

XIN XING

Department of Statistics
Hutcheson Hall, 250 Drillfield Drive
E-mail: xinxing@vt.edu
Website: <https://xin-xing.github.io/>

Virginia Tech
Blacksburg, VA, USA
Phone: +1 (706) 207-9841

APPOINTMENTS

Assistant Professor Department of Statistics, Virginia Tech August 2020 - present

EDUCATION

Postdoc	Statistics	Harvard University Mentor: Dr. Jun S. Liu	2018 - 2020
Ph.D.	Statistics	University of Georgia Advisor: Dr. Wenxuan Zhong Thesis topic: Statistical Methods with Applications in Epigenomics, Metagenomics and Neuroimaging	2013 - 2018
M.S.	Statistics	University of Science and Technology of China (USTC) Advisor: Dr. Yaning Yang	2013
B.S.	Statistics	University of Science and Technology of China (USTC)	2010

RESEARCH INTEREST

Bioinformatics: metagenomics, single cell, epigenomics, neuroimaging
Statistics: minimax nonparametric testing, smoothing spline, dimension reduction, controlled variable selection
Machine learning: computer vision, neural network compression, variable selection via neural network

PAPERS

- [1] Yiwen Liu, **Xin Xing**, Wenxuan Zhong (2018). *Sufficient Dimension Reduction for Tensor Data*. Handbook of Big Data Analytics, Springer.
 - [2] **Xin Xing**, Zuofeng Shang, Pang Du, Ping Ma, Wenxuan Zhong, Jun S. Liu (2019). *Minimax Nonparametric Two-sample Test*, Journal of the American Statistical Association: Theory and Methods, revised. [arXiv:1911.02171](https://arxiv.org/abs/1911.02171)
 - [3] **Xin Xing**, Zhigen Zhao, Jun S. Liu (2019). *Controlling False Discovery Rate Using Gaussian Mirrors*, Journal of the American Statistical Association: Theory and Methods, revised. [arXiv:1911.09761](https://arxiv.org/abs/1911.09761)
 - [4] **Xin Xing**, Yu Gui, Chengguang Dai, Jun S. Liu (2020). *Deep Gaussian Mirror for Controlled Variable Selection*, IEEE ICMLA, 2020.
 - [5] **Xin Xing**, Meimei Liu, Wenxuan Zhong, Ping Ma. (2020). *Minimax Nonparametric Parallelism Test*, Journal of Machine Learning Research, 21(94):1-47.
-

- [6] **Xin Xing**, Meimei Liu, Weiping Zhang (2014). *Joint Semiparametric Mean-Covariance Modeling by Moving Average Cholesky Decomposition for Longitudinal Data*, Journal of University of Science and Technology of China, 43(8), 607-621.
- [7] **Xin Xing**, Long Sha, Pengyu Hong, Zuofeng Shang, Jun S. Liu (2020). *Probabilistic Connection Importance Inference and Lossless Compression of Deep Neural Networks*. International Conference on Learning Representations (ICLR).
- [8] **Xin Xing**, Jun S. Liu, Wenxuan Zhong (2018). *MetaGen: reference-free learning with multiple metagenomic samples*. Genome biology, 18 (1), 187, 2018.
- [9] **Xin Xing**, Jinjin Hu, and Yaning Yang (2014). *Robust minimum variance portfolio with L-infinity constraints*, Journal of Banking and Finance, 46, 107-117.
- [10] **Xin Xing**, Peng Zeng, Wenxuan Zhong (2019). *Parsimonious Tensor Dimension Reduction*, Statistica Sinica, revised.
- [11] **Xin Xing**, Di Xiao, Rui Xie, Wenxuan Zhong (2019). *Model-based Dictionary Learning: Sparse Coding beyond Gaussian Independent Model*, IEEE Transactions on Signal Processing, under review.
- [12] Wenxuan Zhong, **Xin Xing**, Kenneth Suslick (2015). *Tensor Sufficient Dimension Reduction* WIREs Computational Statistics, 7(3), 178-184.
- [13] Terry Ma, Di Xiao, **Xin Xing** (Corresponding author) (2019). *MetaBMF: A Scalable Binning Algorithm for Large-scale Reference-free Metagenomic Studies*. Bioinformatics.
- [14] Terry Ma, **Xin Xing** (2018). *A Scalable Reference-Free Metagenomic Binning Pipeline*. International Symposium on Bioinformatics Research and Applications, 79-83.
- [15] Ping Ma, Xinlian Zhang, **Xin Xing**, Jinyi Ma Michael Mahoney (2020). *Asymptotic Analysis of Sampling Estimators for Randomized Numerical Linear Algebra Algorithms*, International Conference on Artificial Intelligence and Statistics (AISTATS).
- [16] Chengguang Dai, Buyu Lin, **Xin Xing** and Jun S. Liu (2019). *A Scale-free Approach for False Discovery Rate Control in Generalized Linear Models*, Journal of the American Statistical Association: Theory and Methods, under review [arXiv:2007.01237](https://arxiv.org/abs/2007.01237)
- [17] Chengguang Dai, Buyu Lin, **Xin Xing** and Jun S. Liu (2019). *False Discovery Rate Control Via Data Splitting*, Journal of the Royal Statistical Society: Series B, under review.

