

THE ANNALS *of* STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles

- Distributed estimation of principal eigenspaces
JIANQING FAN, DONG WANG, KAIZHENG WANG AND ZIWEI ZHU 3009
- Additive models with trend filtering
VEERANJANEYULU SADHANALA AND RYAN J. TIBSHIRANI 3032
- Sorted concave penalized regression LONG FENG AND CUN-HUI ZHANG 3069
- Active ranking from pairwise comparisons and when parametric assumptions do not
help REINHARD HECKEL, NIHAR B. SHAH, KANNAN RAMCHANDRAN
AND MARTIN J. WAINWRIGHT 3099
- Randomized incomplete U -statistics in high dimensions
XIAOHUI CHEN AND KENGO KATO 3127
- Adaptive estimation of the rank of the coefficient matrix in high-dimensional multivariate
response regression models XIN BING AND MARTEN H. WEGKAMP 3157
- Statistical inference for autoregressive models under heteroscedasticity of unknown
form KE ZHU 3185
- On partial-sum processes of ARMAX residuals
STEFFEN GRØNNEBERG AND BENJAMIN HOLCBLAT 3216
- Quantile regression under memory constraint
XI CHEN, WEIDONG LIU AND YICHEN ZHANG 3244
- Sampling and estimation for (sparse) exchangeable graphs
VICTOR VEITCH AND DANIEL M. ROY 3274
- Hypothesis testing on linear structures of high-dimensional covariance matrix
SHURONG ZHENG, ZHAO CHEN, HENGJIAN CUI AND RUNZE LI 3300
- On optimal designs for nonregular models YI LIN, RYAN MARTIN AND MIN YANG 3335
- A smeary central limit theorem for manifolds with application to high-dimensional
spheres BENJAMIN ELTZNER AND STEPHAN F. HUCKEMANN 3360
- On testing for high-dimensional white noise
ZENG LI, CLIFFORD LAM, JIANFENG YAO AND QIWEI YAO 3382
- Minimax posterior convergence rates and model selection consistency in
high-dimensional DAG models based on sparse Cholesky factors
KYOUNGJAE LEE, JAEYONG LEE AND LIZHEN LIN 3413
- Bootstrapping and sample splitting for high-dimensional, assumption-lean
inference ALESSANDRO RINALDO, LARRY WASSERMAN AND MAX G'SELL 3438
- Joint convergence of sample autocovariance matrices when $p/n \rightarrow 0$ with
application MONIKA BHATTACHARJEE AND ARUP BOSE 3470
- Tracy–Widom limit for Kendall's tau ZHIGANG BAO 3504
- Intrinsic Riemannian functional data analysis ZHENHUA LIN AND FANG YAO 3533
- Detecting relevant changes in the mean of nonstationary processes—A mass excess
approach HOLGER DETTE AND WEICHI WU 3578

THE ANNALS OF STATISTICS

Vol. 47, No. 6, pp. 3009–3608 December 2019

INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.

IMS OFFICERS

President: Susan Murphy, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

President-Elect: Regina Y. Liu, Department of Statistics, Rutgers University, Piscataway, New Jersey 08854-8019, USA

Past President: Xiao-Li Meng, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

Executive Secretary: Edsel Peña, Department of Statistics, University of South Carolina, Columbia, South Carolina 29208-001, USA

Treasurer: Zhengjun Zhang, Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706-1510, USA

Program Secretary: Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027-5927, USA

IMS EDITORS

The Annals of Statistics. *Editors:* Richard J. Samworth, Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WB, UK. Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027, USA

The Annals of Applied Statistics. *Editor-in-Chief:* Karen Kafadar, Department of Statistics, University of Virginia, Heidelberg Institute for Theoretical Studies, Charlottesville, VA 22904-4135, USA

The Annals of Probability. *Editor:* Amir Dembo, Department of Statistics and Department of Mathematics, Stanford University, Stanford, California 94305, USA

The Annals of Applied Probability. *Editors:* François Delarue, Laboratoire J. A. Dieudonné, Université de Nice Sophia-Antipolis, France-06108 Nice Cedex 2. Peter Friz, Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany and Weierstrass-Institut für Angewandte Analysis und Stochastik, 10117 Berlin, Germany

Statistical Science. *Editor:* Cun-Hui Zhang, Department of Statistics, Rutgers University, Piscataway, New Jersey 08854, USA

The IMS Bulletin. *Editor:* Vlada Limic, UMR 7501 de l'Université de Strasbourg et du CNRS, 7 rue René Descartes, 67084 Strasbourg Cedex, France

The Annals of Statistics [ISSN 0090-5364 (print); ISSN 2168-8966 (online)], Volume 47, Number 6, December 2019. Published bimonthly by the Institute of Mathematical Statistics, 3163 Somerset Drive, Cleveland, Ohio 44122, USA. Periodicals postage paid at Cleveland, Ohio, and at additional mailing offices.

POSTMASTER: Send address changes to *The Annals of Statistics*, Institute of Mathematical Statistics, Dues and Subscriptions Office, 9650 Rockville Pike, Suite L 2310, Bethesda, Maryland 20814-3998, USA.

DISTRIBUTED ESTIMATION OF PRINCIPAL EIGENSPACES

BY JIANQING FAN^{*,1}, DONG WANG^{*}, KAIZHENG WANG^{*} AND ZIWEI ZHU[†]

Princeton University^{} and University of Michigan[†]*

Principal component analysis (PCA) is fundamental to statistical machine learning. It extracts latent principal factors that contribute to the most variation of the data. When data are stored across multiple machines, however, communication cost can prohibit the computation of PCA in a central location and distributed algorithms for PCA are thus needed. This paper proposes and studies a distributed PCA algorithm: each node machine computes the top K eigenvectors and transmits them to the central server; the central server then aggregates the information from all the node machines and conducts a PCA based on the aggregated information. We investigate the bias and variance for the resulting distributed estimator of the top K eigenvectors. In particular, we show that for distributions with symmetric innovation, the empirical top eigenspaces are unbiased, and hence the distributed PCA is “unbiased.” We derive the rate of convergence for distributed PCA estimators, which depends explicitly on the effective rank of covariance, eigengap, and the number of machines. We show that when the number of machines is not unreasonably large, the distributed PCA performs as well as the whole sample PCA, even without full access of whole data. The theoretical results are verified by an extensive simulation study. We also extend our analysis to the heterogeneous case where the population covariance matrices are different across local machines but share similar top eigenstructures.

REFERENCES

- ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Stat.* **34** 122–148. [MR0145620](#)
- BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. [MR2165575](#)
- BATTEY, H., FAN, J., LIU, H., LU, J. and ZHU, Z. (2015). Distributed estimation and inference with statistical guarantees. Preprint. Available at [arXiv:1509.05457](#).
- BERTRAND, A. and MOONEN, M. (2014). Distributed adaptive estimation of covariance matrix eigenvectors in wireless sensor networks with application to distributed PCA. *Signal Process.* **104** 120–135.
- BOUTSIDIS, C., WOODRUFF, D. P. and ZHONG, P. (2016). Optimal principal component analysis in distributed and streaming models. In *STOC’16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing* 236–249. ACM, New York. [MR3536569](#)
- CAI, T. T., MA, Z. and WU, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist.* **41** 3074–3110. [MR3161458](#)

MSC2010 subject classifications. Primary 62H25; secondary 62E10.

Key words and phrases. Distributed learning, PCA, one-shot approach, communication efficiency, unbiasedness of empirical eigenspaces, heterogeneity.

- CHEN, X. and XIE, M. (2014). A split-and-conquer approach for analysis of extraordinarily large data. *Statist. Sinica* **24** 1655–1684. [MR3308656](#)
- CHEN, T.-L., CHANG, D. D., HUANG, S.-Y., CHEN, H., LIN, C. and WANG, W. (2016). Integrating multiple random sketches for singular value decomposition. Preprint. Available at [arXiv:1608.08285](#).
- EL KAROUI, N. and D’ASPREMONT, A. (2010). Second order accurate distributed eigenvector computation for extremely large matrices. *Electron. J. Stat.* **4** 1345–1385. [MR2738536](#)
- FAN, J., WANG, W. and ZHU, Z. (2016). Robust low-rank matrix recovery. Preprint. Available at [arXiv:1603.08315](#).
- FAN, J., WANG, D., WANG, K. and ZHU, Z. (2019). Supplement to “Distributed estimation of principal eigenspaces.” DOI:[10.1214/18-AOS1713SUPP](#).
- GARBER, D., SHAMIR, O. and SREBRO, N. (2017). Communication-efficient algorithms for distributed stochastic principal component analysis. Preprint. Available at [arXiv:1702.08169](#).
- GOLUB, G. H. and VAN LOAN, C. F. (2012). *Matrix Computations*, 3rd ed. *Johns Hopkins Studies in the Mathematical Sciences*. Johns Hopkins Univ. Press, Baltimore, MD.
- GUO, Z.-C., LIN, S.-B. and ZHOU, D.-X. (2017). Learning theory of distributed spectral algorithms. *Inverse Probl.* **33** 074009, 29. [MR3672238](#)
- HALKO, N., MARTINSSON, P. G. and TROPP, J. A. (2011). Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **53** 217–288. [MR2806637](#)
- HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24** 417–441.
- JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#)
- JOHNSTONE, I. M. and LU, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *J. Amer. Statist. Assoc.* **104** 682–693. [MR2751448](#)
- JUNG, S. and MARRON, J. S. (2009). PCA consistency in high dimension, low sample size context. *Ann. Statist.* **37** 4104–4130. [MR2572454](#)
- KANNAN, R., VEMPALA, S. and WOODRUFF, D. (2014). Principal component analysis and higher correlations for distributed data. In *Conference on Learning Theory* 1040–1057.
- KARGUPTA, H., HUANG, W., SIVAKUMAR, K. and JOHNSON, E. (2001). Distributed clustering using collective principal component analysis. *Knowledge and Information Systems* **3** 422–448.
- KATO, T. (1966). *Perturbation Theory for Linear Operators*. *Die Grundlehren der Mathematischen Wissenschaften, Band 132*. Springer, New York. [MR0203473](#)
- KNEIP, A. and UTIKAL, K. J. (2001). Inference for density families using functional principal component analysis. *J. Amer. Statist. Assoc.* **96** 519–542. [MR1946423](#)
- KOLTCHINSKII, V. and LOUNICI, K. (2016). Asymptotics and concentration bounds for bilinear forms of spectral projectors of sample covariance. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1976–2013. [MR3573302](#)
- KOLTCHINSKII, V. and LOUNICI, K. (2017). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23** 110–133. [MR3556768](#)
- LEE, J. D., LIU, Q., SUN, Y. and TAYLOR, J. E. (2017). Communication-efficient sparse regression. *J. Mach. Learn. Res.* **18** 1–30. [MR3625709](#)
- LI, L., SCAGLIONE, A. and MANTON, J. H. (2011). Distributed principal subspace estimation in wireless sensor networks. *IEEE J. Sel. Top. Signal Process.* **5** 725–738.
- LIANG, Y., BALCAN, M.-F. F., KANCHANAPALLY, V. and WOODRUFF, D. (2014). Improved distributed principal component analysis. In *Advances in Neural Information Processing Systems* 3113–3121.
- MINSKER, S. (2018). Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries. *Ann. Statist.* **46** 2871–2903. [MR3851758](#)

- MÜCKE, N. and BLANCHARD, G. (2018). Parallelizing spectrally regularized kernel algorithms. *J. Mach. Learn. Res.* **19** Paper No. 30, 29. [MR3862437](#)
- NADLER, B. (2008). Finite sample approximation results for principal component analysis: A matrix perturbation approach. *Ann. Statist.* **36** 2791–2817. [MR2485013](#)
- ONATSKI, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *J. Econometrics* **168** 244–258. [MR2923766](#)
- PAUL, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. [MR2399865](#)
- PAUL, D. and JOHNSTONE, I. M. (2012). Augmented sparse principal component analysis for high dimensional data. Preprint. Available at [arXiv:1202.1242](#).
- PEARSON, K. (1901). On lines and planes of closest fit to systems of point in space. *Philosophical Magazine Series 6* **2** 559–572.
- QU, Y., OSTROUCHOV, G., SAMATOVA, N. and GEIST, A. (2002). Principal component analysis for dimension reduction in massive distributed data sets. In *IEEE International Conference on Data Mining (ICDM)*.
- REISS, M. and WAHL, M. (2016). Non-asymptotic upper bounds for the reconstruction error of PCA. Preprint. Available at [arXiv:1609.03779](#).
- SCHIZAS, I. D. and ADUJOA, A. (2015). A distributed framework for dimensionality reduction and denoising. *IEEE Trans. Signal Process.* **63** 6379–6394. [MR3421046](#)
- SHEN, D., SHEN, H. and MARRON, J. S. (2013). Consistency of sparse PCA in high dimension, low sample size contexts. *J. Multivariate Anal.* **115** 317–333. [MR3004561](#)
- SHEN, D., SHEN, H., ZHU, H. and MARRON, J. S. (2016). The statistics and mathematics of high dimension low sample size asymptotics. *Statist. Sinica* **26** 1747–1770. [MR3586237](#)
- TROPP, J. A., YURTSEVER, A., UDELL, M. and CEVHER, V. (2016). Randomized single-view algorithms for low-rank matrix approximation. Preprint. Available at [arXiv:1609.00048](#).
- VACCARO, R. J. (1994). A second-order perturbation expansion for the SVD. *SIAM J. Matrix Anal. Appl.* **15** 661–671. [MR1266609](#)
- VERSHYNIN, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge Univ. Press, Cambridge. [MR2963170](#)
- VU, V. Q. and LEI, J. (2013). Minimax sparse principal subspace estimation in high dimensions. *Ann. Statist.* **41** 2905–2947. [MR3161452](#)
- WANG, R. (2015). Singular vector perturbation under Gaussian noise. *SIAM J. Matrix Anal. Appl.* **36** 158–177. [MR3310977](#)
- WANG, W. and FAN, J. (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. *Ann. Statist.* **45** 1342–1374. [MR3662457](#)
- WEI, X. and MINSKER, S. (2017). Estimation of the covariance structure of heavy-tailed distributions. In *Advances in Neural Information Processing Systems* 2855–2864.
- XU, Z. (2002). Perturbation analysis for subspace decomposition with applications in subspace-based algorithms. *IEEE Trans. Signal Process.* **50** 2820–2830.
- YU, Y., WANG, T. and SAMWORTH, R. J. (2015). A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika* **102** 315–323. [MR3371006](#)
- ZHANG, Y., DUCHI, J. C. and WAINWRIGHT, M. J. (2013). Divide and conquer kernel ridge regression. In *COLT* 592–617.

ADDITIVE MODELS WITH TREND FILTERING

BY VEERANJANEYULU SADHANALA AND RYAN J. TIBSHIRANI

Carnegie Mellon University

We study additive models built with trend filtering, that is, additive models whose components are each regularized by the (discrete) total variation of their k th (discrete) derivative, for a chosen integer $k \geq 0$. This results in k th degree piecewise polynomial components, (e.g., $k = 0$ gives piecewise constant components, $k = 1$ gives piecewise linear, $k = 2$ gives piecewise quadratic, etc.). Analogous to its advantages in the univariate case, additive trend filtering has favorable theoretical and computational properties, thanks in large part to the localized nature of the (discrete) total variation regularizer that it uses. On the theory side, we derive fast error rates for additive trend filtering estimates, and show these rates are minimax optimal when the underlying function is additive and has component functions whose derivatives are of bounded variation. We also show that these rates are unattainable by additive smoothing splines (and by additive models built from linear smoothers, in general). On the computational side, we use backfitting, to leverage fast univariate trend filtering solvers; we also describe a new backfitting algorithm whose iterations can be run in parallel, which (as far as we can tell) is the first of its kind. Lastly, we present a number of experiments to examine the empirical performance of trend filtering.

REFERENCES

- BOYD, S., PARIKH, N., CHU, E., PELEATO, B. and ECKSTEIN, J. (2011). Distributed optimization and statistical learning via the alternative direction method of multipliers. *Found. Trends Mach. Learn.* **3** 1–122.
- BREIMAN, L. and FRIEDMAN, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *J. Amer. Statist. Assoc.* **80** 580–619. [MR0803258](#)
- BUJA, A., HASTIE, T. and TIBSHIRANI, R. (1989). Linear smoothers and additive models. *Ann. Statist.* **17** 453–555. [MR0994249](#)
- DE BOOR, C. (1978). *A Practical Guide to Splines*. *Applied Mathematical Sciences* **27**. Springer, New York. [MR0507062](#)
- DONOHO, D. L. and JOHNSTONE, I. M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. [MR1635414](#)
- EFRON, B. (1986). How biased is the apparent error rate of a prediction rule? *J. Amer. Statist. Assoc.* **81** 461–470. [MR0845884](#)
- FAHRMEIR, L. and LANG, S. (2001). Bayesian inference for generalized additive mixed models based on Markov random field priors. *J. Roy. Statist. Soc. Ser. C* **50** 201–220. [MR1833273](#)
- FRIEDMAN, J. H. and STUETZLE, W. (1981). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823. [MR0650892](#)

MSC2010 subject classifications. 62G08, 62G20.

Key words and phrases. Additive models, trend filtering, smoothing splines, nonparametric regression, minimax rates, parallel backfitting.

- GU, C. and WAHBA, G. (1991). Minimizing GCV/GML scores with multiple smoothing parameters via the Newton method. *SIAM J. Sci. Statist. Comput.* **12** 383–398. [MR1087766](#)
- HARIS, A., SIMON, N. and SHOJAIE, A. (2018). Generalized sparse additive models. Available at <https://arxiv.org/abs/1903.04641>.
- HASTIE, T. (1983). Non-parametric logistic regression. Technical report, Stanford Univ., Stanford, CA.
- HASTIE, T. J. and TIBSHIRANI, R. J. (1990). *Generalized Additive Models. Monographs on Statistics and Applied Probability* **43**. CRC Press, London. [MR1082147](#)
- KIM, Y.-J. and GU, C. (2004). Smoothing spline Gaussian regression: More scalable computation via efficient approximation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 337–356. [MR2062380](#)
- KIM, S.-J., KOH, K., BOYD, S. and GORINEVSKY, D. (2009). l_1 trend filtering. *SIAM Rev.* **51** 339–360. [MR2505584](#)
- LIN, Y. and ZHANG, H. H. (2006). Component selection and smoothing in multivariate nonparametric regression. *Ann. Statist.* **34** 2272–2297. [MR2291500](#)
- LOU, Y., BIEN, J., CARUANA, R. and GEHRKE, J. (2016). Sparse partially linear additive models. *J. Comput. Graph. Statist.* **25** 1026–1040. [MR3572032](#)
- MAMMEN, E. and VAN DE GEER, S. (1997). Locally adaptive regression splines. *Ann. Statist.* **25** 387–413. [MR1429931](#)
- MEIER, L., VAN DE GEER, S. and BÜHLMANN, P. (2009). High-dimensional additive modeling. *Ann. Statist.* **37** 3779–3821. [MR2572443](#)
- PETERSEN, A., WITTEN, D. and SIMON, N. (2016). Fused lasso additive model. *J. Comput. Graph. Statist.* **25** 1005–1025. [MR3572026](#)
- RAMDAS, A. and TIBSHIRANI, R. J. (2016). Fast and flexible ADMM algorithms for trend filtering. *J. Comput. Graph. Statist.* **25** 839–858. [MR3533641](#)
- RAVIKUMAR, P., LAFFERTY, J., LIU, H. and WASSERMAN, L. (2009). Sparse additive models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 1009–1030. [MR2750255](#)
- RUDIN, L. I., OSHER, S. and FATEMI, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D* **60** 259–268. [MR3363401](#)
- RUE, H., MARTINO, S. and CHOPIN, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 319–392. [MR2649602](#)
- RUPPERT, D., WAND, M. P. and CARROLL, R. J. (2003). *Semiparametric Regression. Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge Univ. Press, Cambridge. [MR1998720](#)
- SADHANALA, V. and TIBSHIRANI, R. J. (2019). Supplement to “Additive Models with Trend Filtering.” DOI:[10.1214/19-AOS1833SUPP](#).
- SARDY, S. and TSENG, P. (2004). AMlet, RAMlet, and GAMlet: Automatic nonlinear fitting of additive models, robust and generalized, with wavelets. *J. Comput. Graph. Statist.* **13** 283–309. [MR2063986](#)
- STEIDL, G., DIDAS, S. and NEUMANN, J. (2006). Splines in higher order TV regularization. *Int. J. Comput. Vis.* **70** 214–255.
- STONE, C. J. (1985). Additive regression and other nonparametric models. *Ann. Statist.* **13** 689–705. [MR0790566](#)
- TAN, Z. and ZHANG, C.-H. (2017). Penalized estimation in additive regression with high-dimensional data. Available at [arXiv:1704.07229](#).
- TIBSHIRANI, R. J. (1983). Non-parametric estimation of relative risk. Technical report, Stanford Univ., Stanford, CA.
- TIBSHIRANI, R. J. (2014). Adaptive piecewise polynomial estimation via trend filtering. *Ann. Statist.* **42** 285–323. [MR3189487](#)
- TIBSHIRANI, R. J. (2015). Degrees of freedom and model search. *Statist. Sinica* **25** 1265–1296. [MR3410308](#)

- TIBSHIRANI, R. J. (2017). Dykstra's algorithm, ADMM, and coordinate descent: Connections, insights, and extensions. *Adv. Neural Inf. Process. Syst.* **30**.
- TIBSHIRANI, R. J. and TAYLOR, J. (2011). The solution path of the generalized lasso. *Ann. Statist.* **39** 1335–1371. [MR2850205](#)
- TIBSHIRANI, R. J. and TAYLOR, J. (2012). Degrees of freedom in lasso problems. *Ann. Statist.* **40** 1198–1232. [MR2985948](#)
- TIBSHIRANI, R., SAUNDERS, M., ROSSET, S., ZHU, J. and KNIGHT, K. (2005). Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 91–108. [MR2136641](#)
- TSENG, P. (2001). Convergence of a block coordinate descent method for nondifferentiable minimization. *J. Optim. Theory Appl.* **109** 475–494. [MR1835069](#)
- VAN DE GEER, S. (2014). On the uniform convergence of empirical norms and inner products, with application to causal inference. *Electron. J. Stat.* **8** 543–574. [MR3211024](#)
- VAN DER BURG, E. and DE LEEUW, J. (1983). Non-linear canonical correlation. *Br. J. Math. Stat. Psychol.* **36** 54–80.
- WANG, Y.-X., SMOLA, A. and TIBSHIRANI, R. J. (2014). The falling factorial basis and its statistical applications. *Int. Conf. Mach. Learn.* **31**.
- WOOD, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 413–428. [MR1749600](#)
- WOOD, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Amer. Statist. Assoc.* **99** 673–686. [MR2090902](#)
- WOOD, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 3–36. [MR2797734](#)
- WOOD, S. N. (2017). *Generalized Additive Models: An Introduction with R*. CRC Press, Boca Raton, FL. [MR3726911](#)
- WOOD, S. N., GOUDE, Y. and SHAW, S. (2015). Generalized additive models for large data sets. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **64** 139–155. [MR3293922](#)
- WOOD, S. N., PYA, N. and SÄFKEN, B. (2016). Smoothing parameter and model selection for general smooth models. *J. Amer. Statist. Assoc.* **111** 1548–1563. [MR3601714](#)
- ZHANG, S. and WONG, M.-Y. (2003). Wavelet threshold estimation for additive regression models. *Ann. Statist.* **31** 152–173. [MR1962502](#)

SORTED CONCAVE PENALIZED REGRESSION

BY LONG FENG AND CUN-HUI ZHANG¹

City University of Hong Kong and Rutgers University

The Lasso is biased. Concave penalized least squares estimation (PLSE) takes advantage of signal strength to reduce this bias, leading to sharper error bounds in prediction, coefficient estimation and variable selection. For prediction and estimation, the bias of the Lasso can be also reduced by taking a smaller penalty level than what selection consistency requires, but such smaller penalty level depends on the sparsity of the true coefficient vector. The sorted ℓ_1 penalized estimation (Slope) was proposed for adaptation to such smaller penalty levels. However, the advantages of concave PLSE and Slope do not subsume each other. We propose sorted concave penalized estimation to combine the advantages of concave and sorted penalizations. We prove that sorted concave penalties adaptively choose the smaller penalty level and at the same time benefits from signal strength, especially when a significant proportion of signals are stronger than the corresponding adaptively selected penalty levels. A local convex approximation for sorted concave penalties, which extends the local linear and quadratic approximations for separable concave penalties, is developed to facilitate the computation of sorted concave PLSE and proven to possess desired prediction and estimation error bounds. Our analysis of prediction and estimation errors requires the restricted eigenvalue condition on the design, not beyond, and provides selection consistency under a required minimum signal strength condition in addition. Thus, our results also sharpens existing results on concave PLSE by removing the upper sparse eigenvalue component of the sparse Riesz condition.

REFERENCES

- [1] AGARWAL, A., NEGAHBAN, S. and WAINWRIGHT, M. J. (2012). Fast global convergence of gradient methods for high-dimensional statistical recovery. *Ann. Statist.* **40** 2452–2482. [MR3097609](#)
- [2] BECK, A. and TEBoulLE, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2** 183–202. [MR2486527](#)
- [3] BELLEC, P. C., LECUÉ, G. and TSYBAKOV, A. B. (2018). Slope meets Lasso: Improved oracle bounds and optimality. *Ann. Statist.* **46** 3603–3642. [MR3852663](#)
- [4] BELLONI, A. and CHERNOZHUKOV, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19** 521–547. [MR3037163](#)
- [5] BICKEL, P. J., RITOV, Y. and TSYBAKOV, A. B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. [MR2533469](#)

MSC2010 subject classifications. Primary 62J05, 62J07; secondary 62H12.

Key words and phrases. Penalized least squares, sorted penalties, concave penalties, Slope, local convex approximation, restricted eigenvalue, minimax rate, signal strength.

