Andreas Buja

Senior Research Scientist, Flatiron Institute
Liem Sioe Liong/First Pacific Company Professor Emeritus, Wharton

Addresses:

Center for Computational Mathematics, Flatiron Institute, Simons Foundation
162 Fifth Avenue, New York, NY 10010
https://www.simonsfoundation.org/flatiron/center-for-computational-mathematics/

Statistics Department, The Wharton School, University of Pennsylvania
Philadelphia, PA 19104-6340
http://stat.wharton.upenn.edu/~buja/

Education:

- Ph.D. in Mathematics/Statistics (1980)
    Swiss Federal Institute of Technology (ETHZ, Zurich)
    Advisors: F. Hampel, P. J. Huber, H. Foellmer
- Diploma in Mathematics (ETH, 1975, won the ETH award medal)
- Systems Course (Digital Equipment Corporation, Marlboro, MA, 1977)

Employment:

- Senior Research Scientist,
    Flatiron Institute, Simons Foundation,
    Jan 2020–
- Liem Sioe Liong/First Pacific Company Professor,
    Statistics Department, The Wharton School, University of Pennsylvania,
    July 2003–June 2021, Emeritus as of July 2021.
- Professor,
    Statistics Department, The Wharton School, University of Pennsylvania,
    Jan 2002– June 2003.
- Technology Consultant (= job rank; not consulting),
    AT&T Labs – Research, Florham Park, New Jersey, 1996–Jan 2002.
- Member of Technical Staff,
    AT&T Bell Laboratories, Murray Hill, NJ, 1994–1996.
- Member of Technical Staff,
    Bell Communications Research, Morristown, NJ, 1987–1994.
- Associate,
    Salomon Brothers Inc, New York, NY, 1987 (first half).

- Associate Professor (with tenure),
    Department of Statistics, University of Washington, Seattle, 1987–1988.
- Assistant Professor, Department of Statistics,
    University of Washington, Seattle, 1982–1987.
- Research Associate,
    ETH Zürich and Children's Hospital, Zürich, 1980–1981.
- Research Assistant,
    Swiss Federal Institute of Technology, Zürich, 1975–80.

Visiting Appointments

- Visiting Faculty (Ph.D. level summer teaching),
    Stanford University, Summer 1997.
- Consultant, Bell Communications Research,
    Morristown, NJ, 1986 (7 months).
- Consultant, Bell Laboratories,
    Murray Hill, NJ, 1985 (1 week).
- Visiting Assistant Professor, Department of Statistics,
    Visiting Scientist, Stanford Linear Accelerator Center,
    Stanford University, Stanford, 1981–1982.
- Visiting Lecturer,
    Stanford University, Stanford, California, Summer 1980.

Supervised Ph.D. Students (*as principal or co-principal advisor*):

- Arun Kumar Kuchibhotla (2020, Wharton, UPenn),
    Co-advisor with Larry D. Brown.
- Xin-Lu Tan (2017, Wharton, UPenn),
    Citadel (Chicago),
    Co-advisor with Zongming Ma.
- Dan Yang (2012, Wharton, UPenn),
    Rutgers University,
    Co-advisor with Zongming Ma.
- Shaun Lysen (2009, Wharton, UPenn),
    Google.
- Kartik Ghia (2007, Wharton, UPenn),
    J.P. Morgan.
- Lisha Chen (2006, Wharton, UPenn),
    Yale University.
- Yi Shen (2005, Wharton, UPenn),
    City Group, St. Louis, MO.

- Young-Seop Lee (1999, Rutgers University, New Brunswick),
    Dongguk University, Korea,
    Co-advisor with J. Cabrera.
- Harry Leung (1998, Rutgers University, New Brunswick),
    Baptist University, Hong Kong,
    Co-advisor with J. Cabrera.
- Dianne Cook (1992, Rutgers University, New Brunswick),
    Iowa State University, Ames, IA,
    Co-advisor with J. Cabrera.
- Catherine Hurley (1987, University of Washington, Seattle),
    National University of Ireland, Maynooth, Ireland,
    Co-Advisor with J. A. McDonald.
- Deborah Donnell (1987, University of Washington, Seattle),
    Fred Hutchinson Cancer Research Center, Seattle,
    Co-advisor with W. Stuetzle.

Service to the Profession:

- Council Member, Institute of Mathematical Statistics (IMS) (2015–2018).
- External Review Committee: Programs in Statistics, Actuarial Science, Statistics and Biostatistics, Statistics and Computing; University of Waterloo, Ontario (June 2009).
- External Review Committee: Department of Statistics, University of California, Berkeley (April 2006).
- Advisory Editor, *Journal of Computational and Graphical Statistics (JCGS)* (2000–2010).
- Associate Editor, *Journal of the American Statistical Association (JASA)* (2005).
- Reviewer, KDD 2004.
- Managing Editor, *Journal of Computational and Graphical Statistics (JCGS)* (1997–2000).
- Co-Chair/Co-Organizer, with D.F. Swayne and M. Theus,
    "Workshop on Data Visualization," sponsored by AT&T Labs,
    Drew University, Madison, NJ (summer 1997).
- Ad hoc committee on electronic publication of journals,
    American Statistical Association (1997).
- Reviewer, *Visualization*, IEEE Computer Society and ACM/SIGGRAPH, (1992–1998).
- Associate Editor, *Journal of Computational and Graphical Statistics* (1990–1997).
- Associate Editor, *The Annals of Statistics* (1989–1992).
- Program Committee, IMS, Joint Statistical Meetings, Boston (1992).
- Program Committee, *Computing Science and Statistics*, 23, Seattle (1991).
- Ad Hoc Committee, IMS representative, on founding of a new journal:
    Journal of Computational and Graphical Statistics (1990).

