

T. Tony Cai
Daniel H. Silberberg Professor
Professor of Statistics and Data Science

Curriculum Vitae

Department of Statistics & Data Science
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-1686

(215)-898-8224 (O)
(215)-898-1280(Fax)
tcai@wharton.upenn.edu
www-stat.wharton.upenn.edu/~tcai

Education:

- Ph.D. Cornell University, Statistics, 1996

Academic Appointments:

University of Pennsylvania:

- Daniel H. Silberberg Professor, The Wharton School, 2018 – present
- Dorothy Silberberg Professor, The Wharton School, 2007-18
- Assistant-Full Professor of Statistics, The Wharton School, 2000-07
- Professor, Applied Mathematics & Computational Science Graduate Group, 2007 – present
- Associate Scholar, Department of Biostatistics, Epidemiology, & Informatics, Perelman School of Medicine, 2006 – present

Other Positions:

- Visiting Fellow, Isaac Newton Institute for Math. Sciences, Cambridge University, 1 & 6/2008
- Visiting Fellow, Centre for Math. and its Applications, Australian National University, 6-7/2004
- Assistant Professor of Statistics, Purdue University, 1996-2000

Administrative Appointment:

- Vice Dean, The Wharton School, 2017-20

Editorial Appointments:

- Co-Editor, *The Annals of Statistics*, 2010-12
- Associate Editor, *Journal of the Royal Statistical Society, Series B*, 2014-18
- Associate Editor, *Journal of the American Statistical Association*, 2005-10
- Associate Editor, *The Annals of Statistics*, 2004-09
- Associate Editor, *Statistica Sinica*, 2005-11
- Associate Editor, *Statistics Surveys*, 2006-09
- Editorial Board, *Frontiers of Statistics* (book series), 2008 - present
- Guest Editor, *Statistica Sinica* Special Issue on Multiscale Methods
- Guest Editor, *J. Nonparametric Statistics* Special Issue for the Inaugural IMS-China Conference

Awards and Honors:

- Wald Memorial Award and Lecture, Institute of Mathematical Statistics, 2027
- Leo Breiman Senior Award, American Statistical Association, 2026
- Best Paper Award (Gold), International Congress of Chinese Mathematicians, 2026
- President-elect, President, Past President, Institute of Mathematical Statistics, 2023-26
- Fellow, American Association for the Advancement of Science (AAAS), 2024
- Noether Distinguished Scholar Award, American Statistical Association, 2023
- Frontiers of Science Award, International Congress of Basic Science, 2023
- Laplace Lecture of the Bernoulli Society, 10th World Congress in Probability & Statistics, 2021
- International Chinese Statistical Association Distinguished Achievement Award, 2019
- International Chinese Statistical Association Outstanding Service Award, 2019
- Peter Whittle Lecture, Cambridge University, 2018
- Best Paper Award, International Congress of Chinese Mathematicians, 2018
- President-elect, President, Past President, International Chinese Statistical Association, 2016-18
- Hermann Otto Hirschfeld Lectures, Humboldt-Universität zu Berlin, 2012
- Forum Lecture, 28th European Meeting of Statisticians, 2010
- Medallion Lecture, Institute of Mathematical Statistics, 2009
- COPSS Presidents' Award, Committee of the Presidents of Statistical Societies (COPSS), 2008
- Fellow, Institute of Mathematical Statistics, 2006

Research Interests:

- Statistical machine learning
- High-dimensional statistics
- Large-scale inference
- Statistical decision theory
- Functional data analysis & nonparametric function estimation
- Applications to genomics, financial engineering, and medical imaging

Research Grants:

- NSF DMS-2413106, Principal Investigator, 2024-27
Multi-source Learning: Data-driven Algorithms, Optimality Theory, and Applications
- NIH R01-GM129781, Co-Investigator, 2023-2027
Methods for integrative genomic data analysis
- NIH R01-GM123056, Co-Investigator, 2022-26
Statistical methods for microbiome and metagenomics
- NSF DMS-2015259, Principal Investigator, 2020-23
Transfer learning for large-scale inference: General framework and data-driven algorithms
- NIH R01-GM129781, Co-Investigator, 2018-22
Methods for integrative genomic data analysis
- NIH R01-GM123056, Co-Investigator, 2017-21

Statistical methods for microbiome and metagenomics

- NSF DMS-1712735, Principal Investigator, 2017-20
Integrative large-scale data analysis and statistical inference
- Wharton School Global Initiatives, Principal Investigator, 2016-17
A quantitative study of Beijing's air pollution and its economic impacts
- NSF DMS-1403708, Principal Investigator, 2014-17
Theory and methods for estimation of nonsmooth functionals and detection of simultaneous signals
- NIH R01 CA-127334, Co-Investigator, 2012-17
Methods for genomic data with graphical structures
- NSF DMS-1208982, Principal Investigator, 2012-15
Random matrix theory and high dimensional statistics
- NSF FRG DMS-0854973, Principal Investigator, 2009-13
Statistical inference for high-dimensional data: Theory, methodology and applications
- NSF DMS-0604954, Principal Investigator, 2006-09
Theory and methodology for sparse inference
- NSF DMS-0306576, Co-Principal Investigator, 2003-06
An adaptation theory for nonparametric function estimation
- NSF DMS-0072578, Principal Investigator, 2000-03
Block thresholding methods for adaptive wavelet estimation: Theory and applications

Books:

1. Cai, T.T. and Shen, X. (eds.) (2010). *High Dimensional Data Analysis*, World Scientific, New Jersey.
2. Berger, J., Cai, T.T., and Johnstone, I. M. (eds.) (2010). *Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown*. Institute of Mathematical Statistics.

Published/Accepted Papers: (Available at www-stat.wharton.upenn.edu/~tcai)

- [1] Cai, T.T., Chakraborty, A., and Vuursteen, L. (2026+). Optimal federated learning for functional mean estimation under heterogeneous privacy constraints. *Journal of the Royal Statistical Society, Series B*, to appear.
- [2] Cai, T.T., Kim, D., Li, Y., and Xiang, D. (2026+). Transfer learning for covariance matrix estimation: Optimality and adaptivity. *Journal of the American Statistical Association*, to appear.
- [3] Cai, T.T., Chakraborty, A., and Vuursteen, L. (2026+). Optimal federated learning for nonparametric regression with heterogeneous distributed differential privacy constraints. *Journal of the American Statistical Association*, to appear.
- [4] Cai, T.T., Chakraborty, A., and Wang, Y. (2026+). Optimal differentially private ranking from pairwise comparisons. *Journal of the American Statistical Association*, to appear.
- [5] Liang, Z., Cai, T.T., Sun, W., and Xia, Y. (2026+). A locally adaptive algorithm for multiple testing with network structure. *Statistica Sinica*, to appear.