- [6] BOGDAN, M., VAN DEN BERG, E., SABATTI, C., SU, W. and CANDÈS, E. J. (2015). SLOPE—Adaptive variable selection via convex optimization. *Ann. Appl. Stat.* **9** 1103–1140. [MR3418717](#)
- [7] CANDÈS, E. and TAO, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n . *Ann. Statist.* **35** 2313–2351. [MR2382644](#)
- [8] CANDÈS, E. J. and TAO, T. (2005). Decoding by linear programming. *IEEE Trans. Inform. Theory* **51** 4203–4215. [MR2243152](#)
- [9] DALALYAN, A. and TSYBAKOV, A. B. (2008). Aggregation by exponential weighting, sharp pac-Bayesian bounds and sparsity. *Mach. Learn.* **72** 39–61.
- [10] EFRON, B., HASTIE, T., JOHNSTONE, I. and TIBSHIRANI, R. (2004). Least angle regression. *Ann. Statist.* **32** 407–499. With discussion, and a rejoinder by the authors. [MR2060166](#)
- [11] FAN, J. and LI, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. [MR1946581](#)
- [12] FAN, J., LIU, H., SUN, Q. and ZHANG, T. (2018). I-LAMM for sparse learning: Simultaneous control of algorithmic complexity and statistical error. *Ann. Statist.* **46** 814–841. [MR3782385](#)
- [13] FENG, L. and ZHANG, C.-H. (2019). Supplement to “Sorted concave penalized regression.” DOI:[10.1214/18-AOS1759SUPP](#).
- [14] HUANG, J. and ZHANG, C.-H. (2012). Estimation and selection via absolute penalized convex minimization and its multistage adaptive applications. *J. Mach. Learn. Res.* **13** 1839–1864. [MR2956344](#)
- [15] LOH, P.-L. and WAINWRIGHT, M. J. (2015). Regularized M -estimators with nonconvexity: Statistical and algorithmic theory for local optima. *J. Mach. Learn. Res.* **16** 559–616. [MR3335800](#)
- [16] MEINSHAUSEN, N. and BÜHLMANN, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. [MR2278363](#)
- [17] NEGAHBAN, S. N., RAVIKUMAR, P., WAINWRIGHT, M. J. and YU, B. (2012). A unified framework for high-dimensional analysis of M -estimators with decomposable regularizers. *Statist. Sci.* **27** 538–557. [MR3025133](#)
- [18] NESTEROV, Y. (2007). Gradient methods for minimizing composite functions. *Math. Program.* **140** 125–161. [MR3071865](#)
- [19] OSBORNE, M. R., PRESNELL, B. and TURLACH, B. A. (2000). A new approach to variable selection in least squares problems. *IMA J. Numer. Anal.* **20** 389–403. [MR1773265](#)
- [20] OSBORNE, M. R., PRESNELL, B. and TURLACH, B. A. (2000). On the LASSO and its dual. *J. Comput. Graph. Statist.* **9** 319–337. [MR1822089](#)
- [21] PARIKH, N. and BOYD, S. (2013). Proximal algorithms. In *Foundations and Trends in Optimization*.
- [22] ROČKOVÁ, V. and GEORGE, E. I. (2018). The spike-and-slab LASSO. *J. Amer. Statist. Assoc.* **113** 431–444. [MR3803476](#)
- [23] RUDELSON, M. and ZHOU, S. (2013). Reconstruction from anisotropic random measurements. *IEEE Trans. Inform. Theory* **59** 3434–3447. [MR3061256](#)
- [24] SU, W. and CANDÈS, E. (2016). SLOPE is adaptive to unknown sparsity and asymptotically minimax. *Ann. Statist.* **44** 1038–1068. [MR3485953](#)
- [25] SUN, T. and ZHANG, C.-H. (2012). Scaled sparse linear regression. *Biometrika* **99** 879–898. [MR2999166](#)
- [26] SUN, T. and ZHANG, C.-H. (2013). Sparse matrix inversion with scaled lasso. *J. Mach. Learn. Res.* **14** 3385–3418. [MR3144466](#)
- [27] TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. [MR1379242](#)
- [28] TROPP, J. A. (2006). Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Trans. Inform. Theory* **52** 1030–1051. [MR2238069](#)

- [29] VAN DE GEER, S. A. and BÜHLMANN, P. (2009). On the conditions used to prove oracle results for the Lasso. *Electron. J. Stat.* **3** 1360–1392. [MR2576316](#)
- [30] WAINWRIGHT, M. J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programming (Lasso). *IEEE Trans. Inform. Theory* **55** 2183–2202. [MR2729873](#)
- [31] WANG, Z., LIU, H. and ZHANG, T. (2014). Optimal computational and statistical rates of convergence for sparse nonconvex learning problems. *Ann. Statist.* **42** 2164–2201. [MR3269977](#)
- [32] YE, F. and ZHANG, C.-H. (2010). Rate minimaxity of the Lasso and Dantzig selector for the ℓ_q loss in ℓ_r balls. *J. Mach. Learn. Res.* **11** 3519–3540. [MR2756192](#)
- [33] ZHANG, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. [MR2604701](#)
- [34] ZHANG, C.-H. and HUANG, J. (2008). The sparsity and bias of the LASSO selection in high-dimensional linear regression. *Ann. Statist.* **36** 1567–1594. [MR2435448](#)
- [35] ZHANG, C.-H. and ZHANG, T. (2012). A general theory of concave regularization for high-dimensional sparse estimation problems. *Statist. Sci.* **27** 576–593. [MR3025135](#)
- [36] ZHANG, T. (2010). Analysis of multi-stage convex relaxation for sparse regularization. *J. Mach. Learn. Res.* **11** 1081–1107. [MR2629825](#)
- [37] ZHAO, P. and YU, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.* **7** 2541–2563. [MR2274449](#)
- [38] ZOU, H. and LI, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Ann. Statist.* **36** 1509–1533. [MR2435443](#)

ACTIVE RANKING FROM PAIRWISE COMPARISONS AND WHEN PARAMETRIC ASSUMPTIONS DO NOT HELP

BY REINHARD HECKEL^{*,1}, NIHAR B. SHAH^{†,2,3}, KANNAN RAMCHANDRAN[‡]
 AND MARTIN J. WAINWRIGHT^{‡,2}

Rice University^{*}, *Carnegie Mellon University*[†] and *University of California,
 Berkeley*[‡]

We consider sequential or active ranking of a set of n items based on noisy pairwise comparisons. Items are ranked according to the probability that a given item beats a randomly chosen item, and ranking refers to partitioning the items into sets of prespecified sizes according to their scores. This notion of ranking includes as special cases the identification of the top- k items and the total ordering of the items. We first analyze a sequential ranking algorithm that counts the number of comparisons won, and uses these counts to decide whether to stop, or to compare another pair of items, chosen based on confidence intervals specified by the data collected up to that point. We prove that this algorithm succeeds in recovering the ranking using a number of comparisons that is optimal up to logarithmic factors. This guarantee does depend on whether or not the underlying pairwise probability matrix, satisfies a particular structural property, unlike a significant body of past work on pairwise ranking based on parametric models such as the Thurstone or Bradley–Terry–Luce models. It has been a long-standing open question as to whether or not imposing these parametric assumptions allows for improved ranking algorithms. For stochastic comparison models, in which the pairwise probabilities are bounded away from zero, our second contribution is to resolve this issue by proving a lower bound for parametric models. This shows, perhaps surprisingly, that these popular parametric modeling choices offer at most logarithmic gains for stochastic comparisons.

REFERENCES

- [1] AGGARWAL, C. C. (2016). *Recommender Systems: The Textbook*. Springer, Berlin.
- [2] AILON, N. (2011). Active learning ranking from pairwise preferences with almost optimal query complexity. In *Advances in Neural Information Processing Systems* 810–818.
- [3] BALLINGER, T. P. and WILCOX, N. T. (1997). Decisions, error and heterogeneity. *Econ. J.* **107** 1090–1105.
- [4] BOYD, S. and VANDENBERGHE, L. (2004). *Convex Optimization*. Cambridge Univ. Press, Cambridge.
- [5] BRADLEY, R. A. and TERRY, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39** 324–345.
- [6] BUBECK, S. and CESA-BIANCHI, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Found. Trends Mach. Learn.* **5**.

MSC2010 subject classifications. 68T99.

Key words and phrases. Active learning, online learning, pairwise comparisons, ranking, multi-armed bandits, Bradley–Terry–Luce model.

- [7] BUBECK, S., WANG, T. and VISWANATHAN, N. (2013). Multiple identifications in multi-armed bandits. In *International Conference on Machine Learning* 258–265.
- [8] BUSA-FEKETE, R., SZORENYI, B., CHENG, W., WENG, P. and HÜLLERMEIER, E. (2013). Top-k selection based on adaptive sampling of noisy preferences. In *International Conference on Machine Learning* 1094–1102.
- [9] CHEN, S., LIN, T., KING, I., LYU, M. R. and CHEN, W. (2014). Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems* 379–387.
- [10] CHEN, X., GOPI, S., MAO, J. and SCHNEIDER, J. (2017). Competitive analysis of the top- K ranking problem. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms* 1245–1264.
- [11] CHEN, Y. and SUH, C. (2015). Spectral MLE: Top-k rank aggregation from pairwise comparisons. In *International Conference on Machine Learning* 371–380.
- [12] DE BORDA, J. C. (1781). Mémoire sur les élections au scrutin. *Histoire de l'Académie Royale des Sciences*.
- [13] ERIKSSON, B. (2013). Learning to top-k search using pairwise comparisons. In *International Conference on Machine Learning* 265–273.
- [14] EVEN-DAR, E., MANNOR, S. and MANSOUR, Y. (2006). Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *J. Mach. Learn. Res.* **7** 1079–1105.
- [15] HAJEK, B., OH, S. and XU, J. (2014). Minimax-optimal inference from partial rankings. In *Advances in Neural Information Processing Systems* 1475–1483.
- [16] HECKEL, R., SHAH, N. B., RAMCHANDRAN, K. and WAINWRIGHT, M. J. (2019). Supplement to “Active ranking from pairwise comparisons and when parametric assumptions do not help.” DOI:[10.1214/18-AOS1772SUPP](https://doi.org/10.1214/18-AOS1772SUPP).
- [17] HUNTER, D. R. (2004). MM algorithms for generalized Bradley–Terry models. *Ann. Statist.* **32** 384–406.
- [18] JAMIESON, K., KATARIYA, S., DESHPANDE, A. and NOWAK, R. (2015). Sparse dueling bandits. In *International Conference on Artificial Intelligence and Statistics* 416–424.
- [19] JAMIESON, K., MALLOY, M., NOWAK, R. and BUBECK, S. (2014). *lil’* UCB: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory* 423–439.
- [20] JAMIESON, K. and NOWAK, R. (2011). Active ranking using pairwise comparisons. In *Advances in Neural Information Processing Systems* 2240–2248.
- [21] JOE, H. (1988). Majorization, entropy and paired comparisons. *Ann. Statist.* **16** 915–925. [MR0947585](https://doi.org/10.1214/aos/1176345585)
- [22] KALYANAKRISHNAN, S., TEWARI, A., AUER, P. and STONE, P. (2012). PAC subset selection in stochastic multi-armed bandits. In *International Conference on Machine Learning* 655–662.
- [23] KAUFMANN, E., CAPPÉ, O. and GARIVIER, A. (2016). On the complexity of best-arm identification in multi-armed bandit models. *J. Mach. Learn. Res.* **17** Paper No. 1, 42.
- [24] LANDAU, H. G. (1953). On dominance relations and the structure of animal societies. III. The condition for a score structure. *Bull. Math. Biophys.* **15** 143–148.
- [25] LUCE, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York.
- [26] MAYSTRE, L. and GROSSGLAUSER, M. (2015). Robust active ranking from sparse noisy comparisons. Available at [arXiv:1502.05556](https://arxiv.org/abs/1502.05556).
- [27] NEGAHBAN, S. N., OH, S. and SHAH, D. (2012). Iterative ranking from pair-wise comparisons. In *Advances in Neural Information Processing Systems* 2474–2482.
- [28] PAULSON, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. *Ann. Math. Stat.* **35** 174–180.

- [29] PIECH, C., HUANG, J., CHEN, Z., DO, C., NG, A. and KOLLER, D. (2013). Tuned models of peer assessment in MOOCs. In *International Conference on Educational Data Mining*.
- [30] SALGANIK, M. J. and LEVY, K. E. C. (2015). Wiki surveys: Open and quantifiable social data collection. *PLoS ONE* **10** e0123483.
- [31] SHAH, N. B., BALAKRISHNAN, S., BRADLEY, J., PAREKH, A., RAMCHANDRAN, K. and WAINWRIGHT, M. J. (2016). Estimation from pairwise comparisons: Sharp minimax bounds with topology dependence. *J. Mach. Learn. Res.* **17** Paper No. 58, 47.
- [32] SHAH, N. B., BALAKRISHNAN, S., GUNTUBOYINA, A. and WAINWRIGHT, M. J. (2017). Stochastically transitive models for pairwise comparisons: Statistical and computational issues. *IEEE Trans. Inform. Theory* **63** 934–959.
- [33] SHAH, N. B., BRADLEY, J. K., PAREKH, A., WAINWRIGHT, M. J. and RAMCHANDRAN, K. (2013). A case for ordinal peer-evaluation in MOOCs. In *NIPS Workshop on Data Driven Education*.
- [34] SHAH, N. B. and WAINWRIGHT, M. J. (2017). Simple, robust and optimal ranking from pairwise comparisons. *J. Mach. Learn. Res.* **18** Paper No. 199, 38.
- [35] SOUFIANI, H., PARKES, D. and XIA, L. (2014). Computing parametric ranking models via rank-breaking. In *International Conference on Machine Learning* 360–368.
- [36] SZÖRÉNYI, B., BUSA-FEKETE, R., PAUL, A. and HÜLLERMEIER, E. (2015). Online rank elicitation for Plackett–Luce: A dueling bandits approach. In *Advances in Neural Information Processing Systems* 604–612.
- [37] THURSTONE, L. (1927). A law of comparative judgment. *Psychol. Rev.* **34** 273–286.
- [38] TVERSKY, A. and EDWARD RUSSO, J. (1969). Substitutability and similarity in binary choices. *J. Math. Psych.* **6** 1–12.
- [39] URVOY, T., CLEROT, F., FÉRAUD, R. and NAAMANE, S. (2013). Generic exploration and K -armed voting bandits. In *International Conference on Machine Learning* 91–99.
- [40] YUE, Y., BRODER, J., KLEINBERG, R. and JOACHIMS, T. (2012). The K -armed dueling bandits problem. *J. Comput. System Sci.* **78** 1538–1556.
- [41] YUE, Y. and JOACHIMS, T. (2011). Beat the mean bandit. In *International Conference on Machine Learning* 241–248.

RANDOMIZED INCOMPLETE U -STATISTICS IN HIGH DIMENSIONS²

BY XIAOHUI CHEN¹ AND KENGO KATO

University of Illinois at Urbana-Champaign and Cornell University

This paper studies inference for the mean vector of a high-dimensional U -statistic. In the era of big data, the dimension d of the U -statistic and the sample size n of the observations tend to be both large, and the computation of the U -statistic is prohibitively demanding. Data-dependent inferential procedures such as the empirical bootstrap for U -statistics is even more computationally expensive. To overcome such a computational bottleneck, incomplete U -statistics obtained by sampling fewer terms of the U -statistic are attractive alternatives. In this paper, we introduce randomized incomplete U -statistics with sparse weights whose computational cost can be made independent of the order of the U -statistic. We derive nonasymptotic Gaussian approximation error bounds for the randomized incomplete U -statistics in high dimensions, namely in cases where the dimension d is possibly much larger than the sample size n , for both nondegenerate and degenerate kernels. In addition, we propose generic bootstrap methods for the incomplete U -statistics that are computationally much less demanding than existing bootstrap methods, and establish finite sample validity of the proposed bootstrap methods. Our methods are illustrated on the application to nonparametric testing for the pairwise independence of a high-dimensional random vector under weaker assumptions than those appearing in the literature.

REFERENCES

- [1] ARCONES, M. A. and GINÉ, E. (1992). On the bootstrap of U and V statistics. *Ann. Statist.* **20** 655–674. [MR1165586](#)
- [2] BERGSMA, W. and DASSIOS, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20** 1006–1028. [MR3178526](#)
- [3] BERTAIL, P. and TRESSOU, J. (2006). Incomplete generalized U -statistics for food risk assessment. *Biometrics* **62** 66–74, 315. [MR2226558](#)
- [4] BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217. [MR0630103](#)
- [5] BLOM, G. (1976). Some properties of incomplete U -statistics. *Biometrika* **63** 573–580. [MR0474582](#)
- [6] BRETAGNOLLE, J. (1983). Lois limites du bootstrap de certaines fonctionnelles. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **19** 281–296. [MR0725561](#)
- [7] BROWN, B. M. and KILDEA, D. G. (1978). Reduced U -statistics and the Hodges–Lehmann estimator. *Ann. Statist.* **6** 828–835. [MR0491556](#)

MSC2010 subject classifications. Primary 62E17, 62F40; secondary 62H15.

Key words and phrases. Incomplete U -statistics, randomized inference, Gaussian approximation, bootstrap, divide and conquer, Bernoulli sampling, sampling with replacement.

- [8] CHEN, X. (2018). Gaussian and bootstrap approximations for high-dimensional U -statistics and their applications. *Ann. Statist.* **46** 642–678. [MR3782380](#)
- [9] CHEN, X. and KATO, K. (2017). Jackknife multiplier bootstrap: Finite sample approximations to the U -process supremum with applications. Available at [arXiv:1708.02705](#).
- [10] CHEN, X. and KATO, K. (2019). Supplement to “Randomized incomplete U -statistics in high dimensions.” DOI:[10.1214/18-AOS1773SUPP](#).
- [11] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. [MR3161448](#)
- [12] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **45** 2309–2352. [MR3693963](#)
- [13] CLÉMENÇON, S., COLIN, I. and BELLET, A. (2016). Scaling-up empirical risk minimization: Optimization of incomplete U -statistics. *J. Mach. Learn. Res.* **17** Paper No. 76, 36. [MR3517099](#)
- [14] DEHLING, H. and MIKOSCH, T. (1994). Random quadratic forms and the bootstrap for U -statistics. *J. Multivariate Anal.* **51** 392–413. [MR1321305](#)
- [15] EMBRECHTS, P., LINDSKOG, F. and MCNEIL, A. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ed.) 8. North-Holland, Amsterdam.
- [16] GEISSER, S. and MANTEL, N. (1962). Pairwise independence of jointly dependent variables. *Ann. Math. Statist.* **33** 290–291. [MR0137188](#)
- [17] GU, Q., CAO, Y., NING, Y. and LIU, H. (2015). Local and global inference for high dimensional nonparanormal graphical models. Available at [arXiv:1502.02347](#).
- [18] HAN, F., CHEN, S. and LIU, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. [MR3737306](#)
- [19] HAN, F. and QIAN, T. (2016). Asymptotics for asymmetric weighted U -statistics: Central limit theorem and bootstrap under data heterogeneity. Preprint.
- [20] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19** 293–325. [MR0026294](#)
- [21] HOEFFDING, W. (1948). A non-parametric test of independence. *Ann. Math. Statistics* **19** 546–557. [MR0029139](#)
- [22] HSING, T. and WU, W. B. (2004). On weighted U -statistics for stationary processes. *Ann. Probab.* **32** 1600–1631. [MR2060311](#)
- [23] HUŠKOVÁ, M. and JANSSEN, P. (1993). Consistency of the generalized bootstrap for degenerate U -statistics. *Ann. Statist.* **21** 1811–1823. [MR1245770](#)
- [24] HUŠKOVÁ, M. and JANSSEN, P. (1993). Generalized bootstrap for studentized U -statistics: A rank statistic approach. *Statist. Probab. Lett.* **16** 225–233. [MR1208512](#)
- [25] JANSON, S. (1984). The asymptotic distributions of incomplete U -statistics. *Z. Wahrsch. Verw. Gebiete* **66** 495–505. [MR0753810](#)
- [26] KLEINER, A., TALWALKAR, A., SARKAR, P. and JORDAN, M. I. (2014). A scalable bootstrap for massive data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 795–816. [MR3248677](#)
- [27] LEE, A. J. (1990). *U-Statistics. Theory and Practice. Statistics: Textbooks and Monographs* **110**. Dekker, New York. [MR1075417](#)
- [28] LEUNG, D. and DRTON, M. (2018). Testing independence in high dimensions with sums of rank correlations. *Ann. Statist.* **46** 280–307. [MR3766953](#)
- [29] MAJOR, P. (1994). Asymptotic distributions for weighted U -statistics. *Ann. Probab.* **22** 1514–1535. [MR1303652](#)
- [30] MENTCH, L. and HOOKER, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *J. Mach. Learn. Res.* **17** Paper No. 26, 41. [MR3491120](#)

- [31] NANDY, P., WEIHS, L. and DRTON, M. (2016). Large-sample theory for the Bergsma–Dassios sign covariance. *Electron. J. Stat.* **10** 2287–2311. [MR3541972](#)
- [32] O’NEIL, K. A. and REDNER, R. A. (1993). Asymptotic distributions of weighted U -statistics of degree 2. *Ann. Probab.* **21** 1159–1169. [MR1217584](#)
- [33] RIFI, M. and UTZET, F. (2000). On the asymptotic behavior of weighted U -statistics. *J. Theoret. Probab.* **13** 141–167. [MR1744988](#)
- [34] RUBIN, H. and VITALE, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** 165–170. [MR0557561](#)
- [35] SHAPIRO, C. P. and HUBERT, L. (1979). Asymptotic normality of permutation statistics derived from weighted sums of bivariate functions. *Ann. Statist.* **7** 788–794. [MR0532242](#)
- [36] SZÉKELY, G. J., RIZZO, M. L. and BAKIROV, N. K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. [MR2382665](#)
- [37] VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. [MR1652247](#)
- [38] VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics. Springer Series in Statistics*. Springer, New York. [MR1385671](#)
- [39] WANG, Q. and JING, B.-Y. (2004). Weighted bootstrap for U -statistics. *J. Multivariate Anal.* **91** 177–198. [MR2087842](#)
- [40] YAO, S., ZHANG, X. and SHAO, X. (2018). Testing mutual independence in high dimension via distance covariance. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 455–480. [MR3798874](#)
- [41] ZHANG, Y., DUCHI, J. and WAINWRIGHT, M. (2015). Divide and conquer kernel ridge regression: A distributed algorithm with minimax optimal rates. *J. Mach. Learn. Res.* **16** 3299–3340. [MR3450540](#)

ADAPTIVE ESTIMATION OF THE RANK OF THE COEFFICIENT MATRIX IN HIGH-DIMENSIONAL MULTIVARIATE RESPONSE REGRESSION MODELS

BY XIN BING AND MARTEN H. WEGKAMP¹

Cornell University

We consider the multivariate response regression problem with a regression coefficient matrix of low, unknown rank. In this setting, we analyze a new criterion for selecting the optimal reduced rank. This criterion differs notably from the one proposed in Bunea, She and Wegkamp (*Ann. Statist.* **39** (2011) 1282–1309) in that it does not require estimation of the unknown variance of the noise, nor does it depend on a delicate choice of a tuning parameter. We develop an iterative, fully data-driven procedure, that adapts to the optimal signal-to-noise ratio. This procedure finds the true rank in a few steps with overwhelming probability. At each step, our estimate increases, while at the same time it does not exceed the true rank. Our finite sample results hold for any sample size and any dimension, even when the number of responses and of covariates grow much faster than the number of observations. We perform an extensive simulation study that confirms our theoretical findings. The new method performs better and is more stable than the procedure of Bunea, She and Wegkamp (*Ann. Statist.* **39** (2011) 1282–1309) in both low- and high-dimensional settings.

REFERENCES

- ANDERSON, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. Math. Stat.* **22** 327–351. [MR0042664](#)
- ANDERSON, T. W. (1999). Asymptotic distribution of the reduced rank regression estimator under general conditions. *Ann. Statist.* **27** 1141–1154. [MR1740118](#)
- ANDERSON, T. W. (2002). Specification and misspecification in reduced rank regression. *Sankhyā Ser. A* **64** 193–205. [MR1981753](#)
- BAI, Z. D. and YIN, Y. Q. (1993). Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.* **21** 1275–1294. [MR1235416](#)
- BING, X. and WEGKAMP, M. H. (2019). Supplement to “Adaptive estimation of the rank of the coefficient matrix in high-dimensional multivariate response regression models.” DOI:[10.1214/18-AOS1774SUPP](#).
- BÜHLMANN, P. and VAN DE GEER, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Series in Statistics. Springer, Heidelberg. [MR2807761](#)
- BUNEA, F., SHE, Y. and WEGKAMP, M. H. (2011). Optimal selection of reduced rank estimators of high-dimensional matrices. *Ann. Statist.* **39** 1282–1309. [MR2816355](#)
- BUNEA, F., SHE, Y. and WEGKAMP, M. H. (2012). Joint variable and rank selection for parsimonious estimation of high-dimensional matrices. *Ann. Statist.* **40** 2359–2388. [MR3097606](#)

MSC2010 subject classifications. 62H15, 62J07.

Key words and phrases. Multivariate response regression, reduced rank estimator, self-tuning, adaptive rank estimation, rank consistency, dimension reduction, oracle inequalities.

- ECKART, C. and YOUNG, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1** 211–218.
- GIRAUD, C. (2011). Low rank multivariate regression. *Electron. J. Stat.* **5** 775–799. [MR2824816](#)
- GIRAUD, C. (2015). *Introduction to High-Dimensional Statistics. Monographs on Statistics and Applied Probability* **139**. CRC Press, Boca Raton, FL. [MR3307991](#)
- IZENMAN, A. J. (1975). Reduced-rank regression for the multivariate linear model. *J. Multivariate Anal.* **5** 248–264. [MR0373179](#)
- IZENMAN, A. J. (2008). *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning. Springer Texts in Statistics*. Springer, New York. [MR2445017](#)
- JOHNSTONE, I. M. (2001). Chi-square oracle inequalities. In *State of the Art in Probability and Statistics: Festschrift for Willem R. van Zwet* (M. de Gunst, C. Klaasen and A. van der Vaart, eds.) 399–418. IMS, Beachwood, OH. [MR1836572](#)
- LOUNICI, K., PONTIL, M., VAN DE GEER, S. and TSYBAKOV, A. B. (2011). Oracle inequalities and optimal inference under group sparsity. *Ann. Statist.* **39** 2164–2204. [MR2893865](#)
- NEGAHBAN, S. and WAINWRIGHT, M. J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.* **39** 1069–1097. [MR2816348](#)
- OBOZINSKI, G., WAINWRIGHT, M. J. and JORDAN, M. I. (2011). Support union recovery in high-dimensional multivariate regression. *Ann. Statist.* **39** 1–47. [MR2797839](#)
- RAO, C. R. (1978). Matrix approximations and reduction of dimensionality in multivariate statistical analysis. In *Multivariate Analysis V* 3–22. North-Holland, Amsterdam.
- REINSEL, G. C. and VELU, R. P. (1998). *Multivariate Reduced-Rank Regression: Theory and Applications. Lecture Notes in Statistics* **136**. Springer, New York. [MR1719704](#)
- ROBINSON, P. M. (1973). Generalized canonical analysis for time series. *J. Multivariate Anal.* **3** 141–160. [MR0326959](#)
- ROBINSON, P. M. (1974). Identification, estimation and large-sample theory for regressions containing unobservable variables. *Internat. Econom. Rev.* **15** 680–692. [MR0356376](#)
- ROHDE, A. and TSYBAKOV, A. B. (2011). Estimation of high-dimensional low-rank matrices. *Ann. Statist.* **39** 887–930. [MR2816342](#)
- RUDELSON, M. and VERSHYNIN, R. (2010). Non-asymptotic theory of random matrices: Extreme singular values. In *Proceedings of the International Congress of Mathematicians. Volume III* 1576–1602. Hindustan Book Agency, New Delhi. [MR2827856](#)
- SCHMIDT, E. (1907). Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener. *Math. Ann.* **63** 433–476.
- STEWART, G. W. (1993). On the early history of the singular value decomposition. *SIAM Rev.* **35** 551–566. [MR1247916](#)
- VERSHYNIN, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge Univ. Press, Cambridge. [MR2963170](#)
- WEYL, H. (1912). Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). *Math. Ann.* **71** 441–479. [MR1511670](#)
- YUAN, M. and LIN, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 49–67. [MR2212574](#)

STATISTICAL INFERENCE FOR AUTOREGRESSIVE MODELS UNDER HETEROSCEDASTICITY OF UNKNOWN FORM

BY KE ZHU¹

University of Hong Kong

This paper provides an entire inference procedure for the autoregressive model under (conditional) heteroscedasticity of unknown form with a finite variance. We first establish the asymptotic normality of the weighted least absolute deviations estimator (LADE) for the model. Second, we develop the random weighting (RW) method to estimate its asymptotic covariance matrix, leading to the implementation of the Wald test. Third, we construct a portmanteau test for model checking, and use the RW method to obtain its critical values. As a special weighted LADE, the feasible adaptive LADE (ALADE) is proposed and proved to have the same efficiency as its infeasible counterpart. The importance of our entire methodology based on the feasible ALADE is illustrated by simulation results and the real data analysis on three U.S. economic data sets.

REFERENCES

- AMADO, C. and TERÄSVIRTA, T. (2013). Modelling volatility by variance decomposition. *J. Econometrics* **175** 142–153. [MR3061936](#)
- AMADO, C. and TERÄSVIRTA, T. (2014). Modelling changes in the unconditional variance of long stock return series. *J. Empir. Finance* **25** 15–35.
- ANDREWS, D. W. K. (1988). Laws of large numbers for dependent nonidentically distributed random variables. *Econometric Theory* **4** 458–467. [MR0985156](#)
- BASSETT, G. JR. and KOENKER, R. (1978). Asymptotic theory of least absolute error regression. *J. Amer. Statist. Assoc.* **73** 618–622. [MR0514166](#)
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. [MR0233396](#)
- BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. [MR0853051](#)
- BUSETTI, F. and TAYLOR, A. M. R. (2003). Variance shifts, structural breaks, and stationarity tests. *J. Bus. Econom. Statist.* **21** 510–531. [MR2008043](#)
- CARRASCO, M. and CHEN, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18** 17–39. [MR1885348](#)
- CARROLL, R. J. (1982). Adapting for heteroscedasticity in linear models. *Ann. Statist.* **10** 1224–1233. [MR0673657](#)
- CAVALIERE, G. (2004). Unit root tests under time-varying variances. *Econometric Rev.* **23** 259–292. [MR2089641](#)
- CAVALIERE, G. and TAYLOR, A. M. R. (2007). Testing for unit roots in time series models with non-stationary volatility. *J. Econometrics* **140** 919–947. [MR2408932](#)

MSC2010 subject classifications. 62F03, 62F12, 62F35, 62M10.

Key words and phrases. Adaptive estimator, autoregressive model, conditional heteroscedasticity, heteroscedasticity, weighted least absolute deviations estimator, wild bootstrap.

