BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Two Number Four November 2026 ISSN: 1350-7265

CONTENTS

(continued)

A list of forthcoming papers can be found online at https://www.bernoullisociety.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Two Number Four November 2026 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 32 Number 4 November 2026 Pages 2543–3381 ISI/BS

Volume 32 Number 4   November 2026   ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope

Issued four times per year, BERNOULLI is the flagship journal of the Bernoulli Society for Mathematical Statistics and Probability. The journal aims at publishing original research contributions of the highest quality in all subfields of Mathematical Statistics and Probability. The main emphasis of Bernoulli is on theoretical work, yet discussion of interesting applications in relation to the proposed methodology is also welcome.

Bernoulli Society for Mathematical Statistics and Probability

The Bernoulli Society was founded in 1973. It is an autonomous Association of the International Statistical Institute, ISI. According to its statutes, the object of the Bernoulli Society is the advancement, through international contacts, of the sciences of probability (including the theory of stochastic processes) and mathematical statistics and of their applications to all those aspects of human endeavour which are directed towards the increase of natural knowledge and the welfare of mankind.

Meetings: https://www.bernoullisociety.org/meetings

The Society holds a World Congress every four years; more frequent meetings, coordinated by the Society's standing committees and often organised in collaboration with other organisations, are the European Meeting of Statisticians, the Conference on Stochastic Processes and their Applications, the CLAPEM meeting (Latin-American Congress on Probability and Mathematical Statistics), the European Young Statisticians Meeting, and various meetings on special topics – in the physical sciences in particular. The Society, as an association of the ISI, also collaborates with other ISI associations in the organization of the biennial ISI World Statistics Congresses (formerly ISI Sessions).

Executive Committee

Detailed information about the members of the Executive Committee can be found on https://www.bernoullisociety.org/who-is-who

---

The papers published in Bernoulli are indexed or abstracted in Mathematical Reviews (*MathSciNet*), Zentralblatt MATH (*zbMATH Open*), Science Citation Index Expanded (*Web of Science*), *SCOPUS* and *Google Scholar*.

In 2026 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society

for Mathematical Statistics

and Probability

Bernoulli **32**(4), 2026, 2543–2568
https://doi.org/10.3150/24-BEJ1813

Estimation and inference for the Wasserstein distance between mixing measures in topic models

XIN BING[1,a], FLORENTINA BUNEA[2,b] and JONATHAN NILES-WEED[3,c]

[1]*Department of Statistical Sciences, University of Toronto, Toronto, Canada,* [a]*xin.bing@utoronto.ca*
[2]*Department of Statistics and Data Science, Cornell University, Ithaca, USA,* [b]*fb238@cornell.edu*
[3]*Courant Institute for Mathematical Sciences and Center for Data Science, New York University, New York, USA,* [c]*jnw@cims.nyu.edu*

The Wasserstein distance between mixing measures has come to occupy a central place in the statistical analysis of mixture models. This work proposes a new canonical interpretation of this distance and provides tools to perform inference on the Wasserstein distance between mixing measures in topic models. We consider the general setting of an identifiable mixture model consisting of mixtures of distributions from a set $\mathcal{A}$ equipped with an arbitrary metric d, and show that the Wasserstein distance between mixing measures is uniquely characterized as the most discriminative convex extension of the metric d to the set of mixtures of elements of $\mathcal{A}$. The Wasserstein distance between mixing measures has been widely used in the study of such models, but without axiomatic justification. Our results establish this metric to be a canonical choice. Specializing our results to topic models, we consider estimation and inference of this distance. Although upper bounds for its estimation have been recently established elsewhere, we prove the first minimax lower bounds for the estimation of the Wasserstein distance between mixing measures, in topic models, when both the mixing weights and the mixture components need to be estimated. Our second main contribution is the provision of fully data-driven inferential tools for estimators of the Wasserstein distance between potentially sparse mixing measures of high-dimensional discrete probability distributions on p points, in the topic model context. These results allow us to obtain the first asymptotically valid, ready to use, confidence intervals for the Wasserstein distance in (sparse) topic models with potentially growing ambient dimension p.

Keywords: De-biased mle; limiting distribution; mixing measure; mixture distribution; optimal rates of convergence; probability distance; sparse mixtures; topic model; Wasserstein distance

References

[1] Agresti, A. (2013). *Categorical Data Analysis*, 3rd ed. *Wiley Series in Probability and Statistics. Wiley-Interscience*. Hoboken, NJ: John Wiley & Sons. MR3087436

[2] Andrews, D.W.K. (2002). Generalized method of moments estimation when a parameter is on a boundary. *J. Bus. Econom. Statist.* **20** 530–544. Twentieth anniversary GMM issue. MR1945607 https://doi.org/10.1198/073500102288618667

[3] Arora, S., Ge, R. and Moitra, A. (2012). Learning topic models - going beyond SVD. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012* 1–10. IEEE Computer Society. https://doi.org/10.1109/FOCS.2012.49

[4] Arora, S., Ge, R., Halpern, Y., Mimno, D.M., Moitra, A., Sontag, D.A., Wu, Y. and Zhu, M. (2013). A practical algorithm for topic modeling with provable guarantees. In *Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16-21 June 2013. JMLR Workshop and Conference Proceedings* **28** 280–288. Available at JMLR.org.

[5] Ba, K.D., Nguyen, H.L., Nguyen, H.N. and Rubinfeld, R. (2011). Sublinear time algorithms for Earth mover's distance. *Theory Comput. Syst.* **48** 428–442. https://doi.org/10.1007/S00224-010-9265-8

[6] Bansal, T., Bhattacharyya, C. and Kannan, R. (2014). A provable SVD-based algorithm for learning topics in dominant admixture corpus. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada* (Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence and K.Q. Weinberger, eds.) 1997–2005.

