Florentina Bunea

Department of Statistics and Data Science
1198 Comstock Hall
Cornell University
Ithaca, New York, 14853-2601
e-mail: fb238@cornell.edu

Education

- Ph.D. in Statistics, University of Washington, Seattle, 2000.
  Advisor: Jon A. Wellner.

- B.S./M.S. in Mathematics, First Class, University of Bucharest, Romania, 1991.

Employment

- **Professor:** Cornell University, Department of Statistics and Data Science, 2011 -

  - Graduate Field Faculty Member: Statistics; Computer Science;
    the Center for Applied Mathematics (CAM).
  - Member of the Advisory Council for Data Science for Enterprise and Society.
  - Member of the Advisory Committee for the Office of Faculty Development.

- Affiliate Professor, NYU Center for Data Science, 2021 - 2025+.

- **Assistant/Associate Professor**: Florida State University, Department of Statistics, 2000 - 2011.

- Research/Teaching Assistant: University of Washington, Department of Statistics, September 1995 - August 2000; University Politehnica Bucharest, Department of Mathematics, September 1991 - September 1995.

Research Interests

- Foundations of data science; statistical machine learning theory.

- High dimensional statistical inference in parametric and non-parametric models: regression, covariance estimation, graphical models, model-based clustering (hard clustering and overlapping clustering), cluster-based models, latent variable models, topic models, networks, new perspectives in prediction problems, statistical aspects regarding Wasserstein-like distances between mixture distributions; statistical analysis of optimal transport plans between structured distributions; theory and methods for discrete mixture models; methods and theory for the interpretable evaluation of LLM / AI output.

- Applications to Systems Biology and Immunology, Text Analysis and other areas.

Editorship

- Series Editor: Monographs on Statistics and Applied Probability (The Green Series), Chapman & Hall/CRC (2009 - ongoing).

- Associate Editor: The Annals of Statistics (2007 - 2010; 2016 - ongoing).

- Associate Editor: The Electronic Journal of Statistics, EJS (2007 - 2022). Guest Editor for a special 2016 issue on High dimensional modeling.

- Associate Editor: The Journal of the American Statistical Association, JASA - Theory and Methods, (2011 - 2017).

- Associate Editor: Bernoulli (2010 - 2016).

- Associate Editor: Journal of the Royal Statistical Society (JRSS)-B (2012- 2014)

- Associate Editor: The Annals of Applied Statistics (2011 - 2012).

- Associate Editor: The International Journal of Biostatistics, Berkeley Press (2007 - 2012).

Grants

- NSF-DMS award 2515156, PI. *Softmax mixture ensembles for leveraging LLM output,* 2025 - 2028.

- NIH subcontract *Modern machine learning methods for inter and intra-species co-evolution,* University of Pittsburgh Medical School, 2023.

- NSF-DMS award 2210563, PI. *Statistical Optimal Transport for High Dimensional Mixtures,* 2022 - 2025. Collaborative research with Jon Niles-Weed, Courant Math Institute, NYU.

- NIH subcontract *Latent space clustering for biological discovery,* via University of Pittsburgh Medical School, 2021.

- NSF-DMS award 2015195, PI, *Learning from hidden signatures in high dimensional models,* 2020 - 2023.

- NSF-DMS award 1712709, PI, *Statistical foundations of model-based variable clustering,* 2017 - 2020.

- NSF-DMS award 1310119, co-PI, *Estimation of high dimensional matrices of low effective rank with applications to structural copula models,* 2013 - 2016.

- NSF-DMS award 10007444, PI, *Matrix estimation under rank constraints for complete and incomplete noisy data,* 2010 - 2013.

- (NSF-DMS) award 0925275, PI, conference grant, *From Probability to Statistics and Back: High-Dimensional Models and Processes,* Seattle, July 28 - 31, 2010.

- (NSF-DMS) award 0706829, co-PI: *Sparsity oracle inequalities via l1 regularization in non-parametric models,* 2007 - 2010.

- (NSF-DMS) 0406049, PI: *Curve aggregation and classification,* 2004 - 2007.

