Yuchen Xu

Cornell University, Department of Statistics & Data Science
Ithaca, NY 14853

yx439@cornell.edu
(+1)6072621106

EDUCATION

2018 - 2023 Ph.D. Candidate in Statistics & Data Science
Cornell University, Ithaca, NY

2014 - 2018 B.S. in Mathematics & Applied Mathematics, Zhiyuan Honored Program
Shanghai Jiao Tong University, Shanghai, China

Fall 2017 Research Intern Exchange
Center for Applied Mathematics, Cornell University, Ithaca, NY

August 2016 Summer Course on Partial Differential Equations
Hertford College, Oxford University, Oxford, UK

RESEARCH THEORY & METHODS

Time Series Analysis:
- Changepoints
- (Hidden) Markov model

Multivariate Analysis:
- Joint matrix diagonalization
- Tensor decomposition
- Blind Source Separation (BSS)

Image Analysis:
- Blob detection
- Ridge detection

RESEARCH APPLICATIONS

- Financial Econometrics
- Medical Images and Signals
- Nanoparticles
- Molecular Dynamics
- Geology

RESEARCH EXPERIENCE

2018 - Present Matteson Lab, Cornell University
Advisor: Prof. David S. Matteson
Tasks: Testing simultaneous diagonalizability.

2019 - Present Atomic-Level Structural Dynamics in Catalysts (ALSDC) Group
Advisor: Prof. David S. Matteson
Tasks: Clustering nanoparticle structures, w/ Prof. Roberto Rivera;
Extraction of TEM atomic columns, w/ Prof. Peter A. Crozier;
Estimating transition rate matrices, w/ Prof. Mahmoud Moradi.

2021 - Present Enterprise Heart Failure Program, New York-Presbyterian (NYP) Hospital
Advisor: Prof. Martin Wells
Tasks: Heart failure inference from ECG data, w/ Ashley N. Beecy, MD.

2022 - Present Department of Surgery, Icahn School of Medicine at Mount Sinai Hospital
Advisor: Prof. David S. Matteson
Tasks: Predicting thyroid cancer recurrence, w/ Denise Lee, MD.

PUBLICATIONS[1]

Manzorro, R., ***Xu,Y.***, Vincent, J. L., Rivera, R., Matteson, D. S., and Crozier, P. A., "Exploring blob detection to determine atomic column positions and intensities in time-resolved TEM images with ultra-low signal-to-noise," *Microscopy and Microanalysis*, vol. 28, no. 6, pp. 1917–1930, Mar. 2022. DOI: 10.1017/s1431927622000356. [Online]. Available: https://doi.org/10.1017%2Fs1431927622000356,

◦ ***The Most Outstanding Students Awards, Bronze Medal, UPSTAT 2021 Conference***.

MANUSCRIPTS UNDER REVIEW OR REVISION

Xu,Y., Düker, M.-C., and Matteson, D. S., *Testing simultaneous diagonalizability*, 2021. DOI: 10.48550/ARXIV.2101.07776. [Online]. Available: https://arxiv.org/abs/2101.07776, Minor Revision resubmitted to Journal of the American Statistical Association.

Xu,Y., Thomas, A. M., Crozier, P. A., and Matteson, D. S., *Dynamic Atomic Column Detection in Transmission Electron Microscopy Videos via Ridge Estimation*, DOI & URL TBA, 2023,

◦ ***The Best Student Poster Award, 2022 IEEE Western New York Image and Signal Processing Workshop (WNYISPW)***.

Thomas, A. M., Crozier, P. A., **Xu,Y.**, and Matteson, D. S., *Detection and hypothesis testing of features in extremely noisy image series using topological data analysis, with applications to nanoparticle videos*, 2022. DOI: 10.48550/ARXIV.2209.13584. [Online]. Available: https://arxiv.org/abs/2209.13584, Submitted to Technometrics.

Goolsby, C., Losey, J., **Xu,Y.**, Düker, M.-C., Sherman, M. G., Matteson, D. S., and Moradi, M., "Addressing the embeddability problem in transition rate estimation," Aug. 2019. DOI: 10.1101/707919. [Online]. Available: https://doi.org/10.1101%2F707919.

SOFTWARE

R package `eigTest` available on Github: Jointly Estimate and Test for Common Eigenvectors.

PRESENTATIONS

Non-parametric ridge recovery of TEM image series given temporal parameterization, 2022 IEEE Western New York Image and Signal Processing Workshop (WNYISPW), (Hybrid) Rochester, NY, Nov. 2022.

Non-parametric ridge recovery of TEM image series given temporal parameterization, Science-Integrated Statistical Learning Section, 2022 INFORMS Annual Meeting, Indianapolis, IN, Oct. 2022.

Recording atomic column positions and intensities via Blob Detection in noise-degraded TEM frames, Data Science in Science Minisymposia, The 37th SIDIM, (Virtual) Puerto Rico, Feb. 2022.

Recording atomic column positions and intensities via Blob Detection in noise-degraded TEM frames, UPSTAT 2021 Conference, (Virtual) Rochester, NY, Apr. 2021.

Testing Simultaneous Diagonalizability, Cornell Celebration of Statistics and Data Science, Ithaca, NY, Sep. 2019.

Testing Simultaneous Diagonalizability, Business and Economic Statistics Section, Speed Session, Joint Statistical Meeting (JSM), Denver, CO, Jul. 2019.

LINKS

Website **Github** **LinkedIn** **Google Scholar** **ORCID**

[1] * First authors contributed equally.

SKILLS

Programming:	R	Matlab	Python	LaTeX
	SQL	AWS	Stan	Java
Language:	English		Mandarin	

TEACHING EXPERIENCE

Spring 2023	Understanding Machine Learning	Teaching Assistant
@ Cornell	Instructor: Andrew M. Thomas	STSCI 4750
Fall 2021	Operations Research Tools for Financial Engineering	Teaching Assistant
@ Cornell	Instructor: David Ruppert	STSCI 4630
Spring 2021	Statistics for Financial Engineering	Teaching Assistant
@ Cornell	Instructor: David S. Matteson	STSCI 5640
Fall 2020	Statistical Sampling	Teaching Assistant
@ Cornell	Instructor: Thomas DiCiccio	STSCI 3100
Spring 2020	Basic Probability	Teaching Assistant
@ Cornell	Instructor: Laurent Saloff-Coste	MATH 4710
Fall 2019	Probability Models and Inference	Teaching assistant
@ Cornell	Instructor: Florentina Bunea	STSCI 3080

SERVICE

January 2023	Reviewer for the Journal of Service Research.
November 2021	Reviewer for the Journal of Econometrics.
January 2021	Reviewer for the Best Student Paper Competition of Joint Statistical Meeting (JSM) Business and Economic Statistics Section (B&E).

INDUSTRY EXPERIENCE

May 2022 — — Aug 2022 Data Scientist Intern
Amazon Web Services (AWS), Seattle, WA
Tasks: Modeling efficacy for internal IT-Services products;
Optimizing data aggregation and interpretation logics.

Mar 2018 — — May 2018 Algorithm & Data Science Intern
China Appraisal Association Data Analysis (CAAD), Shanghai, China
Tasks: Regressing and predicting real estate appraisals;
Optimizing address search algorithms.