- Co-Chair/Co-Organizer, with P.A. Tukey, W. Stahel, IMA 1989 Summer Program, "Computational Statistics, Robustness and Diagnostics," U. of Minnesota (1989).
- Co-Chair/Co-Organizer, with W. Stuetzle, "Graphics, Human Interfaces, and Programming Paradigms for Data Analysis," AMS-SIAM-IMS Summer Research Conference, U. of California, Santa Cruz (1986).
- Associate Editor, *SIAM Journal on Scientific and Statistical Computing* (1983–1985).

Honors:

- Keynote speaker, Classification Society Conference, Milwaukee (2013).
- Fellow, Institute of Mathematical Statistics (2006).
- Keynote speaker, SIAM Conference on Datamining (SDM06) (April 2006).
- IMS Medallion lecture, Joint Statistical Meetings, New York (2002).
- Keynote speaker, European Meeting of the Psychometric Society, Leiden (1995).
- Fellow, American Statistical Association (1994).
- Award Medal for diploma thesis in mathematics, Swiss Federal Institute of Technology (ETH Zurich, 1975).

Special Status Articles:

- "Models as Approximations", Part I and II, *Statistical Science*, 523–620 (2019), with 9 discussions and a rejoinder.

    - "Models as Approximations I: Consequences Illustrated with Linear Regression," Buja, A., Berk, R., Brown, L., George, E., Pitkin, E., Traskin, M., Zhao, L., Zhang, K. (2019).

    - "Models as Approximations, Part II: A General Theory of Model-Robust Regression," Buja, A., Berk, R., Brown, L., George, E., Kuchibhotla A.K., Zhao, L. (2019).

    - "Models as Approximations — Rejoinder," Buja, A., Kuchibhotla A.K., Berk, R., George, E., Tchetgen Tchegen, E., Zhao, L. (2019).

- "Graphical Inference for Infovis," Wickham, H., Cook, D., Hofmann, H., and Buja, A., *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis'10)*, (Nov.-Dec.) 16 (6), 973–979 (2010).
  Best paper award.

- "Quasi-Darwinian Selection in Marketing Relationships," Eyuboglu, N., and Buja, A., *Journal of Marketing*, 17 (Oct), 48–62 (2007).
  Selected as the JM blog article of the issue, and finalist for JM's 2007 Harold H. Maynard Award.

- "Visualization Methodology for Multidimensional Scaling," Buja, A., and Swayne, D.F., *Journal of Classification* 19, 7–43 (2002).
  Special invited lead article.

- "Grand Tour and Projection Pursuit," Cook, D., Buja, A., and Cabrera, J., *Journal of Computational and Graphical Statistics* 4, 155–172 (1995).
  JCGS Invited Paper at the 1995 Symposium on the Interface, Pittsburg.

- "Prosection Views: Dimensional Inference through Sections and Projections," Furnas, G.W., and Buja, A., *Journal of Computational and Graphical Statistics*, 3, 323–385 (1994).
  JCGS Invited Paper at the 1994 Joint Statistical Meetings, Toronto, with discussion.

- "Analysis of Additive Dependencies and Concurvities Using Smallest Additive Principal Components," Donnell, D.J., Buja, A., and Stuetzle, W., *The Annals of Statistics*, 22, 1635–1673 (1994).
  Discussed article.

- "Linear Smoothers and the Additive Model," Buja, A., Hastie, T., and Tibshirani, R., *The Annals of Statistics*, 17, 453–555 (1989).
  Discussed lead article.

Refereed and Solicited Articles:

- "Semi-Supervised Linear Regression," Azriel D., Brown L.D., Sklar M., Berk R., Buja A. and Zhao, L. (2021).
  Journal of the American Statistical Association (to appear)

- "Valid Post-selection Inference in Model-free Linear Regression," Kuchibhotla A.K., Brown L.D., Buja, A., Cai, J., George, E.I., Zhao, L.H. (2020).
  The Annals of Statistics, Vol. 48, No. 5, 2953-2981.

- Kuchibhotla, A.K., Brown, L.D., Buja, A., George, E.I., Zhao, L.H. (2020). "A Model Free Perspective for Linear Regression: Uniform-in-model Bounds for Post Selection Inference." *Econometric Theory*.
  https://arxiv.org/abs/1802.05801, accepted.

