ISSN 0090-5364 (print)
ISSN 2168-8966 (online)

THE ANNALS *of* STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles

Vol. 54, No. 4—August 2026

THE ANNALS OF STATISTICS Vol. 54, No. 4, pp. 1655–2208 August 2026

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.

The Annals of Statistics **[ISSN 0090-5364 (print); ISSN 2168-8966 (online)],** Volume 54, Number 4, August 2026. Published bimonthly by the Institute of Mathematical Statistics, 9760 Smith Road, Waite Hill, OH 44094, 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, PO Box 729, Middletown, MD 21769, USA.

The Annals of Statistics
2026, Vol. 54, No. 4, 1655–1681
https://doi.org/10.1214/25-AOS2612

A NONASYMPTOTIC DISTRIBUTIONAL THEORY OF APPROXIMATE MESSAGE PASSING FOR SPARSE AND ROBUST REGRESSION

BY GEN LI[1,a] AND YUTING WEI[2,b]

[1]*Department of Statistics and Data Science, Chinese University of Hong Kong,* [a]*genli@cuhk.edu.hk*
[2]*Department of Statistics and Data Science, the Wharton School, University of Pennsylvania,* [b]*ytwei@wharton.upenn.edu*

Characterizing the distribution of high-dimensional statistical estimators is a challenging task, due to the breakdown of classical asymptotic theory in high dimension. This paper makes progress towards this by developing nonasymptotic distributional characterizations for approximate message passing (AMP)—a family of iterative algorithms that prove effective as both fast estimators and powerful theoretical machinery—for both sparse and robust regression. Prior AMP theory, which focused on high-dimensional asymptotics for the most part, failed to describe the behavior of AMP when the number of iterations exceeds $o(\log n/\log\log n)$ (with n the sample size). Our AMP theory is further leveraged to derive a refined rate for the Lasso estimator, built upon a nonasymptotic convergence of AMP to Lasso. To the best of our knowledge, this is the first result of its kind, breaking the rate barrier resulting from the widely adopted Gaussian min-max theorem. Our results establish approximate accuracy of Gaussian approximation of the AMP iterates, which improves upon prior results and implies enhanced distributional characterizations for both optimally tuned Lasso and robust M-estimator.

REFERENCES

ADOMAITYTE, U., DEFILIPPIS, L., LOUREIRO, B. and SICURO, G. (2024). High-dimensional robust regression under heavy-tailed data: Asymptotics and universality. *J. Stat. Mech. Theory Exp.* **2024** Paper No. 114002, 36 pp. MR4847056 https://doi.org/10.1088/1742-5468/ad65e6

AMELUNXEN, D., LOTZ, M., MCCOY, M. B. and TROPP, J. A. (2014). Living on the edge: Phase transitions in convex programs with random data. *Inf. Inference* **3** 224–294. MR3311453 https://doi.org/10.1093/imaiai/iau005

BAO, Z., HAN, Q. and XU, X. (2025). A leave-one-out approach to approximate message passing. *Ann. Appl. Probab.* **35** 2716–2766. MR4945090 https://doi.org/10.1214/25-AAP2186

BARBIER, J., KRZAKALA, F., MACRIS, N., MIOLANE, L. and ZDEBOROVÁ, L. (2019). Optimal errors and phase transitions in high-dimensional generalized linear models. *Proc. Natl. Acad. Sci. USA* **116** 5451–5460. MR3939767 https://doi.org/10.1073/pnas.1802705116

BAYATI, M., LELARGE, M. and MONTANARI, A. (2015). Universality in polytope phase transitions and message passing algorithms. *Ann. Appl. Probab.* **25** 753–822. MR3313755 https://doi.org/10.1214/14-AAP1010

BAYATI, M. and MONTANARI, A. (2011). The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Trans. Inf. Theory* **57** 764–785. MR2810285 https://doi.org/10.1109/TIT.2010.2094817

BAYATI, M. and MONTANARI, A. (2012). The LASSO risk for Gaussian matrices. *IEEE Trans. Inf. Theory* **58** 1997–2017. MR2951312 https://doi.org/10.1109/TIT.2011.2174612

BELLEC, P. C. and KORIYAMA, T. (2025). Error estimation and adaptive tuning for unregularized robust M-estimator. *J. Mach. Learn. Res.* **26** Paper No. [16], 40 pp. MR4870572

BELLEC, P. C., SHEN, Y. and ZHANG, C.-H. (2022). Asymptotic normality of robust M-estimators with convex penalty. *Electron. J. Stat.* **16** 5591–5622. MR4497866 https://doi.org/10.1214/22-ejs2065

BELLEC, P. C. and ZHANG, C.-H. (2022). De-biasing the lasso with degrees-of-freedom adjustment. *Bernoulli* **28** 713–743. MR4389062 https://doi.org/10.3150/21-BEJ1348

MSC2020 subject classifications. Primary 62E17; secondary 62J07, 62E10.

Key words and phrases. Distributional theory, linear models, approximate message passing, nonasymptotic analysis, sparse regression, robust regression.

BELLEC, P. C. and ZHANG, C.-H. (2023). Debiasing convex regularized estimators and interval estimation in linear models. *Ann. Statist.* **51** 391–436. MR4600987 https://doi.org/10.1214/22-aos2243

BICKEL, P. J. (1975). One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc.* **70** 428–434. MR0386168

BOGDAN, M., VAN DEN BERG, E., SABATTI, C., SU, W. and CANDES, E. J. (2015). SLOPE—adaptive variable selection via convex optimization. *Ann. Appl. Stat.* **9** 1103–1140. MR3418717 https://doi.org/10.1214/15-AOAS842

BOLTHAUSEN, E. (2009). On the high-temperature phase of the Sherrington–Kirkpatrick model. In *Seminar at EURANDOM, Eindhoven*.

BU, Z., KLUSOWSKI, J. M., RUSH, C. and SU, W. J. (2021). Algorithmic analysis and statistical estimation of SLOPE via approximate message passing. *IEEE Trans. Inf. Theory* **67** 506–537. MR4231969 https://doi.org/10.1109/TIT.2020.3025272

BÜHLMANN, P. and VAN DE GEER, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications. Springer Series in Statistics*. Springer, Heidelberg. MR2807761 https://doi.org/10.1007/978-3-642-20192-9

CADEMARTORI, C. and RUSH, C. (2024). A non-asymptotic analysis of generalized vector approximate message passing algorithms with rotationally invariant designs. *IEEE Trans. Inf. Theory* **70** 5811–5856. MR4773211 https://doi.org/10.1109/tit.2024.3396472

CAI, C., POOR, H. V. and CHEN, Y. (2023). Uncertainty quantification for nonconvex tensor completion: Confidence intervals, heteroscedasticity and optimality. *IEEE Trans. Inf. Theory* **69** 407–452. MR4544966 https://doi.org/10.1109/tit.2022.3205781

CANDES, E. and TAO, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n.

CANDES, E. J., ROMBERG, J. K. and TAO, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.* **59** 1207–1223. MR2230846 https://doi.org/10.1002/cpa.20124

CELENTANO, M., FAN, Z., LIN, L. and MEI, S. (2023). Mean-field variational inference with the TAP free energy: Geometric and statistical properties in linear models. Preprint. Available at arXiv:2311.08442.

CELENTANO, M., FAN, Z. and MEI, S. (2023). Local convexity of the TAP free energy and AMP convergence for $\mathbb{Z}_2$-synchronization. *Ann. Statist.* **51** 519–546. MR4600991 https://doi.org/10.1214/23-aos2257

CELENTANO, M. and MONTANARI, A. (2021). Cad: Debiasing the lasso with inaccurate covariate model. Preprint. Available at arXiv:2107.14172.

CELENTANO, M., MONTANARI, A. and WEI, Y. (2023). The Lasso with general Gaussian designs with applications to hypothesis testing. *Ann. Statist.* **51** 2194–2220. MR4678801 https://doi.org/10.1214/23-aos2327

CHEN, S. S., DONOHO, D. L. and SAUNDERS, M. A. (2001). Atomic decomposition by basis pursuit. *SIAM Rev.* **43** 129–159. MR1854649 https://doi.org/10.1137/S003614450037906X

CHEN, W.-K. and LAM, W.-K. (2021). Universality of approximate message passing algorithms. *Electron. J. Probab.* **26** Paper No. 36, 44 pp. MR4235487 https://doi.org/10.1214/21-EJP604

CHEN, Y., CHI, Y., FAN, J. and MA, C. (2019a). Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval. *Math. Program.* **176** 5–37. MR3960803 https://doi.org/10.1007/s10107-019-01363-6

CHEN, Y., FAN, J., MA, C. and YAN, Y. (2019b). Inference and uncertainty quantification for noisy matrix completion. *Proc. Natl. Acad. Sci. USA* **116** 22931–22937. MR4036123 https://doi.org/10.1073/pnas.1910053116

DALALYAN, A. and CHEN, Y. (2012). Fused sparsity and robust estimation for linear models with unknown variance. *Adv. Neural Inf. Process. Syst.* **25**.

DALALYAN, A. and THOMPSON, P. (2019). Outlier-robust estimation of a sparse linear model using ℓ_1-penalized Huber's m-estimator. *Adv. Neural Inf. Process. Syst.* **32**.

DESHPANDE, Y., ABBE, E. and MONTANARI, A. (2017). Asymptotic mutual information for the balanced binary stochastic block model. *Inf. Inference* **6** 125–170. MR3671474 https://doi.org/10.1093/imaiai/iaw017

DESHPANDE, Y. and MONTANARI, A. (2014). Information-theoretically optimal sparse PCA. In 2014 *IEEE International Symposium on Information Theory* 2197–2201. IEEE.

DONOHO, D. and MONTANARI, A. (2016). High dimensional robust M-estimation: Asymptotic variance via approximate message passing. *Probab. Theory Related Fields* **166** 935–969. MR3568043 https://doi.org/10.1007/s00440-015-0675-z

DONOHO, D. L. (2006). Compressed sensing. *IEEE Trans. Inf. Theory* **52** 1289–1306. MR2241189 https://doi.org/10.1109/TIT.2006.871582

DONOHO, D. L. and HUO, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory* **47** 2845–2862. MR1872845 https://doi.org/10.1109/18.959265

DONOHO, D. L., JAVANMARD, A. and MONTANARI, A. (2013). Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing. *IEEE Trans. Inf. Theory* **59** 7434–7464. MR3124654 https://doi.org/10.1109/TIT.2013.2274513

DONOHO, D. L., MALEKI, A. and MONTANARI, A. (2009). Message-passing algorithms for compressed sensing. *Proc. Natl. Acad. Sci. USA* **106** 18914–18919.

DONOHO, D. L. and MONTANARI, A. (2015). Variance breakdown of Huber (m)-estimators. $n/p \to m$ in $(1, \infty)$. Preprint. Available at arXiv:1503.02106.

DUDEJA, R., LU, Y. M. and SEN, S. (2023). Universality of approximate message passing with semirandom matrices. *Ann. Probab.* **51** 1616–1683. MR4642220 https://doi.org/10.1214/23-aop1628

EL KAROUI, N. (2013). Asymptotic behavior of unregularized and ridge-regularized high-dimensional robust regression estimators: Rigorous results. Preprint. Available at arXiv:1311.2445.

EL KAROUI, N. (2018). On the impact of predictor geometry on the performance on high-dimensional ridge-regularized generalized robust regression estimators. *Probab. Theory Related Fields* **170** 95–175. MR3748322 https://doi.org/10.1007/s00440-016-0754-9

EL KAROUI, N., BEAN, D., BICKEL, P. J., LIM, C. and YU, B. (2013). On robust regression with high-dimensional predictors. *Proc. Natl. Acad. Sci. USA* **110** 14557–14562.

FAN, J., LI, Q. and WANG, Y. (2014). Robust estimation of high-dimensional mean regression. Preprint. Available at arXiv:1410.2150.

FAN, J. and LI, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. MR1946581 https://doi.org/10.1198/016214501753382273

FAN, J., LI, R., ZHANG, C.-H. and ZOU, H. (2020). *Statistical Foundations of Data Science*. CRC Press, Boca Raton.

FAN, Z. (2022). Approximate message passing algorithms for rotationally invariant matrices. *Ann. Statist.* **50** 197–224. MR4382014 https://doi.org/10.1214/21-aos2101

FENG, O. Y., VENKATARAMANAN, R., RUSH, C. and SAMWORTH, R. J. (2022). A unifying tutorial on approximate message passing. *Found. Trends Mach. Learn.* **15** 335–536.

FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **222** 309–368.

HAMPEL, F. R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* **69** 383–393. MR0362657

HAN, Q. and SHEN, Y. (2023). Universality of regularized regression estimators in high dimensions. *Ann. Statist.* **51** 1799–1823. MR4658577 https://doi.org/10.1214/23-aos2309

HASTIE, T., TIBSHIRANI, R. and WAINWRIGHT, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations. Monographs on Statistics and Applied Probability* **143**. CRC Press, Boca Raton, FL. MR3616141

HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101. MR0161415 https://doi.org/10.1214/aoms/1177703732

HUBER, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* **1** 799–821. MR0356373

JAVANMARD, A. and MONTANARI, A. (2013). State evolution for general approximate message passing algorithms, with applications to spatial coupling. *Inf. Inference* **2** 115–144. MR3311445 https://doi.org/10.1093/imaiai/iat004

JAVANMARD, A. and MONTANARI, A. (2014). Hypothesis testing in high-dimensional regression under the Gaussian random design model: Asymptotic theory. *IEEE Trans. Inf. Theory* **60** 6522–6554. MR3265038 https://doi.org/10.1109/TIT.2014.2343629

JAVANMARD, A. and MONTANARI, A. (2018). Debiasing the Lasso: Optimal sample size for Gaussian designs. *Ann. Statist.* **46** 2593–2622. MR3851749 https://doi.org/10.1214/17-AOS1630

LASKA, J. N., DAVENPORT, M. A. and BARANIUK, R. G. (2009). Exact signal recovery from sparsely corrupted measurements through the pursuit of justice. In 2009 *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers* 1556–1560. IEEE.

LE CAM, L. (2012). *Asymptotic Methods in Statistical Decision Theory. Springer Series in Statistics*. Springer, New York. MR0856411 https://doi.org/10.1007/978-1-4612-4946-7

LEI, L., BICKEL, P. J. and EL KAROUI, N. (2018). Asymptotics for high dimensional regression M-estimates: Fixed design results. *Probab. Theory Related Fields* **172** 983–1079. MR3877551 https://doi.org/10.1007/s00440-017-0824-7

LI, G., FAN, W. and WEI, Y. (2023). Approximate message passing from random initialization with applications to $\mathbb{Z}_2$ synchronization. *Proc. Natl. Acad. Sci. USA* **120** Paper No. e2302930120, 7 pp. MR4637851 https://doi.org/10.1073/pnas.2302930120

LI, G. and WEI, Y. (2022). A non-asymptotic framework for approximate message passing in spiked models. Preprint. Available at arXiv:2208.03313.

LI, G. and WEI, Y. (2026). Supplement to "A nonasymptotic distributional theory of approximate message passing for sparse and robust regression." https://doi.org/10.1214/25-AOS2612SUPP

LI, Y., FAN, Z., SEN, S. and WU, Y. (2024). Random linear estimation with rotationally-invariant designs: Asymptotics at high temperature. *IEEE Trans. Inf. Theory* **70** 2118–2153. MR4709776 https://doi.org/10.1109/tit.2023.3321575

LI, Y. and WEI, Y. (2021). Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent. Preprint. Available at arXiv:2110.09502.

LOH, P.-L. (2017). Statistical consistency and asymptotic normality for high-dimensional robust M-estimators. *Ann. Statist.* **45** 866–896. MR3650403 https://doi.org/10.1214/16-AOS1471

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

MA, C., WANG, K., CHI, Y. and CHEN, Y. (2020). Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval, matrix completion, and blind deconvolution. *Found. Comput. Math.* **20** 451–632. MR4099988 https://doi.org/10.1007/s10208-019-09429-9

MA, Z. and NANDY, S. (2023). Community detection with contextual multilayer networks. *IEEE Trans. Inf. Theory* **69** 3203–3239. MR4581286 https://doi.org/10.1109/tit.2023.3238352

MALEKI, A., ANITORI, L., YANG, Z. and BARANIUK, R. G. (2013). Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP). *IEEE Trans. Inf. Theory* **59** 4290–4308. MR3071330 https://doi.org/10.1109/TIT.2013.2252232

MARONNA, R. A., MARTIN, R. D., YOHAI, V. J. and SALIBIÁN-BARRERA, M. (2019). *Robust Statistics: Theory and Methods (with R). Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR3839299

MEINSHAUSEN, N. and BÜHLMANN, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. MR2278363 https://doi.org/10.1214/009053606000000281

MIOLANE, L. and MONTANARI, A. (2021). The distribution of the Lasso: Uniform control over sparse balls and adaptive parameter tuning. *Ann. Statist.* **49** 2313–2335. MR4319252 https://doi.org/10.1214/20-aos2038

MITCHELL, T. J. and BEAUCHAMP, J. J. (1988). Bayesian variable selection in linear regression. *J. Amer. Statist. Assoc.* **83** 1023–1036. MR0997578

MONDELLI, M. and VENKATARAMANAN, R. (2021). PCA initialization for approximate message passing in rotationally invariant models. *Adv. Neural Inf. Process. Syst.* **34** 29616–29629.

MONDELLI, M. and VENKATARAMANAN, R. (2022). Approximate message passing with spectral initialization for generalized linear models. *J. Stat. Mech. Theory Exp.* **2022** Paper No. 114003, 41 pp. MR4535574 https://doi.org/10.1088/1742-5468/ac9828

MONTANARI, A. and VENKATARAMANAN, R. (2021). Estimation of low-rank matrices via approximate message passing. *Ann. Statist.* **49** 321–345. MR4206680 https://doi.org/10.1214/20-AOS1958

MONTANARI, A. and WU, Y. (2023). Posterior sampling from the spiked models via diffusion processes. Preprint. Available at arXiv:2304.11449.

NGUYEN, N. H. and TRAN, T. D. (2013). Robust Lasso with missing and grossly corrupted observations. *IEEE Trans. Inf. Theory* **59** 2036–2058. MR3043781 https://doi.org/10.1109/TIT.2012.2232347

RANGAN, S. (2011). Generalized approximate message passing for estimation with random linear mixing. In 2011 *IEEE International Symposium on Information Theory Proceedings* 2168–2172. IEEE.

RANGAN, S. and FLETCHER, A. K. (2012). Iterative estimation of constrained rank-one matrices in noise. In 2012 *IEEE International Symposium on Information Theory Proceedings* 1246–1250. IEEE.

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 https://doi.org/10.1214/14-AOS1286

ROUSSEEUW, P. J. and LEROY, A. M. (2005). *Robust Regression and Outlier Detection. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. Wiley, New York. MR0914792 https://doi.org/10.1002/0471725382

RUDELSON, M. and VERSHYNIN, R. (2008). On sparse reconstruction from Fourier and Gaussian measurements. *Comm. Pure Appl. Math.* **61** 1025–1045. MR2417886 https://doi.org/10.1002/cpa.20227

RUSH, C. and VENKATARAMANAN, R. (2018). Finite sample analysis of approximate message passing algorithms. *IEEE Trans. Inf. Theory* **64** 7264–7286. MR3876443 https://doi.org/10.1109/TIT.2018.2816681

SCHNITER, P. and RANGAN, S. (2015). Compressive phase retrieval via generalized approximate message passing. *IEEE Trans. Signal Process.* **63** 1043–1055. MR3311635 https://doi.org/10.1109/TSP.2014.2386294

SU, W., BOGDAN, M. and CANDES, E. (2017). False discoveries occur early on the Lasso path. *Ann. Statist.* **45** 2133–2150. MR3718164 https://doi.org/10.1214/16-AOS1521

SUN, Q., ZHOU, W.-X. and FAN, J. (2020). Adaptive Huber regression. *J. Amer. Statist. Assoc.* **115** 254–265. MR4078461 https://doi.org/10.1080/01621459.2018.1543124

SUR, P. and CANDES, E. J. (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proc. Natl. Acad. Sci. USA* **116** 14516–14525. MR3984492 https://doi.org/10.1073/pnas.1810420116

SUR, P., CHEN, Y. and CANDES, E. J. (2019). The likelihood ratio test in high-dimensional logistic regression is asymptotically a *rescaled* chi-square. *Probab. Theory Related Fields* **175** 487–558. MR4009715 https://doi.org/10.1007/s00440-018-00896-9

THRAMPOULIDIS, C., ABBASI, E. and HASSIBI, B. (2018). Precise error analysis of regularized M-estimators in high dimensions. *IEEE Trans. Inf. Theory* **64** 5592–5628. MR3832326 https://doi.org/10.1109/TIT.2018.2840720

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B, Methodol.* **58** 267–288. MR1379242

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 https://doi.org/10.1214/14-AOS1221

VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247 https://doi.org/10.1017/CBO9780511802256

VENKATARAMANAN, R., KÖGLER, K. and MONDELLI, M. (2021). Estimation in rotationally invariant generalized linear models via approximate message passing. Preprint. Available at arXiv:2112.04330.

WAINWRIGHT, M. J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). *IEEE Trans. Inf. Theory* **55** 2183–2202. MR2729873 https://doi.org/10.1109/TIT.2009.2016018

WANG, T., ZHONG, X. and FAN, Z. (2024). Universality of approximate message passing algorithms and tensor networks. *Ann. Appl. Probab.* **34** 3943–3994. MR4783034 https://doi.org/10.1214/24-aap2056

WU, Y. and ZHOU, K. (2024). Sharp analysis of power iteration for tensor PCA. *J. Mach. Learn. Res.* **25** Paper No. [195], 42 pp. MR4777437

XIA, D. and YUAN, M. (2021). Statistical inferences of linear forms for noisy matrix completion. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 58–77. MR4220984 https://doi.org/10.1111/rssb.12400

YAN, Y., CHEN, Y. and FAN, J. (2024). Inference for heteroskedastic PCA with missing data. *Ann. Statist.* **52** 729–756. MR4744194 https://doi.org/10.1214/24-aos2366

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 https://doi.org/10.1111/j.1467-9868.2005.00532.x

ZHANG, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. MR2604701 https://doi.org/10.1214/09-AOS729

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 https://doi.org/10.1111/rssb.12026

ZHANG, Y., JI, H. C., VENKATARAMANAN, R. and MONDELLI, M. (2025). Spectral estimators for structured generalized linear models via approximate message passing. *Math. Stat. Learn.* **8** 193–304. MR4963421 https://doi.org/10.4171/msl/52

ZHAO, P. and YU, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.* **7** 2541–2563. MR2274449

ZHONG, X., WANG, T. and FAN, Z. (2024). Approximate message passing for orthogonally invariant ensembles: Multivariate non-linearities and spectral initialization. *Inf. Inference* **13** Paper No. iaae024, 22 pp. MR4797647 https://doi.org/10.1093/imaiai/iaae024

ZOU, H. and HASTIE, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 301–320. MR2137327 https://doi.org/10.1111/j.1467-9868.2005.00503.x

The Annals of Statistics
2026, Vol. 54, No. 4, 1682–1711
https://doi.org/10.1214/26-AOS2621

LOCAL GEOMETRY OF HIGH-DIMENSIONAL MIXTURE MODELS: EFFECTIVE SPECTRAL THEORY AND DYNAMICAL TRANSITIONS

BY GÉRARD BEN AROUS[1,a], REZA GHEISSARI[2,b], JIAOYANG HUANG[3,c] AND AUKOSH JAGANNATH[4,d]

[1]*Courant Institute, New York University,* [a]*benarous@cims.nyu.edu*

[2]*Department of Mathematics, Northwestern University,* [b]*gheissari@northwestern.edu*

[3]*Department of Statistics and Data Science, University of Pennsylvania,* [c]*huangjy@wharton.upenn.edu*

[4]*Department of Statistics and Actuarial Science, Department of Applied Mathematics, and Cheriton School of Computer Science, University of Waterloo,* [d]*a.jagannath@uwaterloo.ca*

We study the local geometry of empirical risks in high dimensions via the spectral theory of their Hessian and information matrices. We focus on settings where the data, $(Y_\ell)_{\ell=1}^n \in \mathbb{R}^d$, are i.i.d. draws of a k-Gaussian mixture model, and the loss depends on the projection of the data into a fixed number of vectors, namely $\mathbf{x}^\top Y$, where $\mathbf{x} \in \mathbb{R}^{d\times\mathcal{C}}$ are the parameters, and $\mathcal{C}$ need not equal k. This setting captures a broad class of problems such as classification by one and two-layer networks and regression on multi-index models. We provide exact formulas for the limits of the empirical spectral distribution and outlier eigenvalues and eigenvectors of such matrices in the proportional asymptotics limit, where the number of samples and dimension $n, d \to \infty$ and $n/d = \phi \in (0, \infty)$. These limits depend on the parameters $\mathbf{x}$ only through the summary statistic of the $(\mathcal{C}+k)\times(\mathcal{C}+k)$ Gram matrix of the parameters and class means, $\mathbf{G} = (\mathbf{x}, \boldsymbol{\mu})^\top(\mathbf{x}, \boldsymbol{\mu})$.

It is known that under general conditions, when $\mathbf{x}$ is trained by online stochastic gradient descent, the evolution of these same summary statistics along training converges to the solution of an autonomous system of ODEs, called the *effective dynamics*. This enables us to connect the training dynamics to the spectral theory of these matrices generated with test data. We demonstrate our general results by analyzing the *effective spectrum* along the effective dynamics in the case of multiclass logistic regression. In this setting, the empirical Hessian and information matrices have substantially different spectra, each with their own static and dynamical spectral transitions.

REFERENCES

[1] ABBE, E., ADSERA, E. B. and MISIAKIEWICZ, T. (2023). Sgd learning on neural networks: Leap complexity and saddle-to-saddle dynamics. In *The Thirty Sixth Annual Conference on Learning Theory* 2552–2623. PMLR.

[2] AMARI, S. (2016). *Information Geometry and Its Applications*. *Applied Mathematical Sciences* **194**. Springer, Tokyo. MR3495836 https://doi.org/10.1007/978-4-431-55978-8

[3] ARNABOLDI, L., STEPHAN, L., KRZAKALA, F. and LOUREIRO, B. (2023). From high-dimensional & mean-field dynamics to dimensionless ODEs: A unifying approach to SGD in two-layers networks. In *Proceedings of Thirty Sixth Conference on Learning Theory* (G. Neu and L. Rosasco, eds.). *Proceedings of Machine Learning Research* **195** 1199–1227. PMLR.

[4] ASGARI, K., MONTANARI, A. and SAEED, B. (2025). Local minima of the empirical risk in high dimension: General theorems and convex examples.

[5] AUBIN, B., LOUREIRO, B. V., BAKER, A., KRZAKALA, F. and ZDEBOROVÁ, L. (2019). Exact asymptotics for phase retrieval and compressed sensing with random generative priors. ArXiv. Available at arXiv:1912.02008.

[6] BAI, Z. and SILVERSTEIN, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2567175 https://doi.org/10.1007/978-1-4419-0661-8

MSC2020 subject classifications. Primary 62R07, 62H30; secondary 60B20.
Key words and phrases. Gaussian mixture model, random matrix theory, stochastic gradient descent.

[7] BAI, Z. D. and SILVERSTEIN, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051 https://doi.org/10.1214/aop/1022855421

[8] 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 https://doi.org/10.1214/009117905000000233

[9] BARTLETT, P. L., MONTANARI, A. and RAKHLIN, A. (2021). Deep learning: A statistical viewpoint. *Acta Numer.* **30** 87–201. MR4295218 https://doi.org/10.1017/S0962492921000027

[10] BEN AROUS, G., GHEISSARI, R., HUANG, J. and JAGANNATH, A. (2024). High-dimensional SGD aligns with emerging outlier eigenspaces. In *The Twelfth International Conference on Learning Representations*.

[11] BEN AROUS, G., GHEISSARI, R., HUANG, J. and JAGANNATH, A. (2026). Supplement to "Local geometry of high-dimensional mixture models: Effective spectral theory and dynamical transitions." https://doi.org/10.1214/26-AOS2621SUPP

[12] BEN AROUS, G., GHEISSARI, R. and JAGANNATH, A. (2021). Online stochastic gradient descent on nonconvex losses from high-dimensional inference. *J. Mach. Learn. Res.* **22** 106. MR4279757

[13] BEN AROUS, G., GHEISSARI, R. and JAGANNATH, A. (2024). High-dimensional limit theorems for SGD: Effective dynamics and critical scaling. *Comm. Pure Appl. Math.* **77** 2030–2080. MR4692886 https://doi.org/10.1002/cpa.22169

[14] BIETTI, A., BRUNA, J. and PILLAUD-VIVIEN, L. (2025). On learning Gaussian multi-index models with gradient flow part I: General properties and two-timescale learning. *Comm. Pure Appl. Math.* **78** 2354–2435. MR4969078 https://doi.org/10.1002/cpa.70006

[15] BISHOP, C. M. (2006). *Pattern Recognition and Machine Learning*. *Information Science and Statistics*. Springer, New York. MR2247587 https://doi.org/10.1007/978-0-387-45528-0

[16] CANDÈS, E. J., LI, X. and SOLTANOLKOTABI, M. (2015). Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Trans. Inf. Theory* **61** 1985–2007. MR3332993 https://doi.org/10.1109/TIT.2015.2399924

[17] CHEN, Y., CHI, Y., FAN, J. and MA, C. (2019). Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval. *Math. Program.* **176** 5–37. MR3960803 https://doi.org/10.1007/s10107-019-01363-6

[18] CHIZAT, L., OYALLON, E. and BACH, F. (2019). On lazy training in differentiable programming. *Adv. Neural Inf. Process. Syst.* **32**.

[19] COHEN, J., KAUR, S., LI, Y., KOLTER, J. Z. and TALWALKAR, A. (2021). Gradient descent on neural networks typically occurs at the edge of stability. In *International Conference on Learning Representations*.

[20] COLLINS-WOODFIN, E., PAQUETTE, C., PAQUETTE, E. and SEROUSSI, I. (2024). Hitting the high-dimensional notes: An ODE for SGD learning dynamics on GLMs and multi-index models. *Inf. Inference* **13** iaae028. MR4811680 https://doi.org/10.1093/imaiai/iaae028

[21] DAMIAN, A., LEE, J. and SOLTANOLKOTABI, M. (2022). Neural networks can learn representations with gradient descent. In *Proceedings of Thirty Fifth Conference on Learning Theory* (P.-L. Loh and M. Raginsky, eds.). *Proceedings of Machine Learning Research* **178** 5413–5452. PMLR.

[22] DAMIAN, A., NICHANI, E., GE, R. and LEE, J. D. (2023). Smoothing the landscape boosts the signal for SGD: Optimal sample complexity for learning single index models. In *Advances in Neural Information Processing Systems* (A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine, eds.) **36** 752–784. Curran Associates, Red Hook.

[23] DEFILIPPIS, L., DANDI, Y., MERGNY, P., KRZAKALA, F. and LOUREIRO, B. (2025). *Optimal Spectral Transitions in High-Dimensional Multi-Index Models*.

[24] DERRIDA, B. (1981). Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B* (3) **24** 2613–2626. MR0627810 https://doi.org/10.1103/physrevb.24.2613

[25] DUDEJA, R. and HSU, D. (2018). Learning single-index models in Gaussian space. In *Proceedings of the* 31*st Conference on Learning Theory*. *Proceedings of Machine Learning Research* **75** 1887–1930. PMLR.

[26] FAN, Z. and WANG, Z. (2020). Spectra of the conjugate kernel and neural tangent kernel for linear-width neural networks. In *Proceedings of the* 34*th International Conference on Neural Information Processing Systems*. *NIPS'*20 Curran Associates, Red, Hook, NY, USA.

[27] GARROD, C. and KEATING, J. P. (2024). Unifying Low Dimensional Observations in Deep Learning Through the Deep Linear Unconstrained Feature Model.

[28] GHORBANI, B., KRISHNAN, S. and XIAO, Y. (2019). An investigation into neural net optimization via Hessian eigenvalue density. In *Proceedings of the* 36*th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.). *Proceedings of Machine Learning Research* **97** 2232–2241. PMLR.

[29] GOLDT, S., ADVANI, M., SAXE, A. M., KRZAKALA, F. and ZDEBOROVÁ, L. (2019). Dynamics of stochastic gradient descent for two-layer neural networks in the teacher-student setup. *Adv. Neural Inf. Process. Syst.* **32**.

[30] GUR-ARI, G., ROBERTS, D. A. and DYER, E. (2019). Gradient Descent Happens in a Tiny Subspace.

[31] 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 https://doi.org/10.1007/978-0-387-84858-7

[32] JACOT, A., GABRIEL, F. and HONGLER, C. (2020). The asymptotic spectrum of the Hessian of DNN throughout training. In *International Conference on Learning Representations*.

[33] JEONG, H. and GÜNTÜRK, C. S. (2017). Convergence of the randomized Kaczmarz method for phase retrieval. ArXiv. Available at arXiv:1706.10291.

[34] KNOWLES, A. and YIN, J. (2017). Anisotropic local laws for random matrices. *Probab. Theory Related Fields* **169** 257–352. MR3704770 https://doi.org/10.1007/s00440-016-0730-4

[35] KOVACEVIC, F., ZHANG, Y. and MONDELLI, M. (2025). Spectral Estimators for Multi-Index Models: Precise Asymptotics and Optimal Weak Recovery.

[36] LECUN, Y. A., BOTTOU, L., ORR, G. B. and MÜLLER, K.-R. (2012). Efficient BackProp. In *Neural Networks*: *Tricks of the Trade* 9–48 Second ed. Springer, Berlin. https://doi.org/10.1007/978-3-642-35289-8_3

[37] LI, X., GU, Q., ZHOU, Y., CHEN, T. and BANERJEE, A. (2020). Hessian based analysis of SGD for deep nets: Dynamics and generalization. In *Proceedings of the* 2020 *SIAM International Conference on Data Mining* (*SDM*) 190–198. https://doi.org/10.1137/1.9781611976236.22

[38] LIAO, Z. and MAHONEY, M. W. (2021). Hessian eigenspectra of more realistic nonlinear models. In *Advances in Neural Information Processing Systems* (A. Beygelzimer, Y. Dauphin, P. Liang and J. W. Vaughan, eds.).

[39] LU, Y. M. and LI, G. (2020). Phase transitions of spectral initialization for high-dimensional non-convex estimation. *Inf. Inference* **9** 507–541. MR4146345 https://doi.org/10.1093/imaiai/iaz020

[40] LUO, W., ALGHAMDI, W. and LU, Y. M. (2019). Optimal spectral initialization for signal recovery with applications to phase retrieval. *IEEE Trans. Signal Process.* **67** 2347–2456. MR3962211 https://doi.org/10.1109/TSP.2019.2904918

[41] MAILLARD, A., BEN AROUS, G. and BIROLI, G. (2020). Landscape complexity for the empirical risk of generalized linear models. In *Proceedings of the First Mathematical and Scientific Machine Learning Conference* (J. Lu and R. Ward, eds.). *Proceedings of Machine Learning Research* **107** 287–327. PMLR.

[42] MARCHENKO, V. A. and PASTUR, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb.* **114** 507–536.

[43] MARRIOTT, P. (2002). On the local geometry of mixture models. *Biometrika* **89** 77–93. MR1888347 https://doi.org/10.1093/biomet/89.1.77

[44] MARTIN, C. H. and MAHONEY, M. W. (2019). Traditional and Heavy Tailed Self Regularization in Neural Network Models.

[45] MINSKY, M. and PAPERT, S. (1969). An introduction to computational geometry. *Camb. Tiass., HIT* **479** 480.

[46] MONDELLI, M. and MONTANARI, A. (2018). Fundamental limits of weak recovery with applications to phase retrieval. In *Conference on Learning Theory* 1445–1450. PMLR.

[47] MOUSAVI-HOSSEINI, A., PARK, S., GIROTTI, M., MITLIAGKAS, I. and ERDOGDU, M. A. (2023). Neural networks efficiently learn low-dimensional representations with SGD. In *The Eleventh International Conference on Learning Representations*.