[7] Bing, X., Bunea, F. and Niles-Weed, J. (2026). Supplement to "Estimation and inference for the Wasserstein distance between mixing measures in topic models." https://doi.org/10.3150/24-BEJ1813SUPP

[8] Bing, X., Bunea, F. and Wegkamp, M. (2020). A fast algorithm with minimax optimal guarantees for topic models with an unknown number of topics. *Bernoulli* **26** 1765–1796. MR4091091 https://doi.org/10.3150/19-BEJ1166

[9] Bing, X., Bunea, F. and Wegkamp, M. (2020). Optimal estimation of sparse topic models. *J. Mach. Learn. Res.* **21**. Paper No. 177, 45. MR4209463

[10] Bing, X., Bunea, F., Strimas-Mackey, S. and Wegkamp, M. (2022). Likelihood estimation of sparse topic distributions in topic models and its applications to Wasserstein document distance calculations. *Ann. Statist.* **50** 3307–3333. https://doi.org/10.1214/22-AOS2229

[11] Birch, M.W. (1964). A new proof of the Pearson-Fisher theorem. *Ann. Math. Statist.* **35** 817–824. MR169324 https://doi.org/10.1214/aoms/1177703581

[12] Blei, D.M. (2012). Probabilistic topic models. *Commun. ACM* **55** 77–84. https://doi.org/10.1145/2133806.2133826

[13] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003). Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3** 993–1022.

[14] Bravo González-Blas, C., Minnoye, L., Papasokrati, D., Aibar, S., Hulselmans, G., Christiaens, V., Davie, K., Wouters, J. and Aerts, S. (2019). cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat. Methods* **16** 397–400.

[15] Cai, T.T. and Low, M.G. (2011). Testing composite hypotheses, Hermite polynomials and optimal estimation of a nonsmooth functional. *Ann. Statist.* **39** 1012–1041. MR2816346 https://doi.org/10.1214/10-AOS849

[16] Chen, J.H. (1995). Optimal rate of convergence for finite mixture models. *Ann. Statist.* **23** 221–233. MR1331665 https://doi.org/10.1214/aos/1176324464

[17] Chen, S., Rivaud, P., Park, J.H., Tsou, T., Charles, E., Haliburton, J.R., Pichiorri, F. and Thomson, M. (2020). Dissecting heterogeneous cell populations across drug and disease conditions with PopAlign. *Proc. Natl. Acad. Sci. USA* **117** 28784–28794. https://doi.org/10.1073/pnas.2005990117

[18] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999 https://doi.org/10.1214/aop/1022677394

[19] del Barrio, E., González Sanz, A. and Loubes, J.-M. (2024). Central limit theorems for semi-discrete Wasserstein distances. *Bernoulli* **30** 554–580. MR4665589 https://doi.org/10.3150/23-bej1608

[20] del Barrio, E., González-Sanz, A. and Loubes, J.-M. (2024). Central limit theorems for general transportation costs. *Ann. Inst. Henri Poincaré Probab. Stat.* **60** 847–873. MR4757510 https://doi.org/10.1214/22-aihp1356

[21] del Barrio, E., Gordaliza, P. and Loubes, J.-M. (2019). A central limit theorem for L_p transportation cost on the real line with application to fairness assessment in machine learning. *Inf. Inference* **8** 817–849. MR4045479 https://doi.org/10.1093/imaiai/iaz016

[22] del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** 926–951. MR3916938 https://doi.org/10.1214/18-AOP1275

[23] Ding, W., Ishwar, P. and Saligrama, V. (2015). Most large topic models are approximately separable. In *2015 Information Theory and Applications Workshop (ITA)* 199–203. https://doi.org/10.1109/ITA.2015.7308989

[24] Donoho, D.L. and Stodden, V. (2003). When does non-negative matrix factorization give a correct decomposition into parts? In *Advances in Neural Information Processing Systems 16 [Neural Information Processing Systems, NIPS 2003, December 8-13, 2003, Vancouver and Whistler, British Columbia, Canada]* (S. Thrun, L.K. Saul and B. Schölkopf, eds.) 1141–1148. MIT Press.

[25] Dümbgen, L. (1993). On nondifferentiable functions and the bootstrap. *Probab. Theory Related Fields* **95** 125–140. MR1207311 https://doi.org/10.1007/BF01197342

[26] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26. MR515681

[27] Fang, Z. and Santos, A. (2019). Inference on directionally differentiable functions. *Rev. Econ. Stud.* **86** 377–412. MR3936869 https://doi.org/10.1093/restud/rdy049

[28] Fu, X., Huang, K., Sidiropoulos, N.D., Shi, Q. and Hong, M. (2019). Anchor-free correlated topic modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **41** 1056–1071. https://doi.org/10.1109/TPAMI.2018.2827377

[29] Han, Y., Jiao, J. and Weissman, T. (2020). Minimax estimation of divergences between discrete distributions. *IEEE J. Sel. Areas Inf. Theory* **1** 814–823. https://doi.org/10.1109/JSAIT.2020.3041036

[30] Heinrich, P. and Kahn, J. (2018). Strong identifiability and optimal minimax rates for finite mixture estimation. *Ann. Statist.* **46** 2844–2870. MR3851757 https://doi.org/10.1214/17-AOS1641

[31] Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993). *Convex Analysis and Minimization Algorithms. I. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **305**. Berlin: Springer. MR1261420

[32] Ho, N. and Nguyen, X. (2016). Convergence rates of parameter estimation for some weakly identifiable finite mixtures. *Ann. Statist.* **44** 2726–2755. MR3576559 https://doi.org/10.1214/16-AOS1444

[33] Ho, N. and Nguyen, X. (2016). On strong identifiability and convergence rates of parameter estimation in finite mixtures. *Electron. J. Stat.* **10** 271–307. MR3466183 https://doi.org/10.1214/16-EJS1105

[34] Ho, N., Nguyen, X. and Ritov, Y. (2020). Robust estimation of mixing measures in finite mixture models. *Bernoulli* **26** 828–857. MR4058353 https://doi.org/10.3150/18-BEJ1087

[35] Jiao, J., Han, Y. and Weissman, T. (2018). Minimax estimation of the L_1 distance. *IEEE Trans. Inf. Theory* **64** 6672–6706. https://doi.org/10.1109/TIT.2018.2846245

[36] Kirk, W. and Shahzad, N. (2014). *Fixed Point Theory in Distance Spaces*. Cham: Springer. MR3288309 https://doi.org/10.1007/978-3-319-10927-5

[37] Klopp, O., Panov, M., Sigalla, S. and Tsybakov, A.B. (2023). Assigning topics to documents by successive projections. *Ann. Statist.* **51** 1989–2014. MR4678793 https://doi.org/10.1214/23-aos2316

[38] Le Cam, L. and Lo Yang, G. (1990). Locally asymptotically normal families. In *Asymptotics in Statistics: Some Basic Concepts* 52–98. New York, NY: Springer US.