- "Using Recursive Partitioning to Find and Estimate Heterogeneous Treatment Effects in Randomized Clinical Trials," Berk, R., Olson, M., Buja, A., Ouss, A. (2019).
  Journal of Quantitative Criminology.
  https://arxiv.org/abs/1807.04164, accepted.

- "Assumption Lean Regression," Berk, R., Buja, A., Brown, L.D., George, E.I., Kuchibhotla, A.K., Su, W.J., Zhao, L.
  The American Statistician (2019).
  https://doi.org/10.1080/00031305.2019.1592781

- "Damaging de novo Mutations Diminish Motor Skills in Children on the Autism Spectrum," Buja, A., Volfovsky, N., Krieger, A., Lord, C., Lash, A., Wigler, M., Iossifov, I. *Proceedings of the National Academy of Sciences of the United States of America*, 115 (8) 1859–1866 (2018). https://doi.org/10.1073/pnas.1715427115

- "Calibrated Percentile Double Bootstrap For Robust Linear Regression Inference," McCarthy, D., Zhang, K., Brown, L.D., Berk, R., Buja, A., George, E., Zhao, L. *Statistica Sinica*, **28** (4), 2565–2590 (2018). https://doi.org/10.5705/ss.202016.0546

- "Measuring shared variants in cohorts of discordant siblings: application to Autism," Ye, K., Iossifov, I., Levy, D., Yamrom, B., Buja, A., Krieger, A., Wigler, M. *Proceedings of the National Academy of Sciences* 114 (27), 7073–7076 (2017). https://10.1073/pnas.1700439114

- "Working with Misspecified Regression Models," Berk, R., Brown, L., Buja, A., George, E., Zhao L. (2017). *Journal of Quantitative Criminology.* https://doi.org/10.1007/s10940-017-9348-7

- "Managing Multichannel Complexity," Eyuboglu, N., Kabadayi, S., Buja, A. (2017). *Industrial Marketing Management*, **65**, 194–205.

- "A Visualization Tool for Mining Large Correlation Tables: The Association Navigator," Buja, A., Krieger, A., George, E.. *Handbook of Big Data*, eds.: Peter Bühlmann, Petros Drineas, Mark van der Laan, Michael Kane (2016)

- "Rate Optimal Denoising of Simultaneously Sparse and Low Rank Matrices," Yang, D., Ma, Z., Buja, A., *Journal of Machine Learning Research* 17 (92), 1–27 (2016).

- "Misspecified Mean Function Regression: Making Good Use of Regression Models that Are Wrong," Berk, R., Brown, L.D., Buja, A., George, E., Pitkin, E., Zhang, K. Zhao, L., *Sociological Methods & Research* (online), 1–30 (2014).

- "Covariance Adjustments for the Analysis of Randomized Field Experiments," Berk, R., Pitkin, E., Brown, L.D., Buja, A., George, E., Zhao, L., *Evaluation Review* (online), 1–27 (2014).

- "Will the Global Village Fracture into Tribes: Recommender Systems and their Effects on Consumers," Hosanagar, K., Fleder, D. M., Lee, D., Buja, A., *Management Science*, 50 (4), 805–823 (2014).

- "A Sparse SVD Method for High-dimensional Data," Yang, D., Ma, Z., Buja, A., *Journal of Computational and Graphical Statistics*, 23 (4), 923–-942 (2014).

- "The Power to See: A New Graphical Test of Normality," Aldor-Noiman, S., Brown, L. D., Buja, A., Rolke, W., Stine, R.A., *The American Statistician*, 67 (4), 249–260 (2013).

- "Valid Post-Selection Inference," Berk, R., Brown, L., Buja, A., Zhang, K., Zhao, L., *The Annals of Statistics*, 41 (2), 802–837 (2013).

- "Stress Functions for Nonlinear Dimension Reduction, Proximity Analysis, and Graph Drawing," Chen, L., and Buja, A., *Journal of Machine Learning Research*, 14, 1145–73 (2013).

- "Exploring the Relationship Between Anxiety and Insistence on Sameness in Autism Spectrum Disorders," Gotham, K., Bishop, S., Hus, V., Huerta, M., Lund, S., Buja, A., Krieger, A., and Lord, C., *Autism Research*, 6 (1) 33-–41 (2013).

- "Sub-types of Restricted and Repetitive Behaviors in Children with Autism Spectrum Disorders," Bishop, S.L., Hus, V., Duncan, A., Huerta, M., Gotham, K., Pickles, A., Krieger, A., Buja, A., Lund, S., and Lord, C, *Journal of Autism and Developmental Disorders*, 43 (6) 1287—1297 (June 2013).

- "Dosage-dependent phenotypes in models of human 16p11.2 lesions found in autism," Horev, G., Ellegood, J., Lerch, J.P., Son, Y.E., Muthuswamy, L., Vogel, H., Krieger, A.M., Buja, A., Henkelman, M., Wigler, M., and Mills, A.A. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (41), 17076–81 (2011).

- "Rare de novo and transmitted copy number variation in autistic spectrum disorders," Levy, D., Ronemus, M., Yamrom, B., Lee, Y., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Ye, K., Buja, A., Yoon, S., Krieger, A., Troge, J., Rodgers, L., Iossifov, I., and Wigler M. *Neuron*, 70 (5) 886–897 (9 June 2011).