Invited Scholar/Key-Note/Honors/Awards

- The 2026 Puri Memorial Lecturer, April 2026, Purdue University, USA.

- Medallion Award and IMS Medallion Lecture, August 2025, JSM/IMS Nashville.

- Cornell Bowers Research Excellence Award, 2024.

- Key Note at Minghui Yu Memorial Conference, Columbia University, April 2024.

- Key Note talk at NeurIPS, New Orleans, December 2023.

- Invited Visiting Professor, Munich Institute of Technology (TUM), Dept. of Mathematics, Spring 2022.

- Invited Visiting Professor at Columbia University, Department of Statistics, Spring 2022.

- Invited Visiting Researcher, Simons Institute, semester on Computational Complexity of Statistical Inference, November-December 2021.

- Invited Visiting Professor at the Courant Institute and Center for Data Science, NYU, 2021 - 2022.

- Elected Fellow of the Institute of Mathematical Statistics (IMS), 2017
  (For foundational work in model selection and aggregation in parametric, semi-parametric and non-parametric models).

- Key Note at the New Researchers Pre-meeting of the IMS/Bernoulli World Congress for Probability and Statistics, Toronto, 2016.

- Key Note at the Theory of Big Data 2, London, January 2016.

- Fellow of the Newton Institute, Cambridge University, UK, 2008, 2016, 2018.

- SAMSI fellow, Spring 2014.

- Invited Professor, Université de Paris VI and CREST, France, 2003, 2009 (Host: Alexandre Tsybakov.).

- Invited Visiting Scholar, The John Hopkins School of Public Health, Jan/Feb 2008.

- Invited Researcher, L'Institut Henri Poincaré, Paris, France, April - July 2001. (Host: Lucien Birgé.)

- The Centre National de la Recherche Scientifique (CNRS) Award in Statistics for young researchers, 2001 (CNRS is the French equivalent of the National Science Foundation.)

Service to the academic community at large, diversity and education

- Institute of Mathematical Statistics, Committee on Nominations, 2022-2025.

- Institute of Mathematical Statistics (IMS) Council Member (elected), 2014 - 2017.

- IMS representative to the selection committee of the COPSS F.N. David Award for exceptional women in Applied Statistics, 2014 - 2018.

- NSF - DMS Panel member: 2008, 2009, 2011, 2012, 2014, 2015, 2017, 2018, 2020, 2021, 2023, 2024, 2025.

- ERC (European Research Council) panel member: 2024.

Conference and/or Program co-organizer

- *Statistical Foundations of Generative Modeling: Theory, Evaluation, Applications*, Institute for Mathematical and Statistical Innovation (IMSI), Chicago, USA, October 2026.

- *Statistics and Learning Theory in the Era of Artificial Intelligence*, Mathematisches Forschungsinstitut Oberwolfach (MFO), Germany, June/July 2024.

- *Re-thinking high dimensional statistics*, Mathematisches Forschungsinstitut Oberwolfach (MFO), Germany, May 2022.

- *CIS @ 20*, A conference celebrating 20 years since the creation of the college of Computing and Information Science (CIS) at Cornell University, October 2019.

- *Matrix Estimation Meets Statistical Network Analysis: Extracting low-dimensional structures in high dimension,* Mathematisches Forschungsinstitut Oberwolfach (MFO), Germany, 2018.

- *Theory of Big Data*, London, UK, 2017.

- *Cornell Day of Statistics*, Cornell University, September 2015, 2016.

- *LDHD Workshop. Statistical Inference in Sparse High-dimensional Models: theoretical and computational challenges*, SAMSI, North Carolina, February 2014.

- Summer school organizer for the LDHD SAMSI Program, SAMSI, North Carolina, August 2013.

- Program leader for the 2013 - 2014 SAMSI Program on *Low-dimensional Structure in High-dimensional Systems (LDHD)*.

- *From Probability to Statistics and Back: High-Dimensional Models and Processes*, Seattle, Washington, July 2010.

- *Nonparametrics Statistics*, International Conference,Tallahassee, Florida, Jan 2003.

