LANGTIAN MA

langtian.ma@wisc.edu
Homepage GitHub LinkedIn

ABOUT ME

I'm a Statistics PhD student focusing on **Generative AI** (Diffusion Models) and **Computational Biology**, with strong foundations in **Mathematics** and **Software Engineering**. My current research focuses on developing **controllable** and **trustworthy** deep generative models for scientific discovery.

EDUCATION

University of Wisconsin - Madison — Madison, Wisconsin
Ph.D. in Statistics. — August 2024 - Present

Southern University of Science and Technology — Shenzhen, China
B.S. in Statistics. — August 2020 - June 2024

University of Toronto — Toronto, Canada
Exchange student, Faculty of Arts and Science. — September 2023 - December 2023

PUBLICATIONS & PREPRINTS

[1] Tianle Zhang, **Langtian Ma**, et al. "Rethinking Human Evaluation Protocol for Text-to-Video Models: Enhancing Reliability, Reproducibility, and Practicality." *NeurIPS 2024*.

[2] Yirui Huang*, **Langtian Ma***, et al. "Quantifying the Hierarchical Scales of Scientific Mobility." *International Conference on Computational Social Science (IC2S2)*, 2024. *Equal contribution

RESEARCH & PROJECTS

Diffusion Models for Single-Cell Data Generation — University of Wisconsin–Madison
Research Assistant | Advisor: Prof. Kris Sankaran — Jun 2025 – Present

- Built a customizable and reusable pipeline for single-cell data modeling and simulation with **diffusion models**.
- Developing controllable generation methods to synthesize out-of-distribution single-cell data.

scDesigner: Single-Cell Data Simulation Package — University of Wisconsin–Madison
Research Assistant | Advisor: Prof. Kris Sankaran — Jun 2025 – Present

- Designing extensible simulator modules with scikit-learn-style API and **object-oriented** design.
- Developed fast copula modules, achieving 1.2× faster fitting and 35× faster sampling with 20% partial modeling.

Human Evaluation Protocol for Text-to-Video Models
Research Collaboration (with Shanghai AI Lab) | Collaborator: Tianle Zhang — Apr 2024 – Aug 2024

- Co-designed an efficient human evaluation protocol for text-to-video models, reduced annotation cost by 53%.
- Quantified the efficiency–reliability trade-off, showing that the reliability loss does not affect the final conclusions.

Causal Estimation for Logistic Models — Southern University of Science and Technology
Undergraduate Researcher | Advisor: Yuting Ye — Mar 2024 – May 2024

- Analyzed bias of instrumental variable estimators in confounded logistic models under model misspecification.
- Proposed a method to mitigate confounding bias, with theoretical conditions and empirical validations.

Scientific Mobility and Career Dynamics — Southern University of Science and Technology
Undergraduate Researcher | Advisor: Prof. Yifang Ma — Jul 2022 – Dec 2022

- Analyzed hierarchical patterns in academic career mobility using large-scale bibliometric data from OpenAlex.
- Demonstrated the historical variation of scientific mobility across different administrative regions.

Interpretability of Recommender Systems (Course Project) — University of Wisconsin - Madison

- Identified key features that influence the predictions of the LightGCN model using Integrated Gradients method.
- Developed a concept-customizable recommendation method via Concept Activation Vectors.

TECHNICAL STRENGTHS

Programming Languages: Python, Java, R, C, Julia, SQL.

Software Engineering: Git, Docker, Linux, CLI Tools, Object-oriented Design, Package Development.

Machine Learning & Scientific Computing: PyTorch, PyTorch Lightning, Scikit-learn, Scipy, Sympy.