- "tourr: An R Package for Exploring Multivariate Data with Projections," Wickham, H., Cook, D., Hofmann, H., and Buja, A. *Journal of Statistical Software*, 40(2), 1–18. URL http://www.jstatsoft.org/v40/i02/, (2011).

- "Recommender Systems and their Effects on Consumers: The Fragmentation Debate," Fleder, D., Hosanagar, K., and Buja, A., *11th ACM Conference on Electronic Commerce*, Harvard University (June 7–11, 2010).

- "Statistical Inference for Exploratory Data Analysis and Model Diagnostics," Buja, A., Cook, D., Hofmann, H., Lawrence, M., Lee, E.-K., Swayne, D.F., and Wickham, H., *Philosophical Transactions of the Royal Society A.*, 367, 4361–4383 (2009).

- "Principal Component Analysis of Two-Way Functional Data Using Two-Way Regularized Singular Value Decompositions," Huang, J.Z., Shen, H., and Buja, A., *Journal of the American Statistical Association*, 488, 1609–1620 (2009).

- "Local Multidimensional Scaling for Nonlinear Dimension Reduction, Graph Drawing and Proximity Analysis," Chen, L., and Buja, A., *Journal of the American Statistical Association*, Vol. 104, No. 485, 209–219 (March 2009).

- "Functional Principal Components Analysis via Penalized Rank One Approximation," Huang, J., Shen, H., and Buja, A., *Electronic Journal of Statistics*, 2, 678–695 (July 2008).

- "The Plumbing of Interactive Graphics," Wickham, H., Lawrence, M., Cook, D., Buja, A., Hofmann, H., and Swayne, D.F., *Computational Statistics*, (April 2008).

- "Data Visualization with Multidimensional Scaling," Buja, A., Swayne, D.F., Littman, M., Hofmann, H., and Chen, L., *Journal of Computational and Graphical Statistics* (2008).

- "Boosted Classification Trees and Class Probability/Quantile Estimation," Mease, D., Wyner, A.J., and Buja, A., *Journal of Machine Learning Research*, 8 (Mar), 409–439 (2007).

- "Observations on Bagging," Buja, A., and Stuetzle, W., *Statistica Sinica* (special issue on machine learning), 16 (2), 323–352 (2006).

- "Computational Methods for High-Dimensional Rotations in Data Visualization," Buja, A., Cook, D., Asimov, D., and Hurley, C., *Handbook of Statistics*, eds. E.J. Wegman and C.R. Rao (2005).

- "Visual Comparison of Datasets Using Mixture Distributions," Gous, A., and Buja, A., *Journal of Computational and Graphical Statistics*, 13 (1) 1–19 (2004).

- "Exploratory Visual Analysis of Graphs in GGobi," Swayne, D.F., Buja, A., and Temple-Lang, D., refereed proceedings of the *Third Annual Workshop on Distributed Statistical Computing* (DSC 2003), Vienna.

- "GGobi: Evolving from XGobi into an Extensible Framework for Interactive Data Visualization," Buja, A., Lang, D.T., and Swayne, D.F., *Journal of Computational Statistics and Data Analysis*, 43 (4), 423—444 (2003).

- "Robust Phoneme Discrimination Using Acoustic Waveforms," Cvetkovic, Z., Beferull-Lozano, B., and Buja, A., *Proceedings of ICASSP 2002* (May 2002).

- "Data Mining Criteria for Tree-Based Regression and Classification," Buja, A., and Lee, Y.-S. *Proceedings of KDD 2001*, 27–36 (2001).

- "Sampling Schemes for Model Visualization," Smith, A.J., Nelder, J. Buja, A., Malik, Z., Tweedie, L., and Spence, R., *Journal of Computational and Graphical Statistics*, 10, 545–554 (2001).

- "XGobi: Interactive Dynamic Data Visualization in the X Window System," Swayne, D.F., Cook, D., and Buja, A., *Journal of Computational and Graphical Statistics*, 7, 113–130 (1998).

- "Missing Data in Interactive High-Dimensional Data Visualization," Swayne, D.F., and Buja, A., *Computational Statistics*, 13 (1998).

- "Manual Controls for High-Dimensional Data Projections," Cook, D., and Buja, A., *Journal of Computational and Graphical Statistics*, 6, 464–480 (1997).

- "Interactive High-Dimensional Data Visualization," Buja, A., Cook, D., and Swayne, D., *Journal of Computational and Graphical Statistics*, 5, 78–99 (1996).

- "What Criterion for a Power Algorithm?" in: *Robust Statistics, Data Analysis and Computer Intensive Methods; Festschrift on the Occasion of Peter J. Huber's 60'th Birthday*, H. Rieder (ed.), Springer Lecture Notes in Statistics, 109 (1996).

- "Projection Pursuit Indices Based on Orthogonal Function Expansions," Cook, D., Buja, A., Cabrera, J., and Hurley, C., *Journal of Computational and Graphical Statistics* 2, 225–250 (1995).

