

Wenguang Sun

Address: 401W Bridge Hall
Department of Data Sciences and Operations
Marshall School of Business
3670 Trousdale Pkwy
University of Southern California
Los Angeles, CA 90089-0809
Phone: (213) 740-0093
Fax: (213) 740-7313
Email: wenguans@marshall.usc.edu
Website: <http://www-bcf.usc.edu/~wenguans/>

Education:

2008 Ph.D., University of Pennsylvania.
Advisor: T. Tony Cai
2003 B.S., Peking University, China

Employment:

2015-present Associate professor
Department of Data Sciences and Operations
Marshall School of Business, University of Southern California
2011-2015 Assistant professor
Department of Data Sciences and Operations
Marshall School of Business, University of Southern California
2008-2011 Assistant professor
Department of Statistics
North Carolina State University

Selected Awards and Honors:

2013 CAREER Award. National Science Foundation.
2013 Dean's Award for Research Excellence. Marshall School of Business.
2010 Research Grant. National Science Foundation.
2009 Saul Winegrad Dissertation Award. University of Pennsylvania
2007 ENAR Distinguished Student Paper Award.

Research Interests:

- Large-scale multiple testing and false discovery rate control
- Multistage testing and estimation
- Hierarchical inference and pattern classification
- Statistical decision theory
- Large-scale estimation and selective inference

Grants and Contracts:

2013-2018 "Simultaneous and sequential inference of high-dimensional data with sparse structure." CAREER, Division of mathematical sciences, National science foundation. Single PI. \$400K.

2010-2013 “New theory and methodology for large-scale multiple testing.” Division of mathematical sciences, National science foundation. Single PI. \$200K.

Accepted or Published Papers:

1. Cai, T., and **Sun, W.** (2017). Large-Scale Global and Simultaneous Inference: Estimation and Testing in Very High Dimensions. *Annual Review of Economics*, Volume 9, to appear.
2. Cai, T., and **Sun, W.** (2016). Optimal discovery and screening for sparse signals with applications to multistage high-throughput studies. *Journal of the Royal Statistical Society, Series B*, to appear.
3. **Sun, W.**, and Wei, Z. (2015). Hierarchical recognition of sparse patterns in large-scale simultaneous inference. *Biometrika*, 102, 267-280.
4. **Sun, W.**, Reich, B., Cai, T., Guindani, M., and Schwartzman, A. (2015). False discovery control in large-scale spatial multiple testing. *Journal of the Royal Statistical Society, Series B*, 77, 59-83.
5. Cao, H., **Sun, W.**, Kosorok, M. (2013). The optimal power puzzle: scrutiny of the monotone likelihood ratio assumption in multiple testing. *Biometrika*, 100, 495-502.
6. **Sun, W.**, and McLain, A. (2012). Multiple testing of composite null hypotheses in heteroscedastic models. *Journal of the American Statistical Association*, 498, 673-687.
7. **Sun, W.**, and Wei, Z. (2011). Multiple testing for pattern identification, with applications to time-course microarray experiments. *Journal of the American Statistical Association*, 106, 73-88.
8. **Sun, W.**, Joffe, M., Chen, J., and Brunelli, S. (2010). Design and analysis of multiple events case-control studies, *Biometrics*, 66, 1220-1229.
9. Wang, W., Wei, Z., and **Sun, W.** (2010). Simultaneous set-wise testing under dependence, with applications to genome-wide association studies. *Statistics and Its Interface*, 3, 501-511.
10. Cai, T., and **Sun, W.** (2009). Simultaneous testing of grouped hypotheses: finding needles in multiple haystacks. *Journal of the American Statistical Association*, 104, 1467-1481.
11. Wei, Z., **Sun, W.**, Wang, K., and Hakonarson, H. (2009). Multiple testing in genome-wide association studies via hidden Markov models. *Bioinformatics*, 25, 2802-2808.
12. **Sun, W.**, and Cai, T. (2009). Large-scale multiple testing under dependence. *Journal of the Royal Statistical Society, Series B*, 71, 393-424.
13. Shults, J., **Sun, W.**, Tu, X., Kim, H., Amsterdam, J., and Ten Have, T. (2009). On choosing the working correlation structure in the GEE analysis of longitudinal binary data. *Statistics in Medicine*, 28, 2338-2355.
14. **Sun, W.**, Shults, J., and Leonard, M. (2009). A note on the use of unbiased estimating equations to estimate correlation in GEE analysis of longitudinal trials. *Biometrical Journal*, 51, 5-18.
15. **Sun, W.**, and Cai, T. (2007). Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association*, 102, 901-912.
16. Mansson, R., Hennessy, S., **Sun, W.**, and Joffe, M. (2007). On the estimation and use of propensity scores in case-control and case-cohort studies. *American Journal of Epidemiology*, 166, 332-339.