- [6] Auddy, A., Chakraborty, A., and Cai, T.T. (2026). Minimax and adaptive transfer learning for nonparametric classification under distributed differential privacy constraints. *Journal of the Royal Statistical Society, Series B* **88**, 903-929.
- [7] Li, S., Cai, T.T., and Li, H. (2026). Statistical inference for high dimensional regression with proxy data. *Statistica Sinica* **36**, 331-349.
- [8] Cai, T.T., Li, S., and Xiang, D. (2026). Estimation of the unimodal strong stochastic transitivity model and its application in Premier League rankings (in Chinese). *Scientia Sinica Mathematica* **56**, 3-32.
- [9] Chakraborty, A., Auddy, A., and Cai, T.T. (2026). When data can't meet: Estimating correlation across privacy barriers. *Advances in Neural Information Processing Systems* **38**, 95086-95120.
- [10] Rakshit, P., Wang, Z., Cai, T.T., and Guo, Z. (2025). SIHR: Statistical inference in high-dimensional linear and logistic regression models. *The R Journal* **16**, 27-45.
- [11] Tan, J., Zhang, Y., Hong, C., Cai, T.T., Cai, T., and Zhang, A. (2025). Integrated analysis for electronic health records with structured and sporadic missingness. *Journal of Biomedical Informatics* **171**, 104933.
- [12] Wang, S., Yuan, B., Cai, T.T., and Li, H. (2024). Phylogenetic association analysis with conditional rank correlation. *Biometrika* **111**, 881-902.
- [13] Cai, T.T., Ke, Z.T., and Turner, P. (2024). Testing high-dimensional multinomials with applications to text analysis. *Journal of the Royal Statistical Society, Series B* **86**, 922-942.
- [14] Li, S., Zhang, L., Cai, T.T., and Li, H. (2024). Estimation and inference for high-dimensional generalized linear models with knowledge transfer. *Journal of the American Statistical Association* **119**, 1274-1285.
- [15] Cai, T.T. and Wei, H. (2024). Distributed Gaussian mean estimation under communication constraints: Optimal rates and communication-efficient algorithms. *Journal of Machine Learning Research* **25**, 1-63.
- [16] Cai, T.T., Kim, D., and Pu, H. (2024). Transfer learning for functional mean estimation: Phase transition and adaptive algorithms. *The Annals of Statistics* **52**, 654-678.
- [17] Ma, R., Guo, Z., Cai, T.T., and Li, H. (2024). Statistical inference for genetic relatedness based on high-dimensional logistic regression. *Statistica Sinica* **34**, 1023-1043.
- [18] Cai, T.T., Chen, R., and Zhu, Y. (2024). Estimation and inference for minimizer and minimum of convex functions: Optimality, adaptivity, and uncertainty principles. *The Annals of Statistics* **52**, 392-411.
- [19] Cai, C., Cai, T.T., and Li, H. (2024). Transfer learning for contextual multi-armed bandits. *The Annals of Statistics* **52**, 207-232.
- [20] Cai, T.T. and Ma, R. (2024). Matrix reordering for noisy disordered matrices: Optimality and computationally efficient algorithms. *IEEE Transactions on Information Theory* **70**, 509-531.
- [21] Cai, T.T., Guo, Z., and Xia, Y. (2023). Statistical inference and large-scale multiple testing for high-dimensional regression models (invited discussion paper). *Test* **32**, 1135-1171.
- [22] Li, S., Cai, T.T., and Li, H. (2023). Transfer learning in large-scale Gaussian graphical models with false discovery rate control. *Journal of the American Statistical Association* **118**, 2171-2183.

- [23] Wu, R., Zhang, L., and Cai, T.T. (2023). Sparse topic modeling: Computational efficiency, near-optimal algorithms, and statistical inference. *Journal of the American Statistical Association* **118**, 1849-1861.
- [24] Xia, Y. and Cai, T.T. (2023). Discussion of “A Scale-free Approach for False Discovery Rate Control in Generalized Linear Models”. *Journal of the American Statistical Association* **118**, 1569-1572.
- [25] Cai, T.T., Guo, Z., and Ma, R. (2023). Statistical inference for high-dimensional generalized linear models with binary outcomes. *Journal of the American Statistical Association* **118**, 1319-1332.
- [26] Cai, T.T. and Ma, R. (2022). Theoretical foundations of t-SNE for visualizing high-dimensional clustered data. *Journal of Machine Learning Research* **23**, 1-54.
- [27] Li, S., Cai, T.T., and Li, H. (2022). Inference for high-dimensional linear mixed-effects models: A quasi-likelihood approach. *Journal of the American Statistical Association* **117**, 1835-1846.
- [28] Cai, T.T., Zhang, A., and Zhou, Y. (2022). Sparse group Lasso: Sample complexity, optimal rate, and statistical inference. *IEEE Transactions on Information Theory* **68**, 5975-6002.
- [29] Cai, T.T., Sun, W., and Xia, Y. (2022). LAWS: A locally adaptive weighting and screening approach to spatial multiple testing. *Journal of the American Statistical Association* **117**, 1370-1383.
- [30] Li, S., Cai, T.T., and Li, H. (2022). Transfer learning for high-dimensional linear regression: Prediction, estimation, and minimax optimality. *Journal of the Royal Statistical Society, Series B* **84**, 149-173.
- [31] Cai, T.T. and Pu, H. (2022). Stochastic continuum-armed bandits with additive models: Minimax regrets and adaptive algorithm. *The Annals of Statistics* **50**, 2179-2204.
- [32] Cai, T.T. and Wei, H. (2022). Distributed adaptive Gaussian mean estimation with unknown variance: Interactive protocol helps adaptation. *The Annals of Statistics* **50**, 1992-2020.
- [33] Cai, T.T. and Wei, H. (2022). Distributed nonparametric regression: Optimal rate of convergence and cost of adaptation. *The Annals of Statistics* **50**, 698-725.
- [34] Kang, H., Lee, Y., Cai, T.T., and Small, D.S. (2022). Two robust tools for inference about causal effects with invalid instruments. *Biometrics* **78**, 24-34.
- [35] Ma, R., Cai, T.T., and Li, H. (2022). Optimal estimation of simultaneous signals using absolute inner product with applications to integrative genomics. *Statistica Sinica* **32**, 1027-1048.
- [36] Cai, T.T., Han, R., and Zhang, A. (2022). On the non-asymptotic concentration of heteroskedastic Wishart-type matrix. *Electronic Journal of Probability* **27**, 1-40.
- [37] Zhang, A., Cai, T.T., and Wu, Y. (2022). Heteroskedastic PCA: Algorithm, optimality, and applications. *The Annals of Statistics* **50**, 53-80.
- [38] Guo, Z., Renaux, C., Buhlmann, P., and Cai, T.T. (2021). Group inference in high dimensions with applications to hierarchical testing. *Electronic Journal of Statistics* **15**, 6633-6676.
- [39] Cai, T.T., Jiang, T., and Li, X. (2021). Asymptotic analysis for extreme eigenvalues of principal minors of random matrices. *The Annals of Applied Probability* **31**, 2953-2990.
- [40] Cai, T.T., Wang, Y., and Zhang, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics* **49**, 2825-2850.