[39] Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the L_r norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. MR1670867 https://doi.org/10.1007/s004409970006

[40] Leroux, B.G. (1992). Consistent estimation of a mixing distribution. *Ann. Statist.* **20** 1350–1360. MR1186253 https://doi.org/10.1214/aos/1176348772

[41] Liang, T. (2019). On the minimax optimality of estimating the Wasserstein metric.

[42] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. MR733519 https://doi.org/10.1214/aos/1176346412

[43] Maceachern, S.N. and Müller, P. (1998). Estimating mixture of Dirichlet process models. *J. Comput. Graph. Statist.* **7** 223–238. https://doi.org/10.1080/10618600.1998.10474772

[44] Manole, T. and Khalili, A. (2021). Estimating the number of components in finite mixture models via the group-sort-fuse procedure. *Ann. Statist.* **49** 3043–3069. MR4352522 https://doi.org/10.1214/21-aos2072

[45] Manole, T., Balakrishnan, S., Niles-Weed, J. and Wasserman, L. (2024). Plugin estimation of smooth optimal transport maps. *Ann. Statist.* To appear.

[46] McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. *Wiley Series in Probability and Statistics: Applied Probability and Statistics*. New York: Wiley-Interscience. MR1789474 https://doi.org/10.1002/0471721182

[47] Nguyen, X. (2013). Convergence of latent mixing measures in finite and infinite mixture models. *Ann. Statist.* **41** 370–400. MR3059422 https://doi.org/10.1214/12-AOS1065

[48] Niles-Weed, J. and Rigollet, P. (2022). Estimation of Wasserstein distances in the spiked transport model. *Bernoulli* **28** 2663–2688. MR4474558 https://doi.org/10.3150/21-bej1433

[49] Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **185** 71–110.

[50] Rao, C.R. (1957). Maximum likelihood estimation for the multinomial distribution. *Sankhyā* **18** 139–148. MR0105183

[51] Rao, C.R. (1958). Maximum likelihood estimation for the multinomial distribution with infinite number of cells. *Sankhyā* **20** 211–218. MR0107334

[52] Recht, B., Ré, C., Tropp, J.A. and Bittorf, V. (2012). Factoring nonnegative matrices with linear programs. In *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a Meeting Held December 3-6, 2012, Lake Tahoe, Nevada, United States* (P.L. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger, eds.) 1223–1231.

[53] Shapiro, A. (1991). Asymptotic analysis of stochastic programs. *Ann. Oper. Res.* **30** 169–186. Stochastic programming, Part I (Ann Arbor, MI, 1989). MR1118896 https://doi.org/10.1007/BF02204815
[54] Sommerfeld, M. and Munk, A. (2018). Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 219–238. MR3744719 https://doi.org/10.1111/rssb.12236
[55] Takahashi, W. (1970). A convexity in metric space and nonexpansive mappings. I. *Kodai Math. Semin. Rep.* **22** 142–149. MR0267565
[56] Tameling, C., Sommerfeld, M. and Munk, A. (2019). Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications. *Ann. Appl. Probab.* **29** 2744–2781. MR4019874 https://doi.org/10.1214/19-AAP1463
[57] Tracy Ke, Z. and Wang, M. (2024). Using SVD for topic modeling. *J. Amer. Statist. Assoc.* **119** 434–449. MR4713904 https://doi.org/10.1080/01621459.2022.2123813
[58] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original. Translated by Vladimir Zaiats. MR2724359 https://doi.org/10.1007/b13794
[59] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge University Press. MR1652247 https://doi.org/10.1017/CBO9780511802256
[60] Wang, S., Cai, T.T. and Li, H. (2021). Optimal estimation of Wasserstein distance on a tree with an application to microbiome studies. *J. Amer. Statist. Assoc.* **116** 1237–1253. MR4309269 https://doi.org/10.1080/01621459.2019.1699422
[61] Weaver, N. (2018). *Lipschitz Algebras*, 2nd ed. Hackensack, NJ: World Scientific Publishing Co. Pte. Ltd. Second edition of [MR1832645]. MR3792558
[62] Wu, Y. and Yang, P. (2019). Chebyshev polynomials, moment matching, and optimal estimation of the unseen. *Ann. Statist.* **47** 857–883. MR3909953 https://doi.org/10.1214/17-AOS1665
[63] Wu, Y. and Yang, P. (2020). Optimal estimation of Gaussian mixtures via denoised method of moments. *Ann. Statist.* **48** 1981–2007. MR4134783 https://doi.org/10.1214/19-AOS1873
[64] Wu, R., Zhang, L. and Cai, T.T. (2023). Sparse topic modeling: Computational efficiency, near-optimal algorithms, and statistical inference. *J. Amer. Statist. Assoc.* **118** 1849–1861. MR4646611 https://doi.org/10.1080/01621459.2021.2018329

Bernoulli **32**(4), 2026, 2569–2593
https://doi.org/10.3150/24-BEJ1832

Estimation of a score-explained non-randomized treatment effect in fixed and high dimensions

DEBARGHYA MUKHERJEE[1,a], MOULINATH BANERJEE[2,b] and YA'ACOV RITOV[2,c]

[1]*Department of Mathematics and Statistics, Boston University, Boston, USA,* [a]*mdeb@bu.edu*
[2]*Department of Statistics, University of Michigan, Ann Arbor, USA,* [b]*moulib@umich.edu,* [c]*yritov@umich.edu*

In non-randomized treatment allocation models, treatment is assigned to a unit based on a score, e.g., scholarship is allocated based on the score obtained in a merit test, while antihypertensive treatments are allocated based on blood pressure level. In this paper, we present a new model coined SCENTS: Score Explained Non-Randomized Treatment Systems that utilizes the dependency of the score on the explanatory variables to permit efficient estimation. We derive an estimator of the treatment effect which is $\sqrt{n}$ consistent, asymptotically normal, and achieves semiparametric efficiency under normal errors. The analysis is extended to ultra-high dimensional vectors of covariates, where a $\sqrt{n}$ consistent and asymptotically normal debiased estimator is proposed. We analyze two real data sets via our method and compare our results with those obtained by using previous approaches like regression discontinuity design. Some possible extensions are discussed.

