



# Changyu Gao

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

## EDUCATION

---

**University of Wisconsin–Madison**

*Ph.D. Student in Industrial Engineering, Optimization Track*

Madison, WI

*Sep 2019 – Aug 2025 (Expected)*

**University of Wisconsin–Madison**

*M.S., Computer Science*

Madison, WI

*Feb 2021 – June 2024*

**University of Science and Technology of China**

*B.S., Mathematics and Applied Mathematics (Outstanding Graduate)*

Hefei, China

*Aug 2015 – June 2019*

## RESEARCH EXPERIENCE

---

*Advisor:* **Prof. Stephen Wright**

Summary: design algorithms that solves mathematical optimization problems with theoretical guarantees.

**Private Federated Learning**

*Collaborator:* **Dr. Andrew Lowy, Dr. Xingyu Zhou**

- Developed a novel federated learning algorithm with inter-silo record-level DP guarantees, combining the phased localization technique and a private federated learning version of SGD.
- Our algorithm achieves the *optimal* population excess risk bound for heterogeneous data distributions; previous algorithms only achieve suboptimal bounds.
- Significantly improved the communication and gradient complexity over SOTA algorithms.

**Differentially Private Optimization for Smooth Nonconvex ERM**

- Designed differentially private optimization algorithms for finding an approximate second-order stationary point with convergence guarantees.
- Proposed line search, mini-batching, and a two-phase strategy to improve the speed and practicality of the algorithms.
- Implemented the algorithm using Pytorch and demonstrated its performance through experiments.

**Optimization Methods for Probabilistic Soft Logic (PSL)**

- 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 uncertainty function values using soft interpolation and Gaussian Process methods.

## WORK EXPERIENCE

---

### Research Scientist Intern, Meta

Menlo Park, CA

*Team: Meta AI (FAIR) – Reality Labs*, supervised by **Dr. Min Xu**

*Sep 2022 – Jan 2023*

- **Adaptive Training for Transformer-based Models:**

- \* Contributed to the open source project fairscale. Fixed gradient accumulation bugs.
- \* Implemented methods to collect and analyze the gradient statistics during training of the Translation task using the Transformer model. Based on the gradient statistics, designed and implemented an adaptive training method to improve the training efficiency.
- \* Our algorithm uses the gradient statistics to dynamically adjust the gradient accumulation steps and the learning rate, achieving the baseline with better efficiency.

### Applied Scientist Intern, Amazon

Seattle, WA

*Team: Delivery Experience (DEX) – AI*

*May 2021 – Aug 2021*

- **Mining Inconsistency Issues using Semantic Search Model:**

- \* Applied the semantic search model to the customer contact data, facilitating inconsistency detection.
- \* Collected and refined the queries for semantic search; oversaw the data annotation process.
- \* Implemented two fine-tuning schemes of the encoder used in the semantic search model in Tensorflow and thus improved the search model accuracy.
- \* Important inconsistency issues discovered were escalated to the corresponding issue owners.

## PROGRAMMING SKILLS

---

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

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

## PUBLICATIONS

---

**Changyu Gao**, Andrew Lowy, Stephen Wright, Xingyu Zhou. Private Heterogeneous Federated Learning Without a Trusted Server Revisited: Error-Optimal and Communication-Efficient Algorithms for Convex Losses, **ICML 2024**.

**Changyu Gao** and Stephen J Wright. Differentially Private Optimization for Smooth Nonconvex ERM, arXiv preprint. TPDP 2023 Poster.

Charles Andrew Dickens, **Changyu Gao**, Connor Pryor, Stephen Wright, Lise Getoor. Convex and Bilevel Optimization for Neuro-Symbolic Inference and Learning, **ICML 2024**.

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