Publication Record

Published, Accepted and Preprint Papers

(In inverse chronological order)

- F. Bunea, T. Manole, A. Wang and J. Thickstun; *Wasserstein-based metrics for evaluating the quality and diversity of LLM output*, 2026, *Preprint.*

- F. Bunea, X. Bing, M. Wegkamp, W. Wu; *Detecting and leveraging quasi-homogeneity for optimal estimation in softmax mixture ensembles and mixtures-of-experts*, 2026, *Preprint.*

- S. Liu, F. Bunea, J. Niles-Weed; *Beyond entropic regularization: Debiased Gaussian estimators for discrete optimal transport and general linear programs*, 2025, https://arxiv.org/abs/2505.04312

- Xin Bing, Florentina Bunea, Jon Niles-Weed and Marten Wegkamp; *Learning Large Soft-max mixtures with warm start EM*, (2024+) on https://arxiv.org/abs/2409.09903

- Florentina Bunea, Arnak Dalalyan, Robert Nowak, Sara van de Geer, *Statistics and Learning Theory in the Era of Artificial Intelligence*, **Oberwolfach Reports**, 2024.

- J. Rahimikollu, H. 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**, 2024, https://doi.org/10.1038/s41592-024-02175-z

- Xin Bing, Florentina Bunea and Jon Niles-Weed, *Estimation and Inference for the Wasserstein Distance between mixing measures in topic models*, **Bernoulli**, 2024, (forthcoming).

- Shuyu Liu, Florentina Bunea, and Jon Niles-Weed; *Asymptotic confidence sets for random linear programs*, **COLT** (36th Conference on Learning Theory) Proceedings of Machine Learning Research vol 195:1–22, 2023.

- Xin Bing, Florentina Bunea and Marten Wegkamp. *Detecting approximate replicate components of a high-dimensional random vector with latent structure*, **Bernoulli**, 2023, Vol. 29, pages 1368-1392 https://arxiv.org/abs/2010.02288.

- 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*, **Annals of Statistics**, 2022, Vol 50, pages 3307 - 3333.

- Florentina Bunea, Seth Strimas-Mackey, Marten Wegkamp, *Interpolating predictors in high-dimensional factor regression*, **Journal of Machine Learning Research**, Vol. 23, pages 1-60, 2022; https://arxiv.org/abs/2002.02525

- Xin Bing, Tyler Lovelace, Florentina Bunea, Marten Wegkamp, Harinder Singh, Panayiotis V Benos, Jishnu Das. *Essential Regression: a generalizable framework for inferring causal latent factors from multi-omic human datasets*, **Patterns (Cell Press)**, 2022.

- Xin Bing, Florentina Bunea, Seth Strimas-Mackey and Marten Wegkamp. *Prediction in latent factor regression: Adaptive PCR and beyond*, **Journal of Machine Learning Research**, 2021, Vol. 22.

- Xin Bing, Florentina Bunea, Marten Wegkamp, *Inference in latent factor regression with clusterable features*, **Bernoulli**; 2022, Vol 28; *Winner of the Institute of Mathematical Statistics (IMS) Larry Brown best paper award.*

- Florentina Bunea; Robert Nowak; Alexandre Tsybakov. *Re-thinking High-dimensional Mathematical Statistics*; **Oberwolfach Report** (25), 2022, DOI: 10.4171/OWR/2022/25

- Xin Bing, Florentina Bunea, Marten Wegkamp, *Optimal estimation of sparse topic models*, **Journal of Machine Learning Research**, 21 (177), 2020.

- 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), 2020.

- Xin Bing, Florentina Bunea, Yang Ning, Marten Wegkamp, *Adaptive Estimation in Structured Factor Models with Applications to Overlapping Clustering*, **The Annals of Statistics**, 48 (4), 2020.

- C. Eisenach, F. Bunea, Y. Ning and C. Dinicu, *High Dimensional Inference in Cluster-Based Graphical Models*, https://arxiv.org/abs/1806.05139, **Journal of Machine Learning Research**, 21 (53), 2020.

