

# CURRICULUM VITAE

## PERSONAL INFORMATION

---

<b>Name:</b>	Xin Bing	<b>English first name:</b>	Mike
<b>Email:</b>	xin.bing@utoronto.edu		
<b>Phone:</b>	(607)-697-3443		
<b>Current position:</b>	Assistant Professor		
	Department of Statistical Sciences, University of Toronto		

## EDUCATION

---

<b>Cornell University</b> Ph.D. in Statistics with minor in Mathematics <i>Advisors: Florentina Bunea and Marten Wegkamp</i>	09/2016 – 06/2021
<b>University of Washington – Seattle</b> M.S. in Statistics	09/2014 – 06/2016
<b>Shandong University</b> M.F.E. in Financial Mathematics	09/2013 – 06/2014
<b>Shandong University</b> B.S. in Mathematics and Statistics	09/2009 – 06/2013

## EMPLOYMENT

---

**Assistant Professor:** Department of Statistical Sciences at the University of Toronto. 06/2022 – present

**Research Associate:** Department of Pure Mathematics and Mathematical Statistics at the University of Cambridge. Hosted by Richard J. Samworth. 10/2021 – 01/2022

**Computational Systems Immunology Research Associate:** jointly between School of Medicine at the University of Pittsburgh (hosted by Jishnu Das) and Department of Statistics and Data Science at Cornell University. 01/2020 – 08/2020

## RESEARCH OF INTEREST

---

High-dimensional statistics, low-rank matrix estimation, mixture models, multivariate analysis, model-based clustering, latent factor model, topic models, minimax estimation, high-dimensional inference, statistical and computational trade-offs, applications of the Wasserstein distance in statistics, applications to genetics, neuroscience, immunology and other areas.

## HONORS AND AWARDS

---

- 2021 Hsien and Daisy Yen Wu Scholarships
- 2021 IMS Lawrence D. Brown Ph.D. Student Award
- Honor graduate of Shandong University in 2013
- Excellent student scholarship grantee for meticulous academic performance from 2010 to 2012

## JOURNAL PUBLICATIONS

---

1. Xin Bing, Xin He, Dian Jin and Yuqian Zhang.  
Optimal vintage factor analysis with deflation varimax.  
*The Annals of Statistics*, 2025+. [arXiv link](#).
2. Xin Bing, Florentina Bunea and Jonathan Niles-Weed.  
The Sketched Wasserstein Distance for mixture distributions.  
*Bernoulli*, 2025+. [arXiv link](#).
3. Xin Bing and Marten Wegkamp.  
Interpolating Discriminant Functions in High-Dimensional Gaussian Latent Mixtures.  
*Biometrika*, 111(1): 291-308. March, 2024. [Link to the paper](#).
4. Xin Bing, Wei Cheng, Huijie Feng and Yang Ning.  
Inference in High-dimensional Multivariate Response Regression with Hidden Variables.  
*Journal of the American Statistical Association (Theory & Method)*. Sep, 2023. [Link to the paper](#).
5. Xin Bing and Marten Wegkamp.  
Optimal Discriminant Analysis in High-Dimensional Latent Factor Models.  
*The Annals of Statistics* 51(3): 1232-1257. June, 2023. [Link to the paper](#).
6. Xin Bing, Florentina Bunea and Marten Wegkamp.  
Detecting approximate replicate components of a high-dimensional random vector with latent structure.  
*Bernoulli*, 29(2): 1368–1391. May, 2023. [Link to the paper](#).
7. Dian Jin, Xin Bing and Yuqian Zhang.  
Unique sparse decomposition of low rank matrices.  
*IEEE Transactions on Information Theory*, 69(4): 2452–2484, April 2023. [Link to the paper](#).
8. Xin Bing, Florentina Bunea, Seth Strimas-Mackey and Marten Wegkamp.  
Likelihood estimation of sparse topic distributions in topic models and its applications to Wasserstein document distance calculations.  
*The Annals of Statistics*, 50(6): 3307–3333, December 2022. [Link to the paper](#).
9. Xin Bing, Florentina Bunea and Marten Wegkamp.  
Inference in interpretable latent factor regression models.  
*Bernoulli*, 28(2): 997 – 1020, May 2022. [Link to the paper](#).
10. Xin Bing, Yang Ning and Yaosheng Xu.  
Adaptive estimation of multivariate regression with hidden variables.  
*The Annals of Statistics*, 50(2): 640–672, 2022. [Link to the paper](#).
11. Xin Bing, Florentina Bunea, Seth Strimas-Mackey and Marten Wegkamp.  
Prediction in latent factor regression: Adaptive PCR and beyond.  
*The Journal of Machine Learning Research*, 22(177): 1–50, 2021. [Link to the paper](#).
12. Xin Bing, Florentina Bunea and Marten Wegkamp.  
Optimal estimation of sparse topic models.  
*The Journal of Machine Learning Research*, 21(177): 1–45, 2020. [Link to the paper](#).
13. Xin Bing, Florentina Bunea and Marten Wegkamp.  
A fast algorithm with minimax optimal guarantees for topic models with an unknown number of topics.  
*Bernoulli*, 26(3): 1765–1796, 2020. [Link to the paper](#).
14. Xin Bing, Florentina Bunea, Yang Ning and Marten Wegkamp.  
Adaptive estimation in structured factor models with applications to overlapping clustering.  
*The Annals of Statistics*, 48(4): 2055–2081, 2020. [Link to the paper](#).