[48] MOUSAVI-HOSSEINI, A., WU, D., SUZUKI, T. and ERDOGDU, M. A. (2023). Gradient-based feature learning under structured data. In *Advances in Neural Information Processing Systems* (A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt and S. Levine, eds.) **36** 71449–71485. Curran Associates, Red Hook.

[49] NIELSEN, F., RAO, A. SS. and RAO, C. R. (2022). Geometry and Statistics 46, San Diego Academic Press.

[50] OYMAK, S. and SOLTANOLKOTABI, M. (2019). Overparameterized nonlinear learning: Gradient descent takes the shortest path? In *International Conference on Machine Learning* 4951–4960. PMLR.

[51] PAPYAN, V. (2019). Measurements of three-level hierarchical structure in the outliers in the spectrum of deepnet hessians. In *Proceedings of the* 36*th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.). *Proceedings of Machine Learning Research* **97** 5012–5021. PMLR.

[52] PAPYAN, V. (2020). Traces of class/cross-class structure pervade deep learning spectra. *J. Mach. Learn. Res.* **21** 252. MR4213433

[53] PAPYAN, V., HAN, X. Y. and DONOHO, D. L. (2020). Prevalence of neural collapse during the terminal phase of deep learning training. *Proc. Natl. Acad. Sci. USA* **117** 24652–24663. MR4250189 https://doi.org/10.1073/pnas.2015509117

[54] PAQUETTE, C., LEE, K., PEDREGOSA, F. and PAQUETTE, E. (2021). SGD in the large: Average-case analysis, asymptotics, and stepsize criticality. In *Conference on Learning Theory* 3548–3626. PMLR.

[55] PENNINGTON, J. and BAHRI, Y. (2017). Geometry of neural network loss surfaces via random matrix theory. In *Proceedings of the* 34*th International Conference on Machine Learning* (D. Precup and Y. W. Teh, eds.). *Proceedings of Machine Learning Research* **70** 2798–2806. PMLR.

[56] PENNINGTON, J. and WORAH, P. (2018). The spectrum of the Fisher information matrix of a single-hidden-layer neural network. In *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi and R. Garnett, eds.). *Curran Associates* **31**. Curran Associates, Red Hook.

[57] REFINETTI, M., GOLDT, S., KRZAKALA, F. and ZDEBOROVÁ, L. (2021). Classifying high-dimensional Gaussian mixtures: Where kernel methods fail and neural networks succeed. In *International Conference on Machine Learning* 8936–8947. PMLR.

[58] SAAD, D. and SOLLA, S. (1995). Dynamics of on-line gradient descent learning for multilayer neural networks. *Adv. Neural Inf. Process. Syst.* **8**.

[59] SAAD, D. and SOLLA, S. A. (1995). On-line learning in soft committee machines. *Phys. Rev. E* **52** 4225.

[60] SAGUN, L., BOTTOU, L. and LECUN, Y. (2017). *Eigenvalues of the Hessian in Deep Learning: Singularity and Beyond*.

[61] SAGUN, L., EVCI, U., GÜNEY, V. U., DAUPHIN, Y. N. and BOTTOU, L. (2017). Empirical analysis of the Hessian of over-parametrized neural networks. CoRR. Available at arXiv:1706.04454.

[62] SILVERSTEIN, J. W. and BAI, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.* **54** 175–192. MR1345534 https://doi.org/10.1006/jmva.1995.1051

[63] SIMSEK, B., BENDJEDDOU, A. and HSU, D. (2024). *Learning Gaussian Multi-Index Models with Gradient Flow: Time Complexity and Directional Convergence*.

[64] TAN, Y. S. and VERSHYNIN, R. (2019). Phase retrieval via randomized Kaczmarz: Theoretical guarantees. *Inf. Inference* **8** 97–123. MR3922404 https://doi.org/10.1093/imaiai/iay005

[65] TAN, Y. S. and VERSHYNIN, R. (2023). Online stochastic gradient descent with arbitrary initialization solves non-smooth, non-convex phase retrieval. *J. Mach. Learn. Res.* **24** 58. MR4582480

[66] VEIGA, R., STEPHAN, L., LOUREIRO, B., KRZAKALA, F. and ZDEBOROVÁ, L. (2023). Phase diagram of stochastic gradient descent in high-dimensional two-layer neural networks. *J. Stat. Mech. Theory Exp.* **11** 114008. MR4725421 https://doi.org/10.1088/1742-5468/ad01b1

[67] XIE, Z., TANG, Q.-Y., SUN, M. and LI, P. (2023). On the overlooked structure of stochastic gradients. In *Thirty-Seventh Conference on Neural Information Processing Systems*.

[68] ZHANG, H. and LIANG, Y. (2016). Reshaped Wirtinger flow for solving quadratic system of equations. In *Advances in Neural Information Processing Systems* 29 (D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon and R. Garnett, eds.) 2622–2630. Curran Associates, Red Hook.

[69] ZHANG, Y., MONDELLI, M. and VENKATARAMANAN, R. (2022). Precise asymptotics for spectral methods in mixed generalized linear models. arXiv preprint, Available at arXiv:2211.11368.

The Annals of Statistics
2026, Vol. 54, No. 4, 1712–1739
https://doi.org/10.1214/26-AOS2622

HIGH-ORDER ACCURATE INFERENCE ON MANIFOLDS

BY CHENGZHU HUANG[1,a] AND ANRU R. ZHANG[2,b]

[1]*Department of Statistics, Columbia University,* [a]*ch3786@columbia.edu*

[2]*Department of Biostatistics & Bioinformatics and Department of Computer Science, Duke University,* [b]*anru.zhang@duke.edu*

We present a new framework for statistical inference on Riemannian manifolds that achieves high-order accuracy, addressing the challenges posed by non-Euclidean parameter spaces frequently encountered in modern data science. Our approach leverages a novel and computationally efficient procedure to reach higher-order asymptotic precision. In particular, we develop a bootstrap algorithm on Riemannian manifolds that is both computationally efficient and accurate for hypothesis testing and confidence region construction. Although locational hypothesis testing can be reformulated as a standard Euclidean problem, constructing high-order accurate confidence regions necessitates careful treatment of manifold geometry. To this end, we establish high-order asymptotics under an appropriate coordinate representation induced by a second-order retraction, thereby enabling precise expansions that incorporate curvature effects. We demonstrate the versatility of this framework across various manifold settings, including spheres, the Stiefel manifold, fixed-rank matrix manifolds, and rank-one tensor manifolds; for Euclidean submanifolds, we also introduce a class of projection-like coordinate charts with strong consistency properties. Finally, numerical studies confirm the practical merits of the proposed procedure.

REFERENCES

[1] ABSIL, P.-A., MAHONY, R. and SEPULCHRE, R. (2008). *Optimization Algorithms on Matrix Manifolds*. Princeton Univ. Press, Princeton, NJ. MR2364186 https://doi.org/10.1515/9781400830244

[2] ABSIL, P.-A., MAHONY, R. and TRUMPF, J. (2013). An extrinsic look at the Riemannian Hessian. In *Geometric Science of Information*. *Lecture Notes in Computer Science* **8085** 361–368. Springer, Heidelberg. MR3126065 https://doi.org/10.1007/978-3-642-40020-9_39

[3] ABSIL, P.-A. and MALICK, J. (2012). Projection-like retractions on matrix manifolds. *SIAM J. Optim.* **22** 135–158. MR2902688 https://doi.org/10.1137/100802529

[4] ADLER, R. L., DEDIEU, J.-P., MARGULIES, J. Y., MARTENS, M. and SHUB, M. (2002). Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.* **22** 359–390. MR1918655 https://doi.org/10.1093/imanum/22.3.359

[5] AMARI, S. and KUMON, M. (1983). Differential geometry of Edgeworth expansions in curved exponential family. *Ann. Inst. Statist. Math.* **35** 1–24. MR0700711 https://doi.org/10.1007/BF02480959

[6] ANDREWS, D. W. K. (2002). Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica* **70** 119–162. MR1926257 https://doi.org/10.1111/1468-0262.00271

[7] AY, N., JOST, J., LÊ, H. V. and SCHWACHHÖFER, L. (2017). *Inf. Geom.*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete*. 3. *Folge*. *A Series of Modern Surveys in Mathematics* [*Results in Mathematics and Related Areas*. 3*rd Series*. *A Series of Modern Surveys in Mathematics*] **64**. Springer, Cham. MR3701408 https://doi.org/10.1007/978-3-319-56478-4

[8] BHATTACHARYA, R. and PATRANGENARU, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Ann. Statist.* **31** 1–29. MR1962498 https://doi.org/10.1214/aos/1046294456

[9] BHATTACHARYA, R. and PATRANGENARU, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds. II. *Ann. Statist.* **33** 1225–1259. MR2195634 https://doi.org/10.1214/009053605000000093

[10] BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. MR0471142

MSC2020 subject classifications. 62H15, 62F40.
Key words and phrases. Inference, Edgeworth expansion, Riemannian optimization.

[11] BOUMAL, N., ABSIL, P.-A. and CARTIS, C. (2019). Global rates of convergence for nonconvex optimization on manifolds. *IMA J. Numer. Anal.* **39** 1–33. MR4023745 https://doi.org/10.1093/imanum/drx080

[12] BREWIN, L. (2009). Riemann normal coordinate expansions using Cadabra. *Classical Quantum Gravity* **26** 175017. MR2534336 https://doi.org/10.1088/0264-9381/26/17/175017

[13] BRUNEL, V.-E. (2023). Geodesically convex M-estimation in metric spaces. In *The Thirty Sixth Annual Conference on Learning Theory* 2188–2210. PMLR.

[14] CHEN, S., MA, S., SO, A. M.-C. and ZHANG, T. (2020). Proximal gradient method for nonsmooth optimization over the Stiefel manifold. *SIAM J. Optim.* **30** 210–239. MR4049406 https://doi.org/10.1137/18M122457X

[15] DAVIDSON, R. and MACKINNON, J. G. (1999). Bootstrap testing in nonlinear models. *Internat. Econom. Rev.* **40** 487–508. MR1690293 https://doi.org/10.1111/1468-2354.00026

[16] DENNIS, J. E. JR. and SCHNABEL, R. B. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. *Classics in Applied Mathematics* **16**. SIAM, Philadelphia, PA. MR1376139 https://doi.org/10.1137/1.9781611971200

[17] DURRETT, R. (2019). *Probability—Theory and Examples*. *Cambridge Series in Statistical and Probabilistic Mathematics* **49**. Cambridge Univ. Press, Cambridge. MR3930614 https://doi.org/10.1017/9781108591034

[18] EDELMAN, A., ARIAS, T. A. and SMITH, S. T. (1999). The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* **20** 303–353. MR1646856 https://doi.org/10.1137/S0895479895290954

[19] EFRON, B. (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann. Statist.* **3** 1189–1242. MR0428531

[20] EFRON, B. (1978). The geometry of exponential families. *Ann. Statist.* **6** 362–376. MR0471152

[21] EL KAROUI, N. and PURDOM, E. (2018). Can we trust the bootstrap in high-dimensions? The case of linear models. *J. Mach. Learn. Res.* **19** Paper No. 5. MR3862412

[22] FISHER, N. I., LEWIS, T. and EMBLETON, B. J. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge Univ. Press, Cambridge. MR0899958 https://doi.org/10.1017/CBO9780511623059

[23] FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond., Ser. A, Contain. Pap. Math. Phys. Character* **222** 309–368.

[24] FRÉCHET, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. Henri Poincaré* **10** 215–310. MR0027464

[25] GAVRILOV, A. V. (2006). Algebraic properties of the covariant differentiation and composition of exponential mappings. *Mat. Tr.* **9** 3–20. MR2251328

[26] GAVRILOV, A. V. (2007). The double exponential mapping and covariant differentiation. *Sibirsk. Mat. Zh.* **48** 68–74. MR2304878 https://doi.org/10.1007/s11202-007-0006-4

[27] GAWLIK, E. S. and LEOK, M. (2018). High-order retractions on matrix manifolds using projected polynomials. *SIAM J. Matrix Anal. Appl.* **39** 801–828. MR3800232 https://doi.org/10.1137/17M1130459

[28] HALL, P. (2013). *The Bootstrap and Edgeworth Expansion*. Springer Science & Business Media.

[29] HALL, P. and HOROWITZ, J. L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica* **64** 891–916. MR1399222 https://doi.org/10.2307/2171849

[30] HEIDEL, G. and SCHULZ, V. (2018). A Riemannian trust-region method for low-rank tensor completion. *Numer. Linear Algebra Appl.* **25** e2175. MR3890978 https://doi.org/10.1002/nla.2175

[31] HUANG, C. and ZHANG, A. R (2026). Supplement to "High-order accurate inference on manifolds." https://doi.org/10.1214/26-AOS2622SUPP

[32] HUANG, W., ABSIL, P.-A. and GALLIVAN, K. A. (2016). A Riemannian BFGS method for nonconvex optimization problems. In *Numerical Mathematics and Advanced Applications—ENUMATH* 2015. *Lecture Notes in Computational Science and Engineering* **112** 627–634. Springer, Cham. MR3706408 https://doi.org/10.1007/978-3-319-39929-4_6

[33] JOST, J. (2008). *Riemannian Geometry and Geometric Analysis*, 5th ed. *Universitext*. Springer, Berlin. MR2431897

[34] KARCHER, H. (1977). Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.* **30** 509–541. MR0442975 https://doi.org/10.1002/cpa.3160300502

[35] KASAI, H. and MISHRA, B. (2016). Low-rank tensor completion: A Riemannian manifold preconditioning approach. In *International Conference on Machine Learning* 1012–1021. PMLR.

[36] LEE, J. and JUNG, S. (2026). Huber means on Riemannian manifolds. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **88** 444–463. MR5058937 https://doi.org/10.1093/jrsssb/qkaf054

[37] LEE, J. and JUNG, S. (2026). General M-estimators of location on Riemannian manifolds: Existence and uniqueness. *Statist. Probab. Lett.* **233** Paper No. 110670. MR5027575 https://doi.org/10.1016/j.spl.2026.110670

[38] LEE, J. M. (2013). *Introduction to Smooth Manifolds*, 2nd ed. *Graduate Texts in Mathematics* **218**. Springer, New York. MR2954043

[39] LEE, J. M. (2018). *Introduction to Riemannian Manifolds*. *Graduate Texts in Mathematics* **176**. Springer, Cham. MR3887684

[40] LI, X., CHEN, S., DENG, Z., QU, Q., ZHU, Z. and SO, A. M.-C. (2021). Weakly convex optimization over Stiefel manifold using Riemannian subgradient-type methods. *SIAM J. Optim.* **31** 1605–1634. MR4279932 https://doi.org/10.1137/20M1321000

[41] LUO, Y., LI, X. and ZHANG, A. R. (2024). On geometric connections of embedded and quotient geometries in Riemannian fixed-rank matrix optimization. *Math. Oper. Res.* **49** 782–825. MR4755753 https://doi.org/10.1287/moor.2023.1377

[42] MATAIGNE, S., ZIMMERMANN, R. and MIOLANE, N. (2025). An efficient algorithm for the Riemannian logarithm on the Stiefel manifold for a family of Riemannian metrics. *SIAM J. Matrix Anal. Appl.* **46** 879–905. MR4888263 https://doi.org/10.1137/24M1647801

[43] MISHRA, B., APUROOP, K. A. and SEPULCHRE, R. (2012). A Riemannian geometry for low-rank matrix completion. arXiv Preprint. Available at arXiv:1211.1550.

[44] MISHRA, B., MEYER, G., BONNABEL, S. and SEPULCHRE, R. (2014). Fixed-rank matrix factorizations and Riemannian low-rank optimization. *Comput. Statist.* **29** 591–621. MR3261830 https://doi.org/10.1007/s00180-013-0464-z

[45] NEWEY, W. K. and MCFADDEN, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, *Vol. IV*. *Handbooks in Econom.* **2** 2111–2245. North-Holland, Amsterdam. MR1315971

[46] PAINDAVEINE, D. and VERDEBOUT, T. (2020). Inference for spherical location under high concentration. *Ann. Statist.* **48** 2982–2998. MR4152631 https://doi.org/10.1214/19-AOS1918

[47] PENNEC, X. (2018). Barycentric subspace analysis on manifolds. *Ann. Statist.* **46** 2711–2746. MR3851753 https://doi.org/10.1214/17-AOS1636

[48] PENNEC, X. (2019). Curvature effects on the empirical mean in Riemannian and affine Manifolds: a non-asymptotic high concentration expansion in the small-sample regime. arXiv Preprint. Available at arXiv:1906.07418.

[49] SMITH, S. T. (1993). *Geometric Optimization Methods for Adaptive Filtering*. Harvard University.

[50] SMITH, S. T. (1994). Optimization techniques on Riemannian manifolds. In *Hamiltonian and Gradient Flows*, *Algorithms and Control*. *Fields Inst. Commun.* **3** 113–136. Amer. Math. Soc., Providence, RI. MR1297990 https://doi.org/10.1090/fic/003/09

[51] TAGARE, H. D. (2011). Notes on optimization on Stiefel manifolds. Yale Univer., New Haven.

[52] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247 https://doi.org/10.1017/CBO9780511802256

[53] VANDEREYCKEN, B. (2013). Low-rank matrix completion by Riemannian optimization. *SIAM J. Optim.* **23** 1214–1236. MR3069099 https://doi.org/10.1137/110845768

[54] WALD, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Stat.* **20** 595–601. MR0032169 https://doi.org/10.1214/aoms/1177729952

[55] ZHANG, A. (2019). Cross: Efficient low-rank tensor completion. *Ann. Statist.* **47** 936–964. MR3909956 https://doi.org/10.1214/18-AOS1694

[56] ZHENG, Z., ZHOU, X., FAN, Y. and LV, J. (2026). SOFARI: High-dimensional manifold-based inference. *J. Amer. Statist. Assoc.* **121** 513–524. MR5061139 https://doi.org/10.1080/01621459.2025.2529026

The Annals of Statistics
2026, Vol. 54, No. 4, 1740–1767
https://doi.org/10.1214/26-AOS2623

UNIFORM ESTIMATION AND INFERENCE FOR NONPARAMETRIC PARTITIONING-BASED M-ESTIMATORS

BY MATIAS D. CATTANEO[1,a], YINGJIE FENG[2,b] AND BORIS SHIGIDA[3,c]

[1] *Department of Economics, Princeton University,* [a] *cattaneo@princeton.edu*
[2] *School of Economics and Management, Tsinghua University,* [b] *fengyj@sem.tsinghua.edu.cn*
[3] *Department of Operations Research and Financial Engineering, Princeton University,* [c] *bs1624@princeton.edu*

This paper presents uniform estimation and inference theory for a large class of nonparametric partitioning-based M-estimators. The main theoretical results include: (i) uniform consistency for convex and nonconvex objective functions; (ii) rate-optimal uniform Bahadur representations; (iii) rate-optimal uniform (and mean square) convergence rates; (iv) valid strong approximations and feasible uniform inference methods; and (v) extensions to functional transformations of underlying estimators. Uniformity is established over both the evaluation point of the nonparametric functional parameter and a Euclidean parameter indexing the class of loss functions. The results also account explicitly for the smoothness degree of the loss function (if any), and allow for a possibly nonidentity (inverse) link function. We illustrate the theoretical and methodological results in four examples: quantile regression, distribution regression, L_p regression, and logistic regression. Many other possibly nonsmooth, nonlinear, generalized, robust M-estimation settings are covered by our results. We provide detailed comparisons with the existing literature and demonstrate substantive improvements: we achieve the best (in some cases optimal) known results under improved (in some cases minimal) requirements in terms of regularity conditions and side rate restrictions. The Supplementary Material reports complementary technical results that may be of independent interest, including a novel uniform strong approximation result based on Yurinskii's coupling.

REFERENCES

[1] ABADIE, A. and CATTANEO, M. D. (2018). Econometric methods for program evaluation. *Ann. Rev. Econ.* **10** 465–503.

[2] BELLONI, A., CHERNOZHUKOV, V., CHETVERIKOV, D. and FERNÁNDEZ-VAL, I. (2019). Conditional quantile processes based on series or many regressors. *J. Econometrics* **213** 4–29. MR4013213 https://doi.org/10.1016/j.jeconom.2019.04.003

[3] BELLONI, A., CHERNOZHUKOV, V., CHETVERIKOV, D. and KATO, K. (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *J. Econometrics* **186** 345–366. MR3343791 https://doi.org/10.1016/j.jeconom.2015.02.014

[4] BHATIA, R. (1997). *Matrix Analysis*. *Graduate Texts in Mathematics* **169**. Springer, New York. MR1477662 https://doi.org/10.1007/978-1-4612-0653-8

[5] BIRGÉ, L. (2001). An alternative point of view on Lepski's method. In *State of the Art in Probability and Statistics* (*Leiden*, 1999). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **36** 113–133. IMS, Beachwood, OH. MR1836557 https://doi.org/10.1214/lnms/1215090065

[6] CALONICO, S., CATTANEO, M. D. and FARRELL, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *J. Amer. Statist. Assoc.* **113** 767–779. MR3832225 https://doi.org/10.1080/01621459.2017.1285776

[7] CALONICO, S., CATTANEO, M. D. and FARRELL, M. H. (2022). Coverage error optimal confidence intervals for local polynomial regression. *Bernoulli* **28** 2998–3022. MR4474570 https://doi.org/10.3150/21-bej1445

MSC2020 subject classifications. Primary 62G05; secondary 62G20, 62G08.

Key words and phrases. Nonparametric estimation and inference, series methods, partitioning estimators, quantile regression, nonlinear regression, robust regression, generalized linear models, uniform distribution theory.

[8] CATTANEO, M. D., CRUMP, R. K., FARRELL, M. H. and FENG, Y. (2024). On binscatter. *Amer. Econ. Rev.* **114** 1488–1514.

[9] CATTANEO, M. D. and FARRELL, M. H. (2013). Optimal convergence rates, Bahadur representation, and asymptotic normality of partitioning estimators. *J. Econometrics* **174** 127–143. MR3045024 https://doi.org/10.1016/j.jeconom.2013.02.002

[10] CATTANEO, M. D., FARRELL, M. H. and FENG, Y. (2020). Large sample properties of partitioning-based series estimators. *Ann. Statist.* **48** 1718–1741. MR4124341 https://doi.org/10.1214/19-AOS1865

[11] CATTANEO, M. D., FENG, Y. and SHIGIDA, B. (2026). Supplement to "Uniform estimation and inference for nonparametric partitioning-based M-estimators." https://doi.org/10.1214/26-AOS2623SUPP

[12] CHAUDHURI, P. (1991). Global nonparametric estimation of conditional quantile functions and their derivatives. *J. Multivariate Anal.* **39** 246–269. MR1147121 https://doi.org/10.1016/0047-259X(91)90100-G

[13] CHEN, X. and CHRISTENSEN, T. M. (2015). Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions. *J. Econometrics* **188** 447–465. MR3383220 https://doi.org/10.1016/j.jeconom.2015.03.010

[14] CHEN, X. and KATO, K. (2020). Jackknife multiplier bootstrap: Finite sample approximations to the U-process supremum with applications. *Probab. Theory Related Fields* **176** 1097–1163. MR4087490 https://doi.org/10.1007/s00440-019-00936-y

[15] CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. and MELLY, B. (2013). Inference on counterfactual distributions. *Econometrica* **81** 2205–2268. MR3138546 https://doi.org/10.3982/ECTA10582

[16] CHERNOZHUKOV, V., IMBENS, G. W. and NEWEY, W. K. (2007). Instrumental variable estimation of nonseparable models. *J. Econometrics* **139** 4–14. MR2380756 https://doi.org/10.1016/j.jeconom.2006.06.002

[17] DEVROYE, L., GYÖRFI, L. and LUGOSI, G. (1996). *A Probabilistic Theory of Pattern Recognition*. *Applications of Mathematics* (*New York*) **31**. Springer, New York. MR1383093 https://doi.org/10.1007/978-1-4612-0711-5

[18] DUDLEY, R. M. (2014). *Uniform Central Limit Theorems*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **142**. Cambridge Univ. Press, New York. MR3445285

[19] EGGERMONT, P. P. B. and LARICCIA, V. N. (2009). *Maximum Penalized Likelihood Estimation*. *Volume II*: *Regression*. *Springer Series in Statistics*. Springer, Dordrecht. MR2817245 https://doi.org/10.1007/b12285

[20] GYÖRFI, L., KOHLER, M., KRZYŻAK, A. and WALK, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. Springer, Berlin.

[21] HALL, P. (1992). Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. *Ann. Statist.* **20** 675–694. MR1165587 https://doi.org/10.1214/aos/1176348651

[22] HALL, P. and KANG, K.-H. (2001). Bootstrapping nonparametric density estimators with empirically chosen bandwidths. *Ann. Statist.* **29** 1443–1468. MR1873338 https://doi.org/10.1214/aos/1013203461

[23] HUANG, J. Z. (2003). Local asymptotics for polynomial spline regression. *Ann. Statist.* **31** 1600–1635. MR2012827 https://doi.org/10.1214/aos/1065705120

[24] KALLENBERG, O. (2021). *Foundations of Modern Probability*, 3rd ed. *Probability Theory and Stochastic Modelling* **99**. Springer, Cham. MR4226142 https://doi.org/10.1007/978-3-030-61871-1

[25] KOENKER, R. (2005). *Quantile Regression*. *Econometric Society Monographs* **38**. Cambridge Univ. Press, Cambridge. MR2268657 https://doi.org/10.1017/CBO9780511754098

[26] 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 https://doi.org/10.1017/S0266466612000813

[27] LAI, P. Y. and LEE, S. M. S. (2005). An overview of asymptotic properties of L_p regression under general classes of error distributions. *J. Amer. Statist. Assoc.* **100** 446–458. MR2160549 https://doi.org/10.1198/016214504000001385

[28] LEPSKIĬ, O. V. (1992). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Teor. Veroyatn. Primen.* **36** 682–697. MR1147167 https://doi.org/10.1137/1136085

[29] LIN, W. and KULASEKERA, K. B. (2007). Identifiability of single-index models and additive-index models. *Biometrika* **94** 496–501. MR2380574 https://doi.org/10.1093/biomet/asm029

[30] MATZKIN, R. L. (2007). Nonparametric identification. *Handb. Econom.* **6** 5307–5368.

[31] MCCULLAGH, P. and NELDER, J. A. (2019). *Generalized Linear Models*. Routledge, London.

[32] MONRAD, D. and PHILIPP, W. (1991). Nearby variables with nearby conditional laws and a strong approximation theorem for Hilbert space valued martingales. *Probab. Theory Related Fields* **88** 381–404. MR1100898 https://doi.org/10.1007/BF01418867

[33] SHANG, Z. and CHENG, G. (2013). Local and global asymptotic inference in smoothing spline models. *Ann. Statist.* **41** 2608–2638. MR3161439 https://doi.org/10.1214/13-AOS1164

[34] STONE, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.* **8** 1348–1360. MR0594650

[35] STONE, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053. MR0673642

[36] 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 https://doi.org/10.1007/978-1-4757-2545-2

[37] YURINSKII, V. V. (1978). On the error of the Gaussian approximation for convolutions. *Theory Probab. Appl.* **22** 236–247.

[38] ZHANG, H. and SINGER, B. H. (2010). *Recursive Partitioning and Applications*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2674991 https://doi.org/10.1007/978-1-4419-6824-1

[39] ZHOU, S., SHEN, X. and WOLFE, D. A. (1998). Local asymptotics for regression splines and confidence regions. *Ann. Statist.* **26** 1760–1782. MR1673277 https://doi.org/10.1214/aos/1024691356

The Annals of Statistics
2026, Vol. 54, No. 4, 1768–1792
https://doi.org/10.1214/26-AOS2624

DIPMIND: DISTANCE PROFILE BASED MUTUAL INDEPENDENCE TESTING FOR RANDOM OBJECTS

BY YAQING CHEN[1,a] AND PAROMITA DUBEY[2,b1]

[1]*Department of Statistics, Rutgers University,* [a]*yqchen@stat.rutgers.edu*
[2]*Department of Data Sciences and Operations, Marshall School of Business, University of Southern California,* [b]*paromita@marshall.usc.edu*

This paper develops a novel unified framework for testing mutual independence among random objects residing in possibly different metric spaces. The framework generalizes existing methodologies and introduces new measures of mutual independence, and proposes associated tests that achieve minimax rate optimality and exhibit strong empirical power. The foundation of the proposed tests is the new concept of joint distance profiles, which uniquely characterize the joint law of random objects under a mild condition on either the joint law or the metric spaces. Our test statistics quantify the difference of the joint distance profiles of each data point with respect to the joint law and the product of marginal laws of the vector of random objects. To enhance power, we consider integrating this difference with respect to different measures and incorporate flexible data-adaptive weight profiles in the test statistics. We derive the limiting distribution of the test statistics under the null hypothesis of mutual independence and show that the proposed tests with certain weight profiles are asymptotically distribution-free if the marginal distance profiles are continuous. Furthermore, we establish the consistency of the tests under sequences of alternative hypotheses converging to the null. For practical implementations, we employ a permutation scheme to approximate the p-values and provide theoretical guarantees that the permutation-based tests maintain type I error control under the null and achieve consistency under alternatives. We demonstrate the power of the proposed tests across various types of data objects through simulations and real data applications, where the new tests exhibit better performance compared with popular existing approaches.

REFERENCES

AHLFORS, L. (1935). Zur Theorie der Überlagerungsflächen. *Acta Math*. **65** 157–194. MR1555403 https://doi.org/10.1007/BF02420945

BARABÁSI, A.-L. and ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286** 509–512. MR2091634 https://doi.org/10.1126/science.286.5439.509

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 https://doi.org/10.3150/13-BEJ514

BERRETT, T. B. and SAMWORTH, R. J. (2019). Nonparametric independence testing via mutual information. *Biometrika* **106** 547–566. MR3992389 https://doi.org/10.1093/biomet/asz024

BISWAS, M., SARKAR, S. and GHOSH, A. K. (2016). On some exact distribution-free tests of independence between two random vectors of arbitrary dimensions. *J. Statist. Plann. Inference* **175** 78–86. MR3492766 https://doi.org/10.1016/j.jspi.2016.02.007

BLUM, J. R., KIEFER, J. and ROSENBLATT, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist*. **32** 485–498. MR0125690 https://doi.org/10.1214/aoms/1177705055

CAI, Z., LEI, J. and ROEDER, K. (2024). Asymptotic distribution-free independence test for high-dimension data. *J. Amer. Statist. Assoc*. **119** 1794–1804. MR4797902 https://doi.org/10.1080/01621459.2023.2218030

CHEN, Y. and DUBEY, P. (2024). DiPMInd: Distance Profile based Mutual Independence testing for random objects. arXiv preprint. Available at arXiv:2412.06766v1.

MSC2020 subject classifications. Primary 62G10, 62G20; secondary 62G30.
Key words and phrases. Compositional data, distributional data, distribution-free inference, metric spaces, networks.

CHEN, Y. and DUBEY, P. (2026). Supplement to "DiPMInd: Distance profile based mutual independence testing for random objects." https://doi.org/10.1214/26-AOS2624SUPP

CHIQUET, J., DONNET, S. and BARBILLON, P. (2024). sbm: Stochastic blockmodels. R package version 0.4.7.

CSARDI, G. and NEPUSZ, T. (2006). The igraph software package for complex network research. *InterJournal Complex Syst*. **1695**.

DAVID, G., JOURNÉ, J.-L. and SEMMES, S. (1985). Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoam*. **1** 1–56. MR0850408 https://doi.org/10.4171/RMI/17

DEB, N. and SEN, B. (2023). Multivariate rank-based distribution-free nonparametric testing using measure transportation. *J. Amer. Statist. Assoc*. **118** 192–207. MR4571116 https://doi.org/10.1080/01621459.2021.1923508

DRTON, M., HAN, F. and SHI, H. (2020). High-dimensional consistent independence testing with maxima of rank correlations. *Ann. Statist*. **48** 3206–3227. MR4185806 https://doi.org/10.1214/19-AOS1926

DUBEY, P., CHEN, Y. and MÜLLER, H.-G. (2024). Metric statistics: Exploration and inference for random objects with distance profiles. *Ann. Statist*. **52** 757–792. MR4744195 https://doi.org/10.1214/24-aos2368

GAO, L., FAN, Y., LV, J. and SHAO, Q.-M. (2021). Asymptotic distributions of high-dimensional distance correlation inference. *Ann. Statist*. **49** 1999–2020. MR4319239 https://doi.org/10.1214/20-aos2024

GIESER, P. W. and RANDLES, R. H. (1997). A nonparametric test of independence between two vectors. *J. Amer. Statist. Assoc*. **92** 561–567. MR1467849 https://doi.org/10.2307/2965704

GRETTON, A., BOUSQUET, O., SMOLA, A. and SCHÖLKOPF, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory*. *Lecture Notes in Computer Science* **3734** 63–77. Springer, Berlin. MR2255909 https://doi.org/10.1007/11564089_7

HAN, F., CHEN, S. and LIU, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306 https://doi.org/10.1093/biomet/asx050

HEINONEN, J. (2001). *Lectures on Analysis on Metric Spaces*. *Universitext*. Springer, New York. MR1800917 https://doi.org/10.1007/978-1-4613-0131-8

HELLER, R. and HELLER, Y. (2016). Multivariate tests of association based on univariate tests. *Adv. Neural Inf. Process. Syst*. **29**.

HELLER, R., HELLER, Y. and GORFINE, M. (2013). A consistent multivariate test of association based on ranks of distances. *Biometrika* **100** 503–510. MR3068450 https://doi.org/10.1093/biomet/ass070

HOEFFDING, W. (1948). A non-parametric test of independence. *Ann. Math. Statist*. **19** 546–557. MR0029139 https://doi.org/10.1214/aoms/1177730150

HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw*. **5** 109–137. MR0718088 https://doi.org/10.1016/0378-8733(83)90021-7

JIN, Z. and MATTESON, D. S. (2018). Generalizing distance covariance to measure and test multivariate mutual dependence via complete and incomplete V-statistics. *J. Multivariate Anal*. **168** 304–322. MR3858367 https://doi.org/10.1016/j.jmva.2018.08.006

KLEBANOV, L. B. (2005). $\mathfrak{N}$-*Distances and Their Applications*. Charles Univ. in Prague, The Karolinum Press.

LAI, T., ZHANG, Z., WANG, Y. and KONG, L. (2021). Testing independence of functional variables by angle covariance. *J. Multivariate Anal*. **182** 104711. MR4187269 https://doi.org/10.1016/j.jmva.2020.104711

LYONS, R. (2013). Distance covariance in metric spaces. *Ann. Probab*. **41** 3284–3305. MR3127883 https://doi.org/10.1214/12-AOP803

MARRON, J. S. and ALONSO, A. M. (2014). Overview of object oriented data analysis. *Biom. J*. **56** 732–753. MR3258083 https://doi.org/10.1002/bimj.201300072

MARRON, J. S. and DRYDEN, I. L. (2021). *Object Oriented Data Analysis*. CRC Press, Boca Raton.

MOON, H. and CHEN, K. (2022). Interpoint-ranking sign covariance for the test of independence. *Biometrika* **109** 165–179. MR4374647 https://doi.org/10.1093/biomet/asab011

NELSEN, R. B. (2006). *An Introduction to Copulas*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2197664 https://doi.org/10.1007/s11229-005-3715-x

NIU, Z. and BHATTACHARYA, B. B. (2022). Distribution-free joint independence testing and robust independent component analysis using optimal transport. arXiv preprint. Available at arXiv:2211.15639.

PAN, W., WANG, X., ZHANG, H., ZHU, H. and ZHU, J. (2019). Ball covariance: A generic measure of dependence in Banach space. *J. Amer. Statist. Assoc*. **115** 307–317.