- Florentina Bunea, Christophe Giraud, Martin Royer, Nicolas Verzelen, and Xi Luo, *Model-assisted variable clustering: minimax-optimal recovery and algorithms*, **The Annals of Statistics**, 2020, Vol. 48, pp. 111 - 137

- Xin Bing, Florentina Bunea, Martin Royer, Jishnu Das, *Latent model-based clustering for biological discovery*, **iScience**, 125 - 135, 2019.

- Bunea Florentina, Rohde Angelika, Wolfe Patrick, Zhou Harrison, *Matrix Estimation Meets Statistical Network Analysis: Extracting low-dimensional structures in high dimension*, **Oberwolfach Reports** 15 (2018), 1745-1783; DOI: 10.4171/OWR/2018/29

- F. Bunea, C. Giraud, M. Royer and N. Verzelen, PECOK: a convex optimization approach to variable clustering, https://arxiv.org/abs/1606.05100 (2016, 2017).

- F.Bunea, C. Giraud and X. Luo, Minimax optimal variable clustering in G-models via CORD https://arxiv.org/pdf/1508.01939.pdf (2015, 2017).

- J. Bien, F. Bunea and L. Xiao, Convex banding of the covariance matrix, **Journal of American Statistical Association**, Volume 111, 834-845, 2016.

- J. Das, K.M. Gayvert, F. Bunea, M. Wegkamp and H. Yu: *ENCAPP: elastic-net-based prognosis prediction and biomarker discovery for human cancers*, **BMC Genomics**, Vol 16, 2015.

- Florentina Bunea and Luo Xiao, On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA, **Bernoulli**, Volume 21, 1200-1230, 2015.

- L. Xiao and F. Bunea: On the theoretical and practical merits of the banding estimator for large covariance matrices, http://arxiv.org/abs/1402.0844, **EJS** (2014).

- Florentina Bunea, Johannes Lederer and Yiyuan She, The Group Square-Root Lasso: Theoretical Properties and Fast Algorithms, **IEEE-Information Theory**, Vol 60, pages 1313 - 1325, (2014)

- Edited volume. *From Probability to Statistics and Back: High-Dimensional Models and Processes, A Festschrift in Honor of Jon Wellner*; **IMS Collections**, Volume 9, 2012, Banerjee, M., Bunea, F., Huang, J., Koltchinskii, V., and Maathuis, M. H., Editors

- Florentina Bunea, Yiyuan She and Marten Wegkamp *Joint variable and rank selection for parsimonious estimation of high dimensional matrices.* **The Annals of Statistics**, Vol 40, 2359-2388, (2012).

- Florentina Bunea, Yiyuan She and Marten Wegkamp *Optimal selection of reduced rank estimators of high dimensional matrices.* **The Annals of Statistics**, Vol 39, 1282- 1309, (2011).

- Florentina Bunea, Andrada Ivanescu and Marten Wegkamp *Adaptive inference for the mean of a Gaussian process in functional data.* **J.R. Statist. Soc. B**, Vol 73 (4), 531 - 558, (2011).

- Florentina Bunea, Yiyuan She, Hernando Ombao et al. *Penalized Least Squares Regression Methods and Applications to Neuroimaging.* **NeuroImage**, Vol. 55, 1519 - 1527 (2011).

- Florentina Bunea, Alexandre Tsybakov, Marten Wegkamp and Adrian Barbu *SPADES and mixture models*, **The Annals of Statistics**, Vol 38, 2525 - 2558, (2010).

- Florentina Bunea and Adrian Barbu *Dimension reduction and variable selection in case-control studies via regularized likelihood optimization*, **Electronic Journal of Statistics**, Vol. 3, 1257 - 1287, (2009)

- Florentina Bunea *Honest variable selection in linear and logistic regression models via ℓ_1 and $\ell_1 + \ell_2$ penalization*, **Electronic Journal of Statistics**, Vol 2, pp 1153 - 1194, (2008).

- Florentina Bunea and Andrew Nobel *Online prediction algorithms for aggregation of arbitrary estimators of a conditional mean.* **IEEE Transactions on Information Theory**, Vol 54 (4), pp 1725 - 1735, (2008).