Keywords: Confounding variables; endogeneity; global effect; high dimensional analysis; semiparametric efficiency

References

Angrist, J.D. and Imbens, G.W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J. Amer. Statist. Assoc.* **90** 431–442.

Angrist, J.D. and Rokkanen, M. (2015). Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff. *J. Amer. Statist. Assoc.* **110** 1331–1344.

Arai, Y., Otsu, T., Seo, M.H. et al. (2019). Causal inference on regression discontinuity designs by high-dimensional methods. LSE, STICERD.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Rev. Econ. Stud.* **81** 608–650.

Berk, R.A. and Rauma, D. (1983). Capitalizing on nonrandom assignment to treatments: A regression-discontinuity evaluation of a crime-control program. *J. Amer. Statist. Assoc.* **78** 21–27.

Bhattacharya, P. and Zhao, P.L. (1997). Semiparametric inference in a partial linear model. *Ann. Statist.* **25** 244–262.

Bickel, P.J., Ritov, Y., Tsybakov, A.B. et al. (2009). Simultaneous analysis of Lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732.

Calonico, S. (2020). Package 'rdrobust'.

Calonico, S., Cattaneo, M.D. and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* **82** 2295–2326.

Calonico, S., Cattaneo, M.D., Farrell, M.H. and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Rev. Econ. Stat.* **101** 442–451.

Campbell, D.T. and Stanley, J. (1963). *Experimental Designs and Quasi-Experimental Designs for Research.* Boston: Houghton Mifflin Company.

Cattaneo, M.D., Idrobo, N. and Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **21** C1–C68.

Erickson, M.R. and Cromack, T. (1972). Evaluating a tutoring program. *J. Exp. Educ.* **41** 27–31.
Finkelstein, M.O., Levin, B. and Robbins, H. (1996). Clinical and prophylactic trials with assured new treatment for those at greater risk: I. A design proposal. *Amer. J. Publ. Health* **86** 691–695.
Guo, W., Zhou, X.H. and Ma, S. (2021). Estimation of optimal individualized treatment rules using a covariate-specific treatment effect curve with high-dimensional covariates. *J. Amer. Statist. Assoc.* **116** 309–321.
Han, F., Ren, Z. and Zhu, Y. (2017). Adaptive estimation of high dimensional partially linear model. arXiv preprint, arXiv:1705.08930.
Holland, J.L. and Stalnaker, J.M. (1957). An honorary scholastic award. *J. High. Educ.* **28** 361–368.
Honoré, B.E. and Powell, J. (2005). Pairwise difference estimators for nonlinear models. In *Identification and Inference for Econometric Models 520–553* Cambridge University Press.
Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Rev. Econ. Stud.* **79** 933–959.
Lian, H., Zhao, K. and Lv, S. (2019). Projected spline estimation of the nonparametric function in high-dimensional partially linear models for massive data. *Ann. Statist.* **47** 2922–2949.
Lindo, J.M., Sanders, N.J. and Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *Amer. Econ. J. Appl. Econ.* **2** 95–117.
Lohr, B. (1972). An historical view of the research on the factors related to the utilization of health services. Bureau for Health Services Research and Evaluation, Social and Economic Analysis Division, Rockville, MD.
Lousdal, M.L. (2018). An introduction to instrumental variable assumptions, validation and estimation. *Emerg. Themes Epidemiol.* **15** 1–7.
Lv, S. and Lian, H. (2017). Debiased distributed learning for sparse partial linear models in high dimensions. arXiv preprint, arXiv:1708.05487.
Meyersson, E. (2014). Islamic rule and the empowerment of the poor and pious. *Econometrica* **82** 229–269.
Mukherjee, D., Banerjee, M. and Ritov, Y. (2026). Supplement to "Estimation of a score-explained non-randomized treatment effect in fixed and high dimensions." https://doi.org/10.3150/24-BEJ1832SUPP
Müller, P. and Van de Geer, S. (2015). The partial linear model in high dimensions. *Scand. J. Stat.* **42** 580–608.
Newey, W.K., Powell, J.L. and Vella, F. (1999). Nonparametric estimation of triangular simultaneous equations models. *Econometrica* **67** 565–603.
Pinkse, J. (2000). Nonparametric two-step regression estimation when regressors and error are dependent. *Canad. J. Statist.* **28** 289–300.
Porter, J. (2003). Estimation in the regression discontinuity model. 5–19. Department of Economics, University of Wisconsin at Madison. Unpublished Manuscript, 2003.
Qiu, Y., Tao, J. and Zhou, X.H. (2021). Inference of heterogeneous treatment effects using observational data with high-dimensional covariates. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 1016–1043.
Robinson, P.M. (1988). Root-N-consistent semiparametric regression. *Econometrica* **56** 931–954.
Rudelson, M. and Zhou, S. (2012). Reconstruction from anisotropic random measurements. In *COLT 10–1. JMLR Workshop and Conference Proceedings*.
Schick, A. (1986). On asymptotically efficient estimation in semiparametric models. *Ann. Statist.* **14** 1139–1151.
Stadthaus, A.M. (1972). A comparison of the subsequent academic achievement of marginal selectees and rejectees for the Cincinnati Public Schools Special College Preparatory Program: An application of Campbell's regression discontinuity design PhD thesis, ProQuest Information & Learning.
Thistlethwaite, D.L. (1959). Effects of social recognition upon the educational motivation of talented youth. *J. Educ. Psychol.* **50** 111.
Thistlethwaite, D.L. and Campbell, D.T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *J. Educ. Psychol.* **51** 309.
Yatchew, A. (1997). An elementary estimator of the partial linear model. *Econom. Lett.* **57** 135–143.
Zhu, Y. (2017). Nonasymptotic analysis of semiparametric regression models with high-dimensional parametric coefficients. *Ann. Statist.* **45** 2274–2298.
Zhu, Y., Yu, Z. and Cheng, G. (2019). High dimensional inference in partially linear models. In *The 22nd AISTATS* 2760–2769. PMLR.