PETERSEN, A. and MÜLLER, H.-G. (2016). Functional data analysis for density functions by transformation to a Hilbert space. *Ann. Statist*. **44** 183–218. MR3449766 https://doi.org/10.1214/15-AOS1363

PFISTER, N., BÜHLMANN, P., SCHÖLKOPF, B. and PETERS, J. (2018). Kernel-based tests for joint independence. *J. R. Stat. Soc. Ser. B. Stat. Methodol*. **80** 5–31. MR3744710 https://doi.org/10.1111/rssb.12235

PFISTER, N. and PETERS, J. (2019). dHSIC: Independence Testing via Hilbert Schmidt Independence Criterion. R package version 2.1.

RIZZO, M. and SZEKELY, G. (2022). Energy: E-statistics: Multivariate inference via the energy of data. R package version 1.7-11.

SCHWEIZER, B. and WOLFF, E. F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9** 879–885. MR0619291

SEJDINOVIC, D., SRIPERUMBUDUR, B., GRETTON, A. and FUKUMIZU, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.* **41** 2263–2291. MR3127866 https://doi.org/10.1214/13-AOS1140

SHI, H., DRTON, M. and HAN, F. (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *J. Amer. Statist. Assoc.* **117** 395–410. MR4399094 https://doi.org/10.1080/01621459.2020.1782223

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 https://doi.org/10.1214/009053607000000505

WANG, X., ZHU, J., PAN, W., ZHU, J. and ZHANG, H. (2024). Nonparametric statistical inference via metric distribution function in metric spaces. *J. Amer. Statist. Assoc.* **119** 2772–2784. MR4833914 https://doi.org/10.1080/01621459.2023.2277417

WEIHS, L., DRTON, M. and MEINSHAUSEN, N. (2018). Symmetric rank covariances: A generalized framework for nonparametric measures of dependence. *Biometrika* **105** 547–562. MR3842884 https://doi.org/10.1093/biomet/asy021

ZHANG, Y. and ZHU, L. (2024). Projective independence tests in high dimensions: The curses and the cures. *Biometrika* **111** 1013–1027. MR4794156 https://doi.org/10.1093/biomet/asad070

ZHOU, H. and MÜLLER, H.-G. (2025). Association and independence test for random objects. arXiv preprint. Available at arXiv:2505.01983v1.

ZHU, C., YAO, J. and WANG, J.-L. (2024). Testing independence for sparse longitudinal data. *Biometrika* **111** 1187–1199. MR4829509 https://doi.org/10.1093/biomet/asae035

ZHU, L., XU, K., LI, R. and ZHONG, W. (2017). Projection correlation between two random vectors. *Biometrika* **104** 829–843. MR3737307 https://doi.org/10.1093/biomet/asx043

The Annals of Statistics
2026, Vol. 54, No. 4, 1793–1818
https://doi.org/10.1214/26-AOS2625

ESTIMATING THE FALSE DISCOVERY RATE OF VARIABLE SELECTION

BY YIXIANG LUO[1,a], WILLIAM FITHIAN[2,b] AND LIHUA LEI[3,c]

[1]*Department of Mathematics, University of California, Berkeley,* [a]*yixiangluo@berkeley.edu*
[2]*Department of Statistics, University of California, Berkeley,* [b]*wfithian@berkeley.edu*
[3]*Graduate School of Business, Stanford University,* [c]*lihualei@stanford.edu*

We introduce a generic estimator for the false discovery rate of any model selection procedure, in common statistical modeling settings including the Gaussian linear model, Gaussian graphical model, and model-X setting. We prove that our method has a conservative (nonnegative) bias in finite samples under standard statistical assumptions, and provide a bootstrap method for assessing its standard error. For methods like the Lasso, forward-stepwise regression, and the graphical Lasso, our estimator serves as a valuable companion to cross-validation, illuminating the tradeoff between prediction error and variable selection accuracy as a function of the model complexity parameter.

REFERENCES

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-AOS1337

BAUMANN, K. (2003). Cross-validation as the objective function for variable-selection techniques. *TrAC, Trends Anal. Chem.* **22** 395–406.

BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B, Methodol.* **57** 289–300. MR1325392

BENJAMINI, Y. and YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29** 1165–1188. MR1869245 https://doi.org/10.1214/aos/1013699998

BLACKWELL, D. (1947). Conditional expectation and unbiased sequential estimation. *Ann. Math. Stat.* **18** 105–110. MR0019903 https://doi.org/10.1214/aoms/1177730497

CANDÈS, E., FAN, Y., JANSON, 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 https://doi.org/10.1111/rssb.12265

DRTON, M. and PERLMAN, M. D. (2007). Multiple testing and error control in Gaussian graphical model selection. *Statist. Sci.* **22** 430–449. MR2416818 https://doi.org/10.1214/088342307000000113

DUNNETT, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50** 1096–1121.

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 https://doi.org/10.1214/009053604000000067

FITHIAN, W. and LEI, L. (2022). Conditional calibration for false discovery rate control under dependence. *Ann. Statist.* **50** 3091–3118. MR4524490 https://doi.org/10.1214/21-aos2137

FRIEDMAN, J., HASTIE, T. and TIBSHIRANI, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9** 432–441.

GOEMAN, J. J. and SOLARI, A. (2011). Multiple testing for exploratory research. *Statist. Sci.* **26** 584–597. MR2951390 https://doi.org/10.1214/11-STS356

HOCKING, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics* **32** 1–49. MR0398008 https://doi.org/10.2307/2529336

HUANG, D. and JANSON, L. (2020). Relaxing the assumptions of knockoffs by conditioning. *Ann. Statist.* **48** 3021–3042. MR4152633 https://doi.org/10.1214/19-AOS1920

LEHMANN, E. L. and SCHEFFÉ, H. (1950). Completeness, similar regions, and unbiased estimation. I. *Sankhyā* **10** 305–340. MR0039201 https://doi.org/10.1007/978-1-4614-1412-4_23

MSC2020 subject classifications. Primary 62H12; secondary 62J07.
Key words and phrases. False discovery rate, Lasso, graphical Lasso, multiple hypotheses testing.

LEI, J. (2019). Fast exact conformalization of the lasso using piecewise linear homotopy. *Biometrika* **106** 749–764. MR4031197 https://doi.org/10.1093/biomet/asz046

LI, X. and FITHIAN, W. (2021). Whiteout: when do fixed-X knockoffs fail? arXiv preprint. Available at arXiv:2107.06388.

LUO, Y., FITHIAN, W. and LEI, L. (2026). Supplement to "Estimating the false discovery rate of variable selection." https://doi.org/10.1214/26-AOS2625SUPPA, https://doi.org/10.1214/26-AOS2625SUPPB, https://doi.org/10.1214/26-AOS2625SUPPC

MEINSHAUSEN, N. and BÜHLMANN, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. MR2278363 https://doi.org/10.1214/009053606000000281

RADHAKRISHNA RAO, C. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37** 81–91. MR0015748

RHEE, S.-Y., FESSEL, W. J., ZOLOPA, A. R., HURLEY, L., LIU, T., TAYLOR, J., NGUYEN, D. P., SLOME, S., KLEIN, D. et al. (2005). HIV-1 protease and reverse-transcriptase mutations: Correlations with antiretroviral therapy in subtype B isolates and implications for drug-resistance surveillance. *J. Infect. Dis.* **192** 456–465.

RHEE, S.-Y., TAYLOR, J., WADHERA, G., BEN-HUR, A., BRUTLAG, D. L. and SHAFER, R. W. (2006). Genotypic predictors of human immunodeficiency virus type 1 drug resistance. *Proc. Natl. Acad. Sci. USA* **103** 17355–17360.

SACHS, K., PEREZ, O., PE'ER, D., LAUFFENBURGER, D. A. and NOLAN, G. P. (2005). Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308** 523–529.

STOREY, J. D. (2002). A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 479–498. MR1924302 https://doi.org/10.1111/1467-9868.00346

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B, Methodol.* **58** 267–288. MR1379242

YUAN, M. and LIN, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika* **94** 19–35. MR2367824 https://doi.org/10.1093/biomet/asm018

ZOU, H. and HASTIE, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 301–320. MR2137327 https://doi.org/10.1111/j.1467-9868.2005.00503.x

The Annals of Statistics
2026, Vol. 54, No. 4, 1819–1842
https://doi.org/10.1214/26-AOS2626

IMPROVED THRESHOLDS FOR E-VALUES

BY CHRISTOPHER BLIER-WONG[1,a] AND RUODU WANG[2,b]

[1]*Department of Statistical Sciences, University of Toronto,* [a]*christopher.blierwong@utoronto.ca*
[2]*Department of Statistics and Actuarial Science, University of Waterloo,* [b]*wang@uwaterloo.ca*

The rejection threshold used for e-values and e-processes is by default set to $1/\alpha$ for a guaranteed type-I error control at α, based on Markov's and Ville's inequalities. This threshold can be wasteful in practical applications. We discuss how this threshold can be improved under additional distributional assumptions on the e-values; some of these assumptions are naturally plausible and empirically observable, without knowing explicitly the form or model of the e-values. For small values of α, the threshold can roughly be improved (divided) by a factor of 2 for decreasing or unimodal densities, and by a factor of e for decreasing or unimodal-symmetric densities of log-transformed e-values. Moreover, we propose to use the supremum of comonotonic e-values, which is shown to preserve the type-I error guarantee. We also propose some preliminary methods to boost e-values in the e-BH procedure under some distributional assumptions while controlling the false discovery rate. Through a series of simulation studies, we demonstrate the effectiveness of our proposed methods in various testing scenarios, showing enhanced power.

REFERENCES

[1] BAGNOLI, M. and BERGSTROM, T. (2005). Log-concave probability and its applications. *Econom. Theory* **26** 445–469. MR2213177 https://doi.org/10.1007/s00199-004-0514-4
[2] 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-AOS1337
[3] BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B, Methodol.* **57** 289–300. MR1325392
[4] BERNARD, C., KAZZI, R. and VANDUFFEL, S. (2020). Range Value-at-Risk bounds for unimodal distributions under partial information. *Insurance Math. Econom.* **94** 9–24. MR4117414 https://doi.org/10.1016/j.insmatheco.2020.05.013
[5] BLIER-WONG, C. and WANG, R. (2026). Supplement to "Improved thresholds for e-values." https://doi.org/10.1214/26-AOS2626SUPP
[6] CANDÈS, E., FAN, Y., JANSON, 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 https://doi.org/10.1111/rssb.12265
[7] GRÜNWALD, P., DE HEIDE, R. and KOOLEN, W. (2024). Safe testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 1091–1128. MR4825000 https://doi.org/10.1093/jrsssb/qkae011
[8] GUAN, Y., JIAO, Z. and WANG, R. (2024). A reverse ES (CVaR) optimization formula. *N. Amer. Actuar. J.* **28** 611–625. MR4793867 https://doi.org/10.1080/10920277.2023.2249524
[9] HUBER, M. (2019). Halving the bounds for the Markov, Chebyshev, and Chernoff inequalities using smoothing. *Amer. Math. Monthly* **126** 915–927. MR4033226 https://doi.org/10.1080/00029890.2019.1656484
[10] IGNATIADIS, N., WANG, R. and RAMDAS, A. (2024). E-values as unnormalized weights in multiple testing. *Biometrika* **111** 417–439. MR4745574 https://doi.org/10.1093/biomet/asad057
[11] IGNATIADIS, N., WANG, R. and RAMDAS, A. (2025). Asymptotic and compound e-values: multiple testing and empirical Bayes. Available at arXiv:2409.19812.
[12] JEFFREYS, H. (1961). *Theory of Probability*, 3rd ed. *Oxford Classic Texts in the Physical Sciences*. Oxford Univ. Press, London. MR1647885
[13] LEE, J. and REN, Z. (2024). Boosting e-BH via conditional calibration. Available at arXiv:2404.17562.

MSC2020 subject classifications. Primary 62G10, 62F03; secondary 60E15.
Key words and phrases. Markov's inequality, e-variables, unimodality, calibrators, false discovery rate.

[14] LI, L., SHAO, H., WANG, R. and YANG, J. (2018). Worst-case range value-at-risk with partial information. *SIAM J. Financial Math.* **9** 190–218. MR3755662 https://doi.org/10.1137/17M1126138

[15] RAMDAS, A., GRÜNWALD, P., VOVK, V. and SHAFER, G. (2023). Game-theoretic statistics and safe anytime-valid inference. *Statist. Sci.* **38** 576–601. MR4665027 https://doi.org/10.1214/23-sts894

[16] RAMDAS, A. and WANG, R. (2025). Hypothesis testing with E-values. *Found. Trends Stat.* **1** 1–390.

[17] REN, Z. and BARBER, R. F. (2024). Derandomised knockoffs: Leveraging e-values for false discovery rate control. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 122–154. MR4754084 https://doi.org/10.1093/jrsssb/qkad085

[18] SHAFER, G. (2021). The language of betting as a strategy for statistical and scientific communication. *J. Roy. Statist. Soc., Ser. A* **184** 407–431. MR4255905 https://doi.org/10.1111/rssa.12647

[19] 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 https://doi.org/10.1111/j.1467-9868.2011.01034.x

[20] SHAKED, M. and SHANTHIKUMAR, J. G. (2007). *Stochastic Orders. Springer Series in Statistics*. Springer, New York. MR2265633 https://doi.org/10.1007/978-0-387-34675-5

[21] TSE, T. and DAVISON, A. C. (2022). A note on universal inference. *Stat* **11** Paper No. e501, 9. MR4529724 https://doi.org/10.1002/sta4.501

[22] VOVK, V. and WANG, R. (2021). E-values: Calibration, combination and applications. *Ann. Statist.* **49** 1736–1754. MR4298879 https://doi.org/10.1214/20-aos2020

[23] VOVK, V. and WANG, R. (2024). Merging sequential e-values via martingales. *Electron. J. Stat.* **18** 1185–1205. MR4718470 https://doi.org/10.1214/24-ejs2228

[24] VOVK, V. and WANG, R. (2024). Nonparametric e-tests of symmetry. *New Engl. J. Stat. Data Sci.* **2** 261–270.

[25] WANG, R. (2024). Testing with p*-values: Between p-values, mid p-values, and e-values. *Bernoulli* **30** 1313–1346. MR4699554 https://doi.org/10.3150/23-bej1633

[26] WANG, R. and RAMDAS, A. (2022). False discovery rate control with e-values. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 822–852. MR4460577 https://doi.org/10.1111/rssb.12489

[27] WASSERMAN, L., RAMDAS, A. and BALAKRISHNAN, S. (2020). Universal inference. *Proc. Natl. Acad. Sci. USA* **117** 16880–16890. MR4242731 https://doi.org/10.1073/pnas.1922664117

[28] WAUDBY-SMITH, I. and RAMDAS, A. (2024). Estimating means of bounded random variables by betting. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 1–27. MR4716192 https://doi.org/10.1093/jrsssb/qkad009

The Annals of Statistics
2026, Vol. 54, No. 4, 1843–1869
https://doi.org/10.1214/26-AOS2627

MEASURING EVIDENCE AGAINST EXCHANGEABILITY AND GROUP INVARIANCE WITH E-VALUES

BY NICK W. KONING[a]

Econometric Institute, Erasmus University Rotterdam, [a]*n.w.koning@ese.eur.nl*

We study e-values for quantifying evidence against exchangeability and general invariance of a random variable under a compact group. We start by characterizing such e-values, and explaining how they nest traditional group invariance tests as a special case. We show they can be easily designed for an arbitrary test statistic, and computed through Monte Carlo sampling. We prove a result that characterizes optimal e-values for group invariance against optimality targets that satisfy a mild orbit-wise decomposition property. We apply this to design expected-utility-optimal e-values for group invariance, which include both Neyman–Pearson-optimal tests and log-optimal e-values. Moreover, we generalize the notion of rank- and sign-based testing to compact groups, by using a representative inversion kernel. In addition, we characterize e-processes for group invariance for arbitrary filtrations, and provide tools to construct them. We also describe test martingales under a natural filtration, which are simpler to construct. Peeking beyond compact groups, we encounter e-values and e-processes based on ergodic theorems. These nest e-processes based on de Finetti's theorem for testing exchangeability.

REFERENCES

ANGELOPOULOS, A. N., BARBER, R. F. and BATES, S. (2025). Theoretical Foundations of Conformal Prediction. arXiv preprint. Available at arXiv:2411.11824.

CHIU, K. and BLOEM-REDDY, B. (2023). Non-parametric Hypothesis Tests for Distributional Group Symmetry. arXiv preprint. Available at arXiv:2307.15834.

CHUNG, J. H. and FRASER, D. A. S. (1958). Randomization tests for a multivariate two-sample problem. *J. Amer. Statist. Assoc.* **53** 729–735.

DE LA PEÑA, V. H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** 537–564. MR1681153 https://doi.org/10.1214/aop/1022677271

EATON, M. L. (1989). *Group Invariance Applications in Statistics. NSF-CBMS Regional Conference Series in Probability and Statistics* **1**. IMS, Hayward, CA. MR1089423

FISCHER, L. and RAMDAS, A. (2025). Sequential Monte Carlo testing by betting. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **87** 1200–1220. MR4961268 https://doi.org/10.1093/jrsssb/qkaf014

GILOVICH, T., VALLONE, R. and TVERSKY, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cogn. Psychol.* **17** 295–314.

GRÜNWALD, P., DE HEIDE, R. and KOOLEN, W. (2024). Safe testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 1091–1128. MR4825000 https://doi.org/10.1093/jrsssb/qkae011

GRÜNWALD, P. D. (2024). Beyond Neyman-Pearson: E-values enable hypothesis testing with a data-driven alpha. *Proc. Natl. Acad. Sci. USA* **121** Paper No. e2302098121. MR4826968 https://doi.org/10.1073/pnas.2302098121

HEMERIK, J. and GOEMAN, J. (2018). Exact testing with random permutations. *TEST* **27** 811–825. MR3878362 https://doi.org/10.1007/s11749-017-0571-1

HOWARD, S. R., RAMDAS, A., MCAULIFFE, J. and SEKHON, J. (2021). Time-uniform, nonparametric, nonasymptotic confidence sequences. *Ann. Statist.* **49** 1055–1080. MR4255119 https://doi.org/10.1214/20-aos1991

IGNATIADIS, N., WANG, R. and RAMDAS, A. (2024). *E*-values as unnormalized weights in multiple testing. *Biometrika* **111** 417–439. MR4745574 https://doi.org/10.1093/biomet/asad057

KALLENBERG, O. (2011). Invariant Palm and related disintegrations via skew factorization. *Probab. Theory Related Fields* **149** 279–301. MR2773033 https://doi.org/10.1007/s00440-009-0254-2

MSC2020 subject classifications. 62G10, 62L10.

Key words and phrases. Permutation test, e-values, sequential testing, group invariance test, anytime valid inference, post-hoc valid inference.

KALLENBERG, O. (2017). *Random Measures, Theory and Applications*. *Probability Theory and Stochastic Modelling* **77**. Springer, Cham. MR3642325 https://doi.org/10.1007/978-3-319-41598-7

KALLENBERG, O. (2021). *Foundations of Modern Probability*. *Probability Theory and Stochastic Modelling* **99**. Springer, Cham. MR4226142 https://doi.org/10.1007/978-3-030-61871-1

KASHLAK, A. B. (2022). Asymptotic Invariance and Robustness of Randomization Tests. arXiv Preprint. Available at arXiv:2211.00144.

KONING, N. W. (2024a). Continuous Testing: Unifying Tests and E-values. arXiv Preprint. Available at arXiv: 2409.05654.

KONING, N. W. (2024b). Post-hoc α hypothesis testing and post-hoc p-values. arXiv Preprint. Available at arXiv: 2312.08040.

KONING, N. W. (2024c). More power by using fewer permutations. *Biometrika* **111** 1405–1412. MR4829514 https://doi.org/10.1093/biomet/asae031

KONING, N. W. (2026). Supplement to "Measuring evidence against exchangeability and group invariance with e-values." https://doi.org/10.1214/26-AOS2627SUPPA, https://doi.org/10.1214/26-AOS2627SUPPB

KONING, N. W. and HEMERIK, J. (2024). More efficient exact group invariance testing: Using a representative subgroup. *Biometrika* **111** 441–458. MR4745575 https://doi.org/10.1093/biomet/asad050

KONING, N. W. and VAN MEER, S. (2026). Anytime validity is free: Inducing sequential tests. *J. R. Stat. Soc. Ser. B. Stat. Methodol*. https://doi.org/10.1093/jrsssb/qkag050

LARDY, T., GRÜNWALD, P. and HARREMOËS, P. (2024). Reverse information projections and optimal e-statistics. *IEEE Trans. Inf. Theory* **70** 7616–7631. MR4818350 https://doi.org/10.1109/tit.2024.3444458

LARDY, T. and PÉREZ-ORTIZ, M. F. (2024). Anytime-Valid Tests of Group Invariance through Conformal Prediction. arXiv Preprint. Available at arXiv:2401.15461.

LARSSON, M., RAMDAS, A. and RUF, J. (2025). The numeraire e-variable and reverse information projection. *Ann. Statist.* **53** 1015–1043. MR4925114 https://doi.org/10.1214/24-aos2487

LEHMANN, E. L. and STEIN, C. (1949). On the theory of some nonparametric hypotheses. *Ann. Math. Statist.* **20** 28–45. MR0030168

MILLER, J. B. and SANJURJO, A. (2018). Surprised by the hot hand fallacy? A truth in the law of small numbers. *Econometrica* **86** 2019–2047. MR3901245 https://doi.org/10.3982/ECTA14943

PÉREZ-ORTIZ, M. F., LARDY, T., DE HEIDE, R. and GRÜNWALD, P. D. (2024). e-statistics, group invariance and anytime-valid testing. *Ann. Statist.* **52** 1410–1432. MR4804814 https://doi.org/10.1214/24-aos2394

RAMDAS, A., BARBER, R. F., CANDÈS, E. J. and TIBSHIRANI, R. J. (2023a). Permutation tests using arbitrary permutation distributions. *Sankhya A* **85** 1156–1177. MR4619064 https://doi.org/10.1007/s13171-023-00308-8

RAMDAS, A., GRÜNWALD, P., VOVK, V. and SHAFER, G. (2023b). Game-theoretic statistics and safe anytime-valid inference. *Statist. Sci.* **38** 576–601. MR4665027 https://doi.org/10.1214/23-sts894

RAMDAS, A., RUF, J., LARSSON, M. and KOOLEN, W. (2022a). Admissible anytime-valid sequential inference must rely on nonnegative martingales. arXiv preprint. Available at arXiv:2009.03167.

RAMDAS, A., RUF, J., LARSSON, M. and KOOLEN, W. M. (2022b). Testing exchangeability: Fork-convexity, supermartingales and e-processes. *Internat. J. Approx. Reason.* **141** 83–109. MR4364897 https://doi.org/10.1016/j.ijar.2021.06.017

RITZWOLLER, D. M. and ROMANO, J. P. (2022). Uncertainty in the hot hand fallacy: Detecting streaky alternatives to random Bernoulli sequences. *Rev. Econ. Stud.* **89** 976–1007. MR4409195 https://doi.org/10.1093/restud/rdab020

SAHA, A. and RAMDAS, A. (2024). Testing exchangeability by pairwise betting. In *International Conference on Artificial Intelligence and Statistics* 4915–4923. PMLR.

SHAFER, G. (2021). Testing by betting: A strategy for statistical and scientific communication. *J. Roy. Stat. Soc. Ser. A* **184** 407–478. MR4255905 https://doi.org/10.1111/rssa.12647

SHAFER, G. and VOVK, V. (2008). A tutorial on conformal prediction. *J. Mach. Learn. Res.* **9** 371–421. MR2417240

STOEPKER, I. V. and CASTRO, R. M. (2024). Inference with Sequential Monte-Carlo Computation of p-values: Fast and Valid Approaches. arXiv Preprint. Available at arXiv:2409.18908.

VOVK, V. (2021). Testing randomness online. *Statist. Sci.* **36** 595–611. MR4323055 https://doi.org/10.1214/20-sts817

VOVK, V. (2025). Testing exchangeability in the batch mode with e-values and Markov alternatives. *Mach. Learn.* **114** Paper No. 99. MR4868057 https://doi.org/10.1007/s10994-024-06720-x

VOVK, V. and WANG, R. (2021). E-values: Calibration, combination and applications. *Ann. Statist.* **49** 1736–1754. MR4298879 https://doi.org/10.1214/20-aos2020

VOVK, V. and WANG, R. (2024). Nonparametric e-tests of symmetry. *N. Engl. J. Stat. Data Sci.* **2** 261–270.

WANG, R. and RAMDAS, A. (2022). False discovery rate control with e-values. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 822–852. MR4460577 https://doi.org/10.1111/rssb.12489

The Annals of Statistics
2026, Vol. 54, No. 4, 1870–1894
https://doi.org/10.1214/26-AOS2628
This research was funded, in whole or in part, by UKRI, EP/X027872/1; EP/X028100/1; EP/X028712/1; EP/X028119/1.

PRIVACY GUARANTEES IN POSTERIOR SAMPLING UNDER CONTAMINATION

BY SHENGGANG HU[1,a], LOUIS ASLETT[2,c], HONGSHENG DAI[3,d], MURRAY POLLOCK[3,e] AND GARETH O. ROBERTS[1,b]

[1]*Department of Statistics, University of Warwick,* [a]*shenggang.hu@warwick.ac.uk,* [b]*gareth.o.roberts@warwick.ac.uk*
[2]*Department of Mathematical Sciences, Durham University,* [c]*louis.aslett@durham.ac.uk*
[3]*School of Mathematics, Statistics and Physics, Newcastle University,* [d]*hongsheng.dai@newcastle.ac.uk,* [e]*murray.pollock@newcastle.ac.uk*

In recent years, differential privacy has been adopted by tech-companies and governmental agencies as the standard for measuring privacy in algorithms. In this article, we study differential privacy in Bayesian posterior sampling settings. We begin by considering differential privacy in the most common privatisation setting in which Laplace or Gaussian noise is injected into the output. In an effort to achieve better differential privacy, we consider adopting *Huber's contamination model* for use within privacy settings, and replace at random data points with samples from a heavy-tailed distribution (*instead* of injecting noise into the output). We derive bounds for the differential privacy level (ϵ, δ) of our approach, without requiring bounded observation and parameter spaces, a restriction commonly imposed in the literature. We further consider for our approach the effect of sample size on the privacy level and the rate at which (ϵ, δ) converges to zero. Asymptotically, our contamination approach is fully private with no information loss. We also provide examples of inference models for which our approach applies, with theoretical convergence rate analysis and simulation studies.

REFERENCES

ABOWD, J. M. (2018). The US Census Bureau adopts differential privacy. In *Proceedings of the* 24*th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* 2867–2867.

AKTAY, A., BAVADEKAR, S., COSSOUL, G., DAVIS, J., DESFONTAINES, D., FABRIKANT, A., GABRILOVICH, E., GADEPALLI, K., GIPSON, B. et al. (2020). Google COVID-19 community mobility reports: Anonymization process description (version 1.1). arXiv Preprint. Available at arXiv:2004.04145.

ANDRIEU, C., LEE, A., POWER, S. and WANG, A. Q. (2022). Comparison of Markov chains via weak Poincaré inequalities with application to pseudo-marginal MCMC. *Ann. Statist.* **50** 3592–3618. MR4524509 https://doi.org/10.1214/22-aos2241

ANDRIEU, C., LEE, A., POWER, S. and WANG, A. Q. (2026). Weak Poincaré inequalities for Markov chains: Theory and applications. *Ann. Appl. Probab.* **36** 46–107. MR5033571 https://doi.org/10.1214/25-AAP2185

APPLE (2017). Learning with Privacy at Scale. Available at https://docs-assets.developer.apple.com/ml-research/papers/learning-with-privacy-at-scale.pdf. [Online; accessed 13-September-2023].

ARTIN, E. (1964). *The Gamma Function. Athena Series: Selected Topics in Mathematics*. Holt, Rinehart and Winston, New York-Toronto-London. MR0165148

BERNSTEIN, G. and SHELDON, D. R. (2019). Differentially private Bayesian linear regression. *Adv. Neural Inf. Process. Syst.* **32**.

BISWAS, S., DONG, Y., KAMATH, G. and ULLMAN, J. (2020). Coinpress: Practical private mean and covariance estimation. *Adv. Neural Inf. Process. Syst.* **33** 14475–14485.

BUN, M. and STEINKE, T. (2016a). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography. Part I. Lecture Notes in Computer Science* **9985** 635–658. Springer, Berlin. MR3591832 https://doi.org/10.1007/978-3-662-53641-4_24

MSC2020 subject classifications. Primary 62F15; secondary 62J12.
Key words and phrases. Huber's contamination model, posterior sampling, Bayesian inference, differential privacy.

CAI, T. T., WANG, Y. and ZHANG, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *Ann. Statist.* **49** 2825–2850. MR4338894 https://doi.org/10.1214/21-aos2058

COVINGTON, C., HE, X., HONAKER, J. and KAMATH, G. (2025). Unbiased statistical estimation and valid confidence intervals under differential privacy. *Statist. Sinica* **35** 651–670. MR4861984 https://doi.org/10.5705/ss.202022.0276

DIMITRAKAKIS, C., NELSON, B., ZHANG, Z., MITROKOTSA, A. and RUBINSTEIN, B. I. P. (2017). Differential privacy for Bayesian inference through posterior sampling. *J. Mach. Learn. Res.* **18** Paper No. 11. MR3634878

DING, B., KULKARNI, J. and YEKHANIN, S. (2017). Collecting telemetry data privately. *Adv. Neural Inf. Process. Syst.* **30**.

DUCHI, J., HAQUE, S. and KUDITIPUDI, R. (2023). A fast algorithm for adaptive private mean estimation. arXiv Preprint. Available at arXiv:2301.07078.

DURMUS, A. and MOULINES, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479 https://doi.org/10.1214/16-AAP1238

DWORK, C., MCSHERRY, F., NISSIM, K. and SMITH, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. *Lecture Notes in Computer Science* **3876** 265–284. Springer, Berlin. MR2241676 https://doi.org/10.1007/11681878_14

DWORK, C. and ROTH, A. (2013). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* **9** 211–487. MR3254020 https://doi.org/10.1561/0400000042

DWORK, C. and ROTHBLUM, G. N. (2016). Concentrated differential privacy. arXiv preprint. Available at arXiv:1603.01887.

ERLINGSSON, Ú., PIHUR, V. and KOROLOVA, A. (2014). Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the* 2014 *ACM SIGSAC Conference on Computer and Communications Security* 1054–1067.

FOULDS, J., GEUMLEK, J., WELLING, M. and CHAUDHURI, K. (2016). On the theory and practice of privacy-preserving Bayesian data analysis. In *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*. *UAI'*16 192–201. AUAI Press, Arlington, VA, USA.

GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge Univ. Press, Cambridge. MR3587782 https://doi.org/10.1017/9781139029834

HALL, R., WASSERMAN, L. and RINALDO, A. (2013). Random differential privacy. *J. Priv. Confid.* **4**.

HEIKKILÄ, M., JÄLKÖ, J., DIKMEN, O. and HONKELA, A. (2019). Differentially private Markov chain Monte Carlo. *Adv. Neural Inf. Process. Syst.* **32**.

HU, S., ASLETT, L., DAI, H., POLLOCK, M. and ROBERTS, G. O. (2026). Supplement to "Privacy guarantees in posterior sampling under contamination." https://doi.org/10.1214/26-AOS2628SUPP

HUANG, Z., LIANG, Y. and YI, K. (2021). Instance-optimal mean estimation under differential privacy. *Adv. Neural Inf. Process. Syst.* **34** 25993–26004.

HUBER, P. J. (2004). *Robust Statistics* **523**. Wiley, New York.

KROLL, M. (2021). On density estimation at a fixed point under local differential privacy. *Electron. J. Stat.* **15** 1783–1813. MR4255316 https://doi.org/10.1214/21-ejs1830

KULKARNI, T., JÄLKÖ, J., KOSKELA, A., KASKI, S. and HONKELA, A. (2021). Differentially private Bayesian inference for generalized linear models. In *International Conference on Machine Learning* 5838–5849. PMLR.

LI, M., BERRETT, T. B. and YU, Y. (2023). On robustness and local differential privacy. *Ann. Statist.* **51** 717–737. MR4600999 https://doi.org/10.1214/23-aos2267

MACHANAVAJJHALA, A., KIFER, D., ABOWD, J., GEHRKE, J. and VILHUBER, L. (2008). Privacy: Theory meets practice on the map. In 2008 *IEEE* 24*th International Conference on Data Engineering* 277–286. IEEE Press, New York.

MINAMI, K., ARAI, H., SATO, I. and NAKAGAWA, H. (2016). Differential privacy without sensitivity. *Adv. Neural Inf. Process. Syst.* **29**.

MU, W. and XIONG, S. (2023). On Huber's contaminated model. *J. Complexity* **77** Paper No. 101745. MR4560511 https://doi.org/10.1016/j.jco.2023.101745

MURPHY, K. P. (2022). *Probabilistic Machine Learning*: *An Introduction*. MIT Press, Cambridge.

PEREIRA, M., KIM, A., ALLEN, J., WHITE, K., FERRES, J. L. and DODHIA, R. (2021). US broadband coverage data set: a differentially private data release. arXiv Preprint. Available at arXiv:2103.14035.

ROBERT, C. P. and CASELLA, G. (1999). *Monte Carlo Statistical Methods*. *Springer Texts in Statistics*. Springer, New York. MR1707311 https://doi.org/10.1007/978-1-4757-3071-5

ROHDE, A. and STEINBERGER, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.* **48** 2646–2670. MR4152116 https://doi.org/10.1214/19-AOS1901

SHEFFET, O. (2019). Old techniques in differentially private linear regression. In *Algorithmic Learning Theory* 2019. *Proc. Mach. Learn. Res.* (*PMLR*) **98** [788–826]. Proceedings of Machine Learning Research PMLR. MR3932869

SHEN, X. and WASSERMAN, L. (2001). Rates of convergence of posterior distributions. *Ann. Statist.* **29** 687–714. MR1865337 https://doi.org/10.1214/aos/1009210686

SHEN, X. and WONG, W. H. (1994). Convergence rate of sieve estimates. *Ann. Statist.* **22** 580–615. MR1292531 https://doi.org/10.1214/aos/1176325486

TRIASTCYN, A. and FALTINGS, B. (2019). Federated learning with Bayesian differential privacy. In 2019 *IEEE International Conference on Big Data* (*Big Data*) 2587–2596. IEEE Press, New York.

TRIASTCYN, A. and FALTINGS, B. (2020). Bayesian differential privacy for machine learning. In *International Conference on Machine Learning* 9583–9592. PMLR.

WANG, Y.-X., FIENBERG, S. and SMOLA, A. (2015). Privacy for free: Posterior sampling and stochastic gradient Monte Carlo. In *International Conference on Machine Learning* 2493–2502. PMLR.

WASSERMAN, L. (2004). *All of Statistics*: *A Concise Course in Statistical Inference*. *Springer Texts in Statistics*. Springer, New York. MR2055670 https://doi.org/10.1007/978-0-387-21736-9

WASSERMAN, L. (2020). Lecture note 15, intermediate statistics, fall 2020. [Online; accessed 19-October-2023].

WONG, W. H. and SHEN, X. (1995). Probability inequalities for likelihood ratios and convergence rates of sieve MLEs. *Ann. Statist.* **23** 339–362. MR1332570 https://doi.org/10.1214/aos/1176324524

YILDIRIM, S. and ERMIŞ, B. (2019). Exact MCMC with differentially private moves: Revisiting the penalty algorithm in a data privacy framework. *Stat. Comput.* **29** 947–963. MR3994611 https://doi.org/10.1007/s11222-018-9847-x

ZHANG, W. and ZHANG, R. (2023). DP-fast MH: Private, fast, and accurate Metropolis–Hastings for large-scale Bayesian inference. In *International Conference on Machine Learning* 41847–41860. PMLR.

ZHANG, Z., RUBINSTEIN, B. and DIMITRAKAKIS, C. (2016). On the differential privacy of Bayesian inference. In *Proceedings of the AAAI Conference on Artificial Intelligence* **30**.