- Florentina Bunea *Consistent selection via the Lasso for high dimensional approximating models.* **IMS Lecture Notes-Monograph Series**, Vol 123, pp. 123 - 137, (2008).

- Florentina Bunea, Alexandre Tsybakov and Marten Wegkamp *Sparsity oracle inequalities for the lasso.* **Electronic Journal of Statistics,** pp. 169 - 194, (2007).

- Florentina Bunea, Alexandre Tsybakov and Marten Wegkamp *Sparse density estimation with l1 penalties.* Annual Conference on Learning Theory, **COLT: Lecture Notes in Artificial Intelligence**, pp. 530 - 544, Springer, (2007).

- Florentina Bunea, Alexandre Tsybakov and Marten Wegkamp *Aggregation for Gaussian regression.* **Annals of Statistics**, Vol 35, 1674 - 1697, (2007).

- Florentina Bunea, Alexandre Tsybakov and Marten Wegkamp *Aggregation and sparsity via l_1 penalized least squares.* **COLT: Lecture Notes in Artificial Intelligence**, Springer, pp. 379 - 391, (2006).

- Florentina Bunea, Marten Wegkamp and Anna Auguste *Consistent covariate selection in high dimensional regression via multiple testing.* **Journal of Statistical Planning and Inference**, Elsevier, Vol 136, pp 4349-4364, (2006).

- Florentina Bunea, Hernando Ombao and Anna Auguste *Minimax adaptive spectral estimation from an ensemble of signals.* **IEEE Transactions on Signal Processing**, Vol. 54, pp. 2865 - 2874, (2006).

- Gerard Biau, Florentina Bunea and Marten Wegkamp *Function Classification in Hilbert Spaces.* **IEEE Transactions on Information Theory,** Vol. 51, pp 2163 - 2172, (2005).

- Florentina Bunea and Ian McKeague *Covariate selection for semiparametric hazard function regression models.* **The Journal of Multivariate Analysis,** Vol. 92, pp. 186 - 204, (2005).

- Florentina Bunea *Consistent covariate selection and postmodel selection inference in semiparametric regression.* **The Annals of Statistics,** Vol. 32, No. 3, pp. 898 - 927, (2004).

- Florentina Bunea and Marten Wegkamp *Two-stage model selection procedures in partially linear regression.* **The Canadian Journal of Statistics,** Vol. 32, pp. 105 - 118, (2004).

- Florentina Bunea and Marten Wegkamp. *A Note on Penalized Minimum Distance Estimation in Nonparametric Regression.* **The Canadian Journal of Statistics**, Vol. 31, No 3, pp. 267-274, (2003).

- Florentina Bunea and Julian Besag *Markov Chain Monte Carlo in $I \times J \times K$ contingency tables*, **Fields Institute Communications**, AMS, Providence, Rhode Island. Vol. 26, pp. 25 - 36, (2000).

- Florentina Bunea *Geometrical restrictions in self deconvolution*, The Annals of the University of Bucharest, Mathematics and Informatics Series, Vol. 44, pp. 41-50, (1995).

Invited conference session organizer

- Invited chair and organizer of the session "High Dimensional Data Inference" IMS/Bernoulli World Congress of Probability and Statistics, Seoul, South Korea, 2021.

- Invited chair and organizer of the session "High Dimensional Inference in Structured Models", the European Meeting of Statisticians, Palermo, Italy, 2019.

- Invited chair and organizer of the session "Model selection in high dimensions", 9th World Congress of Probability and Statistics, Fields Institute, Toronto, July 11-15, 2016.

- Invited chair and organizer of the session "New directions in big data theory", 3rd conference of the International Society for Non-Parametric Statistics (ISNPS), Avignon, France 11-16 June 2016.

- Invited chair and organizer of the session "High-Dimensional Inference", the European Meeting of Statisticians, Amsterdam, 2015.

- Invited chair and organizer of the session "High-Dimensional Statistical Inference for Matrix Models", at WNAR, Colorado, June 2012.

- Invited chair and organizer of the session "High dimensional inference and matrix models", at the 1st conference of the International Society for Non-Parametric Statistics, Greece, June 2012.