15. Xin Bing and Marten Wegkamp.  
Adaptive estimation of the rank of the coefficient matrix in high-dimensional multivariate response regression models.  
*The Annals of Statistics*, 47(6): 3157–3184, 2019. [Link to the paper.](#)

## CONFERENCE PUBLICATION

---

1. Chao Wang, Xin Bing, Xin He and Caixing Wang.  
Towards Theoretical Understanding of Learning Large-scale Dependent Data via Random Features.  
*International Conference on Machine Learning (ICML)*, spotlight, 2024. [Link to the paper.](#)
2. Dian Jin, Xin Bing and Yuqian Zhang.  
Unique sparse decomposition of low rank matrices.  
*Neural Information Processing Systems (NeurIPS)*, 2021. [Link to the paper.](#)

## COLLABORATIVE PUBLICATIONS AND DISCUSSIONS

---

1. Javad Rahimikollu, Hanxi Xiao, Anna E. Rosengart, Tracy Tabib, Paul Zdinak, Kun He, Xin Bing, Florentina Bunea, Marten Wegkamp, Amanda C. Poholek, Alok V Joglekar, Robert A Lafyatis, Jishnu Das.  
SLIDE: Significant Latent Factor Interaction Discovery and Exploration across biological domains.  
*Nature Methods*. Feb, 2024. [Link to the paper.](#)
2. Xin Bing, Tyler Lovelace, Florentina Bunea, Marten Wegkamp, Harinder Singh, Panayiotis Benos, Jishnu Das.  
Essential Regression – a generalizable framework for inferring causal latent factors from multi-omic human datasets.  
*Patterns* (Cell press), 3(5): 100473, 2022. [Link to the paper.](#)
3. Xin Bing, Florentina Bunea, Martin Royer and Jishnu Das.  
Latent model-based clustering for biological discovery.  
*iScience* 14 (2019), 125–135. [Link to the paper.](#)
4. Xin Bing and Marten Wegkamp.  
Discussion of random-projection ensemble classification by Timothy I. Cannings and Richard J. Samworth.  
*J. R. Statist. Soc. B* 79 (2017), no. 4, 1006–1007.

## PAPERS UNDER REVIEW

---

1. Xin Bing, Bingqing Li and Marten Wegkamp. Linear Discriminant Regularized Regression. [arXiv link.](#)
2. Xin Bing and Derek Latremouille. High-Dimensional Invariant Tests of Multivariate Normality Based on Radial Concentration. [arXiv link.](#)
3. Xin Bing, Xin He and Chao Wang. Kernel Ridge Regression with Predicted Feature Inputs and Applications to Factor-Based Nonparametric Regression. [arXiv link.](#)
4. Eugen Pircalabelu and Xin Bing. Overlapping clustering of time dependent variables for fMRI data.
5. Xin Bing, Florentina Bunea, Jonathan Niles-Weed and Marten Wegkamp. Learning large softmax mixtures with warm start EM. [arXiv link.](#)
6. Xin Bing, Dehan Kong and Bingqing Li. Convergence and Optimality of the EM Algorithm Under Multi-Component Gaussian Mixture Models. [arXiv link.](#)

## WORK IN PROGRESS

---

1. The Behaviour of High-Dimensional Euclidean Space: Concentration, Geometric Representation, Projection, & Intrinsic Dimension, with Application to Outlier Detection. Joint work with Derek Latremouille. (Preprint available upon request)
2. Iterative Clustering in Low-dimensional Subspace via Multiple Dimension Reductions. Joint work with Dehan Kong and Bingqing Li.
3. Statistically Optimal and Computationally Efficient Non-parametric Latent Factor Regression. Joint work with Chao Wang and Xin He.
4. Double denoising k-means clustering. Joint work with Xin He and Shangkai Zhu.

## GRANTS

---

- Graduate Student Exchange Scholarship from the Canadian Statistical Sciences Institute (CANSSI). (\$15K) 2024 – 2025
- NSERC Discovery Grant Individual. (\$30K per year) 2023 – 2028
- NSERC Discovery Launch Supplement. (\$12.5K in total) Apr 2023 – Mar 2024