The Annals of Statistics
2026, Vol. 54, No. 4, 1895–1922
https://doi.org/10.1214/26-AOS2630

UNBIASED KINETIC LANGEVIN MONTE CARLO WITH INEXACT GRADIENTS

BY NEIL K. CHADA[1,a], BENEDICT LEIMKUHLER[2,b], DANIEL PAULIN[3,c] AND PETER A. WHALLEY[4,d]

[1]*Department of Mathematics, City University of Hong Kong,* [a]*neilchada123@gmail.com*
[2]*School of Mathematics, University of Edinburgh,* [b]*b.leimkuhler@ed.ac.uk*
[3]*School of Physical and Mathematical Sciences, Nanyang Technological University,* [c]*paulindani@gmail.com*
[4]*Seminar for Statistics, ETH Zurich,* [d]*peter.a.whalley@gmail.com*

We present an unbiased method for Bayesian posterior means based on kinetic Langevin dynamics that combines advanced splitting methods with enhanced gradient approximations. Our approach avoids Metropolis correction by coupling Markov chains at different discretization levels in a multilevel Monte Carlo approach. Theoretical analysis demonstrates that our proposed estimator is unbiased, attains finite variance, and satisfies a central limit theorem. It can achieve accuracy $\epsilon > 0$ for estimating expectations of Lipschitz functions in d dimensions with $\mathcal{O}(d^{1/4}\epsilon^{-2})$ expected gradient evaluations, without assuming warm start. We exhibit similar bounds using both approximate and stochastic gradients, and our method's computational cost is shown to scale independently of the size of the dataset. The proposed method is tested using a multinomial regression problem on the MNIST dataset and a Poisson regression model for soccer scores. Experiments indicate that the number of gradient evaluations per effective sample is independent of dimension, even when using inexact gradients. For product distributions, we give dimension-independent variance bounds. Our results demonstrate that in large-scale applications, the unbiased algorithm we present can be 2–3 orders of magnitude more efficient than the "gold-standard" randomized Hamiltonian Monte Carlo.

REFERENCES

[1] ALAMO, A. and SANZ-SERNA, J. M. (2016). A technique for studying strong and weak local errors of splitting stochastic integrators. *SIAM J. Numer. Anal.* **54** 3239–3257. MR3570281 https://doi.org/10.1137/16M1058765

[2] ÁLAMO ZAPATERO, A. (2021). Word series for the numerical integration of stochastic differential equations. Ph.D. thesis, Universidad de Valladolid.

[3] APERS, S., GRIBLING, S. and SZILÁGYI, D. (2024). Hamiltonian Monte Carlo for efficient Gaussian sampling: Long and random steps. *J. Mach. Learn. Res.* **25** 348. MR4829192

[4] BAKER, J., FEARNHEAD, P., FOX, E. B. and NEMETH, C. (2019). Control variates for stochastic gradient MCMC. *Stat. Comput.* **29** 599–615. MR3969063 https://doi.org/10.1007/s11222-018-9826-2

[5] BESAG, J. E. (1994). Comments on 'Representations of knowledge in complex systems' by U. Grenander and M.I. Miller. *J. Roy. Statist. Soc. Ser. B, Methodol.* **56** 591–592.

[6] BESKOS, A., PILLAI, N., ROBERTS, G., SANZ-SERNA, J.-M. and STUART, A. (2013). Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli* **19** 1501–1534. MR3129023 https://doi.org/10.3150/12-BEJ414

[7] BIERKENS, J., FEARNHEAD, P. and ROBERTS, G. (2019). The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *Ann. Statist.* **47** 1288–1320. MR3911113 https://doi.org/10.1214/18-AOS1715

[8] BOND, S. D. and LEIMKUHLER, B. J. (2007). Molecular dynamics and the accuracy of numerically computed averages. *Acta Numer.* **16** 1–65. MR2417927 https://doi.org/10.1017/S0962492906280012

MSC2020 subject classifications. 65C05, 65C30, 65C40, 62F15.
Key words and phrases. Unbiased estimation, kinetic Langevin dynamics, multilevel Monte Carlo, stochastic gradient.

[9] BOU-RABEE, N. and OWHADI, H. (2010). Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.* **48** 278–297. MR2608370 https://doi.org/10.1137/090758842

[10] BOU-RABEE, N. and SANZ-SERNA, J. M. (2017). Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **27** 2159–2194. MR3693523 https://doi.org/10.1214/16-AAP1255

[11] BRADBURY, J., FROSTIG, R., HAWKINS, P., JOHNSON, M. J., LEARY, C., MACLAURIN, D., NECULA, G., PASZKE, A., VANDERPLAS, J. et al. (2018). JAX: Composable transformations of Python+NumPy programs. Available at http://github.com/jax-ml/jax.

[12] BROSSE, N., DURMUS, A. and MOULINES, E. (2018). The promises and pitfalls of stochastic gradient Langevin dynamics. *Adv. Neural Inf. Process. Syst.* 8268–8278.

[13] BRÜNGER, A., BROOKS, C. L. and KARPLUS, M. (1984). Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.* **105** 495–500.

[14] CAO, Y., LU, J. and WANG, L. (2023). On explicit L^2-convergence rate estimate for underdamped Langevin dynamics. *Arch. Ration. Mech. Anal.* **247** 90. MR4632836 https://doi.org/10.1007/s00205-023-01922-4

[15] CHADA, N. K., FRANKS, J., JASRA, A., LAW, K. J. and VIHOLA, M. (2021). Unbiased inference for discretely observed hidden Markov model diffusions. *SIAM/ASA J. Uncertain. Quantificat.* **9** 763–787. MR4269502 https://doi.org/10.1137/20M131549X

[16] CHADA, N. K., LEIMKUHLER, B., PAULIN, D. and WHALLEY, P. A. (2026). Supplement to "Unbiased kinetic Langevin Monte Carlo with inexact gradients." https://doi.org/10.1214/26-AOS2630SUPP

[17] CHAK, M., KANTAS, N., LELIÈVRE, T. and PAVLIOTIS, G. A. (2023). Optimal friction matrix for underdamped Langevin sampling. *ESAIM. Math. Model. Numer. Anal.* **57** 3335–3371. MR4671454 https://doi.org/10.1051/m2an/2023083

[18] CHAK, M. and MONMARCHÉ, P. (2023). Reflection coupling for unadjusted generalized Hamiltonian Monte Carlo in the nonconvex stochastic gradient case. arXiv preprint. Available at arXiv:2310.18774.

[19] CHEN, Y., CHENG, X., NILES-WEED, J. and WEARE, J. (2026). Convergence of unadjusted Langevin in high dimensions: Delocalization of bias. *Comm. Pure Appl. Math.* **79** 1467–1491. MR5058539 https://doi.org/10.1002/cpa.70032

[20] CHEN, Y. and GATMIRY, K. (2023). When does Metropolized Hamiltonian Monte Carlo provably outperform Metropolis-adjusted Langevin algorithm? arXiv preprint. Available at arXiv:2304.04724.

[21] CHENG, X., CHATTERJI, N. S., BARTLETT, P. L. and JORDAN, M. I. (2018). Underdamped Langevin MCMC: A non-asymptotic analysis. In *Conference on Learning Theory* 300–323. PMLR.

[22] CORENFLOS, A., SUTTON, M. and CHOPIN, N. (2025). Debiasing piecewise deterministic Markov process samplers using couplings. *Scand. J. Stat.* **52** 1932–1974. MR4986919 https://doi.org/10.1111/sjos.70015

[23] CORNISH, R., VANETTI, P., BOUCHARD-CÔTÉ, A., DELIGIANNIDIS, G. and DOUCET, A. (2019). Scalable Metropolis–Hastings for exact Bayesian inference with large datasets. In *International Conference on Machine Learning* 1351–1360. PMLR.

[24] DALALYAN, A. S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183

[25] DALALYAN, A. S. and RIOU-DURAND, L. (2020). On sampling from a log-concave density using kinetic Langevin diffusions. *Bernoulli* **26** 1956–1988. MR4091098 https://doi.org/10.3150/19-BEJ1178

[26] DICICCIO, T. J. and EFRON, B. (1996). Bootstrap confidence intervals. *Statist. Sci.* **11** 189–228. MR1436647 https://doi.org/10.1214/ss/1032280214

[27] DOUC, R., JACOB, P. E., LEE, A. and VATS, D. (2026). Solving the Poisson equation using coupled Markov chains. *Ann. Statist.* **54** 201–225. MR5039616 https://doi.org/10.1214/25-AOS2564

[28] DUANE, S., KENNEDY, A. D., PENDLETON, B. J. and ROWETH, D. (1987). Hybrid Monte Carlo. *Phys. Lett. B* **195** 216–222. MR3960671 https://doi.org/10.1016/0370-2693(87)91197-x

[29] DURMUS, A. and MOULINES, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479 https://doi.org/10.1214/16-AAP1238

[30] DWIVEDI, R., CHEN, Y., WAINWRIGHT, M. J. and YU, B. (2019). Log-concave sampling: Metropolis–Hastings algorithms are fast. *J. Mach. Learn. Res.* **20** 183. MR4048994

[31] EBERLE, A., GUILLIN, A. and ZIMMER, R. (2019). Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47** 1982–2010. MR3980913 https://doi.org/10.1214/18-AOP1299

[32] FOSTER, J., LYONS, T. and OBERHAUSER, H. (2021). The shifted ODE method for underdamped Langevin MCMC. arXiv preprint, arXiv:2101.03446.

[33] FOSTER, J. M., DOS REIS, G. and STRANGE, C. (2024). High order splitting methods for SDEs satisfying a commutativity condition. *SIAM J. Numer. Anal.* **62** 500–532. MR4706044 https://doi.org/10.1137/23M155147X

[34] GILES, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numer.* **24** 259–328. MR3349310 https://doi.org/10.1017/S096249291500001X

[35] GILES, M. B., MAJKA, M. B., SZPRUCH, L., VOLLMER, S. J. and ZYGALAKIS, K. C. (2020). Multi-level Monte Carlo methods for the approximation of invariant measures of stochastic differential equations. *Stat. Comput.* **30** 507–524. MR4065216 https://doi.org/10.1007/s11222-019-09890-0

[36] GLYNN, P. W. and RHEE, C.-H. (2014). Exact estimation for Markov chain equilibrium expectations. *J. Appl. Probab.* **51A** 377–389. MR3317370 https://doi.org/10.1239/jap/1417528487

[37] HENG, J. and JACOB, P. E. (2019). Unbiased Hamiltonian Monte Carlo with couplings. *Biometrika* **106** 287–302. MR3949304 https://doi.org/10.1093/biomet/asy074

[38] IZAGUIRRE, J. A., CATARELLO, D. P., WOZNIAK, J. M. and SKEEL, R. D. (2001). Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **114** 2090–2098.

[39] JACOB, P. E., O'LEARY, J. and ATCHADÉ, Y. F. (2020). Unbiased Markov chain Monte Carlo methods with couplings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 543–600. MR4112777 https://doi.org/10.1111/rssb.12336

[40] JOHNSON, R. and ZHANG, T. (2013). Accelerating stochastic gradient descent using predictive variance reduction. *Adv. Neural Inf. Process. Syst.* **26**.

[41] KAHALÉ, N. (2024). Unbiased time-average estimators for Markov chains. *Math. Oper. Res.* **49** 2136–2165. MR4838256 https://doi.org/10.1287/moor.2022.0326

[42] KOOPMAN, S. J. and LIT, R. (2015). A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League. *J. R. Stat. Soc. Ser. A. Stat. Soc.* **178** 167–186. MR3291766 https://doi.org/10.1111/rssa.12042

[43] LECUN, Y., CORTES, C. and BURGES, C. (2010). MNIST handwritten digit database.

[44] LEE, Y. T., SHEN, R. and TIAN, K. (2021). Lower bounds on Metropolized sampling methods for well-conditioned distributions. *Adv. Neural Inf. Process. Syst.* **34** 18812–18824.

[45] LEIMKUHLER, B. and MATTHEWS, C. (2013). Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express. AMRX* **1** 34–56. MR3040887 https://doi.org/10.1093/amrx/abs010

[46] LEIMKUHLER, B. and MATTHEWS, C. (2013). Robust and efficient configurational molecular sampling via Langevin dynamics. *J. Chem. Phys.* **138** 174102.

[47] LEIMKUHLER, B. and MATTHEWS, C. (2015). *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods. Interdisciplinary Applied Mathematics* **39**. Springer, Cham. MR3362507

[48] LEIMKUHLER, B., MATTHEWS, C. and STOLTZ, G. (2016). The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.* **36** 13–79. MR3463433 https://doi.org/10.1093/imanum/dru056

[49] LEIMKUHLER, B., PAULIN, D. and WHALLEY, P. A. (2024). Contraction rate estimates of stochastic gradient kinetic Langevin integrators. *ESAIM. Math. Model. Numer. Anal.* **58** 2255–2286. MR4834894 https://doi.org/10.1051/m2an/2024038

[50] LEIMKUHLER, B. J., PAULIN, D. and WHALLEY, P. A. (2024). Contraction and convergence rates for discretized kinetic Langevin dynamics. *SIAM J. Numer. Anal.* **62** 1226–1258. MR4748799 https://doi.org/10.1137/23M1556289

[51] LU, J. and WANG, L. (2022). On explicit L^2-convergence rate estimate for piecewise deterministic Markov processes in MCMC algorithms. *Ann. Appl. Probab.* **32** 1333–1361. MR4414707 https://doi.org/10.1214/21-aap1710

[52] MAHER, M. J. (1982). Modelling association football scores. *Stat. Neerl.* **36** 109–118.

[53] MAJKA, M. B., MIJATOVIĆ, A. and SZPRUCH, Ł. (2020). Nonasymptotic bounds for sampling algorithms without log-concavity. *Ann. Appl. Probab.* **30** 1534–1581. MR4132634 https://doi.org/10.1214/19-AAP1535

[54] MIDDLETON, L., DELIGIANNIDIS, G., DOUCET, A. and JACOB, P. E. (2020). Unbiased Markov chain Monte Carlo for intractable target distributions. *Electron. J. Stat.* **14** 2842–2891. MR4132645 https://doi.org/10.1214/20-EJS1727

[55] MONMARCHÉ, P. (2021). High-dimensional MCMC with a standard splitting scheme for the underdamped Langevin diffusion. *Electron. J. Stat.* **15** 4117–4166. MR4309974 https://doi.org/10.1214/21-ejs1888

[56] MÜLLER, E. H., SCHEICHL, R. and SHARDLOW, T. (2015). Improving multilevel Monte Carlo for stochastic differential equations with application to the Langevin equation. *Proc. A* **471** 20140679. MR3325195 https://doi.org/10.1098/rspa.2014.0679

[57] NEAL, R. M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC Handb. Mod. Stat. Methods* 113–162. CRC Press, Boca Raton, FL. MR2858447

[58] PARK, T. and CASELLA, G. (2008). The Bayesian lasso. *J. Amer. Statist. Assoc.* **103** 681–686. MR2524001 https://doi.org/10.1198/016214508000000337

[59] PAULIN, D. and WHALLEY, P. A. (2024). Correction to "Wasserstein distance estimates for the distributions of numerical approximations to ergodic stochastic differential equations". *J. Mach. Learn. Res.* **25** 376. MR4847440

[60] PAVLIOTIS, G. A. (2014). *Stochastic Processes and Applications*. *Texts in Applied Mathematics* **60**. Springer, New York. Diffusion processes, the Fokker-Planck and Langevin equations. MR3288096 https://doi.org/10.1007/978-1-4939-1323-7

[61] QUIROZ, M., KOHN, R., VILLANI, M. and TRAN, M.-N. (2019). Speeding up MCMC by efficient data subsampling. *J. Amer. Statist. Assoc.* **114** 831–843. MR3963184 https://doi.org/10.1080/01621459.2018.1448827

[62] RHEE, C.-H. and GLYNN, P. W. (2015). Unbiased estimation with square root convergence for SDE models. *Oper. Res.* **63** 1026–1043. MR3422533 https://doi.org/10.1287/opre.2015.1404

[63] RICHARDSON, L. F. (1911). IX. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philos. Trans. R. Soc. Lond. A.* **210** 307–357.

[64] ROBERT, C. P. and CASELLA, G. (2004). *Monte Carlo Statistical Methods*, 2nd ed. *Springer Texts in Statistics*. Springer, New York. MR2080278 https://doi.org/10.1007/978-1-4757-4145-2

[65] ROBERTS, G. O. and ROSENTHAL, J. S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 255–268. MR1625691 https://doi.org/10.1111/1467-9868.00123

[66] ROBERTS, G. O. and ROSENTHAL, J. S. (2001). Optimal scaling for various Metropolis–Hastings algorithms. *Statist. Sci.* **16** 351–367. MR1888450 https://doi.org/10.1214/ss/1015346320

[67] RUZAYQAT, H., CHADA, N. K. and JASRA, A. (2023). Unbiased estimation using underdamped Langevin dynamics. *SIAM J. Sci. Comput.* **45** A3047–A3070. MR4673688 https://doi.org/10.1137/22M1503385

[68] SANZ-SERNA, J. M. and ZYGALAKIS, K. C. (2021). Wasserstein distance estimates for the distributions of numerical approximations to ergodic stochastic differential equations. *J. Mach. Learn. Res.* **22** 242. MR4329821

[69] SCHUH, K. and WHALLEY, P. A. (2024). Convergence of kinetic Langevin samplers for non-convex potentials. arXiv preprint, arXiv:2405.09992.

[70] SKEEL, R. D. and IZAGUIRRE, J. A. (2002). An impulse integrator for Langevin dynamics. *Mol. Phys.* **100** 3885–3891.

[71] TEH, Y. W., THIERY, A. H. and VOLLMER, S. J. (2016). Consistency and fluctuations for stochastic gradient Langevin dynamics. *J. Mach. Learn. Res.* **17** 7. MR3482927

[72] TELATOVICH, A. and LI, X. (2017). The strong convergence of operator-splitting methods for the Langevin dynamics model. arXiv preprint. Available at arXiv:1706.04237.

[73] VOLLMER, S. J., ZYGALAKIS, K. C. and TEH, Y. W. (2016). Exploration of the (non-)asymptotic bias and variance of stochastic gradient Langevin dynamics. *J. Mach. Learn. Res.* **17** 159. MR3555050

[74] WANG, T. and WANG, G. (2023). Unbiased multilevel Monte Carlo methods for intractable distributions: MLMC meets MCMC. *J. Mach. Learn. Res.* **24** 249. MR4633976

[75] WEISS, C. H., ZHU, F. and HOSHIYAR, A. (2022). Softplus INGARCH models. *Statist. Sinica* **32** 1099–1120. MR4405111 https://doi.org/10.5705/ss.202020.0353

[76] WELLING, M. and TEH, Y. W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the* 28*th International Conference on Machine Learning* 681–688.

[77] WIEMANN, P. F. V., KNEIB, T. and HAMBUCKERS, J. (2024). Using the softplus function to construct alternative link functions in generalized linear models and beyond. *Statist. Papers* **65** 3155–3180. MR4767659 https://doi.org/10.1007/s00362-023-01509-x

[78] ZHENGMIAN, H., HUANG, F. and HUANG, H. (2021). Optimal underdamped Langevin MCMC method. *Adv. Neural Inf. Process. Syst.* **34** 19363–19374.

The Annals of Statistics
2026, Vol. 54, No. 4, 1923–1949
https://doi.org/10.1214/26-AOS2632

ASSOCIATION AND INDEPENDENCE TEST FOR RANDOM OBJECTS

BY HANG ZHOU[1,a] AND HANS-GEORG MÜLLER[2,b]

[1]*Department of Statistics and Operations Research & School of Data and Information Sciences, University of North Carolina at Chapel Hill,* [a]*hzh@unc.edu*

[2]*Department of Statistics, University of California, Davis,* [b]*hgmueller@ucdavis.edu*

We develop a unified framework for testing independence and quantifying association between random objects that are located in general metric spaces. Special cases include functional and high-dimensional data as well as networks, covariance matrices and data on Riemannian manifolds, among other metric space-valued data. A key concept is the profile association, a measure based on distance profiles that intrinsically characterize the distributions of random objects in metric spaces. We establish a connection between Hoeffding's D statistic and the profile association and derive a permutation test with theoretical guarantees for consistency and power under alternatives to the null hypothesis of independence (equivalently, no association). We extend this framework to the conditional setting, where the independence between random objects given a Euclidean predictor is of interest. In simulations across various metric spaces, the proposed profile independence test is found to outperform existing approaches. The practical utility of this framework is demonstrated with applications to brain connectivity networks derived from magnetic resonance imaging and age-at-death distributions for males and females obtained from human mortality data.

REFERENCES

BANERJEE, B., BHATTACHARYA, B. B. and GHOSH, A. K. (2026). A ball divergence based measure for conditional independence testing. *Biometrika* **103** asag027. MR5097446 https://doi.org/10.1093/biomet/asag027

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 https://doi.org/10.3150/13-BEJ514

BERRETT, T. B., WANG, Y., BARBER, R. F. and SAMWORTH, R. J. (2020). The conditional permutation test for independence while controlling for confounders. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 175–197. MR4060981

BEURLING, A. and AHLFORS, L. (1956). The boundary correspondence under quasiconformal mappings. *Acta Math.* **96** 125–142. MR0086869 https://doi.org/10.1007/BF02392360

BHATTACHARJEE, S., LI, B. and XUE, L. (2025). Nonlinear global Fréchet regression for random objects via weak conditional expectation. *Ann. Statist.* **53** 117–143. MR4865010 https://doi.org/10.1214/24-aos2457

BHATTACHARJEE, S. and MÜLLER, H.-G. (2023). Single index Fréchet regression. *Ann. Statist.* **51** 1770–1798. MR4658576 https://doi.org/10.1214/23-aos2307

BICKEL, P. J. (2022). Measures of independence and functional dependence. arXiv Preprint. Available at arXiv:2206.13663.

BIGOT, J., GOUET, R., KLEIN, T. and LÓPEZ, A. (2017). Geodesic PCA in the Wasserstein space by convex PCA. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1–26. MR3606732 https://doi.org/10.1214/15-AIHP706

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 https://doi.org/10.1006/aama.2001.0759

BOLSTAD, B. M., IRIZARRY, R. A., ÅSTRAND, M. and SPEED, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19** 185–193.

BROCK, W. A., SCHEINKMAN, J. A., DECHERT, W. D. and LEBARON, B. (1996). A test for independence based on the correlation dimension. *Econometric Rev.* **15** 197–235. MR1410877 https://doi.org/10.1080/07474939608800353

MSC2020 subject classifications. Primary 62R20, 62G10; secondary 62G20.

Key words and phrases. Conditional association, conditional independence test, correlation, distributional data, functional data, high-dimensional data, inference for networks, non-Euclidean data, permutation test.

BROWNING, S. R. and BROWNING, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Amer. J. Hum. Genet.* **81** 1084–1097.

BUET, B. and LEONARDI, G. P. (2016). Recovering measures from approximate values on balls. *Ann. Acad. Sci. Fenn. Math.* **41** 947–972. MR3525409 https://doi.org/10.5186/aasfm.2016.4160

CHATTERJEE, S. (2021). A new coefficient of correlation. *J. Amer. Statist. Assoc.* **116** 2009–2022. MR4353729 https://doi.org/10.1080/01621459.2020.1758115

CHEN, X., COOK, R. D. and ZOU, C. (2015). Diagnostic studies in sufficient dimension reduction. *Biometrika* **102** 545–558. MR3394274 https://doi.org/10.1093/biomet/asv016

CHEN, Y. and DUBEY, P. (2026). DiPMInd: Distance profile based mutual independence testing for random objects. *Ann. Statist.* To appear.

CHEN, Y., LIN, Z. and MÜLLER, H.-G. (2023). Wasserstein regression. *J. Amer. Statist. Assoc.* **118** 869–882. MR4595462 https://doi.org/10.1080/01621459.2021.1956937

CHRISTENSEN, J. P. R. (1970). On some measures analogous to Haar measure. *Math. Scand.* **26** 103–106. MR0260979 https://doi.org/10.7146/math.scand.a-10969

COOK, R. D. (2004). Testing predictor contributions in sufficient dimension reduction. *Ann. Statist.* **32** 1062–1092. MR2065198 https://doi.org/10.1214/009053604000000292

DAVIES, R. O. (1971). Measures not approximable or not specifiable by means of balls. *Mathematika* **18** 157–160. MR0310162 https://doi.org/10.1112/S0025579300005386

DRYDEN, I. L. and KENT, J. T. (2015). *Geometry Driven Statistics*. Wiley, New York.

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 https://doi.org/10.1214/09-AOAS249

DUBEY, P., CHEN, Y. and MÜLLER, H.-G. (2024). Metric statistics: Exploration and inference for random objects with distance profiles. *Ann. Statist.* **52** 757–792. MR4744195 https://doi.org/10.1214/24-aos2368

DUBEY, P. and MÜLLER, H.-G. (2020). Functional models for time-varying random objects. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 275–327. MR4084166

ELAM, J. S., GLASSER, M. F., HARMS, M. P., SOTIROPOULOS, S. N., ANDERSSON, J. L., BURGESS, G. C., CURTISS, S. W., OOSTENVELD, R., LARSON-PRIOR, L. J. et al. (2021). The human connectome project: A retrospective. *NeuroImage* **244** 118543.

FAN, J. and GIJBELS, I. (1992). Variable bandwidth and local linear regression smoothers. *Ann. Statist.* **20** 2008–2036. MR1193323 https://doi.org/10.1214/aos/1176348900

FAN, J., LIAO, Y. and MINCHEVA, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 603–680. MR3091653 https://doi.org/10.1111/rssb.12016

FAN, J., LOU, Z. and ZHANG, D. (2026). Test of independence using generalized distance correlation. *Ann. Statist.* **54** 1451–1473. MR5088655 https://doi.org/10.1214/25-AOS2613

FRÉCHET, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. Henri Poincaré* **10** 215–310. MR0027464

GAO, Z., WANG, R., WANG, X. and ZHANG, H. (2025). Studentized tests of independence: Random-lifter approach. *Ann. Statist.* **53** 703–723. MR4900163 https://doi.org/10.1214/24-aos2478

GAO, Z., ZHU, J., HU, Y., PAN, W. and WANG, X. (2026). Identification of genetic factors associated with corpus callosum morphology: Conditional strong independence screening for non-Euclidean responses. *Ann. Appl. Stat.* **20** 560–577. MR5047674 https://doi.org/10.1214/25-AOAS2119

GUNSILIUS, F. F. (2023). Distributional synthetic controls. *Econometrica* **91** 1105–1117. MR4598717 https://doi.org/10.3982/ecta18260

HAN, F. and HUANG, Z. (2024). Azadkia-Chatterjee's correlation coefficient adapts to manifold data. *Ann. Appl. Probab.* **34** 5172–5210. MR4840483 https://doi.org/10.1214/24-aap2088

HEINONEN, J. (2001). *Lectures on Analysis on Metric Spaces*. *Universitext*. Springer, New York. MR1800917 https://doi.org/10.1007/978-1-4613-0131-8

HOEFFDING, W. (1948). A non-parametric test of independence. *Ann. Math. Statist.* **19** 546–557. MR0029139 https://doi.org/10.1214/aoms/1177730150

HOFFMANN-JØRGENSEN, J. (1975). Measures which agree on balls. *Math. Scand.* **37** 319–326. MR0409757 https://doi.org/10.7146/math.scand.a-11610

HUANG, T.-M. (2010). Testing conditional independence using maximal nonlinear conditional correlation. *Ann. Statist.* **38** 2047–2091. MR2676883 https://doi.org/10.1214/09-AOS770

JIANG, F., ZHU, C. and SHAO, X. (2024). Two-sample and change-point inference for non-Euclidean valued time series. *Electron. J. Stat.* **18** 848–894. MR4718463 https://doi.org/10.1214/24-ejs2218

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93.

KLEBANOV, L. B. (2005). *N-Distances and Their Applications*. Charles Univ. Prague, Karolinum Press.

KOLLER, D. and FRIEDMAN, N. (2009). *Probabilistic Graphical Models*: *Principles and Techniques*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. MR2778120

LEE, C. E., ZHANG, X. and SHAO, X. (2020). Testing conditional mean independence for functional data. *Biometrika* **107** 331–346. MR4108934 https://doi.org/10.1093/biomet/asz070

LI, B. (2018). *Sufficient Dimension Reduction*: *Methods and Applications with R*. *Monographs on Statistics and Applied Probability* **161**. CRC Press, Boca Raton, FL. MR3838449 https://doi.org/10.1201/9781315119427

LILA, E. and ASTON, J. A. D. (2020). Statistical analysis of functions on surfaces, with an application to medical imaging. *J. Amer. Statist. Assoc.* **115** 1420–1434. MR4143475 https://doi.org/10.1080/01621459.2019.1635479

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 https://doi.org/10.1214/16-AOAS975

LIN, Z. (2019). Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition. *SIAM J. Matrix Anal. Appl.* **40** 1353–1370. MR4032859 https://doi.org/10.1137/18M1221084

LIN, Z. and HAN, F. (2023). On boosting the power of Chatterjee's rank correlation. *Biometrika* **110** 283–299. MR4589063 https://doi.org/10.1093/biomet/asac048

LIN, Z. and HAN, F. (2024). On the failure of the bootstrap for Chatterjee's rank correlation. *Biometrika* **111** 1063–1070. MR4794428 https://doi.org/10.1093/biomet/asae004

LIN, Z. and YAO, F. (2019). Intrinsic Riemannian functional data analysis. *Ann. Statist.* **47** 3533–3577. MR4025751 https://doi.org/10.1214/18-AOS1787

LINTON, O. and GOZALO, P. (1996). Conditional independence restrictions: Testing and estimation. Technical Report, Cowles Foundation for Research in Economics, Yale Univ.

LUNDBORG, A. R., SHAH, R. D. and PETERS, J. (2022). Conditional independence testing in Hilbert spaces with applications to functional data analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1821–1850. MR4515559 https://doi.org/10.1111/rssb.12544

LYONS, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305. MR3127883 https://doi.org/10.1214/12-AOP803

MARRON, J. S. and DRYDEN, I. L. (2022). *Object Oriented Data Analysis*. *Monographs on Statistics and Applied Probability* **169**. CRC Press, Boca Raton, FL. MR4880420

MITCHELL, E. G., DRYDEN, I. L., FALLAIZE, C. J., ANDERSEN, R., BRADLEY, A. V., LARGE, D. J. and SOWTER, A. (2025). Object oriented data analysis of surface motion time series in peatland landscapes. *J. R. Stat. Soc.*, *Ser. C*, *Appl. Stat.* **74** 406–428. MR4878915 https://doi.org/10.1093/jrsssc/qlae060

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 https://doi.org/10.1137/S0895479803436937

MÜLLER, H.-G. (2016). Peter Hall, functional data analysis and random objects. *Ann. Statist.* **44** 1867–1887. MR3546436 https://doi.org/10.1214/16-AOS1492

O'NEILL, B. (1983). *Semi-Riemannian Geometry*: *With Applications to Relativity*. *Pure and Applied Mathematics* **103**. Academic Press [Harcourt Brace Jovanovich, Publishers], New York. MR0719023

PAN, W., WANG, X., ZHANG, H., ZHU, H. and ZHU, J. (2020). Ball covariance: A generic measure of dependence in Banach space. *J. Amer. Statist. Assoc.* **115** 307–317. MR4078465 https://doi.org/10.1080/01621459.2018.1543600

PEARSON, K. (1900). I. Mathematical contributions to the theory of evolution.-VII. On the correlation of characters not quantitatively measurable. *Philos. Trans. R. Soc. Lond. A* **195** 1–47.

PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2017). *Elements of Causal Inference*: *Foundations and Learning Algorithms*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. MR3822088

PETERSEN, A. and MÜLLER, H.-G. (2016). Functional data analysis for density functions by transformation to a Hilbert space. *Ann. Statist.* **44** 183–218. MR3449766 https://doi.org/10.1214/15-AOS1363

PETERSEN, A. and MÜLLER, H.-G. (2019). Fréchet regression for random objects with Euclidean predictors. *Ann. Statist.* **47** 691–719. MR3909947 https://doi.org/10.1214/17-AOS1624

PRITCHARD, J. K., STEPHENS, M. and DONNELLY, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155** 945–959.

SCHÖTZ, C. (2025). Variance inequalities for transformed Fréchet means in Hadamard spaces. *Electron. J. Probab.* **30** Paper No. 15. MR4853997 https://doi.org/10.1214/25-ejp1273

SEJDINOVIC, D., SRIPERUMBUDUR, B., GRETTON, A. and FUKUMIZU, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.* **41** 2263–2291. MR3127866 https://doi.org/10.1214/13-AOS1140

SERFLING, R. J. (2009). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

SHI, H., DRTON, M. and HAN, F. (2022). On the power of Chatterjee's rank correlation. *Biometrika* **109** 317–333. MR4430960 https://doi.org/10.1093/biomet/asab028

SPEARMAN, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15** 72–101.
STOREY, J. D. and TIBSHIRANI, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100** 9440–9445. MR1994856 https://doi.org/10.1073/pnas.1530509100
SU, L. and WHITE, H. (2007). A consistent characteristic function-based test for conditional independence. *J. Econometrics* **141** 807–834. MR2413488 https://doi.org/10.1016/j.jeconom.2006.11.006
SU, L. and WHITE, H. (2008). A nonparametric Hellinger metric test for conditional independence. *Econometric Theory* **24** 829–864. MR2428851 https://doi.org/10.1017/S0266466608080341
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 https://doi.org/10.1214/009053607000000505
VAN ESSEN, D. C., SMITH, S. M., BARCH, D. M., BEHRENS, T. E., YACOUB, E., UGURBIL, K., CONSORTIUM, W.-M. H. et al. (2013). The WU-Minn human connectome project: An overview. *NeuroImage* **80** 62–79.
WANG, X. and HONG, Y. (2018). Characteristic function based testing for conditional independence: A nonparametric regression approach. *Econometric Theory* **34** 815–849. MR3815468 https://doi.org/10.1017/S026646661700010X
WANG, X., PAN, W., HU, W., TIAN, Y. and ZHANG, H. (2015). Conditional distance correlation. *J. Amer. Statist. Assoc.* **110** 1726–1734. MR3449068 https://doi.org/10.1080/01621459.2014.993081
WANG, X., ZHU, J., PAN, W., ZHU, J. and ZHANG, H. (2024). Nonparametric statistical inference via metric distribution function in metric spaces. *J. Amer. Statist. Assoc.* **119** 2772–2784. MR4833914 https://doi.org/10.1080/01621459.2023.2277417
WEIHS, L., DRTON, M. and MEINSHAUSEN, N. (2018). Symmetric rank covariances: A generalized framework for nonparametric measures of dependence. *Biometrika* **105** 547–562. MR3842884 https://doi.org/10.1093/biomet/asy021
WILLIAMS, D. R. and MULDER, J. (2020). Bayesian hypothesis testing for Gaussian graphical models: Conditional independence and order constraints. *J. Math. Psych.* **99** 102441. MR4156963 https://doi.org/10.1016/j.jmp.2020.102441
WU-MINN, H. (2017). 1200 *Subjects Data Release Reference Manual*. Human Connectome Project.
YANAGIMOTO, T. (1970). On measures of association and a related problem. *Ann. Inst. Statist. Math.* **22** 57–63.
ZHANG, H., ZHOU, S., ZHANG, K. and GUAN, J. (2017). Causal discovery using regression-based conditional independence tests. In *Proceedings of the AAAI Conference on Artificial Intelligence* **31**.
ZHANG, K., PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2011). Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence* 804–813.
ZHOU, H. and MÜLLER, H.-G. (2025). Conformal inference for random objects. *Ann. Statist.* **53** 1123–1149. MR4925118 https://doi.org/10.1214/25-aos2495
ZHOU, H. and MÜLLER, H.-G. (2026). Supplement to "Association and independence test for random objects." https://doi.org/10.1214/26-AOS2632SUPP
ZHOU, Y., LIU, J. and ZHU, L. (2020). Test for conditional independence with application to conditional screening. *J. Multivariate Anal.* **175** 104557. MR4017960 https://doi.org/10.1016/j.jmva.2019.104557
ZHOU, Y. and MÜLLER, H.-G. (2022). Network regression with graph Laplacians. *J. Mach. Learn. Res.* **23** Paper No. [320]. MR4577759
ZHU, C. and MÜLLER, H.-G. (2023). Autoregressive optimal transport models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 1012–1033. MR4726980 https://doi.org/10.1093/jrsssb/qkad051
ZHU, J., PAN, W., ZHENG, W. and WANG, X. (2021). Ball: An R package for detecting distribution difference and association in metric spaces. *J. Stat. Softw.* **97** 1–31.