- "Penalized Discriminant Analysis," Hastie, T., Buja, A., and Tibshirani, R., *The Annals of Statistics*, 23, 73–102 (1995).

- "Flexible Discriminant Analysis," Hastie, T., Tibshirani, R., and Buja, A., *Journal of the American Statistical Association*, 89, 1255–1270 (1994).

- "Inequalities and Positive-Definite Functions Arising from a Problem in Multidimensional Scaling," Buja, A., Logan, B.F., Reeds, J.R., and Shepp, L.A., *The Annals of Statistics*, 22, 406–438 (1994).

- "Remarks on Parallel Analysis," Buja, A., and Eyuboglu, N., *Multivariate Behavioral Research* 27, 509–540 (1993).

- "Dynamics of Channel Negotiations: Tactics, Outcomes, and Contention Level," Eyuboglu, N., and Buja, A., *Psychology and Marketing* 10, 47–65 (1992).

- "Structural Power in Channel Relationships," Eyuboglu, N., Buja, A., and Didow, N., *Advances in Distribution Channel Research*, 1, 81–113 (1992).

- "Interactive Data Visualization using Focusing and Linking," Buja, A., McDonald, J.A., and Stuetzle, W., *Visualization '91* (IEEE), 156–163 (1991).

- "Painting Multiple Views of Complex Objects," McDonald, J.A., Stuetzle, W., and Buja, A., *OOPSLA/ECOOP '90 Proceedings*, 245–257 (1990).

- "Analyzing High-Dimensional Data with Motion Graphics," Hurley, C.B., and Buja, A., *SIAM Journal on Scientific and Statistical Computing*, 11, 1193–1211 (1990).

- "Remarks on Functional Canonical Variates, Alternating Least Squares Methods, and ACE," Buja, A., *The Annals of Statistics*, 18, 1032–1069 (1990).

- "Elements of a Viewing Pipeline for Data Analysis," Buja, A., Asimov, D., Hurley, C.B., and McDonald, J.A., in: *Dynamic Graphics for Statistics*, W.S. Cleveland, M.E. McGill (eds.), Wadsworth, Statistics/Probability Series, 277–297 (1988).

- "On the Huber-Strassen Theorem," Buja, A., *Probability Theory and Related Fields*, 73, 149–152 (1986).

- "Simultaneously Least Favorable Experiments, Part II: Standard loss functions and their applications," Buja, A., *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 69, 387–420 (1985).

- "Simultaneously Least Favorable Experiments, Part I: Upper standard functionals and sufficiency," Buja, A., *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 65, 367–384 (1984).

Under Revision or Unpublished:

- "Rates of contributory de novo mutation in high and low risk autism families," Yoon S., Munoz A., Yamrom B., Lee Y.-h., Andrews P., Marks S., Wang1 Z., Buja A., Pradhan K., Ronemus M., Baldwin K.K., Levy D., Wigler M., and Iossifov I.

- "Hole or grain? A Section Pursuit Index for Finding Hidden Structure in Multiple Dimensions," Laa, U., Cook, D., Buja, A., Valencia, G.
  `https://arxiv.org/abs/2004.13327v1`

- "Model-free Study of Ordinary Least Squares Linear Regression," Kuchibhotla, A.K., Brown, L.D., Buja, A..
  `https://arxiv.org/abs/1809.10538`

- "Valid Post-selection Inference in Assumption-Lean Linear Regression," Kuchibhotla, A.K., Brown, L.D., Buja, A., George, E.I., Zhao, L. (2018).
  `https://arxiv.org/abs/1806.04119`

- "Smoothing Effects of Bagging: Von Mises Expansions of Bagged Statistical Functionals," Buja, A., Stuetzle, W..
  `http://arxiv.org/abs/1612.02528`

- "Semi-supervised linear regression," Azriel, D., Brown, L.D., Sklar, M., Berk, R., Buja, A., Zhao, L. (2016).
  `https://arxiv.org/abs/1612.02391`

- "Kernel Additive Principal Components," Tan, X.L., Buja, A., Ma, Z..
  `https://arXiv.org/abs/1511.06821`

- "Theory of Dynamic Projections in High-Dimensional Data Visualization," Buja, A., Cook, D., Asimov, D., and Hurley, C.
  `www-stat.wharton.upenn.edu/~buja/PAPERS/paper-dyn-proj-math.pdf`

- "Loss Functions for Binary Class Probability Estimation and Classification: Structure and Applications," Buja, A., Stuetzle, W., and Shen, Y.

- "Sampling/Resampling Methods for Simultaneous Inference with Applications to Function Estimation and Functional Data," Buja, A., and Rolke, W.
  `www-stat.wharton.upenn.edu/~buja/PAPERS/paper-sim.pdf`

Invited Discussions:

- With L.D. Brown: On "A Significance Test for the Lasso," by Lockhart, R., Taylor, J., Tibshirani, R.J., and Tibshirani, R. *The Annals of Statistics* 42 (2), 509–517 (2014).

- With W. Stuetzle: On "Evidence Contrary to the Statistical View of Boosting," by Wyner, A., and Mease, D., *Journal of Machine Learning Research*, 9, 41–46 (2008).