17. Tu, X., Zhang, J., Kowalski J., Shults J., Feng C., **Sun, W.**, Tan W. (2007). Power analyses for longitudinal study designs with missing data. *Statistics in Medicine*, 26, 2958-2981.

Other Publications:

18. James, G., **Sun, W.**, and Qiao, X. (2012). Discussion of “Clustering random curves under dependence” by Serban and Jiang. *Technometrics*, 54, 123-126.
19. Cai, T., and **Sun, W.** (2010). A compound decision-theoretic approach to large-scale multiple testing. Book chapter for *Analysis of High Dimensional Data*, edited by Cai, T. and Shen, X..
20. Shults, J., **Sun, W.**, Tu, X., and Amsterdam, J. (2006). On the violation of bounds for the correlation in GEE analyses of binary data from longitudinal trials. *BEPRESS, University of Pennsylvania, Biostatistics Paper 8*.

Papers under Review or Revision

21. Cai, T., **Sun, W.** and Wang, W.* (2015). CARS: a covariate-assisted approach to large-scale two-sample inference with sparse means. Under revision for JRSSB.
22. Feng, T.*, Basu, P.*, **Sun, W.**, and Mack, W. (2016). Control of false discovery rate and optimal design for high-throughput screening. Under revision for Statistics in Medicine.
23. Basu, P.*, Cai, T., and **Sun, W.** (2016). Weighted false discovery control in large-scale multiple testing. Under revision for JASA.

Working Papers Close to Submission

24. Wang, W.*, and **Sun, W.** (2016). SMART: simultaneous multistage adaptive ranking and testing of sparse signals. To be submitted to the Annals of Statistics.
25. Fu, L.*, James, G., **Sun, W.** (2016). Nonparametric empirical Bayes estimation of a high-dimensional normal mean vector with heteroscedastic errors. To be submitted to JRSSB.
26. Xia, Y. Cai, T., and **Sun, W.** (2016). GAP: a general framework for information pooling in two-sample sparse inference. To be submitted to JASA.

Presentations:

1. August 2017. Upcoming invited talk at Joint Statistical Meetings. Baltimore, MD.
2. June 2017. Upcoming invited talk at ICSA Applied Statistics Symposium, Chicago, IL.
3. June 2017. Upcoming invited talk at International Conference on Statistics and Econometrics. Hong Kong.
4. 12/21/2016. Invited talk at International Conference on Data Science. Shanghai, China.
5. 12/19/2016. Invited talk at The 10th ICSA International Conference. Shanghai, China.
6. 10/24/2016. Invited talk at USC Dornsife INET Conference on Big Data in Economics.

7. 6/1/2016. Invited talk at Canadian Statistical Conference. St. Catharines, Ontario, Canada.
8. 3/4/2016. Presentation at USC Marshall Research Fair.
9. 11/13/2015. Invited talk at Department of Information, Operation and Management department, Stern School of Business, NYU.
10. 7/2/2015. Invited talk at the IMS-China International Conference on Statistics and Probability. Kunming, China.
11. 6/27/2015. Invited talk at the Conference of High-Dimensional Statistics in the Age of Big Data. Beijing, China.
12. 6/26/2015. Invited seminar. Capital University of Economics and Business, Beijing, China.
13. 6/25/2015. Invited talk at the 10th International Conference on “Frontiers of Statistics.” Beijing, China.
14. 6/15/2015. Invited talk at ICSA conference. Weighted False Discovery Rate Control in Large-scale Multiple Testing. Fort Collins, CO.
15. June 2015. Invited talk at the IMS-China International Conference on Statistics and Probability. Beijing, China.
16. 11/5/2014. Invited seminar in the department of Probability and Statistics at UCSB. Santa Barbara, CA.
17. 7/9/2014. Invited talk at the Institute of Mathematical Statistics Annual meeting. False discovery control in large-scale spatial multiple testing. Sydney, Australia.
18. 4/7/2014. Invited seminar at the Department of Statistics and Operations Research. False discovery control in large-scale spatial multiple testing. University of North Carolina at Chapel Hill.
19. 3/9/2014. Invited talk at the ENAR conference. False discovery control in large-scale spatial multiple testing. Baltimore, MD.
20. 12/4/2013. Invited seminar at the Department of Statistics, Wharton School, University of Pennsylvania. Multiple testing for pattern classification.
21. 9/5/2013. Invited seminar at USC, Molecular & Computational Biology section. Multiple testing for pattern classification.
22. 8/4/2013. Topic contributed talk at Joint Statistical Meetings. Simultaneous inference on a decision tree.
23. 3/12/2013. Invited talk at the ENAR conference. Simultaneous inference on a decision tree.
24. 10/5/2012. Invited talk at The International Conference on Advances in Interdisciplinary Statistics. False discovery control in large-scale spatial multiple testing. Greensboro, NC.
25. 5/20/2012. Invited talk at FACM conference. False discovery control in large-scale spatial multiple testing. Newark, NJ.
26. 5/16/2012. Invited seminar at the Department of Biostatistics. University of Chicago. False discovery control under dependence. Chicago, IL.
27. 10/13/2011. Invited seminar at the Department of Statistics, University of Toronto. Some topics on multiple testing under dependence. Toronto, Canada.
28. 6/27/2011. Invited talk at ICSA Conference. Spatial multiple hypothesis testing. New York City.
29. 6/20/2011. Invited talk at WNAR Summer Meeting. Spatial multiple hypothesis testing. San Luis Obispo, CA.
30. 3/31/2011. Seminar at the Department of Statistics (Biostatistics seminar series),