- [41] Ma, R., Cai, T.T., and Li, H. (2021). Optimal estimation of bacterial growth rates based on permuted monotone matrix. *Biometrika* **108**, 693-708.
- [42] Cai, T., Cai, T.T., and Guo, Z. (2021). Individualized treatment selection: An optimal hypothesis testing approach in high-dimensional models. *Journal of the Royal Statistical Society, Series B* **83**, 669-719.
- [43] Ma, R., Cai, T.T., and Li, H. (2021). Optimal permutation recovery in permuted monotone matrix model. *Journal of the American Statistical Association* **116**, 1358-1372.
- [44] Wang, S., Cai, T.T., and Li, H. (2021). Optimal estimation of Wasserstein distance on a tree with an application to microbiome studies. *Journal of the American Statistical Association* **116**, 1237-1253.
- [45] Cai, T.T. and Zhang, L. (2021). A convex optimization approach to high-dimensional sparse quadratic discriminant analysis. *The Annals of Statistic* **49**, 1537-1568.
- [46] Wang, S., Cai, T.T., and Li, H. (2021). Hypothesis testing for phylogenetic composition: A minimum-cost flow perspective. *Biometrika* **108**, 17-36.
- [47] Ma, R., Cai, T.T., and Li, H. (2021). Global and simultaneous testing for high-dimensional logistic regression models. *Journal of the American Statistical Association* **116**, 984-998.
- [48] Cai, T.T., Kim, D., Song, X., and Wang, Y. (2021). Optimal sparse eigenspace and low-rank density matrix estimation for quantum systems. *Journal of Statistical Planning and Inference* **213**, 50-71.
- [49] Cai, T.T., Li, H. and Ma, R. (2021). Optimal structured principal subspace estimation: Metric entropy and minimax rates. *Journal of Machine Learning Research* **22**, 1-45.
- [50] Cai, T.T. and Wei, H. (2021). Transfer learning for nonparametric classification: Minimax rate and adaptive classifier. *The Annals of Statistics* **49**, 100-128.
- [51] Xia, Y., Cai, T.T., and Sun, W. (2020). GAP: A general framework for information pooling in two-sample sparse inference. *Journal of the American Statistical Association* **115**, 1236-1250.
- [52] Cai, T.T. and Wu, Y. (2020). Statistical and computational limits for sparse matrix detection. *The Annals of Statistics* **48**, 1593-1614.
- [53] Cai, T.T., Han, X., and Pan, G. (2020). Limiting laws for divergent spiked eigenvalues and largest non-spiked eigenvalue of sample covariance matrices. *The Annals of Statistics* **48**, 1255-1280.
- [54] Cai, T.T. and Guo, Z. (2020). Semi-supervised inference for explained variance in high-dimensional regression and its applications. *Journal of the Royal Statistical Society, Series B* **82**, 391-419.
- [55] Cai, T.T., Liang, T., and Rakhlin, A. (2020). Weighted message passing and minimum energy flow for heterogeneous stochastic block models with side information. *Journal of Machine Learning Research* **21**, 1-34.
- [56] Cai, T.T., Hu, J., Li, Y., and Zheng, X. (2020). High-dimensional minimum variance portfolio estimation based on high-frequency data. *Journal of Econometrics* **214**, 482-494.
- [57] Cai, T.T. (2019). Gaussianization Machines for non-Gaussian function estimation models. *Statistical Science* **34**, 635-656.
- [58] Cai, T.T., Li, H., Ma, J., and Xia, Y. (2019). Differential Markov random field analysis with an application to detecting differential microbial community networks. *Biometrika* **106**, 401-416.

- [59] Xiang, D., Zhao, S. D., and Cai, T.T. (2019). Signal classification for the integrative analysis of multiple sequences of multiple tests. *Journal of the Royal Statistical Society, Series B* **81**, 707-734.
- [60] Cai, T.T. and Zhang, L. (2019). High-dimensional linear discriminant analysis: Optimality, adaptive algorithm, and missing data. *Journal of the Royal Statistical Society, Series B* **81**, 675-705.
- [61] Zhang, A., Brown, L.D., and Cai, T.T. (2019). Semi-supervised inference: General theory and estimation of means. *The Annals of Statistics* **47**, 2538-2566.
- [62] Guo, Z., Wang, W., Cai, T.T., and Li, H. (2019). Optimal estimation of genetic relatedness in high-dimensional linear models. *Journal of the American Statistical Association* **114**, 358-369.
- [63] Cai, T.T., Sun, W., and Wang, W. (2019). CARS: Covariate assisted ranking and screening for large-scale two-sample inference (with discussion). *Journal of the Royal Statistical Society, Series B* **81**, 187-234.
- [64] Cai, T., Cai, T.T., Liao, K., and Liu, W. (2019). Large-scale simultaneous testing of cross-covariance matrix with applications to PheWAS. *Statistica Sinica* **29**, 983-1005.
- [65] Cai, T.T., Ma, J. and Zhang, L. (2019). CHIME: Clustering of high-dimensional Gaussian mixtures with EM algorithm and its optimality. *The Annals of Statistics* **47**, 1234-1267.
- [66] Guo, Z., Kang, H., Cai, T.T., and Small, D.S. (2018). Testing endogeneity with high dimensional covariates. *Journal of Econometrics* **207**, 175-187.
- [67] Basu, P., Cai, T.T., Das, K., and Sun, W. (2018). Weighted false discovery rate control in large-scale multiple testing. *Journal of the American Statistical Association* **113**, 1172-1183.
- [68] Guo, Z., Kang, H., Cai, T.T., and Small, D.S. (2018). Confidence intervals for causal effects with invalid instruments using two-stage hard thresholding with voting. *Journal of the Royal Statistical Society, Series B* **80**, 793-815.
- [69] Cai, T.T., Zhang, L., and Zhou, H.H. (2018). Adaptive functional linear regression via functional principal component analysis and block thresholding. *Statistica Sinica* **28**, 2455-2468.
- [70] Cai, T.T. and Guo, Z. (2018). Accuracy assessment for high-dimensional linear regression. *The Annals of Statistics* **46**, 1807-1836.
- [71] Xia, Y., Cai, T.T., and Li, H. (2018). Joint testing and false discovery rate control for high-dimensional multivariate regression. *Biometrika* **105**, 249-269.
- [72] Xia, Y., Cai, T., and Cai, T.T. (2018). Multiple testing of submatrices of a precision matrix with applications to identification of between pathway interactions. *Journal of the American Statistical Association* **113**, 328-339.
- [73] Cai, T.T., Guntuboyina, A., and Wei, Y. (2018). Adaptive estimation of planar convex sets. *The Annals of Statistics* **46**, 1018-1049.
- [74] Cai, T.T. and Zhang, L. (2018). High-dimensional Gaussian regression: Adaptive estimation and statistical inference. *Statistica Sinica* **28**, 963-993.
- [75] Cai, T.T. and Zhang, A. (2018). Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics. *The Annals of Statistics* **46**, 60-89.