PAPERS IN PREPARATION

- [1] Representation Learning of T-cell Receptor in cancer immunotherapy (with Songpeng Zu, Xiaole Liu and Jun S. Liu)
- [2] Bayesian Slicing Methods for Detecting Variable Dependency (with Jiexing Wu and Jun S. Liu)
- [3] Variable Hunting: New Promise in Binary Predictor Selection (with Wenxuan Zhong and Ping Ma)

RESEARCH HIGHLIGHTS

- Research is highlighted on the website of graduate school of UGA.
Link: <https://grad.uga.edu/index.php/xin-2017/>
- Research is highlighted on the website of Pittsburgh Supercomputer Center.
Link: <https://www.psc.edu/161-news/psc-highlights/2414-bridges-reveals-diabetes-gut-microbe-links-2>

GRANT

- NVIDIA GPU Grant for Accelerated Data Science, 2019.

HONORS AND AWARDS

- Best Senior PhD Student, UGA, May. 2018.
- Best Beginning PhD Student, UGA, Jul. 2014.
- Outstanding Dissertation For Bachelor's Degree, USTC, Jul. 2010.
- Third Prize of The Second Statistical Contest in Modeling, USTC, May. 2010.
- Third Prize of The First Statistical Contest in Modeling, USTC, May. 2009.
- Outstanding Student Scholarship (Silver Award), USTC, Oct. 2006.

TEACHING

- CMDA 3654 Lecture: Intro to Data Analytics & Visualization, Fall, 2020.
- STAT 221 Guest Lecturer : Monte Carlo Methods for Statistical Learning and Intro to Deep Learning, Harvard, Fall, 2019.
- STAT 8200 TA : Experimental Designs, University of Georgia, Fall, 2017.
- STAT 6210 TA : Introduction to Statistics I, University of Georgia, Fall, 2016.
- MSIT 3000 TA : Stat Analysis for Business I, University of Georgia, Spring, 2015.

CONFERENCES & INVITED TALKS

- Neural Gaussian Mirror for Controlled Feature Selection in Neural Networks
The International Conference on Learning Representations (ICLR), Online, April, 2020 (Invited talk)
- To Knockoff or To Disturb? Controlled Variable Selection in Regression Problems
Department of Statistics, Virginia Tech, VA, Feb. , 2020 (Invited talk)
- PCII: Probabilistic Connection Importance Inference for Neural Network Lossless Compression
International Chinese Statistical Association (ICSA), NJ, 2019 (Invited talk)
- Minimax Nonparametric Test for Density Comparison
Department of Mathematical Science, IUPUI, IN, Mar 2018 (Invited talk)
- Reference-free Learning with Multiple Metagenomic Samples
Department of Statistics, Harvard University, MA, Feb 2017 (Invited talk)
- Identifying Nonparallel Differentially Methylated DNA Regions
University of Science and Technology of China, Hefei, China, Jan. 2017 (Invited talk)
- Identifying Nonparallel Differentially Methylated DNA Regions
Fudan University, Shanghai, China, Jan. 2017 (Invited talk)
- Minimax Nonparametric Test for Density Comparison
The Georgia Statistics Day Conference, GA, Oct 2015 (Invited talk)
- Nonnegative Matrix Factorization for Metagenomic Deconvolution
Algorithm of Threat Detection Conference, CO, Sep 2014 (Invited talk)
- Nonnegative Matrix Factorization for Metagenomic Deconvolution
Department of Automation, Tsinghua University, Beijing, China, Jun. 2014 (Invited talk)

SOFTWARE

- Parallelism: R package "Parallelism" is to provide a nonparametric testing approach to test whether the spatial or temporal signals in treatment and control groups are parallel or not.
Link: <https://github.com/BioAlgs/Parallelism>

- MetaMat: a fast algorithm for large-scale reference-free metagenomic studies. The pipeline outputs all binned species in multiple metagenomic samples and their estimated relative abundances.
Link: <https://github.com/didi10384/MetaBMF>
- MetaGen: a statistically based algorithm to simultaneously identify microbial species and estimate their abundances in multiple metagenomic samples without using any reference genome.
Link: <https://github.com/BioAlgs/MetaGen>

PROFESSIONAL SERVICE

- Referee Service, *Journal of the American Statistical Association* 2018-present
- Referee Service, *Electronic Journal of Statistics* 2018-present
- Organizer, ICSA 2018 Applied Statistics Symposium 2018
- Referee Service, *IEEE Transactions on Computational Biology and Bioinformatics* 2018-present
- Referee Service, *Statistical Analysis and Data Mining* 2018-present
- Referee Service, *Frontiers in Genetics* 2016-present
- Organizer, The Georgia Statistics Day Conference, Athens, GA Oct. 2016
- Member, American Statistical Association Since 2015
- Member, International Chinese Statistical Association Since 2015
- Judge, Georgia Science and Engineering Fair 2015