- CHEN, B. and HONG, Y. (2016). Detecting for smooth structural changes in GARCH models. *Econometric Theory* **32** 740–791. [MR3506438](#)
- CRAGG, J. G. (1983). More efficient estimation in the presence of heteroscedasticity of unknown form. *Econometrica* **51** 751–763. [MR0712368](#)
- DAVIDSON, J. (2002). Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *J. Econometrics* **106** 243–269. [MR1885370](#)
- DAVIDSON, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *J. Bus. Econom. Statist.* **22** 16–29. [MR2028205](#)
- DAVIS, R. A. (1996). Gauss–Newton and M -estimation for ARMA processes with infinite variance. *Stochastic Process. Appl.* **63** 75–95. [MR1411191](#)
- DAVIS, R. A. and DUNSMUIR, W. T. M. (1997). Least absolute deviation estimation for regression with ARMA errors. Dedicated to Murray Rosenblatt. *J. Theoret. Probab.* **10** 481–497. [MR1455154](#)
- DIEBOLD, F. X. (1986). Modeling the persistence of conditional variances: A comment. *Econom. Rev.* **5** 51–56.
- DING, Z., GRANGER, C. W. J. and ENGLE, R. F. (1993). A long memory property of stock market returns and a new model. *J. Empir. Finance* **1** 83–106.
- DROST, F. C. and NIJMAN, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica* **61** 909–927. [MR1231681](#)
- ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. [MR0666121](#)
- ESCANCIANO, J. C. (2006). Goodness-of-fit tests for linear and nonlinear time series models. *J. Amer. Statist. Assoc.* **101** 531–541. [MR2256173](#)
- FAN, J. and YAO, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* **85** 645–660. [MR1665822](#)
- FRANCQ, C. and ZAKOÏAN, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley, Chichester. [MR3186556](#)
- FRYZLEWICZ, P., SAPATINAS, T. and SUBBA RAO, S. (2006). A Haar–Fisz technique for locally stationary volatility estimation. *Biometrika* **93** 687–704. [MR2261451](#)
- GIRAITIS, L., KOKOSZKA, P. and LEIPUS, R. (2000). Stationary ARCH models: Dependence structure and central limit theorem. *Econometric Theory* **16** 3–22. [MR1749017](#)
- GONÇALVES, S. and KILIAN, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *J. Econometrics* **123** 89–120. [MR2125439](#)
- GUTENBRUNNER, C. and JUREČKOVÁ, J. (1992). Regression rank scores and regression quantiles. *Ann. Statist.* **20** 305–330. [MR1150346](#)
- HONG, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64** 837–864. [MR1399220](#)
- HONG, Y. and LEE, Y.-J. (2005). Generalized spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form. *Rev. Econ. Stud.* **72** 499–541. [MR2129829](#)
- JIN, Z., YING, Z. and WEI, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika* **88** 381–390. [MR1844838](#)
- KATO, K. (2009). Asymptotics for argmin processes: Convexity arguments. *J. Multivariate Anal.* **100** 1816–1829. [MR2535389](#)
- KNIGHT, K. (1998). Limiting distributions for L_1 regression estimators under general conditions. *Ann. Statist.* **26** 755–770. [MR1626024](#)
- KUERSTEINER, G. M. (2002). Efficient IV estimation for autoregressive models with conditional heteroskedasticity. *Econometric Theory* **18** 547–583. [MR1906326](#)
- LIU, R. Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Ann. Statist.* **16** 1696–1708. [MR0964947](#)

- MIKOSCH, T. and STĂRICĂ, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Rev. Econ. Stat.* **86** 378–390.
- NICHOLLS, D. F. and PAGAN, A. R. (1983). Heteroscedasticity in models with lagged dependent variables. *Econometrica* **51** 1233–1242. [MR0710229](#)
- PAN, J., WANG, H. and YAO, Q. (2007). Weighted least absolute deviations estimation for ARMA models with infinite variance. *Econometric Theory* **23** 852–879. [MR2395837](#)
- PHILLIPS, P. C. B. and XU, K.-L. (2006). Inference in autoregression under heteroskedasticity. *J. Time Series Anal.* **27** 289–308. [MR2235847](#)
- POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer Series in Statistics. Springer, New York. [MR0762984](#)
- ROBINSON, P. M. (1987). Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* **55** 875–891. [MR0906567](#)
- ROBINSON, P. M. (1989). Nonparametric estimation of time-varying parameters. In *Statistical Analysis and Forecasting of Economic Structural Change* (P. Hackl, ed.) 253–264. Springer, Berlin.
- ROBINSON, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *J. Econometrics* **47** 67–84. [MR1087207](#)
- ROBINSON, P. M. (2012). Nonparametric trending regression with cross-sectional dependence. *J. Econometrics* **169** 4–14. [MR2935258](#)
- SENSIER, M. and VAN DIJK, D. (2004). Testing for volatility changes in U.S. macroeconomic time series. *Rev. Econ. Stat.* **86** 833–839.
- SHAO, Q.-M. and YU, H. (1996). Weak convergence for weighted empirical processes of dependent sequences. *Ann. Probab.* **24** 2098–2127. [MR1415243](#)
- STĂRICĂ, C. and GRANGER, C. (2005). Nonstationarities in stock returns. *Rev. Econ. Stat.* **87** 503–522.
- TSAY, R. S. (1988). Outliers, level shifts and variance changes in time series. *J. Forecast.* **7** 1–20.
- WATSON, M. W. (1999). Explaining the increased variability in long-term interest rates. *Econ. Q.-Fed. Reserve Bank Richmond* **85** 71–96.
- WU, C.-F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* **14** 1261–1350. [MR0868303](#)
- WU, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. [MR2172215](#)
- XU, K.-L. and PHILLIPS, P. C. B. (2008). Adaptive estimation of autoregressive models with time-varying variances. *J. Econometrics* **142** 265–280. [MR2408736](#)
- YAO, Q. and BROCKWELL, P. J. (2006). Gaussian maximum likelihood estimation for ARMA models. I. Time series. *J. Time Series Anal.* **27** 857–875. [MR2328545](#)
- YU, K. and JONES, M. C. (2004). Likelihood-based local linear estimation of the conditional variance function. *J. Amer. Statist. Assoc.* **99** 139–144. [MR2054293](#)
- ZHANG, R. and LING, S. (2015). Asymptotic inference for AR models with heavy-tailed G-GARCH noises. *Econometric Theory* **31** 880–890. [MR3377272](#)
- ZHAO, Q. (2001). Asymptotically efficient median regression in the presence of heteroskedasticity of unknown form. *Econometric Theory* **17** 765–784. [MR1862375](#)
- ZHOU, Z. and WU, W. B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. [MR2541444](#)
- ZHU, K. (2016). Bootstrapping the portmanteau tests in weak auto-regressive moving average models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 463–485. [MR3454205](#)
- ZHU, K. (2019). Supplement to “Statistical inference for autoregressive models under heteroscedasticity of unknown form.” DOI:[10.1214/18-AOS1775SUPP](#).
- ZHU, K. and LI, W. K. (2015). A bootstrapped spectral test for adequacy in weak ARMA models. *J. Econometrics* **187** 113–130. [MR3347298](#)
- ZHU, K. and LING, S. (2012). The global weighted LAD estimators for finite/infinite variance ARMA(p, q) models. *Econometric Theory* **28** 1065–1086. [MR2981698](#)

ZHU, K. and LING, S. (2015). LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises. *J. Amer. Statist. Assoc.* **110** 784–794. [MR3367264](#)

ON PARTIAL-SUM PROCESSES OF ARMAX RESIDUALS

BY STEFFEN GRØNNEBERG AND BENJAMIN HOLCBLAT

BI Norwegian Business School and Université du Luxembourg, LSF

We establish general and versatile results regarding the limit behavior of the partial-sum process of ARMAX residuals. Illustrations include ARMA with seasonal dummies, misspecified ARMAX models with autocorrelated errors, nonlinear ARMAX models, ARMA with a structural break, a wide range of ARMAX models with infinite-variance errors, weak GARCH models and the consistency of kernel estimation of the density of ARMAX errors. Our results identify the limit distributions, and provide a general algorithm to obtain pivot statistics for CUSUM tests.

REFERENCES

- [1] ANDREOU, E. and WERKER, B. J. M. (2012). An alternative asymptotic analysis of residual-based statistics. *Rev. Econ. Stat.* **94** 88–99.
- [2] BAI, J. (1993). On the partial sums of residuals in autoregressive and moving average models. *J. Time Series Anal.* **14** 247–260. [MR1221937](#)
- [3] BAI, J. (1993). Estimation of structural change based on Wald-type statistics. Technical Report. Dept. Economics, Massachusetts Institute of Technology, Cambridge, MA.
- [4] BAI, J. (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.* **22** 2051–2061. [MR1329182](#)
- [5] BAI, J. (1994). Least squares estimation of a shift in linear processes. *J. Time Series Anal.* **15** 453–472. [MR1292161](#)
- [6] BASRAK, B. and KRIZMANIĆ, D. (2014). A limit theorem for moving averages in the α -stable domain of attraction. *Stochastic Process. Appl.* **124** 1070–1083. [MR3138607](#)
- [7] BERENGUER-RICO, V. and NIELSEN, B. (2015). Cumulated sum of squares statistics for non-linear and non-stationary regressions. Technical Report. Nuffield Discussion Paper 2015-W09, Univ. Oxford.
- [8] BEVERIDGE, S. and NELSON, C. (1981). A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycles. *J. Monet. Econ.* **7** 151–174.
- [9] BIERENS, H. J. (1994). *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-Section and Time Series Models*. Cambridge Univ. Press, Cambridge. [MR1291390](#)
- [10] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. [MR0233396](#)
- [11] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. [MR1324786](#)

MSC2010 subject classifications. Primary 62M10, 60F17, 62J20; secondary 91B84.

Key words and phrases. ARMAX residuals, change-point problems, CUSUM test, misspecified ARMAX, autocorrelated errors, infinite-variance errors, heteroscedastic errors, linear and nonlinear regression, seasonal dummy, kernel density estimation of ARMAX errors, nonlinear ARMAX, weak GARCH.

- [12] BOLDIN, M. V. (1982). An estimate of the distribution of the noise in an autoregressive scheme. *Teor. Veroyatn. Primen.* **27** 805–810. [MR0681475](#)
- [13] BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. Springer, New York. [MR1093459](#)
- [14] BROWN, R. L., DURBIN, J. and EVANS, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. Roy. Statist. Soc. Ser. B* **37** 149–192. [MR0378310](#)
- [15] CHAN, N. H. and LING, S. (2008). Residual empirical processes for long and short memory time series. *Ann. Statist.* **36** 2453–2470. [MR2458194](#)
- [16] DAHLHAUS, R. and WEFELMEYER, W. (1996). Asymptotically optimal estimation in mis-specified time series models. *Ann. Statist.* **24** 952–974. [MR1401832](#)
- [17] DAVIDSON, J. (1994). *Stochastic Limit Theory. Advanced Texts in Econometrics*. Oxford Univ. Press, New York. [MR1430804](#)
- [18] DUFOUR, J.-M. and STEVANOVIĆ, D. (2013). Factor-augmented VARMA models with macroeconomic applications. *J. Bus. Econom. Statist.* **31** 491–506. [MR3173696](#)
- [19] ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. [MR0666121](#)
- [20] ENGLER, E. and NIELSEN, B. (2009). The empirical process of autoregressive residuals. *Econom. J.* **12** 367–381. [MR2562392](#)
- [21] EPANEČNIKOV, V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Teor. Veroyatn. Primen.* **14** 156–162. [MR0250422](#)
- [22] FRANCO, C. and SUCARRAT, G. (2017). An equation-by-equation estimator of a multivariate log-GARCH-X model of financial returns. *J. Multivariate Anal.* **153** 16–32. [MR3578836](#)
- [23] FRANCO, C. and ZAKOÏAN, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference* **68** 145–165. [MR1627612](#)
- [24] FRANCO, C. and ZAKOÏAN, J.-M. (2000). Estimating weak GARCH representations. *Econometric Theory* **16** 692–728. [MR1802838](#)
- [25] FRANCO, C. and ZAKOÏAN, J.-M. (2010). *GARCH Models*. Wiley, Chichester. [MR3186556](#)
- [26] GHOUDI, K. and RÉMILLARD, B. (2015). Diagnostic tests for innovations of ARMA models using empirical processes of residuals. In *Asymptotic Laws and Methods in Stochastics. Fields Inst. Commun.* **76** 239–282. [MR3409835](#)
- [27] GRØNNEBERG, S. and HOLCBLAT, B. (2019). Supplement to “On partial-sum processes of ARMAX residuals.” DOI:[10.1214/18-AOS1776SUPP](#).
- [28] HANNAN, E. J. and DEISTLER, M. (2012). *The Statistical Theory of Linear Systems. Classics in Applied Mathematics* **70**. SIAM, Philadelphia, PA. [MR3397291](#)
- [29] HANNAN, E. J., DUNSMUIR, W. T. M. and DEISTLER, M. (1980). Estimation of vector AR-MAX models. *J. Multivariate Anal.* **10** 275–295. [MR0588073](#)
- [30] HANNAN, E. J. and QUINN, B. G. (1979). The determination of the order of an autoregression. *J. Roy. Statist. Soc. Ser. B* **41** 190–195. [MR0547244](#)
- [31] HAYASHI, F. (2000). *Econometrics*. Princeton Univ. Press, Princeton, NJ. [MR1881537](#)
- [32] KLÜPPELBERG, C. and MIKOSCH, T. (1996). Parameter estimation for a misspecified ARMA model with infinite variance innovations. *J. Math. Sci.* **78** 60–65. [MR1381036](#)
- [33] KOREISHA, S. G. and FANG, Y. (2001). Generalized least squares with misspecified serial correlation structures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 515–531. [MR1858400](#)
- [34] KOUL, H. L. and LEVENTAL, S. (1989). Weak convergence of the residual empirical process in explosive autoregression. *Ann. Statist.* **17** 1784–1794. [MR1026313](#)
- [35] KRÄMER, W., PLOBERGER, W. and ALT, R. (1988). Testing for structural change in dynamic models. *Econometrica* **56** 1355–1369. [MR0973984](#)
- [36] LEE, S. (1997). A note on the residual empirical process in autoregressive models. *Statist. Probab. Lett.* **32** 405–411. [MR1602223](#)
- [37] LEEB, H. and PÖTSCHER, B. M. (2008). Can one estimate the unconditional distribution of post-model-selection estimators? *Econometric Theory* **24** 338–376. [MR2422862](#)

- [38] MACNEILL, I. B. (1978). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statist.* **6** 422–433. [MR0474645](#)
- [39] MIKOSCH, T., GADRICH, T., KLÜPPELBERG, C. and ADLER, R. J. (1995). Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.* **23** 305–326. [MR1331670](#)
- [40] NIELSEN, B. and SOHKANEN, J. S. (2011). Asymptotic behavior of the CUSUM of squares test under stochastic and deterministic time trends. *Econometric Theory* **27** 913–927. [MR2822369](#)
- [41] PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Stat.* **33** 1065–1076. [MR0143282](#)
- [42] PHILLIPS, P. C. B. and SOLO, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20** 971–1001. [MR1165602](#)
- [43] PLOBERGER, W. and KRÄMER, W. (1987). Mean adjustment and the CUSUM test for structural change. *Econom. Lett.* **25** 255–258. [MR0938813](#)
- [44] PLOBERGER, W. and KRÄMER, W. (1992). The CUSUM test with OLS residuals. *Econometrica* **60** 271–285. [MR1162619](#)
- [45] PLOBERGER, W. and KRÄMER, W. (1996). A trend-resistant test for structural change based on OLS residuals. *J. Econometrics* **70** 175–185. [MR1378773](#)
- [46] PÖTSCHER, B. M. (1991). Effects of model selection on inference. *Econometric Theory* **7** 163–185. [MR1128410](#)
- [47] ROBINSON, P. M. (1987). Time series residuals with application to probability density estimation. *J. Time Series Anal.* **8** 329–344. [MR0903762](#)
- [48] SENSIER, M. and VAN DIJK, D. (2004). Testing for volatility changes in US macroeconomic time series. *Rev. Econ. Stat.* **86** 833–839.
- [49] SHIN, D. W. and LEE, J. H. (2000). Consistency of the maximum likelihood estimators for nonstationary ARMA regressions with time trends. *J. Statist. Plann. Inference* **87** 55–68. [MR1772039](#)
- [50] SHORACK, G. R. and WELLNER, J. A. (2009). *Empirical Processes with Applications to Statistics. Classics in Applied Mathematics* **59**. SIAM, Philadelphia, PA. [MR3396731](#)
- [51] SUCARRAT, G., GRØNNEBERG, S. and ESCRIBANO, A. (2016). Estimation and inference in univariate and multivariate log-GARCH-X models when the conditional density is unknown. *Comput. Statist. Data Anal.* **100** 582–594. [MR3505819](#)
- [52] TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. Springer, New York. [MR2724359](#)
- [53] WRIGHT, J. H. (1993). The CUSUM test based on least squares residuals in regressions with integrated variables. *Econom. Lett.* **41** 353–358.
- [54] XIAO, Z. and PHILLIPS, P. C. B. (2002). A CUSUM test for cointegration using regression residuals. *J. Econometrics* **108** 43–61. [MR1894218](#)
- [55] YU, H. (2007). High moment partial sum processes of residuals in ARMA models and their applications. *J. Time Series Anal.* **28** 72–91. [MR2332851](#)

QUANTILE REGRESSION UNDER MEMORY CONSTRAINT

BY XI CHEN^{*,1} WEIDONG LIU^{†,2} AND YICHEN ZHANG^{*}

New York University^{} and Shanghai Jiao Tong University[†]*

This paper studies the inference problem in quantile regression (QR) for a large sample size n but under a limited memory constraint, where the memory can only store a small batch of data of size m . A natural method is the naive divide-and-conquer approach, which splits data into batches of size m , computes the local QR estimator for each batch and then aggregates the estimators via averaging. However, this method only works when $n = o(m^2)$ and is computationally expensive. This paper proposes a computationally efficient method, which only requires an initial QR estimator on a small batch of data and then successively refines the estimator via multiple rounds of aggregations. Theoretically, as long as n grows polynomially in m , we establish the asymptotic normality for the obtained estimator and show that our estimator with only a few rounds of aggregations achieves the same efficiency as the QR estimator computed on all the data. Moreover, our result allows the case that the dimensionality p goes to infinity. The proposed method can also be applied to address the QR problem under distributed computing environment (e.g., in a large-scale sensor network) or for real-time streaming data.

REFERENCES

- BANERJEE, M., DUROT, C. and SEN, B. (2018). Divide and conquer in non-standard problems and the super-efficiency phenomenon. *Ann. Statist.* To appear.
- BANG, H. and TSIATIS, A. A. (2002). Median regression with censored cost data. *Biometrics* **58** 643–649. [MR1926117](#)
- BATTEY, H., FAN, J., LIU, H., LU, J. and ZHU, Z. (2018). Distributed estimation and inference with statistical guarantees. *Ann. Statist.* **46** 1352–1382.
- BECK, A. (2014). *Introduction to Nonlinear Optimization*. MOS-SIAM Series on Optimization **19**. SIAM, Philadelphia, PA. [MR3288060](#)
- BELLONI, A. and CHERNOZHUKOV, V. (2011). ℓ_1 -penalized quantile regression in high-dimensional sparse models. *Ann. Statist.* **39** 82–130. [MR2797841](#)
- BELLONI, A., CHERNOZHUKOV, V., CHETVERIKOV, D. and FERNANDEZ-VAL, I. (2011). Conditional Quantile Processes based on Series or Many Regressors. Technical report. Preprint. Available at [arXiv:1105.6154v3](#).
- CHEN, X., LIU, W. and ZHANG, Y. (2019). Supplement to “Quantile regression under memory constraint.” DOI:[10.1214/18-AOS1777SUPP](#).
- CHEN, X. and XIE, M. (2014). A split-and-conquer approach for analysis of extraordinarily large data. *Statist. Sinica* **24** 1655–1684. [MR3308656](#)
- FALK, M. (1999). A simple approach to the generation of uniformly distributed random variables with prescribed correlations. *Comm. Statist. Simulation Comput.* **28** 785–791. [MR1713632](#)

MSC2010 subject classifications. Primary 62F12; secondary 62J02.

Key words and phrases. Quantile regression, sample quantile, divide-and-conquer, distributed inference, streaming data.

- FAN, J., XUE, L. and ZOU, H. (2016). Multitask quantile regression under the transnormal model. *J. Amer. Statist. Assoc.* **111** 1726–1735. [MR3601731](#)
- GALVAO, A. F. and KATO, K. (2016). Smoothed quantile regression for panel data. *J. Econometrics* **193** 92–112. [MR3500178](#)
- GAMA, J., SEBASTIÃO, R. and RODRIGUES, P. P. (2013). On evaluating stream learning algorithms. *Mach. Learn.* **90** 317–346. [MR3021877](#)
- GREENWALD, M. B. and KHANNA, S. (2004). Power-conserving computation of order-statistics over sensor networks. In *Proceedings of the ACM Symposium on Principles of Database Systems*.
- GU, Y., FAN, J., KONG, L., MA, S. and ZOU, H. (2018). ADMM for high-dimensional sparse penalized quantile regression. *Technometrics* **60** 319–331. [MR3847169](#)
- GUHA, S. and MCGREGOR, A. (2008/09). Stream order and order statistics: Quantile estimation in random-order streams. *SIAM J. Comput.* **38** 2044–2059. [MR2476286](#)
- HE, X. and SHAO, Q.-M. (2000). On parameters of increasing dimensions. *J. Multivariate Anal.* **73** 120–135. [MR1766124](#)
- HESTENES, M. R. and STIEFEL, E. (1952). Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand., B Math. Sci.* **49** 409–436. [MR0060307](#)
- HOROWITZ, J. L. (1998). Bootstrap methods for median regression models. *Econometrica* **66** 1327–1351. [MR1654307](#)
- HUANG, Z., WANG, L., YI, K. and LIU, Y. (2011). Sampling based algorithms for quantile computation in sensor networks. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*.
- JOHNSON, R. and ZHANG, T. (2013). Accelerating stochastic gradient descent using predictive variance reduction. In *Proceedings of the Advances in Neural Information Processing Systems*.
- JORDAN, M. I., LEE, J. D. and YANG, Y. (2018). Communication-efficient distributed statistical inference. *J. Amer. Statist. Assoc.* To appear.
- KLEINER, A., TALWALKAR, A., SARKAR, P. and JORDAN, M. I. (2014). A scalable bootstrap for massive data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 795–816. [MR3248677](#)
- KOENKER, R. (2005). *Quantile Regression. Econometric Society Monographs* **38**. Cambridge Univ. Press, Cambridge. [MR2268657](#)
- KOENKER, R., CHERNOZHUKOV, V., HE, X. and PENG, L., eds. (2018). *Handbook of Quantile Regression. Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. CRC Press, Boca Raton, FL. [MR3728340](#)
- KONG, E., LINTON, O. and XIA, Y. (2013). Global Bahadur representation for nonparametric censored regression quantiles and its applications. *Econometric Theory* **29** 941–968. [MR3148821](#)
- LEE, J. D., LIU, Q., SUN, Y. and TAYLOR, J. E. (2017). Communication-efficient sparse regression. *J. Mach. Learn. Res.* **18** 1–30. [MR3625709](#)
- LENG, C. and TONG, X. (2014). Censored quantile regression via Box–Cox transformation under conditional independence. *Statist. Sinica* **24** 221–249. [MR3183682](#)
- LI, R., LIN, D. K. J. and LI, B. (2013). Statistical inference in massive data sets. *Appl. Stoch. Models Bus. Ind.* **29** 399–409. [MR3117826](#)
- LUO, X., HUANG, C.-Y. and WANG, L. (2013). Quantile regression for recurrent gap time data. *Biometrics* **69** 375–385. [MR3071056](#)
- MANKU, G. S., RAJAGOPALAN, S. and LINDSAY, B. G. (1998). Approximate medians and other quantiles in one pass and with limited memory. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*.
- MUNRO, J. I. and PATERSON, M. S. (1980). Selection and sorting with limited storage. *Theoret. Comput. Sci.* **12** 315–323. [MR0589312](#)
- PANG, L., LU, W. and WANG, H. J. (2012). Variance estimation in censored quantile regression via induced smoothing. *Comput. Statist. Data Anal.* **56** 785–796. [MR2888725](#)
- PORTNOY, S. and KOENKER, R. (1997). The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Statist. Sci.* **12** 279–300. [MR1619189](#)

- RAJAGOPAL, J., WAINWRIGHT, M. and VARAIYA, P. (2006). Universal quantile estimation with feedback in the communication-constrained setting. In *Proceedings of the IEEE International Symposium on Information Theory*.
- SHAMIR, O., SREBRO, N. and ZHANG, T. (2014). Communication efficient distributed optimization using an approximate Newton-type method. In *Proceedings of the International Conference on Machine Learning*.
- SHERWOOD, B., WANG, L. and ZHOU, X.-H. (2013). Weighted quantile regression for analyzing health care cost data with missing covariates. *Stat. Med.* **32** 4967–4979. [MR3127188](#)
- SHI, C., LU, W. and SONG, R. (2017). A massive data framework for M-estimators with cubic-rate. *J. Amer. Statist. Assoc.* To appear.
- SHRIVASTAVA, N., BURAGOHAIN, C., AGRAWAL, D. and SURI, S. (2004). Medians and beyond: New aggregation techniques for sensor networks. In *Proceedings of the International Conference on Embedded Networked Sensor Systems*.
- SIEGMUND, D. (1969). On moments of the maximum of normed partial sums. *Ann. Math. Stat.* **40** 527–531. [MR0239695](#)
- VOLGUSHEV, S., CHAO, S.-K. and CHENG, G. (2018). Distributed inference for quantile regression processes. *Ann. Statist.* To appear.
- VOLGUSHEV, S., WAGENER, J. and DETTE, H. (2014). Censored quantile regression processes under dependence and penalization. *Electron. J. Stat.* **8** 2405–2447. [MR3278338](#)
- WANG, X. and DUNSON, D. B. (2014). Parallelizing MCMC via Weierstrass sampler. Technical report. Preprint. Available at [arXiv:1312.4605](#).
- WANG, H. and LI, C. (2018). Distributed quantile regression over sensor networks. *IEEE Trans. Signal Inform. Process. Netw.* **4** 338–348. [MR3808198](#)
- WANG, H. J., STEFANSKI, L. A. and ZHU, Z. (2012). Corrected-loss estimation for quantile regression with covariate measurement errors. *Biometrika* **99** 405–421. [MR2931262](#)
- WANG, H. J. and WANG, L. (2009). Locally weighted censored quantile regression. *J. Amer. Statist. Assoc.* **104** 1117–1128. [MR2562007](#)
- WANG, H. J. and WANG, L. (2014). Quantile regression analysis of length-biased survival data. *Stat* **3** 31–47.
- WANG, L., WU, Y. and LI, R. (2012). Quantile regression for analyzing heterogeneity in ultra-high dimension. *J. Amer. Statist. Assoc.* **107** 214–222. [MR2949353](#)
- WANG, L., LUO, G., YI, K. and CORMODE, G. (2013). Quantiles over data streams: An experimental study. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*.
- WANG, X., GUO, F., HELLER, K. and DUNSON, D. (2015). Parallelizing MCMC with random partition trees. In *Proceedings of the Advances in Neural Information Processing Systems*.
- WANG, J., KOLAR, M., SREBRO, N. and ZHANG, T. (2017). Efficient distributed learning with sparsity. In: *Proceedings of the International Conference on Machine Learning*.
- WHANG, Y.-J. (2006). Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory* **22** 173–205. [MR2230386](#)
- WU, Y., MA, Y. and YIN, G. (2015). Smoothed and corrected score approach to censored quantile regression with measurement errors. *J. Amer. Statist. Assoc.* **110** 1670–1683. [MR3449063](#)
- XU, G., SIT, T., WANG, L. and HUANG, C.-Y. (2017). Estimation and inference of quantile regression for survival data under biased sampling. *J. Amer. Statist. Assoc.* **112** 1571–1586. [MR3750882](#)
- YU, L., LIN, N. and WANG, L. (2017). A parallel algorithm for large-scale nonconvex penalized quantile regression. *J. Comput. Graph. Statist.* **26** 935–939. [MR3765357](#)
- ZHANG, Y., DUCHI, J. and WAINWRIGHT, M. (2015). Divide and conquer kernel ridge regression: A distributed algorithm with minimax optimal rates. *J. Mach. Learn. Res.* **16** 3299–3340. [MR3450540](#)

- ZHANG, Q. and WANG, W. (2007). A fast algorithm for approximate quantiles in high speed data streams. In *Proceedings of the International Conference on Scientific and Statistical Database Management*.
- ZHAO, T., CHENG, G. and LIU, H. (2016). A partially linear framework for massive heterogeneous data. *Ann. Statist.* **44** 1400–1437. [MR3519928](#)
- ZHENG, S. (2011). Gradient descent algorithms for quantile regression with smooth approximation. *International Journal of Machine Learning and Cybernetics* **2** 191.
- ZHENG, Q., PENG, L. and HE, X. (2018). High dimensional censored quantile regression. *Ann. Statist.* **46** 308–343. [MR3766954](#)

SAMPLING AND ESTIMATION FOR (SPARSE) EXCHANGEABLE GRAPHS¹

BY VICTOR VEITCH AND DANIEL M. ROY

Columbia University and University of Toronto

Sparse exchangeable graphs on $\mathbb{R}_+$, and the associated graphex framework for sparse graphs, generalize exchangeable graphs on $\mathbb{N}$, and the associated graphon framework for dense graphs. We develop the graphex framework as a tool for statistical network analysis by identifying the sampling scheme that is naturally associated with the models of the framework, formalizing two natural notions of consistent estimation of the parameter (the graphex) underlying these models, and identifying general consistent estimators in each case. The sampling scheme is a modification of independent vertex sampling that throws away vertices that are isolated in the sampled subgraph. The estimators are variants of the empirical graphon estimator, which is known to be a consistent estimator for the distribution of dense exchangeable graphs; both can be understood as graph analogues to the empirical distribution in the i.i.d. sequence setting. Our results may be viewed as a generalization of consistent estimation via the empirical graphon from the dense graph regime to also include sparse graphs.