- [76] Xia, Y., Cai, T., and Cai, T.T. (2018). Two-sample tests for high-dimensional linear regression with an application to detecting interactions. *Statistica Sinica* **28**, 63-92.
- [77] Zhao, S. D., Cai, T.T., and Li, H. (2017). Optimal detection of positive latent dependence between two sequences of multiple tests. *Journal of Multivariate Analysis* **160**, 169-184.
- [78] Cai, T.T. and Sun, W. (2017). Large-scale global and simultaneous inference: Estimation and testing in very high dimensions. *Annual Review of Economics* **9**, 411-439.
- [79] Zhao, S.D., Cai, T.T., Cappola, T.P., Margulies, K.B., and Li, H. (2017). Sparse simultaneous signal detection for identifying genetically controlled disease genes. *Journal of the American Statistical Association* **112**, 1032-1046.
- [80] Cai, T.T., Liang, T., and Rakhlin, A. (2017). On detection and structural reconstruction of small world random networks. *IEEE Transactions on Network Science and Engineering* **4**, 165-176.
- [81] Cai, T.T., Liang, T., and Rakhlin, A. (2017). Computational and statistical boundaries for submatrix localization in a large noisy matrix. *The Annals of Statistics* **45**, 1403-1430.
- [82] Liao, K.P., Sparks, J.A., Hejblum, B., Kuo, I-H., Cui, J., Lahey, L., Cagan, A., Gainer, V.S., Liu, W., Cai, T.T., Sokolove, J., and Cai, T. (2017). Phenome-wide association study of autoantibodies to citrullinated and non-citrullinated epitopes in rheumatoid arthritis. *Arthritis & Rheumatology* **69**, 742-749.
- [83] Cai, T.T. and Guo, Z. (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *The Annals of Statistics* **45**, 615-646.
- [84] Cai, T.T. and Tan, X.L. (2017). Optimal estimation of a quadratic functional under the Gaussian two-sequence model. *Statistica Sinica* **27**, 879-906.
- [85] Cai, T.T. (2017). Global testing and large-scale multiple testing for high-dimensional covariance structures. *Annual Review of Statistics and Its Application* **4**, 423-446.
- [86] Cai, T.T. and Sun, W. (2017). Optimal screening of sparse signals with applications to multistage high-throughput studies. *Journal of the Royal Statistical Society, Series B* **79**, 197-223.
- [87] Cai, T.T. and Zhang, L. (2016). Discussion of “Influential Feature PCA for High Dimensional Clustering”. *The Annals of Statistics* **44**, 2372-2381.
- [88] Cai, T.T., Li, X. and Ma, Z. (2016). Optimal rates of convergence for noisy sparse phase retrieval via thresholded Wirtinger flow. *The Annals of Statistics* **44**, 2221-2251.
- [89] Cai, T., Cai, T.T., and Zhang, A. (2016). Structured matrix completion with applications to genomic data integration. *Journal of the American Statistical Association* **111**, 621-633.
- [90] Cai, T.T., Liang, T., and Rakhlin, A. (2016). Geometric inference for general high-dimensional linear inverse problems. *The Annals of Statistics* **44**, 1536-1563.
- [91] Cai, T.T. and Zhang, A. (2016). Minimax rate-optimal estimation of high-dimensional covariance matrices with incomplete data. *Journal of Multivariate Analysis* **150**, 55-74.
- [92] Cai, T.T., Eldar, Y.C., and Li, X. (2016). Global testing against sparse alternatives in time-frequency analysis. *The Annals of Statistics* **44**, 1438–1466.
- [93] Cai, T.T. and Zhou, W. (2016). Matrix completion via max-norm constrained optimization. *Electronic Journal of Statistics* **10**, 1493-1525.

- [94] Cai, T.T. and Yuan, M. (2016). Minimax and adaptive estimation of covariance operator for random variables observed on a lattice graph. *Journal of the American Statistical Association* **111**, 253-265.
- [95] Cai, T.T. and Liu, W. (2016). Large-scale multiple testing of correlations. *Journal of the American Statistical Association* **111**, 229-240.
- [96] Kang, H., Zhang, A., Cai, T.T., and Small, D.S. (2016). Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. *Journal of the American Statistical Association* **111**, 132-144.
- [97] Cai, T.T., Liu, W., and Zhou, H.H. (2016). Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *The Annals of Statistics* **44**, 455-488.
- [98] Cai, T.T., Kim, D., Wang, Y., Yuan, M., and Zhou, H.H. (2016). Optimal large-scale quantum state tomography with Pauli measurements. *The Annals of Statistics* **44**, 681-712.
- [99] Cai, T.T., Li, H., Liu, W., and Xie, J. (2016). Joint estimation of multiple high-dimensional precision matrices. *Statistica Sinica* **26**, 445-464.
- [100] Cai, T.T., Ren, Z., and Zhou, H.H. (2016). Estimating structured high-dimensional covariance and precision matrices: Optimal rates and adaptive estimation (with discussion). *Electronic Journal of Statistics* **10**, 1-59.
- [101] Cai, T.T. and Zhang, A. (2016). Inference for high-dimensional differential matrices. *Journal of Multivariate Analysis* **143**, 107-126.
- [102] Liao, K., Sparks, J.A., Hejblum, B., Kuo, I., Cui, J., Lahey, L.J., Cagan, A., Gainer, V., Liu, W., Cai, T.T. and Sokolove, J. (2015). Phenome-wide association study of novel anti-citrullinated peptide antibodies in rheumatoid arthritis. In *Arthritis & Rheumatology* (Vol. 67), Wiley-Blackwell, New Jersey.
- [103] Cai, T.T. and Li, X. (2015). Robust and computationally feasible community detection algorithm in the presence of outlier nodes. *The Annals of Statistics* **43**, 1027-1059.
- [104] Cai, T.T., Ma, Z., and Wu, Y. (2015). Optimal estimation and rank detection for sparse spiked covariance matrices. *Probability Theory and Related Fields* **161**, 781-815.
- [105] Xia, Y., Cai, T., and Cai, T.T. (2015). Testing differential network with applications to detecting gene by gene interactions. *Biometrika* **102**, 247-266.
- [106] Cai, T.T. and Low, M.G. (2015). A framework for estimation of convex functions. *Statistica Sinica* **25**, 423-456.
- [107] Cai, T.T. and Zhang, A. (2015). ROP: Matrix recovery via rank-one projections. *The Annals of Statistics* **43**, 102-138.
- [108] Jeng, J., Cai, T.T., and Li, H. (2015). Sparse segment identifications with applications to DNA copy number variation analysis. In *Advanced Medical Statistics, 2nd Edition*, Y. Lu, J. Fang, L., Tian, and H. Jin, eds., World Scientific, New Jersey, 863-887.
- [109] Sun, W., Reich, B., Cai, T.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.
- [110] Cai, T.T., Liang, T., and Zhou, H.H. (2015). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *Journal of Multivariate Analysis* **137**, 161-172.