The Annals of Statistics
2026, Vol. 54, No. 4, 1950–1978
https://doi.org/10.1214/26-AOS2634

LINEAR METHODS FOR NONLINEAR INVERSE PROBLEMS

BY GEERTEN KOERS[1,a], BOTOND SZABÓ[2,c] AND AAD VAN DER VAART[1,b]

[1]*Delft University of Technology, DIAM,* [a]*geertenkoers@outlook.com,* [b]*a.w.vandervaart@tudelft.nl*
[2]*Bocconi University, Department of Data Sciences and BIDSA,* [c]*botond.szabo@unibocconi.it*

We consider the recovery of an unknown function f from a noisy observation of the solution u_f to a partial differential equation that can be written in the form $\mathcal{L}u_f = c(f, u_f)$, for a differential operator $\mathcal{L}$ that is rich enough to recover f from $\mathcal{L}u_f$. Examples include the time-independent Schrödinger equation $\Delta u_f = 2u_f f$, the heat equation with absorption term $(\partial_t - \Delta_x/2)u_f = fu_f$ and the Darcy problem $\nabla \cdot (f\nabla u_f) = h$. We transform this problem into the linear inverse problem of recovering $\mathcal{L}u_f$ under the Dirichlet boundary condition, and show that Bayesian methods with priors placed either on u_f or $\mathcal{L}u_f$ for this problem yield optimal recovery rates not only for u_f, but also for f. We also derive frequentist coverage guarantees for the corresponding Bayesian credible sets. Adaptive priors are shown to yield adaptive contraction rates for f, thus eliminating the need to know the smoothness of this function. The results are illustrated by numerical experiments on synthetic data sets.

REFERENCES

[1] AGAPIOU, S., LARSSON, S. and STUART, A. M. (2013). Posterior contraction rates for the Bayesian approach to linear ill-posed inverse problems. *Stoch. Process. Appl.* **123** 3828–3860. MR3084161 https://doi.org/10.1016/j.spa.2013.05.001

[2] AGAPIOU, S. and SAVVA, A. (2024). Adaptive inference over Besov spaces in the white noise model using p-exponential priors. *Bernoulli* **30** 2275–2300. MR4746608 https://doi.org/10.3150/23-bej1673

[3] ALBERTI, G. S. and CAPDEBOSCQ, Y. (2018). *Lectures on Elliptic Methods for Hybrid Inverse Problems. Cours Spécialisés* [*Specialized Courses*] **25**. Société Mathématique de France, Paris. MR3751364

[4] BOURGAIN, J. and WOLFF, T. (1990). A remark on gradients of harmonic functions in dimension ≥ 3. *Colloq. Math.* **60/61** 253–260. MR1096375 https://doi.org/10.4064/cm-60-61-1-253-260

[5] CAI, T. T. and LOW, M. G. (2004). An adaptation theory for nonparametric confidence intervals. *Ann. Statist.* **32** 1805–1840. MR2102494 https://doi.org/10.1214/009053604000000049

[6] CASTILLO, I. (2012). A semiparametric Bernstein–von Mises theorem for Gaussian process priors. *Probab. Theory Related Fields* **152** 53–99. MR2875753 https://doi.org/10.1007/s00440-010-0316-5

[7] CASTILLO, I. and NICKL, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. MR3262473 https://doi.org/10.1214/14-AOS1246

[8] CAVALIER, L. (2011). Inverse problems in statistics. In *Inverse Problems and High-Dimensional Estimation. Lect. Notes Stat. Proc.* **203** 3–96. Springer, Heidelberg. MR2868199 https://doi.org/10.1007/978-3-642-19989-9_1

[9] CHAKRABORTY, M. and GHOSAL, S. (2022). Rates and coverage for monotone densities using projection-posterior. *Bernoulli* **28** 1093–1119. MR4388931 https://doi.org/10.3150/21-bej1379

[10] CHUNG, K. L. and ZHAO, Z. X. (1995). *From Brownian Motion to Schrödinger's Equation. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **312**. Springer, Berlin. MR1329992 https://doi.org/10.1007/978-3-642-57856-4

[11] DI NEZZA, E., PALATUCCI, G. and VALDINOCI, E. (2012). Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** 521–573. MR2944369 https://doi.org/10.1016/j.bulsci.2011.12.004

[12] DOOB, J. L. (1955). A probability approach to the heat equation. *Trans. Amer. Math. Soc.* **80** 216–280. MR0079376 https://doi.org/10.2307/1993013

[13] ENGL, H. W., HANKE, M. and NEUBAUER, A. (1996). *Regularization of Inverse Problems. Mathematics and Its Applications* **375**. Kluwer Academic, Dordrecht. MR1408680

MSC2020 subject classifications. Primary 62G05, 62G15; secondary 62G20.
Key words and phrases. Posterior distribution, Inverse problem, nonlinear, contraction rate, adaptation, uncertainty quantification.

[14] EVANS, L. C. (2010). *Partial Differ. Equ.*, 2nd ed. *Graduate Studies in Mathematics* **19**. Amer. Math. Soc., Providence, RI. MR2597943 https://doi.org/10.1090/gsm/019

[15] FLEMMING, J. (2014). Regularization of autoconvolution and other ill-posed quadratic equations by decomposition. *J. Inverse Ill-Posed Probl.* **22** 551–567. MR3259109 https://doi.org/10.1515/jip-2013-0038

[16] GHOSAL, S., GHOSH, J. K. and VAN DER VAART, A. W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228

[17] GILBARG, D. and TRUDINGER, N. S. (2015). *Elliptic Partial Differential Equations of Second Order*. *Grundlehren der Mathematischen Wissenschaften* **224**. Springer Berlin / Heidelberg, Berlin.

[18] GIORDANO, M. and NICKL, R. (2020). Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. *Inverse Probl.* **36** 085001, 35. MR4151406 https://doi.org/10.1088/1361-6420/ab7d2a

[19] GUGUSHVILI, S., VAN DER VAART, A. and YAN, D. (2020). Bayesian linear inverse problems in regularity scales. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2081–2107. MR4116718 https://doi.org/10.1214/19-AIHP1029

[20] IVRII, V. (2016). 100 years of Weyl's law. *Bull. Math. Sci.* **6** 379–452. MR3556544 https://doi.org/10.1007/s13373-016-0089-y

[21] KEKKONEN, H. (2022). Consistency of Bayesian inference with Gaussian process priors for a parabolic inverse problem. *Inverse Probl.* **38** Paper No. 035002, 29. MR4385425 https://doi.org/10.1088/1361-6420/ac4839

[22] KINDERLEHRER, D. and STAMPACCHIA, G. (2000). *An Introduction to Variational Inequalities and Their Applications*. *Classics in Applied Mathematics* **31**. SIAM, Philadelphia, PA. Reprint of the 1980 original. MR1786735 https://doi.org/10.1137/1.9780898719451

[23] KNAPIK, B. T., SZABÓ, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probab. Theory Related Fields* **164** 771–813. MR3477780 https://doi.org/10.1007/s00440-015-0619-7

[24] KNAPIK, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.* **39** 2626–2657. MR2906881 https://doi.org/10.1214/11-AOS920

[25] KNAPIK, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2013). Bayesian recovery of the initial condition for the heat equation. *Comm. Statist. Theory Methods* **42** 1294–1313. MR3031282 https://doi.org/10.1080/03610926.2012.681417

[26] KOERS, G., SZABÓ, B. and VAN DER VAART, A. (2026). Supplement to "Linear methods for nonlinear inverse problems." https://doi.org/10.1214/26-AOS2634SUPP

[27] MONARD, F., NICKL, R. and PATERNAIN, G. P. (2021). Statistical guarantees for Bayesian uncertainty quantification in nonlinear inverse problems with Gaussian process priors. *Ann. Statist.* **49** 3255–3298. MR4352530 https://doi.org/10.1214/21-aos2082

[28] MONARD, F., NICKL, R. and PATERNAIN, G. P. (2021). Consistent inversion of noisy non-Abelian X-ray transforms. *Comm. Pure Appl. Math.* **74** 1045–1099. MR4230066 https://doi.org/10.1002/cpa.21942

[29] NICKL, R. (2020). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc.* (*JEMS*) **22** 2697–2750. MR4118619 https://doi.org/10.4171/JEMS/975

[30] NICKL, R. (2023). *Bayesian Non-linear Statistical Inverse Problems*. *Zurich Lectures in Advanced Mathematics*. EMS Press, Berlin. MR4604099 https://doi.org/10.4171/zlam/30

[31] NICKL, R., VAN DE GEER, S. and WANG, S. (2020). Convergence rates for penalized least squares estimators in PDE constrained regression problems. *SIAM/ASA J. Uncertain. Quantificat.* **8** 374–413. MR4074017 https://doi.org/10.1137/18M1236137

[32] NICKL, R. and WANG, S. (2020). On polynomial-time computation of high-dimensional posterior measures by Langevin-type algorithms.

[33] NIEMAN, D., SZABO, B. and VAN ZANTEN, H. (2023). Uncertainty quantification for sparse spectral variational approximations in Gaussian process regression. *Electron. J. Stat.* **17** 2250–2288. MR4649981 https://doi.org/10.1214/23-ejs2155

[34] RANDRIANARISOA, T. and SZABO, B. (2023). Variational Gaussian processes for linear inverse problems. *Adv. Neural Inf. Process. Syst.* **36** 28960–28972.

[35] RAY, K. (2013). Bayesian inverse problems with non-conjugate priors. *Electron. J. Stat.* **7** 2516–2549. MR3117105 https://doi.org/10.1214/13-EJS851

[36] RAY, K. (2017). Adaptive Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **45** 2511–2536. MR3737900 https://doi.org/10.1214/16-AOS1533

[37] RICHTER, G. R. (1981). An inverse problem for the steady state diffusion equation. *SIAM J. Appl. Math.* **41** 210–221. MR0628945 https://doi.org/10.1137/0141016

[38] RICHTER, G. R. (1981). Numerical identification of a spatially varying diffusion coefficient. *Math. Comp.* **36** 375–386. MR0606502 https://doi.org/10.2307/2007648

[39] ROBINS, J. and VAN DER VAART, A. (2006). Adaptive nonparametric confidence sets. *Ann. Statist.* **34** 229–253. MR2275241 https://doi.org/10.1214/009053605000000877

[40] ROUSSEAU, J. and SZABO, B. (2017). Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator. *Ann. Statist.* **45** 833–865. MR3650402 https://doi.org/10.1214/16-AOS1469

[41] ROUSSEAU, J. and SZABO, B. (2020). Asymptotic frequentist coverage properties of Bayesian credible sets for sieve priors. *Ann. Statist.* **48** 2155–2179. MR4134790 https://doi.org/10.1214/19-AOS1881

[42] SEELEY, R. T. (1965). Integro-differential operators on vector bundles. *Trans. Amer. Math. Soc.* **117** 167–204. MR0173174 https://doi.org/10.2307/1994203

[43] SNIEKERS, S. and VAN DER VAART, A. (2015). Adaptive Bayesian credible sets in regression with a Gaussian process prior. *Electron. J. Stat.* **9** 2475–2527. MR3425364 https://doi.org/10.1214/15-EJS1078

[44] SNIEKERS, S. and VAN DER VAART, A. (2020). Adaptive Bayesian credible bands in regression with a Gaussian process prior. *Sankhya A* **82** 386–425. MR4136240 https://doi.org/10.1007/s13171-019-00185-0

[45] SZABO, B., HADJI, A. and VAN DER VAART, A. (2025). Adaptation using spatially distributed Gaussian processes. *J. Amer. Statist. Assoc.* **120** 2784–2795. MR5006881 https://doi.org/10.1080/01621459.2025.2501717

[46] SZABÓ, B., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428. MR3357861 https://doi.org/10.1214/14-AOS1270

[47] SZABÓ, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2013). Empirical Bayes scaling of Gaussian priors in the white noise model. *Electron. J. Stat.* **7** 991–1018. MR3044507 https://doi.org/10.1214/13-EJS798

[48] TAYLOR, M. E. (2011). *Partial Differential Equations I. Basic Theory*, 2nd ed. *Applied Mathematical Sciences* **115**. Springer, New York. MR2744150 https://doi.org/10.1007/978-1-4419-7055-8

[49] TITSIAS, M. (2009). Variational learning of inducing variables in sparse Gaussian processes. In *Artificial Intelligence and Statistics* 567–574. PMLR.

[50] TRAVIS, L. and RAY, K. (2024). Pointwise uncertainty quantification for sparse variational Gaussian process regression with a Brownian motion prior. *Adv. Neural Inf. Process. Syst.* **36**.

[51] TRIEBEL, H. (2008). *Function Spaces and Wavelets on Domains*. *EMS Tracts in Mathematics* **7**. European Mathematical Society (EMS), Zürich. MR2455724 https://doi.org/10.4171/019

[52] TRÖLTZSCH, F. (2010). *Optimal Control of Partial Differential Equations*: *Theory*, *Methods and Applications*. *Graduate Studies in Mathematics* **112**. Amer. Math. Soc., Providence, RI. Translated from the 2005 German original by Jürgen Sprekels. MR2583281 https://doi.org/10.1090/gsm/112

[53] WANG, S. (2026). Global polynomial-time estimation in statistical nonlinear inverse problems via generalized stability.

[54] YAN, D. (2020). Bayesian inference for Gaussian models. PhD Thesis, Univ. Leiden, https://hdl.handle.net/1887/86070.

[55] YAN, D., VAN DER VAART, A. and GUGUSHVILI, S. (2024). Bayesian linear inverse problems in regularity scales with discrete observations. *Sankhya A* **86** S228–S254. MR4827150 https://doi.org/10.1007/s13171-024-00342-0

The Annals of Statistics
2026, Vol. 54, No. 4, 1979–2004
https://doi.org/10.1214/26-AOS2636

ON THE EFFICIENCY OF HIGHLY STRATIFIED EXPERIMENTS

BY YUEHAO BAI[1,a], JIZHOU LIU[2,b], AZEEM M. SHAIKH[3,c] AND MAX TABORD-MEEHAN[4,d]

[1]*Department of Economics, University of Southern California,* [a]*yuehao.bai@usc.edu*
[2]*HSBC Business School, Peking University,* [b]*jizhou.liu@phbs.pku.edu.cn*
[3]*Department of Economics, University of Chicago,* [c]*amshaikh@uchicago.edu*
[4]*Department of Economics, University of Toronto,* [d]*m.tabordmeehan@utoronto.ca*

This paper studies the use of highly stratified designs for the efficient estimation of a large class of treatment effect parameters that arise in the analysis of experiments. By a "highly stratified" design, we mean experiments in which units are divided into blocks of a fixed size and a proportion within each block is assigned to a binary treatment uniformly at random. The class of parameters considered are those that can be expressed as the solution to a set of moment conditions constructed using a known function of the observed data. They include, among other things, average treatment effects, quantile treatment effects and local average treatment effects as well as the counterparts to these quantities in experiments in which the unit is itself a cluster. In this setting, we establish three results. First, we show that under a highly stratified design, the naïve method of moments estimator achieves the same asymptotic variance as what could typically be attained under alternative treatment assignment mechanisms only through *ex post* covariate adjustment. Second, we argue that the naïve method of moments estimator under a highly stratified design is asymptotically efficient by deriving a lower bound on the asymptotic variance of regular estimators of the parameter of interest in the form of a convolution theorem. In this sense, highly stratified experiments are attractive because they lead to efficient estimators of treatment effect parameters "by design." Finally, we strengthen this conclusion by establishing conditions under which a "fast-balancing" property of highly stratified designs is in fact necessary for the naïve method of moments estimator to attain the efficiency bound.

REFERENCES

[1] ABADIE, A. and IMBENS, G. W. (2008). Estimation of the conditional variance in paired experiments. *Ann. Écon. Stat.* **1** 175–187.

[2] ANGRIST, J. and LAVY, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *Amer. Econ. Rev.* **99** 1384–1414.

[3] ARMSTRONG, T. B. (2022). Asymptotic efficiency bounds for a class of experimental designs. arXiv:2205.02726 [stat].

[4] BAI, Y. (2022). Optimality of matched-pair designs in randomized controlled trials. *Amer. Econ. Rev.* **112** 3911–3940.

[5] BAI, Y., GUO, H., SHAIKH, A. M. and TABORD-MEEHAN, M. (2025). Inference in experiments with matched pairs and imperfect compliance. *J. Bus. Econom. Statist.* **43** 627–642. MR4918706 https://doi.org/10.1080/07350015.2024.2416972

[6] BAI, Y., HUANG, X., ROMANO, J. P., SHAIKH, A. M. and TABORD-MEEHAN, M. (2025). A new design-based variance estimator for finely stratified experiments. arXiv preprint arXiv:2503.10851.

[7] BAI, Y., JIANG, L., ROMANO, J. P., SHAIKH, A. M. and ZHANG, Y. (2024). Covariate adjustment in experiments with matched pairs. *J. Econometrics* **241** Paper No. 105740, 17. MR4735186 https://doi.org/10.1016/j.jeconom.2024.105740

MSC2020 subject classifications. Primary 62K05, 62K10; secondary 62K99.
Key words and phrases. Convolution theorem, efficiency, experiment, experimental design, highly stratified experiment, matched pairs, randomized controlled trial.

[8] BAI, Y., LIU, J., SHAIKH, A. M. and TABORD-MEEHAN, M. (2024). Inference in cluster randomized trials with matched pairs. *J. Econometrics* **245** Paper No. 105873, 19. MR4814644 https://doi.org/10.1016/j.jeconom.2024.105873

[9] BAI, Y., LIU, J., SHAIKH, A. M. and TABORD-MEEHAN, M. (2026). Supplement to "On the efficiency of highly stratified experiments." https://doi.org/10.1214/26-AOS2636SUPP

[10] BAI, Y., LIU, J. and TABORD-MEEHAN, M. (2024). Inference for matched tuples and fully blocked factorial designs. *Quant. Econ.* **15** 279–330. MR4750953 https://doi.org/10.3982/qe2354

[11] BAI, Y., ROMANO, J. P. and SHAIKH, A. M. (2022). Inference in experiments with matched pairs. *J. Amer. Statist. Assoc.* **117** 1726–1737. MR4528466 https://doi.org/10.1080/01621459.2021.1883437

[12] BANERJEE, A., DUFLO, E., GLENNERSTER, R. and KINNAN, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. *Amer. Econ. J. Appl. Econ.* **7** 22–53.

[13] BELLONI, A., CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. and HANSEN, C. (2017). Program evaluation and causal inference with high-dimensional data. *Econometrica* **85** 233–298. MR3611771 https://doi.org/10.3982/ECTA12723

[14] BRUHN, M., LEÃO, L. D. S., LEGOVINI, A., MARCHETTI, R. and ZIA, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *Amer. Econ. J. Appl. Econ.* **8** 256–295.

[15] BRUHN, M. and MCKENZIE, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *Amer. Econ. J. Appl. Econ.* **1** 200–232.

[16] BUGNI, F., CANAY, I., SHAIKH, A. and TABORD-MEEHAN, M. (2022). Inference for cluster randomized experiments with non-ignorable cluster sizes. arXiv:2204.08356 [econ, stat].

[17] BUGNI, F. A., CANAY, I. A. and SHAIKH, A. M. (2018). Inference under covariate-adaptive randomization. *J. Amer. Statist. Assoc.* **113** 1784–1796. MR3902246 https://doi.org/10.1080/01621459.2017.1375934

[18] BUGNI, F. A., CANAY, I. A. and SHAIKH, A. M. (2019). Inference under covariate-adaptive randomization with multiple treatments. *Quant. Econ.* **10** 1747–1785. MR4048253 https://doi.org/10.3982/qe1150

[19] CAI, Y. and RAFI, A. (2024). On the performance of the Neyman allocation with small pilots. *J. Econometrics* **242** Paper No. 105793, 20. MR4755483 https://doi.org/10.1016/j.jeconom.2024.105793

[20] CASABURI, L. and REED, T. (2022). Using individual-level randomized treatment to learn about market structure. *Amer. Econ. J. Appl. Econ.* **14** 58–90.

[21] CATTANEO, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *J. Econometrics* **155** 138–154. MR2607191 https://doi.org/10.1016/j.jeconom.2009.09.023

[22] CHEN, X., HONG, H. and TAROZZI, A. (2008). Semiparametric efficiency in GMM models with auxiliary data. *Ann. Statist.* **36** 808–843. MR2396816 https://doi.org/10.1214/009053607000000947

[23] CHERNOZHUKOV, V., CHETVERIKOV, D., DEMIRER, M., DUFLO, E., HANSEN, C. and NEWEY, W. (2017). Double/debiased/Neyman machine learning of treatment effects. *Amer. Econ. Rev.* **107** 261–265.

[24] CHERNOZHUKOV, V., CINELLI, C., NEWEY, W., SHARMA, A. and SYRGKANIS, V. (2024). Long story short: Omitted variable bias in causal machine learning. arXiv:2112.13398 [cs, econ, stat].

[25] CYTRYNBAUM, M. (2023). Designing representative and balanced experiments by local randomization. arXiv preprint arXiv:2111.08157.

[26] CYTRYNBAUM, M. (2024). Covariate adjustment in stratified experiments. *Quant. Econ.* **15** 971–998. MR4840764 https://doi.org/10.3982/qe2475

[27] CYTRYNBAUM, M. (2024). Finely stratified rerandomization designs. arXiv preprint arXiv:2407.03279.

[28] EFRON, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika* **58** 403–417. MR0312660 https://doi.org/10.1093/biomet/58.3.403

[29] FARRELL, M. H. (2015). Robust inference on average treatment effects with possibly more covariates than observations. *J. Econometrics* **189** 1–23. MR3397349 https://doi.org/10.1016/j.jeconom.2015.06.017

[30] FIRPO, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica* **75** 259–276. MR2284743 https://doi.org/10.1111/j.1468-0262.2007.00738.x

[31] FOGARTY, C. B. (2018). On mitigating the analytical limitations of finely stratified experiments. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 1035–1056. MR3874309 https://doi.org/10.1111/rssb.12290

[32] FOGARTY, C. B. (2018). Regression-assisted inference for the average treatment effect in paired experiments. *Biometrika* **105** 994–1000. MR3877880 https://doi.org/10.1093/biomet/asy034

[33] FREEDMAN, D. A. (2008). On regression adjustments to experimental data. *Adv. in Appl. Math.* **40** 180–193. MR2388610 https://doi.org/10.1016/j.aam.2006.12.003

[34] HAHN, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica* **66** 315–331. MR1612242 https://doi.org/10.2307/2998560

[35] HAN, Q. and WELLNER, J. A. (2021). Complex sampling designs: Uniform limit theorems and applications. *Ann. Statist.* **49** 459–485. MR4206686 https://doi.org/10.1214/20-AOS1964

[36] HANNEKE, S., KONTOROVICH, A., SABATO, S. and WEISS, R. (2021). Universal Bayes consistency in metric spaces. *Ann. Statist.* **49** 2129–2150. MR4319244 https://doi.org/10.1214/20-aos2029

[37] HIRANO, K. and IMBENS, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Serv. Outcomes Res. Methodol.* **2** 259–278.
[38] HIRANO, K., IMBENS, G. W. and RIDDER, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71** 1161–1189. MR1995826 https://doi.org/10.1111/1468-0262.00442
[39] HU, Y. and HU, F. (2012). Asymptotic properties of covariate-adaptive randomization. *Ann. Statist.* **40** 1794–1815. MR3015044 https://doi.org/10.1214/12-AOS983
[40] IMAI, K. (2008). Variance identification and efficiency analysis in randomized experiments under the matched-pair design. *Stat. Med.* **27** 4857–4873. MR2528770 https://doi.org/10.1002/sim.3337
[41] IMAI, K., KING, G. and NALL, C. (2009). The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation. *Statist. Sci.* **24** 29–53. MR2561126 https://doi.org/10.1214/08-STS274
[42] IMBENS, G. W. (2011). Experimental design for unit and cluster randomized trials. Technical Report.
[43] IMBENS, G. W. and ANGRIST, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62** 467–475.
[44] JIANG, L., LINTON, O. B., TANG, H. and ZHANG, Y. (2022). Improving estimation efficiency via regression-adjustment in covariate-adaptive randomizations with imperfect compliance. arXiv:2201.13004 [econ, stat].
[45] JIANG, L., LIU, X., PHILLIPS, P. C. and ZHANG, Y. (2021). Bootstrap inference for quantile treatment effects in randomized experiments with matched pairs. *Rev. Econ. Stat.* 1–47.
[46] JIANG, L., PHILLIPS, P. C. B., TAO, Y. and ZHANG, Y. (2023). Regression-adjusted estimation of quantile treatment effects under covariate-adaptive randomizations. *J. Econometrics* **234** 758–776. MR4574121 https://doi.org/10.1016/j.jeconom.2022.08.010
[47] LEHMANN, E. L. and ROMANO, J. P. (2021). *Testing Statistical Hypotheses. Springer Texts in Statistics*. Springer, Cham. Fourth edition [of 0107933]. MR4489085
[48] LI, X. and DING, P. (2017). General forms of finite population central limit theorems with applications to causal inference. *J. Amer. Statist. Assoc.* **112** 1759–1769. MR3750897 https://doi.org/10.1080/01621459.2017.1295865
[49] LI, X. and DING, P. (2020). Rerandomization and regression adjustment. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 241–268. MR4060984
[50] LI, X., DING, P. and RUBIN, D. B. (2018). Asymptotic theory of rerandomization in treatment-control experiments. *Proc. Natl. Acad. Sci. USA* **115** 9157–9162. MR3856113 https://doi.org/10.1073/pnas.1808191115
[51] LI, X., DING, P. and RUBIN, D. B. (2020). Rerandomization in 2^K factorial experiments. *Ann. Statist.* **48** 43–63. MR4065152 https://doi.org/10.1214/18-AOS1790
[52] LIN, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *Ann. Appl. Stat.* **7** 295–318. MR3086420 https://doi.org/10.1214/12-AOAS583
[53] LIU, H. and YANG, Y. (2020). Regression-adjusted average treatment effect estimates in stratified randomized experiments. *Biometrika* **107** 935–948. MR4186497 https://doi.org/10.1093/biomet/asaa038
[54] MA, W., LI, P., ZHANG, L.-X. and HU, F. (2024). A new and unified family of covariate adaptive randomization procedures and their properties. *J. Amer. Statist. Assoc.* **119** 151–162. MR4713882 https://doi.org/10.1080/01621459.2022.2102986
[55] MA, W., QIN, Y., LI, Y. and HU, F. (2020). Statistical inference for covariate-adaptive randomization procedures. *J. Amer. Statist. Assoc.* **115** 1488–1497. MR4143480 https://doi.org/10.1080/01621459.2019.1635483
[56] NEWEY, W. K. and MCFADDEN, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics, Vol. IV. Handbooks in Econom.* **2** 2111–2245. North-Holland, Amsterdam. MR1315971
[57] PASHLEY, N. E. and MIRATRIX, L. W. (2021). Insights on variance estimation for blocked and matched pairs designs. *J. Educ. Behav. Stat.* **46** 271–296.
[58] POCOCK, S. J. and SIMON, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 103–115.
[59] RAFI, A. (2023). Efficient semiparametric estimation of average treatment effects under covariate adaptive randomization. arXiv preprint arXiv:2305.08340.
[60] ROBINS, J. M., ROTNITZKY, A. and ZHAO, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J. Amer. Statist. Assoc.* **90** 106–121. MR1325118
[61] TSIATIS, A. A., DAVIDIAN, M., ZHANG, M. and LU, X. (2008). Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: A principled yet flexible approach. *Stat. Med.* **27** 4658–4677. MR2528575 https://doi.org/10.1002/sim.3113
[62] VAN DER VAART, A. (1989). On the asymptotic information bound. *Ann. Statist.* **17** 1487–1500. MR1026295 https://doi.org/10.1214/aos/1176347377

[63] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247 https://doi.org/10.1017/CBO9780511802256

[64] VAN DER VAART, A. W. and WELLNER, J. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*. *Springer Series in Statistics*. Springer-Verlag, New York. MR4628026 https://doi.org/10.1007/978-1-4757-2545-2

[65] WEI, L. J. (1978). The adaptive biased coin design for sequential experiments. *Ann*. *Statist*. **6** 92–100. MR0471205

[66] YE, T., YI, Y. and SHAO, J. (2022). Inference on the average treatment effect under minimization and other covariate-adaptive randomization methods. *Biometrika* **109** 33–47. MR4374639 https://doi.org/10.1093/biomet/asab015

[67] ZELEN, M. (1974). The randomization and stratification of patients to clinical trials. *J*. *Chronic Dis*. **27** 365–375.

[68] ZHANG, M., TSIATIS, A. A. and DAVIDIAN, M. (2008). Improving efficiency of inferences in randomized clinical trials using auxiliary covariates. *Biometrics* **64** 707–715. MR2526620 https://doi.org/10.1111/j.1541-0420.2007.00976.x

The Annals of Statistics
2026, Vol. 54, No. 4, 2005–2030
https://doi.org/10.1214/26-AOS2638

GENERALIZED MULTIVARIATE THRESHOLD AUTOREGRESSIVE MODELS WITH LINEARLY PARTITIONED THRESHOLD SPACE

BY GAN YUAN[1,a] AND CHUN YIP YAU[2,b]

[1] *Department of Biostatistics, City University of Hong Kong,* [a] *ganyuan@cityu.edu.hk*
[2] *Department of Statistics and Data Science, The Chinese University of Hong Kong,* [b] *cy.yau@cuhk.edu.hk*

We consider a k-dimensional multiple-regime vector threshold autoregressive model, in which the regime-switching mechanism is governed by a bivariate threshold variable. Specifically, the regimes are induced by a partition of the threshold space by an unknown number of threshold lines. Within each regime, the process follows a specific vector autoregressive (VAR) model. We formulate the model selection and parameter estimation into a minimization problem based on the Minimum Description Length (MDL) principle and estimate the number of threshold lines, parametric forms of threshold lines and VAR model parameters in each regime simultaneously. Theoretically, we show that the MDL estimators of threshold lines are n-consistent and characterize their limiting distribution. This requires novel proving techniques of introducing a new functional space $\mathbb{G}$ for the local MDL difference functions and establishing weak convergence results therein. Finally, we conduct some empirical studies with simulated datasets and perform real data analyses on U.S. interest rates and U.S. GNP Data.

REFERENCES

[1] ADDO, P. M. (2014). Multivariate self-exciting threshold autoregressive models with exogenous input. Available at arXiv:1407.7738.

[2] BAI, Q. and LING, S. (2024). On ergodicity of threshold ARMA(m, p, q) models. *Jpn. J. Stat. Data Sci.* **7** 667–675. MR4816074 https://doi.org/10.1007/s42081-024-00248-z

[3] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

[4] CHAN, K.-S. and GORACCI, G. (2019). On the ergodicity of first-order threshold autoregressive moving-average processes. *J. Time Series Anal.* **40** 256–264. MR3915530 https://doi.org/10.1111/jtsa.12440

[5] CHAN, K. S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Ann. Statist.* **21** 520–533. MR1212191 https://doi.org/10.1214/aos/1176349040

[6] CHAN, K. S. and TSAY, R. S. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika* **85** 413–426. MR1649122 https://doi.org/10.1093/biomet/85.2.413

[7] CHEN, C. W. S., SO, M. K. P. and LIU, F.-C. (2011). A review of threshold time series models in finance. *Stat. Interface* **4** 167–181. MR2812813 https://doi.org/10.4310/SII.2011.v4.n2.a12

[8] CHEN, H., CHONG, T. T.-L. and BAI, J. (2012). Theory and applications of TAR model with two threshold variables. *Econometric Rev.* **31** 142–170. MR2855216 https://doi.org/10.1080/07474938.2011.607100

[9] DAVIS, R. A., LEE, T. C. M. and RODRIGUEZ-YAM, G. A. (2008). Break detection for a class of nonlinear time series models. *J. Time Series Anal.* **29** 834–867. MR2450899 https://doi.org/10.1111/j.1467-9892.2008.00585.x

[10] DAVIS, R. A. and YAU, C. Y. (2013). Consistency of minimum description length model selection for piecewise stationary time series models. *Electron. J. Stat.* **7** 381–411. MR3020426 https://doi.org/10.1214/13-EJS769

[11] EMILSSON, V., THORLEIFSSON, G., ZHANG, B., LEONARDSON, A. S., ZINK, F., ZHU, J., CARLSON, S., HELGASON, A., WALTERS, G. B. et al. (2008). Genetics of gene expression and its effect on disease. *Nature* **452** 423–428.

MSC2020 subject classifications. 62M10.
Key words and phrases. Multivariate threshold model, nonlinear time series, asymptotic analysis, weak convergence.

[12] GIANNERINI, S., GORACCI, G. and RAHBEK, A. (2024). The validity of bootstrap testing for threshold autoregression. *J. Econometrics* **239** Paper No. 105379. MR4711065 https://doi.org/10.1016/j.jeconom.2023.01.004

[13] GIANNERINI, S., MAASOUMI, E. and BEE DAGUM, E. (2015). Entropy testing for nonlinear serial dependence in time series. *Biometrika* **102** 661–675. MR3394282 https://doi.org/10.1093/biomet/asv007

[14] GORACCI, G. (2021). An empirical study on the parsimony and descriptive power of TARMA models. *Stat. Methods Appl.* **30** 109–137. MR4221676 https://doi.org/10.1007/s10260-020-00516-8

[15] GORACCI, G., FERRARI, D., GIANNERINI, S. and RAVAZZOLO, F. (2025). Robust estimation for threshold autoregressive moving-average models. *J. Bus. Econom. Statist.* **43** 579–591. MR4918702 https://doi.org/10.1080/07350015.2024.2412011

[16] HANSEN, B. E. (2011). Threshold autoregression in economics [discussion of MR2812802]. *Stat. Interface* **4** 123–127. MR2812805 https://doi.org/10.4310/SII.2011.v4.n2.a4

[17] HANSEN, M. H. and YU, B. (2001). Model selection and the principle of minimum description length. *J. Amer. Statist. Assoc.* **96** 746–774. MR1939352 https://doi.org/10.1198/016214501753168398

[18] LEE, T. C. M. (2001). An introduction to coding theory and the two-part minimum description length principle. *Int. Stat. Rev.* **69** 169–183.

[19] LEWIS, P. A. W. and STEVENS, J. G. (1991). Nonlinear modeling of time series using multivariate adaptive regression splines. *J. Amer. Statist. Assoc.* **86** 864–877.

[20] LI, D., LI, W. K. and LING, S. (2011). On the least squares estimation of threshold autoregressive and moving-average models. *Stat. Interface* **4** 183–196. MR2812814 https://doi.org/10.4310/SII.2011.v4.n2.a13

[21] LI, D. and LING, S. (2012). On the least squares estimation of multiple-regime threshold autoregressive models. *J. Econometrics* **167** 240–253. MR2885449 https://doi.org/10.1016/j.jeconom.2011.11.006

[22] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295. MR0293706 https://doi.org/10.1214/aoms/1177693241

[23] QIAN, L. (1998). On maximum likelihood estimators for a threshold autoregression. *J. Statist. Plann. Inference* **75** 21–46. MR1671666 https://doi.org/10.1016/S0378-3758(98)00113-X

[24] RISSANEN, J. (1989). *Stochastic Complexity in Statistical Inquiry*. *World Scientific Series in Computer Science* **15**. World Scientific Co., Inc., Teaneck, NJ. MR1082556

[25] SEIJO, E. and SEN, B. (2011). Change-point in stochastic design regression and the bootstrap. *Ann. Statist.* **39** 1580–1607. MR2850213 https://doi.org/10.1214/11-AOS874

[26] TIAO, G. C. and TSAY, R. S. (1994). Some advances in non-linear and adaptive modelling in time-series. *J. Forecast.* **13** 109–131.