- On "Boosting Algorithms: Regularization, Prediction and Model Fitting," by Bühlmann, P., and Hothorn, T., *Statistical Science*, 22 (4), 506–512 (2007).

- On "Tukey's Paper After 40 Years," by Mallows, C., *Technometrics*, 48 (3), 327–330 (2006).

- On "Exploratory Data Analysis for Complex Models," by Gelman, A., *Journal of Computational and Graphical Statistics*, 13 (4), 780–784 (2004).

- On "Additive Logistic Regression: A Statistical View of Boosting," by Friedman, J.H., Hastie, T., and Tibshirani, R., *The Annals of Statistics*, 28, 387–391 (2000).
- On "Interactive Graphical Methods in the Analysis of Customer Panel Data," by Koschat, M., and Swayne, D., *Journal of Business and Economics Statistics* (1996).
- On "The Use of Polynomial Splines and their Tensor Products in Multivariate Function Estimation," by Stone, Ch.J., *The Annals of Statistics*, 22, 171–176 (1994).
- On "Multivariate Adaptive Regression Splines," by J.H. Friedman, (Buja, A., Duffy, D., Hastie, T., and Tibshirani, R.) *The Annals of Statistics*, 19, 93–98 (1991).
- On "OMEGA: Online Multivariate Exploratory Graphical Analysis," by Weihs, C., and Schmidli, H., (Buja, A., and Hurley, C.B.), *Statistical Science*, 5, 208–211 (1990).
- On "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data," by Cleveland, W.S., and McGill, R., *Journal of the Royal Statistical Society*, Series A, 150, 215–216 (1987).
- On "Computers in Statistical Research," by W. Eddy et al. (Buja, A., Fowlkes, E.B., Kettenring, J.R.), *Statistical Sciences*, 1, 440–442 (1986).
- On "Projection Pursuit," by Huber, P.J., (Buja, A., and Stuetzle, W.), *The Annals of Statistics*, 13, 484–490 (1985).
- On "Estimating Optimal Transformations for Multiple Regression and Correlation," by Breiman, L., and Friedman, J.H., (Buja, A., and Kass, R.E.), *Journal of the American Statistical Society*, 80, 602–607 (1985).

Interviews:

- "Interview with Andreas Buja," Symanzik, J., *Computational Statistics*, 23, 177–184 (2008).
- "A Conversation with Peter Huber," Buja, A., and Künsch, H. R., *Statistical Science*, Vol. 23, No. 1, 120–135 (2008).

Bookchapter:

- " Grand Tours, Projection Pursuit Guided Tours, and Manual Controls," Cook, Dianne, Buja, A., Lee, E.-K., Wickham, H., in *Handbook of Data Visualization*, eds. Chen, Ch.-H., Härdle, W., Unwin, A. Springer: Heidelberg and Berlin, 295–314 (2008).

Invited Talks with Proceedings Papers (*unrefereed*):

- "GGobi: XGobi Redesigned and Extended," Swayne, D.F., Temple Lang, D., Buja, A., and Cook, D., in: *Computing Science and Statistics*, 33 (2001).

- "Flexible Discriminant and Mixture Models," Hastie, T., Tibshirani, R., and Buja, A., in: "Neural Networks and Statistics," J. Kay, D. Titterington (eds.), Oxford University Press (1997).

- "Visualizing the Embedding of Objects in Euclidean Space," Littman, M.L., Swayne, D.F., Dean, N., and Buja, A., in: *Computing Science and Statistics*, 24, 208–217 (1992).

- "XGobi: Interactive Dynamic Graphics in the X Window System with a Link to S," Swayne, D.F., Cook, D., and Buja, A., in: *Proceedings of the 1991 Joint Statistical Meetings, American Statistical Association* (1992).

- "XGobi Meets S: Integrating Software for Data Analysis," Swayne, D.F., Hubbell, N., and Buja, A., in: *Computing Science and Statistics*, 23, 430–434 (1991).

- "A Data Viewer for Multivariate Data," Buja, A., Hurley, C.B., and McDonald, J.A., in: *Computing Science and Statistics*, 18, 171–174 (1986).

- "Grand Tour Methods: An Outline," Buja, A., and Asimov, D., in: *Computing Science and Statistics*, 17, 63–67 (1985).

Proceedings Papers (*unrefereed*):

- "The Grand Tour via Geodesic Interpolation of 2-Frames," Asimov, D., and Buja, A., in: *Visual Data Exploration and Analysis, Symposium on Electronic Imaging Science and Technology*, IS&T/SPIE (The Society for Imaging Science and Technology/ International Society for Optical Engineering) (1994).

- "Empirical Comparisons of Neural Networks and Statistical Methods for Classification and Regression," Buja, A., Duffy, D., Yuhas, B., and Jain, A., in: *Neural Networks in Telecommunications*, B. Yuhas, N. Ansari (eds.), Kluwer Academic Publishers, Norwell MA (1993).

- "An Analysis of Polynomial-based Projection Pursuit," Cook, D., Buja, A., and Cabrera, J., *Computing Science and Statistics*, 24, 478–482 (1992).