REFERENCES

- [1] AIROLDI, E. M., BLEI, D. M., FEINBERG, S. E. and XING, E. P. (2008). Mixed-membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.
- [2] AIROLDI, E. M., COSTA, T. B. and CHAN, S. H. (2013). Stochastic blockmodel approximation of a graphon: Theory and consistent estimation. In *Advances in Neural Information Processing Systems* 26 692–700.
- [3] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. [MR2337396](#)
- [4] BOLLOBÁS, B. and RIORDAN, O. (2007). Metrics for sparse graphs.
- [5] BORGS, C., CHAYES, J., COHN, H. and VEITCH, V. (2017). Sampling perspectives on sparse exchangeable graphs. *Ann. Probab.* To appear.
- [6] BORGS, C., CHAYES, J. and SMITH, A. (2015). Private graphon estimation for sparse graphs. In *Advances in Neural Information Processing Systems* **28** (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett, eds.) 1369–1377. Curran Associates, Red Hook, NY.
- [7] BORGS, C., CHAYES, J. T., COHN, H. and GANGULY, S. (2015). Consistent nonparametric estimation for heavy-tailed sparse graphs. Preprint. Available at [arXiv:1508.06675](#).
- [8] BORGS, C., CHAYES, J. T., COHN, H. and HOLDEN, N. (2017). Sparse exchangeable graphs and their limits via graphon processes. *J. Mach. Learn. Res.* **18** 210. [MR3827098](#)
- [9] BORGS, C., CHAYES, J. T., COHN, H. and ZHAO, Y. (2018). An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence. *Ann. Probab.* **46** 337–396. [MR3758733](#)

MSC2010 subject classifications. Primary 62G05; secondary 62D05, 62G09, 62M99.

Key words and phrases. Network analysis, sampling, nonparametric estimation.

- [10] BORGES, C., CHAYES, J. T., COHN, H. and ZHAO, Y. (2019). An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Trans. Amer. Math. Soc.* **372** 3019–3062. [MR3988601](#)
- [11] CARON, F. and FOX, E. B. (2017). Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1295–1366. [MR3731666](#)
- [12] CHAN, S. and AIROLDI, E. (2014). A consistent histogram estimator for exchangeable graph models. In *Proceedings of the 31st International Conference on Machine Learning* (E. P. Xing and T. Jebara, eds.). *Proceedings of Machine Learning Research* **32** 208–216. PMLR, Beijing.
- [13] CRANE, H. and DEMPSEY, W. (2015). A framework for statistical network modeling. Preprint. Available at [arXiv:1509.08185](#).
- [14] DALEY, J. and VERE-JONES, (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, 2nd ed. Springer, New York.
- [15] DEN HOLLANDER, F. (2012). *Probability Theory: The Coupling Method*.
- [16] GAO, C., LU, Y. and ZHOU, H. H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43** 2624–2652. [MR3405606](#)
- [17] HERLAU, T., SCHMIDT, M. N. and MØRUP, M. (2016). Completely random measures for modelling block-structured sparse networks. In *Advances in Neural Information Processing Systems* **29** 4260–4268.
- [18] HOFF, P. D., RAFTERY, A. E. and HANDCOCK, M. S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. [MR1951262](#)
- [19] JANSON, S. (2016). Graphons and cut metric on sigma-finite measure spaces. Preprint. Available at [arXiv:1608.01833](#).
- [20] JANSON, S. (2017). On convergence for graphexes. Preprint. Available at [arXiv:1702.06389](#).
- [21] KALLENBERG, O. (1990). Exchangeable random measures in the plane. *J. Theoret. Probab.* **3** 81–136. [MR1031426](#)
- [22] KALLENBERG, O. (1997). *Foundations of Modern Probability. Probability and Its Applications (New York)*. Springer, New York. [MR1464694](#)
- [23] KALLENBERG, O. (1999). Multivariate sampling and the estimation problem for exchangeable arrays. *J. Theoret. Probab.* **12** 859–883. [MR1702867](#)
- [24] KALLENBERG, O. (2005). *Probabilistic Symmetries and Invariance Principles. Probability and Its Applications (New York)*. Springer, New York. [MR2161313](#)
- [25] LATOUCHE, P. and ROBIN, S. (2013). Variational Bayes model averaging for graphon functions and motif frequencies inference in W-graph models. Available at [arxiv.org/abs/1310.6150](#).
- [26] LLOYD, J. R., ORBANZ, P., GHAHRAMANI, Z. and ROY, D. M. (2012). Random function priors for exchangeable arrays. In *Advances in Neural Information Processing Systems* **25** 1007–1015.
- [27] MILLER, K., JORDAN, M. I. and GRIFFITHS, T. L. (2009). Nonparametric latent feature models for link prediction. In *Advances in Neural Information Processing Systems* **22** 1276–1284.
- [28] NOWICKI, K. and SNIJDERS, T. A. B. (2001). Estimation and prediction for stochastic block-structures. *J. Amer. Statist. Assoc.* **96** 1077–1087. [MR1947255](#)
- [29] ORBANZ, P. (2017). Subsampling large graphs and invariance in networks. Preprint. Available at [arXiv:1710.04217](#).
- [30] ORBANZ, P. and ROY, D. M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 437–461. DOI:10.1109/TPAMI.2014.2334607.
- [31] TODESCHINI, A., MISCOURIDOU, X. and CARON, F. (2016). Exchangeable random measures for sparse and modular graphs with overlapping communities. Preprint. Available at [arXiv:1602.02114](#).

- [32] VEITCH, V. and ROY, D. M. (2015). The class of random graphs arising from exchangeable random measures. Preprint. Available at [arXiv:1512.03099](https://arxiv.org/abs/1512.03099).
- [33] VEITCH, V. and ROY, D. M (2019). Supplement to “Sampling and estimation for (sparse) exchangeable graphs.” DOI:[10.1214/18-AOS1778SUPP](https://doi.org/10.1214/18-AOS1778SUPP).
- [34] WOLFE, P. J. and OLHEDE, S. C. (2013). Nonparametric graphon estimation. Preprint. Available at [arXiv:1309.5936](https://arxiv.org/abs/1309.5936).
- [35] YANG, J., OHAN, C. and AIROLDI, E. (2014). Nonparametric estimation and testing of exchangeable graph models. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics* (S. Kaski and J. Corander, eds.). *Proceedings of Machine Learning Research* **33** 1060–1067. PMLR, Reykjavik.

HYPOTHESIS TESTING ON LINEAR STRUCTURES OF HIGH-DIMENSIONAL COVARIANCE MATRIX

BY SHURONG ZHENG^{*,1}, ZHAO CHEN^{†,2}, HENGJIAN CUI^{‡,3} AND RUNZE LI^{§,4}

Northeast Normal University^{}, Fudan University[†], Pennsylvania State
University[‡] and Capital Normal University[§]*

This paper is concerned with test of significance on high-dimensional covariance structures, and aims to develop a unified framework for testing commonly used linear covariance structures. We first construct a consistent estimator for parameters involved in the linear covariance structure, and then develop two tests for the linear covariance structures based on entropy loss and quadratic loss used for covariance matrix estimation. To study the asymptotic properties of the proposed tests, we study related high-dimensional random matrix theory, and establish several highly useful asymptotic results. With the aid of these asymptotic results, we derive the limiting distributions of these two tests under the null and alternative hypotheses. We further show that the quadratic loss based test is asymptotically unbiased. We conduct Monte Carlo simulation study to examine the finite sample performance of the two tests. Our simulation results show that the limiting null distributions approximate their null distributions quite well, and the corresponding asymptotic critical values keep Type I error rate very well. Our numerical comparison implies that the proposed tests outperform existing ones in terms of controlling Type I error rate and power. Our simulation indicates that the test based on quadratic loss seems to have better power than the test based on entropy loss.

REFERENCES

- ANDERSON, T. W.T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. *Ann. Statist.* **1** 135–141. [MR0331612](#)
- ANDERSON, T. W.T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley Series in Probability and Statistics. Wiley Interscience, Hoboken, NJ. [MR1990662](#)
- BAI, ZHIDONGZ. and SARANADASA, H. (1996). Effect of high dimension: By an example of a two sample problem. *Statist. Sinica* **6** 311–329. [MR1399305](#)
- BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. [MR2040792](#)
- BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. Springer, New York. [MR2567175](#)
- BAI, Z., JIANG, D., YAO, J.-F. and ZHENG, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* **37** 3822–3840. [MR2572444](#)
- BIRKE, M. and DETTE, H. (2005). A note on testing the covariance matrix for large dimension. *Statist. Probab. Lett.* **74** 281–289. [MR2189467](#)

MSC2010 subject classifications. Primary 62H15; secondary 62H10.

Key words and phrases. Covariance matrix structure, random matrix theory, test of compound symmetric structure, test of banded structure, test of sphericity.

- 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. *Ann. Statist.* **39** 1496–1525. [MR2850210](#)
- CAI, T. T. and MA, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* **19** 2359–2388. [MR3160557](#)
- CHEN, S. X. and QIN, Y.-L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Ann. Statist.* **38** 808–835. [MR2604697](#)
- CHEN, S. X., ZHANG, L.-X. and ZHONG, P.-S. (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105** 810–819. [MR2724863](#)
- FAN, J. and LI, R. (2006). Statistical challenges with high dimensionality: Feature selection in knowledge discovery. In *International Congress of Mathematicians, Vol. III* 595–622. Eur. Math. Soc., Zürich. [MR2275698](#)
- HAFF, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Ann. Statist.* **8** 586–597. [MR0568722](#)
- JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 361–379. Univ. California Press, Berkeley, CA. [MR0133191](#)
- JIANG, D., JIANG, T. and YANG, F. (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *J. Statist. Plann. Inference* **142** 2241–2256. [MR2911842](#)
- JIANG, T. and QI, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scand. J. Stat.* **42** 988–1009. [MR3426306](#)
- JIANG, T. and YANG, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Statist.* **41** 2029–2074. [MR3127857](#)
- JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#)
- KATO, N., YAMADA, T. and FUJIKOSHI, Y. (2010). High-dimensional asymptotic expansion of LR statistic for testing intraclass correlation structure and its error bound. *J. Multivariate Anal.* **101** 101–112. [MR2557621](#)
- LEDOIT, O. and WOLF, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. [MR1926169](#)
- MCDONALD, R. P. (1974). The measurement of factor indeterminacy. *Psychometrika* **39** 203–222. [MR0350973](#)
- MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York. [MR0652932](#)
- OLKIN, I. and SELLIH, J. B. (1977). Estimating covariances in a multivariate normal distribution. In *Statistical Decision Theory and Related Topics, II* (S. S. Gupta and D. S. Moore, eds.) 313–326. Academic Press, New York. [MR0436407](#)
- QIU, Y. and CHEN, S. X. (2012). Test for bandedness of high-dimensional covariance matrices and bandwidth estimation. *Ann. Statist.* **40** 1285–1314. [MR3015026](#)
- SRIVASTAVA, M. S. (2005). Some tests concerning the covariance matrix in high dimensional data. *J. Japan Statist. Soc.* **35** 251–272. [MR2328427](#)
- SRIVASTAVA, M. S. and REID, N. (2012). Testing the structure of the covariance matrix with fewer observations than the dimension. *J. Multivariate Anal.* **112** 156–171. [MR2957293](#)
- WANG, C. (2014). Asymptotic power of likelihood ratio tests for high dimensional data. *Statist. Probab. Lett.* **88** 184–189. [MR3178349](#)
- WANG, Q. and YAO, J. (2013). On the sphericity test with large-dimensional observations. *Electron. J. Stat.* **7** 2164–2192. [MR3104916](#)
- WANG, C., YANG, J., MIAO, B. and CAO, L. (2013). Identity tests for high dimensional data using RMT. *J. Multivariate Anal.* **118** 128–137. [MR3054095](#)
- ZHENG, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F -matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 444–476. [MR2954263](#)

- ZHENG, S., CHEN, Z., CUI, H. and LI, R. (2019). Supplement to “Hypothesis testing on linear structures of high-dimensional covariance matrix.” DOI:[10.1214/18-AOS1779SUPP](https://doi.org/10.1214/18-AOS1779SUPP).
- ZHONG, P.-S., LAN, W., SONG, P. X. K. and TSAI, C.-L. (2017). Tests for covariance structures with high-dimensional repeated measurements. *Ann. Statist.* **45** 1185–1213. [MR3662452](#)
- ZWIERNIK, P., UHLER, C. and RICHARDS, D. (2017). Maximum likelihood estimation for linear Gaussian covariance models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1269–1292. [MR3689318](#)

ON OPTIMAL DESIGNS FOR NONREGULAR MODELS

BY YI LIN^{*}, RYAN MARTIN[†] AND MIN YANG^{*}

University of Illinois at Chicago^{} and North Carolina State University[†]*

Classically, Fisher information is the relevant object in defining optimal experimental designs. However, for models that lack certain regularity, the Fisher information does not exist, and hence, there is no notion of design optimality available in the literature. This article seeks to fill the gap by proposing a so-called *Hellinger information*, which generalizes Fisher information in the sense that the two measures agree in regular problems, but the former also exists for certain types of nonregular problems. We derive a Hellinger information inequality, showing that Hellinger information defines a lower bound on the local minimax risk of estimators. This provides a connection between features of the underlying model—in particular, the design—and the performance of estimators, motivating the use of this new Hellinger information for nonregular optimal design problems. Hellinger optimal designs are derived for several nonregular regression problems, with numerical results empirically demonstrating the efficiency of these designs compared to alternatives.

REFERENCES

- BERNARDO, J.-M. (1979). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. Ser. B* **41** 113–147. [MR0547240](#)
- BIEDERMANN, S., DETTE, H. and HOFFMANN, P. (2009). Constrained optimal discrimination designs for Fourier regression models. *Ann. Inst. Statist. Math.* **61** 143–157. [MR2481031](#)
- BIEDERMANN, S., DETTE, H. and WOODS, D. C. (2011). Optimal design for additive partially nonlinear models. *Biometrika* **98** 449–458. [MR2806440](#)
- BIEDERMANN, S., DETTE, H. and ZHU, W. (2006). Optimal designs for dose-response models with restricted design spaces. *J. Amer. Statist. Assoc.* **101** 747–759. [MR2256185](#)
- CHERNOZHUKOV, V. and HONG, H. (2004). Likelihood estimation and inference in a class of nonregular econometric models. *Econometrica* **72** 1445–1480. [MR2077489](#)
- COUSINEAU, D. (2009). Fitting the three-parameter Weibull distribution: Review and evaluation of existing and new methods. *IEEE Trans. Dielectr. Electr. Insul.* **16** 281–288.
- DETTE, H., KONSTANTINOU, M. and ZHIGLJAVSKY, A. (2017). A new approach to optimal designs for correlated observations. *Ann. Statist.* **45** 1579–1608. [MR3670189](#)
- DETTE, H. and MELAS, V. B. (2011). A note on the de la Garza phenomenon for locally optimal designs. *Ann. Statist.* **39** 1266–1281. [MR2816354](#)
- DETTE, H., PEPELYSHEV, A. and ZHIGLJAVSKY, A. (2016). Optimal designs in regression with correlated errors. *Ann. Statist.* **44** 113–152. [MR3449764](#)
- DETTE, H. and SCHORNING, K. (2013). Complete classes of designs for nonlinear regression models and principal representations of moment spaces. *Ann. Statist.* **41** 1260–1267. [MR3113810](#)

MSC2010 subject classifications. Primary 62K05; secondary 62B10.

Key words and phrases. E-optimality, experimental design, Fisher information, Hellinger distance, information inequality.

- DETTE, H. and TITOFF, S. (2009). Optimal discrimination designs. *Ann. Statist.* **37** 2056–2082. [MR2533479](#)
- DETTE, H., BRETZ, F., PEPELYSHEV, A. and PINHEIRO, J. (2008). Optimal designs for dose-finding studies. *J. Amer. Statist. Assoc.* **103** 1225–1237. [MR2462895](#)
- DETTE, H., GUCHENKO, R., MELAS, V. B. and WONG, W. K. (2018). Optimal discrimination designs for semiparametric models. *Biometrika* **105** 185–197. [MR3768873](#)
- DE LA GARZA, A. (1954). Spacing of information in polynomial regression. *Ann. Math. Stat.* **25** 123–130. [MR0060777](#)
- DROR, H. A. and STEINBERG, D. M. (2006). Robust experimental design for multivariate generalized linear models. *Technometrics* **48** 520–529. [MR2328621](#)
- FELLER, C., SCHORNING, K., DETTE, H., BERMANN, G. and BORNKAMP, B. (2017). Optimal designs for dose response curves with common parameters. *Ann. Statist.* **45** 2102–2132. [MR3718163](#)
- GOTWALT, C. M., JONES, B. A. and STEINBERG, D. M. (2009). Fast computation of designs robust to parameter uncertainty for nonlinear settings. *Technometrics* **51** 88–95. [MR2664866](#)
- HARMAN, R. and BENKOVÁ, E. (2017). Barycentric algorithm for computing D -optimal size- and cost-constrained designs of experiments. *Metrika* **80** 201–225. [MR3597582](#)
- HIROSE, H. and LAI, T. L. (1997). Inference from grouped data in three-parameter Weibull models with applications to breakdown-voltage experiments. *Technometrics* **39** 199–210. [MR1452348](#)
- IBRAGIMOV, I. A. and HAS'MINSKIĬ, R. Z. (1981). *Statistical Estimation. Asymptotic Theory. Applications of Mathematics* **16**. Springer, New York–Berlin. Translated from the Russian by Samuel Kotz. [MR0620321](#)
- KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879. [MR0356386](#)
- KOENKER, R. (2013). quantreg: Quantile regression. R package version 5.05. Available at <http://CRAN.R-project.org/package=quantreg>.
- KOENKER, R. and HALLOCK, K. (2001). Quantile regression: An introduction. *J. Econ. Perspect.* **15** 43–56.
- LEHMANN, E. L. and CASELLA, G. (1998). *Theory of Point Estimation*, 2nd ed. *Springer Texts in Statistics*. Springer, New York. [MR1639875](#)
- LIN, Y., MARTIN, R. and YANG, M. (2019). Supplement to “On optimal designs for nonregular models.” DOI:[10.1214/18-AOS1780SUPP](https://doi.org/10.1214/18-AOS1780SUPP).
- LINDLEY, D. V. (1956). On a measure of the information provided by an experiment. *Ann. Math. Stat.* **27** 986–1005. [MR0083936](#)
- LÓPEZ-FIDALGO, J., TOMMASI, C. and TRANDAFIR, P. C. (2007). An optimal experimental design criterion for discriminating between non-normal models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 231–242. [MR2325274](#)
- POLLARD, D. (1997). Another look at differentiability in quadratic mean. In *Festschrift for Lucien Le Cam* 305–314. Springer, New York. [MR1462953](#)
- POLLARD, D. (2005). Asymptopia. In *Hellinger Differentiability*, Chapter 6, <http://www.stat.yale.edu/~pollard/Courses/607.spring05/handouts/DQM.pdf>.
- SAGNOL, G. and HARMAN, R. (2015). Computing exact D -optimal designs by mixed integer second-order cone programming. *Ann. Statist.* **43** 2198–2224. [MR3396983](#)
- SCHORNING, K., DETTE, H., KETTELHAKE, K., WONG, W. K. and BRETZ, F. (2017). Optimal designs for active controlled dose-finding trials with efficacy-toxicity outcomes. *Biometrika* **104** 1003–1010. [MR3737318](#)
- SHEMYAKIN, A. E. (1991). Multidimensional integral inequalities of Rao–Cramér type for parametric families with singularities. *Sibirsk. Mat. Zh.* **32** 204–215, 230. [MR1142080](#)
- SHEMYAKIN, A. E. (1992). On information inequalities in parametric estimation theory. *Teor. Veroyatn. Primen.* **37** 121–123. [MR1211210](#)

- SHEMYAKIN, A. (2014). Hellinger distance and non-informative priors. *Bayesian Anal.* **9** 923–938. [MR3293962](#)
- SMITH, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72** 67–90. [MR0790201](#)
- SMITH, R. L. (1994). Nonregular regression. *Biometrika* **81** 173–183. [MR1279665](#)
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. [MR1652247](#)
- VAN DER VAART, A. (2002). The statistical work of Lucien Le Cam. *Ann. Statist.* **30** 631–682. Dedicated to the memory of Lucien Le Cam. [MR1922537](#)
- WATERHOUSE, T. H., WOODS, D. C., ECCLESTON, J. A. and LEWIS, S. M. (2008). Design selection criteria for discrimination/estimation for nested models and a binomial response. *J. Statist. Plann. Inference* **138** 132–144. [MR2369620](#)
- WOODROOFE, M. (1974). Maximum likelihood estimation of translation parameter of truncated distribution. II. *Ann. Statist.* **2** 474–488. Collection of articles dedicated to Jerzy Neyman on his 80th birthday. [MR0356343](#)
- YANG, M. (2010). On the de la Garza phenomenon. *Ann. Statist.* **38** 2499–2524. [MR2676896](#)
- YANG, M., BIEDERMANN, S. and TANG, E. (2013). On optimal designs for nonlinear models: A general and efficient algorithm. *J. Amer. Statist. Assoc.* **108** 1411–1420. [MR3174717](#)
- YANG, M. and STUFKEN, J. (2009). Support points of locally optimal designs for nonlinear models with two parameters. *Ann. Statist.* **37** 518–541. [MR2488361](#)
- YANG, M. and STUFKEN, J. (2012). Identifying locally optimal designs for nonlinear models: A simple extension with profound consequences. *Ann. Statist.* **40** 1665–1681. [MR3015039](#)
- YU, Y. (2011). D-optimal designs via a cocktail algorithm. *Stat. Comput.* **21** 475–481. [MR2826686](#)

A SMEARY CENTRAL LIMIT THEOREM FOR MANIFOLDS WITH APPLICATION TO HIGH-DIMENSIONAL SPHERES¹

BY BENJAMIN ELTZNER AND STEPHAN F. HUCKEMANN

*Felix-Bernstein-Institut für Mathematische Statistik in den Biowissenschaften and
 Georg-August-Universität Göttingen*

The (CLT) central limit theorems for generalized Fréchet means (data descriptors assuming values in manifolds, such as intrinsic means, geodesics, etc.) on manifolds from the literature are only valid if a certain empirical process of Hessians of the Fréchet function converges suitably, as in the proof of the prototypical BP-CLT [*Ann. Statist.* **33** (2005) 1225–1259]. This is not valid in many realistic scenarios and we provide for a new very general CLT. In particular, this includes scenarios where, in a suitable chart, the sample mean fluctuates asymptotically at a scale n^α with exponents $\alpha < 1/2$ with a nonnormal distribution. As the BP-CLT yields only fluctuations that are, rescaled with $n^{1/2}$, asymptotically normal, just as the classical CLT for random vectors, these lower rates, somewhat loosely called smeariness, had to date been observed only on the circle. We make the concept of smeariness on manifolds precise, give an example for two-smeariness on spheres of arbitrary dimension, and show that smeariness, although “almost never” occurring, may have serious statistical implications on a continuum of sample scenarios nearby. In fact, this effect increases with dimension, striking in particular in high dimension low sample size scenarios.

REFERENCES

- AFSARI, B. (2011). Riemannian L^p center of mass: Existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.* **139** 655–673. [MR2736346](#)
- ARNAUDON, M. and MICLO, L. (2014). Means in complete manifolds: Uniqueness and approximation. *ESAIM Probab. Stat.* **18** 185–206. [MR3230874](#)
- BARDEN, D., LE, H. and OWEN, M. (2013). Central limit theorems for Fréchet means in the space of phylogenetic trees. *Electron. J. Probab.* **18** 1–25. [MR3035753](#)
- BARDEN, D., LE, H. and OWEN, M. (2018). Limiting behaviour of Fréchet means in the space of phylogenetic trees. *Ann. Inst. Statist. Math.* **70** 99–129. [MR3742820](#)
- BHATTACHARYA, A. and BHATTACHARYA, R. (2008). Statistics on Riemannian manifolds: Asymptotic distribution and curvature. *Proc. Amer. Math. Soc.* **136** 2959–2967. [MR2399064](#)
- BHATTACHARYA, R. and LIN, L. (2017). Omnibus CLTs for Fréchet means and nonparametric inference on non-Euclidean spaces. *Proc. Amer. Math. Soc.* **145** 413–428. [MR3565392](#)
- BHATTACHARYA, R. and PATRANGENARU, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Ann. Statist.* **31** 1–29. [MR1962498](#)
- BHATTACHARYA, R. and PATRANGENARU, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds. II. *Ann. Statist.* **33** 1225–1259. [MR2195634](#)

MSC2010 subject classifications. Primary 62G20, 62H11; secondary 53C22, 58K45.

Key words and phrases. Fréchet means, asymptotic consistency and normality, asymptotics on manifolds, lower asymptotic rate, high dimension low sample size.

- BHATTACHARYA, R. and PATRANGENARU, V. (2014). Statistics on manifolds and landmarks based image analysis: A nonparametric theory with applications. *J. Statist. Plann. Inference* **145** 1–22. [MR3125343](#)
- BILLERA, L. J., HOLMES, S. P. and VOGTMANN, K. (2001). Geometry of the space of phylogenetic trees. *Adv. in Appl. Math.* **27** 733–767. [MR1867931](#)
- DRYDEN, I. L., KOLOYDENKO, A. and ZHOU, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3** 1102–1123. [MR2750388](#)
- DRYDEN, I. L. and MARDIA, K. V. (1998). *Statistical Shape Analysis. Wiley Series in Probability and Statistics: Probability and Statistics*. Wiley, Chichester. [MR1646114](#)
- ELLINGSON, L., PATRANGENARU, V. and RUYMGAART, F. (2013). Nonparametric estimation of means on Hilbert manifolds and extrinsic analysis of mean shapes of contours. *J. Multivariate Anal.* **122** 317–333. [MR3189326](#)
- FLETCHER, P. T. and JOSHI, S. C. (2004). Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. *ECCV Workshops CVAMIA and MMBIA* 87–98.
- FRÉCHET, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. Henri Poincaré* **10** 215–310. [MR0027464](#)
- GROSSER, D. (2005). On the convergence of some Procrustean averaging algorithms. *Stochastics* **77** 31–60. [MR2138772](#)
- HOTZ, T. and HUCKEMANN, S. (2015). Intrinsic means on the circle: Uniqueness, locus and asymptotics. *Ann. Inst. Statist. Math.* **67** 177–193. [MR3297863](#)
- HOTZ, T., HUCKEMANN, S., LE, H., MARRON, J. S., MATTINGLY, J. C., MILLER, E., NOLEN, J., OWEN, M., PATRANGENARU, V. et al. (2013). Sticky central limit theorems on open books. *Ann. Appl. Probab.* **23** 2238–2258. [MR3127934](#)
- HUCKEMANN, S. (2011a). Inference on 3D Procrustes means: Tree bole growth, rank deficient diffusion tensors and perturbation models. *Scand. J. Stat.* **38** 424–446. [MR2833839](#)
- HUCKEMANN, S. F. (2011b). Intrinsic inference on the mean geodesic of planar shapes and tree discrimination by leaf growth. *Ann. Statist.* **39** 1098–1124. [MR2816349](#)
- HUCKEMANN, S. F. (2012). On the meaning of mean shape: Manifold stability, locus and the two sample test. *Ann. Inst. Statist. Math.* **64** 1227–1259. [MR2981621](#)
- HUCKEMANN, S. (2015). (Semi-)intrinsic statistical analysis on non-Euclidean spaces. In *Advances in Complex Data Modeling and Computational Methods in Statistics*. Springer, Berlin.
- HUCKEMANN, S. F. and ELTZNER, B. (2018). Backward nested descriptors asymptotics with inference on stem cell differentiation. *Ann. Statist.* **46** 1994–2019. [MR3845008](#)
- HUCKEMANN, S., HOTZ, T. and MUNK, A. (2010). Intrinsic shape analysis: Geodesic principal component analysis for Riemannian manifolds modulo Lie group actions (with discussion). *Statist. Sinica* **20** 1–100.
- HUCKEMANN, S. and ZIEZOLD, H. (2006). Principal component analysis for Riemannian manifolds, with an application to triangular shape spaces. *Adv. in Appl. Probab.* **38** 299–319. [MR2264946](#)
- HUCKEMANN, S., MATTINGLY, J. C., MILLER, E. and NOLEN, J. (2015). Sticky central limit theorems at isolated hyperbolic planar singularities. *Electron. J. Probab.* **20** no. 78, 34. [MR3371437](#)
- JUNG, S., DRYDEN, I. L. and MARRON, J. S. (2012). Analysis of principal nested spheres. *Biometrika* **99** 551–568. [MR2966769](#)
- JUNG, S., FOSKEY, M. and MARRON, J. S. (2011). Principal arc analysis on direct product manifolds. *Ann. Appl. Stat.* **5** 578–603. [MR2810410](#)
- KARCHER, H. (1977). Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.* **30** 509–541. [MR0442975](#)
- KENDALL, D. G. (1984). Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. Lond. Math. Soc.* **16** 81–121. [MR0737237](#)

- KENDALL, W. S. (1990). Probability, convexity, and harmonic maps with small image. I. Uniqueness and fine existence. *Proc. Lond. Math. Soc.* (3) **61** 371–406. [MR1063050](#)
- KENDALL, D. G., BARDEN, D., CARNE, T. K. and LE, H. (1999). *Shape and Shape Theory*. Wiley Series in Probability and Statistics. Wiley, Chichester. [MR1891212](#)
- LE, H. (2001). Locating Fréchet means with application to shape spaces. *Adv. in Appl. Probab.* **33** 324–338. [MR1842295](#)
- LE, H. and BARDEN, D. (2014). On the measure of the cut locus of a Fréchet mean. *Bull. Lond. Math. Soc.* **46** 698–708. [MR3239607](#)
- MARDIA, K. V. and JUPP, P. E. (2000). *Directional Statistics*. Wiley Series in Probability and Statistics. Wiley, Chichester. [MR1828667](#)
- MARDIA, K. V. and PATRANGENARU, V. (2005). Directions and projective shapes. *Ann. Statist.* **33** 1666–1699. [MR2166559](#)
- MARRON, J. S. and ALONSO, A. M. (2014). Overview of object oriented data analysis. *Biom. J.* **56** 732–753. [MR3258083](#)
- McKILLIAM, R. G., QUINN, B. G. and CLARKSON, I. V. L. (2012). Direction estimation by minimum squared arc length. *IEEE Trans. Signal Process.* **60** 2115–2124. [MR2954196](#)
- MOULTON, V. and STEEL, M. (2004). Peeling phylogenetic ‘oranges’. *Adv. in Appl. Math.* **33** 710–727. [MR2095862](#)
- MUNK, A., PAIGE, R., PANG, J., PATRANGENARU, V. and RUYMGAART, F. (2008). The one- and multi-sample problem for functional data with application to projective shape analysis. *J. Multivariate Anal.* **99** 815–833. [MR2405093](#)
- PATRANGENARU, V. and ELLINGSON, L. (2016). *Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis*. CRC Press, Boca Raton, FL. [MR3444169](#)
- PENNEC, X. (2018). Barycentric subspace analysis on manifolds. *Ann. Statist.* **46** 2711–2746. [MR3851753](#)
- SOMMER, S. (2016). Anisotropically weighted and nonholonomically constrained evolutions on manifolds. *Entropy* **18** Paper No. 425, 21. [MR3594070](#)
- TURNER, K., MILEYKO, Y., MUKHERJEE, S. and HARER, J. (2014). Fréchet means for distributions of persistence diagrams. *Discrete Comput. Geom.* **52** 44–70. [MR3231030](#)
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics **3**. Cambridge Univ. Press, Cambridge. [MR1652247](#)
- ZIEZOLD, H. (1977). On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the 1974 European Meeting of Statisticians* 591–602. [MR0501230](#)

ON TESTING FOR HIGH-DIMENSIONAL WHITE NOISE

BY ZENG LI^{†,1}, CLIFFORD LAM^{*}, JIANFENG YAO^{‡,2} AND QIWEI YAO^{*}

London School of Economics^{}, Pennsylvania State University[†] and The University of Hong Kong[‡]*

Testing for white noise is a classical yet important problem in statistics, especially for diagnostic checks in time series modeling and linear regression. For high-dimensional time series in the sense that the dimension p is large in relation to the sample size T , the popular omnibus tests including the multivariate Hosking and Li–McLeod tests are extremely conservative, leading to substantial power loss. To develop more relevant tests for high-dimensional cases, we propose a portmanteau-type test statistic which is the sum of squared singular values of the first q lagged sample autocovariance matrices. It, therefore, encapsulates all the serial correlations (up to the time lag q) within and across all component series. Using the tools from random matrix theory and assuming both p and T diverge to infinity, we derive the asymptotic normality of the test statistic under both the null and a specific VMA(1) alternative hypothesis. As the actual implementation of the test requires the knowledge of three characteristic constants of the population cross-sectional covariance matrix and the value of the fourth moment of the standardized innovations, nontrivial estimations are proposed for these parameters and their integration leads to a practically usable test. Extensive simulation confirms the excellent finite-sample performance of the new test with accurate size and satisfactory power for a large range of finite (p, T) combinations, therefore, ensuring wide applicability in practice. In particular, the new tests are consistently superior to the traditional Hosking and Li–McLeod tests.