- Program committee member and chair of the "Machine Learning" session for SRCOS (Southern Regional Conference on Statistics), Norfolk, Virginia, June 2010.

- Invited chair and organizer of the session "Recent results in theoretical machine learning" at the eighth ICSA International Conference: Frontiers of Interdisciplinary and Methodological Statistical Research, Guangzhou University, China, Dec 2010.

- Invited chair and organizer of the invited IMS session "Advances in Statistical Aggregation" at the Joint Statistical Meeting, Salt Lake City, August 2007.

- Invited chair and organizer of the invited IMS session "Semiparametric Inference in Practice" at the Joint Statistical Meeting, Seattle, August 2006.

- Invited chair and organizer of the invited IMS session "Aggregation in non-parametric models" at the Joint Statistical Meeting in Minneapolis, August 2005.

Invited conference and workshop talks

- JSM Annual Meeting, Boston, August, 2026.

- IMS Annual Meeting, Salzburg, Austria, July, 2026.

- The Puri Memorial Lecture, Purdue University, April 2026.

- Theoretical Foundations: From the Early Days of Neural Networks to the Modern Deep Learning Era, the Simons Institute, Berkeley, April 2026.

- 2025 IMS International Conference on Statistics and Data Science (ICSDS), Seville Spain, December 2025.

- Meetings in Mathematical Statistics, Centre International de Rencontres Mathématiques, Marseilles, France, December 2025.

- Mathematical Statistics in the Information Age - Statistical efficiency and computational tractability" Vienna, September 2025

- JSM, Medallion Lecture, August 2025.

- New challenges in high-dimensional statistics, Centre International de Rencontres Mathématiques, Marseilles, France, December 2024.

- IMS International Conference on Statistics and Data Science (ICSDS) Nice, France, December 2024.

- 11th World Congress in Probability and Statistics, a joint Bernoulli and IMS congress, Germany, August 2024.

- The Mathematics of Machine Learning, ETH, Zurich, Switzerland, June 18 - 21, 2024.

- Minghui Yu Memorial Conference, Columbia University, April 2024 (Keynote).

- Harnessing the power of latent structure models and modern Big Data learning, The Banff Institute, CA and the Institute for Advanced Study in Mathematics, Hangzhou, China (Zoom), December 2023.

- Optimal Transport and Machine Learning, NeurIPS 2023, December 2023, New Orleans (keynote).

- Data Science and Dependence, July 2023, The International Academic Forum, Heidelberg, Germany.

- Workshop on Statistical Network Analysis and Beyond, Alaska, June 2023.

- Annual Conference of the International Indian Statistical Association (IISA2023), Colorado, June 2023 (Special Invited Talk)

- Princeton University, Statistical Foundations of Data Science and their Applications, May 2023.

- Sampling, Transport, and Diffusions, Flatiron Institute, NYC, November 2022.

- The Joint Statistical Meeting, IMS invited sessions, Washington DC, August 2022.

- The Institute of Mathematical Statistics annual meting, London, UK, June 2022.

- Mathematical Methods of Modern Statistics 3, Marseille, France, June 2022.

- Statistics in the Big Data Era, the Simons Institute, June 2022.

- Oberwolfach meeting on "Re-thinking high-dimensional statistical inference", Oberwolfach, Germany, May 2022.

- The ICSA Applied Statistics Symposium, September 2021 (Virtual).

- The Joint Statistical Meeting, IMS invited sessions, Seattle, August 2021(Virtual).

- Oberwolfach meeting on Mathematical Foundations of Machine Learning, March 2021. (Virtual/Hybrid)

- Mathematical Methods of Modern Statistics 2, Marseille, France, June 2020 (virtual).

- Fourth Workshop on High Dimensional Asymptotics and Post-Selection Inference, St. Louis, August, 2019.

- Workshop on Machine Learning and Data Science, Columbia University, June 2019.

- Oberwolfach meeting on Statistical and Computational Aspects of Learning with Complex Structure, May 2019

- Joint Statistical Meeting, Vancouver, 2018.