Bernoulli **32**(4), 2026, 2594–2618
https://doi.org/10.3150/25-BEJ1884

Asymptotics of numerical integration for two-level mixed models

BLAIR BILODEAU[1,a], ALEX STRINGER[2,b] and YANBO TANG[3,c]

[1] *Department of Statistics, University of Toronto, Toronto, Canada,* [a]*blair.bilodeau@mail.utoronto.ca*
[2] *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada,* [b]*alex.stringer@uwaterloo.ca*
[3] *Department of Mathematics, Imperial College London, London, UK,* [c]*yanbo.tang@imperial.ac.uk*

We study mixed models with a single grouping factor where inference about unknown parameters requires optimizing a marginal likelihood defined by an intractable integral. Low-dimensional numerical integration techniques are regularly used to approximate these integrals with inferences about parameters based on the resulting approximate marginal likelihood. For a generic class of mixed models that satisfy explicit regularity conditions we derive the stochastic relative error rate incurred for both the likelihood and maximum likelihood estimator when adaptive numerical integration is used to approximate the marginal likelihood. We then specialize the analysis to well-specified generalized linear mixed models having exponential family response and multivariate Gaussian random effects, verifying that the regularity conditions hold and hence that the convergence rates apply. We also prove that for models with likelihoods satisfying very weak concentration conditions that the maximum likelihood estimators from non-adaptive numerical integration approximations of the marginal likelihood are not consistent, further motivating adaptive numerical integration as the preferred tool for inference in mixed models. Code to reproduce the simulations in this paper is provided at https://github.com/awstringer1/aq-theory-paper-code.

Keywords: Adaptive quadrature; approximate inference; generalized linear models

References

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67** 1–48.

Bhaskaran, A. and Wand, M.P. (2023). Dispersion parameter extension of precise generalized linear mixed model asymptotics. *Statist. Probab. Lett.* **193** 109691.

Bianconcini, S. (2014). Asymptotic properties of adaptive maximum likelihood estimators in latent variable models. *Bernoulli* **20** 1507–1531.

Bilodeau, B., Stringer, A. and Tang, Y. (2024). Stochastic convergence rates and applications of adaptive quadrature in Bayesian inference. *J. Amer. Statist. Assoc.* **119** 690–700.

Bilodeau, B., Stringer, A. and Tang, Y. (2026). Supplement to “Asymptotics of numerical integration for two-level mixed models.” https://doi.org/10.3150/25-BEJ1884SUPPA, https://doi.org/10.3150/25-BEJ1884SUPPB

Breslow, N.E. and Clayton, D.G. (1993). Approximate inference in generalized linear mixed models. *J. Amer. Statist. Assoc.* **88** 9–25.

Davis, P.J. and Rabinowitz, P. (1984). *Methods of Numerical Integration*. Academic Press.

Diggle, P., Moraga, P., Rowlingson, B. and Taylor, B. (2013). Spatial and spatio-temporal log-Gaussian Cox processes: Extending the geostatistical paradigm. *Statist. Sci.* **28** 542–563.

Jiang, J., Wand, M.P. and Bhaskaran, A. (2021). Usable and precise asymptotics for generalized linear mixed model analysis and design. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 55–82.

Jin, S. and Andersson, B. (2020). A note on the accuracy of adaptive Gauss-Hermite quadrature. *Biometrika* **107** 737–744.

Kauermann, G., Krivobokova, T. and Fahrmeir, L. (2009). Some asymptotic results on generalized penalized spline smoothing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 487–503.

1350-7265

Kleijn, B.J.K. and van der Vaart, A.W. (2012). The Bernstein von-Mises theorem under misspecification. *Electron. J. Stat.* **6** 354–381.

Lesaffre, E. and Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random-effects model: An example. *J. R. Stat. Soc., Ser. C, Appl. Stat.* **50** 325–335.

Liu, Q. and Pierce, D.A. (1994). A note on Gauss-Hermite quadrature. *Biometrika* **81** 624–629.

Maestrini, L., Bhaskaran, A. and Wand, M.P. (2024). Second term improvement to generalized linear mixed model asymptotics. *Biometrika* **111** 1077–1084.

Naylor, J.C. and Smith, A.F.M. (1982). Applications of a method for the efficient computation of posterior distributions. *J. R. Stat. Soc., Ser. C, Appl. Stat.* **31** 214–225.

Nie, L. (2007). Convergence rate of MLE in generalized linear and nonlinear mixed-effects models: Theory and applications. *J. Statist. Plann. Inference* **137** 1787–1804.

Ogden, H. (2017). On asymptotic validity of naive inference with an approximate likelihood. *Biometrika* **104** 153–164.

Owen, A.B. (2019). Monte Carlo Book: The Quasi-Monte Carlo Parts. Unpublished manuscript.

Pinheiro, J.C. and Bates, D.M. (1995). Approximations to the log-likelihood function in the nonlinear mixed effects models. *J. Comput. Graph. Statist.* **4** 12–35.

Rizopoulos, D. (2020). GLMMadaptive: generalized linear mixed models using adaptive Gaussian quadrature R package version 0.7-15.

Sanz-Alonso, D. and Yang, R. (2022). Finite element representations of Gaussian processes: Balancing numerical and statistical accuracy. *SIAM/ASA J. Uncertain. Quantificat.* **10** 1323–1349.