[27] TONG, H. (1978). On a threshold model. *Pattern Recogn. Signal Process.* **29** 575–586.

[28] TONG, H. (1983). *Threshold Models in Nonlinear Time Series Analysis*. *Lecture Notes in Statistics* **21**. Springer, New York. MR0717388 https://doi.org/10.1007/978-1-4684-7888-4

[29] TONG, H. (2015). Threshold models in time series analysis—some reflections. *J. Econometrics* **189** 485–491. MR3414916 https://doi.org/10.1016/j.jeconom.2015.03.039

[30] TONG, H. and LIM, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **42** 245–292.

[31] TSAY, R. S. (1998). Testing and modeling multivariate threshold models. *J. Amer. Statist. Assoc.* **93** 1188–1202. MR1649212 https://doi.org/10.2307/2669861

[32] TSAY, R. S. (2021). Multivariate hysteretic autoregressive models. *Statist. Sinica* **31** 2257–2274. MR4338084 https://doi.org/10.1002/rnc.5392

[33] YAU, C. Y., TANG, C. M. and LEE, T. C. M. (2015). Estimation of multiple-regime threshold autoregressive models with structural breaks. *J. Amer. Statist. Assoc.* **110** 1175–1186. MR3420693 https://doi.org/10.1080/01621459.2014.954706

[34] YU, P. (2014). The bootstrap in threshold regression. *Econometric Theory* **30** 676–714. MR3205610 https://doi.org/10.1017/S0266466614000012

[35] YUAN, G. and YAU, C. Y. (2026). Supplement to "Generalized multivariate threshold autoregressive models with linearly partitioned threshold space." https://doi.org/10.1214/26-AOS2638SUPP

[36] ZHANG, X., LI, D. and TONG, H. (2024). On the least squares estimation of multiple-threshold-variable autoregressive models. *J. Bus. Econom. Statist.* **42** 215–228. MR4684502 https://doi.org/10.1080/07350015.2023.2174124

The Annals of Statistics
2026, Vol. 54, No. 4, 2031–2054
https://doi.org/10.1214/26-AOS2639

PSEUDO-MAXIMUM LIKELIHOOD THEORY FOR HIGH-DIMENSIONAL RANK ONE INFERENCE

BY CURTIS GRANT[1,a], AUKOSH JAGANNATH[2,b] AND JUSTIN KO[2,c]

[1]*Department of Mathematics, Northwestern University,* [a]*curtisgrant2026@u.northwestern.edu*
[2]*Department of Statistics and Actuarial Science, University of Waterloo,* [b]*a.jagannath@uwaterloo.ca,* [c]*jko104@syr.edu*

We develop a pseudo-likelihood theory for rank one matrix estimation problems in the high-dimensional limit. We prove a variational principle for the limiting pseudo-maximum likelihood, which also characterizes the performance of the corresponding pseudo-maximum likelihood estimator. We show that this variational principle is universal and depends only on four parameters determined by the corresponding null model. Through this universality, we introduce a notion of equivalence for estimation problems of this type and, in particular, show that a broad class of estimation tasks, including community detection, sparse submatrix detection and nonlinear spiked matrix models, are equivalent to spiked matrix models. As an application, we obtain a complete description of the performance of the least-squares (or "best rank one") estimator for any rank one matrix estimation problem.

REFERENCES

[1] ABBE, E. (2017). Community detection and stochastic block models: Recent developments. *J. Mach. Learn. Res.* **18** Paper No. 177. MR3827065

[2] ABBE, E., BANDEIRA, A. S. and HALL, G. (2016). Exact recovery in the stochastic block model. *IEEE Trans. Inf. Theory* **62** 471–487. MR3447993 https://doi.org/10.1109/TIT.2015.2490670

[3] AUFFINGER, A. and CHEN, W.-K. (2015). The Parisi formula has a unique minimizer. *Comm. Math. Phys.* **335** 1429–1444. MR3320318 https://doi.org/10.1007/s00220-014-2254-z

[4] AUFFINGER, A. and CHEN, W.-K. (2017). Parisi formula for the ground state energy in the mixed p-spin model. *Ann. Probab.* **45** 4617–4631. MR3737919 https://doi.org/10.1214/16-AOP1173

[5] 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 https://doi.org/10.1214/009117905000000233

[6] BARBIER, J., CHEN, W.-K., PANCHENKO, D. and SÁENZ, M. (2025). Performance of Bayesian linear regression in a model with mismatch. *Inf. Inference* **14** Paper No. iaaf019. MR4934628 https://doi.org/10.1093/imaiai/iaaf019

[7] BARBIER, J., DIA, M., MACRIS, N., KRZAKALA, F., LESIEUR, T. and ZDEBOROVÁ, L. (2016). Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Proceedings of the* 30*th International Conference on Neural Information Processing Systems*. *NIPS*'16 424–432. Curran Associates, Red Hook, NY, USA.

[8] BARBIER, J., HOU, T., MONDELLI, M. and SÁENZ, M. (2022). The price of ignorance: How much does it cost to forget noise structure in low-rank matrix estimation?

[9] BARBIER, J. and REEVES, G. (2020). Information-theoretic limits of a multiview low-rank symmetric spiked matrix model. In 2020 *IEEE International Symposium on Information Theory* (*ISIT*) 2771–2776. IEEE Press, New York.

[10] BEN AROUS, G. and GUIONNET, A. (2008). The spectrum of heavy tailed random matrices. *Comm. Math. Phys.* **278** 715–751. MR2373441 https://doi.org/10.1007/s00220-007-0389-x

[11] BEN AROUS, G. and JAGANNATH, A. (2018). Spectral gap estimates in mean field spin glasses. *Comm. Math. Phys.* **361** 1–52. MR3825934 https://doi.org/10.1007/s00220-018-3152-6

[12] BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. MR2782201 https://doi.org/10.1016/j.aim.2011.02.007

MSC2020 subject classifications. Primary 62F12, 62H12; secondary 60K35, 82D30.

Key words and phrases. Pseudo-maximum likelihood theory, high-dimensional rank one inference, rank one, Gaussian equivalence, spin glasses, Parisi formula.

[13] BHAMIDI, S., DEY, P. S. and NOBEL, A. B. (2017). Energy landscape for large average submatrix detection problems in Gaussian random matrices. *Probab. Theory Related Fields* **168** 919–983. MR3663635 https://doi.org/10.1007/s00440-017-0766-0

[14] BUTUCEA, C. and INGSTER, Y. I. (2013). Detection of a sparse submatrix of a high-dimensional noisy matrix. *Bernoulli* **19** 2652–2688. MR3160567 https://doi.org/10.3150/12-BEJ470

[15] CAMILLI, F., CONTUCCI, P. and MINGIONE, E. (2022). An inference problem in a mismatched setting: A spin-glass model with Mattis interaction. *SciPost Phys*. **12** Paper No. 125. MR4409513 https://doi.org/10.21468/scipostphys.12.4.125

[16] CELENTANO, M., FAN, Z. and MEI, S. (2023). Local convexity of the TAP free energy and AMP convergence for $\mathbb{Z}_2$-synchronization. *Ann. Statist*. **51** 519–546. MR4600991 https://doi.org/10.1214/23-aos2257

[17] CHEN, H., MOURRAT, J.-C. and XIA, J. (2022). Statistical inference of finite-rank tensors. *Ann. Henri Lebesgue* **5** 1161–1189. MR4526249 https://doi.org/10.5802/ahl.146

[18] CHEN, H.-B. and XIA, J. (2022). Hamilton-Jacobi equations for inference of matrix tensor products. *Ann. Inst. Henri Poincaré Probab. Stat*. **58** 755–793. MR4421607 https://doi.org/10.1214/21-aihp1183

[19] COUILLET, R. and BENAYCH-GEORGES, F. (2016). Kernel spectral clustering of large dimensional data. *Electron. J. Stat*. **10** 1393–1454. MR3507369 https://doi.org/10.1214/16-EJS1144

[20] DE ALMEIDA, J. R. L. and THOULESS, D. J. (1978). Stability of the Sherrington-Kirkpatrick solution of a spin glass model. *J. Phys. A: Math. Gen*. **11** 983.

[21] DESHPANDE, Y. and MONTANARI, A. (2014). Information-theoretically optimal sparse PCA. In 2014 *IEEE International Symposium on Information Theory* 2197–2201. IEEE Press, New York.

[22] DOMINGUEZ, T. and MOURRAT, J.-C. (2024). Mutual information for the sparse stochastic block model. *Ann. Probab*. **52** 434–501. MR4718399 https://doi.org/10.1214/23-aop1665

[23] DONOHO, D. L., MALEKI, A. and MONTANARI, A. (2009). Message-passing algorithms for compressed sensing. *Proc. Natl. Acad. Sci. USA* **106** 18914–18919.

[24] ECKART, C. and YOUNG, G. M. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1** 211–218.

[25] EL ALAOUI, A., KRZAKALA, F. and JORDAN, M. (2020). Fundamental limits of detection in the spiked Wigner model. *Ann. Statist*. **48** 863–885. MR4102679 https://doi.org/10.1214/19-AOS1826

[26] EL ALAOUI, A., MONTANARI, A. and SELLKE, M. (2021). Optimization of mean-field spin glasses. *Ann. Probab*. **49** 2922–2960. MR4348682 https://doi.org/10.1214/21-aop1519

[27] EL KAROUI, N. (2010). The spectrum of kernel random matrices. *Ann. Statist*. **38** 1–50. MR2589315 https://doi.org/10.1214/08-AOS648

[28] EL KAROUI, N. (2010). On information plus noise kernel random matrices. *Ann. Statist*. **38** 3191–3216. MR2722468 https://doi.org/10.1214/10-AOS801

[29] FAN, Z. (2022). Approximate message passing algorithms for rotationally invariant matrices. *Ann. Statist*. **50** 197–224. MR4382014 https://doi.org/10.1214/21-aos2101

[30] FAN, Z., MEI, S. and MONTANARI, A. (2021). TAP free energy, spin glasses and variational inference. *Ann. Probab*. **49** 1–45. MR4203332 https://doi.org/10.1214/20-AOP1443

[31] FELDMAN, M. J. (2025). Spectral properties of elementwise-transformed spiked matrices. *SIAM J. Math. Data Sci*. **7** 542–571. MR4898690 https://doi.org/10.1137/23M1627985

[32] FENG, O. Y., VENKATARAMANAN, R., RUSH, C. and SAMWORTH, R. J. (2022). A unifying tutorial on approximate message passing. *Found. Trends Mach. Learn*. **15** 335–536.

[33] FÉRAL, D. and PÉCHÉ, S. (2007). The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys*. **272** 185–228. MR2291807 https://doi.org/10.1007/s00220-007-0209-3

[34] FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond., Ser. A, Contain. Pap. Math. Phys. Character* **222** 309–368.

[35] FLETCHER, A. K. and RANGAN, S. (2018). Iterative reconstruction of rank-one matrices in noise. *Inf. Inference* **7** 531–562. MR3858334 https://doi.org/10.1093/imaiai/iax014

[36] FU, T., LIU, Y., BARBIER, J., MONDELLI, M., LIANG, S. and HOU, T. (2023). Mismatched estimation of non-symmetric rank-one matrices corrupted by structured noise. In 2023 *IEEE International Symposium on Information Theory* (*ISIT*) 1178–1183.

[37] GAMARNIK, D., JAGANNATH, A. and SEN, S. (2021). The overlap gap property in principal submatrix recovery. *Probab. Theory Related Fields* **181** 757–814. MR4344133 https://doi.org/10.1007/s00440-021-01089-7

[38] GAMARNIK, D., JAGANNATH, A. and WEIN, A. S. (2024). Hardness of random optimization problems for Boolean circuits, low-degree polynomials, and Langevin dynamics. *SIAM J. Comput*. **53** 1–46. MR4704592 https://doi.org/10.1137/22M150263X

[39] GOURIÉROUX, C., MONFORT, A. and TROGNON, A. (1984). Pseudomaximum likelihood methods: Theory. *Econometrica* **52** 681–700. MR0740305 https://doi.org/10.2307/1913471

[40] GRANT, C., JAGANNATH, A. and KO, J. (2026). Supplement to "Pseudo-maximum likelihood theory for high-dimensional rank one inference." https://doi.org/10.1214/26-AOS2639SUPP

[41] GRANT, C., JAGANNATH, A. and KO, J. (2026). Pseudo-maximum likelihood theory for high-dimensional rank one inference II: Algorithms and hardness. In preparation.

[42] GUIONNET, A., KO, J., KRZAKALA, F., MERGNY, P. and ZDEBOROVÁ, L. (2023). Spectral Phase Transitions in Non-Linear Wigner Spiked Models. arXiv Preprint. Available at arXiv:2310.14055.

[43] GUIONNET, A., KO, J., KRZAKALA, F. and ZDEBOROVÁ, L. (2025). Estimating rank-one matrices with mismatched prior and noise: Universality and large deviations. *Comm. Math. Phys.* **406** Paper No. 9. MR4837935 https://doi.org/10.1007/s00220-024-05179-0

[44] GUIONNET, A., KO, J., KRZAKALA, F. and ZDEBOROVÁ, L. (2025). Low-rank matrix estimation with inhomogeneous noise. *Inf. Inference* **14** Paper No. iaaf010. MR4893427 https://doi.org/10.1093/imaiai/iaaf010

[45] HAJEK, B., WU, Y. and XU, J. (2017). Information limits for recovering a hidden community. *IEEE Trans. Inf. Theory* **63** 4729–4745. MR3683533 https://doi.org/10.1109/TIT.2017.2653804

[46] HAJEK, B., WU, Y. and XU, J. (2018). Recovering a hidden community beyond the Kesten-Stigum threshold in $O(|E|\log^*|V|)$ time. *J. Appl. Probab.* **55** 325–352. MR3832891 https://doi.org/10.1017/jpr.2018.22

[47] HOPKINS, S. B., KOTHARI, P. K., POTECHIN, A., RAGHAVENDRA, P., SCHRAMM, T. and STEURER, D. (2017). The power of sum-of-squares for detecting hidden structures. In 58*th Annual IEEE Symposium on Foundations of Computer Science—FOCS* 2017 720–731. IEEE Computer Soc., Los Alamitos, CA. MR3734275 https://doi.org/10.1109/FOCS.2017.72

[48] HOPKINS, S. B., SHI, J., SCHRAMM, T. and STEURER, D. (2016). Fast spectral algorithms from sum-of-squares proofs: Tensor decomposition and planted sparse vectors. In *STOC'*16*—Proceedings of the* 48*th Annual ACM SIGACT Symposium on Theory of Computing* 178–191. ACM, New York. MR3536564 https://doi.org/10.1145/2897518.2897529

[49] HOPKINS, S. B., SHI, J. and STEURER, D. (2015). Tensor principal component analysis via sum-of-square proofs. In *Conference on Learning Theory* 956–1006. PMLR.

[50] HUANG, B. and SELLKE, M. (2025). Tight Lipschitz hardness for optimizing mean field spin glasses. *Comm. Pure Appl. Math.* **78** 60–119. MR4822907 https://doi.org/10.1002/cpa.22222

[51] IBA, Y. (1999). The Nishimori line and Bayesian statistics. *J. Phys. A* **32** 3875–3888. MR1696814 https://doi.org/10.1088/0305-4470/32/21/302

[52] JAGANNATH, A., LOPATTO, P. and MIOLANE, L. (2020). Statistical thresholds for tensor PCA. *Ann. Appl. Probab.* **30** 1910–1933. MR4132641 https://doi.org/10.1214/19-AAP1547

[53] JAGANNATH, A. and SEN, S. (2021). On the unbalanced cut problem and the generalized Sherrington-Kirkpatrick model. *Annales de l'Institut Henri Poincaré D* **8** 35–88. MR4228619 https://doi.org/10.4171/aihpd/97

[54] JAGANNATH, A. and TOBASCO, I. (2016). A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.* **144** 3135–3150. MR3487243 https://doi.org/10.1090/proc/12968

[55] JAGANNATH, A. and TOBASCO, I. (2017). Low temperature asymptotics of spherical mean field spin glasses. *Comm. Math. Phys.* **352** 979–1017. MR3631397 https://doi.org/10.1007/s00220-017-2864-3

[56] JAVANMARD, A., MONTANARI, A. and RICCI-TERSENGHI, F. (2016). Phase transitions in semidefinite relaxations. *Proc. Natl. Acad. Sci. USA* **113** E2218–E2223. MR3494080 https://doi.org/10.1073/pnas.1523097113

[57] JAVANMARD, A., MONTANARI, A. and RICCI-TERSENGHI, F. (2016). Phase transitions in semidefinite relaxations. *Proc. Natl. Acad. Sci. USA* **113** E2218–E2223. MR3494080 https://doi.org/10.1073/pnas.1523097113

[58] 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 https://doi.org/10.1198/jasa.2009.0121

[59] KARRER, B. and NEWMAN, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* **83** 016107. MR2788206 https://doi.org/10.1103/PhysRevE.83.016107

[60] KUNISKY, D. (2025). Low coordinate degree algorithms I: Universality of computational thresholds for hypothesis testing. *Ann. Statist.* **53** 774–801. MR4900166 https://doi.org/10.1214/24-aos2484

[61] LELARGE, M. and MIOLANE, L. (2019). Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Related Fields* **173** 859–929. MR3936148 https://doi.org/10.1007/s00440-018-0845-x

[62] LESIEUR, T., KRZAKALA, F. and ZDEBOROVÁ, L. (2015). MMSE of probabilistic low-rank matrix estimation: Universality with respect to the output channel. In 2015 53*rd Annual Allerton Conference on Communication, Control, and Computing* (*Allerton*) 680–687. IEEE Press, New York.

[63] LESIEUR, T., KRZAKALA, F. and ZDEBOROVÁ, L. (2017). Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *J. Stat. Mech. Theory Exp.* **7** 073403. MR3683819 https://doi.org/10.1088/1742-5468/aa7284

[64] LOUART, C., LIAO, Z. and COUILLET, R. (2018). A random matrix approach to neural networks. *Ann. Appl. Probab.* **28** 1190–1248. MR3784498 https://doi.org/10.1214/17-AAP1328
[65] LUNEAU, C., BARBIER, J. and MACRIS, N. (2021). Mutual information for low-rank even-order symmetric tensor estimation. *Inf. Inference* **10** 1167–1207. MR4353122 https://doi.org/10.1093/imaiai/iaaa022
[66] MADEIRA, S. C. and OLIVEIRA, A. L. (2004). Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **1** 24–45.
[67] MAYYA, V. and REEVES, G. (2019). Mutual information in community detection with covariate information and correlated networks. In 2019 *57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* 602–607. IEEE Press, New York.
[68] MERGNY, P., KO, J., KRZAKALA, F. and ZDEBOROVÁ, L. (2024). Fundamental limits of non-linear low-rank matrix estimation. In *Proceedings of Thirty Seventh Conference on Learning Theory* (S. Agrawal and A. Roth, eds.). *Proceedings of Machine Learning Research* **247** 3873–3873. PMLR.
[69] MIOLANE, L. (2018). Phase transitions in spiked matrix estimation: Information-theoretic analysis. arXiv Preprint. Available at arXiv:1806.04343.
[70] MONTANARI, A. and WEIN, A. S. (2025). Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation. *Probab. Theory Related Fields* **191** 181–233. MR4869255 https://doi.org/10.1007/s00440-024-01322-z
[71] PANCHENKO, D. (2013). The Parisi ultrametricity conjecture. *Ann. of Math.* (2) **177** 383–393. MR2999044 https://doi.org/10.4007/annals.2013.177.1.8
[72] PARISI, G. (1979). Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.* **43** 1754.
[73] PENNINGTON, J. and WORAH, P. (2017). Nonlinear random matrix theory for deep learning. In *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds.) **30**. Curran Associates, Red Hook.
[74] PERRY, A., WEIN, A. S., BANDEIRA, A. S. and MOITRA, A. (2018). Optimality and sub-optimality of PCA I: Spiked random matrix models. *Ann. Statist.* **46** 2416–2451. MR3845022 https://doi.org/10.1214/17-AOS1625
[75] PICCOLO, V. and SCHRÖDER, D. (2021). Analysis of one-hidden-layer neural networks via the resolvent method. In *Advances in Neural Information Processing Systems* (M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang and J. W. Vaughan, eds.) **34** 5225–5235. Curran Associates, Red Hook.
[76] POURKAMALI, F. and MACRIS, N. (2022). Mismatched estimation of non-symmetric rank-one matrices under Gaussian noise. In 2022 *IEEE International Symposium on Information Theory (ISIT)* 1288–1293.
[77] REEVES, G. (2020). Information-theoretic limits for the matrix tensor product. *IEEE J. Sel. Areas Inf. Theory* **1** 777–798.
[78] RENFREW, D. and SOSHNIKOV, A. (2013). On finite rank deformations of Wigner matrices II: Delocalized perturbations. *Random Matrices Theory Appl.* **2** 1250015. MR3039820 https://doi.org/10.1142/S2010326312500153
[79] SCHERVISH, M. J. (1995). *Theory of Statistics. Springer Series in Statistics*. Springer, New York. MR1354146 https://doi.org/10.1007/978-1-4612-4250-5
[80] SHERRINGTON, D. and KIRKPATRICK, S. (1975). Solvable model of a spin-glass. *Phys. Rev. Lett.* **35** 1792–1796.
[81] SINGER, A. (2011). Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.* **30** 20–36. MR2737931 https://doi.org/10.1016/j.acha.2010.02.001
[82] SINGER, A. and SHKOLNISKY, Y. (2011). Three-dimensional structure determination from common lines in cryo-EM by eigenvectors and semidefinite programming. *SIAM J. Imaging Sci.* **4** 543–572. MR2810897 https://doi.org/10.1137/090767777
[83] TALAGRAND, M. (2006). The Parisi formula. *Ann. of Math.* (2) **163** 221–263. MR2195134 https://doi.org/10.4007/annals.2006.163.221
[84] TONINELLI, F. L. (2002). About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean-field spin glass model. *Europhys. Lett.* **60** 764.
[85] ZOU, H., HASTIE, T. and TIBSHIRANI, R. (2006). Sparse principal component analysis. *J. Comput. Graph. Statist.* **15** 265–286. MR2252527 https://doi.org/10.1198/106186006X113430

The Annals of Statistics
2026, Vol. 54, No. 4, 2055–2082
https://doi.org/10.1214/26-AOS2640
This research was funded, in whole or in part, by EPSRC, EP/T02772X/1. A CC BY 4.0 license is applied to this article arising from this submission, in accordance with the grant's open access conditions.

ESTIMATION BEYOND MISSING (COMPLETELY) AT RANDOM

BY TIANYI MA[1,a], KABIR A. VERCHAND[2,c], THOMAS B. BERRETT[3,d], TENGYAO WANG[4,e] AND RICHARD J. SAMWORTH[1,b]

[1]*Statistical Laboratory, University of Cambridge,* [a]*tm681@cam.ac.uk,* [b]*r.samworth@statslab.cam.ac.uk*

[2]*Department of Data Sciences and Operations, University of Southern California,* [c]*verchand@usc.edu*

[3]*Department of Statistics, University of Warwick,* [d]*tom.berrett@warwick.ac.uk*

[4]*Department of Statistics, London School of Economics and Political Science,* [e]*t.wang59@lse.ac.uk*

We study the effects of missingness on the estimation of population parameters. Moving beyond restrictive missing completely at random (MCAR) assumptions, we first formulate a missing data analogue of Huber's arbitrary ϵ-contamination model. For mean estimation with respect to squared Euclidean error loss, we show that the minimax quantiles decompose as a sum of the corresponding minimax quantiles under a heterogeneous, MCAR assumption and a robust error term, depending on ϵ, that reflects the additional error incurred by departure from MCAR.

We next introduce natural classes of *realisable ϵ-contamination models*, where an MCAR version of a base distribution P is contaminated by an arbitrary missing not at random (MNAR) version of P. These classes are rich enough to capture various notions of biased sampling and sensitivity conditions, yet we show that they enjoy improved minimax performance relative to our earlier arbitrary contamination classes for both parametric and nonparametric classes of base distributions. For instance, with a univariate Gaussian base distribution, consistent mean estimation over realisable ϵ-contamination classes is possible even when ϵ and the proportion of missingness converge (slowly) to 1. We extend our results to the setting of departures from missing at random (MAR) in normal linear regression with a realisable missing response, and also demonstrate that our methods can be made adaptive to the case of unknown ϵ.

REFERENCES

ARONOW, P. M. and LEE, D. K. K. (2013). Interval estimation of population means under unknown but bounded probabilities of sample selection. *Biometrika* **100** 235–240. MR3034337 https://doi.org/10.1093/biomet/ass064

BAKSHI, A. and PRASAD, A. (2021). Robust linear regression: Optimal rates in polynomial time. In *STOC '21—Proceedings of the* 53*rd Annual ACM SIGACT Symposium on Theory of Computing* 102–115. ACM, New York. MR4398827

BELLONI, A., ROSENBAUM, M. and TSYBAKOV, A. B. (2017). Linear and conic programming estimators in high dimensional errors-in-variables models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 939–956. MR3641415 https://doi.org/10.1111/rssb.12196

BERKELAAR, M. et al. (2023). lpSolve: Interface to 'Lp_solve' v. 5.5 to Solve Linear/Integer Programs. R package version 5.6.20.

BERRETT, T. B. and SAMWORTH, R. J. (2023). Optimal nonparametric testing of missing completely at random and its connections to compatibility. *Ann. Statist.* **51** 2170–2193. MR4678800 https://doi.org/10.1214/23-aos2326

BICKEL, P. J. and RITOV, J. (1991). Large sample theory of estimation in biased sampling regression models. I. *Ann. Statist.* **19** 797–816. MR1105845 https://doi.org/10.1214/aos/1176348121

BOLLINGER, C. R., HIRSCH, B. T., HOKAYEM, C. M. and ZILIAK, J. P. (2019). Trouble in the tails? What we know about earnings nonresponse 30 years after Lillard, Smith, and Welch. *J. Polit. Econ.* **127** 2143–2185.

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series*: *Theory and Methods*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4

MSC2020 subject classifications. Primary 62D10; secondary 62C20.
Key words and phrases. Missing data, Huber contamination model, missing not at random, robust estimation.

CAI, T. T. and WEI, H. (2021). Transfer learning for nonparametric classification: Minimax rate and adaptive classifier. *Ann*. *Statist*. **49** 100–128. MR4206671 https://doi.org/10.1214/20-AOS1949

CAI, T. T. and ZHANG, L. (2019). High dimensional linear discriminant analysis: Optimality, adaptive algorithm and missing data. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **81** 675–705. MR3997097

CHEN, M., GAO, C. and REN, Z. (2018). Robust covariance and scatter matrix estimation under Huber's contamination model. *Ann*. *Statist*. **46** 1932–1960. MR3845006 https://doi.org/10.1214/17-AOS1607

CRAVEN, B. D. and KOLIHA, J. J. (1977). Generalizations of Farkas' theorem. *SIAM J*. *Math*. *Anal*. **8** 983–997. MR0471302 https://doi.org/10.1137/0508076

CRESSIE, N. A. C. (2015). *Statistics for Spatial Data*, Revised ed. *Wiley Classics Library*. Wiley, New York. MR3559472

DALALYAN, A. S. and MINASYAN, A. (2022). All-in-one robust estimator of the Gaussian mean. *Ann*. *Statist*. **50** 1193–1219. MR4404933 https://doi.org/10.1214/21-aos2145

DASKALAKIS, C., GOULEAKIS, T., TZAMOS, C. and ZAMPETAKIS, M. (2018). Efficient statistics, in high dimensions, from truncated samples. In *59th Annual IEEE Symposium on Foundations of Computer Science—FOCS* 2018 639–649. IEEE Computer Soc., Los Alamitos, CA. MR3899629 https://doi.org/10.1109/FOCS.2018.00067

DEPERSIN, J. and LECUÉ, G. (2022a). Robust sub-Gaussian estimation of a mean vector in nearly linear time. *Ann*. *Statist*. **50** 511–536. MR4382026 https://doi.org/10.1214/21-aos2118

DEPERSIN, J. and LECUÉ, G. (2022b). Optimal robust mean and location estimation via convex programs with respect to any pseudo-norms. *Probab*. *Theory Related Fields* **183** 997–1025. MR4453320 https://doi.org/10.1007/s00440-022-01127-y

DEVROYE, L., LERASLE, M., LUGOSI, G. and OLIVEIRA, R. I. (2016). Sub-Gaussian mean estimators. *Ann*. *Statist*. **44** 2695–2725. MR3576558 https://doi.org/10.1214/16-AOS1440

DIAKONIKOLAS, I. and KANE, D. M. (2023). *Algorithmic High-Dimensional Robust Statistics*. Cambridge Univ. Press, Cambridge.

DIAKONIKOLAS, I., KANE, D. M. and PENSIA, A. (2020). Outlier robust mean estimation with subgaussian rates via stability. In *Advances in Neural Information Processing Systems* 1830–1840.

DIAKONIKOLAS, I., KANE, D. M., PITTAS, T. and ZARIFIS, N. (2024). Statistical query lower bounds for learning truncated Gaussians. In *Conference on Learning Theory* 1336–1363.

DO, K. T., WAHL, S., RAFFLER, J., MOLNOS, S., LAIMIGHOFER, M., ADAMSKI, J., SUHRE, K., STRAUCH, K., PETERS, A. et al. (2018). Characterization of missing values in untargeted MS-based metabolomics data and evaluation of missing data handling strategies. *Metabolomics* **14** 1–18.

ELSENER, A. and VAN DE GEER, S. (2019). Sparse spectral estimation with missing and corrupted measurements. *Stat* **8** e229. MR3978409 https://doi.org/10.1002/sta4.229

FAREWELL, D. M., DANIEL, R. M. and SEAMAN, S. R. (2022). Missing at random: A stochastic process perspective. *Biometrika* **109** 227–241. MR4374651 https://doi.org/10.1093/biomet/asab002

FARKAS, J. (1902). Theorie der einfachen Ungleichungen. *J*. *Reine Angew*. *Math*. **124** 1–27. MR1580578 https://doi.org/10.1515/crll.1902.124.1

FOLLAIN, B., WANG, T. and SAMWORTH, R. J. (2022). High-dimensional changepoint estimation with heterogeneous missingness. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **84** 1023–1055. MR4460584 https://doi.org/10.1111/rssb.12540

GAO, C. (2020). Robust regression via mutivariate regression depth. *Bernoulli* **26** 1139–1170. MR4058363 https://doi.org/10.3150/19-BEJ1144

GILL, R. D., VARDI, Y. and WELLNER, J. A. (1988). Large sample theory of empirical distributions in biased sampling models. *Ann*. *Statist*. **16** 1069–1112. MR0959189 https://doi.org/10.1214/aos/1176350948

HIRANO, K., IMBENS, G. W. and RIDDER, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71** 1161–1189. MR1995826 https://doi.org/10.1111/1468-0262.00442

HU, L. and REINGOLD, O. (2021). Robust mean estimation on highly incomplete data with arbitrary outliers. In *Conference on Artificial Intelligence and Statistics* 1558–1566.

HUA, V., NGUYEN, T., DAO, M.-S., NGUYEN, H. D. and NGUYEN, B. T. (2024). The impact of data imputation on air quality prediction problem. *PLoS ONE* **19** e0306303.

HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann*. *Math*. *Statist*. **35** 73–101. MR0161415 https://doi.org/10.1214/aoms/1177703732

JANA, S., FAN, J. and KULKARNI, S. (2024). A general theory for robust clustering via trimmed mean. arXiv Preprint. Available at arXiv:2401.05574.

KENNEDY, C., BLUMENTHAL, M., CLEMENT, S., CLINTON, J. D., DURAND, C., FRANKLIN, C., MCGEENEY, K., MIRINGOFF, L., OLSON, K. et al. (2018). An evaluation of the 2016 election polls in the United States. *Public Opin*. *Q*. **82** 1–33.

KONTONIS, V., TZAMOS, C. and ZAMPETAKIS, M. (2019). Efficient truncated statistics with unknown truncation. In 2019 *IEEE* 60*th Annual Symposium on Foundations of Computer Science* 1578–1595. IEEE Comput. Soc. Press, Los Alamitos, CA. MR4228243 https://doi.org/10.1109/FOCS.2019.00093

KWAK, S. K. and KIM, J. H. (2017). Statistical data preparation: Management of missing values and outliers. *Korean J. Anesthesiol.* **70** 407.

LERASLE, M. and OLIVEIRA, R. I. (2011). Robust empirical mean estimators. arXiv preprint. Available at arXiv:1112.3914.

LITTLE, R. J. A. and RUBIN, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR1925014 https://doi.org/10.1002/9781119013563

LIU, A. and MOITRA, A. (2023). Robustly learning general mixtures of Gaussians. *J. ACM* **70** Art. 21. MR4596518 https://doi.org/10.1145/3583680

LOH, P.-L. and TAN, X. L. (2018). High-dimensional robust precision matrix estimation: Cellwise corruption under ϵ-contamination. *Electron. J. Stat.* **12** 1429–1467. MR3804842 https://doi.org/10.1214/18-EJS1427

LOH, P.-L. and WAINWRIGHT, M. J. (2012). High-dimensional regression with noisy and missing data: Provable guarantees with nonconvexity. *Ann. Statist.* **40** 1637–1664. MR3015038 https://doi.org/10.1214/12-AOS1018

LOUNICI, K. (2014). High-dimensional covariance matrix estimation with missing observations. *Bernoulli* **20** 1029–1058. MR3217437 https://doi.org/10.3150/12-BEJ487

LUGOSI, G. and MENDELSON, S. (2021). Robust multivariate mean estimation: The optimality of trimmed mean. *Ann. Statist.* **49** 393–410. MR4206683 https://doi.org/10.1214/20-AOS1961

MA, T., VERCHAND, K. A., BERRETT, T. B., WANG, T. and SAMWORTH, R. J. (2026). Supplement to "Estimation beyond missing (completely) at random." https://doi.org/10.1214/26-AOS2640SUPP

MA, T., VERCHAND, K. A. and SAMWORTH, R. J. (2026). High-probability minimax lower bounds. *Statist. Sci.* **41** 637–658. MR5109066 https://doi.org/10.1214/25-STS996

MCCAFFREY, D. F. and LOCKWOOD, J. R. (2011). Missing data in value-added modeling of teacher effects. *Ann. Appl. Stat.* **5** 773–797. MR2840175 https://doi.org/10.1214/10-AOAS405

MCKENNAN, C., OBER, C. and NICOLAE, D. (2020). Estimation and inference in metabolomics with nonrandom missing data and latent factors. *Ann. Appl. Stat.* **14** 789–808. MR4117830 https://doi.org/10.1214/20-AOAS1328

MINSKER, S. and NDAOUD, M. (2021). Robust and efficient mean estimation: An approach based on the properties of self-normalized sums. *Electron. J. Stat.* **15** 6036–6070. MR4355703 https://doi.org/10.1214/21-ejs1925

OTTOSEN, T.-B. and KUMAR, P. (2019). Outlier detection and gap filling methodologies for low-cost air quality measurements. *Environ. Sci.: Process. Impacts* **21** 701–713.

PENSIA, A., JOG, V. and LOH, P.-L. (2025). Robust regression with covariate filtering: Heavy tails and adversarial contamination. *J. Amer. Statist. Assoc.* **120** 1002–1013. MR4917273 https://doi.org/10.1080/01621459.2024.2392906

PRINCE, M. (2012). Epidemiology. In *Core Psychiatry* (P. Wright, J. Stern and M. Phelan, eds.) 115–129. Elsevier Health Sciences.

REEVE, H. W. J., CANNINGS, T. I. and SAMWORTH, R. J. (2021). Adaptive transfer learning. *Ann. Statist.* **49** 3618–3649. MR4352543 https://doi.org/10.1214/21-aos2102

ROSENBAUM, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika* **74** 13–26. MR0885915 https://doi.org/10.1093/biomet/74.1.13

RUBIN, D. (1983). Imputing income in the CPS: Comments on "Measures of aggregate labor cost in the United States". In *The Measurement of Labor Cost* 333–344. Univ. Chicago Press, Chicago.