- [111] Zhao, S.D., Cai, T.T., and Li, H. (2014). Direct estimation of differential networks. *Biometrika* **101**, 253-268.
- [112] Zhao, S.D., Cai, T.T., and Li, H. (2014). More powerful genetic association testing via a new statistical framework for integrative genomics. *Biometrics* **70**, 881-890.
- [113] Cai, T.T. and Xia, Y. (2014). High-dimensional sparse MANOVA. *Journal of Multivariate Analysis* **131**, 174-196.
- [114] Cai, T.T., Low, M.G., and Ma, Z. (2014). Adaptive confidence bands for nonparametric regression functions. *Journal of the American Statistical Association* **109**, 1054-1070.
- [115] Cai, T.T. and Zhang, A. (2014). Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Transactions on Information Theory* **60**, 122-132.
- [116] Cai, T.T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society, Series B* **76**, 349-372.
- [117] Cai, T.T. and Wu, Y. (2014). Optimal detection for sparse mixtures against a given null distribution. *IEEE Transactions on Information Theory* **60**, 2217-2232.
- [118] Cai, T.T. and Yuan, M. (2014). Comments on “A Significance Test for the Lasso”. *The Annals of Statistics* **42**, 478-482.
- [119] Cai, T.T. Ma, Z., and Wu, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *The Annals of Statistics* **41**, 3074-3110.
- [120] Cai, T.T. and Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* **19**, 2359-2388.
- [121] Cai, T.T. and Zhou, W. (2013). A max-norm constrained optimization approach to 1-bit matrix completion. *Journal of Machine Learning Research* **14**, 3619-3647.
- [122] Cai, T.T. and Zhang, A. (2013). Compressed sensing and affine rank minimization under restricted isometry. *IEEE Transactions on Signal Processing* **61**, 3279-3290.
- [123] Cai, T.T., Low, M.G., and Xia, Y. (2013). Adaptive confidence intervals for regression functions under shape constraints. *The Annals of Statistics* **42**, 722-750.
- [124] Cai, T.T., Fan, J., and Jiang, T. (2013). Distributions of angles in random packing on spheres. *Journal of Machine Learning Research* **14**, 1837-1864.
- [125] Cai, T.T., Ren, Z., and Zhou, H.H. (2013). Optimal rates of convergence for estimating Toeplitz covariance matrices. *Probability Theory and Related Fields* **156**, 101-143.
- [126] Cai, T.T. and Zhang, A. (2013). Sharp RIP bound for sparse signal and low-rank matrix recovery. *Applied and Computational Harmonic Analysis* **35**, 74-93.
- [127] Cai, T.T., Ma, Z., and Wu, Y. (2013). Recent results on sparse principal component analysis. *IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 181-183.
- [128] Jeng, J., Cai, T.T., and Li, H. (2013). Simultaneous discovery of rare and common segment variants. *Biometrika* **100**, 157-172.
- [129] Cai, T.T., Li, H., Liu, W., and Xie, J. (2013). Covariate adjusted precision matrix estimation with an application in genetical genomics. *Biometrika* **100**, 139-156.
- [130] Cai, T.T., Liu, W., and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association* **108**, 265-277.

- [131] Cai, T.T. and Zhou, H.H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. *The Annals of Statistics* **40**, 2389-2420.
- [132] Cai, T.T., Jeng, J., and Li, H. (2012). Robust detection and identification of sparse segments in ultra-high dimensional data analysis. *Journal of the Royal Statistical Society, Series B* **74**, 773-797.
- [133] Cai, T.T. and Zhou, H.H. (2012). Minimax estimation of large covariance matrices under l_1 norm (with discussion). *Statistica Sinica* **22**, 1319-1378.
- [134] Cai, T.T. and Yuan, M. (2012). Adaptive covariance matrix estimation through block thresholding. *The Annals of Statistics* **40**, 2014-2042.
- [135] Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *Journal of the American Statistical Association* **107**, 1201-1216.
- [136] Cai, T.T. (2012). Minimax and adaptive inference in nonparametric function estimation. *Statistical Science* **27**, 31-50.
- [137] Cai, T.T. and Jiang, T. (2012). Phase transition in limiting distributions of coherence of high-dimensional random matrices. *Journal of Multivariate Analysis* **107**, 24-39.
- [138] Cai, T.T. and Liu, W. (2011). A direct estimation approach to sparse linear discriminant analysis. *Journal of the American Statistical Association* **106**, 1566-1577.
- [139] Xie, J., Cai, T.T., Maris, J., and Li, H. (2011). Optimal false discovery rate control for dependent data. *Statistics and Its Interface* **4**, 417-430.
- [140] Cai, T.T. and Yuan, M. (2011). Optimal estimation of the mean function based on discretely sampled functional data: Phase transition. *The Annals of Statistics* **39**, 2330-2355.
- [141] Cai, T.T. and Jiang, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *The Annals of Statistics* **39**, 1496-1525.
- [142] Cai, T.T., Jeng, J., and Jin, J. (2011). Optimal detection of heterogeneous and heteroscedastic mixtures. *Journal of the Royal Statistical Society, Series B* **73**, 629-662.
- [143] Cai, T.T. and Wang, L. (2011). Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information Theory* **57**, 4680-4688.
- [144] Cai, T.T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association* **106**, 672-684.
- [145] Cai, T.T., Liu, W., and Luo, X. (2011). A constrained l_1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association* **106**, 594-607.
- [146] Wang, L., Brown, L.D., and Cai, T.T. (2011). A difference-based approach to semiparametric partial linear model. *Electronic Journal of Statistics* **6**, 619-641.
- [147] Xie, J., Cai, T.T., and Li, H. (2011). Sample size and power analysis for sparse signal recovery in genome-wide association studies. *Biometrika* **98**, 273-290.
- [148] Cai, T.T. and Low, M.G. (2011). Testing composite hypotheses, Hermite polynomials, and optimal estimation of a nonsmooth functional. *The Annals of Statistics* **39**, 1012-1041.
- [149] Jeng, J., Cai, T.T., and Li, H. (2010). Optimal sparse segment identification with application in copy number variation analysis. *Journal of the American Statistical Association* **105**, 1156-1166.