REFERENCES

- BAI, Z., CHEN, J. and YAO, J. (2010). On estimation of the population spectral distribution from a high-dimensional sample covariance matrix. *Aust. N. Z. J. Stat.* **52** 423–437. [MR2791528](#)
- BAI, J. and NG, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* **70** 191–221. [MR1926259](#)
- BAI, Z., JIANG, D., YAO, J.-F. and ZHENG, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* **37** 3822–3840. [MR2572444](#)
- BASU, S. and MICHAELIDIS, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567. [MR3357870](#)
- BHATTACHARJEE, M. and BOSE, A. (2016). Large sample behaviour of high dimensional autocovariance matrices. *Ann. Statist.* **44** 598–628. [MR3476611](#)
- BICKEL, P. J. and GEL, Y. R. (2011). Banded regularization of autocovariance matrices in application to parameter estimation and forecasting of time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 711–728. [MR2867455](#)

MSC2010 subject classifications. Primary 62M10, 62H15; secondary 15A52.

Key words and phrases. Large autocovariance matrix, Hosking test, Li–McLeod test, high-dimensional time series, random matrix theory.

- CHANG, J., GUO, B. and YAO, Q. (2015). High dimensional stochastic regression with latent factors, endogeneity and nonlinearity. *J. Econometrics* **189** 297–312. [MR3414901](#)
- CHANG, J., YAO, Q. and ZHOU, W. (2017). Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika* **104** 111–127. [MR3626482](#)
- FORNI, M., HALLIN, M., LIPPI, M. and REICHLIN, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Rev. Econ. Stat.* **82** 540–554.
- FORNI, M., HALLIN, M., LIPPI, M. and REICHLIN, L. (2005). The generalized dynamic factor model: One-sided estimation and forecasting. *J. Amer. Statist. Assoc.* **100** 830–840. [MR2201012](#)
- GUO, S., WANG, Y. and YAO, Q. (2016). High-dimensional and banded vector autoregressions. *Biometrika* **103** 889–903. [MR3620446](#)
- HAN, F., LU, H. and LIU, H. (2015). A direct estimation of high dimensional stationary vector autoregressions. *J. Mach. Learn. Res.* **16** 3115–3150. [MR3450535](#)
- HAUFE, S., NOLTE, G., MUELLER, K. R. and KRÄMER, N. (2009). Sparse causal discovery in multivariate time series. In *Proceedings of the 2008th International Conference on Causality: Objectives and Assessment* **6** 97–106. [JMLR.org](#).
- HOSKING, J. R. M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.* **75** 602–608. [MR0590689](#)
- HSU, N.-J., HUNG, H.-L. and CHANG, Y.-M. (2008). Subset selection for vector autoregressive processes using Lasso. *Comput. Statist. Data Anal.* **52** 3645–3657. [MR2427370](#)
- JOHNSTONE, I. M. (2007). High dimensional statistical inference and random matrices. In *International Congress of Mathematicians. Vol. I* 307–333. Eur. Math. Soc., Zürich. [MR2334195](#)
- LAM, C. and YAO, Q. (2012). Factor modeling for high-dimensional time series: Inference for the number of factors. *Ann. Statist.* **40** 694–726. [MR2933663](#)
- LI, W. K. (2004). *Diagnostic Checks in Time Series*. CRC Press/CRC, Boca Raton.
- LI, W. K. and MCLEOD, A. I. (1981). Distribution of the residual autocorrelations in multivariate ARMA time series models. *J. Roy. Statist. Soc. Ser. B* **43** 231–239. [MR0626770](#)
- LI, Z., LAM, C., YAO, J. and YAO, Q. (2019). Supplement to “On testing for high-dimensional white noise.” DOI:[10.1214/18-AOS1782SUPP](#).
- LIU, H., AUE, A. and PAUL, D. (2015). On the Marčenko–Pastur law for linear time series. *Ann. Statist.* **43** 675–712. [MR3319140](#)
- LÜTKEPOHL, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin. [MR2172368](#)
- PAUL, D. and AUE, A. (2014). Random matrix theory in statistics: A review. *J. Statist. Plann. Inference* **150** 1–29. [MR3206718](#)
- SHOJAIE, A. and MICHAILIDIS, G. (2010). Discovering graphical Granger causality using the truncating LASSO penalty. *Bioinformatics* **26** 517–523.
- STOCK, J. H. and WATSON, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroecon. Annu.* **4** 351–394.
- STOCK, J. H. and WATSON, M. W. (1998). Diffusion Indexes (No. W6702). National Bureau of Economic Research.
- STOCK, J. H. and WATSON, M. W. (1999). Forecasting inflation. *J. Monet. Econ.* **44** 293–335.
- TSAY, R. (2017). Testing for serial correlations in high-dimensional time series via extreme value theory. Preprint.
- WANG, Q. and YAO, J. (2013). On the sphericity test with large-dimensional observations. *Electron. J. Stat.* **7** 2164–2192. [MR3104916](#)
- YAO, J., ZHENG, S. and BAI, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis. Cambridge Series in Statistical and Probabilistic Mathematics* **39**. Cambridge Univ. Press, New York. [MR3468554](#)
- ZHENG, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F -matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 444–476. [MR2954263](#)

MINIMAX POSTERIOR CONVERGENCE RATES AND MODEL SELECTION CONSISTENCY IN HIGH-DIMENSIONAL DAG MODELS BASED ON SPARSE CHOLESKY FACTORS

BY KYOUNGJAE LEE^{*,1}, JAEYONG LEE^{†,2} AND LIZHEN LIN^{‡,1}

Inha University, ^{*}*Seoul National University*[†] and *University of Notre Dame*[‡]

In this paper we study the high-dimensional sparse directed acyclic graph (DAG) models under the empirical sparse Cholesky prior. Among our results, strong model selection consistency or graph selection consistency is obtained under more general conditions than those in the existing literature. Compared to Cao, Khare and Ghosh [*Ann. Statist.* (2019) **47** 319–348], the required conditions are weakened in terms of the dimensionality, sparsity and lower bound of the nonzero elements in the Cholesky factor. Furthermore, our result does not require the irrerepresentable condition, which is necessary for Lasso-type methods. We also derive the posterior convergence rates for precision matrices and Cholesky factors with respect to various matrix norms. The obtained posterior convergence rates are the fastest among those of the existing Bayesian approaches. In particular, we prove that our posterior convergence rates for Cholesky factors are the minimax or at least nearly minimax depending on the relative size of true sparseness for the entire dimension. The simulation study confirms that the proposed method outperforms the competing methods.

REFERENCES

- BANERJEE, S. and GHOSAL, S. (2014). Posterior convergence rates for estimating large precision matrices using graphical models. *Electron. J. Stat.* **8** 2111–2137. [MR3273620](#)
- BANERJEE, S. and GHOSAL, S. (2015). Bayesian structure learning in graphical models. *J. Multivariate Anal.* **136** 147–162. [MR3321485](#)
- BEN-DAVID, E., LI, T., MASSAM, H. and RAJARATNAM, B. (2015). High dimensional Bayesian inference for Gaussian directed acyclic graph models. Available at [arXiv:1109.4371v5](#).
- BHATTACHARYA, A., PATI, D. and YANG, Y. (2019). Bayesian fractional posteriors. *Ann. Statist.* **47** 39–66. [MR3909926](#)
- BICKEL, P. J. and LEVINA, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227. [MR2387969](#)
- BÜHLMANN, P. and VAN DE GEER, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Series in Statistics. Springer, Heidelberg. [MR2807761](#)
- CAI, T. T., LIU, W. and ZHOU, H. H. (2016). Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *Ann. Statist.* **44** 455–488. [MR3476606](#)
- CAI, T., MA, Z. and WU, Y. (2015). Optimal estimation and rank detection for sparse spiked covariance matrices. *Probab. Theory Related Fields* **161** 781–815. [MR3334281](#)

MSC2010 subject classifications. Primary 62C20; secondary 62F15, 62C12.

Key words and phrases. DAG model, precision matrix, Cholesky factor, posterior convergence rate, strong model selection consistency.

- CAI, T. T. and YUAN, M. (2012). Adaptive covariance matrix estimation through block thresholding. *Ann. Statist.* **40** 2014–2042. [MR3059075](#)
- CAI, T. T., ZHANG, C.-H. and ZHOU, H. H. (2010). Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.* **38** 2118–2144. [MR2676885](#)
- CAI, T. T. and ZHOU, H. H. (2012a). Minimax estimation of large covariance matrices under ℓ_1 -norm. *Statist. Sinica* **22** 1319–1349. [MR3027084](#)
- CAI, T. T. and ZHOU, H. H. (2012b). Optimal rates of convergence for sparse covariance matrix estimation. *Ann. Statist.* **40** 2389–2420. [MR3097607](#)
- CAO, X., KHARE, K. and GHOSH, M. (2019). Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models. *Ann. Statist.* **47** 319–348. [MR3909935](#)
- CASTILLO, I., SCHMIDT-HIEBER, J. and VAN DER VAART, A. (2015). Bayesian linear regression with sparse priors. *Ann. Statist.* **43** 1986–2018. [MR3375874](#)
- FAN, J., FAN, Y. and LV, J. (2008). High dimensional covariance matrix estimation using a factor model. *J. Econometrics* **147** 186–197. [MR2472991](#)
- GAO, C. and ZHOU, H. H. (2015). Rate-optimal posterior contraction for sparse PCA. *Ann. Statist.* **43** 785–818. [MR3325710](#)
- GRÜNWARD, P. and VAN OMMEN, T. (2017). Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Anal.* **12** 1069–1103. [MR3724979](#)
- HUANG, J. Z., LIU, N., POURAHMADI, M. and LIU, L. (2006). Covariance matrix selection and estimation via penalised normal likelihood. *Biometrika* **93** 85–98. [MR2277742](#)
- JOHNSTONE, I. M. and LU, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *J. Amer. Statist. Assoc.* **104** 682–693. [MR2751448](#)
- KALISCH, M. and BÜHLMANN, P. (2007). Estimating high-dimensional directed acyclic graphs with the PC-algorithm. *J. Mach. Learn. Res.* **8** 613–636.
- KHARE, K., OH, S., RAHMAN, S. and RAJARATNAM, B. (2016). A convex framework for high-dimensional sparse Cholesky based covariance estimation. Preprint. Available at [arxiv:1610.02436](#).
- LEE, K. and LEE, J. (2017). Estimating large precision matrices via modified cholesky decomposition. Available at [arXiv:1707.01143](#).
- LEE, K. and LEE, J. (2018). Optimal Bayesian minimax rates for unconstrained large covariance matrices. *Bayesian Anal.* **13** 1211–1229. [MR3855369](#)
- LEE, K., LEE, J. and LIN, L. (2019). Supplement to “Minimax posterior convergence rates and model selection consistency in high-dimensional DAG models based on sparse Cholesky factors.” DOI:[10.1214/18-AOS1783SUPP](#).
- LIANG, F., PAULO, R., MOLINA, G., CLYDE, M. A. and BERGER, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *J. Amer. Statist. Assoc.* **103** 410–423. [MR2420243](#)
- MARTIN, R., MESS, R. and WALKER, S. G. (2017). Empirical Bayes posterior concentration in sparse high-dimensional linear models. *Bernoulli* **23** 1822–1847. [MR3624879](#)
- MARTIN, R. and WALKER, S. G. (2014). Asymptotically minimax empirical Bayes estimation of a sparse normal mean vector. *Electron. J. Stat.* **8** 2188–2206. [MR3273623](#)
- MILLER, J. W. and DUNSON, D. B. (2018). Robust Bayesian inference via coarsening. *J. Amer. Statist. Assoc.* DOI:[10.1080/01621459.2018.1469995](#).
- NARISSETTY, N. N. and HE, X. (2014). Bayesian variable selection with shrinking and diffusing priors. *Ann. Statist.* **42** 789–817. [MR3210987](#)
- PATI, D., BHATTACHARYA, A., PILLAI, N. S. and DUNSON, D. (2014). Posterior contraction in sparse Bayesian factor models for massive covariance matrices. *Ann. Statist.* **42** 1102–1130. [MR3210997](#)
- REID, S., TIBSHIRANI, R. and FRIEDMAN, J. (2016). A study of error variance estimation in Lasso regression. *Statist. Sinica* **26** 35–67. [MR3468344](#)
- REN, Z., SUN, T., ZHANG, C.-H. and ZHOU, H. H. (2015). Asymptotic normality and optimalities in estimation of large Gaussian graphical models. *Ann. Statist.* **43** 991–1026. [MR3346695](#)

- ROTHMAN, A. J., LEVINA, E. and ZHU, J. (2010). A new approach to Cholesky-based covariance regularization in high dimensions. *Biometrika* **97** 539–550. [MR2672482](#)
- ROVERATO, A. (2000). Cholesky decomposition of a hyper inverse Wishart matrix. *Biometrika* **87** 99–112. [MR1766831](#)
- RÜTIMANN, P. and BÜHLMANN, P. (2009). High dimensional sparse covariance estimation via directed acyclic graphs. *Electron. J. Stat.* **3** 1133–1160. [MR2566184](#)
- SHANG, Z. and CLAYTON, M. K. (2011). Consistency of Bayesian linear model selection with a growing number of parameters. *J. Statist. Plann. Inference* **141** 3463–3474. [MR2817355](#)
- SHIN, M., BHATTACHARYA, A. and JOHNSON, V. E. (2018). Scalable Bayesian variable selection using nonlocal prior densities in ultrahigh-dimensional settings. *Statist. Sinica* **28** 1053–1078. [MR3791100](#)
- SHOJAIE, A. and MICHAELIDIS, G. (2010). Penalized likelihood methods for estimation of sparse high-dimensional directed acyclic graphs. *Biometrika* **97** 519–538. [MR2672481](#)
- SYRING, N. A. and MARTIN, R. (2016). *Scaling the Gibbs Posterior Credible Regions*. Preprint. Available at [arxiv:1509.00922](#).
- VAN DE GEER, S. and BÜHLMANN, P. (2013). ℓ_0 -penalized maximum likelihood for sparse directed acyclic graphs. *Ann. Statist.* **41** 536–567. [MR3099113](#)
- WAINWRIGHT, M. J. (2009a). Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting. *IEEE Trans. Inform. Theory* **55** 5728–5741. [MR2597190](#)
- WAINWRIGHT, M. J. (2009b). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programming (Lasso). *IEEE Trans. Inform. Theory* **55** 2183–2202. [MR2729873](#)
- WALKER, S. and HJORT, N. L. (2001). On Bayesian consistency. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 811–821. [MR1872068](#)
- XIANG, R., KHARE, K. and GHOSH, M. (2015). High dimensional posterior convergence rates for decomposable graphical models. *Electron. J. Stat.* **9** 2828–2854. [MR3439186](#)
- YANG, Y., WAINWRIGHT, M. J. and JORDAN, M. I. (2016). On the computational complexity of high-dimensional Bayesian variable selection. *Ann. Statist.* **44** 2497–2532. [MR3576552](#)
- YU, G. and BIEN, J. (2017). Learning local dependence in ordered data. *J. Mach. Learn. Res.* **18** 42. [MR3655307](#)
- ZELLNER, A. (1986). On assessing prior distributions and Bayesian regression analysis with g -prior distributions. In *Bayesian Inference and Decision Techniques. Stud. Bayesian Econometrics Statist.* **6** 233–243. North-Holland, Amsterdam. [MR0881437](#)

BOOTSTRAPPING AND SAMPLE SPLITTING FOR HIGH-DIMENSIONAL, ASSUMPTION-LEAN INFERENCE

BY ALESSANDRO RINALDO, LARRY WASSERMAN AND MAX G'SELL

Carnegie Mellon University

Several new methods have been recently proposed for performing valid inference after model selection. An older method is sample splitting: use part of the data for model selection and the rest for inference. In this paper, we revisit sample splitting combined with the bootstrap (or the Normal approximation). We show that this leads to a simple, assumption-lean approach to inference and we establish results on the accuracy of the method. In fact, we find new bounds on the accuracy of the bootstrap and the Normal approximation for general nonlinear parameters with increasing dimension which we then use to assess the accuracy of regression inference. We define new parameters that measure variable importance and that can be inferred with greater accuracy than the usual regression coefficients. Finally, we elucidate an inference-prediction trade-off: splitting increases the accuracy and robustness of inference but can decrease the accuracy of the predictions.

REFERENCES

- [1] ANASTASIOU, A. and GAUNT, R. E. (2016). Multivariate normal approximation of the maximum likelihood estimator via the delta method. Preprint. Available at [arXiv:1609.03970](https://arxiv.org/abs/1609.03970).
- [2] ANASTASIOU, A. and LEY, C. (2015). New simpler bounds to assess the asymptotic normality of the maximum likelihood estimator. Preprint. Available at [arXiv:1508.04948](https://arxiv.org/abs/1508.04948).
- [3] ANASTASIOU, A. and REINERT, G. (2017). Bounds for the normal approximation of the maximum likelihood estimator. *Bernoulli* **23** 191–218. [MR3556771](https://doi.org/10.1017/S0005321717000071)
- [4] ANDREWS, D. W. K. and GUGGENBERGER, P. (2009). Hybrid and size-corrected subsampling methods. *Econometrica* **77** 721–762. [MR2531360](https://doi.org/10.1111/j.1467-9892.00600.x)
- [5] BACHOC, F., LEEB, H. and PÖTSCHER, B. M. (2014). Valid confidence intervals for post-model-selection predictors. Available at [arXiv:1412.4605](https://arxiv.org/abs/1412.4605).
- [6] BACHOC, F., PREINERSTORFER, D. and STEINBERGER, L. (2016). Uniformly valid confidence intervals post-model-selection. Available at [arXiv:1611.01043](https://arxiv.org/abs/1611.01043).
- [7] BARBER, R. F. and CANDÈS, E. J. (2015). Controlling the false discovery rate via knockoffs. *Ann. Statist.* **43** 2055–2085. [MR3375876](https://doi.org/10.1214/15-AOS1274)
- [8] BARNARD, G. A. (1974). Discussion of “Cross-validatory choice and assessment of statistical predictions,” by M. Stone. *J. Roy. Statist. Soc. Ser. B* 133–135.
- [9] BELLONI, A., CHERNOZHUKOV, V. and HANSEN, C. B. (2013). *Inference for High-Dimensional Sparse Econometric Models*. vol. 3 245–295. Cambridge Univ. Press.
- [10] BELLONI, A., CHERNOZHUKOV, V. and KATO, K. (2015). Uniform post-selection inference for least absolute deviation regression and other Z-estimation problems. *Biometrika* **102** 77–94. [MR3335097](https://doi.org/10.1093/biomet/asv007)

MSC2010 subject classifications. Primary 62F40, 62F35; secondary 62J05, 62G09, 62G20.

Key words and phrases. Sample splitting, bootstrap, regression, assumption-lean.

- [11] BENTKUS, V. Y. (1985). Lower bounds for the rate of convergence in the central limit theorem in Banach spaces. *Lith. Math. J.* **25** 312–320.
- [12] BERK, R., BROWN, L., BUJA, A., ZHANG, K. and ZHAO, L. (2013). Valid post-selection inference. *Ann. Statist.* **41** 802–837. [MR3099122](#)
- [13] BÜHLMANN, P. (2013). Statistical significance in high-dimensional linear models. *Bernoulli* **19** 1212–1242. [MR3102549](#)
- [14] BÜHLMANN, P. and VAN DE GEER, S. (2015). High-dimensional inference in misspecified linear models. *Electron. J. Stat.* **9** 1449–1473. [MR3367666](#)
- [15] BUJA, A., BERK, R., BROWN, L., GEORGE, E., PITKIN, E., TRASKIN, M., ZHAO, L. and ZHANG, K. (2015). Models as approximations—A conspiracy of random regressors and model deviations against classical inference in regression. *Statist. Sci.* **1460**.
- [16] CANDÈS, E., FAN, Y., JANSO, L. and LV, J. (2018). Panning for gold: “model-X” knockoffs for high dimensional controlled variable selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 551–577. [MR3798878](#)
- [17] CHATTERJEE, A. and LAHIRI, S. N. (2011). Bootstrapping lasso estimators. *J. Amer. Statist. Assoc.* **106** 608–625. [MR2847974](#)
- [18] CHATTERJEE, A. and LAHIRI, S. N. (2013). Rates of convergence of the adaptive LASSO estimators to the oracle distribution and higher order refinements by the bootstrap. *Ann. Statist.* **41** 1232–1259. [MR3113809](#)
- [19] CHEN, L. H. Y. and SHAO, Q.-M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** 581–599. [MR2331265](#)
- [20] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. [MR3161448](#)
- [21] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields* **162** 47–70. [MR3350040](#)
- [22] CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **45** 2309–2352. [MR3693963](#)
- [23] COX, D. R. (1975). A note on data-splitting for the evaluation of significance levels. *Biometrika* **62** 441–444. [MR0378189](#)
- [24] DEZEURE, R., BÜHLMANN, P., MEIER, L. and MEINSHAUSEN, N. (2015). High-dimensional inference: Confidence intervals, p -values and R-software hdi. *Statist. Sci.* **30** 533–558. [MR3432840](#)
- [25] DEZEURE, R., BÜHLMANN, P. and ZHANG, C.-H. (2017). High-dimensional simultaneous inference with the bootstrap. *TEST* **26** 685–719. [MR3713586](#)
- [26] EFRON, B. (2014). Estimation and accuracy after model selection. *J. Amer. Statist. Assoc.* **109** 991–1007. [MR3265671](#)
- [27] FARAWAY, J. J. (1995). Data splitting strategies for reducing the effect of model selection on inference. Technical report, Citeseer.
- [28] FITHIAN, W., SUN, D. L. and TAYLOR, J. (2014). Optimal inference after model selection. Available at [arXiv:1410.2597](#).
- [29] HARTIGAN, J. A. (1969). Using subsample values as typical values. *J. Amer. Statist. Assoc.* **64** 1303–1317. [MR0261737](#)
- [30] HJORT, N. L. and CLAESKENS, G. (2003). Frequentist model average estimators. *J. Amer. Statist. Assoc.* **98** 879–899. [MR2041481](#)
- [31] HSU, D., KAKADE, S. M. and ZHANG, T. (2014). Random design analysis of ridge regression. *Found. Comput. Math.* **14** 569–600. [MR3201956](#)
- [32] HURVICH, C. M. and TSAI, C. (1990). The impact of model selection on inference in linear regression. *Amer. Statist.* **44** 214–217.

- [33] JAVANMARD, A. and MONTANARI, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909. [MR3277152](#)
- [34] LEE, J. D., SUN, D. L., SUN, Y. and TAYLOR, J. E. (2016). Exact post-selection inference, with application to the lasso. *Ann. Statist.* **44** 907–927. [MR3485948](#)
- [35] LEEB, H. and PÖTSCHER, B. M. (2008). Can one estimate the unconditional distribution of post-model-selection estimators? *Econometric Theory* **24** 338–376. [MR2422862](#)
- [36] LEI, J., G’SSELL, M., RINALDO, A., TIBSHIRANI, R. J. and WASSERMAN, L. (2018). Distribution-free predictive inference for regression. *J. Amer. Statist. Assoc.* **113** 1094–1111. [MR3862342](#)
- [37] LI, K.-C. (1989). Honest confidence regions for nonparametric regression. *Ann. Statist.* **17** 1001–1008. [MR1015135](#)
- [38] LOCKHART, R., TAYLOR, J., TIBSHIRANI, R. J. and TIBSHIRANI, R. (2014). A significance test for the lasso. *Ann. Statist.* **42** 413–468. [MR3210970](#)
- [39] LOFTUS, J. R. and TAYLOR, J. E. (2015). Selective inference in regression models with groups of variables. Preprint. Available at [arXiv:1511.01478](#).
- [40] MARKOVIC, J. and TAYLOR, J. (2016). Bootstrap inference after using multiple queries for model selection. Available at [arXiv:1612.07811](#).
- [41] MARKOVIC, J., XIA, L. and TAYLOR, J. (2017). Comparison of prediction errors: Adaptive p-values after cross-validation. Available at [arXiv:1703.06559](#).
- [42] MEINSHAUSEN, N. (2015). Group bound: Confidence intervals for groups of variables in sparse high dimensional regression without assumptions on the design. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 923–945. [MR3414134](#)
- [43] MEINSHAUSEN, N. and BÜHLMANN, P. (2010). Stability selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 417–473. [MR2758523](#)
- [44] MEINSHAUSEN, N., MEIER, L. and BÜHLMANN, P. (2009). *p*-values for high-dimensional regression. *J. Amer. Statist. Assoc.* **104** 1671–1681. [MR2750584](#)
- [45] MENTCH, L. and HOOKER, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *J. Mach. Learn. Res.* **17** Paper No. 26, 41. [MR3491120](#)
- [46] MILLER, A. J. (1990). *Subset Selection in Regression. Monographs on Statistics and Applied Probability* **40**. CRC Press, London. [MR1072361](#)
- [47] MORAN, P. A. P. (1973). Dividing a sample into two parts. A statistical dilemma. *Sankhyā Ser. A* **35** 329–333. [MR0518783](#)
- [48] MOSTELLER, F. and TUKEY, J. W. (1977). *Data Analysis and Regression: A Second Course in Statistics*. Addison-Wesley Series in Behavioral Science: Quantitative Methods.
- [49] NAZAROV, F. (2003). On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a Gaussian measure. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1807** 169–187. Springer, Berlin. [MR2083397](#)
- [50] NICKL, R. and VAN DE GEER, S. (2013). Confidence sets in sparse regression. *Ann. Statist.* **41** 2852–2876. [MR3161450](#)
- [51] PICARD, R. R. and BERK, K. N. (1990). Data splitting. *Amer. Statist.* **44** 140–147.
- [52] PINELIS, I. and MOLZON, R. (2016). Optimal-order bounds on the rate of convergence to normality in the multivariate delta method. *Electron. J. Stat.* **10** 1001–1063. [MR3486424](#)
- [53] PORTNOY, S. (1987). A central limit theorem applicable to robust regression estimators. *J. Multivariate Anal.* **22** 24–50. [MR0890880](#)
- [54] POUZO, D. (2015). Bootstrap consistency for quadratic forms of sample averages with increasing dimension. *Electron. J. Stat.* **9** 3046–3097. [MR3450756](#)
- [55] RINALDO, A., WASSERMAN, L. and G’SSELL, M. (2019). Supplement to “Bootstrapping and sample splitting for high-dimensional, assumption-lean inference.” DOI:[10.1214/18-AOS1784SUPP](#).