- Isaac Newton Institute workshop: Future challenges in statistical scalability, Cambridge, UK, June 2018.

- Oberwolfach workshop Matrix estimation meets network analysis, Germany, June 2018.

- International workshop on New Aspects of Statistics, Financial Econometrics and Data Science, University of Chicago, May 2018.

- Oberwolfach workshop on Statistical Inference for Structured High-dimensional Models, Germany, March 2018.

- Meeting in Mathematical Statistics, Marseilles, France, 2017.

- Second Workshop on High Dimensional Asymptotics and Post-Selection Inference, St. Louis, 2017.

- Joint Statistical Meeting, Baltimore, 2017.

- Conference on Foundations of Computational Mathematics, Barcelona, Spain, 2017.

- Meeting in Mathematical Statistics: Advances in nonparametric and high-dimensional Statistics, Fréjus, France, 2016.

- Workshop on High Dimensional Asymptotics and Post-Selection Inference, St. Louis, 2016.

- JSM, Seattle, August 2015.

- Oberwolfach Workshop, Probabilistic Techniques in Modern Statistics, Oberwolfach, Germany, May 2015.

- American Institute of Mathematics (AIM), Palo Alto, CA, Workshop on Inference in high dimensional regression, 2015.

- Joint Statistical Meeting, Boston, August 2014.

- Dimension Reduction and High Dimensional Inference Workshop, University of Florida, 2014.

- International Symposium on Business and Industrial Statistics/Conference of the ASA Section on Statistical Learning and Data Mining, Durham, June 2014.

- Oberwolfach Workshop, Adaptive Statistical Inference, Oberwolfach, G ermany, March 2014.

- 3rd Princeton Day of Statistics Workshop, Oct. 19, 2012.

- International Chinese Statistics Association Applied Statistics Symposium, Boston, June 2012, *Simultaneous variable and rank selection for optimal estimation of high dimensional matrices.*

- High-dimensional problems in statistics workshop, ETH, Zurich, Switzerland, September 2011, *Simultaneous variable and rank selection for optimal estimation of high dimensional matrices.*

- Instantaneous Frequencies and Trends for Nonstationary Nonlinear Data, IMA, Minneapolis, September 2011, *Simultaneous variable and rank selection for optimal estimation of high dimensional matrices.*

- Sparse Statistics, Optimization and Machine Learning BIRS Workshop, Banff, Canada, January 2011, *Optimal selection of reduced rank estimators of high dimensional matrices.*

- The Applied Statistics Symposium of the International Chinese Statistical Association, Indianapolis, June 2010, *Adaptive Rank Penalized Estimators in Multivariate Regression.*

- Conference on Resampling Methods and High Dimensional Data, College Station, Texas, March 2010. *Optimal dimension reduction in high dimensional matrix models via the Rank Selection Criterion.*

- Sparsity in high dimensions conference, The Mathematical Institute at Oberwolfach, Germany, March 2009, *Model selection and sparsity in case control studies.*

- Understanding the New Statistics: Expanding Core Statistical Theory, Banff, Canada, September 2008. *Honest variable selection.*

- International Workshop on Sparsity in High Dimensional Statistics and Learning Theory, March 2008, Georgia Institute of Technology, Atlanta. *On ℓ_1 regularization in functional data and binary response regression models.*

- The Isaac Newton Institute, Cambridge University, UK, Programme on Statistical Theory and Methods for Complex, High-Dimensional Data, May 2008. *Non-asymptotic variable selection.*

- The Joint Statistical Meeting, the Institute of Mathematical Statistics section, Salt Lake City, August 2007. *Sparse oracle inequalities for the Lasso.*

- The Annual Conference on Learning Theory (COLT), San Diego, June 2007. *Identifying components in sparse mixtures: the Spades estimator.*

- Statistical and Probabilistic Methods in Model Selection, The Mathematical Institute at Oberwolfach, Germany, October 2005. *On-line prediction algorithms in aggregation.*

- Kickoff Workshop of the SAMSI program on High Dimensional Inference and Random Matrices, September 2006.