- [150] Cai, T.T. and Zhou, H.H. (2010). Nonparametric regression in natural exponential families. In *Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown*, IMS Collections Vol. 6, 199–215.
- [151] Cai, T.T., Wang, L., and Xu, G. (2010). New bounds for restricted isometry constants. *IEEE Transactions on Information Theory* **56**, 4388-4394.
- [152] Cai, T.T. and Sun, W. (2010). A compound decision-theoretic approach to large-scale multiple testing. In *High-Dimensional Data Analysis*, T.T. Cai and X. Shen, eds., World Scientific, New Jersey, 75-116.
- [153] Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *The Annals of Statistics* **38**, 3412-3444.
- [154] Cai, T.T., Wang, L., and Xu, G. (2010). Stable recovery of sparse signals and an oracle inequality. *IEEE Transactions on Information Theory* **56**, 3516-3522.
- [155] Cai, T.T., Munk, A., and Schmidt-Hieber, J. (2010). Sharp minimax estimation of the variance of Brownian motion corrupted with Gaussian noise. *Statistica Sinica* **20**, 1011-1024.
- [156] Cai, T.T. (2010). Comments on "Correlated z-values and the Accuracy of Large-Scale Statistical Estimates" by Bradley Efron. *Journal of the American Statistical Association* **105**, 1055-1056.
- [157] Cai, T.T., Zhang, C.-H., and Zhou, H.H. (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics* **38**, 2118-2144.
- [158] Brown, L.D., Cai, T.T., and Zhou, H.H. (2010). Nonparametric regression in exponential families. *The Annals of Statistics* **38**, 2005-2046.
- [159] Brown, L.D., Cai, T.T., Zhang, R., Zhao, L., and Zhou, H.H. (2010). The root-unroot algorithm for density estimation as implemented via wavelet block thresholding. *Probability Theory and Related Fields* **146**, 401-433.
- [160] Cai, T.T., Wang, L., and Xu, G. (2010). Shifting Inequality and recovery of sparse signals. *IEEE Transactions on Signal Processing* **58**, 1300-1308.
- [161] Cai, T.T. and Jin, J. (2010). Optimal rates of convergence for estimating the null and proportion of non-null effects in large-scale multiple testing. *The Annals of Statistics* **38**, 100-145.
- [162] Cai, T.T., Low, M.G., and Zhao, L.H. (2009). Sharp adaptive estimation by a blockwise method. *Journal of Nonparametric Statistics* **21**, 839-850.
- [163] Cai, T.T. and Sun, W. (2009). Simultaneous testing of grouped hypotheses: Finding needles in multiple haystacks. *Journal of the American Statistical Association* **104**, 1467-1481.
- [164] Cai, T.T. and Zhou, H.H. (2009). Asymptotic equivalence and adaptive estimation for robust nonparametric regression. *The Annals of Statistics* **37**, 3204-3235.
- [165] Cai, T.T. and Wang, H. (2009). Probability matching tolerance intervals for distributions in exponential families. *Statistica Sinica* **19**, 905-923.
- [166] Cai, T.T., Xu, G., and Zhang, J. (2009). On recovery of sparse signals via l_1 minimization. *IEEE Transactions on Information Theory* **55**, 3388-3397.
- [167] Cai, T.T. and Zhou, H.H. (2009). A data-driven block thresholding approach to wavelet estimation. *The Annals of Statistics* **37**, 569-595.

- [168] Cai, T.T., Levine, M., and Wang, L. (2009). Variance function estimation in multivariate nonparametric regression. *Journal of Multivariate Analysis* **100**, 126-136.
- [169] Sun, W. and Cai, T.T. (2009). Large-scale multiple testing under dependency. *Journal of the Royal Statistical Society, Series B* **71**, 393-424.
- [170] Cai, T.T. and Wang, L. (2008). Adaptive variance function estimation in heteroscedastic nonparametric regression. *The Annals of Statistics* **36**, 2025-2054.
- [171] Brown, L.D., Cai, T.T., and Zhou, H.H. (2008). Robust nonparametric estimation via wavelet median regression. *The Annals of Statistics* **36**, 2055-2084.
- [172] Wang, L., Brown, L.D., Cai, T.T., and Levine, M. (2008). Effect of mean on variance function estimation in nonparametric regression. *The Annals of Statistics* **36**, 646-664.
- [173] Cai, T.T. (2008). Discussion of “Microarrays, empirical Bayes, and the two-groups model” by Bradley Efron. *Statistical Science* **23**, 29-33.
- [174] Cai, T.T. (2008). On information pooling, adaptability and superefficiency in nonparametric function estimation. *Journal of Multivariate Analysis* **99**, 412-436.
- [175] Sun, W. and Cai, T.T. (2007). Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association* **102**, 901-912.
- [176] Jin, J. and Cai, T.T. (2007). Estimating the null and the proportion of non-null effects in large-scale multiple comparisons. *Journal of the American Statistical Association* **102**, 495-506.
- [177] Cai, T.T., Low, M.G., and Zhao, L. H. (2007). Tradeoffs between global and local risks in nonparametric function estimation. *Bernoulli* **13**, 1-19.
- [178] Cai, T.T. and Lv, J. (2007). Discussion of “The Dantzig Selector: Statistical estimation when p is much larger than n ” by E. J. Candes and T. Tao. *The Annals of Statistics* **35**, 2365-2369.
- [179] Cai, T.T., Jin, J., and Low, M.G. (2007). Estimation and confidence sets for sparse normal mixtures. *The Annals of Statistics* **35**, 2421-2449.
- [180] Cai, T.T. and Low, M.G. (2006). Optimal adaptive estimation of a quadratic functional. *The Annals of Statistics* **34**, 2298-2325.
- [181] Cai, T.T. and Hall, P. (2006). Prediction in functional linear regression. *The Annals of Statistics* **34**, 2159-2179.
- [182] Cai, T.T. and Low, M.G. (2006). Adaptive confidence balls. *The Annals of Statistics* **34**, 202-228.
- [183] Cai, T.T. and Low, M.G. (2006). Adaptation under probabilistic error for estimating linear functionals. *Journal of Multivariate Analysis* **97**, 231-245.
- [184] Cai, T.T. and Low, M.G. (2005). Non-quadratic estimators of a quadratic functional. *The Annals of Statistics* **33**, 2930-2956.
- [185] Cai, T.T. and Low, M.G. (2005). On adaptive estimation of linear functionals. *The Annals of Statistics* **33**, 2311-2343.
- [186] Cai, T.T. and Low, M.G. (2005). Nonparametric function estimation over shrinking neighborhoods: Superefficiency and adaptation. *The Annals of Statistics* **33**, 184-213.
- [187] Brown, L.D., Cai, T.T., and DasGupta, A. (2005). Discussion of “Fuzzy And Randomized Confidence Intervals And P -values” by Charles J. Geyer and Glen D. Meeden. *Statistical Science* **20**, 375-379.

