: music, guitar, hiking, tech, society, language, reading, podcasts