- North Carolina State University. Large-scale multiple testing: a survey of recent developments and open problems. Raleigh, NC.
31. 3/24/2011. Invited seminar at the Department of Information and Operation Management, Marshall School of Business, University of Southern California. Large-scale multiple testing under dependence. Los Angeles, CA.
 32. 1/5/2011. Invited seminar at the Department of Statistics, School of Mathematics, Peking University. Large-scale multiple testing under dependence. Beijing, China.
 33. 12/20/2010. Invited talk at ICSA Applied Statistical Symposium. Large-scale multiple testing under dependence. Guangzhou, China.
 34. 8/5/2010. Invited talk at Joint Statistical Meetings. A false important discovery approach to testing large non-null effects. Vancouver, Canada.
 35. 7/27/2010. Invited talk at the ASA New Researcher's Conference. Optimal screening for sparse signals. Vancouver, Canada.
 36. 6/22/2010. Invited talk at ICSA Conference. Optimal screening for sparse signals. Indianapolis, IA.
 37. 3/22/2010. Invited talk at ENAR Spring Meeting. Optimal screening for sparse signals. New Orleans, LA.
 38. Invited talk. UNC.
 39. 2/29/2008. Invited seminar at Department of Statistics, Harvard University. Compound decision theory for large-scale multiple testing.
 40. 2/27/2008. Invited seminar at Department of Biostatistics, Boston University. Compound decision theory for large-scale multiple testing.
 41. 2/22/2008. Invited seminar at Department of Statistics, Brown University. Compound decision theory for large-scale multiple testing.
 42. 2/14/2008. Invited seminar at Department of Biostatistics and Medical Informatics, University of Wisconsin. Compound decision theory for large-scale multiple testing.
 43. 2/5/2008. Invited seminar at School of Statistics, University of Minnesota. Compound decision theory for large-scale multiple testing.
 44. 1/31/2008. Invited seminar at Department of Statistics, Rutgers University. Compound decision theory for large-scale multiple testing.
 45. 1/28/2008. Invited seminar at Department of Biostatistics, University of Iowa. Compound decision theory for large-scale multiple testing.
 46. 1/24/2008. Invited seminar at Department of Biostatistics, University of Michigan. Compound decision theory for large-scale multiple testing.
 47. 1/14/2008. Invited seminar at Department of Statistics, North Carolina State University. Compound decision theory for large-scale multiple testing.
 48. 8/1/2007. Contributed talk at JSM. Oracle and Adaptive Compound Decision Rules for False Discovery Rate Control, Salt Lake City, UT.
 49. 3/18/2007. Contributed talk at ENAR Spring Meeting. Oracle and Adaptive Compound Decision Rules for False Discovery Rate Control. Atlanta, GA.
 50. 3/28/2006. Contributed talk at ENAR Spring Meeting. Case-Control Follow-up Study: a New Scheme for Sampling from a Cohort. Tampa, FL.
 51. 8/9/2005. Contributed talk at JSM. Comparison of Wang-Carey Estimation versus Quasi-Least Squares, Minneapolis, MN.
 52. 3/21/2005. Contributed talk at ENAR Spring Meeting. Comparison of Wang-Carey Estimation versus Quasi-Least Squares. Austin, TX.