- [56] SHAH, R. D. and BÜHLMANN, P. (2018). Goodness-of-fit tests for high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 113–135. [MR3744714](#)
- [57] SHAH, R. D. and SAMWORTH, R. J. (2013). Variable selection with error control: Another look at stability selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 55–80. [MR3008271](#)
- [58] SHAO, J. (1993). Linear model selection by cross-validation. *J. Amer. Statist. Assoc.* **88** 486–494. [MR1224373](#)
- [59] SHAO, Q.-M., ZHANG, K. and ZHOU, W.-X. (2016). Stein’s method for nonlinear statistics: A brief survey and recent progress. *J. Statist. Plann. Inference* **168** 68–89. [MR3412222](#)
- [60] SHORACK, G. R. (2000). *Probability for Statisticians. Springer Texts in Statistics.* Springer, New York. [MR1762415](#)
- [61] TIAN, X. and TAYLOR, J. (2018). Selective inference with a randomized response. *Ann. Statist.* **46** 679–710. [MR3782381](#)
- [62] TIBSHIRANI, R. J., RINALDO, A., TIBSHIRANI, R. and WASSERMAN, L. (2018). Uniform asymptotic inference and the bootstrap after model selection. *Ann. Statist.* **46** 1255–1287. [MR3798003](#)
- [63] TIBSHIRANI, R. J., TAYLOR, J., LOCKHART, R. and TIBSHIRANI, R. (2016). Exact post-selection inference for sequential regression procedures. *J. Amer. Statist. Assoc.* **111** 600–620. [MR3538689](#)
- [64] VAN DE GEER, S., BÜHLMANN, P., RITOV, Y. and DEZEURE, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42** 1166–1202. [MR3224285](#)
- [65] WAGER, S., HASTIE, T. and EFRON, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *J. Mach. Learn. Res.* **15** 1625–1651. [MR3225243](#)
- [66] WASSERMAN, L. (2014). Discussion: “A significance test for the lasso” [MR3210970]. *Ann. Statist.* **42** 501–508. [MR3210975](#)
- [67] WASSERMAN, L. and ROEDER, K. (2009). High-dimensional variable selection. *Ann. Statist.* **37** 2178–2201. [MR2543689](#)
- [68] ZHANG, C.-H. and ZHANG, S. S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 217–242. [MR3153940](#)
- [69] ZHANG, X. and CHENG, G. (2017). Simultaneous inference for high-dimensional linear models. *J. Amer. Statist. Assoc.* **112** 757–768. [MR3671768](#)

JOINT CONVERGENCE OF SAMPLE AUTOCOVARANCE MATRICES WHEN $p/n \rightarrow 0$ WITH APPLICATION

BY MONIKA BHATTACHARJEE¹ AND ARUP BOSE²

University of Florida and Indian Statistical Institute

Consider a high-dimensional linear time series model where the dimension p and the sample size n grow in such a way that $p/n \rightarrow 0$. Let $\hat{\Gamma}_u$ be the u th order sample autocovariance matrix. We first show that the LSD of any symmetric polynomial in $\{\hat{\Gamma}_u, \hat{\Gamma}_u^*, u \geq 0\}$ exists under independence and moment assumptions on the driving sequence together with weak assumptions on the coefficient matrices. This LSD result, with some additional effort, implies the asymptotic normality of the trace of any polynomial in $\{\hat{\Gamma}_u, \hat{\Gamma}_u^*, u \geq 0\}$. We also study similar results for several independent MA processes.

We show applications of the above results to statistical inference problems such as in estimation of the unknown order of a high-dimensional MA process and in graphical and significance tests for hypotheses on coefficient matrices of one or several such independent processes.

REFERENCES

- BAI, Z. D. and SILVERSTEIN, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. [MR2040792](#)
- BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. Springer, New York. [MR2567175](#)
- BAI, Z., WANG, X. and ZHOU, W. (2009). CLT for linear spectral statistics of Wigner matrices. *Electron. J. Probab.* **14** 2391–2417. [MR2556016](#)
- BAI, Z. D. and YIN, Y. Q. (1988). Convergence to the semicircle law. *Ann. Probab.* **16** 863–875. [MR0929083](#)
- BAI, Z., JIANG, D., YAO, J. and ZHENG, S. (2013). Testing linear hypotheses in high-dimensional regressions. *Statistics* **47** 1207–1223. [MR3175743](#)
- BAO, Z., LIN, L.-C., PAN, G. and ZHOU, W. (2015). Spectral statistics of large dimensional Spearman’s rank correlation matrix and its application. *Ann. Statist.* **43** 2588–2623. [MR3405605](#)
- BASU, S. and MICHAILIDIS, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567. [MR3357870](#)
- BHATTACHARJEE, M. and BOSE, A. (2014). Estimation of autocovariance matrices for infinite dimensional vector linear process. *J. Time Series Anal.* **35** 262–281. [MR3194965](#)
- BHATTACHARJEE, M. and BOSE, A. (2016a). Large sample behaviour of high dimensional autocovariance matrices. *Ann. Statist.* **44** 598–628. [MR3476611](#)
- BHATTACHARJEE, M. and BOSE, A. (2016b). Polynomial generalizations of the sample variance-covariance matrix when $pn^{-1} \rightarrow 0$. *Random Matrices Theory Appl.* **5** 1650014. [MR3569842](#)

MSC2010 subject classifications. Primary 62M10; secondary 37M10, 58C40.

Key words and phrases. Moving average process, sample autocovariance matrices, limiting spectral distribution, trace, asymptotic normality, estimation, testing of hypothesis.

- BHATTACHARJEE, M. and BOSE, A. (2019). Supplement to “Joint convergence of sample autocovariance matrices when $p/n \rightarrow 0$ with application.” DOI:[10.1214/18-AOS1785SUPP](https://doi.org/10.1214/18-AOS1785SUPP).
- BHATTACHARJEE, M. (2016). Asymptotics of Large Variance-Covariance and Autocovariance Matrices. PhD Thesis, Indian Statistical Institute.
- BROCKWELL, P. J. and DAVIS, R. A. (2006). *Time Series: Theory and Methods*. Springer Series in Statistics. Springer, New York. [MR2839251](#)
- DIACONIS, P. and EVANS, S. N. (2001). Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** 2615–2633. [MR1828463](#)
- HOSKING, J. R. M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.* **75** 602–608. [MR0590689](#)
- LIU, H., AUE, A. and PAUL, D. (2015). On the Marčenko–Pastur law for linear time series. *Ann. Statist.* **43** 675–712. [MR3319140](#)
- MARČENKO, V. and PASTUR, L. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR, Sb.* **1** 457–483.
- NICA, A. and SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability*. London Mathematical Society Lecture Note Series **335**. Cambridge Univ. Press, Cambridge. [MR2266879](#)
- PFÄFFEL, O. and SCHLEMM, E. (2011). Eigenvalue distribution of large sample covariance matrices of linear processes. *Probab. Math. Statist.* **31** 313–329. [MR2853681](#)
- SHAO, X. (2011). Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models. *Econometric Theory* **27** 312–343. [MR2782041](#)
- WANG, L., AUE, A. and PAUL, D. (2017). Spectral analysis of sample autocovariance matrices of a class of linear time series in moderately high dimensions. *Bernoulli* **23** 2181–2209. [MR3648029](#)
- XIAO, H. and WU, W. B. (2014). Portmanteau test and simultaneous inference for serial covariances. *Statist. Sinica* **24** 577–599. [MR3235390](#)
- ZHENG, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F -matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 444–476. [MR2954263](#)
- ZHENG, S., BAI, Z. and YAO, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Statist.* **43** 546–591. [MR3316190](#)

TRACY–WIDOM LIMIT FOR KENDALL’S TAU

BY ZHIGANG BAO¹

Hong Kong University of Science and Technology

In this paper, we study a high-dimensional random matrix model from nonparametric statistics called the Kendall rank correlation matrix, which is a natural multivariate extension of the Kendall rank correlation coefficient. We establish the Tracy–Widom law for its largest eigenvalue. It is the first Tracy–Widom law for a nonparametric random matrix model, and also the first Tracy–Widom law for a high-dimensional U-statistic.

REFERENCES

- [1] BAI, Z. and ZHOU, W. (2008). Large sample covariance matrices without independence structures in columns. *Statist. Sinica* **18** 425–442. [MR2411613](#)
- [2] BAIK, J., BEN AROUS, G. and PÉCHÉ, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. [MR2165575](#)
- [3] BANDEIRA, A. S., LODHIA, A. and RIGOLLET, P. (2017). Marčenko–Pastur law for Kendall’s tau. *Electron. Commun. Probab.* **22** Paper no. 32. [MR3663103](#)
- [4] BAO, Z., LIN, L.-C., PAN, G. and ZHOU, W. (2015). Spectral statistics of large dimensional Spearman’s rank correlation matrix and its application. *Ann. Statist.* **43** 2588–2623. [MR3405605](#)
- [5] BAO, Z., PAN, G. and ZHOU, W. (2012). Tracy–Widom law for the extreme eigenvalues of sample correlation matrices. *Electron. J. Probab.* **17** Paper no. 88. [MR2988403](#)
- [6] BAO, Z., PAN, G. and ZHOU, W. (2015). Universality for the largest eigenvalue of sample covariance matrices with general population. *Ann. Statist.* **43** 382–421. [MR3311864](#)
- [7] BAO, Z. (2019). Supplement to “Tracy–Widom limit for Kendall’s tau.” DOI:[10.1214/18-AOS1786SUPP](https://doi.org/10.1214/18-AOS1786SUPP).
- [8] BAO, Z. G. (2018). Tracy–Widom limit for Spearman’s rho. Preprint.
- [9] BLOEMENDAL, A., KNOWLES, A., YAU, H.-T. and YIN, J. (2016). On the principal components of sample covariance matrices. *Probab. Theory Related Fields* **164** 459–552. [MR3449395](#)
- [10] DING, X. and YANG, F. (2018). A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices. *Ann. Appl. Probab.* **28** 1679–1738. [MR3809475](#)
- [11] EL KAROUI, N. (2006). A rate of convergence result for the largest eigenvalue of complex white Wishart matrices. *Ann. Probab.* **34** 2077–2117. [MR2294977](#)
- [12] EL KAROUI, N. (2007). Tracy–Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.* **35** 663–714. [MR2308592](#)
- [13] ERDŐS, L., KNOWLES, A. and YAU, H.-T. (2013). Averaging fluctuations in resolvents of random band matrices. *Ann. Henri Poincaré* **14** 1837–1926. [MR3119922](#)

MSC2010 subject classifications. Primary 60B20, 62G10; secondary 62H10, 15B52, 62H25.

Key words and phrases. Tracy–Widom law, largest eigenvalue, nonparametric statistics, U-statistics, random matrices.

- [14] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** 341–407. [MR2981427](#)
- [15] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229** 1435–1515. [MR2871147](#)
- [16] FAN, Z. and JOHNSTONE, I. (2017). Tracy–Widom at each edge of real covariance estimators. Available at [arXiv:1707.02352](#).
- [17] GAO, J., HAN, X., PAN, G. and YANG, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 677–693. [MR3641402](#)
- [18] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325. [MR0026294](#)
- [19] JIANG, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880. [MR2052906](#)
- [20] JOHANSSON, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** 437–476. [MR1737991](#)
- [21] JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#)
- [22] JOHNSTONE, I. M. and MA, Z. (2012). Fast approach to the Tracy–Widom law at the edge of GOE and GUE. *Ann. Appl. Probab.* **22** 1962–1988. [MR3025686](#)
- [23] KNOWLES, A. and YIN, J. (2017). Anisotropic local laws for random matrices. *Probab. Theory Related Fields* **169** 257–352. [MR3704770](#)
- [24] LEE, J. O. and SCHNELLI, K. (2015). Edge universality for deformed Wigner matrices. *Rev. Math. Phys.* **27** 1550018. [MR3405746](#)
- [25] LEE, J. O. and SCHNELLI, K. (2016). Tracy–Widom distribution for the largest eigenvalue of real sample covariance matrices with general population. *Ann. Appl. Probab.* **26** 3786–3839. [MR3582818](#)
- [26] LEE, J. O. and SCHNELLI, K. (2018). Local law and Tracy–Widom limit for sparse random matrices. *Probab. Theory Related Fields* **171** 543–616. [MR3800840](#)
- [27] MA, Z. (2012). Accuracy of the Tracy–Widom limits for the extreme eigenvalues in white Wishart matrices. *Bernoulli* **18** 322–359. [MR2888709](#)
- [28] MARCHENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR, Sb.* **1** 457.
- [29] ONATSKI, A. (2008). The Tracy–Widom limit for the largest eigenvalues of singular complex Wishart matrices. *Ann. Appl. Probab.* **18** 470–490. [MR2398763](#)
- [30] PILLAI, N. S. and YIN, J. (2012). Edge universality of correlation matrices. *Ann. Statist.* **40** 1737–1763. [MR3015042](#)
- [31] PILLAI, N. S. and YIN, J. (2014). Universality of covariance matrices. *Ann. Appl. Probab.* **24** 935–1001. [MR3199978](#)
- [32] WANG, K. (2012). Random covariance matrices: Universality of local statistics of eigenvalues up to the edge. *Random Matrices Theory Appl.* **1** 1150005. [MR2930383](#)
- [33] ZHOU, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363. [MR2327033](#)

INTRINSIC RIEMANNIAN FUNCTIONAL DATA ANALYSIS¹

BY ZHENHUA LIN^{*} AND FANG YAO[†]

National University of Singapore^{} and Peking University[†]*

In this work we develop a novel and foundational framework for analyzing general Riemannian functional data, in particular a new development of tensor Hilbert spaces along curves on a manifold. Such spaces enable us to derive Karhunen–Loève expansion for Riemannian random processes. This framework also features an approach to compare objects from different tensor Hilbert spaces, which paves the way for asymptotic analysis in Riemannian functional data analysis. Built upon intrinsic geometric concepts such as vector field, Levi-Civita connection and parallel transport on Riemannian manifolds, the developed framework applies to not only Euclidean submanifolds but also manifolds without a natural ambient space. As applications of this framework, we develop intrinsic Riemannian functional principal component analysis (iRFPCA) and intrinsic Riemannian functional linear regression (iRFLR) that are distinct from their traditional and ambient counterparts. We also provide estimation procedures for iRFPCA and iRFLR, and investigate their asymptotic properties within the intrinsic geometry. Numerical performance is illustrated by simulated and real examples.

REFERENCES

- AFSARI, B. (2011). Riemannian L^p center of mass: Existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.* **139** 655–673. [MR2736346](#)
- ALDOUS, D. J. (1976). A characterisation of Hilbert space using the central limit theorem. *J. Lond. Math. Soc.* (2) **14** 376–380. [MR0443017](#)
- ARSIGNY, V., FILLARD, P., PENNEC, X. and AYACHE, N. (2006/07). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **29** 328–347. [MR2288028](#)
- BALAKRISHNAN, A. V. (1960). Estimation and detection theory for multiple stochastic processes. *J. Math. Anal. Appl.* **1** 386–410. [MR0131323](#)
- BHATTACHARYA, R. and PATRANGENARU, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Ann. Statist.* **31** 1–29. [MR1962498](#)
- BINDER, J. R., FROST, J. A., HAMMEKE, T. A., COX, R. W., RAO, S. M. and PRIETO, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *J. Neurosci.* **17** 353–362.
- BOSQ, D. (2000). *Linear Processes in Function Spaces. Lecture Notes in Statistics* **149**. Springer, New York. [MR1783138](#)
- CARDOT, H., FERRATY, F. and SARDA, P. (2003). Spline estimators for the functional linear model. *Statist. Sinica* **13** 571–591. [MR1997162](#)

MSC2010 subject classifications. 62G05, 62J05.

Key words and phrases. Functional principal component, functional linear regression, intrinsic Riemannian Karhunen–Loève expansion, parallel transport, tensor Hilbert space.

- CARDOT, H., MAS, A. and SARDA, P. (2007). CLT in functional linear regression models. *Probab. Theory Related Fields* **138** 325–361. [MR2299711](#)
- CHEN, D. and MÜLLER, H.-G. (2012). Nonlinear manifold representations for functional data. *Ann. Statist.* **40** 1–29. [MR3013177](#)
- CHENG, G., HO, J., SALEHIAN, H. and VEMURI, B. C. (2016). Recursive computation of the Fréchet mean on non-positively curved Riemannian manifolds with applications. In *Riemannian Computing in Computer Vision* 21–43. Springer, Cham. [MR3444345](#)
- CORNEA, E., ZHU, H., KIM, P. and IBRAHIM, J. G. (2017). Regression models on Riemannian symmetric spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 463–482. [MR3611755](#)
- DAI, X. and MÜLLER, H.-G. (2018). Principal component analysis for functional data on Riemannian manifolds and spheres. *Ann. Statist.* **46** 3334–3361. [MR3852654](#)
- DAUXOIS, J., POUSSE, A. and ROMAIN, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *J. Multivariate Anal.* **12** 136–154. [MR0650934](#)
- DAYAN, E. and COHEN, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron* **72** 443–454.
- DRYDEN, I. L., KOLOYDENKO, A. and ZHOU, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3** 1102–1123. [MR2750388](#)
- ESSEN, D. C. V., SMITH, S. M., BARCH, D. M., BEHRENS, T. E. J., YACCOUB, E., UGURBIL, K. and CONSORTIUM, W.-M. H. (2013). The WU-Minn human connectome project: An overview. *NeuroImage* **80** 62–79.
- FERRATY, F. and VIEU, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. *Springer Series in Statistics*. Springer, New York. [MR2229687](#)
- FLETCHER, P. T. and JOSHIB, S. (2007). Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Process.* **87** 250–262.
- FRISTON, K. J. (2011). Functional and effective connectivity: A review. *Brain Connect.* **1** 13–36.
- HALL, P. and HOROWITZ, J. L. (2005). Nonparametric methods for inference in the presence of instrumental variables. *Ann. Statist.* **33** 2904–2929. [MR2253107](#)
- HALL, P. and HOROWITZ, J. L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. [MR2332269](#)
- HALL, P. and HOSSEINI-NASAB, M. (2006). On properties of functional principal components analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 109–126. [MR2212577](#)
- HAPP, C. and GREVEN, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J. Amer. Statist. Assoc.* **113** 649–659. [MR3832216](#)
- HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. Springer, New York. [MR2722294](#)
- HSING, T. and EUBANK, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Wiley, Chichester. [MR3379106](#)
- KELLY, E. J. and ROOT, W. L. (1960). A representation of vector-valued random processes. *J. Math. Phys.* **39** 211–216. [MR0125624](#)
- KENDALL, W. S. and LE, H. (2011). Limit theorems for empirical Fréchet means of independent and non-identically distributed manifold-valued random variables. *Braz. J. Probab. Stat.* **25** 323–352. [MR2832889](#)
- KLEFFE, J. (1973). Principal components of random variables with values in a separable Hilbert space. *Statistics* **4** 391–406. [MR0391402](#)
- KOKOSZKA, P. and REIMHERR, M. (2017). *Introduction to Functional Data Analysis*. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. [MR3793167](#)

- KONG, D., XUE, K., YAO, F. and ZHANG, H. H. (2016). Partially functional linear regression in high dimensions. *Biometrika* **103** 147–159. [MR3465827](#)
- LANG, S. (1995). *Differential and Riemannian Manifolds*, 3rd ed. *Graduate Texts in Mathematics* **160**. Springer, New York. [MR1335233](#)
- LANG, S. (1999). *Fundamentals of Differential Geometry*. *Graduate Texts in Mathematics* **191**. Springer, New York. [MR1666820](#)
- LEE, J. M. (1997). *Riemannian Manifolds: An Introduction to Curvature*. *Graduate Texts in Mathematics* **176**. Springer, New York. [MR1468735](#)
- LEE, J. M. (2013). *Introduction to Smooth Manifolds*, 2nd ed. *Graduate Texts in Mathematics* **218**. Springer, New York. [MR2954043](#)
- LILA, E., ASTON, J. A. D. and SANGALLI, L. M. (2016). Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. *Ann. Appl. Stat.* **10** 1854–1879. [MR3592040](#)
- LIN, Z. and YAO, F. (2019). Functional regression with unknown manifold structures. Available at [arXiv:1704.03005](#).
- MOAKHER, M. (2005). A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **26** 735–747. [MR2137480](#)
- PARK, J. E., JUNG, S. C., RYU, K. H., OH, J. Y., KIM, H. S., CHOI, C. G., KIM, S. J. and SHIM, W. H. (2017). Differences in dynamic and static functional connectivity between young and elderly healthy adults. *Neuroradiol.* **59** 781–789.
- PETERSEN, A. and MÜLLER, H.-G. (2019). Fréchet regression for random objects with Euclidean predictors. *Ann. Statist.* **47** 691–719. [MR3909947](#)
- PHANA, K. L., WAGER, T., TAYLOR, S. F. and LIBERZON, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* **16** 331–348.
- RAICHLEN, D. A., BHARADWAJ, P. K., FITZHUGH, M. C., HAWS, K. A., TORRE, G.-A., TROUARD, T. P. and ALEXANDER, G. E. (2016). Differences in resting state functional connectivity between young adult endurance athletes and healthy controls. *Front. Human Neurosci.* **10** 610.
- RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. [MR2168993](#)
- RAO, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics* **14** 1–17.
- SALEHIAN, H., CHAKRABORTY, R., OFORI, E., VAILLANCOURT, D. and VEMURI, B. C. (2015). An efficient recursive estimator of the Fréchet mean on hypersphere with applications to Medical Image Analysis. In *5th MICCAI Workshop on Mathematical Foundations of Computational Anatomy (MFCA)*.
- SASAKI, S. (1958). On the differential geometry of tangent bundles of Riemannian manifolds. *Tôhoku Math. J. (2)* **10** 338–354. [MR0112152](#)
- SHI, X., STYNER, M., LIEBERMAN, J., IBRAHIM, J. G., LIN, W. and ZHU, H. (2009). Intrinsic regression models for manifold-valued data. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI* **12** 192–199.
- SILVERMAN, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *Ann. Statist.* **24** 1–24. [MR1389877](#)
- STEINKE, F., HEIN, M. and SCHÖLKOPF, B. (2010). Nonparametric regression between general Riemannian manifolds. *SIAM J. Imaging Sci.* **3** 527–563. [MR2736019](#)
- WANG, L. (2008). Karhunen–Loeve expansions and their applications Ph.D. thesis The London School of Economics and Political Science.
- WANG, J.-L., CHIOU, J.-M. and MÜLLER, H.-G. (2016). Review of functional data analysis. *Ann. Rev. Statist. Appl.* **3** 257–295.
- WANG, H. and MARRON, J. S. (2007). Object oriented data analysis: Sets of trees. *Ann. Statist.* **35** 1849–1873. [MR2363955](#)

- YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005a). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. [MR2160561](#)
- YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005b). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. [MR2253106](#)
- YUAN, M. and CAI, T. T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. [MR2766857](#)
- YUAN, Y., ZHU, H., LIN, W. and MARRON, J. S. (2012). Local polynomial regression for symmetric positive definite matrices. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 697–719. [MR2965956](#)
- ZHANG, X. and WANG, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321. [MR3546451](#)

DETECTING RELEVANT CHANGES IN THE MEAN OF NONSTATIONARY PROCESSES—A MASS EXCESS APPROACH¹

BY HOLGER DETTE* AND WEICHI WU^{†,*}

Ruhr-Universität Bochum and Tsinghua University[†]*

This paper considers the problem of testing if a sequence of means $(\mu_t)_{t=1,\dots,n}$ of a nonstationary time series $(X_t)_{t=1,\dots,n}$ is stable in the sense that the difference of the means μ_1 and μ_t between the initial time $t = 1$ and any other time is smaller than a given threshold, that is $|\mu_1 - \mu_t| \leq c$ for all $t = 1, \dots, n$. A test for hypotheses of this type is developed using a bias corrected monotone rearranged local linear estimator and asymptotic normality of the corresponding test statistic is established. As the asymptotic variance depends on the location of the roots of the equation $|\mu_1 - \mu_t| = c$ a new bootstrap procedure is proposed to obtain critical values and its consistency is established. As a consequence we are able to quantitatively describe relevant deviations of a nonstationary sequence from its initial value. The results are illustrated by means of a simulation study and by analyzing data examples.

REFERENCES

- [1] ÁLVAREZ-ESTEBAN, P. C., DEL BARRIO, E., CUESTA-ALBERTOS, J. A. and MATRÁN, C. (2008). Trimmed comparison of distributions. *J. Amer. Statist. Assoc.* **103** 697–704. [MR2435470](#)
- [2] ÁLVAREZ-ESTEBAN, P. C., DEL BARRIO, E., CUESTA-ALBERTOS, J. A. and MATRÁN, C. (2012). Similarity of samples and trimming. *Bernoulli* **18** 606–634. [MR2922463](#)
- [3] ANDREWS, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* **61** 821–856. [MR1231678](#)
- [4] AUE, A., HÖRMANN, S., HORVÁTH, L. and REIMHERR, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. [MR2572452](#)
- [5] AUE, A. and HORVÁTH, L. (2013). Structural breaks in time series. *J. Time Series Anal.* **34** 1–16. [MR3008012](#)
- [6] BAI, J. and PERRON, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66** 47–78. [MR1616121](#)
- [7] BAÍLLO, A. (2003). Total error in a plug-in estimator of level sets. *Statist. Probab. Lett.* **65** 411–417. [MR2039885](#)
- [8] BERMAN, S. M. (1992). *Sojourns and Extremes of Stochastic Processes. The Wadsworth & Brooks/Cole Statistics/Probability Series*. Wadsworth & Brooks/Cole, Pacific Grove, CA. [MR1126464](#)
- [9] BROWN, R. L., DURBIN, J. and EVANS, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. Roy. Statist. Soc. Ser. B* **37** 149–163. [MR0378310](#)

MSC2010 subject classifications. 62M10, 62F05, 62G08, 62G09.

Key words and phrases. Locally stationary process, change point analysis, relevant change points, local linear estimation, Gaussian approximation, rearrangement estimators.