## INVITED TALKS AND PRESENTATIONS

---

- The 2025 Joint Statistical Meetings (JSM). 2–7, Aug, 2025. Nashville, TN, USA. [Invited talk]
- Seminar in the Department of Statistical Sciences at Cornell University. 22 Apr, 2025. [Departmental seminar talk]
- Seminar in the Department of Statistics and Probability at Michigan State University. 10 Apr, 2025. [Departmental seminar talk]
- Seminar in the Siebel School of Computing and Data Science at the University of Illinois Urbana-Champaign. 7 Apr, 2025. [Departmental seminar talk]
- The 7th International Conference on Econometrics and Statistics: 17-19, July 2024. Beijing Normal University, Beijing, China. [Invited talk]
- The 2nd Joint Conference on Statistics and Data Science in China: 12-14 July, 2024. Haigeng Convention Center, Kunming, Yunnan, China. [Invited talk]
- 2024 Hangzhou International Conference on Frontiers of Data Science: 8-10 July, 2024. Center for Data Science of Zhejiang University, Hangzhou, Zhejiang, China. [Invited talk]
- Oberwolfach Workshop: Statistics and Learning Theory in the Era of Artificial Intelligence. 23–28 June, 2024. Oberwolfach Research Institute for Mathematics, Germany. [Invited talk]
- The 2024 WNAR/IMS/Graybill meeting: 9-12 June, 2024. Colorado State University, Fort Collins, Colorado, U.S.A.. [Invited talk]
- The Sixth ICSA-Canada Chapter Symposium: 7-9 June, 2024. Niagara Falls, Canada. [Invited talk]
- 16th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-Statistics 2023). HTW Berlin, University of Applied Sciences. Berlin, Germany. 16-18 December, 2023. [Invited talk and organizer of an invited session]
- Seminar in the Department of Mathematics and Statistics at the McMaster University. 31 Oct, 2023. [Departmental seminar talk]
- International Chinese Statistical Association (ICSA): 2023 applied statistics symposium. 11–14 Jun, 2023. Ann Arbor, Michigan. [Invited talk]

- Seminar in the School of Mathematical Sciences at Shanghai Jiao Tong University. 05 June, 2023. [Departmental seminar talk]
- Seminar in the Department of Mathematics at Shandong University. 22 May, 2023. [Departmental seminar talk]
- Seminar in the Department of Statistics and Management at Shanghai University of Finance and Economics. 28 April, 2023. [Departmental seminar talk]
- Seminar in the Department of Mathematics and Statistics at Auburn University. 19 April, 2023. [Departmental seminar talk]
- Seminar in the Department of Statistics and Actuarial Science at the University of Waterloo. 08 March, 2023. [Departmental seminar talk]
- 15th International Conference of the ERCIM WG on Computational and Methodological Statistics (CM-Statistics 2022). King's College London, UK. 17–19 December 2022. [Invited talk]
- International Chinese Statistical Association (ICSA): 2022 applied statistics symposium. 19–22 Jun, 2022. Gainesville, Florida. [Invited talk]
- Oberwolfach Workshop: Re-thinking High-dimensional Mathematical Statistics. 15–21 May, 2022. Oberwolfach Research Institute for Mathematics, Germany. [Postponed due to pandemic]
- Seminar in the Department of Statistics and Management at Shanghai University of Finance and Economics. 28 April, 2022. [Departmental seminar talk]
- Seminar in the Institute of Data Analysis and Modeling in economics and statistics, UCLouvain. Nov 26, 2021. [Departmental seminar talk]
- Joint Statistical Meetings / IMS Annual Meeting. Seattle, Washington, USA, August 7–12, 2021. [Special talk: IMS Lawrence D. Brown Ph.D. Student Award]
- International Chinese Statistical Association (ICSA): 2020 applied statistics symposium. Houston, TX. [Postponed due to pandemic]
- Workshop at Institute for Advanced Study (IAS, Princeton University): Missing Data Challenges in Computation, Statistics, and Applications. March 16–18, 2020. [Postponed due to pandemic]
- International Chinese Statistical Association (ICSA): 2019 applied statistics symposium. 09 – 12 Jun, 2019. Raleigh, NC. [Invited talk]
- International Center for Mathematical Science (ICMS) workshop: Computational strategies for large-scale statistical data analysis. Date: 02–07 July 2018. Edinburgh, UK. [Invited talk]
- Oberwolfach Workshop: Matrix Estimation Meets Statistical Network Analysis: Extracting low-dimensional structures in high dimension. 17–23 Jun 2018. Oberwolfach Research Institute for Mathematics, Germany. [Contributed talk]

## TEACHING

---

Teaching at the University of Toronto

- |   |                                      |
|---|--------------------------------------|
| • STA314H: Statistical Methods for Machine Learning I | Falls in 2022 – 2024, spring in 2026 |
| • STA2211HS: Graduate Probability II                  | Springs in 2025 and 2026             |
| • STA3000Y: Advanced Theory of Statistics             | Spring in 2026                       |

## SERVICE TO THE ACADEMIC COMMUNITY

---

Referee for the Electronic Journal of Statistics, the Annals of Statistics, Biometrika, Statistica Sinica, Journal of the Royal Statistical Society - Series B, Journal of American Statistical Association, Journal of Computational and Graphical Statistics, Journal of Machine Learning Research, the SIAM Journal on Imaging Sciences, the Journal of Econometrics