- [188] Cai, T.T. and Low, M.G. (2005). Adaptive estimation of linear functionals under different performance measures. *Bernoulli* **11**, 341-358.
- [189] Chicken, E. and Cai, T.T. (2005). Block thresholding for density estimation: Local and global adaptivity. *Journal of Multivariate Analysis* **95**, 76-106.
- [190] Cai, T.T. (2005). One-sided confidence intervals in discrete distributions. *Journal of Statistical Planning and Inference* **131**, 63-88.
- [191] Cai, T.T. and Low, M.G. (2004). An adaptation theory for nonparametric confidence intervals. *The Annals of Statistics* **32**, 1805-1840.
- [192] Cai, T.T. and Low, M.G. (2004). Minimax estimation of linear functionals over nonconvex parameter spaces. *The Annals of Statistics* **32**, 552-576.
- [193] Cai, T.T. (2003). Rates of convergence and adaptation over Besov spaces under pointwise risk. *Statistica Sinica* **13**, 881-902
- [194] Cai, T.T. and Low, M.G. (2003). A note on nonparametric estimation of linear functionals. *The Annals of Statistics* **31**, 1140-1153.
- [195] Brown, L.D., Cai, T.T., and DasGupta, A. (2003). Interval estimation in exponential families. *Statistica Sinica* **13**, 19-49.
- [196] Brown, L.D., Cai, T.T., Low, M.G., and Zhang, C-H. (2002). Asymptotic equivalence theory for nonparametric regression with random design. *The Annals of Statistics* **30**, 688-707.
- [197] Cai, T.T. (2002). On adaptive estimation of a derivative and other related linear inverse problems. *Journal of Statistical Planning and Inference* **108**, 329-349.
- [198] Cai, T.T. (2002). On block thresholding in wavelet regression: adaptivity, block size, and threshold level. *Statistica Sinica* **12**, 1241-1273.
- [199] Brown, L.D., Cai, T.T., and DasGupta, A. (2002). Confidence intervals for a binomial proportion and asymptotic expansions. *The Annals of Statistics* **30**, 160-201.
- [200] Cai, T.T. and Silverman, B.W. (2001). Incorporating information on neighboring coefficients into wavelet estimation. *Sankhya (ser. B)* **63**, 127-148.
- [201] Cai, T.T. (2001). Discussion of "Regularization of Wavelets Approximations" by A. Antoniadis and J. Fan. *Journal of the American Statistical Association* **96**, 960-962.
- [202] Cai, T.T., Zhang, D., and Ben-Amotz, D. (2001). Enhanced chemical classification of Raman images with wavelet transformation. *Applied Spectroscopy* **55**, 1124-1130.
- [203] Brown, L.D., Cai, T.T., and DasGupta, A. (2001). Interval estimation for a binomial proportion (with discussion). *Statistical Science* **16**, 106-133.
- [204] Cai, T.T. and Brown, L.D. (1999). Wavelet regression for random uniform design. *Statistics and Probability Letters* **42**, 313-321.
- [205] Shyu, C-R., Cai, T.T., and Broderick, L.S. (1999). On archiving and retrieval of sequential images from tomographic databases in PACS. *Proc. SPIE - the International Society for Optical Engineering* **v3656**, 33-40.
- [206] Cai, T.T. (1999). Adaptive wavelet estimation: a block thresholding and oracle inequality approach. *The Annals of Statistics* **27**, 898-924.
- [207] Cai, T.T. and Brown, L.D. (1998). Wavelet shrinkage for nonequispaced samples. *The Annals of Statistics* **26**, 1783-1799.

- [208] Cai, T.T. (1991). Constructions of BIB designs using orthogonal geometries. *Journal of Shanghai Jiao Tong University* **25**, 113-121 (in Chinese).
- [209] Cai, T.T. (1990). On the existence of Kirkman systems $KS(2, 4; v)$ containing Kirkman subsystems. In W.D. Wallis (ed) *Combinatorial Designs and Applications*, Marcel Dekker, New York.
- [210] Cai, T.T. (1989). Embeddings of Kirkman systems $KS(2, 4; v)$. *Advances in Mathematics* **18**(3), 374-375.

Service on National/International Committees:

- Executive Committee, Institute of Mathematical Statistics, 2023-26
- Board of Advisors, Institute for Mathematical and Statistical Innovation, 2020-24
- Council of the Institute of Mathematical Statistics, 2020-23
- COPSS Distinguished Achievement Award & Lecture Committee, 2020-23
- Program Chair, IMS Annual Meeting at the Joint Statistical Meetings, Baltimore, 2017
- Publications Committee, Bernoulli Society, 2011-14
- COPSS Presidents' Award Committee, 2009-2012, 2014; Chair, 2012
- Program Committee, Joint Meeting between the IMS and Australian Statistical Conference, 2014
- Council of the Institute of Mathematical Statistics, 2010-12
- Program Committee, 8th World Congress in Probability and Statistics, Istanbul, Turkey, 2012
- Committee on Special Lectures, Institute of Mathematical Statistics, 2008-12
- Director of the Board of International Chinese Statistical Association, 2009-11
- Committee on Nominations, Institute of Mathematical Statistics, 2008-09
- Co-Chair, Scientific Committee, IMS-China International Conference, Weihai, China, 2009
- Scientific Committee, IMS-China International Conference, Hangzhou, China, 2008
- Program Chair, IMS Annual Meeting, Salt Lake City, 2007
- Committee to Select Editors, Institute of Mathematical Statistics, 2000-03
- National Science Foundation CAREER Award Panelist
- National Science Foundation Panelist

Professional Service:

Reviewer for National Science Foundation, National Security Agency, Israel Science Foundation.
 Referee for the Annals of Statistics, Journal of the American Statistical Association, Journal of the Royal Statistical Society, Series B, Biometrika, Bernoulli, Probability Theory and Related Fields, Applied and Computational Harmonic Analysis, Statistica Sinica, Scandinavian Journal of Statistics, Journal of Computational and Graphical Statistics, The American Statistician, Journal of Statistical Planning and Inference, Statistics and Probability Letters, Statistical Science, Psychometrika, Science in China, IEEE Transactions on Information Theory, IEEE Transactions on Image Processing, IEEE Transactions on Signal Processing, Journal of Multivariate Analysis, Proceedings of the National Academy of Sciences, Journal of Machine Learning Research.

University and Departmental Service:

Lawrence D. Brown Distinguished Lecture and Ph.D. Student Workshop Committee, 2021-2024
Wharton Faculty IT Steering Committee, 2022-23
Vice Dean for China Initiatives, The Wharton School, 2017–20
Hiring Committee, Department of Statistics, 2017-18
Wharton Dean's Advisory Council, 2016-17
Executive Committee, Applied Math and Computational Science Grad. Group, 2015-17
Wharton Faculty Survey Committee, 2013-14
Hiring Committee, Department of Statistics, 2012-14
Quinquennial Departmental Review Committee, 2010-11
Admission Committee, Applied Math and Computational Science Grad. Group, 2008-20
Doctoral Program Executive Committee, The Wharton School, 2005-08
Co-director for the Ph.D. Program in Statistics, 2005-08
Board for the Student Disciplinary System, University of Pennsylvania, 2002-04
Board for the Code of Academic Integrity, University of Pennsylvania, 2002-04
Colloquium Chair, Department of Statistics, Spring, 2002
Computing Committee, Department of Statistics, 2001-02
Design of Departmental Website, Spring, 2002

Professional Society Membership:

Institute of Mathematical Statistics
American Statistical Association
International Chinese Statistical Association
American Association for the Advancement of Science

Special Invited Lectures:

1. Plenary Lecture, International Congress of Chinese Mathematicians, Shanghai, 1/2026
2. Plenary Lecture, Zhejiang University Symposium on Probability and Statistics, Hangzhou, 12/2025
3. IMS Presidential Address, Nashville, 8/2025
4. Plenary Lecture, IMS New Researchers Conference, Nashville, 8/2025
5. Plenary Lecture, International Congress of Basic Science, Beijing, 7/2025
6. Plenary Lecture, International Conference on Statistical Optimization and Learning, Beijing, 7/2025
7. Plenary Lecture, International Congress of Basic Science, Beijing, 7/2024
8. Bahadur Lecture, Conference on Theory and Foundations of Statistics in the Era of Big Data, Tallahassee, 4/2024
9. Distinguished Scientist Hall Lecture, Institute of Systems Science, Chinese Academy of Sciences, Beijing, 10/2023
10. Gottfried E. Noether Award Lecture, American Statistical Association, Joint Statistical Meetings, Toronto, 2023
11. Frontiers of Science Award Lecture, International Congress of Basic Science, Beijing, 7/2023