- [10] CADRE, B. (2006). Kernel estimation of density level sets. *J. Multivariate Anal.* **97** 999–1023. [MR2256570](#)
- [11] CHAMP, C. W. and WOODALL, W. H. (1987). Exact results for Shewhart control charts with supplementary runs rules. *Technometrics* **29** 393–399.
- [12] CHANDLER, G. and POLONIK, W. (2006). Discrimination of locally stationary time series based on the excess mass functional. *J. Amer. Statist. Assoc.* **101** 240–253. [MR2268042](#)
- [13] CHENG, M.-Y. and HALL, P. (1998). Calibrating the excess mass and dip tests of modality. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 579–589. [MR1625938](#)
- [14] CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. and GALICHON, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika* **96** 559–575. [MR2538757](#)
- [15] CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. and GALICHON, A. (2010). Quantile and probability curves without crossing. *Econometrica* **78** 1093–1125. [MR2667913](#)
- [16] CHOW, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica* **28** 591–605. [MR0141193](#)
- [17] CUEVAS, A., GONZÁLEZ-MANTEIGA, W. and RODRÍGUEZ-CASAL, A. (2006). Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.* **48** 7–19. [MR2234775](#)
- [18] DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. [MR1429916](#)
- [19] DETTE, H., NEUMEYER, N. and PILZ, K. F. (2006). A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli* **12** 469–490. [MR2232727](#)
- [20] DETTE, H. and VOLGUSHEV, S. (2008). Non-crossing non-parametric estimates of quantile curves. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 609–627. [MR2420417](#)
- [21] DETTE, H. and WIED, D. (2016). Detecting relevant changes in time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 371–394. [MR3454201](#)
- [22] DETTE, H. and WU, W. (2019). Supplement to “Detecting relevant changes in the mean of nonstationary processes—A mass excess approach.” DOI:[10.1214/19-AOS1811SUPP](#).
- [23] GIJBELS, I., LAMBERT, A. and QIU, P. (2007). Jump-preserving regression and smoothing using local linear fitting: A compromise. *Ann. Inst. Statist. Math.* **59** 235–272. [MR2394168](#)
- [24] HARTIGAN, J. A. and HARTIGAN, P. M. (1985). The dip test of unimodality. *Ann. Statist.* **13** 70–84. [MR0773153](#)
- [25] JANDHYALA, V., FOTOPOULOS, S., MACNEILL, I. and LIU, P. (2013). Inference for single and multiple change-points in time series. *J. Time Series Anal.* **34** 423–446. [MR3070866](#)
- [26] KRÄMER, W., PLOBERGER, W. and ALT, R. (1988). Testing for structural change in dynamic models. *Econometrica* **56** 1355–1369. [MR0973984](#)
- [27] MASON, D. M. and POLONIK, W. (2009). Asymptotic normality of plug-in level set estimates. *Ann. Appl. Probab.* **19** 1108–1142. [MR2537201](#)
- [28] MERCURIO, D. and SPOKOINY, V. (2004). Statistical inference for time-inhomogeneous volatility models. *Ann. Statist.* **32** 577–602. [MR2060170](#)
- [29] MÜLLER, D. W. and SAWITZKI, G. (1991). Excess mass estimates and tests for multimodality. *J. Amer. Statist. Assoc.* **86** 738–746. [MR1147099](#)
- [30] NASON, G. P., VON SACHS, R. and KROISANDT, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 271–292. [MR1749539](#)
- [31] OMBAO, H., VON SACHS, R. and GUO, W. (2005). SLEX analysis of multivariate nonstationary time series. *J. Amer. Statist. Assoc.* **100** 519–531. [MR2160556](#)
- [32] PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115. [MR0088850](#)
- [33] POLONIK, W. (1995). Measuring mass concentrations and estimating density contour clusters—An excess mass approach. *Ann. Statist.* **23** 855–881. [MR1345204](#)

- [34] POLONIK, W. and WANG, Z. (2005). Estimation of regression contour clusters—An application of the excess mass approach to regression. *J. Multivariate Anal.* **94** 227–249. [MR2167913](#)
- [35] QIU, P. (2003). A jump-preserving curve fitting procedure based on local piecewise-linear kernel estimation. *J. Nonparametr. Stat.* **15** 437–453. [MR2017479](#)
- [36] RINALDO, A. and WASSERMAN, L. (2010). Generalized density clustering. *Ann. Statist.* **38** 2678–2722. [MR2722453](#)
- [37] SAMWORTH, R. J. and WAND, M. P. (2010). Asymptotics and optimal bandwidth selection for highest density region estimation. *Ann. Statist.* **38** 1767–1792. [MR2662359](#)
- [38] SCHUCANY, W. R. and SOMMERS, J. P. (1977). Improvement of kernel type density estimators. *J. Amer. Statist. Assoc.* **72** 420–423. [MR0448691](#)
- [39] SPOKOINY, V. (2009). Multiscale local change point detection with applications to value-at-risk. *Ann. Statist.* **37** 1405–1436. [MR2509078](#)
- [40] TAKÁCS, L. (1996). Sojourn times. *J. Appl. Math. Stoch. Anal.* **9** 415–426. [MR1429264](#)
- [41] TSYBAKOV, A. B. (1997). On nonparametric estimation of density level sets. *Ann. Statist.* **25** 948–969. [MR1447735](#)
- [42] VOGT, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. [MR3097614](#)
- [43] WELLEK, S. (2010). *Testing Statistical Hypotheses of Equivalence and Noninferiority*, 2nd ed. CRC Press, Boca Raton, FL. [MR2676002](#)
- [44] WOODALL, W. H. and MONTGOMERY, D. C. (1999). Research issues and ideas in statistical process control. *J. Qual. Technol.* **31** 376–386.
- [45] WU, W. B. and POURAHMADI, M. (2009). Banding sample autocovariance matrices of stationary processes. *Statist. Sinica* **19** 1755–1768. [MR2589209](#)
- [46] WU, W. B. and ZHAO, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. [MR2323759](#)
- [47] ZHOU, Z. and WU, W. B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. [MR2541444](#)
- [48] ZHOU, Z. and WU, W. B. (2010). Simultaneous inference of linear models with time varying coefficients. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 513–531. [MR2758526](#)

THE ANNALS
of
STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

VOLUME 47

2019

CONTENTS OF VOLUME 47

Articles

AAMARI, EDDIE AND LEVRARD, CLÉMENT. Nonasymptotic rates for manifold, tangent space and curvature estimation.....	177–204
ALQUIER, PIERRE, COTTET, VINCENT AND LECUÉ, GUILLAUME. Estimation bounds and sharp oracle inequalities of regularized procedures with Lipschitz loss functions.....	2117–2144
ATHEY, SUSAN, TIBSHIRANI, JULIE AND WAGER, STEFAN. Generalized random forests	1148–1178
BACHOC, FRANÇOIS, LEEB, HANNES AND PÖTSCHER, BENEDIKT M. Valid confidence intervals for post-model-selection predictors.....	1475–1504
BALAKRISHNAN, SIVARAMAN AND WASSERMAN, LARRY. Hypothesis testing for densities and high-dimensional multinomials: Sharp local minimax rates	1893–1927
BANERJEE, MOULINATH, DUROT, CÉCILE AND SEN, BODHISATTVA. Divide and conquer in nonstandard problems and the super-efficiency phenomenon.....	720–757
BAO, ZHIGANG. Tracy–Widom limit for Kendall’s tau	3504–3532
BAO, ZHIGANG, HU, JIANG, PAN, GUANGMING AND ZHOU, WANG. Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case.....	612–640
BARBER, RINA F., WAINWRIGHT, MARTIN J., JORDAN, MICHAEL I. AND RAMDAS, AADITYA K. A unified treatment of multiple testing with prior knowledge using the p-filter	2790–2821
BARBER, RINA FOYGEL AND CANDÈS, EMMANUEL J. A knock-off filter for high-dimensional selective inference	2504–2537
BASU, KINJAL, ZHAO, QINGYUAN, OWEN, ART B. AND HE, HERA Y. Permutation p -value approximation via generalized Stolarsky invariance	583–611
BAUER, BENEDIKT AND KOHLER, MICHAEL. On deep learning as a remedy for the curse of dimensionality in nonparametric regression	2261–2285
BERRETT, THOMAS B., SAMWORTH, RICHARD J. AND YUAN, MING. Efficient multivariate entropy estimation via k -nearest neighbour distances	288–318
BERTHET, QUENTIN, RIGOLLET, PHILIPPE AND SRIVASTAVA, PIYUSH. Exact recovery in the Ising blockmodel.....	1805–1834
BEVILACQUA, MORENO, FAOUZI, TARIK, FURRER, REINHARD AND PORCU, EMILIO. Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics.....	828–856

BHATTACHARJEE, MONIKA AND BOSE, ARUP. Joint convergence of sample autocovariance matrices when $p/n \rightarrow 0$ with application	3470–3503
BHATTACHARYA, ANIRBAN, PATI, DEBDEEP AND YANG, YUN. Bayesian fractional posteriors	39–66
BIERKENS, JORIS, FEARNHEAD, PAUL AND ROBERTS, GARETH. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data	1288–1320
BING, XIN AND WEGKAMP, MARTEN H. Adaptive estimation of the rank of the coefficient matrix in high-dimensional multivariate response regression models	3157–3184
BOBKOV, SERGEY G. Khinchine’s theorem and Edgeworth approximations for weighted sums	1616–1633
BODNAR, TARAS, DETTE, HOLGER AND PAROLYA, NESTOR. Testing for independence of large dimensional vectors	2977–3008
BOSE, ARUP AND BHATTACHARJEE, MONIKA. Joint convergence of sample autocovariance matrices when $p/n \rightarrow 0$ with application	3470–3503
BÖTTCHER, BJÖRN, KELLER-RESSEL, MARTIN AND SCHILLING, RENÉ L. Distance multivariate: New dependence measures for random vectors	2757–2789
BOUCHARD-CÔTÉ, ALEXANDRE, DOUCET, ARNAUD AND DELIGIANNIDIS, GEORGE. Exponential ergodicity of the bouncy particle sampler	1268–1287
BROWN, LAWRENCE D., CAI, T. TONY AND ZHANG, ANRU. Semi-supervised inference: General theory and estimation of means	2538–2566
BÜHLMANN, PETER, MEINSHAUSEN, NICOLAI AND ROTHENHÄUSLER, DOMINIK. Causal Dantzig: Fast inference in linear structural equation models with hidden variables under additive interventions	1688–1722
CAI, DINGJIAO, SUN, JIANGUO, TONG, XINGWEI, GAO, FUQING AND CHEN, KANI. Maximum likelihood estimation in transformed linear regression with nonnormal errors	1864–1892
CAI, T. TONY, MA, JING AND ZHANG, LINJUN. CHIME: Clustering of high-dimensional Gaussian mixtures with EM algorithm and its optimality	1234–1267
CAI, T. TONY, ZHANG, ANRU AND BROWN, LAWRENCE D. Semi-supervised inference: General theory and estimation of means	2538–2566

CAMERLENGHI, FEDERICO, LIJOI, ANTONIO, ORBANZ, PETER AND PRÜNSTER, IGOR. Distribution theory for hierarchical processes	67–92
CANDÈS, EMMANUEL J. AND BARBER, RINA FOYGEL. A knock- off filter for high-dimensional selective inference	2504–2537
CAO, XUAN, KHARE, KSHITIJ AND GHOSH, MALAY. Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models	319–348
CAPE, JOSHUA, TANG, MINH AND PRIEBE, CAREY E. The two- to-infinity norm and singular subspace geometry with applica- tions to high-dimensional statistics	2405–2439
CARPENTIER, ALEXANDRA AND VERZELEN, NICOLAS. Adap- tive estimation of the sparsity in the Gaussian vector model	93–126
CARROLL, RAYMOND J., MA, SHUJIE, ZHU, LIPING, ZHANG, ZHIWEI AND TSAI, CHIH-LING. A robust and efficient ap- proach to causal inference based on sparse sufficient dimension reduction	1505–1535
CHAKRABORTY, RUDRASIS AND VEMURI, BABA C. Statistics on the Stiefel manifold: Theory and applications	415–438
CHANG, HSIN-WEN AND MCKEAGUE, IAN W. Nonparametric testing for multiple survival functions with noninferiority mar- gins	205–232
CHANG, MING-CHUNG, CHENG, SHAO-WEI AND CHENG, CHING- SHUI. Signal aliasing in Gaussian random fields for experi- ments with qualitative factors	909–935
CHAO, SHIH-KANG, CHENG, GUANG AND VOLGUSHEV, STANISLAV. Distributed inference for quantile regression pro- cesses	1634–1662
CHATTERJEE, SABYASACHI, SAMWORTH, RICHARD J., HAN, QIYANG AND WANG, TENG YAO. Isotonic regression in gen- eral dimensions	2440–2471
CHEN, HAN, RASKUTTI, GARVESH AND YUAN, MING. Convex regularization for high-dimensional multiresponse tensor regres- sion	1554–1584
CHEN, HAO. Sequential change-point detection based on nearest neighbors	1381–1407
CHEN, HAO AND CHU, LYNNA. Asymptotic distribution-free change-point detection for multivariate and non-Euclidean data	382–414
CHEN, KANI, CAI, DINGJIAO, SUN, JIANGUO, TONG, XINGWEI AND GAO, FUQING. Maximum likelihood estimation in trans- formed linear regression with nonnormal errors	1864–1892

CHEN, NINGYUAN, LEE, DONALD K. K. AND NEGAHBAN, SAHAND N. Super-resolution estimation of cyclic arrival rates . .	1754–1775
CHEN, SONG XI, LI, JUN AND ZHONG, PING-SHOU. Two-sample and ANOVA tests for high dimensional means	1443–1474
CHEN, WEI-KUO. Phase transition in the spiked random tensor with Rademacher prior	2734–2756
CHEN, XI, LIU, WEIDONG AND ZHANG, YICHEN. Quantile regression under memory constraint	3244–3273
CHEN, XIAOHUI AND KATO, KENGO. Randomized incomplete U -statistics in high dimensions	3127–3156
CHEN, YEN-CHI. Generalized cluster trees and singular measures .	2174–2203
CHEN, YUXIN, FAN, JIANQING, MA, CONG AND WANG, KAIZHENG. Spectral method and regularized MLE are both optimal for top- K ranking	2204–2235
CHEN, ZHAO, CUI, HENGJIAN, LI, RUNZE AND ZHENG, SHURONG. Hypothesis testing on linear structures of high-dimensional covariance matrix	3300–3334
CHEN, ZHAO, LI, RUNZE, SHI, CHENGCHUN AND SONG, RUI. Linear hypothesis testing for high dimensional generalized linear models	2671–2703
CHENG, CHING-SHUI, CHANG, MING-CHUNG AND CHENG, SHAO-WEI. Signal aliasing in Gaussian random fields for experiments with qualitative factors	909–935
CHENG, GUANG, VOLGUSHEV, STANISLAV AND CHAO, SHIH-KANG. Distributed inference for quantile regression processes	1634–1662
CHENG, GUANGHUI, GUO, JIANHUA, ZHU, HONGTU AND ZHENG, SHURONG. Test for high-dimensional correlation matrices . .	2887–2921
CHENG, SHAO-WEI, CHENG, CHING-SHUI AND CHANG, MING-CHUNG. Signal aliasing in Gaussian random fields for experiments with qualitative factors	909–935
CHÉTELAT, DIDIER AND WELLS, MARTIN T. The middle-scale asymptotics of Wishart matrices	2639–2670
CHOPIN, NICOLAS, WHITELEY, NICK AND GERBER, MATHIEU. Negative association, ordering and convergence of resampling methods	2236–2260
CHU, LYNNA AND CHEN, HAO. Asymptotic distribution-free change-point detection for multivariate and non-Euclidean data	382–414
COOK, R. DENNIS AND FORZANI, LILIANA. Partial least squares prediction in high-dimensional regression	884–908
COTTET, VINCENT, LECUÉ, GUILLAUME AND ALQUIER, PIERRE. Estimation bounds and sharp oracle inequalities of regularized procedures with Lipschitz loss functions	2117–2144

CUESTA-ALBERTOS, JUAN A., GARCÍA-PORTUGUÉS, EDUARDO, FEBRERO-BANDE, MANUEL AND GONZÁLEZ-MANTEIGA, WENCESLAO. Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes	439–467
CUI, HENGJIAN, LI, RUNZE, ZHENG, SHURONG AND CHEN, ZHAO. Hypothesis testing on linear structures of high-dimen- sional covariance matrix	3300–3334
DAHLHAUS, RAINER AND RICHTER, STEFAN. Cross validation for locally stationary processes	2145–2173
DAS, DEBRAJ, GREGORY, KARL AND LAHIRI, S. N. Perturbation bootstrap in adaptive Lasso	2080–2116
DASGUPTA, SAYAN, GOLDBERG, YAIR AND KOSOROK, MICHAEL R. Feature elimination in kernel machines in moderately high dimensions	497–526
DE CASTRO, YOHANN, GAMBOA, FABRICE, HENRION, DIDIER, HESS, ROXANA AND LASSERRE, JEAN-BERNARD. Approx- imate optimal designs for multivariate polynomial regression . .	127–155
DELIGIANNIDIS, GEORGE, BOUCHARD-CÔTÉ, ALEXANDRE AND DOUCET, ARNAUD. Exponential ergodicity of the bouncy par- ticle sampler	1268–1287
DESCARY, MARIE-HÉLÈNE AND PANARETOS, VICTOR M. Func- tional data analysis by matrix completion	1–38
DETTE, HOLGER, PAROLYA, NESTOR AND BODNAR, TARAS. Testing for independence of large dimensional vectors	2977–3008
DETTE, HOLGER, PEPELYSHEV, ANDREY AND ZHIGLJAVSKY, ANATOLY. The BLUE in continuous-time regression models with correlated errors	1928–1959
DETTE, HOLGER AND WU, WEICHI. Detecting relevant changes in the mean of nonstationary processes—A mass excess approach	3578–3608
DOSS, CHARLES R. AND WELLNER, JON A. Inference for the mode of a log-concave density	2950–2976
DOUCET, ARNAUD, DELIGIANNIDIS, GEORGE AND BOUCHARD- CÔTÉ, ALEXANDRE. Exponential ergodicity of the bouncy particle sampler	1268–1287
DRTON, MATHIAS, FOX, CHRISTOPHER, KÄUFL, ANDREAS AND POULIOT, GUILLAUME. The maximum likelihood threshold of a path diagram	1536–1553
DRTON, MATHIAS, FOX, CHRISTOPHER AND WANG, Y. SAMUEL. Computation of maximum likelihood estimates in cyclic struc- tural equation models	663–690

DUROT, CÉCILE, SEN, BODHISATTVA AND BANERJEE, MOULI-NATH. Divide and conquer in nonstandard problems and the super-efficiency phenomenon.....	720–757
EENDEBAK, PIETER T., GOOS, PETER AND SCHOEN, ERIC D. A classification criterion for definitive screening designs	1179–1202
ELTZNER, BENJAMIN AND HUCKEMANN, STEPHAN F. A smeary central limit theorem for manifolds with application to high-dimensional spheres	3360–3381
ENIKEEVA, FARIDA AND HARCHAOUI, ZAID. High-dimensional change-point detection under sparse alternatives	2051–2079
FAN, JIANQING, MA, CONG, WANG, KAIZHENG AND CHEN, YUXIN. Spectral method and regularized MLE are both optimal for top- K ranking	2204–2235
FAN, JIANQING, WANG, DONG, WANG, KAIZHENG AND ZHU, ZIWEI. Distributed estimation of principal eigenspaces	3009–3031
FAN, ZHOU AND JOHNSTONE, IAIN M. Eigenvalue distributions of variance components estimators in high-dimensional random effects models	2855–2886
FAOUZI, TARIK, FURRER, REINHARD, PORCU, EMILIO AND BEVILACQUA, MORENO. Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics	828–856
FEARNHEAD, PAUL, ROBERTS, GARETH AND BIERKENS, JORIS. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data	1288–1320
FEBRERO-BANDE, MANUEL, GONZÁLEZ-MANTEIGA, WENCESLAO, CUESTA-ALBERTOS, JUAN A. AND GARCÍA-PORTUGUÉS, EDUARDO. Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes	439–467
FELLOURIS, GEORGIOS AND SONG, YANGLEI. Sequential multiple testing with generalized error control: An asymptotic optimality theory	1776–1803
FENG, LONG AND ZHANG, CUN-HUI. Sorted concave penalized regression	3069–3098
FITHIAN, WILLIAM AND HUNG, KENNETH. Rank verification for exponential families	758–782
FORZANI, LILIANA AND COOK, R. DENNIS. Partial least squares prediction in high-dimensional regression	884–908
FOX, CHRISTOPHER, KÄUFL, ANDREAS, POULIOT, GUILLAUME AND DRTON, MATHIAS. The maximum likelihood threshold of a path diagram	1536–1553

FOX, CHRISTOPHER, WANG, Y. SAMUEL AND DRTON, MATHIAS. Computation of maximum likelihood estimates in cyclic structural equation models	663–690
FURRER, REINHARD, PORCU, EMILIO, BEVILACQUA, MORENO AND FAOUZI, TARIK. Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics	828–856
GAMBOA, FABRICE, HENRION, DIDIER, HESS, ROXANA, LASSERRE, JEAN-BERNARD AND DE CASTRO, YOHANN. Approximate optimal designs for multivariate polynomial regression	127–155
GAO, FUQING, CHEN, KANI, CAI, DINGJIAO, SUN, JIANGUO AND TONG, XINGWEI. Maximum likelihood estimation in transformed linear regression with nonnormal errors	1864–1892
GARCÍA-PORTUGUÉS, EDUARDO, FEBRERO-BANDE, MANUEL, GONZÁLEZ-MANTEIGA, WENCESLAO AND CUESTA-ALBERTOS, JUAN A. Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes .	439–467
GERBER, MATHIEU, CHOPIN, NICOLAS AND WHITELEY, NICK. Negative association, ordering and convergence of resampling methods	2236–2260
GHOSH, MALAY, CAO, XUAN AND KHARE, KSHITIJ. Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models	319–348
GOLDBERG, YAIR, KOSOROK, MICHAEL R. AND DASGUPTA, SAYAN. Feature elimination in kernel machines in moderately high dimensions	497–526
GONZÁLEZ-MANTEIGA, WENCESLAO, CUESTA-ALBERTOS, JUAN A., GARCÍA-PORTUGUÉS, EDUARDO AND FEBRERO-BANDE, MANUEL. Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes	439–467
GOOS, PETER, SCHOEN, ERIC D. AND EENDEBAK, PIETER T. A classification criterion for definitive screening designs	1179–1202
GREGORY, KARL, LAHIRI, S. N. AND DAS, DEBRAJ. Perturbation bootstrap in adaptive Lasso	2080–2116
GRØNNEBERG, STEFFEN AND HOLCBLAT, BENJAMIN. On partial-sum processes of ARMAX residuals	3216–3243
G’SELL, MAX, RINALDO, ALESSANDRO AND WASSERMAN, LARRY. Bootstrapping and sample splitting for high-dimensional, assumption-lean inference	3438–3469

GUNTUBOYINA, ADITYANAND, WEI, YUTING AND WAINWRIGHT, MARTIN J. The geometry of hypothesis testing over convex cones: Generalized likelihood ratio tests and minimax radii . . .	994–1024
GUO, JIANHUA, ZHU, HONGTU, ZHENG, SHURONG AND CHENG, GUANGHUI. Test for high-dimensional correlation matrices . .	2887–2921
HAN, QIYANG, WANG, TENG YAO, CHATTERJEE, SABYASACHI AND SAMWORTH, RICHARD J. Isotonic regression in general dimensions	2440–2471
HAN, QIYANG AND WELLNER, JON A. Convergence rates of least squares regression estimators with heavy-tailed errors	2286–2319
HAN, XU. Nonparametric screening under conditional strictly convex loss for ultrahigh dimensional sparse data	1995–2022
HANNIG, JAN AND WILLIAMS, JONATHAN P. Nonpenalized variable selection in high-dimensional linear model settings via generalized fiducial inference	1723–1753
HARCHAOUI, ZAID AND ENIKEEVA, FARIDA. High-dimensional change-point detection under sparse alternatives	2051–2079
HE, HERA Y., BASU, KINJAL, ZHAO, QINGYUAN AND OWEN, ART B. Permutation p -value approximation via generalized Stolarsky invariance	583–611
HECKEL, REINHARD, SHAH, NIHAR B., RAMCHANDRAN, KANNAN AND WAINWRIGHT, MARTIN J. Active ranking from pairwise comparisons and when parametric assumptions do not help	3099–3126
HENRION, DIDIER, HESS, ROXANA, LASSERRE, JEAN-BERNARD, DE CASTRO, YOHANN AND GAMBOA, FABRICE. Approximate optimal designs for multivariate polynomial regression . .	127–155
HESS, ROXANA, LASSERRE, JEAN-BERNARD, DE CASTRO, YOHANN, GAMBOA, FABRICE AND HENRION, DIDIER. Approximate optimal designs for multivariate polynomial regression	127–155
HOBERT, JAMES P. AND QIN, QIAN. Convergence complexity analysis of Albert and Chib’s algorithm for Bayesian probit regression	2320–2347
HOLCBLAT, BENJAMIN AND GRØNNEBERG, STEFFEN. On partial-sum processes of ARMAX residuals	3216–3243
HSU, HSIANG-LING, ING, CHING-KANG AND TONG, HOWELL. On model selection from a finite family of possibly misspecified time series models	1061–1087
HU, JIANG, LI, WEIMING, LIU, ZHI AND ZHOU, WANG. High-dimensional covariance matrices in elliptical distributions with application to spherical test	527–555

HU, JIANG, PAN, GUANGMING, ZHOU, WANG AND BAO, ZHI-GANG. Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case.....	612–640
HUCKEMANN, STEPHAN F. AND ELTZNER, BENJAMIN. A smeary central limit theorem for manifolds with application to high-dimensional spheres.....	3360–3381
HUNG, KENNETH AND FITHIAN, WILLIAM. Rank verification for exponential families.....	758–782
ING, CHING-KANG, TONG, HOWELL AND HSU, HSIANG-LING. On model selection from a finite family of possibly misspecified time series models.....	1061–1087
JOHNSTONE, IAIN M. AND FAN, ZHOU. Eigenvalue distributions of variance components estimators in high-dimensional random effects models.....	2855–2886
JORDAN, MICHAEL I., RAMDAS, AADITYA K., BARBER, RINA F. AND WAINWRIGHT, MARTIN J. A unified treatment of multiple testing with prior knowledge using the p-filter.....	2790–2821
KATO, KENGO AND CHEN, XIAOHUI. Randomized incomplete U -statistics in high dimensions.....	3127–3156
KÄUFL, ANDREAS, POULIOT, GUILLAUME, DRTON, MATHIAS AND FOX, CHRISTOPHER. The maximum likelihood threshold of a path diagram.....	1536–1553
KELLER-RESSEL, MARTIN, SCHILLING, RENÉ L. AND BÖTTCHER, BJÖRN. Distance multivariate: New dependence measures for random vectors.....	2757–2789
KHARE, KSHITIJ, GHOSH, MALAY AND CAO, XUAN. Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models.....	319–348
KOHLER, MICHAEL AND BAUER, BENEDIKT. On deep learning as a remedy for the curse of dimensionality in nonparametric regression.....	2261–2285
KOIKE, YUTA. Gaussian approximation of maxima of Wiener functionals and its application to high-frequency data.....	1663–1687
KONTOROVICH, ARYEH AND PINELIS, IOSIF. Exact lower bounds for the agnostic probably-approximately-correct (PAC) machine learning model.....	2822–2854
KOSOROK, MICHAEL R., DASGUPTA, SAYAN AND GOLDBERG, YAIR. Feature elimination in kernel machines in moderately high dimensions.....	497–526
LAHIRI, S. N., DAS, DEBRAJ AND GREGORY, KARL. Perturbation bootstrap in adaptive Lasso.....	2080–2116

LAM, CLIFFORD, YAO, JIANFENG, YAO, QIWEI AND LI, ZENG. On testing for high-dimensional white noise	3382–3412
LASSERRE, JEAN-BERNARD, DE CASTRO, YOHANN, GAMBOA, FABRICE, HENRION, DIDIER AND HESS, ROXANA. Approx- imate optimal designs for multivariate polynomial regression ..	127–155
LAURITZEN, STEFFEN, UHLER, CAROLINE AND ZWIERNIK, PIOTR. Maximum likelihood estimation in Gaussian models under total positivity.....	1835–1863
LECUÉ, GUILLAUME, ALQUIER, PIERRE AND COTTET, VINCENT. Estimation bounds and sharp oracle inequalities of regularized procedures with Lipschitz loss functions	2117–2144
LEE, DONALD K. K., NEGAHBAN, SAHAND N. AND CHEN, NINGYUAN. Super-resolution estimation of cyclic arrival rates	1754–1775
LEE, JAEYONG, LIN, LIZHEN AND LEE, KYOUNGJAE. Minimax posterior convergence rates and model selection consistency in high-dimensional DAG models based on sparse Cholesky fac- tors	3413–3437
LEE, KYOUNGJAE, LEE, JAEYONG AND LIN, LIZHEN. Minimax posterior convergence rates and model selection consistency in high-dimensional DAG models based on sparse Cholesky fac- tors	3413–3437
LEEB, HANNES, PÖTSCHER, BENEDIKT M. AND BACHOC, FRANÇOIS. Valid confidence intervals for post-model- selection predictors	1475–1504
LEEB, HANNES AND STEINBERGER, LUKAS. Prediction when fit- ting simple models to high-dimensional data	1408–1442
LEPSKI, O. V. AND WILLER, T. Oracle inequalities and adaptive estimation in the convolution structure density model	233–287
LEVRARD, CLÉMENT AND AAMARI, EDDIE. Nonasymptotic rates for manifold, tangent space and curvature estimation.....	177–204
LI, JIA, LIU, YUNXIAO AND XIU, DACHENG. Efficient estimation of integrated volatility functionals via multiscale jackknife	156–176
LI, JUN, ZHONG, PING-SHOU AND CHEN, SONG XI. Two-sample and ANOVA tests for high dimensional means	1443–1474
LI, RUNZE, SHI, CHENGCHUN, SONG, RUI AND CHEN, ZHAO. Linear hypothesis testing for high dimensional generalized lin- ear models.....	2671–2703
LI, RUNZE, ZHENG, SHURONG, CHEN, ZHAO AND CUI, HENGJIAN. Hypothesis testing on linear structures of high-dimensional co- variance matrix	3300–3334

