

Erchi Wang

📍 La Jolla, CA ✉ erw011@ucsd.edu ☎ 217-693-8227 🔗 erchiw.github.io in erchi-wang 🌐 erchiw

Summary

I am a Ph.D. student in Data Science at UC San Diego, specializing in privacy-preserving machine learning and responsible AI. Experienced in developing data-adaptive and practical differentially private algorithms with provable guarantees[1, 2]. Currently, I am also exploring the use of differential privacy techniques to address safety risks in foundation models, such as protecting privacy in Retrieval-Augmented Generation (RAG) [3], and building benchmarks to evaluate frontier LLMs' capability to verify differential privacy guarantees.

Education

University of California, San Diego, San Diego, US
Ph.D. in Data Science, GPA: 3.90/4.0

Jul. 2024 – Jul. 2027 (Est.)
Advised by Prof. Yu-Xiang Wang

University of California, Santa Barbara, Santa Barbara, US
M.A. in Statistics, GPA: 3.91/4.0

Aug. 2021 – Jul. 2024
Advised by Prof. Yu-Xiang Wang

University of Illinois at Urbana-Champaign, Urbana, US
M.S. in Statistics, GPA: 3.82/4.0

Aug. 2018 – Dec. 2020

Ocean University of China, Qingdao, China
B.S. in Applied Math and Biological Science, GPA: 3.76/4.0

Aug. 2013 – Jul. 2018

Publications & Manuscripts (* denotes equal contribution)

- [1] **Erchi Wang**, Yuqing Zhu, Yu-Xiang Wang. Adapting to Linear Separable Subsets with Large Margin in Differentially Private Learning. *ICML-2025*. (selected as **Oral Presentations** at *TPDP 2025* and *Crypto-PPML 2025*.) [Arxiv link](#)
- [2] **Erchi Wang***, Yingyu Lin*, Yi-An Ma, Yu-Xiang Wang. Purifying Approximate Differential Privacy with Randomized Post-processing. *NeurIPS 2025, Spotlight*. (selected as **Oral Presentation** at *TPDP 2025*.) [Arxiv link](#)
- [3] Ruihan Wu*, **Erchi Wang***, Yu-Xiang Wang. Beyond Per-Question Privacy: Multi-Query Differential Privacy for RAG Systems. *NeurIPS 2025 Workshop: Reliable ML from Unreliable Data* [Arxiv link](#)
- [4] Erchi Wang, Arinbjörn Kolbeinsson, Luca Foschini, Yu-Xiang Wang. Revisiting Differentially Private XG-boost: Are Random Decision Trees Really Better than Greedy Ones? *In Submission*.

Selected Projects

(On going) DP Checker: Evaluation of frontier LLMs' Ability in Verifying Differential Privacy

- Designing a comprehensive benchmark covering a wide range of differentially private algorithms and their associated privacy-guarantee proofs.
- Evaluating frontier LLMs on this benchmark (including GPT-5, Gemini, and many others) and analyzing their failure modes, laying the groundwork for the design of a differentially private agent.

Multi-Query Retrieval-Augmented Generation with Differential Privacy Guarantee

- Designed a novel DP-RAG framework enabling multi-query retrieval-augmented generation with up to $100\times$ reduction in privacy budget while improving generation utility on sensitive tasks.
- Demonstrated strong performance across four QA benchmarks and three LLMs (OPT-1.3B, Pythia-1.4B, Mistral-7B), outperforming non-private LLM baselines without RAG on privacy-sensitive evaluation.
- Fine-tuned models using DP-SGD and generated baseline results using private evolution, showcasing the practicality of DP-RAG in low-budget privacy settings.
- Evaluated empirical privacy leakage via jailbreak and membership inference attacks, demonstrating that our DP-RAG methods provide substantially stronger privacy guarantees.

Differential Private Adaptive Margin Learning

- Designed a computationally efficient differentially private algorithm for classification problems. Implemented advanced private hyperparameter tuning methods and refined the analysis of DP-SGD, allowing the algorithm to adapt to large data margins without requiring prior knowledge of the margin value. Theoretically, the proposed method guarantees utility adaptation to both separable and non-separable cases.

Converting Approximate DP Mechanisms into Pure DP Mechanisms

- Developed a black-box converter from approximate to pure differential privacy and leveraged it to design efficient pure DP optimization and data-dependent algorithms that were previously difficult to construct.

Differential Private Greedy XGBoost on Tabular Data

- Designed and implemented an enhanced differentially private greedy XGBoost algorithm, leveraging modern privacy accounting techniques, including Rényi Differential Privacy-based composition and bounded range analysis for the exponential mechanism. ([GitHub Repo](#))
- Conducted extensive empirical studies on 18 UCI tabular datasets, achieving state-of-the-art performance with DP-XGBoost by reducing the number of trees by 30% to 50%, thereby enhancing model explainability and accelerating inference speed.

Programming Skills

Languages: Python, R, Bash, Git

Libraries & Frameworks: Pytorch, Huggingface, vLLM. Pandas, SciPy, Scikit-learn, Opacus, AutoDP

Professional Service

Reviewers for NeurIPS (2024, 2025), ICLR (2025, 2026), AISTATS (2025, 2026), ICML (2025), NeurIPS Reliable ML Workshop (2025), EurIPS PPML Workshop (2025), TMLR