12. Keynote Lecture, Joint Conference of Statistics and Data Science in China, Beijing, 7/2023
13. Plenary Lecture, International Conference on High-dimensional Statistics, Changsha, 5/2023
14. Distinguished Lecture, Institute for Advanced Study, Hong Kong University of Science and Technology, 5/2023
15. AMA Distinguished Lecture, Hong Kong Polytechnic University, 5/2023
16. Herman Rubin Distinguished Lecture, Purdue University, 4/2023
17. Distinguished Lecture, National University of Singapore, 5/2022
18. Plenary Lecture, 2021 Shenzhen International Conference on Frontiers of Statistics and Data Science, 12/2021
19. Laplace Lecture of the Bernoulli Society, 10th World Congress in Probability & Statistics, 7/2021
20. Plenary Lecture, The 3rd CSIAM Conference on Big Data and Artificial Intelligence, 7/2021
21. Keynote Lecture, 4th Symposium of the ICSA-Canada Chapter, Queen's University in Kingston, 8/2019
22. Keynote Lecture, 12th Annual International Conference of the Chinese Scholars Association for Management Science and Engineering, Chengdu, 7/2019
23. Keynote Lecture, 2019 Lilac International Conference on Application of Statistics, 6/2019
24. Distinguished Lecture, Zhejiang Gongshang University, 5/2019
25. Plenary Lecture, DAMO Academy Workshop on Extremely Large-Scale Graph and Graph Representation in Practice, Hangzhou, 3/2019
26. Joe Talk, 53rd Wharton Global Forum, Shanghai, 3/2019
27. Plenary Lecture, 2019 ICSA Conference on Data Science, 1/2019
28. Distinguished Lecture, 2018 International Conference on Data Science, Shanghai, 12/2018
29. Peter Whittle Lecture, Cambridge University, 10/2018
30. Keynote Lecture, 5th IBS-China Joint Biostatistics Conference, Guangzhou, China, 7/2018
31. Plenary Lecture, 9th International Purdue Symposium on Statistics, 6/2018
32. Keynote Lecture, Computational and Statistical Interface to Big Data Conference, King Abdullah University of Science and Technology (KAUST), Saudi Arabia, 3/2018
33. Distinguished Lecture, Department of Statistics, Chinese University of Hong Kong, 1/2018
34. JRI Distinguished lecture, Hong Kong Polytechnic University, 1/2018
35. Special Invited Lecture, 6th IMS-China Conference, Nanning, 7/2017
36. Distinguished Lecture, Chongqing Normal University, 6/2017
37. Keynote Lecture, 5th Workshop on Biostatistics and Bioinformatics, Atlanta, 5/2017
38. Distinguished Lecture, Southern University of Science and Technology, 3/2017
39. JRI Distinguished lecture, Hong Kong Polytechnic University, 1/2017
40. Keynote Lecture, 10th ICSA International Conference, Shanghai, 12/2016
41. Plenary Lecture, International Conference on "Some Mathematical Approximation Approaches in Data Science", Hangzhou, 12/2016
42. Distinguished Lecture, Institute for Advanced Study & Department of ISOM, Hong Kong University of Science and Technology, 6/2016
43. Keynote Lecture, Symposium: Frontiers of Statistics and Data Science, Hong Kong, 6/2016

44. Keynote Lecture, Workshop on High-Dimensional Statistical Modeling and Analysis in the Era of Big Data, Beijing, 5/2016
45. The Annals of Statistics Editors' special invited talk at the Joint Statistical Meetings, Seattle, 8/2015
46. Keynote Lecture, 2nd International Workshop on Frontiers of Statistics with Applications to Finance, Shanghai, 7/2015
47. Plenary Lecture, 2015 International Workshop on Mathematical Issues in Information Sciences, Xi'An, 7/2015
48. Plenary Lecture, 10th International Conference on Frontiers of Statistics, Beijing, 6/2015
49. Plenary Lecture, International Conference on Robust Statistics, Kolkata, India, 1/2015
50. Distinguished Lecture, Department of Statistics, Chinese University of Hong Kong, 1/2015
51. Distinguished Lecture, Department of Mathematics, Hong Kong Baptist University, 1/2015
52. Plenary Lecture, Workshop on "Statistics and Computational Interface to Big Data", Institute for Advanced Studies, Hong Kong University of Science and Technology, 1/2015
53. Plenary Lecture, International Conference on Learning and Approximation, Shanghai, 12/2014
54. Keynote Lecture, The 2014 ICSA Shanghai Statistics Conference, Shanghai, 7/2014
55. Distinguished Lecture, Department of Mathematics, Zhejiang University, 6/2014
56. Keynote Lecture, Conference on Big Data and Statistics, Tianjin, 4/2014
57. Special Invited Lecture, 4th IMS-China Conference, Chengdu, 7/2013
58. Keynote Lecture, International Conference on Statistics and Optimization, Beijing, 6/2013
59. Plenary Lecture, 7th Annual Information Theory and Applications Conference, San Diego, 2/2013
60. Distinguished Lecture, Department of Statistics, University of Washington, 2/2013
61. Hermann Otto Hirschfeld Lectures, Humboldt-Universität zu Berlin, 10/2012
62. Distinguished Lecture, 2nd Institute of Mathematical Statistics - Asia Pacific Rim Meeting, Tokyo, 7/2012
63. Distinguished Lecture, Department of Mathematics, Hong Kong Baptist University, 1/2012
64. Plenary Talk, International Research Symposium on Frontiers of Statistics, Hefei, 7/2011
65. Distinguished Lecture, 3rd IMS-China Conference, Xi'An, 7/2011
66. Plenary Lecture, International Conference on High Dimensional Statistics: Advances and Challenges, Singapore, 5/2011
67. Plenary Talk, 4th Lehmann Symposium, Houston, 5/2011
68. Puri Lecture, Department of Statistics, Purdue University, 4/2011
69. Forum Lecture, 28th European Meeting of Statisticians, Piraeus, Greece, 8/2010
70. Plenary Lecture, International Conference on Statistical Analysis of Complex Data, Kunming, 7/2010
71. Keynote Lecture, 9th German Open Conference on Probability and Statistics, Leipzig, 3/2010
72. Medallion Lecture, Institute of Mathematical Statistics Annual Meeting, Washington DC, 8/2009
73. Distinguished Lecture, 1st Institute of Mathematical Statistics - Asia Pacific Rim Meeting, Seoul, 6/2009
74. Keynote Lecture, Symposium on New Directions in Asymptotic Statistics, Athens, 5/2009