- "Direction and Motion Control in the Grand Tour," Cook, D., Buja, A., and Cabrera, J., *Computing Science and Statistics*, 23, 180–183 (1991).

- "Discovering Features of multivariate Data through Statistical Graphics," Buja, A., Fowlkes, E.B., Keramidas, E.M., Kettenring, J.R., Lee, J.C., Swayne, D.F., and Tukey, P.A., *ASA Proceedings in Statistical Graphics*, 98–103 (1986).

- "Designing an Intelligent System for Spectral Analysis," Percival, D.B., Buja, A., Martin, R.D., Belcher, E.O., Kerr, R.K., Yee, S.D., and Hurley, C.B., *Computing Science and Statistics*, 17, 29–37 (1985).

Reports for Foundation and Company Projects:

- "Different Worlds: Do Recommender Systems Fragment Consumers' Interests?" Fleder, D., Hosanagar, K., Buja, A.; *Knowledge@Wharton*,

http://knowledge.wharton.upenn.edu/article.cfm?articleid=2835 (2011).

- "A Tool for Mining Large Correlation Tables: The Association Navigator," Buja, A., Krieger, A., George, E.; report to the *Simons Foundation Autism Research Initiative* (2010).

- "2009 Report on the Analysis of the SSC Phenotype Data," Buja, A., Krieger, A., George, E.; report to the *Simons Foundation Autism Research Initiative* (2010).

- "Exploratory Analysis of a Large Telecom Network," Abney, S., Buja, A., and Volinsky, C.; AT&T Labs internal report (2004).

- "Finding Clusters in Common Channel Signaling Links with Correlated Failure Patterns," in: Bellcore Common Channel Signaling Network Notes (1994).

- "An Analysis of Common Channel Signaling Link Outages and Retransmission Data," Buja, A., and Dunlap, K., in: Bellcore Common Channel Signaling Network Notes (1994).

- "An Empirical Analysis of Link Oscillations in a Common Channel Signaling Network," Buja, A., and Hung, N., in: Bellcore Common Channel Signaling Network Notes (1992).

- "Methodology for Relational Database Analysis – a Study in Digital Switch Memory," Buja, A., Duffy, D., and Kagan, J., Bellcore Technical Memorandum (1990).

- "Analysis of Data on Customer Satisfaction with Residence Billing Inquiries," Buja, A., and Duffy, D., Bellcore Technical Memorandum (1989).

Data Visualization Software:

- "GGvis," a plug-in for "GGobi" for graph-layout and multidimensional scaling, successor to "XGvis".

- "GGobi," a platform-independent system for interactive graphical data analysis, successor of XGobi (co-designer). Swayne, D.F., Temple-Lang, D., Cook, D., Hofmann, H., and Buja, A. (2000-).

- "XGvis," a system for visualizing graphs and proximity data with multidimensional scaling (co-designer and co-implementor). Buja, A. Swayne, D.F., Littman, M.L., Dean, N., and Hofmann, H. (1992-).

- "XGobi," a system for interactive graphical data analysis based on the X window system (co-designer). Swayne, D.F., Cook, D., and Buja, A. (1990-2000).

- "Data Viewer," a Lisp machine-based prototype system for interactive graphical data analysis (sole designer and implementor) (1985-90).

Films and Videos:

The first, second and last item are available from the Video Lending Library of the ASA Graphics Section.

- "XGobi: Dynamic graphics for data analysis," Swayne, D.F., Cook, D., and Buja, A. Video, 20 min., Bellcore (l991).

- "Visualization of quantitative data," Buja, A., Stuetzle, W., McDonald, J.A., Michalak, J., and Willis S. Video, 30 min., Bellcore and Univ. of Washington (l990).

- "Grand tour methods," Buja, A., and McDonald, J.A.. Film, 5 min., Stanford University (l984).

- "Finding structure in unstructured data," Buja, A., and Asimov D. Film, 12 min., Stanford Linear Accelerator Center (l983).

- "Multidimensional scaling in a new environment," Buja, A. Film, 30 min., Stanford Linear Accelerator Center (1983).

Edited Book:

- "Computing and Graphics in Statistics," Buja, A., and Tukey, P.A. (eds.), *The IMA Volumes in Mathematics and its Applications*, vol. 36, Springer-Verlag (1991).

Grants:

- Simons Foundation Autism Research Initiative (SFARI), Directors' Award, 2019-2022, co-PI, with A. Krieger (University of Pennsylvania)

- Simons Foundation Autism Research Initiative (SFARI), Directors' Award, 2016-2019, co-PI, with A. Krieger (University of Pennsylvania)

- Simons Foundation Autism Research Initiative (SFARI 296012), $418,285, 2013-2015, "Statistical Analysis and Methodology Stemming from the SFARI Phenotype and Genotype Data," co-PI, with A. Krieger (co-PI) (University of Pennsylvania).

- NSF, DMS-1310795, $199,395, 2013-2014, "Collab. Research: Inference for Linear Model Parameters in Model-Free Populations," with L. Brown (PI), L. Zhao (University of Pennsylvania).

