Jiaxin Hu

Email: jhu267@wisc.edu **Tel:** +1 (608) 9600989 **Address:** 1300 University Ave, Madison, WI, 53706
Personal website: https://jiaxin-hu.com

Education

University of Wisconsin-Madison — **Madison, U.S.**
Ph.D. in Statistics with minor in Computer Science, Advisor: Prof. Miaoyan Wang — *Sep. 2020 - present*

University of Wisconsin-Madison — **Madison, U.S.**
M.S. in Statistics — *Sep. 2018 - May. 2020*

Wuhan University — **Wuhan, China**
B.S. in Statistics — *Sep. 2015 - Jun. 2019*

Research Interests

Statistical Machine Learning, Matrix/Tensor Data Analysis, Network Analysis, Real-world Application

Publications

- **Jiaxin Hu** and Miaoyan Wang. Multiway Spherical Clustering via Degree-Corrected Tensor Block Models. *IEEE Transactions on Information Theory, 2023.*
  - Short version of this work was published in International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2022.
  - This work won **Best Student Paper Award – honorable mention** from the Statistical Learning and Data Science Section of the American Statistical Association (ASA), 2022.
  - Part of the work was accepted into Advances in Neural Information Processing Systems 34 (**NeurIPS**) Workshop on Quantum Tensor Networks in Machine Learning, 2021.
- **Jiaxin Hu**, Chanwoo Lee, and Miaoyan Wang. Generalized Tensor Decomposition With Features on Multiple Modes. *Journal of Computational and Graphical Statistics, 2021.*
  - This work won **Best Student Paper Award** from the Statistical Computing and Graphics Section of ASA, 2021.
  - Part of the work was accepted into **NeurIPS** Workshop on Machine Learning and the Physical Sciences, 2020.
- Miao J., Song G., Wu Y., **Hu J.**, Wu Y., Basu S., Andrews J., Schaumberg K., Fletcher J., Schmitz L., Lu Q. Reimagining Gene-Environment Interaction Analysis for Human Complex Traits. *Under review of Cell, 2023.*

Awards

- Honorable Mention Graduate Course Teaching Assistant (TA) Award, University of Wisconsin-Madison 2023
- Student Research Grants Competition Award, University of Wisconsin-Madison 2022
- Best Student Paper Award – honorable mention, Statistical Learning and Data Science Section, ASA 2022
- Best Student Paper Award, Statistical Computing and Graphics Section, ASA 2021
- Exchange & Visiting International Student Academic Excellence Award, University of Wisconsin-Madison 2019
- Second-Class Scholarship for Excellent Students (top 10%), Wuhan University 2016, 2017, 2018

Skills & Core Graduate Courses

- **Programming Skills:** Proficient in R, Matlab, Python, LaTeX; familiar with Shell, High Performance Computing
- **Statistics courses:** Mathematical Statistics, Statistical Methods, Statistical Learning Theory, Statistical Computing
- **Machine Learning courses:** Machine Learning, Theoretical Foundations of Machine Learning
- **Theoretical Data Science courses:** Nonlinear Optimization, Nonparametric Method in Data Science

Software

- `dTBM`. R package to implement the multiway clustering via degree-corrected tensor block models.
- `tensorregress`. R package to implement the generalized tensor decomposition with multiple feature matrices.

Services

- **Invited Session Chair** for Statistics in Social Sciences and Survey Research at Joint Statistical Meetings, 2022
- **Invited Session Chair** for Recent Methods for Tensor Data at New England Statistical Society symposium, 2022
- **Invited Session Chair** for Software on Statistical Computing at Joint Statistical Meetings, 2021
- **Reviewer** for AISTATS 2021, 2022, Statistics and Its Interface, Computational Statistics and Data Analysis
- **Research Assistant** at Institute for Foundation of Data Science, University of Wisconsin-Madison, Spring 2022
- **Discussion TA** for STAT 850 (Ph.D. compulsory course), Theory and Application of Regression, Spring 2023
- **Discussion TA** for STAT 601 (core graduate course), Statistical Method, Fall 2020-2022, Spring 2021

Selected Projects

- **Multiway Spherical Clustering via Degree-Corrected Tensor Block Models** 2021 - 2023
  - Proposed a novel degree-corrected tensor block model (dTBM) accounting for the individual heterogeneity in multiway clustering tasks.
  - Established minimax lower bounds and maximum likelihood estimator (MLE) upper bounds of dTBM based on the notion of angle separability; discovered the special phase transition for the statistical-computational behavior of dTBM clustering with higher-order tensors.
  - Designed an efficient two-step algorithm that provably achieves exact recovery under the mild signal condition with polynomial-time complexity.
  - Created an R package, `dTBM`; performed massive simulations with High Performance Computing to verify theoretical results and compare with competitive methods.
  - Applied dTBM to Human Brain Connectome project dataset and uncovered brain node clusters that align with the functional partition in the human brain; applied dTBM to Peru Legislation network datasets demonstrating that dTBM outperforms competitive methods in identifying legislators' party affiliations.
- **Generalized Tensor Decomposition With Features on Multiple Modes** 2019 - 2021
  - Proposed a novel supervised tensor decomposition (STD) that incorporates side-information from multiple feature matrices and handles a broad range of data types.
  - Designed an efficient alternating optimization algorithm with accuracy guarantees for global minimizers and local algorithm outpus.
  - Created an R package, `tensorregress`; performed massive simulations with High Performance Computing to evaluate the finite-sample performance of STD and compare with multiple competitive methods.
  - Applied STD to Human Brain Connectome project dataset and multi-relational political network data identifying the key global connectivity pattern and the local region associated with interested features.
- **A Spectral Framework to Map QTLs Affecting Joint Differential Networks of Gene Coexpression** (ongoing)
  - Designed a novel spectral framework (snQTL) to map quantitative trait loci (QTLs) affecting the joint differential networks of gene coexpression.
  - Creatively adopted matrix and tensor spectral decomposition to resolve the multiple testing burden and information loss issues in coexpression QTL studies.
  - Proposed a simulation framework to generate synthetic genotype and expression data that effectively mimics the real genetics data; performed massive simulations confirming the efficacy of snQTL.
  - Conducted snQTL analysis on real high-dimensional sequencing data of three-spined sticklebacks; identified three clusters of network QTLs in the region with strong genomic signatures of natural selection and the joint differential networks that regulate the oxygen transportation.