Zhengyan Wan

Email: zhengyanwan066@gmail.com | Github: https://github.com/WanZhengyan

EDUCATION

East China Normal University **Shanghai, China**

M.S. in Statistics. Advised by Prof. Fang Fang Sept 2023 – Jun 2026

- GPA: 3.88/4.0
- Major Courses: Advanced Probability Theory, Advanced Mathematical Statistics, Statistical Learning Theory, Convex Optimization, Survival Analysis, Empirical Processes, Stochastic Analysis

Jilin University **Changchun, China**

B.S. in Statistics Sept 2019 – Jun 2023

- GPA: 3.85/4.0
- Major Courses: Ordinary Differential Equations, Functional Analysis, Time Series Analysis, Multivariate Statistical Analysis, Nonparametric Statistics

PREPRINTS

[1] **Zhengyan Wan**, Yidong Ouyang, Qiang Yao, Liyan Xie, Fang Fang, Hongyuan Zha, and Guang Cheng. "Error Analysis of Discrete Flow with Generator Matching". In: *arXiv preprint* (2025). (Under review at ICLR). arXiv: 2509.21906. URL: https://arxiv.org/abs/2509.21906.

[2] **Zhengyan Wan***, Yidong Ouyang*, Liyan Xie, Fang Fang, Hongyuan Zha, and Guang Cheng. "Discrete Guidance Matching: Exact Guidance for Discrete Flow Matching". In: *arXiv preprint* (2025). (*: Equal Contribution; Under review at ICLR). arXiv: 2509.21912. URL: https://arxiv.org/abs/2509.21912.

[3] **Zhengyan Wan**, Fang Fang, and Binyan Jiang. "High-Dimensional Model Averaging via Cross-Validation". In: *arXiv preprint* (2025). (Under review at JMLR). arXiv: 2506.08451. URL: https://arxiv.org/abs/2506.08451.

RESEARCH EXPERIENCE

Error Analysis of Discrete Flow with Generator Matching July 2025 – Sept 2025

https://arxiv.org/pdf/2509.21906

- We derive a Girsanov-type theorem for continuous-time Markov chains (CTMCs) and the KL divergence between the path measures of two CTMCs.
- We establish a non-asymptotic error bound for the total variation between data distribution and estimated distribution in discrete flow-based models, taking estimation error into account.
- We provide a comprehensive error analysis for discrete flow-based models using neural networks with ReLU activation functions, simultaneously examining three sources of error: stochastic error, approximation error, and early stopping error.

Discrete Guidance Matching: Exact Guidance for Discrete Flow Matching Jun 2025 – Sept 2025

https://arxiv.org/pdf/2509.21912

- We introduce a guidance framework for discrete flow matching, achieving sampling efficiency without first-order approximation.
- We learn the guidance model by minimizing a Bregman divergence. We further propose a regularization technique that leverages samples from the target distribution.
- We demonstrate the effectiveness of the proposed framework through simulations, as well as preference alignment on text-to-image and multimodal understanding benchmarks.

High-Dimensional Model Averaging via Cross-Validation Jun 2024 – May 2025

https://arxiv.org/pdf/2506.08451v1

- We propose a high-dimensional model averaging method under a general framework.
- We establish rigorous theoretical guarantees for the proposed estimator, including the non-asymptotic upper bound, minimax lower bound and asymptotic optimality.

- We introduce a post-averaging one-step debiased estimator and establish its Gaussian and Bootstrap approximation results to construct simultaneous confidence intervals.
- We develop a first-order algorithm to solve simplex-constrained optimization problem, with a convergence analysis demonstrating a convergence rate matching the optimal rate of first-order algorithms.

Distributed Model Averaging with Nonignorable Nonresponse - Third author   Jan 2025 – Present
- We develop a weight estimator that minimizes a penalized squared-loss criterion built on an IPW-type empirical cumulative distribution function, under a distributed learning framework with nonignorable nonresponse.
- My contribution: verifying the correctness of the technical proofs & Developing an efficient ADMM algorithm for solving a quadratic programming with simplex constraints.

Visiting Experience

The Chinese University of Hong Kong, Shenzhen   **Shenzhen, China**
School of Data Science. Advised by Prof. Hongyuan Zha   Jun 2025 – Sept 2025

Teaching Experience

Teaching Assistant - Course: Stochastic Processes (Graduate-level Course)   Sept 2024 – Jan 2025
- Graded weekly assignments and provided feedback to students & Answered student queries

Scholarships

Panshi Scholarship (East China Normal University)   2024
First Class Scholarship of the School of Statistics (East China Normal University)   2024
First Class Scholarship & University Excellent Student (Jilin University)   2021

Skills

Programming languages: R, Python, C, MATLAB, SQL
Languages: Mandarin (native), English