LI, WEIMING, LIU, ZHI, ZHOU, WANG AND HU, JIANG. High-dimensional covariance matrices in elliptical distributions with application to spherical test.....	527–555
LI, ZENG, LAM, CLIFFORD, YAO, JIANFENG AND YAO, QIWEI. On testing for high-dimensional white noise	3382–3412
LIAN, HENG, ZHAO, KAIFENG AND LV, SHAOGAO. Projected spline estimation of the nonparametric function in high-dimensional partially linear models for massive data.....	2922–2949
LIJOI, ANTONIO, ORBANZ, PETER, PRÜNSTER, IGOR AND CAMERLENGHI, FEDERICO. Distribution theory for hierarchical processes	67–92
LIN, LIZHEN, LEE, KYOUNGJAE AND LEE, JAEYONG. Minimax posterior convergence rates and model selection consistency in high-dimensional DAG models based on sparse Cholesky factors	3413–3437
LIN, YI, MARTIN, RYAN AND YANG, MIN. On optimal designs for nonregular models.....	3335–3359
LIN, ZHENHUA AND YAO, FANG. Intrinsic Riemannian functional data analysis	3533–3577
LIU, HAN AND NEYKOV, MATEY. Property testing in high-dimensional Ising models	2472–2503
LIU, HAN, NEYKOV, MATEY AND LU, JUNWEI. Combinatorial inference for graphical models	795–827
LIU, WEIDONG, ZHANG, YICHEN AND CHEN, XI. Quantile regression under memory constraint.....	3244–3273
LIU, YUNXIAO, XIU, DACHENG AND LI, JIA. Efficient estimation of integrated volatility functionals via multiscale jackknife	156–176
LIU, ZHI, ZHOU, WANG, HU, JIANG AND LI, WEIMING. High-dimensional covariance matrices in elliptical distributions with application to spherical test.....	527–555
LOPES, MILES E. Estimating the algorithmic variance of randomized ensembles via the bootstrap	1088–1112
LU, JUNWEI, LIU, HAN AND NEYKOV, MATEY. Combinatorial inference for graphical models	795–827
LU, WENBIN, SHI, CHENGCHUN AND SONG, RUI. On testing conditional qualitative treatment effects.....	2348–2377
LUGOSI, GÁBOR AND MENDELSON, SHAHAR. Sub-Gaussian estimators of the mean of a random vector	783–794
LV, SHAOGAO, LIAN, HENG AND ZHAO, KAIFENG. Projected spline estimation of the nonparametric function in high-dimensional partially linear models for massive data.....	2922–2949

MA, CONG, WANG, KAIZHENG, CHEN, YUXIN AND FAN, JIANQING. Spectral method and regularized MLE are both optimal for top- K ranking	2204–2235
MA, JING, ZHANG, LINJUN AND CAI, T. TONY. CHIME: Clustering of high-dimensional Gaussian mixtures with EM algorithm and its optimality	1234–1267
MA, SHUJIE, ZHU, LIPING, ZHANG, ZHIWEI, TSAI, CHIH-LING AND CARROLL, RAYMOND J. A robust and efficient approach to causal inference based on sparse sufficient dimension reduction	1505–1535
MARTIN, RYAN, YANG, MIN AND LIN, YI. On optimal designs for nonregular models	3335–3359
MASSAM, HÉLÈNE, WANG, NANWEI AND RAUH, JOHANNES. Approximating faces of marginal polytopes in discrete hierarchical models	1203–1233
McKEAGUE, IAN W. AND CHANG, HSIN-WEN. Nonparametric testing for multiple survival functions with noninferiority margins	205–232
MEINSHAUSEN, NICOLAI, ROTHENHÄUSLER, DOMINIK AND BÜHLMANN, PETER. Causal Dantzig: Fast inference in linear structural equation models with hidden variables under additive interventions	1688–1722
MENDELSON, SHAHAR AND LUGOSI, GÁBOR. Sub-Gaussian estimators of the mean of a random vector	783–794
MONARD, FRANÇOIS, NICKL, RICHARD AND PATERNAIN, GABRIEL P. Efficient nonparametric Bayesian inference for X -ray transforms	1113–1147
MÜLLER, HANS-GEORG AND PETERSEN, ALEXANDER. Fréchet regression for random objects with Euclidean predictors	691–719
NAN, BIN AND SHU, HAI. Estimation of large covariance and precision matrices from temporally dependent observations	1321–1350
NEGAHBAN, SAHAND N., CHEN, NINGYUAN AND LEE, DONALD K. K. Super-resolution estimation of cyclic arrival rates	1754–1775
NEYKOV, MATEY AND LIU, HAN. Property testing in high-dimensional Ising models	2472–2503
NEYKOV, MATEY, LU, JUNWEI AND LIU, HAN. Combinatorial inference for graphical models	795–827
NICKL, RICHARD, PATERNAIN, GABRIEL P. AND MONARD, FRANÇOIS. Efficient nonparametric Bayesian inference for X -ray transforms	1113–1147

NORDMAN, DANIEL J., TEWES, JOHANNES AND POLITIS, DIMITRIS N. Convolved subsampling estimation with applications to block bootstrap	468–496
ORBANZ, PETER, PRÜNSTER, IGOR, CAMERLENGHI, FEDERICO AND LIJOI, ANTONIO. Distribution theory for hierarchical processes	67–92
OWEN, ART B., HE, HERA Y., BASU, KINJAL AND ZHAO, QINGYUAN. Permutation p -value approximation via generalized Stolarsky invariance	583–611
PAN, GUANGMING, ZHOU, WANG, BAO, ZHIGANG AND HU, JIANG. Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case	612–640
PANARETOS, VICTOR M. AND DESCARY, MARIE-HÉLÈNE. Functional data analysis by matrix completion	1–38
PAROLYA, NESTOR, BODNAR, TARAS AND DETTE, HOLGER. Testing for independence of large dimensional vectors	2977–3008
PATERNAIN, GABRIEL P., MONARD, FRANÇOIS AND NICKL, RICHARD. Efficient nonparametric Bayesian inference for X -ray transforms	1113–1147
PATI, DEBDEEP, YANG, YUN AND BHATTACHARYA, ANIRBAN. Bayesian fractional posteriors	39–66
PATSCHKOWSKI, TIM AND ROHDE, ANGELIKA. Locally adaptive confidence bands	349–381
PENSKY, MARIANNA. Dynamic network models and graphon estimation	2378–2403
PEPELYSHEV, ANDREY, ZHIGLJAVSKY, ANATOLY AND DETTE, HOLGER. The BLUE in continuous-time regression models with correlated errors	1928–1959
PETERSEN, ALEXANDER AND MÜLLER, HANS-GEORG. Fréchet regression for random objects with Euclidean predictors	691–719
PINELIS, IOSIF AND KONTOROVICH, ARYEH. Exact lower bounds for the agnostic probably-approximately-correct (PAC) machine learning model	2822–2854
POLITIS, DIMITRIS N., NORDMAN, DANIEL J. AND TEWES, JOHANNES. Convolved subsampling estimation with applications to block bootstrap	468–496
PORCU, EMILIO, BEVILACQUA, MORENO, FAOUZI, TARIK AND FURRER, REINHARD. Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics	828–856

PÖTSCHER, BENEDIKT M., BACHOC, FRANÇOIS AND LEEB, HANNES. Valid confidence intervals for post-model-selection predictors	1475–1504
POULIOT, GUILLAUME, DRTON, MATHIAS, FOX, CHRISTOPHER AND KÄUFL, ANDREAS. The maximum likelihood threshold of a path diagram	1536–1553
PRIEBE, CAREY E., CAPE, JOSHUA AND TANG, MINH. The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics	2405–2439
PRÜNSTER, IGOR, CAMERLENGHI, FEDERICO, LIJOI, ANTONIO AND ORBANZ, PETER. Distribution theory for hierarchical processes	67–92
QIN, LIKUAN AND TODOROV, VIKTOR. Nonparametric implied Lévy densities	1025–1060
QIN, QIAN AND HOBERT, JAMES P. Convergence complexity analysis of Albert and Chib’s algorithm for Bayesian probit regression	2320–2347
RAMCHANDRAN, KANNAN, WAINWRIGHT, MARTIN J., HECKEL, REINHARD AND SHAH, NIHAR B. Active ranking from pairwise comparisons and when parametric assumptions do not help	3099–3126
RAMDAS, AADITYA K., BARBER, RINA F., WAINWRIGHT, MARTIN J. AND JORDAN, MICHAEL I. A unified treatment of multiple testing with prior knowledge using the p-filter	2790–2821
RASKUTTI, GARVESH, YUAN, MING AND CHEN, HAN. Convex regularization for high-dimensional multiresponse tensor regression	1554–1584
RAUH, JOHANNES, MASSAM, HÉLÈNE AND WANG, NANWEI. Approximating faces of marginal polytopes in discrete hierarchical models	1203–1233
RICHTER, STEFAN AND DAHLHAUS, RAINER. Cross validation for locally stationary processes	2145–2173
RIGOLLET, PHILIPPE, SRIVASTAVA, PIYUSH AND BERTHET, QUENTIN. Exact recovery in the Ising blockmodel	1805–1834
RINALDO, ALESSANDRO, WASSERMAN, LARRY AND G’SSELL, MAX. Bootstrapping and sample splitting for high-dimensional, assumption-lean inference	3438–3469
ROBERTS, GARETH, BIERKENS, JORIS AND FEARNHEAD, PAUL. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data	1288–1320
ROHDE, ANGELIKA AND PATSCHKOWSKI, TIM. Locally adaptive confidence bands	349–381

ROHE, KARL. A critical threshold for design effects in network sampling	556–582
ROTHENHÄUSLER, DOMINIK, BÜHLMANN, PETER AND MEINSHAUSEN, NICOLAI. Causal Dantzig: Fast inference in linear structural equation models with hidden variables under additive interventions	1688–1722
ROY, DANIEL M. AND VEITCH, VICTOR. Sampling and estimation for (sparse) exchangeable graphs	3274–3299
SADHANALA, VEERANJANEYULU AND TIBSHIRANI, RYAN J. Additive models with trend filtering	3032–3068
SAEGUSA, TAKUMI. Large sample theory for merged data from multiple sources	1585–1615
SAMWORTH, RICHARD J., HAN, QIYANG, WANG, TENG YAO AND CHATTERJEE, SABYASACHI. Isotonic regression in general dimensions	2440–2471
SAMWORTH, RICHARD J., YUAN, MING AND BERRETT, THOMAS B. Efficient multivariate entropy estimation via k -nearest neighbour distances	288–318
SCHILLING, RENÉ L., BÖTTCHER, BJÖRN AND KELLER-RESSEL, MARTIN. Distance multivariate: New dependence measures for random vectors	2757–2789
SCHOEN, ERIC D., EENDEBAK, PIETER T. AND GOOS, PETER. A classification criterion for definitive screening designs	1179–1202
SCOTT, CLAYTON D. AND VANDERMEULEN, ROBERT A. An operator theoretic approach to nonparametric mixture models	2704–2733
SEN, BODHISATTVA, BANERJEE, MOULINATH AND DUROT, CÉCILE. Divide and conquer in nonstandard problems and the super-efficiency phenomenon	720–757
SHAH, NIHAR B., RAMCHANDRAN, KANNAN, WAINWRIGHT, MARTIN J. AND HECKEL, REINHARD. Active ranking from pairwise comparisons and when parametric assumptions do not help	3099–3126
SHI, CHENGCHUN, SONG, RUI, CHEN, ZHAO AND LI, RUNZE. Linear hypothesis testing for high dimensional generalized linear models	2671–2703
SHI, CHENGCHUN, SONG, RUI AND LU, WENBIN. On testing conditional qualitative treatment effects	2348–2377
SHU, HAI AND NAN, BIN. Estimation of large covariance and precision matrices from temporally dependent observations	1321–1350
SONG, RUI, CHEN, ZHAO, LI, RUNZE AND SHI, CHENGCHUN. Linear hypothesis testing for high dimensional generalized linear models	2671–2703

SONG, RUI, LU, WENBIN AND SHI, CHENGCHUN. On testing conditional qualitative treatment effects	2348–2377
SONG, YANGLEI AND FELLOURIS, GEORGIOS. Sequential multiple testing with generalized error control: An asymptotic optimality theory	1776–1803
SPOKOINY, VLADIMIR AND WILLRICH, NIKLAS. Bootstrap tuning in Gaussian ordered model selection	1351–1380
SRIVASTAVA, PIYUSH, BERTHET, QUENTIN AND RIGOLLET, PHILIPPE. Exact recovery in the Ising blockmodel	1805–1834
STEINBERGER, LUKAS AND LEEB, HANNES. Prediction when fitting simple models to high-dimensional data	1408–1442
SUN, FASHENG, WANG, YAPING AND XU, HONGQUAN. Uniform projection designs	641–661
SUN, JIANGUO, TONG, XINGWEI, GAO, FUQING, CHEN, KANI AND CAI, DINGJIAO. Maximum likelihood estimation in transformed linear regression with nonnormal errors	1864–1892
TAN, FALONG AND ZHU, LIXING. Adaptive-to-model checking for regressions with diverging number of predictors	1960–1994
TAN, ZHIQIANG AND ZHANG, CUN-HUI. Doubly penalized estimation in additive regression with high-dimensional data	2567–2600
TANG, MINH, PRIEBE, CAREY E. AND CAPE, JOSHUA. The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics	2405–2439
TEWES, JOHANNES, POLITIS, DIMITRIS N. AND NORDMAN, DANIEL J. Convolved subsampling estimation with applications to block bootstrap	468–496
TIBSHIRANI, JULIE, WAGER, STEFAN AND ATHEY, SUSAN. Generalized random forests	1148–1178
TIBSHIRANI, RYAN J. AND SADHANALA, VEERANJANEYULU. Additive models with trend filtering	3032–3068
TODOROV, VIKTOR AND QIN, LIKUAN. Nonparametric implied Lévy densities	1025–1060
TONG, HOWELL, HSU, HSIANG-LING AND ING, CHING-KANG. On model selection from a finite family of possibly misspecified time series models	1061–1087
TONG, XINGWEI, GAO, FUQING, CHEN, KANI, CAI, DINGJIAO AND SUN, JIANGUO. Maximum likelihood estimation in transformed linear regression with nonnormal errors	1864–1892
TRUQUET, LIONEL. Local stationarity and time-inhomogeneous Markov chains	2023–2050

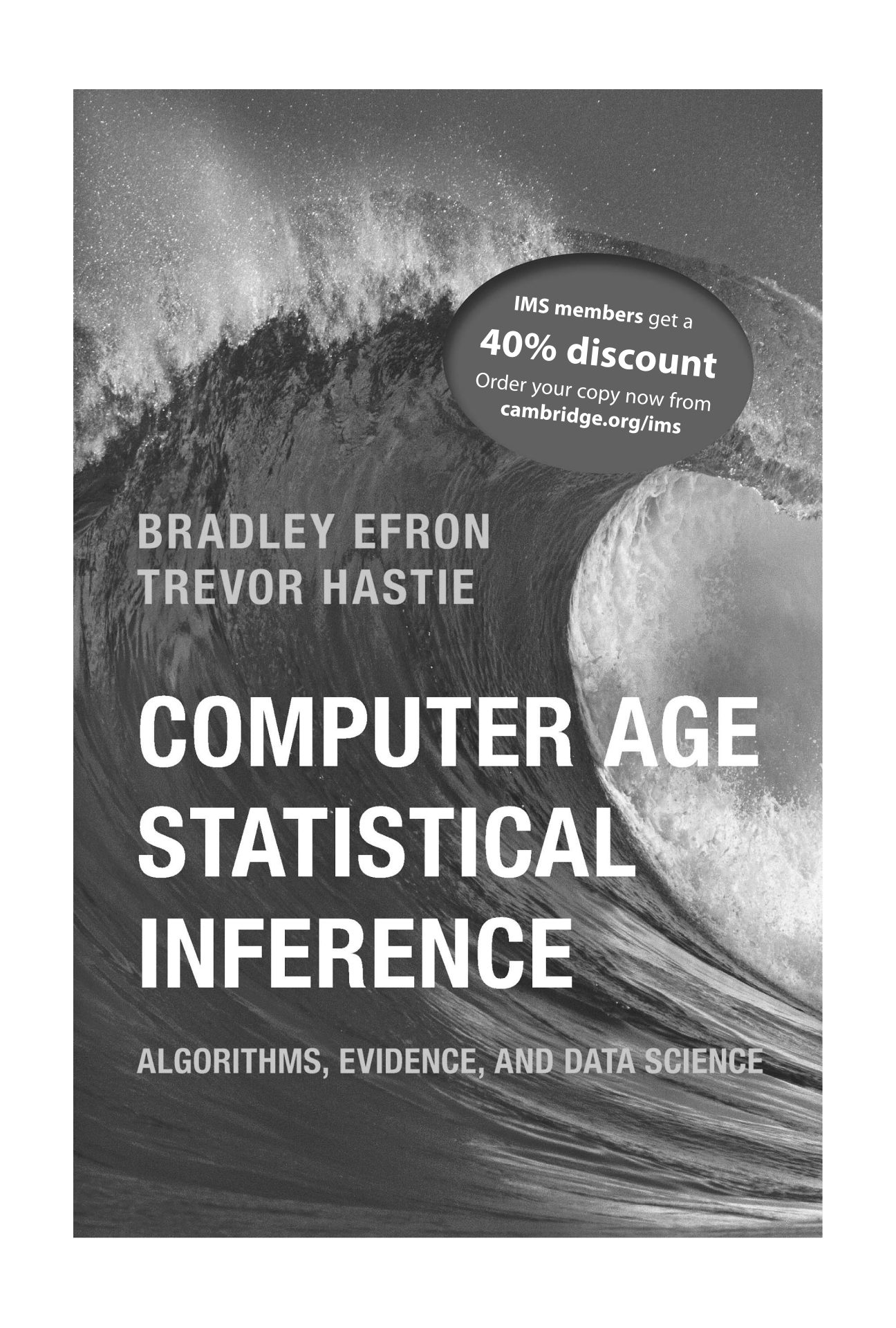
TSAI, CHIH-LING, CARROLL, RAYMOND J., MA, SHUIE, ZHU, LIPING AND ZHANG, ZHIWEI. A robust and efficient approach to causal inference based on sparse sufficient dimension reduction	1505–1535
UHLER, CAROLINE, ZWIERNIK, PIOTR AND LAURITZEN, STEFFEN. Maximum likelihood estimation in Gaussian models under total positivity	1835–1863
VAN DEN AKKER, RAMON, WERKER, BAS J. M. AND ZHOU, BO. Semiparametrically point-optimal hybrid rank tests for unit roots	2601–2638
VANDERMEULEN, ROBERT A. AND SCOTT, CLAYTON D. An operator theoretic approach to nonparametric mixture models	2704–2733
VEITCH, VICTOR AND ROY, DANIEL M. Sampling and estimation for (sparse) exchangeable graphs	3274–3299
VEMURI, BABA C. AND CHAKRABORTY, RUDRASI. Statistics on the Stiefel manifold: Theory and applications	415–438
VERZELEN, NICOLAS AND CARPENTIER, ALEXANDRA. Adaptive estimation of the sparsity in the Gaussian vector model	93–126
VOLGUSHEV, STANISLAV, CHAO, SHIH-KANG AND CHENG, GUANG. Distributed inference for quantile regression processes	1634–1662
WAGER, STEFAN, ATHEY, SUSAN AND TIBSHIRANI, JULIE. Generalized random forests	1148–1178
WAINWRIGHT, MARTIN J., GUNTUBOYINA, ADITYANAND AND WEI, YUTING. The geometry of hypothesis testing over convex cones: Generalized likelihood ratio tests and minimax radii	994–1024
WAINWRIGHT, MARTIN J., HECKEL, REINHARD, SHAH, NIHAR B. AND RAMCHANDRAN, KANNAN. Active ranking from pairwise comparisons and when parametric assumptions do not help	3099–3126
WAINWRIGHT, MARTIN J., JORDAN, MICHAEL I., RAMDAS, AADITYA K. AND BARBER, RINA F. A unified treatment of multiple testing with prior knowledge using the p-filter	2790–2821
WANG, DONG, WANG, KAIZHENG, ZHU, ZIWEI AND FAN, JIANQING. Distributed estimation of principal eigenspaces	3009–3031
WANG, KAIZHENG, CHEN, YUXIN, FAN, JIANQING AND MA, CONG. Spectral method and regularized MLE are both optimal for top- K ranking	2204–2235
WANG, KAIZHENG, ZHU, ZIWEI, FAN, JIANQING AND WANG, DONG. Distributed estimation of principal eigenspaces	3009–3031
WANG, NANWEI, RAUH, JOHANNES AND MASSAM, HÉLÈNE. Approximating faces of marginal polytopes in discrete hierarchical models	1203–1233

WANG, TENG YAO, CHATTERJEE, SABYASACHI, SAMWORTH, RICHARD J. AND HAN, QIYANG. Isotonic regression in general dimensions	2440–2471
WANG, Y. SAMUEL, DRTON, MATHIAS AND FOX, CHRISTOPHER. Computation of maximum likelihood estimates in cyclic structural equation models	663–690
WANG, YAPING, XU, HONGQUAN AND SUN, FASHENG. Uniform projection designs	641–661
WASSERMAN, LARRY AND BALAKRISHNAN, SIVARAMAN. Hypothesis testing for densities and high-dimensional multinomials: Sharp local minimax rates	1893–1927
WASSERMAN, LARRY, G'SELL, MAX AND RINALDO, ALESSANDRO. Bootstrapping and sample splitting for high-dimensional, assumption-lean inference	3438–3469
WEGKAMP, MARTEN H. AND BING, XIN. Adaptive estimation of the rank of the coefficient matrix in high-dimensional multivariate response regression models	3157–3184
WEI, YUTING, WAINWRIGHT, MARTIN J. AND GUNTUBOYINA, ADITYANAND. The geometry of hypothesis testing over convex cones: Generalized likelihood ratio tests and minimax radii	994–1024
WELLNER, JON A. AND DOSS, CHARLES R. Inference for the mode of a log-concave density	2950–2976
WELLNER, JON A. AND HAN, QIYANG. Convergence rates of least squares regression estimators with heavy-tailed errors	2286–2319
WELLS, MARTIN T. AND CHÉTELAT, DIDIER. The middle-scale asymptotics of Wishart matrices	2639–2670
WERKER, BAS J. M., ZHOU, BO AND VAN DEN AKKER, RAMON. Semiparametrically point-optimal hybrid rank tests for unit roots	2601–2638
WHITELEY, NICK, GERBER, MATHIEU AND CHOPIN, NICOLAS. Negative association, ordering and convergence of resampling methods	2236–2260
WILLER, T. AND LEPSKI, O. V. Oracle inequalities and adaptive estimation in the convolution structure density model	233–287
WILLIAMS, JONATHAN P. AND HANNIG, JAN. Nonpenalized variable selection in high-dimensional linear model settings via generalized fiducial inference	1723–1753
WILLRICH, NIKLAS AND SPOKOINY, VLADIMIR. Bootstrap tuning in Gaussian ordered model selection	1351–1380
WU, WEICHI AND DETTE, HOLGER. Detecting relevant changes in the mean of nonstationary processes—A mass excess approach	3578–3608

WU, YIHONG AND YANG, PENGKUN. Chebyshev polynomials, moment matching, and optimal estimation of the unseen	857–883
XIU, DACHENG, LI, JIA AND LIU, YUNXIAO. Efficient estimation of integrated volatility functionals via multiscale jackknife	156–176
XU, HONGQUAN, SUN, FASHENG AND WANG, YAPING. Uniform projection designs	641–661
YANG, MIN, LIN, YI AND MARTIN, RYAN. On optimal designs for nonregular models	3335–3359
YANG, PENGKUN AND WU, YIHONG. Chebyshev polynomials, moment matching, and optimal estimation of the unseen	857–883
YANG, YUN, BHATTACHARYA, ANIRBAN AND PATI, DEBDEEP. Bayesian fractional posteriors	39–66
YAO, FANG AND LIN, ZHENHUA. Intrinsic Riemannian functional data analysis	3533–3577
YAO, JIANFENG, YAO, QIWEI, LI, ZENG AND LAM, CLIFFORD. On testing for high-dimensional white noise	3382–3412
YAO, QIWEI, LI, ZENG, LAM, CLIFFORD AND YAO, JIANFENG. On testing for high-dimensional white noise	3382–3412
YUAN, MING, BERRETT, THOMAS B. AND SAMWORTH, RICHARD J. Efficient multivariate entropy estimation via k -nearest neighbour distances	288–318
YUAN, MING, CHEN, HAN AND RASKUTTI, GARVESH. Convex regularization for high-dimensional multiresponse tensor regression	1554–1584
ZHANG, ANRU. Cross: Efficient low-rank tensor completion	936–964
ZHANG, ANRU, BROWN, LAWRENCE D. AND CAI, T. TONY. Semi-supervised inference: General theory and estimation of means	2538–2566
ZHANG, CUN-HUI AND FENG, LONG. Sorted concave penalized regression	3069–3098
ZHANG, CUN-HUI AND TAN, ZHIQIANG. Doubly penalized estimation in additive regression with high-dimensional data	2567–2600
ZHANG, LINJUN, CAI, T. TONY AND MA, JING. CHIME: Clustering of high-dimensional Gaussian mixtures with EM algorithm and its optimality	1234–1267
ZHANG, YICHEN, CHEN, XI AND LIU, WEIDONG. Quantile regression under memory constraint	3244–3273
ZHANG, ZHIWEI, TSAI, CHIH-LING, CARROLL, RAYMOND J., MA, SHUJIE AND ZHU, LIPING. A robust and efficient approach to causal inference based on sparse sufficient dimension reduction	1505–1535

ZHAO, KAIFENG, LV, SHAOGAO AND LIAN, HENG. Projected spline estimation of the nonparametric function in high-dimensional partially linear models for massive data	2922–2949
ZHAO, QINGYUAN. Covariate balancing propensity score by tailored loss functions	965–993
ZHAO, QINGYUAN, OWEN, ART B., HE, HERA Y. AND BASU, KINJAL. Permutation p -value approximation via generalized Stolarsky invariance	583–611
ZHENG, SHURONG, CHEN, ZHAO, CUI, HENGJIAN AND LI, RUNZE. Hypothesis testing on linear structures of high-dimensional covariance matrix	3300–3334
ZHENG, SHURONG, CHENG, GUANGHUI, GUO, JIANHUA AND ZHU, HONGTU. Test for high-dimensional correlation matrices	2887–2921
ZHIGLJAVSKY, ANATOLY, DETTE, HOLGER AND PEPELYSHEV, ANDREY. The BLUE in continuous-time regression models with correlated errors	1928–1959
ZHONG, PING-SHOU, CHEN, SONG XI AND LI, JUN. Two-sample and ANOVA tests for high dimensional means	1443–1474
ZHOU, BO, VAN DEN AKKER, RAMON AND WERKER, BAS J. M. Semiparametrically point-optimal hybrid rank tests for unit roots	2601–2638
ZHOU, WANG, BAO, ZHIGANG, HU, JIANG AND PAN, GUANGMING. Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case	612–640
ZHOU, WANG, HU, JIANG, LI, WEIMING AND LIU, ZHI. High-dimensional covariance matrices in elliptical distributions with application to spherical test	527–555
ZHU, HONGTU, ZHENG, SHURONG, CHENG, GUANGHUI AND GUO, JIANHUA. Test for high-dimensional correlation matrices	2887–2921
ZHU, KE. Statistical inference for autoregressive models under heteroscedasticity of unknown form	3185–3215
ZHU, LIPING, ZHANG, ZHIWEI, TSAI, CHIH-LING, CARROLL, RAYMOND J. AND MA, SHUJIE. A robust and efficient approach to causal inference based on sparse sufficient dimension reduction	1505–1535
ZHU, LIXING AND TAN, FALONG. Adaptive-to-model checking for regressions with diverging number of predictors	1960–1994
ZHU, ZIWEI, FAN, JIANQING, WANG, DONG AND WANG, KAIZHENG. Distributed estimation of principal eigenspaces	3009–3031

ZWIERNIK, PIOTR, LAURITZEN, STEFFEN AND UHLER, CAROLINE. Maximum likelihood estimation in Gaussian models under total positivity	1835–1863
--	-----------



IMS members get a
40% discount

Order your copy now from
cambridge.org/ims

BRADLEY EFRON
TREVOR HASTIE

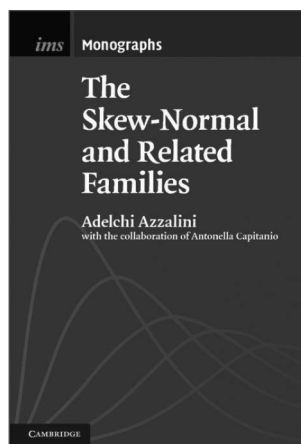
COMPUTER AGE STATISTICAL INFERENCE

ALGORITHMS, EVIDENCE, AND DATA SCIENCE



The Institute of Mathematical Statistics presents

IMS MONOGRAPHS



The Skew-Normal and Related Families

Adelchi Azzalini

in collaboration with Antonella Capitanio

Interest in the skew-normal and related families of distributions has grown enormously over recent years, as theory has advanced, challenges of data have grown, and computational tools have made substantial progress. This comprehensive treatment, blending theory and practice, will be the standard resource for statisticians and applied researchers. Assuming only basic knowledge of (non-measure-theoretic) probability and statistical inference, the book is accessible to the wide range of researchers who use statistical modelling techniques. Guiding readers through the main concepts and results, it covers both the probability and the statistics sides of the subject, in the univariate and multivariate settings. The theoretical development is complemented by numerous illustrations and applications to a range of fields including quantitative finance, medical statistics, environmental risk studies, and industrial and business efficiency.

The author's freely available R package *sn*, available from CRAN, equips readers to put the methods into action with their own data.

**IMS member? Claim
your 40% discount:
www.cambridge.org/ims**

**Hardback price
US\$48.00
(non-member price
\$80.00)**

Cambridge University Press, in conjunction with the Institute of Mathematical Statistics, established the IMS Monographs and IMS Textbooks series of high-quality books. The Series Editors are Xiao-Li Meng, Susan Holmes, Ben Hambly, D. R. Cox and Alan Agresti.