

Florentina Bunea

Department of Statistical 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.
Dissertation: A Model Selection Approach to Partially Linear Regression.
Advisor: Jon A. Wellner.
- M.S. in Mathematics, First Class, University of Bucharest, Romania, 1991.
- B.S. in Mathematics, First Class, University of Bucharest, Romania, 1989.

Employment

- **Professor:** Cornell University, Department of Statistical Science, 2011 -
- **Faculty member of the Center of Applied Mathematics (CAM),** Cornell University, 2011-
- **Assistant/Associate Professor:** Florida State University, Department of Statistics, 2000 - 2011.
- **Research/Teaching Assistant :** University of Washington, Department of Statistics and the National Research. September 1995 - August 2000.
- **Assistant Professor:** University Politehnica Bucharest, Department of Mathematics (Romania). September 1991 - September 1995.

Research Interests

- Methods and theory for model selection and aggregation in high dimensions. High dimensional inference. Nonparametrics. Post-model selection inference.
 - High dimensional random matrix models: matrix regression, covariance estimation, clustering, graphical models, clustering, networks.
 - Dimension reduction methods.
 - Adaptation to unknown sparsity in high dimensions.
- The theoretical analysis of computationally efficient procedures in high dimensional statistics. Applications to Neuroscience, Genetics and Medicine.

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: JASA (2011 - ongoing).
- Associate Editor: Bernoulli (2010 - ongoing).
- Associate Editor: The Electronic Journal of Statistics, IMS (2007 - ongoing); Guest Editor for a special 2016 issue on "High dimensional modeling".
- Associate Editor: Journal of the Royal Statistical Society, JRSS-B, (2012 - June 2014).
- Associate Editor: The Annals of Applied Statistics (2011 - 2012).
- Associate Editor: The International Journal of Biostatistics, Berkeley Press (2007 - 2012).

Service to the academic community at large, diversity and education

- Institute of Mathematical Statistics (IMS) Council member (elected), 2014 -2017.
- IMS representative to the selection committee of the COPSS F.N. David Award, 2014-2018.
- NSF - DMS Panel member, January 2008, 2009, 2011, 2012, 2014, 2015.
- Member of the advisory council of the FSU Women in Mathematics, Science and Engineering Program (2001 - 2010).
- Member of the FSU Council on Research and Creativity, an internal grant awarding unit (2006 - 2008).

Grants

- 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 l_1 regularization in non-parametric models*, 2007 - 2010.
- (NSF-DMS) 0406049, PI: *Curve aggregation and classification*, 2004 - 2007.

Publication Record

Published, Accepted and Preprint Papers

(In inverse chronological order)

- F. Bunea, C. Eisenach, Y. Ning et al. *Cluster-based graphical models*, Preprint, 2016.
- F. Bunea, Y. Ning, M. Wegkamp *Overlapping variable clustering*, Preprint, 2016.
- F. Bunea, C. Giraud, M. Royer and N. Verzelen, *PECOK: a convex optimization approach to variable clustering*, <https://arxiv.org/abs/1606.05100>, under revision with the Annals of Statistics, 2016.
- F. Bunea, C. Giraud and X. Luo, *Minimax optimal variable clustering in G-models via CORD* (2015+), <https://arxiv.org/pdf/1508.01939.pdf>, under revision with the Annals of Statistics, (2015+)
- J. Bien, F. Bunea and L. Xiao, *Convex banding of the covariance matrix*, Journal of American Statistical Association, Volume 111, 834-845, 2016.
- *ENCAPP: elastic-net-based prognosis prediction and biomarker discovery for human cancers*, J. Das, K.M. Gayvert, F. Bunea, M. Wegkamp and H. Yu; 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.
- *On the theoretical and practical merits of the banding estimator for large covariance matrices*, <http://arxiv.org/abs/1402.0844>, L. Xiao and F. Bunea, Accepted by 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*. Journal of the Royal Statistical Society 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 l_1 penalties*. Annual Conference on Learning Theory, 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*. 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 University of Bucharest, Mathematics and Informatics Series, Vol. 44, pp. 41-50, (1995).