SAHOO, R., LEI, L. and WAGER, S. (2022). Learning from a biased sample. arXiv preprint. Available at arXiv:2209.01754.

SEAMAN, S., GALATI, J., JACKSON, D. and CARLIN, J. (2013). What is meant by "missing at random"? *Statist. Sci.* **28** 257–268. MR3112409 https://doi.org/10.1214/13-sts415

SELL, T., BERRETT, T. B. and CANNINGS, T. I. (2024). Nonparametric classification with missing data. *Ann. Statist.* **52** 1178–1200. MR4784074 https://doi.org/10.1214/24-aos2389

TUKEY, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians* (*Vancouver, B.C.*, 1974), *Vol.* 2 523–531. Canad. Math. Congr., Montreal, QC. MR0426989

VARDI, Y. (1985). Empirical distributions in selection bias models. *Ann. Statist.* **13** 178–205. MR0773161 https://doi.org/10.1214/aos/1176346585

VERSHYNIN, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge Univ. Press, Cambridge. MR3837109 https://doi.org/10.1017/9781108231596

XIE, Y., HUANG, J. and WILLETT, R. (2012). Change-point detection for high-dimensional time series with missing data. *IEEE J. Sel. Top. Signal Process.* **7** 12–27.

YAN, Y., CHEN, Y. and FAN, J. (2024). Inference for heteroskedastic PCA with missing data. *Ann. Statist.* **52** 729–756. MR4744194 https://doi.org/10.1214/24-aos2366

GREENLEES, J. S., REECE, W. S. and ZIESCHANG, K. D. (1982). Imputation of missing values when the probability of response depends on the variable being imputed. *J. Amer. Statist. Assoc.* **77** 251–261.

ZHAO, Q., SMALL, D. S. and BHATTACHARYA, B. B. (2019). Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 735–761. MR3997099

ZHU, Z., WANG, T. and SAMWORTH, R. J. (2022). High-dimensional principal component analysis with heterogeneous missingness. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 2000–2031. MR4515564 https://doi.org/10.1111/rssb.12550

The Annals of Statistics
2026, Vol. 54, No. 4, 2083–2109
https://doi.org/10.1214/26-AOS2641

FAST MIXING OF DATA AUGMENTATION ALGORITHMS: BAYESIAN PROBIT, LOGIT AND LASSO REGRESSION

BY HOLDEN LEE[a] AND KEXIN ZHANG[b]

Department of Applied Mathematics and Statistics, Johns Hopkins University, [a]*hlee283@jhu.edu,* [b]*kzhang91@jhu.edu*

We propose using a modified conductance-based method to study the mixing time of an important class of two-block Gibbs samplers, the data augmentation (DA) algorithm. Using this method, we prove the first non-asymptotic polynomial upper bounds on mixing times of three important DA algorithms: DA algorithms for Bayesian Probit regression (ProbitDA) and Bayesian Logit regression (LogitDA), and Bayesian Lasso Regression (LassoDA). Concretely, for ProbitDA and LogitDA, we demonstrate a tight bound that explicitly depends on the design matrix and prior covariance matrix. Under the assumption that data are independently generated from either a sub-Gaussian or log-concave distribution and properly scaled, the bound implies that with η-warm start, parameter dimension d, and sample size n, with high probability over data, the two algorithms require $\mathcal{O}(n\log(\frac{\log\eta}{\epsilon}))$ steps to obtain samples with at most ϵ error in TV, KL or χ^2 distance. Meanwhile, we show that under minimal data assumptions, LassoDA requires $\mathcal{O}(d^2(d\log d + n\log n)^2\log(\frac{\eta}{\epsilon}))$ steps to achieve ϵ-accuracy in TV distance. The results are generally applicable to settings with large n and large d, including settings with highly imbalanced response data in Probit and Logit regression. We compare them with the best known guarantees of Langevin Monte Carlo and Metropolis Adjusted Langevin Algorithm. We evaluate our theoretical results using numerical examples, and discuss the mixing times of the three algorithms under feasible initialization.

REFERENCES

[1] ADAMCZAK, R., LITVAK, A. E., PAJOR, A. and TOMCZAK-JAEGERMANN, N. (2010). Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles. *J. Amer. Math. Soc.* **23** 535–561. MR2601042 https://doi.org/10.1090/S0894-0347-09-00650-X

[2] ADAMCZAK, R., LITVAK, A. E., PAJOR, A. and TOMCZAK-JAEGERMANN, N. (2011). Sharp bounds on the rate of convergence of the empirical covariance matrix. *C. R. Math. Acad. Sci. Paris* **349** 195–200. MR2769907 https://doi.org/10.1016/j.crma.2010.12.014

[3] ALBERT, J. H. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. MR1224394

[4] ALONSO-GUTIÉRREZ, D. and BASTERO, J. (2015). *Approaching the Kannan-Lovász-Simonovits and Variance Conjectures*. *Lecture Notes in Math.* **2131**. Springer, Cham. MR3308635 https://doi.org/10.1007/978-3-319-13263-1

[5] ANDREWS, D. F. and MALLOWS, C. L. (1974). Scale mixtures of normal distributions. *J. Roy. Statist. Soc. Ser. B, Methodol.* **36** 99–102. MR0359122

[6] ASCOLANI, F., LAVENANT, H. and ZANELLA, G. (2024). Entropy contraction of the Gibbs sampler under log-concavity. arXiv Preprint. Available at arXiv:2410.00858.

[7] BARTHE, F. and KLARTAG, B. (2020). Spectral gaps, symmetries and log-concave perturbations. *Bull. Hellenic Math. Soc.* **64** 1–31. MR4075908

[8] BARTHE, F. and MILMAN, E. (2013). Transference principles for log-Sobolev and spectral-gap with applications to conservative spin systems. *Comm. Math. Phys.* **323** 575–625. MR3096532 https://doi.org/10.1007/s00220-013-1782-2

MSC2020 subject classifications. Primary 60J05, 62F15, 65C05; secondary 62J07, 62J12.
Key words and phrases. MCMC algorithm, Gibbs sampling, data augmentation algorithm, log-concave sampling, nonlog-concave sampling, conductance method, isoperimetric inequality.

[9] BOBKOV, S. G. (1999). Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.* **27** 1903–1921. MR1742893 https://doi.org/10.1214/aop/1022874820

[10] BOBKOV, S. G. and HOUDRÉ, C. (1997). Isoperimetric constants for product probability measures. *Ann. Probab.* **25** 184–205. MR1428505 https://doi.org/10.1214/aop/1024404284

[11] CAFFARELLI, L. A. (2000). Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.* **214** 547–563. MR1800860 https://doi.org/10.1007/s002200000257

[12] CATTIAUX, P. and GUILLIN, A. (2020). On the Poincaré constant of log-concave measures. In *Geometric Aspects of Functional Analysis. Vol. I. Lecture Notes in Math.* **2256** 171–217. Springer, Cham. MR4175748 https://doi.org/10.1007/978-3-030-36020-7_9

[13] CATTIAUX, P. and GUILLIN, A. (2022). Functional inequalities for perturbed measures with applications to log-concave measures and to some Bayesian problems. *Bernoulli* **28** 2294–2321. MR4474544 https://doi.org/10.3150/21-bej1419

[14] CHEN, Y., DWIVEDI, R., WAINWRIGHT, M. J. and YU, B. (2018). Fast MCMC sampling algorithms on polytopes. *J. Mach. Learn. Res.* **19** Paper No. 55. MR3874163

[15] CHEN, Y., DWIVEDI, R., WAINWRIGHT, M. J. and YU, B. (2020). Fast mixing of metropolized Hamiltonian Monte Carlo: Benefits of multi-step gradients. *J. Mach. Learn. Res.* **21** Paper No. 92. MR4119160

[16] CHEWI, S. (2023). Log-concave sampling. Book Draft Available at https://chewisinho.github.io.

[17] CHOI, H. M. and HOBERT, J. P. (2013). The Polya-gamma Gibbs sampler for Bayesian logistic regression is uniformly ergodic. *Electron. J. Stat.* **7** 2054–2064. MR3091616 https://doi.org/10.1214/13-EJS837

[18] COHEN, A. and EINAV, L. (2007). Estimating risk preferences from deductible choice. *Amer. Econ. Rev.* **97** 745–788.

[19] COURTADE, T. A. (2020). Bounds on the Poincaré constant for convolution measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 566–579. MR4059000 https://doi.org/10.1214/19-AIHP973

[20] COUSINS, B. and VEMPALA, S. (2014). A cubic algorithm for computing Gaussian volume. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms* 1215–1228. ACM, New York. MR3376450 https://doi.org/10.1137/1.9781611973402.90

[21] DAI, Y., GAO, Y., HUANG, J., JIAO, Y., KANG, L. and LIU, J. (2023). Lipschitz Transport Maps via the Follmer Flow. arXiv Preprint. Available at arXiv:2309.03490.

[22] DALALYAN, A. (2017). Further and stronger analogy between sampling and optimization: Langevin Monte Carlo and gradient descent. In *Conference on Learning Theory* 678–689. PMLR.

[23] DIEBOLT, J. and ROBERT, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *J. Roy. Statist. Soc. Ser. B, Methodol.* **56** 363–375. MR1281940

[24] DURANTE, D. and DUNSON, D. B. (2018). Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* **13** 29–58. MR3737942 https://doi.org/10.1214/16-BA1030

[25] DVORZAK, M. and WAGNER, H. (2016). Sparse Bayesian modelling of underreported count data. *Stat. Model.* **16** 24–46. MR3457686 https://doi.org/10.1177/1471082X15588398

[26] DWIVEDI, R., CHEN, Y., WAINWRIGHT, M. J. and YU, B. (2019). Log-concave sampling: Metropolis–Hastings algorithms are fast. *J. Mach. Learn. Res.* **20** Paper No. 183. MR4048994

[27] DYER, M., FRIEZE, A. and KANNAN, R. (1991). A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.* **38** 1–17. MR1095916 https://doi.org/10.1145/102782.102783

[28] ELDAN, R. (2013). Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.* **23** 532–569. MR3053755 https://doi.org/10.1007/s00039-013-0214-y

[29] FRÜHWIRTH-SCHNATTER, S. and FRÜHWIRTH, R. (2007). Auxiliary mixture sampling with applications to logistic models. *Comput. Statist. Data Anal.* **51** 3509–3528. MR2367817 https://doi.org/10.1016/j.csda.2006.10.006

[30] FUGLSTAD, G.-A., HEM, I. G., KNIGHT, A., RUE, H. and RIEBLER, A. (2020). Intuitive joint priors for variance parameters. *Bayesian Anal.* **15** 1109–1137. MR4171145 https://doi.org/10.1214/19-BA1185

[31] GRANT, E. H. C., MILLER, D. A., SCHMIDT, B. R., ADAMS, M. J., AMBURGEY, S. M., CHAMBERT, T., CRUICKSHANK, S. S., FISHER, R. N., GREEN, D. M. et al. (2016). Quantitative evidence for the effects of multiple drivers on continental-scale amphibian declines. *Sci. Rep.* **6** 25625.

[32] GRIFFIN, J. E., MATECHOU, E., BUXTON, A. S., BORMPOUDAKIS, D. and GRIFFITHS, R. A. (2020). Modelling environmental DNA data; Bayesian variable selection accounting for false positive and false negative errors. *J. R. Stat. Soc., Ser. C, Appl. Stat.* **69** 377–392. MR4098953 https://doi.org/10.1111/rssc.12390

[33] HANS, C. (2009). Bayesian lasso regression. *Biometrika* **96** 835–845. MR2564494 https://doi.org/10.1093/biomet/asp047

[34] HOBERT, J. P. (2011). The data augmentation algorithm: Theory and methodology. In *Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC Handb. Mod. Stat. Methods* 253–293. CRC Press, Boca Raton, FL. MR2858452

[35] HOLLEY, R. and STROOCK, D. (1987). Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.* **46** 1159–1194. MR0893137 https://doi.org/10.1007/BF01011161

[36] HOLMES, C. C. and HELD, L. (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Anal.* **1** 145–168. MR2227368 https://doi.org/10.1214/06-BA105

[37] JERRUM, M. and SINCLAIR, A. (1989). Approximating the permanent. *SIAM J. Comput.* **18** 1149–1178. MR1025467 https://doi.org/10.1137/0218077

[38] JOHNDROW, J. E., SMITH, A., PILLAI, N. and DUNSON, D. B. (2019). MCMC for imbalanced categorical data. *J. Amer. Statist. Assoc.* **114** 1394–1403. MR4011787 https://doi.org/10.1080/01621459.2018.1505626

[39] JONES, G. L. and HOBERT, J. P. (2001). Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statist. Sci.* **16** 312–334. MR1888447 https://doi.org/10.1214/ss/1015346317

[40] JOSEPH, L., GYORKOS, T. W. and COUPAL, L. (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Amer. J. Epidemiol.* **141** 263–272.

[41] JUSTINIANO, A. and PRIMICERI, G. E. (2008). The time-varying volatility of macroeconomic fluctuations. *Amer. Econ. Rev.* **98** 604–641.

[42] KANNAN, R., LOVÁSZ, L. and SIMONOVITS, M. (1995). Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* **13** 541–559. MR1318794 https://doi.org/10.1007/BF02574061

[43] KANNAN, R., LOVÁSZ, L. and SIMONOVITS, M. (1997). Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures Algorithms* **11** 1–50. MR1608200 https://doi.org/10.1002/(SICI)1098-2418(199708)11:1<1::AID-RSA1>3.0.CO;2-X

[44] KHARE, K. and HOBERT, J. P. (2013). Geometric ergodicity of the Bayesian lasso. *Electron. J. Stat.* **7** 2150–2163. MR3104915 https://doi.org/10.1214/13-EJS841

[45] KIM, Y.-H. and MILMAN, E. (2012). A generalization of Caffarelli's contraction theorem via (reverse) heat flow. *Math. Ann.* **354** 827–862. MR2983070 https://doi.org/10.1007/s00208-011-0749-x

[46] KLARTAG, B. (2023). Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inven. Anal.* Paper No. 4. MR4603941 https://doi.org/10.15781/jsjy-0b06

[47] KOLESNIKOV, A. V. (2011). Mass transportation and contractions. arXiv Preprint. Available at arXiv:1103.1479.

[48] LEE, H. and ZHANG, K. (2026). Supplement to "Fast mixing of data augmentation algorithms: Bayesian probit, logit and lasso regression." https://doi.org/10.1214/26-AOS2641SUPPA, https://doi.org/10.1214/26-AOS2641SUPPB

[49] LEE, Y. T. and VEMPALA, S. S. (2017). Eldan's stochastic localization and the KLS hyperplane conjecture: An improved lower bound for expansion. In *58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017* 998–1007. IEEE Computer Soc., Los Alamitos, CA. MR3734299 https://doi.org/10.1109/FOCS.2017.96

[50] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR3726904 https://doi.org/10.1090/mbk/107

[51] LIU, J. S., WONG, W. H. and KONG, A. (1994). Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika* **81** 27–40. MR1279653 https://doi.org/10.1093/biomet/81.1.27

[52] LOVÁSZ, L. (1999). Hit-and-run mixes fast. *Math. Program.* **86** 443–461. MR1733749 https://doi.org/10.1007/s101070050099

[53] LOVÁSZ, L. and SIMONOVITS, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms* **4** 359–412. MR1238906 https://doi.org/10.1002/rsa.3240040402

[54] LOVÁSZ, L. and VEMPALA, S. (2004). Hit-and-run is fast and fun. Available at https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-2003-05.pdf.

[55] LOVÁSZ, L. and VEMPALA, S. (2004). Hit-and-run from a corner. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing* 310–314. ACM, New York. MR2121613 https://doi.org/10.1145/1007352.1007403

[56] LOVÁSZ, L. and VEMPALA, S. (2007). The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms* **30** 307–358. MR2309621 https://doi.org/10.1002/rsa.20135

[57] MALLICK, H. and YI, N. (2014). A new Bayesian lasso. *Stat. Interface* **7** 571–582. MR3302383 https://doi.org/10.4310/SII.2014.v7.n4.a12

[58] MANGOUBI, O. and SMITH, A. (2021). Mixing of Hamiltonian Monte Carlo on strongly log-concave distributions: Continuous dynamics. *Ann. Appl. Probab.* **31** 2019–2045. MR4332690 https://doi.org/10.1214/20-aap1640

[59] MIKULINCER, D. and SHENFELD, Y. (2024). The Brownian transport map. *Probab. Theory Related Fields* **190** 379–444. MR4797372 https://doi.org/10.1007/s00440-024-01286-0

[60] MILMAN, E. (2010). Isoperimetric and concentration inequalities: Equivalence under curvature lower bound. *Duke Math. J.* **154** 207–239. MR2682183 https://doi.org/10.1215/00127094-2010-038

[61] MILMAN, E. (2012). Properties of isoperimetric, functional and transport-entropy inequalities via concentration. *Probab. Theory Related Fields* **152** 475–507. MR2892954 https://doi.org/10.1007/s00440-010-0328-1

[62] MILMAN, E. and SODIN, S. (2008). An isoperimetric inequality for uniformly log-concave measures and uniformly convex bodies. *J. Funct. Anal.* **254** 1235–1268. MR2386937 https://doi.org/10.1016/j.jfa.2007.12.002

[63] MONTANARI, A. and WU, Y. (2026). Provably efficient posterior sampling for sparse linear regression via measure decomposition. *J. Amer. Statist. Assoc.* **121** 636–654. MR5061149 https://doi.org/10.1080/01621459.2025.2537461

[64] MOU, W., HO, N., WAINWRIGHT, M. J., BARTLETT, P. L. and JORDAN, M. I. (2019). Sampling for Bayesian mixture models: Mcmc with polynomial-time mixing. arXiv Preprint. Available at arXiv:1912.05153.

[65] NARAYANAN, H. (2016). Randomized interior point methods for sampling and optimization. *Ann. Appl. Probab.* **26** 597–641. MR3449327 https://doi.org/10.1214/15-AAP1104

[66] NGUYEN, H. H., NGO, V. M. and TRAN, A. N. T. (2021). Financial performances, entrepreneurial factors and coping strategy to survive in the COVID-19 pandemic: Case of Vietnam. *Res. Int. Bus. Finance* **56** 101380.

[67] PARK, T. and CASELLA, G. (2008). The Bayesian lasso. *J. Amer. Statist. Assoc.* **103** 681–686. MR2524001 https://doi.org/10.1198/016214508000000337

[68] POLSON, N. G., SCOTT, J. G. and WINDLE, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *J. Amer. Statist. Assoc.* **108** 1339–1349. MR3174712 https://doi.org/10.1080/01621459.2013.829001

[69] QIN, Q. and HOBERT, J. P. (2019). Convergence complexity analysis of Albert and Chib's algorithm for Bayesian probit regression. *Ann. Statist.* **47** 2320–2347. MR3953453 https://doi.org/10.1214/18-AOS1749

[70] QIN, Q. and HOBERT, J. P. (2021). On the limitations of single-step drift and minorization in Markov chain convergence analysis. *Ann. Appl. Probab.* **31** 1633–1659. MR4312841 https://doi.org/10.1214/20-aap1628

[71] QIN, Q. and HOBERT, J. P. (2022). Wasserstein-based methods for convergence complexity analysis of MCMC with applications. *Ann. Appl. Probab.* **32** 124–166. MR4386523 https://doi.org/10.1214/21-aap1673

[72] RAJARATNAM, B. and SPARKS, D. (2015). MCMC-based inference in the era of big data: a fundamental analysis of the convergence complexity of high-dimensional chains. arXiv Preprint. Available at arXiv:1508.00947.

[73] RAJARATNAM, B., SPARKS, D., KHARE, K. and ZHANG, L. (2015). Scalable Bayesian shrinkage and uncertainty quantification for high-dimensional regression. arXiv preprint. Available at arXiv:1509.03697.

[74] ROBERTS, G. O. and ROSENTHAL, J. S. (2001). Markov chains and de-initializing processes. *Scand. J. Stat.* **28** 489–504. MR1858413 https://doi.org/10.1111/1467-9469.00250

[75] ROSENTHAL, J. S. (1995). Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **90** 558–566. MR1340509

[76] ROY, V. and HOBERT, J. P. (2007). Convergence rates and asymptotic standard errors for Markov chain Monte Carlo algorithms for Bayesian probit regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 607–623. MR2370071 https://doi.org/10.1111/j.1467-9868.2007.00602.x

[77] ROY, V., KHARE, K. and HOBERT, J. P. (2024). The data augmentation algorithm. arXiv preprint. Available at arXiv:2406.10464.

[78] SAMPFORD, M. R. (1953). Some inequalities on Mill's ratio and related functions. *Ann. Math. Stat.* **24** 130–132. MR0054890 https://doi.org/10.1214/aoms/1177729093

[79] SIMPSON, D., RUE, H., RIEBLER, A., MARTINS, T. G. and SØRBYE, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statist. Sci.* **32** 1–28. MR3634300 https://doi.org/10.1214/16-STS576

[80] SINCLAIR, A. and JERRUM, M. (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. Comput.* **82** 93–133. MR1003059 https://doi.org/10.1016/0890-5401(89)90067-9

[81] TALAGRAND, M. (1991). A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric Aspects of Functional Analysis* (1989–90). *Lecture Notes in Math.* **1469** 94–124. Springer, Berlin. MR1122615 https://doi.org/10.1007/BFb0089217

[82] TALAGRAND, M. (1996). Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** 587–600. MR1392331 https://doi.org/10.1007/BF02249265

[83] TANNER, M. A. and WONG, W. H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82** 528–550. MR0898357

[84] TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B, Methodol.* **58** 267–288. MR1379242

[85] VAN DYK, D. A. and MENG, X.-L. (2001). The art of data augmentation. *J. Comput. Graph. Statist.* **10** 1–111. MR1936358 https://doi.org/10.1198/10618600152418584

[86] VEMPALA, S. and WIBISONO, A. (2019). Rapid convergence of the unadjusted Langevin algorithm: Isoperimetry suffices. *Adv. Neural Inf. Process. Syst.* **32**.

[87] VERSHYNIN, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science. Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge Univ. Press, Cambridge. MR3837109 https://doi.org/10.1017/9781108231596

[88] WIBISONO, A. (2019). Proximal Langevin algorithm: Rapid convergence under isoperimetry. arXiv Preprint. Available at arXiv:1911.01469.

[89] XUN, G., LI, Y., ZHAO, W. X., GAO, J. and ZHANG, A. (2017). A correlated topic model using word embeddings. In *IJCAI* **17** 4207–4213.

[90] ZENS, G., FRÜHWIRTH-SCHNATTER, S. and WAGNER, H. (2024). Ultimate Pólya gamma samplers—efficient MCMC for possibly imbalanced binary and categorical data. *J. Amer. Statist. Assoc.* **119** 2548–2559. MR4833896 https://doi.org/10.1080/01621459.2023.2259030

[91] ZHANG, Z., SINHA, S., MAITI, T. and SHIPP, E. (2018). Bayesian variable selection in the accelerated failure time model with an application to the surveillance, epidemiology, and end results breast cancer data. *Stat. Methods Med. Res.* **27** 971–990. MR3770130 https://doi.org/10.1177/0962280215626947

The Annals of Statistics
2026, Vol. 54, No. 4, 2110–2135
https://doi.org/10.1214/26-AOS2643

STATISTICAL IMPOSSIBILITY AND POSSIBILITY OF ALIGNING LLMS WITH HUMAN PREFERENCES: FROM CONDORCET PARADOX TO NASH EQUILIBRIUM

BY KAIZHAO LIU[1,a], QI LONG[2,b], ZHEKUN SHI[3,d], WEIJIE J. SU[4,e] AND JIANCONG XIAO[2,c]

[1]*Department of Mathematics, Massachusetts Institute of Technology,* [a]*mrzt@mit.edu*

[2]*Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania,* [b]*qlong@upenn.edu,* [c]*jcxiao@upenn.edu*

[3]*Department of Operations Research and Financial Engineering, Princeton University,* [d]*zs3417@princeton.edu*

[4]*Department of Statistics and Data Science, University of Pennsylvania,* [e]*suw@wharton.upenn.edu*

Aligning large language models (LLMs) with diverse human preferences is critical for ensuring fairness and informed outcomes when deploying these models for decision-making. In this paper, we seek to uncover fundamental statistical limits concerning aligning LLMs with human preferences, with a focus on the probabilistic representation of human preferences and the preservation of diverse preferences in aligned LLMs. We first show that human preferences can be represented by a reward model if and only if the preference among LLM-generated responses is free of any Condorcet cycle. Moreover, we prove that Condorcet cycles exist with probability converging to one exponentially fast under a general probabilistic preference model called the Luce model, thereby demonstrating the *impossibility* of fully aligning human preferences using reward-based approaches such as reinforcement learning from human feedback. Next, we explore the conditions under which LLMs would employ mixed strategies—meaning they do not collapse to a single response—when aligned in the limit using a nonreward-based approach, such as Nash learning from human feedback. We identify a necessary and sufficient condition for mixed strategies: the *absence* of a response that is preferred over all others by a majority. As a blessing, we prove that this condition holds with high probability under the Luce model, thereby highlighting the statistical *possibility* of preserving minority preferences without explicit regularization in aligning LLMs.

REFERENCES

[1] ANTHROPIC, A. (2024). The Claude 3 model family: Opus, Sonnet, Haiku. Claude-3 Model Card.

[2] AZAR, M. G., GUO, Z. D., PIOT, B., MUNOS, R., ROWLAND, M., VALKO, M. and CALANDRIELLO, D. (2024). A general theoretical paradigm to understand learning from human preferences. In *International Conference on Artificial Intelligence and Statistics* 4447–4455. PMLR.

[3] BECKER, G. M. and MCCLINTOCK, C. G. (1967). Value: Behavioral decision theory. *Annu. Rev. Psychol.* **18** 239–286.

[4] BELL, C. E. (1978). What happens when majority rule breaks down? Some probability calculations. *Public Choice* **33** 121–126.

[5] BENDER, E. M., GEBRU, T., MCMILLAN-MAJOR, A. and SHMITCHELL, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the* 2021 *ACM Conference on Fairness*, *Accountability*, *and Transparency* 610–623.

[6] BLIN, J.-M. (1973). Intransitive social orderings and the probability of the Condorcet effect. *Kyklos* **26** 25–35.

[7] BORGA, J., CHATTERJEE, S. and DIACONIS, P. (2025). Permuton and local limits for the Luce model. Preprint. Available at arXiv:2509.07729.

MSC2020 subject classifications. Primary 62F07; secondary 62P25.

Key words and phrases. Large language models, preference alignment, reward model, Condorcet paradox, Nash equilibrium.

[8] BRADLEY, R. A. and TERRY, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39** 324–345. MR0070925 https://doi.org/10.2307/2334029

[9] BUBECK, S., CHANDRASEKARAN, V., ELDAN, R., GEHRKE, J., HORVITZ, E., KAMAR, E., LEE, P., LEE, Y. T., LI, Y. et al. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. Preprint. Available at arXiv:2303.12712.

[10] CALANDRIELLO, D., GUO, Z. D., MUNOS, R., ROWLAND, M., TANG, Y., AVILA PIRES, B., RICHEMOND, P. H., LE LAN, C., VALKO, M. et al. (2024). Human alignment of large language models through online preference optimisation. In *Forty-First International Conference on Machine Learning*.

[11] CASPER, S., DAVIES, X., SHI, C., GILBERT, T. K., SCHEURER, J., RANDO, J., FREEDMAN, R., KORBAK, T., LINDNER, D. et al. (2023). Open problems and fundamental limitations of reinforcement learning from human feedback. *Trans. Mach. Learn. Res.*.

[12] CHAKRABORTY, S., QIU, J., YUAN, H., KOPPEL, A., HUANG, F., MANOCHA, D., BEDI, A. S. and WANG, M. (2024). MaxMin-RLHF: Towards equitable alignment of large language models with diverse human preferences. Preprint. Available at arXiv:2402.08925.

[13] CHOWDHERY, A., NARANG, S., DEVLIN, J., BOSMA, M., MISHRA, G., ROBERTS, A., BARHAM, P., CHUNG, H. W., SUTTON, C. et al. (2023). Palm: Scaling language modeling with pathways. *J. Mach. Learn. Res.* **24** 1–113.

[14] DONG, H., XIONG, W., PANG, B., WANG, H., ZHAO, H., ZHOU, Y., JIANG, N., SAHOO, D., XIONG, C. et al. (2024). RLHF workflow: From reward modeling to online RLHF. *Trans. Mach. Learn. Res.*.

[15] DUERSCH, P., OECHSSLER, J. and SCHIPPER, B. C. (2012). Pure strategy equilibria in symmetric two-player zero-sum games. *Internat. J. Game Theory* **41** 553–564. MR2950851 https://doi.org/10.1007/s00182-011-0302-x

[16] ELOUNDOU, T., MANNING, S., MISHKIN, P. and ROCK, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science* **384** 1306–1308.

[17] ERDŐS, P. and RÉNYI, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. MR0120167 https://doi.org/10.5486/pmd.1959.6.3-4.12

[18] FAN, J., HOU, J. and YU, M. (2024). Uncertainty quantification of MLE for entity ranking with covariates. *J. Mach. Learn. Res.* **25** Paper No. [358], 83 pp. MR4847422

[19] FAN, J., LOU, Z., WANG, W. and YU, M. (2025). Ranking inferences based on the top choice of multiway comparisons. *J. Amer. Statist. Assoc.* **120** 237–250.

[20] FAN, J., LOU, Z., WANG, W. and YU, M. (2026). Spectral ranking inferences based on general multiway comparisons. *Oper. Res.* **74** 161–180. MR5016999 https://doi.org/10.1287/opre.2023.0439

[21] FENG, S., PARK, C. Y., LIU, Y. and TSVETKOV, Y. (2023). From pretraining data to language models to downstream tasks: Tracking the trails of political biases leading to unfair NLP models. In *Proceedings of the* 61*st Annual Meeting of the Association for Computational Linguistics* (*Volume* 1: *Long Papers*) 11737–11762.

[22] GARMAN, M. B. and KAMIEN, M. I. (1968). The paradox of voting: Probability calculations. *Behav. Sci.* **13** 306–316.

[23] GE, L., HALPERN, D., MICHA, E., PROCACCIA, A. D., SHAPIRA, I., VOROBEYCHIK, Y. and WU, J. (2024). Axioms for ai alignment from human feedback. *Adv. Neural Inf. Process. Syst.* **37** 80439–80465.

[24] GEHRLEIN, W. V. (2006). *Condorcet's Paradox*. Springer, Berlin.

[25] GEHRLEIN, W. V. and LEPELLEY, D. (2004). Probability calculations in voting theory: An overview of recent results. In *Reasoned Choices*: *Essays in Honor of Academy Professor Hannu Nurmi*.

[26] GÖLZ, P., HAGHTALAB, N. and YANG, K. (2025). Distortion of AI alignment: Does preference optimization optimize for preferences? In *The Thirty-Ninth Annual Conference on Neural Information Processing Systems*.

[27] GUILBAUD, G. T. (1952). Les théories de l'intérêt général et le problème logique de l'agrégation. *Econ. Appl.* **5** 501–584.

[28] HURST, A., LERER, A., GOUCHER, A. P., PERELMAN, A., RAMESH, A., CLARK, A., OSTROW, A., WELIHINDA, A., HAYES, A. et al. (2024). Gpt-4o system card. Preprint. Available at arXiv:2410.21276.

[29] ISMAIL, M. S. and PEETERS, R. (2025). Existence of pure equilibria in symmetric two-player zero-sum games. *Internat. J. Game Theory* **54** Paper No. 13, 18 pp. MR4903720 https://doi.org/10.1007/s00182-025-00938-2

[30] JI, W., YUAN, W., GETZEN, E., CHO, K., JORDAN, M. I., MEI, S., WESTON, J. E., SU, W. J., XU, J. et al. (2025). An overview of large language models for statisticians. Preprint. Available at arXiv:2502.17814.

[31] KELLY, J. S. (1974). Voting anomalies, the number of voters, and the number of alternatives. *Econometrica* **42** 239–251. MR0441220 https://doi.org/10.2307/1911974

[32] KOTEK, H., DOCKUM, R. and SUN, D. (2023). Gender bias and stereotypes in large language models. In *Proceedings of the ACM Collective Intelligence Conference* 12–24.

[33] KURRILD-KLITGAARD, P. (2001). An empirical example of the Condorcet paradox of voting in a large electorate. *Public Choice* **107** 135–145.

[34] LI, Z., CHEN, C., XU, T., QIN, Z., XIAO, J., SUN, R. and LUO, Z.-Q. (2024a). Entropic distribution matching for supervised fine-tuning of LLMs: Less overfitting and better diversity. In *NeurIPS* 2024 *Workshop on Fine-Tuning in Modern Machine Learning*: *Principles and Scalability*.

[35] LI, Z., XU, T., ZHANG, Y., LIN, Z., YU, Y., SUN, R. and LUO, Z.-Q. (2024b). ReMax: A simple, effective, and efficient reinforcement learning method for aligning large language models. In *Proceedings of the* 41*st International Conference on Machine Learning* 29128–29163.

[36] LIU, K., LONG, Q., SHI, Z., SU, W. J and XIAO, J. (2026). Supplement to "Statistical impossibility and possibility of aligning LLMs with human preferences: From Condorcet paradox to Nash equilibrium." https://doi.org/10.1214/26-AOS2643SUPP

[37] LUCE, R. D. (2012). *Individual Choice Behavior*: *A Theoretical Analysis*. Courier Corporation. MR0108411

[38] LUO, Q., PUETT, M. J. and SMITH, M. D. (2024). A perspectival mirror of the elephant. *Commun*. *ACM* **67** 98–105.

[39] MAURA-RIVERO, R.-R., LANCTOT, M., VISIN, F. and LARSON, K. (2025). Jackpot! Alignment as a maximal lottery. Preprint. Available at arXiv:2501.19266.

[40] MAY, R. M. (1971). Some mathematical remarks on the paradox of voting. *Behav. Sci*. **16** 143–151.

[41] MUNOS, R., VALKO, M., CALANDRIELLO, D., GHESHLAGHI AZAR, M., ROWLAND, M., GUO, Z. D., TANG, Y., GEIST, M., MESNARD, T. et al. (2024). Nash learning from human feedback. In *Forty-First International Conference on Machine Learning*.

[42] OUYANG, L., WU, J., JIANG, X., ALMEIDA, D., WAINWRIGHT, C. L., MISHKIN, P., ZHANG, C., AGARWAL, S., SLAMA, K. et al. (2022). Training language models to follow instructions with human feedback. *Adv. Neural Inf. Process*. *Syst*. **35** 27730–27744.

[43] RAFAILOV, R., SHARMA, A., MITCHELL, E., MANNING, C. D., ERMON, S. and FINN, C. (2023). Direct preference optimization: Your language model is secretly a reward model. *Adv. Neural Inf. Process*. *Syst*. **36** 53728–53741.

[44] SANTURKAR, S., DURMUS, E., LADHAK, F., LEE, C., LIANG, P. and HASHIMOTO, T. (2023). Whose opinions do language models reflect? In *International Conference on Machine Learning* 29971–30004. PMLR.

[45] SAUERMANN, L. (2022). On the probability of a Condorcet winner among a large number of alternatives. Preprint. Available at arXiv:2203.13713.

[46] SCHULMAN, J., WOLSKI, F., DHARIWAL, P., RADFORD, A. and KLIMOV, O. (2017). Proximal policy optimization algorithms. Preprint. Available at arXiv:1707.06347.

[47] SHI, Z., LIU, K., LONG, Q., SU, W. J. and XIAO, J. (2025). Fundamental limits of game-theoretic llm alignment: Smith consistency and preference matching. Preprint. Available at arXiv:2505.20627.

[48] SITHTHARANJAN, A., LAIDLAW, C. and HADFIELD-MENELL, D. (2023). Distributional preference learning: Understanding and accounting for hidden context in RLHF. Preprint. Available at arXiv:2312.08358.

