Puqian Wang

Email : e.puqian.wang@gmail.com
Personal Website : http://puqianwang.github.io
Google Scholar : Puqian-GoogleScholar

Education

- **University of Wisconsin - Madison** — Madison, WI, US
  Ph.D. in Computer Sciences — *Sep 2021 - 2027(Expected)*
  Advisor: Prof. Jelena Diakonikolas
- **University of Wisconsin - Madison** — Madison, WI, US
  M.Sc. in Computer Sciences — *Sep 2021 - May 2024*
- **Shandong University** — Jinan, Shandong, China
  B.S. in Mathematics, Major: Information and Computing Science, — *Sep 2017 - June 2021*

Research Interests

- Optimization theory and applications
- Statistical and computational machine learning theory

Honors

- Honorable Mention of Ivanisevic Award for Computer Sciences Dissertators, UW Madison, 2026
- IFDS Research Assistant Fellowship, UW Madison, Department of Computer Science, 2024, 2025
- Summer Research Assistant Fellowship, UW Madison, Department of Computer Science, Summer 2022
- The 2nd Prize of Shandong Province in Chinese Mathematical Modeling Competition, 10/2019
- The 1st, 2nd Prize in Academic Performance Scholarship of School of Mathematics, 09/2018, 09/2019
- The 1st Prize in Physics Competition for College Students in Shandong Province, 12/2018
- The 2nd Prize of Nation (1st Prize of Shandong Province) in Chinese Mathematics Competition, 11/2018

Internship

- **Research Intern – Microsoft Research Cambridge** — Cambridgeshire, UK
  Mentor: Chao Ma — *July 2025 - September 2025*
  Worked on adaptively rotated optimization methods for large language model pre-training. The project showed that rotation-based preconditioning can improve AdamW and other widely used optimizers by up to 1.3×. My theoretical contribution was to connect this phenomenon to mirror descent in ℓ_∞ geometry, proving that a related method can converge faster than gradient descent.

Selected Publications

1. **ARO: A New Lens On Matrix Optimization For Large Models**

   Wenbo Gong, Javier Zazo, Qijun Luo, **Puqian Wang**, James Hensman, Chao Ma. *In submission, 2026*

   Introduced a rotation-preconditioned perspective on optimizer design for large-model training. The key insight is that widely used adaptive optimizers are not rotation invariant, and choosing appropriate coordinate systems can improve training performance by up to 1.3×.

2. **Robustly Learning Monotone Single-Index Models**

   Puqian Wang*, Nikos Zarifis*, Ilias Diakonikolas, Jelena Diakonikolas. In *NeurIPS, 2025*

   Developed the first polynomial-time robust algorithm for learning monotone single-index models with arbitrary unknown monotone link function under Gaussian covariates and strong adversarial label noise. The main technical idea is a gradient-like vector field that provably aligns with the hidden signal without explicitly estimating the unknown link function.

3. **Robustly Learning Monotone Generalized Linear Models via Data Augmentation**

   Nikos Zarifis*, **Puqian Wang***, Ilias Diakonikolas, Jelena Diakonikolas. In *COLT, 2025*

   Gave the first polynomial-time algorithm for robustly learning monotone Gaussian generalized linear models with an arbitrary known monotone link under adversarial label noise. The key technique is to smooth the gradient of a nonconvex loss via data augmentation, revealing a reliable direction toward the hidden signal despite adversarial noise; analytically, this relies on bounding the Fourier tail of monotone functions by the norm of the smoothed gradient.

4. **Robustly Learning a Single Neuron Via Sharpness**

   Puqian Wang*, Nikos Zarifis*, Ilias Diakonikolas, Jelena Diakonikolas. In *ICML 2023, Oral Presentation*

   Studied the problem of learning a single neuron with respect to the L_2^2-loss in the presence of adversarial label noise. We give an efficient algorithm that, for a broad family of activations including ReLUs, approximates the optimal L_2^2-error within a constant factor. Notably, our algorithm succeeds under much milder distributional assumptions compared to prior work. The key ingredient enabling our results is a novel connection to local error bounds from optimization theory.

5. **Potential Function-based Framework for Minimizing Gradients in Convex and Min-Max Optimization**

   (α-β) Jelena Diakonikolas, **Puqian Wang**. In *SIAM Journal on Optimization, 2022*

   Introduced a novel potential function-based framework to study the convergence of standard methods for making the gradients small in smooth convex optimization and convex-concave min-max optimization.

Publications

1. **ARO: A New Lens On Matrix Optimization For Large Models**

   Wenbo Gong, Javier Zazo, Qijun Luo, **Puqian Wang**, James Hensman, Chao Ma. *In submission, 2026*

2. **Robustly Learning Monotone Single-Index Models**

   Puqian Wang*, Nikos Zarifis*, Ilias Diakonikolas, Jelena Diakonikolas. In *NeurIPS, 2025*

3. **Robustly Learning Monotone Generalized Linear Models via Data Augmentation**

   Nikos Zarifis*, **Puqian Wang***, Ilias Diakonikolas, Jelena Diakonikolas. In *COLT 2025*

4. **Sample and Computationally Efficient Robust Learning of Gaussian Single-Index Models**

   Puqian Wang, Nikos Zarifis, Ilias Diakonikolas, Jelena Diakonikolas. In *NeurIPS 2024*

5. **Robustly Learning Single-Index Models via Alignment Sharpness**

   Nikos Zarifis*, **Puqian Wang***, Ilias Diakonikolas, Jelena Diakonikolas. In *ICML 2024*

6. **Near-Optimal Bounds for Learning Gaussian Halfspaces with Random Classification Noise**

   (α-β) Ilias Diakonikolas, Jelena Diakonikolas, Daniel Kane, **Puqian Wang**, Nikos Zarifis. In *Neurips 2023*

7. **Information-Computation Tradeoffs for Learning Margin Halfspaces with Random Classification Noise**

   (α-β) Ilias Diakonikolas, Jelena Diakonikolas, **Puqian Wang**, Nikos Zarifis. In *COLT 2023*

8. **Robustly Learning a Single Neuron Via Sharpness**

   Puqian Wang*, Nikos Zarifis*, Ilias Diakonikolas, Jelena Diakonikolas. In *ICML 2023, Oral Presentation*

9. **Potential Function-based Framework for Minimizing Gradients in Convex and Min-Max Optimization**

   (α-β) Jelena Diakonikolas, **Puqian Wang**. In *SIAM Journal on Optimization, 2022*

Talks

- Informs Annual Meeting 2025, Atlanta, October 2025: Robustly Learning Single-Index Models
- IFDS Forum, UW-Madison, October 2024: Sample and Computationally Efficient Robust Learning of Gaussian Single-Index Models
- 60th Allerton Conference, Urbana-Champaign, September 2024: Robustly Learning a Single Neuron via Alignment Sharpness
- 25th International Symposium on Mathematical Programming (ISMP), Montreal, Canada, July 2024: Robustly Learning a Single Neuron via Alignment Sharpness
- IFDS Forum, UW Madison, February 2023: Robustly Learning a Single Neuron via Sharpness
- SIAM Annual Meeting, Undergraduate Research Presentation, Online, July 2021: How fast can you make the gradients small?

Service

- **Reviewer**: ICML 2024, 2025, ICLR 2024, Neurips 2023, 2024, 2025, 2026, FOCS 2026
- **Teaching Assitant**: CS 577 Introduction to Algorithms, Fall 2021; CS 726 Nonlinear Optimization, Fall 2022.