Tang, Y. and Reid, N. (2025). Laplace and saddlepoint approximations in high dimensions. *Bernoulli.* **31** 1759–1788.

van der Vaart, A. (1998). *Asymptotic Statistics*. Cambridge University Press.

Vonesh, E.F. (1996). A note on the use of Laplace's approximation for nonlinear mixed effect models. *Biometrika* **83** 447–452.

Wood, S. (2016). Smoothing parameter and model selection for general smooth models. *J. Amer. Statist. Assoc.* **111** 1548–1575.

Bernoulli **32**(4), 2026, 2619–2638
https://doi.org/10.3150/25-BEJ1915

Seriation of Tœplitz and latent position matrices: Optimal rates and computational trade-offs

CLÉMENT BERENFELD[1,a], ALEXANDRA CARPENTIER[2,b] and NICOLAS VERZELEN[3,c]

[1] *INRIA, PreMeDICaL Team, Univ. Montpellier, Montpellier, France,* [a] *clement.berenfeld@inria.fr*
[2] *Institut für Mathematik – Universität Potsdam, Potsdam, Germany,* [b] *carpentier@uni-potsdam.de*
[3] *INRAE, Institut Agro, MISTEA, Univ. Montpellier, Montpellier, France,* [c] *nicolas.verzelen@inrae.fr*

In this paper, we consider the problem of seriation of a permuted structured matrix based on noisy observations. The entries of the matrix relate to an expected quantification of interaction between two objects: the higher the value, the closer the objects. A popular structured class for modeling such matrices is the permuted Robinson class, namely the set of matrices whose coefficients are decreasing away from its diagonal, up to a permutation of its rows and columns. We consider in this paper two submodels of Robinson matrices: the Tœplitz model, and the latent position model. We provide a computational lower bound based on the low-degree paradigm, which hints that there is a statistical-computational gap for seriation when measuring the error based on the Frobenius norm. We also provide a simple and polynomial-time algorithm that achieves this lower bound. Along the way, we also characterize the information-theory optimal risk thereby giving evidence for the extent of the computation/information gap for this problem.

Keywords: Computational barrier; geometric graph; seriation; Tœplitz matrix

References

[1] Atkins, J.E., Boman, E.G. and Hendrickson, B. (1998). A spectral algorithm for seriation and the consecutive ones problem. *SIAM J. Comput.* **28** 297–310.
[2] Barnard, S.T., Pothen, A. and Simon, H. (1995). A spectral algorithm for envelope reduction of sparse matrices. *Numer. Linear Algebra Appl.* **2** 317–334.
[3] Barthélemy, J.-P. and Brucker, F. (2001). NP-hard approximation problems in overlapping clustering. *J. Classification* **18** 159–183.
[4] Berenfeld, C., Carpentier, A. and Verzelen, N. (2026). Supplement to "Seriation of Tœplitz and latent position matrices: optimal rates and computational trade-offs." https://doi.org/10.3150/25-BEJ1915SUPP
[5] Cai, T.T. and Ma, R. (2023). Matrix reordering for noisy disordered matrices: Optimality and computationally efficient algorithms. *IEEE Trans. Inf. Theory*.
[6] Carmona, M., Chepoi, V., Naves, G. and Préa, P. (2024). Modules in Robinson spaces. *SIAM J. Discrete Math.* **38** 190–224.
[7] Chepoi, V., Fichet, B. and Seston, M. (2009). Seriation in the presence of errors: NP-hardness of l_∞-fitting Robinson structures to dissimilarity matrices. *J. Classification* **26** 279–296.
[8] Chepoi, V. and Seston, M. (2011). Seriation in the presence of errors: A factor 16 approximation algorithm for l_∞-fitting Robinson structures to distances. *Algorithmica* **59** 521–568.
[9] Diaz, J., McDiarmid, C. and Mitsche, D. (2020). Learning random points from geometric graphs or orderings. *Random Structures Algorithms* **57** 339–370.
[10] Elson, J., Karp, R.M., Papadimitriou, C.H. and Shenker, S. (2004). Global synchronization in sensornets. In *LATIN 2004: Theoretical Informatics* (M. Farach-Colton, ed.) 609–624. Berlin, Heidelberg: Springer.
[11] Flammarion, N., Mao, C. and Rigollet, P. (2019). Optimal rates of statistical seriation. *Bernoulli* **25** 623–653.
[12] Fogel, F., Jenatton, R., Bach, F. and d'Aspremont, A. (2013). Convex relaxations for permutation problems. In *Adv. Neural Inf. Process* 1016–1024.

