Kaiwen Wu

Philadelphia, PA, USA https://kayween.github.io kaiwenwu@seas.upenn.edu

Research

I am interested in both theory and applications of machine learning. My PhD research focuses on various areas in probabilistic machine learning including Gaussian processes, variational inference, Bayesian optimization, and active learning. Before starting my PhD, I worked on generative modeling, convex optimization, and adversarial robustness of deep learning models.

Education

University of Pennsylvania — Philadelphia, PA
Doctor of Philosophy in Computer and Information Science — 2021–Present
Advisor: Jacob R. Gardner

University of Waterloo — Waterloo, ON
Master of Mathematics in Computer Science (Thesis Option) — 2018–2020
Advisor: Yaoliang Yu

Nanjing University — Nanjing, China
Bachelor of Science in Computer Science — 2014–2018

Work Experience

Meta — New York, NY
Research Scientist Intern — Jun–Oct 2025
- Worked on Bayesian optimization for automated machine learning (AutoML).

Meta — Redmond, WA
Research Scientist Intern — Aug–Dec 2024
- Worked on active learning for VR/AR hardware design.

Borealis AI — Waterloo, ON
Research Intern — Jan–May 2019
- Worked on generative adversarial networks (GANs) for robust mean estimation on noisy data.

Open Source Contributions

Ax (`https://github.com/facebook/ax`)
Adaptive experimentation platform

BoTorch (`https://github.com/pytorch/botorch`)
Python package for Bayesian optimization in PyTorch

GPyTorch (`https://github.com/cornellius-gp/gpytorch`)
Python package for Gaussian processes in PyTorch

LinearOperator (`https://github.com/cornellius-gp/linear_operator`)
Linear algebra package for implementing structured matrices in PyTorch

Publications

See also at my Google Scholar 🎓.

Preprints

[P1] Haydn Jones, Yimeng Zeng, Alden Rose, Li S. Yifei, Yining Huang, **Kaiwen Wu**, Jiaming Liang, Maggie Ziyu Huan, Yoseph Barash, Cesar de la Fuente-Nunez, Osbert Bastani, Zachary Ives, Mark Yatskar, and Jacob R. Gardner
Self-Driving Datasets: From 20 Million Papers to Nuanced Biomedical Knowledge at Scale
arXiv preprint arXiv:2605.07022. 2026.

[P2] **Kaiwen Wu** and Jacob R Gardner
Knowledge Gradient for Preference Learning
arXiv preprint arXiv:2601.22335. 2026.

[P3] **Kaiwen Wu**, Craig Sanders, Benjamin Letham, and Phillip Guan
Mixed Likelihood Variational Gaussian Processes
arXiv preprint arXiv:2503.04138. 2025.

Refereed Conference Proceedings

[C1] Jonathan Wenger, **Kaiwen Wu**, Philipp Hennig, Jacob R. Gardner, Geoff Pleiss, and John P. Cunningham
Computation-Aware Gaussian Processes: Model Selection And Linear-Time Inference
Advances in Neural Information Processing Systems (NeurIPS). 2024.

[C2] **Kaiwen Wu** and Jacob R. Gardner
Understanding Stochastic Natural Gradient Variational Inference
Proceedings of the 41st International Conference on Machine Learning (ICML). 2024.

[C3] **Kaiwen Wu**, Jonathan Wenger, Haydn T Jones, Geoff Pleiss, and Jacob Gardner
Large-Scale Gaussian Processes via Alternating Projection
Proceedings of the 27th International Conference on Artificial Intelligence and Statistics (AISTATS). 2024.

[C4] Kyurae Kim, Jisu Oh, **Kaiwen Wu**, Yian Ma, and Jacob R. Gardner
On the Convergence of Black-Box Variational Inference
Advances in Neural Information Processing Systems (NeurIPS). 2023.

[C5] Kyurae Kim, **Kaiwen Wu**, Jisu Oh, and Jacob R. Gardner
Practical and Matching Gradient Variance Bounds for Black-Box Variational Bayesian Inference
Proceedings of the 40th International Conference on Machine Learning (ICML). 2023.

[C6] Natalie Maus, **Kaiwen Wu**, David Eriksson, and Jacob Gardner
Discovering Many Diverse Solutions with Bayesian Optimization
Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS). 2023.

[C7] **Kaiwen Wu**, Kyurae Kim, Roman Garnett, and Jacob R. Gardner
The Behavior and Convergence of Local Bayesian Optimization
Advances in Neural Information Processing Systems (NeurIPS). 2023.

[C8] Xinran Zhu, **Kaiwen Wu**, Natalie Maus, Jacob Gardner, and David Bindel
Variational Gaussian Processes with Decoupled Conditionals
Advances in Neural Information Processing Systems (NeurIPS). 2023.

[C9] Quan Nguyen, **Kaiwen Wu**, Jacob Gardner, and Roman Garnett
Local Bayesian optimization via maximizing probability of descent
Advances in Neural Information Processing Systems (NeurIPS). 2022.

[C10] **Kaiwen Wu**, Gavin Weiguang Ding, Ruitong Huang, and Yaoliang Yu
On Minimax Optimality of GANs for Robust Mean Estimation
Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS). 2020.

[C11] **Kaiwen Wu**, Allen Wang, and Yaoliang Yu
Stronger and Faster Wasserstein Adversarial Attacks
Proceedings of the 37th International Conference on Machine Learning (ICML). 2020.

Workshop Papers

[W1] **Kaiwen Wu** and Jacob R. Gardner
A Fast, Robust Elliptical Slice Sampling Method for Linearly Truncated Multivariate Normal Distributions
NeurIPS Workshop on Bayesian Decision-making and Uncertainty. 2024.

[W2] Guojun Zhang, **Kaiwen Wu**, Pascal Poupart, and Yaoliang Yu
Newton-type Methods for Minimax Optimization
ICML Workshop on Beyond First-order Methods in Machine Learning Systems. 2021.

Teaching Experience

University of Pennsylvania

- Teaching Assistant, CIS 520 Machine Learning — Fall 2023
- Teaching Assistant, ESE 204 Decision Models — Fall 2022

University of Waterloo

- Teaching Assistant, CS 370 Numerical Computation — Spring 2020
- Teaching Assistant, CS 486/686 Introduction to Artificial Intelligence — Spring 2019
- Teaching Assistant, CS 480/680 Introduction to Machine Learning — Fall 2018

Honors & Awards

- Top Reviewer, Conference on Neural Information Processing Systems — 2024
- Graduate Fellowship, University of Pennsylvania — 2021
- Vector Scholarship in Artificial Intelligence, Vector Institute — 2018
- Entrance Scholarship, University of Waterloo — 2018
- Jimin Liu Scholarship, Nanjing University — 2017

Academic Service

Conference Reviews

- Program Committee Member, AAAI Conference on Artificial Intelligence (2021, 2026)
- Reviewer, Conference on Neural Information Processing Systems (2023–2025)
- Reviewer, International Conference on Artificial Intelligence and Statistics (2021, 2024)
- Reviewer, International Conference on Learning Representations (2024–2026)
- Reviewer, International Conference on Machine Learning (2023–2026)

Journal Reviews

- Reviewer, Transactions on Machine Learning Research (2025)
- Reviewer, Journal of Machine Learning Research (2025)