[49] SONG, Z., CAI, T., LEE, J. D. and SU, W. J. (2023). Reward collapse in aligning large language models: A prompt-aware approach to preference rankings. In *ICML* 2023 *Workshop the Many Facets of Preference-Based Learning*.

[50] TADELIS, S. (2013). *Game Theory*: *An Introduction*. Princeton Univ. Press, Princeton, NJ. MR3235473

[51] TANG, Y., GUO, D. Z., ZHENG, Z., CALANDRIELLO, D., CAO, Y., TARASSOV, E., MUNOS, R., PIRES, B. Á., VALKO, M. et al. (2024). Understanding the performance gap between online and offline alignment algorithms. Preprint. Available at arXiv:2405.08448.

[52] TOUVRON, H., LAVRIL, T., IZACARD, G., MARTINET, X., LACHAUX, M.-A., LACROIX, T., ROZIÈRE, B., GOYAL, N., HAMBRO, E. et al. (2023). Llama: Open and efficient foundation language models. Preprint. Available at arXiv:2302.13971.

[53] WILSON, R. (1972). Social choice theory without the Pareto principle. *J. Econom*. *Theory* **5** 478–486. MR0449494 https://doi.org/10.1016/0022-0531(72)90051-8

[54] XIAO, J., LI, Z., XIE, X., GETZEN, E., FANG, C., LONG, Q. and SU, W. J. (2025). On the algorithmic bias of aligning large language models with RLHF: Preference collapse and matching regularization. *J. Amer. Statist*. *Assoc*. **120** 2154–2164. MR5006831 https://doi.org/10.1080/01621459.2025.2555067

[55] XIAO, J., SHI, Z., LIU, K., LONG, Q. and SU, W. J. (2025). Theoretical tensions in rlhf: Reconciling empirical success with inconsistencies in social choice theory. Preprint. Available at arXiv:2506.12350.

[56] XIONG, W., DONG, H., YE, C., WANG, Z., ZHONG, H., JI, H., JIANG, N. and ZHANG, T. (2023). Iterative preference learning from human feedback: Bridging theory and practice for rlhf under kl-constraint. In *ICLR* 2024 *Workshop on Mathematical and Empirical Understanding of Foundation Models*.

[57] YE, C., XIONG, W., ZHANG, Y., DONG, H., JIANG, N. and ZHANG, T. (2024). Online iterative reinforcement learning from human feedback with general preference model. *Adv. Neural Inf. Process. Syst.* **37** 81773–81807.

[58] ZHANG, Y., ZHANG, G., WU, Y., XU, K. and GU, Q. (2024). General preference modeling with preference representations for aligning language models. Preprint. Available at arXiv:2410.02197.

The Annals of Statistics
2026, Vol. 54, No. 4, 2136–2163
https://doi.org/10.1214/26-AOS2644

HIGHER-ORDER GRAPHON THEORY: FLUCTUATIONS, DEGENERACIES AND INFERENCE

BY ANIRBAN CHATTERJEE[1,a], SOHAM DAN[2,b] AND BHASWAR B. BHATTACHARYA[3,c]

[1]*Department of Statistics, University of Chicago,* [a]*anirbanc@uchicago.edu*
[2]*Scale AI, San Francisco,* [b]*sohamdan@seas.upenn.edu*
[3]*Department of Statistics and Data Science, University of Pennsylvania,* [c]*bhaswar@wharton.upenn.edu*

Exchangeable random graphs, which include some of the most widely studied network models, have emerged as the mainstay of statistical network analysis in recent years. Graphons, which are the central objects in graph limit theory, provide a natural way to sample exchangeable random graphs. It is well known that network moments (motif/subgraph counts) identify a graphon (up to an isomorphism); hence, understanding the sampling distribution of subgraph counts in random graphs sampled from a graphon is pivotal for nonparametric network inference. Although there are a few results regarding the asymptotic normality of subgraph counts in graphon models, for many commonly appearing graphons this distribution is degenerate. This degeneracy phenomenon was overlooked until very recently and its consequences in network inference have remained unexplored. Toward this, we obtain the following results: We derive the joint asymptotic distribution of any finite collection of network moments in random graphs sampled from a graphon, which includes both the nondegenerate case (where the distribution is Gaussian) as well as the degenerate case (where the distribution has both Gaussian or non-Gaussian components). This provides the higher-order fluctuation theory for subgraph counts in the graphon model. Furthermore, we develop a novel multiplier bootstrap for graphons that consistently approximates the limiting distribution of the network moments (both in the Gaussian and non-Gaussian regimes). Using this and a procedure for testing degeneracy, we construct joint confidence sets for any finite collection of motif densities. This provides a general framework for statistical inference based on network moments in the graphon model. Examples and simulations are provided to validate the general theory. To illustrate the broad scope of our results, we also consider the problem of detecting global structure (i.e., testing whether the graphon is a constant function) based on small subgraphs. We propose a consistent test for this problem, invoking celebrated results on quasirandom graphs, and derive its limiting distribution both under the null and the alternative.

REFERENCES

[1] ALDOUS, D. J. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** 581–598. MR0637937 https://doi.org/10.1016/0047-259X(81)90099-3
[2] ALON, U. (2007). Network motifs: Theory and experimental approaches. *Nat. Rev. Genet.* **8** 450–461.
[3] ATHREYA, A., FISHKIND, D. E., TANG, M., PRIEBE, C. E., PARK, Y., VOGELSTEIN, J. T., LEVIN, K., LYZINSKI, V., QIN, Y. et al. (2017). Statistical inference on random dot product graphs: A survey. *J. Mach. Learn. Res.* **18** Paper No. 226, 92 pp. MR3827114
[4] AUSTERN, M. and ORBANZ, P. (2022). Limit theorems for distributions invariant under groups of transformations. *Ann. Statist.* **50** 1960–1991. MR4474479 https://doi.org/10.1214/21-aos2165
[5] BARBOUR, A. D., KAROŃSKI, M. and RUCIŃSKI, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** 125–145. MR1047781 https://doi.org/10.1016/0095-8956(89)90014-2

MSC2020 subject classifications. 05C80, 62E20, 60F05.
Key words and phrases. Inhomogeneous random graphs, network analysis, generalized U-statistics, motif estimation, subgraph counts.

[6] BHATTACHARYA, A., BISHNU, A., GHOSH, A. and MISHRA, G. (2021). On triangle estimation using tripartite independent set queries. *Theory Comput. Syst.* **65** 1165–1192. MR4324413 https://doi.org/10.1007/s00224-021-10043-y

[7] BHATTACHARYA, B. B., CHATTERJEE, A. and JANSON, S. (2023). Fluctuations of subgraph counts in graphon based random graphs. *Combin. Probab. Comput.* **32** 428–464. MR4574856 https://doi.org/10.1017/s0963548322000335

[8] BHATTACHARYA, B. B., DAS, S. and MUKHERJEE, S. (2022). Motif estimation via subgraph sampling: The fourth-moment phenomenon. *Ann. Statist.* **50** 987–1011. MR4404926 https://doi.org/10.1214/21-aos2134

[9] BHATTACHARYYA, S. and BICKEL, P. J. (2015). Subsampling bootstrap of count features of networks. *Ann. Statist.* **43** 2384–2411. MR3405598 https://doi.org/10.1214/15-AOS1338

[10] BICKEL, P. J. and CHEN, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci. USA* **106** 21068–21073.

[11] BICKEL, P. J., CHEN, A. and LEVINA, E. (2011). The method of moments and degree distributions for network models. *Ann. Statist.* **39** 2280–2301. MR2906868 https://doi.org/10.1214/11-AOS904

[12] BOGUNÁ, M. and PASTOR-SATORRAS, R. (2003). Class of correlated random networks with hidden variables. *Phys. Rev. E* **68** 036112.

[13] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. MR2337396 https://doi.org/10.1002/rsa.20168

[14] BORGS, C., CHAYES, J. and LOVÁSZ, L. (2010). Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** 1597–1619. MR2594615 https://doi.org/10.1007/s00039-010-0044-0

[15] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626 https://doi.org/10.1016/j.aim.2008.07.008

[16] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382 https://doi.org/10.4007/annals.2012.176.1.2

[17] BRAVO-HERMSDORFF, G., GUNDERSON, L. M., MAUGIS, P.-A. and PRIEBE, C. E. (2023). Quantifying network similarity using graph cumulants. *J. Mach. Learn. Res.* **24** Paper No. [187], 27 pp. MR4633576

[18] CHATTERJEE, A., DAN, S. and BHATTACHARYA, B. B. (2026). Supplement to "higher-order graphon theory: Fluctuations, degeneracies and inference." https://doi.org/10.1214/26-AOS2644SUPP

[19] CHATTERJEE, S., CHATTERJEE, A., CHAKRABORTY, A. and BHATTACHARYA, B. B. (2025). Asymptotic normality of subgraph counts in sparse inhomogeneous random graphs. Preprint. Available at arXiv:2512.12937.

[20] CHATTERJEE, S. and DIACONIS, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41** 2428–2461. MR3127871 https://doi.org/10.1214/13-AOS1155

[21] CHUNG, F. R. K., GRAHAM, R. L. and WILSON, R. M. (1989). Quasi-random graphs. *Combinatorica* **9** 345–362. MR1054011 https://doi.org/10.1007/BF02125347

[22] CRANE, H. (2018). *Probabilistic Foundations of Statistical Network Analysis*. *Monographs on Statistics and Applied Probability* **157**. CRC Press, Boca Raton, FL. MR3791467

[23] DELMAS, J.-F., DHERSIN, J.-S. and SCIAUVEAU, M. (2021). Asymptotic for the cumulative distribution function of the degrees and homomorphism densities for random graphs sampled from a graphon. *Random Structures Algorithms* **58** 94–149. MR4180255 https://doi.org/10.1002/rsa.20965

[24] DIACONIS, P. and FREEDMAN, D. (1981). On the statistics of vision: The Julesz conjecture. *J. Math. Psych.* **24** 112–138. MR0640207 https://doi.org/10.1016/0022-2496(81)90039-0

[25] DIACONIS, P. and JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* (7) **28** 33–61. MR2463439

[26] EDEN, T., LEVI, A., RON, D. and SESHADHRI, C. (2017). Approximately counting triangles in sublinear time. *SIAM J. Comput.* **46** 1603–1646. MR3709896 https://doi.org/10.1137/15M1054389

[27] FANG, X. and RÖLLIN, A. (2015). Rates of convergence for multivariate normal approximation with applications to dense graphs and doubly indexed permutation statistics. *Bernoulli* **21** 2157–2189. MR3378463 https://doi.org/10.3150/14-BEJ639

[28] FÉRAY, V., MÉLIOT, P.-L. and NIKEGHBALI, A. (2020). Graphons, permutons and the Thoma simplex: Three mod-Gaussian moduli spaces. *Proc. Lond. Math. Soc.* (3) **121** 876–926. MR4105789 https://doi.org/10.1112/plms.12344

[29] GAO, C. and LAFFERTY, J. (2017). Testing for global network structure using small subgraph statistics. Preprint. Available at arXiv:1710.00862.

[30] GAO, C. and LAFFERTY, J. (2017). Testing network structure using relations between small subgraph probabilities. Preprint. Available at arXiv:1704.06742.

[31] GAO, C., LU, Y. and ZHOU, H. H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43** 2624–2652. MR3405606 https://doi.org/10.1214/15-AOS1354

[32] GHOSHDASTIDAR, D., GUTZEIT, M., CARPENTIER, A. and VON LUXBURG, U. (2017). Two-sample tests for large random graphs using network statistics. In *Conference on Learning Theory* 954–977. PMLR.

[33] GREEN, A. and SHALIZI, C. R. (2022). Bootstrapping exchangeable random graphs. *Electron. J. Stat.* **16** 1058–1095. MR4377133 https://doi.org/10.1214/21-ejs1896

[34] HLADKÝ, J., PELEKIS, C. and ŠILEIKIS, M. (2021). A limit theorem for small cliques in inhomogeneous random graphs. *J. Graph Theory* **97** 578–599. MR4313198 https://doi.org/10.1002/jgt.22673

[35] 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 https://doi.org/10.1198/016214502388618906

[36] HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. MR0718088 https://doi.org/10.1016/0378-8733(83)90021-7

[37] HOLLAND, P. W. and LEINHARDT, S. (1971). Transitivity in structural models of small groups. *Comp. Group Stud.* **2** 107–124.

[38] HOOVER, D. N. (1982). Row-column exchangeability and a generalized model for probability. In *Exchangeability in Probability and Statistics* (*Rome*, 1981) 281–291. North-Holland, Amsterdam. MR0675982

[39] HUANG, K. H., AUSTERN, M. and ORBANZ, P. (2024). Gaussian universality for approximately polynomial functions of high-dimensional data. Preprint. Available at arXiv:2403.10711.

[40] HUNTER, D. R., HANDCOCK, M. S., BUTTS, C. T., GOODREAU, S. M. and MORRIS, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *J. Stat. Softw.* **24** nihpa54860.

[41] JANSON, S. (1990). A functional limit theorem for random graphs with applications to subgraph count statistics. *Random Structures Algorithms* **1** 15–37. MR1068489 https://doi.org/10.1002/rsa.3240010103

[42] JANSON, S. (1994). Orthogonal decompositions and functional limit theorems for random graph statistics. *Mem. Amer. Math. Soc.* **111** vi+78. MR1219708 https://doi.org/10.1090/memo/0534

[43] JANSON, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge Univ. Press, Cambridge. MR1474726 https://doi.org/10.1017/CBO9780511526169

[44] JANSON, S., ŁUCZAK, T. and RUCINSKI, A. (2011). *Random Graphs*. *Wiley-Interscience Series in Discrete Mathematics and Optimization*. Wiley Interscience, New York. MR1782847 https://doi.org/10.1002/9781118032718

[45] JANSON, S. and NOWICKI, K. (1991). The asymptotic distributions of generalized U-statistics with applications to random graphs. *Probab. Theory Related Fields* **90** 341–375. MR1133371 https://doi.org/10.1007/BF01193750

[46] KAUR, G. and RÖLLIN, A. (2021). Higher-order fluctuations in dense random graph models. *Electron. J. Probab.* **26** Paper No. 139, 36 pp. MR4346670 https://doi.org/10.1214/21-ejp708

[47] KLUSOWSKI, J. M. and WU, Y. (2018). Counting motifs with graph sampling. In *Conference on Learning Theory* 1966–2011.

[48] LE MINH, T. (2023). U-statistics on bipartite exchangeable networks. *ESAIM Probab. Stat.* **27** 576–620. MR4602388 https://doi.org/10.1051/ps/2023010

[49] LEI, J. (2016). A goodness-of-fit test for stochastic block models. *Ann. Statist.* **44** 401–424. MR3449773 https://doi.org/10.1214/15-AOS1370

[50] LEI, J. (2021). Network representation using graph root distributions. *Ann. Statist.* **49** 745–768. MR4255106 https://doi.org/10.1214/20-aos1976

[51] LEVIN, K. and LEVINA, E. (2025). Bootstrapping networks with latent space structure. *Electron. J. Stat.* **19** 745–791. MR4870301 https://doi.org/10.1214/25-ejs2347

[52] LIN, Q., LUNDE, R. and SARKAR, P. (2022). Trading off accuracy for speedup: Multiplier bootstraps for subgraph counts. Preprint. Available at arXiv:2009.06170v5.

[53] LIU, Q. and PRIVAULT, N. (2026). Gaussian fluctuations of generalized U-statistics and subgraph counting in the binomial random-connection model. *Stoch. Process. Appl.* **192** Paper No. 104825, 21 pp. MR4986652 https://doi.org/10.1016/j.spa.2025.104825

[54] LOVÁSZ, L. (2012). *Large Networks and Graph Limits*. *American Mathematical Society Colloquium Publications* **60**. Amer. Math. Soc., Providence, RI. MR3012035 https://doi.org/10.1090/coll/060

[55] LOVÁSZ, L. and SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085 https://doi.org/10.1016/j.jctb.2006.05.002

[56] LUCE, R. D. and PERRY, A. D. (1949). A method of matrix analysis of group structure. *Psychometrika* **14** 95–116. MR0035974 https://doi.org/10.1007/BF02289146

[57] LUNDE, R. and SARKAR, P. (2023). Subsampling sparse graphons under minimal assumptions. *Biometrika* **110** 15–32. MR4565441 https://doi.org/10.1093/biomet/asac032

[58] LYU, H., MÉMOLI, F. and SIVAKOFF, D. (2023). Sampling random graph homomorphisms and applications to network data analysis. *J. Mach. Learn. Res.* **24** Paper No. [9], 79 pp. MR4582431
[59] MAUGIS, P., PRIEBE, C. E., OLHEDE, S. C. and WOLFE, P. J. (2017). Statistical inference for network samples using subgraph counts. Preprint. Available at arXiv:1701.00505.
[60] MAUGIS, P. A. (2024). Central limit theorems for local network statistics. *Biometrika* **111** 743–754. MR4794147 https://doi.org/10.1093/biomet/asad080
[61] MILO, R., SHEN-ORR, S., ITZKOVITZ, S., KASHTAN, N., CHKLOVSKII, D. and ALON, U. (2002). Network motifs: Simple building blocks of complex networks. *Science* **298** 824–827.
[62] MOSSEL, E., NEEMAN, J. and SLY, A. (2016). Consistency thresholds for the planted bisection model. *Electron. J. Probab.* **21** Paper No. 21, 24 pp. MR3485363 https://doi.org/10.1214/16-EJP4185
[63] MUKHERJEE, S. (2020). Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics. *Bernoulli* **26** 1016–1043. MR4058359 https://doi.org/10.3150/19-BEJ1135
[64] MUKHERJEE, S. and XU, Y. (2023). Statistics of the two star ERGM. *Bernoulli* **29** 24–51. MR4497238 https://doi.org/10.3150/21-bej1448
[65] NOWICKI, K. (1989). Asymptotic normality of graph statistics. *J. Statist. Plann. Inference* **21** 209–222. MR0985458 https://doi.org/10.1016/0378-3758(89)90005-0
[66] NOWICKI, K. and WIERMAN, J. C. (1988). Subgraph counts in random graphs using incomplete U-statistics methods. *Discrete Math.* **72** 299–310. MR0975550 https://doi.org/10.1016/0012-365X(88)90220-8
[67] OUADAH, S., LATOUCHE, P. and ROBIN, S. (2022). Motif-based tests for bipartite networks. *Electron. J. Stat.* **16** 293–330. MR4360230 https://doi.org/10.1214/21-ejs1944
[68] PENROSE, M. (2003). *Random Geometric Graphs. Oxford Studies in Probability* **5**. Oxford Univ. Press, Oxford. MR1986198 https://doi.org/10.1093/acprof:oso/9780198506263.001.0001
[69] PRŽULJ, N. (2007). Biological network comparison using graphlet degree distribution. *Bioinformatics* **23** e177–e183.
[70] RUBIN-DELANCHY, P., CAPE, J., TANG, M. and PRIEBE, C. E. (2022). A statistical interpretation of spectral embedding: The generalised random dot product graph. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1446–1473. MR4494166 https://doi.org/10.1111/rssb.12509
[71] RUCIŃSKI, A. (1988). When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields* **78** 1–10. MR0940863 https://doi.org/10.1007/BF00718031
[72] SHAO, M., XIA, D., ZHANG, Y., WU, Q. and CHEN, S. (2026). Higher-order accurate two-sample network inference and network hashing. *J. Amer. Statist. Assoc.* **121** 373–388. MR5061128 https://doi.org/10.1080/01621459.2025.2520459
[73] SHEN-ORR, S. S., MILO, R., MANGAN, S. and ALON, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet.* **31** 64–68.
[74] TANG, M., ATHREYA, A., SUSSMAN, D. L., LYZINSKI, V. and PRIEBE, C. E. (2017). A nonparametric two-sample hypothesis testing problem for random graphs. *Bernoulli* **23** 1599–1630. MR3624872 https://doi.org/10.3150/15-BEJ789
[75] WASSERMAN, S. and FAUST, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge Univ. Press, Cambridge.
[76] WATTS, D. J. and STROGATZ, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* **393** 440–442.
[77] WEGNER, A. E., OSPINA-FORERO, L., GAUNT, R. E., DEANE, C. M. and REINERT, G. (2018). Identifying networks with common organizational principles. *J. Complex Netw.* **6** 887–913. MR3892299 https://doi.org/10.1093/comnet/cny003
[78] ZHANG, Y. and XIA, D. (2022). Edgeworth expansions for network moments. *Ann. Statist.* **50** 726–753. MR4404918 https://doi.org/10.1214/21-aos2125
[79] ZHANG, Z.-S. (2022). Berry-Esseen bounds for generalized U-statistics. *Electron. J. Probab.* **27** Paper No. 134, 36 pp. MR4492985 https://doi.org/10.1214/22-ejp860

The Annals of Statistics
2026, Vol. 54, No. 4, 2164–2190
https://doi.org/10.1214/26-AOS2647

GAUSSIAN AND BOOTSTRAP APPROXIMATION FOR MATCHING-BASED AVERAGE TREATMENT EFFECT ESTIMATORS

BY ZHAOYANG SHI[1,a], CHINMOY BHATTACHARJEE[2,b], KRISHNAKUMAR BALASUBRAMANIAN[3,c] AND WOLFGANG POLONIK[3,d]

[1]*Department of Statistics, Harvard University,* [a]*zyshi10m@gmail.com*
[2]*Department of Mathematics, Universität Hamburg,* [b]*chinmoy.bhattacharjee@uni-hamburg.de*
[3]*Department of Statistics, University of California, Davis,* [c]*kbala@ucdavis.edu,* [d]*wpolonik@ucdavis.edu*

We establish Gaussian approximation bounds for covariate and rank-matching-based Average Treatment Effect (ATE) estimators. By analyzing these estimators through the lens of stabilization theory, we employ the Malliavin–Stein method to derive our results. Our bounds precisely quantify the impact of key problem parameters, including the number of matches and treatment balance, on the accuracy of the Gaussian approximation. Additionally, we develop multiplier bootstrap procedures to estimate the limiting distribution in a fully data-driven manner, and we leverage the derived Gaussian approximation results to further obtain bootstrap approximation bounds. Our work not only introduces a novel theoretical framework for commonly used ATE estimators, but also provides data-driven methods for constructing nonasymptotically valid confidence intervals.

REFERENCES

[1] ABADIE, A. and IMBENS, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica* **74** 235–267. MR2194325 https://doi.org/10.1111/j.1468-0262.2006.00655.x
[2] ABADIE, A. and IMBENS, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica* **76** 1537–1557. MR2468559 https://doi.org/10.3982/ECTA6474
[3] ABADIE, A. and IMBENS, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *J. Bus. Econom. Statist.* **29** 1–11. MR2789386 https://doi.org/10.1198/jbes.2009.07333
[4] ABADIE, A. and SPIESS, J. (2022). Robust post-matching inference. *J. Amer. Statist. Assoc.* **117** 983–995. MR4436327 https://doi.org/10.1080/01621459.2020.1840383
[5] BANG, H. and ROBINS, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics* **61** 962–973. MR2216189 https://doi.org/10.1111/j.1541-0420.2005.00377.x
[6] BHATTACHARJEE, C. and MOLCHANOV, I. (2022). Gaussian approximation for sums of region-stabilizing scores. *Electron. J. Probab.* **27** Paper No. 111, 27. MR4474533 https://doi.org/10.1214/22-ejp832
[7] CATTANEO, M. D., HAN, F. and LIN, Z. (2025). On Rosenbaum's rank-based matching estimator. *Biometrika* **112** Paper No. asae062, 12. MR4869184 https://doi.org/10.1093/biomet/asae062
[8] GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548 https://doi.org/10.1198/016214506000001437
[9] IMBENS, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econ. Stat.* **86** 4–29.
[10] IMBENS, G. W. and WOOLDRIDGE, J. M. (2009). Recent developments in the econometrics of program evaluation. *J. Econ. Lit.* **47** 5–86.
[11] KOSKO, M., WANG, L. and SANTACATTERINA, M. (2024). A fast bootstrap algorithm for causal inference with large data. *Stat. Med.* **43** 2894–2927. MR4761089 https://doi.org/10.1002/sim.10075
[12] LACHIÈZE-REY, R., PECCATI, G. and YANG, X. (2022). Quantitative two-scale stabilization on the Poisson space. *Ann. Appl. Probab.* **32** 3085–3145. MR4474528 https://doi.org/10.1214/21-aap1768
[13] LACHIÈZE-REY, R., SCHULTE, M. and YUKICH, J. E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 https://doi.org/10.1214/18-AAP1405

MSC2020 subject classifications. Primary 62D20; secondary 62G09, 62E17, 62F40.
Key words and phrases. Nonasymptotic confidence interval, average treatment effect, residual-based multiplier bootstrap, Gaussian approximation.

[14] LIN, Z., DING, P. and HAN, F. (2023). Estimation based on nearest neighbor matching: From density ratio to average treatment effect. *Econometrica* **91** 2187–2217. MR4680059 https://doi.org/10.3982/ecta20598

[15] LIN, Z. and HAN, F. (2026). On the consistency of bootstrap for matching estimators. *Biometrika* **113** Paper No. asag005, 8. MR5051418 https://doi.org/10.1093/biomet/asag005

[16] MORGAN, S. L. and HARDING, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociol. Methods Res.* **35** 3–60. MR2247150 https://doi.org/10.1177/0049124106289164

[17] OTSU, T. and RAI, Y. (2017). Bootstrap inference of matching estimators for average treatment effects. *J. Amer. Statist. Assoc.* **112** 1720–1732. MR3750894 https://doi.org/10.1080/01621459.2016.1231613

[18] PENROSE, M. D. and YUKICH, J. E. (2005). Normal approximation in geometric probability. In *Stein's Method and Applications*. *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 37–58. Singapore Univ. Press, Singapore. MR2201885 https://doi.org/10.1142/9789812567673_0003

[19] PENROSE, M. D. and YUKICH, J. E. (2013). Limit theory for point processes in manifolds. *Ann. Appl. Probab.* **23** 2161–2211. MR3127932 https://doi.org/10.1214/12-AAP897

[20] ROSENBAUM, P. R. (1995). *Obs. Stud. Springer Series in Statistics*. Springer, New York. MR1353914 https://doi.org/10.1007/978-1-4757-2443-1

[21] ROSENBAUM, P. R. (2005). An exact distribution-free test comparing two multivariate distributions based on adjacency. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 515–530. MR2168202 https://doi.org/10.1111/j.1467-9868.2005.00513.x

[22] ROSENBAUM, P. R. (2010). *Design of Observational Studies*. *Springer Series in Statistics*. Springer, New York. MR2561612 https://doi.org/10.1007/978-1-4419-1213-8

[23] ROY, A. and BALASUBRAMANIAN, K. (2023). Online covariance estimation for stochastic gradient descent under Markovian sampling. arXiv preprint. Available at arXiv:2308.01481.

[24] RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66** 688.

[25] SCHARFSTEIN, D. O., ROTNITZKY, A. and ROBINS, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *J. Amer. Statist. Assoc.* **94** 1096–1120. With comments and a rejoinder by the authors. MR1731478 https://doi.org/10.2307/2669923

[26] SHI, L. and DING, P. (2026). Berry–Esseen bounds for design-based causal inference with possibly diverging treatment levels and varying group sizes. *Ann. Statist.* **54** 324–349. MR5039621 https://doi.org/10.1214/25-AOS2569

[27] SHI, Z., BALASUBRAMANIAN, K. and POLONIK, W. (2024). A flexible approach for normal approximation of geometric and topological statistics. *Bernoulli* **30** 3029–3058. MR4779858 https://doi.org/10.3150/23-bej1705

[28] SHI, Z., BHATTACHARJEE, C., BALASUBRAMANIAN, K. and POLONIK, W. (2024). Multivariate Gaussian approximation for random forest via region-based stabilization. arXiv preprint. Available at arXiv:2403.09960.

[29] SHI, Z., BHATTACHARJEE, C., BALASUBRAMANIAN, K. and POLONIK, W. (2026). Supplement to "Gaussian and bootstrap approximation for matching-based average treatment effect estimators." https://doi.org/10.1214/26-AOS2647SUPP

[30] STUART, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statist. Sci.* **25** 1–21. MR2741812 https://doi.org/10.1214/09-STS313

[31] TRAUTHWEIN, T. (2025). Quantitative CLTs on the Poisson space via Skorohod estimates and p-Poincaré inequalities. *Ann. Appl. Probab.* **35** 1716–1754. MR4921341 https://doi.org/10.1214/25-AAP2153

[32] WALSH, C. and JENTSCH, C. (2025). Nearest neighbor matching: M-out-of-N bootstrapping without bias correction vs. the naive bootstrap. *Econom. Stat.* **36** 81–89. MR4967022 https://doi.org/10.1016/j.ecosta.2023.04.005

The Annals of Statistics
2026, Vol. 54, No. 4, 2191–2208
https://doi.org/10.1214/26-AOS2649

SAMPLE SIZE AND POWER CALCULATIONS FOR CAUSAL INFERENCE IN OBSERVATIONAL STUDIES

BY BO LIU[1,a], CHENGXIN YANG[2,c] AND FAN LI[1,b]

[1]*Department of Statistical Science, Duke University,* [a]*boliu1997@alumni.duke.edu,* [b]*fl35@duke.edu*
[2]*Department of Biostatistics and Bioinformatics, Duke University,* [c]*chengxin.yang@duke.edu*

This paper investigates the theoretical foundation and develops analytical formulas for sample size and power calculations for causal inference with observational data. By analyzing the variance of an inverse probability weighting estimator of the average treatment effect, we decompose the power calculation into three components: propensity score distribution, potential outcome distribution and their correlation. We show that to determine the minimal sample size of an observational study, in addition to the standard inputs in the power calculation of randomized trials, it is sufficient to have two parameters: an overlap coefficient and a confounding coefficient, which quantify the strength of the confounder-treatment and the confounder-outcome association, respectively. For the overlap coefficient, we propose using the Bhattacharyya coefficient, which measures the covariate overlap and, together with the treatment proportion, leads to a uniquely identifiable and easily computable propensity score distribution. For the confounding coefficient, we propose a sensitivity parameter bounded by the R-squared statistic of the regression of the outcome on covariates. Our procedure relies on a parametric propensity score model and a semiparametric restricted mean outcome model, but does not require distributional assumptions on the multivariate covariates. We develop an associated R package `PSpower` and an online calculator.

REFERENCES

AITCHISON, J. and SHEN, S. M. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika* **67** 261–272. MR0581723 https://doi.org/10.2307/2335470

BENTKUS, V. (2005). A Lyapunov type bound in R^d. *Teor. Veroyatn. Primen.* **49** 311–323. MR2144310 https://doi.org/10.1137/S0040585X97981123

BHATTACHARYYA, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* **35** 99–110. MR0010358

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. A Wiley-Interscience Publication. MR1324786

BRANSON, Z. and DASGUPTA, T. (2020). Sampling-based randomised designs for causal inference under the potential outcomes framework. *Int. Stat. Rev.* **88** 101–121. MR4088012 https://doi.org/10.1111/insr.12339

BRANSON, Z., LI, X. and DING, P. (2024). Power and sample size calculations for rerandomization. *Biometrika* **111** 355–363. MR4704575 https://doi.org/10.1093/biomet/asad027

CHOW, S.-C., SHAO, J., WANG, H. and LOKHNYGINA, Y. (2018). *Sample Size Calculations in Clinical Research*. *Chapman & Hall/CRC Biostatistics Series*. CRC Press, Boca Raton, FL. Third edition of [MR2356591]. MR3753624

CONNORS, A. F., SPEROFF, T., DAWSON, N. V., THOMAS, C., HARRELL, F. E., WAGNER, D., DESBIENS, N., GOLDMAN, L., WU, A. W. et al. (1996). The effectiveness of right heart catheterization in the initial care of critically ill patients. *J. Amer. Med. Assoc.* **276** 889–897.

DING, P., FELLER, A. and MIRATRIX, L. (2016). Randomization inference for treatment effect variation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 655–671. MR3506797 https://doi.org/10.1111/rssb.12124

DING, P., FELLER, A. and MIRATRIX, L. (2019). Decomposing treatment effect variation. *J. Amer. Statist. Assoc.* **114** 304–317. MR3941256 https://doi.org/10.1080/01621459.2017.1407322

MSC2020 subject classifications. 62D20.
Key words and phrases. Causal inference, observational study, overlap, power, sample size, weighting.

HERNÁN, M. A. (2022). Causal analyses of existing databases: No power calculations required. *J. Clin. Epidemiol.* **144** 203–205.

HSIEH, F. and LAVORI, P. W. (2000). Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. *Control. Clin. Trials* **21** 552–560.

INMAN, H. F. and BRADLEY, E. L. JR. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Comm. Statist. Theory Methods* **18** 3851–3874. MR1040680 https://doi.org/10.1080/03610928908830127

KELCEY, B., DONG, N., SPYBROOK, J. and COX, K. (2017). Statistical power for causally defined indirect effects in group-randomized trials with individual-level mediators. *J. Educ. Behav. Stat.* **42** 499–530.

LI, F., MORGAN, K. L. and ZASLAVSKY, A. M. (2018). Balancing covariates via propensity score weighting. *J. Amer. Statist. Assoc.* **113** 390–400. MR3803473 https://doi.org/10.1080/01621459.2016.1260466

LI, F., THOMAS, L. E. and LI, F. (2019). Addressing extreme propensity scores via the overlap weights. *Amer. J. Epidemiol.* **188** 250–257.

LI, X., DING, P. and RUBIN, D. B. (2018). Asymptotic theory of rerandomization in treatment-control experiments. *Proc. Natl. Acad. Sci. USA* **115** 9157–9162. MR3856113 https://doi.org/10.1073/pnas.1808191115

LIU, B., YANG, C. and LI, F. (2026). Supplement to "Sample size and power calculations for causal inference in observational studies." https://doi.org/10.1214/26-AOS2649SUPP

LUNCEFORD, J. K. and DAVIDIAN, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Stat. Med.* **23** 2937–2960.

MUKHERJEE, S., TAMAYO, P., ROGERS, S., RIFKIN, R., ENGLE, A., CAMPBELL, C., GOLUB, T. R. and MESIROV, J. P. (2003). Estimating dataset size requirements for classifying DNA microarray data. *J. Comput. Biol.* **10** 119–142.

QIN, X. (2024). Sample size and power calculations for causal mediation analysis: A tutorial and shiny app. *Behav. Res. Methods* **56** 1738–1769.

RAUDENBUSH, S. W. (1997). Statistical analysis and optimal design for cluster randomised trials. *Psychol. Methods* **2** 173–185.

ROBINS, J. M., ROTNITZKY, A. and ZHAO, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *J. Amer. Statist. Assoc.* **89** 846–866. MR1294730

ROSENBAUM, P. R. and RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70** 41–55. MR0742974 https://doi.org/10.1093/biomet/70.1.41

RUBIN, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *J. Amer. Statist. Assoc.* **75** 591–593.

SHOOK-SA, B. E. and HUDGENS, M. G. (2022). Power and sample size for observational studies of point exposure effects. *Biometrics* **78** 388–398. MR4408596 https://doi.org/10.1111/biom.13405

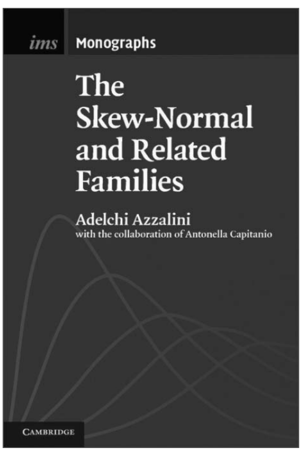
TSIATIS, A. A. (2006). *Semiparametric Theory and Missing Data. Springer Series in Statistics*. Springer, New York. MR2233926

YANG, C., LIU, B. and LI, F. (2026). Sample size and power calculations for causal inference with time-to-event outcomes. arXiv preprint. Available at arXiv:2605.10088.

YANG, S. and DING, P. (2018). Asymptotic inference of causal effects with observational studies trimmed by the estimated propensity scores. *Biometrika* **105** 487–493. MR3804418 https://doi.org/10.1093/biomet/asy008

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.