[13] Fulkerson, D. and Gross, O. (1965). Incidence matrices and interval graphs. *Pacific J. Math.* **15** 835–855.
[14] Garriga, G.C., Junttila, E. and Mannila, H. (2011). Banded structure in binary matrices. *Knowl. Inf. Syst.* **28** 197–226.
[15] Gilbert, E.N. (1961). Random plane networks. *SIAM J. Appl.* **9** 533–543.
[16] Giraud, C., Issartel, Y. and Verzelen, N. (2023). Localization in 1D non-parametric latent space models from pairwise affinities. *Electron. J. Stat.* **17** 1587–1662.
[17] Giridhar, A. and Kumar, P.R. (2006). Distributed clock synchronization over wireless networks: Algorithms and analysis. In *45th Proc. IEEE Conf. Decis. Control* 4915–4920.
[18] Han, Y., Rigollet, P. and Stepaniants, G. (2023). Covariance alignment: from maximum likelihood estimation to Gromov-Wasserstein. arXiv preprint, arXiv:2311.13595.
[19] Hoff, P.D., Raftery, A.E. and Handcock, M.S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist.* **97** 1090–1098.
[20] Issartel, Y., Giraud, C. and Verzelen, N. (2024). Minimax optimal seriation in polynomial time. arXiv preprint, arXiv:2405.08747.
[21] Janssen, J. and Smith, A. (2022). Reconstruction of line-embeddings of graphons. *Electron. J. Stat.* **16** 331–407.
[22] Karin, J., Bornfeld, Y. and Nitzan, M. (2023). scPrisma infers, filters and enhances topological signals in single-cell data using spectral template matching. *Nat. Biotechnol.* **41** 1645–1654.
[23] Kunisky, D., Wein, A.S. and Bandeira, A.S. (2019). Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *ISAAC Congress* 1–50. Springer.
[24] Luo, Y. and Gao, C. (2024). Computational lower bounds for graphon estimation via low-degree polynomials. *Ann. Statist.* **52** 2318–2348.
[25] Natik, A. and Smith, A. (2021). Consistency of spectral seriation. arXiv preprint. Available at arXiv:2112.04408.
[26] Penrose, M. (2003). *Random Geometric Graphs 5*. Oxford university press.
[27] Pilliat, E., Carpentier, A. and Verzelen, N. (2023). Optimal permutation estimation in crowdsourcing problems. *Ann. Statist.* **51** 935–961.
[28] Pilliat, E., Carpentier, A. and Verzelen, N. (2024). Optimal rates for ranking a permuted isotonic matrix in polynomial time. In *Proceedings of the 2024 Annual ACM-SIAM SODA* 3236–3273. SIAM.
[29] Recanati, A., Brüls, T. and d'Aspremont, A. (2017). A spectral algorithm for fast de novo layout of uncorrected long nanopore reads. *Bioinformatics* **33** 3188–3194.
[30] Recanati, A., Kerdreux, T. and d'Aspremont, A. (2018). Reconstructing latent orderings by spectral clustering. arXiv preprint. Available at arXiv:1807.07122.
[31] Robinson, W.S. (1951). A method for chronologically ordering archaeological deposits. *Amer. Antiq.* **16** 293–301.
[32] Schramm, T. and Wein, A.S. (2022). Computational barriers to estimation from low-degree polynomials. *Ann. Statist.* **50** 1833–1858.

Bernoulli **32**(4), 2026, 2639–2659
https://doi.org/10.3150/25-BEJ1921

Learning latent tree models with small query complexity

LUC DEVROYE[1,a], GÁBOR LUGOSI[2,3,4,b] and PIOTR ZWIERNIK[2,3,c]

[1] *School of Computer Science, McGill University, Montreal, Canada,* [a] *lucdevroye@gmail.com*
[2] *Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain*
[3] *Barcelona School of Economics, Barcelona, Spain,* [b] *gabor.lugosi@upf.edu,* [c] *piotr.zwiernik@upf.edu*
[4] *Department, ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

We consider the problem of structure recovery in a graphical model of a tree where some variables are latent. Specifically, we focus on the Gaussian case, which can be reformulated as a well-studied problem: recovering a semi-labeled tree from a distance metric. We introduce randomized procedures that achieve query complexity of optimal order. Additionally, we provide statistical analysis for scenarios where the tree distances are noisy. The Gaussian setting can be extended to other situations, including the binary case and non-paranormal distributions.

Keywords: Latent tree models; phylogenetics; query complexity; probabilistic graphical models; structure learning; phylogenetics

References

Abdi, H. (1990). Additive-tree representations. In *Trees and Hierarchical Structures: Proceedings of a Conference Held at Bielefeld, FRG, Oct. 5-9th, 1987. Lecture Notes in Biomathematics* **84** 43–59. Springer.

Afshar, R., Goodrich, M.T., Matias, P. and Osegueda, M.C. (2020). Reconstructing biological and digital phylogenetic trees in parallel. In *28th Annual European Symposium on Algorithms (ESA 2020)* 3:1–3:24. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Aizenbud, Y., Jaffe, A., Wang, M., Hu, A., Amsel, N., Nadler, B., Chang, J.T. and Kluger, Y. (2023). Spectral top-down recovery of latent tree models. *Inf. Inference* **12** iaad032. https://doi.org/10.1093/imaiai/iaad032

Anandkumar, A., Chaudhuri, K., Hsu, D.J., Kakade, S.M., Song, L. and Zhang, T. (2011). Spectral methods for learning multivariate latent tree structure. In *NIPS'11: Proceedings of the 24th International Conference on Neural Information Processing Systems* **24** 2025–2033.

Bastide, P. and Groenland, C. (2024). Optimal distance query reconstruction for graphs without long induced cycles. Available at arXiv:2306.05979.

Batagelj, V., Pisanski, T. and Simoes-Pereira, J.M.S. (1990). An algorithm for tree-realizability of distance matrices. *Int. J. Comput. Math.* **34** 171–176.

Boesch, F.T. (1968). Properties of the distance matrix of a tree. *Quart. Appl. Math.* **26** 607–609.

Brodal, G.S., Fagerberg, R., Pedersen, C.N. and Östlin, A. (2001). The complexity of constructing evolutionary trees using experiments. In *Automata, Languages and Programming: 28th International Colloquium, ICALP 2001 Crete, Greece, July 8–12, 2001. Proceedings* **28** 140–151. Springer.

Buneman, P. (1971). The recovery of trees from measures of dissimilarity. In *Mathematics in the Archaeological and Historical Sciences* (F. Hodson, D. Kendall and P. Tautu, eds.) 387–395. Edinburgh University Press.

Buneman, P. (1974). A note on the metric properties of trees. *J. Combin. Theory Ser. B* **17** 48–50.

Chen, P., Chen, Z. and Zhang, N.L. (2019). A novel document generation process for topic detection based on hierarchical latent tree models. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty: 15th European Conference, ECSQARU 2019, Belgrade, Serbia, September 18-20, 2019. Proceedings* **15** 265–276. Springer.

Choi, M.J., Tan, V.Y.F., Anandkumar, A. and Willsky, A.S. (2011). Learning latent tree graphical models. *J. Mach. Learn. Res.* **12** 1771–1812.

1350-7265

Côme, E., Jouvin, N., Latouche, P. and Bouveyron, C. (2021). Hierarchical clustering with discrete latent variable models and the integrated classification likelihood. *Adv. Data Anal. Classif.* **15** 957–986.

Csürös, M. (2002). Fast recovery of evolutionary trees with thousands of nodes. *J. Comput. Biol.* **9** 277–297.