Invited Fellow/Scholar/Honors

- Fellow of the Newton Institute, Cambridge University, UK, 2008, 2016, 2018.
- SAMSI fellow, Spring 2014.
- Invited Professor, Université de Paris VI, June 2009.
- Invited Visiting Scholar, The John Hopkins School of Public Health, Jan/Feb 2008.
- Invited Research Professor, Université de Paris VI (Jussieu) and CREST, France, May 2003 (Host: Alexandre Tsybakov.)
- Invited Researcher, L'Institut Henri Poincaré, Paris, France, April/July 2001. (Host: Lucien Birgé.)
- Invited Researcher, Université Rennes II, France, July 2001. (Host: Marian Hristache.)
- 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.)
- The FSU First Year Assistant Professor Award, 2001.

Student supervision

Current Ph.D students: Martin Royer (jointly with Christophe Giraud), Claudiu Dinicu.

Conference and/or Program co-organizer

- Matrix Estimation Meets Statistical Network Analysis: Extracting low-dimensional structures in high dimension, Mathematisches Forschungsinstitut Oberwolfach (MFO Oberwolfach), Germany, 2018.
- Theory of Big Data Conference, London, UK, 2017.
- Cornell Day of Statistics, September 2015, 2016.
- LDHD Workshop: Statistical Inference in Sparse High-dimensional Models: theoretical and computational challenges, February 2014.
- Summer school organizer for the LDHD SAMSI Program, August 2013.
- Program leader for the 2013 - 2014 SAMSI Program on Low-dimensional Structure in High-dimensional Systems (LDHD).

- The international conference *From Probability to Statistics and Back: High-Dimensional Models and Processes*, Seattle, July 2010.
- Nonparametrics Research International Conference, Tallahassee, FL, Jan 2003.

Conference session organizer

- 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 Talks at Conferences and Workshops

- 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.
- New Researchers Pre-meeting of the IMS/Bernoulli World Congress for Probability and Statistics, Toronto, 2016, Plenary speaker.
- Theory of Big Data 2, London, January 2016, Plenary speaker.
- 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.
- Oberwolfach Workshop, Adaptive Statistical Inference, Oberwolfach, Germany, March 2014.
- 3rd Princeton Day of Statistics Workshop, Oct. 19, 2012.
- International Chinese Statistics Association 2012 Applied Statistics Symposium will take place in Boston, June 23-June 26, *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.*
- The Isaac Newton Institute, Cambridge University, UK, Programme on Statistical Theory and Methods for Complex, High-Dimensional Data, May 2008. *Non-asymptotic 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 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.*

- The Annual Conference on Learning Theory (COLT), Pittsburgh, June 2006.
- 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.*
- Statistical and Probabilistic Methods in Model Selection, The Mathematical Institute at Oberwolfach, Germany, October 2005. *On-line prediction algorithms in aggregation.*
- International Workshop on Applied Probability, Greece, March 2004. *One and multiple-sample aggregation in nonparametric regression models.*
- SRCOS/ASA Summer Research Conference, Virginia Tech, June 2004. *Oracle inequalities in aggregation.*
- Nonparametrics Research Conference, Tallahassee, FL, January 2003. *The consistency of the FDR procedure with applications to semiparametric models.*
- New Researchers Conference, Davis, CA, July 2003. *Consistent model selection via the FDR procedure.*
- International conference on current advances and trends in nonparametric statistics, Crete, July 2002. *Post model selection inference in semiparametric regression models.*
- 1999 SIAM Annual Meeting, Atlanta. *MCMC for multi-dimensional contingency tables.*

Invited Individual Seminars

- *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*
 2016: Princeton, Columbia
 2014: Brown
 2010, 2012: Cornell University
 2011: Princeton
 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.

2004: The Pennsylvania State University, Duke University.

2002: University of Florida.

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

Professional Affiliations

- Member of the Institute of Mathematical Statistics.

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: 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: vector and matrix models.