- Simons Foundation Autism Research Initiative (SFARI), $324,464, June 2011 - Dec 2012 (NCE to Dec 2013), "A Study in Autism, Phase 3", with A. Krieger (PI), E. George (University of Pennsylvania).

- Simons Foundation Autism Research Initiative (SFARI), $291,461, Jan 2010 - Dec 2010, "A Study in Autism, Phase 2", with A. Krieger (PI), E. George (University of Pennsylvania).

- Simons Foundation Autism Research Initiative (SFARI),
  $217,402, Aug 2008 - Dec 2009,
  "A Study in Autism, Phase 1",
  with A. Krieger, E. George (University of Pennsylvania).

- NSF, DMS-1007689, $129,999, 2010-2013
  "Collaborative Research: Inference for Statistical Graphics,"
  PI (University of Pennsylvania), collaborative with D. Cook, H. Hofmann (Iowa State
  U.) and H. Wickham (Rice U.)

- NSF, DMS-1007657, $400,000, 2010-2013,
  "Post Model Selection Inference & Empirical Bayes Selection,"
  with L. Brown (PI), L. Zhao (University of Pennsylvania).

- NSF, DMS-8504359, 1985-1988,
  $33,400 in FY85, $36,000 in FY86, $38,000 in FY87,
  "Non-parametric Techniques in Multiple Regression,"
  with W. Stuetzle (University of Washington).

- DoE, DE-FG06-85-ER25006, July 1985-June 1988,
  $250,000 in FY85, $260,000 in FY86, $270,000 in FY87,
  "Non-parametric Techniques in Multiple Regression,"
  with W. Stuetzle (University of Washington).

- DoD, FY84/FY85 $245,920
  "DoD-University Research Instrumentation Program,"
  with R.D. Martin, E. Belcher, D. Percival (University of Washington).

- ONR, N00014-81-k-0095, Jan 1984-Dec 1986, $139,044
  "Expert Systems and Real-Time Graphics for Time Series Analysis,"
  with R.D. Martin, E. Belcher, D. Percival (University of Washington).

- NSF, MCS-8304484, Aug 1983-Jan 1985, $30,000
  "Mathematical Sciences Research Equipment,"
  with R.D. Martin (University of Washington).

- NSF, MCS-8304234, June 1983-Nov 1985, $24,350
  "Mathematical Sciences: Statistical Methods Based on Real-Time Graphics"
  (University of Washington).

Recent Talks:

- ETHZ, Zurich, Switzerland, 2019/06/13
  "A Model-Free Theory of Parametric Regression"

- 80th Birthday Celebration for Jerome H. Friedman, Stanford University, 2019/05/15
  "How Jerry ACEd Regression"
- Washington University, St Louis, 2019/05/10
  "A Model-Free Theory of Parametric Regression"
- Flatiron Institute, NYC, Conference on Operator Splitting, 2019/03/21
  "A Probability Coverage Problem that Could Solve the Reproducibility Crisis in the Empirical Sciences"
- Memorial Conference in Honor of Larry Brown, University of Pennsylvania, 2018/12/01
  "A Model-Free Theory of Parametric Regression"
- Workshop on Higher-Order Asymptotics and Post-Selection Inference (WHOA-PSI) III, Washington University, St Louis, 2018/09/08-10
  "Construction of PoSI Statistics"
- Conference on Statistical Learning and Data Science / Nonparametric Statistics, Columbia University, 2018/06/04-06
  "Post-Selection Inference and Misspecification"
- Children's Hospital, Zurich, Switzerland, 2018/02/13
  "De Novo Mutations and Motor Skills"
- Simons Foundation Autism Research Initiative (SFARI), New York, 2018/01/16,25
  "De Novo Mutations and Motor Skills"
- Workshop on "Post Selection Inference and Multiple Testing", Mathematics Institute of Toulouse, France, 2018/02/7-9
  "From Post-Selection to Misspecification"
- Workshop on Higher-Order Asymptotics and Post-Selection Inference (WHOA-PSI) II, Washington University, St Louis, 2017/08/14
  "Higher-Order von Mises Expansions, Bagging and Assumption-Lean Inference"
- Workshop on Higher-Order Asymptotics and Post-Selection Inference (WHOA-PSI) I, Washington University, St Louis, 2016/10/02
  "Why Do Statisticians Treat Predictors as Fixed? A Conspiracy Theory"
- New York University, 2016/09/23
  "Why Do Statisticians Treat Predictors as Fixed? A Conspiracy Theory"
- Joint Statistical Meetings (JSM), Chicago, 2016/08/02
  "Why Do Statisticians Treat Predictors as Fixed? A Conspiracy Theory"
- World Congress in Probability and Statistics, Toronto, 2016/07/12
  "Post-selection inference and misspecification"
- Rutgers University, Statistics Department, 2016/04/20
  "Inference when Models are Approximations Rather than Truths"
- Deutsche Arbeitsgemeinschaft für Statistik (DAGSTAT), Göttingen, 2016/03/18
  "Combining Data Visualization and Analytics — Examples and Problems"