- The Joint Statistical Meeting, the Institute of Mathematical Statistics Section, Seattle, August 2006.

- The Joint Statistical Meeting, the Institute of Mathematical Statistics section, Minneapolis, August 2005. *Aggregation for regression learning.*

- SRCOS/ASA Summer Research Conference, Virginia Tech, June 2004. *Oracle inequalities in aggregation.*

- International Workshop on Applied Probability, Greece, March 2004. *One and multiple-sample aggregation in nonparametric regression models.*

- New Researchers Conference, Davis, CA, July 2003. *Consistent model selection via the FDR procedure.*

- Nonparametrics Research Conference, Tallahassee, FL, January 2003. *The consistency of the FDR procedure with applications to semiparametric models.*

- International conference on current advances and trends in nonparametric statistics, Crete, July 2002. *Post model selection inference in semiparametric regression models.*

- IMS meeting in Guanajuato, Mexico, 2000. *Partially linear regression.*

- SIAM Annual Meeting, Atlanta, 1999. *MCMC for multi-dimensional contingency tables.*

Invited Individual Seminars

- *Germany*: Technical University of Munich, Department of Mathematics, May 2022.

- *Canada*: University of Toronto (2002), University of British Columbia (2000).

- *France*: Université de Paris VI and CREST (Paris, 2009), Université de Paris VI (Paris, 2003), ENSAI (Rennes, 2001), L'Institut Henri Poincaré (Paris, 2001).

- *United Kingdom*: Bath University (2001), Newton Institute, Programme on Statistical Theory and Methods for Complex High-dimensional Data, Cambridge (2008).

- *Switzerland*: The Swiss Institute of Technology (ETH), Zurich (2008).

- *USA*

  2025: Harvard University

  2024: Yale University; Columbia University.

  2022: Columbia University; Princeton University; University of Chicago Booth.

  2021: University of Michigan, NYU (Currant Institute), Cornell Tech., Cornell-Ithaca.

  2019: University of Washington, Temple University, Harvard University.

  2017: Columbia

  2016: Princeton

  2014: Brown

  2010, 2012: Cornell University

  2008: University of Chicago.

  2007: Carnegie Mellon University, Johns Hopkins University, University of Minnesota, Rutgers University.

  2006: Texas A & M University, North Carolina State University.

  2005: University of Georgia, University of Illinois at Urbana-Champaign.

  2004: The Pennsylvania State University, Duke University.

  2002: University of Florida.

  2000/2001: Yale University, University of Pennsylvania, Stanford University, University of British Columbia.

Ph.D. Student supervision

Former Ph. D. students: Anna Auguste, Andrada Ivanescu, Shuva Gupta, Jennifer Geis, Claudiu Dinicu, Martin Royer, Mike Bing, Seth Strimas-Mackey.

Recent Ph.D committees/collaborations/visiting: Amanda Wang (CS), Xinyi Wang (EE), Jerry Chee (CS), Daniel Lee (Statistics); Shuyu Liu (Courant Institute, NYU); Nayel Betache (ENSAE, France); Ruoqi Zhang (Mathematics).

Postodocs; Visiting post-docs

Kangjie Zhou, Zong Shang (2026); Nayel Betache, 2024-2025, Johannes Lederer, 2013 (currently at the University of Bochum, Germany), Eugen Parcalabelu, 2019 (currently at Universite Louvaine La Neuve, Belgium), Merle Behr, 2019 (currently at the University Regensburg, Germany).

Courses (Sample)

- Statistics through examples

- Introduction to Probability

- Mathematical Statistics

- Probability Models and Inference

- Probability and Measure

- Probability Theory

- Topics in Stochastic Processes

- Advanced Topics in Probability and Statistics: Statistical optimal transport for high dimensional mixtures

- Advanced Topics in Probability and Statistics: Model Selection

- Advanced Topics in Probability and Statistics: Semiparametric models and empirical processes

- Advanced Topics in Probability and Statistics: Aggregation in non-parametric models

- Advanced Topics in Probability and Statistics: Model selection and aggregation in high dimensions

- Advanced Topics in Probability and Statistics: Inference in large scale problems