Culberson, J.C. and Rudnicki, P. (1989). A fast algorithm for constructing trees from distance matrices. *Inform. Process. Lett.* **30** 215–220.

Daskalakis, C., Mossel, E. and Roch, S. (2006). Optimal phylogenetic reconstruction. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing (STOC)* 159–168.

Daskalakis, C., Mossel, E. and Roch, S. (2009). Phylogenies without branch bounds: Contracting the short, pruning the deep. In *Proceedings of the Thirteenth Annual International Conference on Computational Molecular Biology (RECOMB)* 451–465.

Felsenstein, J. (2004). *Inferring Phylogenies*. Sinauer Associates, Inc.

Gąsieniec, L., Jansson, J., Lingas, A. and Östlin, A. (1999). On the complexity of constructing evolutionary trees. *J. Comb. Optim.* **3** 183–197.

Gronau, I., Moran, S. and Snir, S. (2012). Fast and reliable reconstruction of phylogenetic trees with indistinguishable edges. *Random Structures Algorithms* **40** 350–384. https://doi.org/10.1002/rsa.20372

Gusfield, D.M. (1997). *Algorithms on Strings, Trees and Sequences*. Cambridge University Press.

Hakimi, S.L. and Yau, S.S. (1964). Distance matrix of a graph and its realizability. *Quart. Appl. Math.* **22** 305–317.

Hein, J.J. (1989). An optimal algorithm to reconstruct trees from additive distance data. *Bull. Math. Biol.* **51** 597–603.

Huang, F., Naresh, N.U., Perros, I., Chen, R., Sun, J. and Anandkumar, A. (2020). Guaranteed scalable learning of latent tree models. In *Uncertainty in Artificial Intelligence* 883–893. PMLR.

Jaffe, A., Amsel, N., Aizenbud, Y., Nadler, B., Chang, J.T. and Kluger, Y. (2021). Spectral neighbor joining for reconstruction of latent tree models. *SIAM J. Math. Data Sci.* **3** 113–141.

Jansson, J. (2016). Phylogenetic tree construction from a distance matrix. In *Encyclopedia of Algorithms*, 2nd ed. (M.-Y. Kao, ed.) 1564–1567. Springer Science.

Kandiros, V., Daskalakis, C., Dagan, Y. and Choo, D. (2023). Learning and testing latent-tree Ising models efficiently. In *The Thirty Sixth Annual Conference on Learning Theory 1666–1729. PMLR.*

Kannan, S.K., Lawler, E.L. and Warnow, T.J. (1996). Determining the evolutionary tree using experiments. *J. Algebra* **21** 26–50.

Kao, M.-Y., Lingas, A. and Östlin, A. (1999). Balanced randomized tree splitting with applications to evolutionary tree constructions. In *STACS 99: 16th Annual Symposium on Theoretical Aspects of Computer Science Trier, Germany, March 4–6, 1999. Proceedings* **16** 184–196. Springer.

King, V., Zhang, L. and Zhout, Y. (2003). On the complexity of distance-based evolutionary tree reconstruction. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* 444–453. SIAM.

Lauritzen, S.L. (1996). *Graph. Models*. Oxford, United Kingdom: Clarendon Press.

Liu, H., Han, F., Yuan, M., Lafferty, J. and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. *Ann. Statist.* **40** 2293–2326.

Lugosi, G. and Mendelson, S. (2019). Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.* **19** 1145–1190.

Mourad, R., Sinoquet, C., Zhang, N.L., Liu, T. and Leray, P. (2013). A survey on latent tree models and applications. *J. Artificial Intelligence Res.* **47** 157–203.

Nakhleh, L., Warnow, T., Ringe, D. and Evans, S.N. (2005). A comparison of phylogenetic reconstruction methods on an indo-European dataset. *Trans. Philol. Soc.* **103** 171–192. https://doi.org/10.1111/j.1467-968X.2005.00155.x

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. *The Morgan Kaufmann Series in Representation and Reasoning*. San Mateo, CA: Morgan Kaufmann.

Reyzin, L. and Srivastava, N. (2007). On the longest path algorithm for reconstructing trees from distance matrices. *Inform. Process. Lett.* **101** 98–100.

Semple, C. and Steel, M.A. (2003). *Phylogenetics*. *Oxford Lecture Series in Mathematics and Its Applications* **24**. Oxford University Press.

Shiers, N., Zwiernik, P., Aston, J.A.D. and Smith, J.Q. (2016). The correlation space of Gaussian latent tree models and model selection without fitting. *Biometrika* **103** 531–545.

Sturma, N., Drton, M. and Leung, D. (2022). Testing many and possibly singular polynomial constraints. Available at arXiv:2208.11756.

Warnow, T. (1999). Some combinatorial optimization problems in phylogenetics. In *Graph Theory and Combinatorial Biology*. *Bolyai Society Mathematical Studies* **7** 363–413. Bolyai János Matematikai Társulat.

Waterman, M.S., Smith, T.F., Singh, M. and Beyer, W.A. (1977). Additive evolutionary trees. *J. Theoret. Biol.* **64** 199–213.

Wu, B.Y., Chao, K.M. and Tang, C.Y. (1999). Approximation and exact algorithms for constructing minimum ultrametric trees from distance matrices. *J. Comb. Optim.* **3** 199–211.

Zhang, N.L. (2004). Hierarchical latent class models for cluster analysis. *J. Mach. Learn. Res.* **5** 697–723.

Zhang, N. and Poon, L. (2017). Latent tree analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence* **31** 4891–4897.

Zhang, H., Zhao, J., Wang, X. and Xuan, Y. (2022). Low-voltage distribution grid topology identification with latent tree model. *IEEE Trans. Smart Grid* **13** 2158–2169.

Zhou, C., Wang, X. and Guo, J. (2020). Learning mixed latent tree models. *J. Mach. Learn. Res.* **21** 1–35.

Zwiernik, P. (2018). Latent tree models. In *Handbook of Graphical Models* (M. Drton, M.H. Maathuis, S.L. Lauritzen and M.J. Wainwright, eds.) 283–306. CRC Press.

Bernoulli **32**(4), 2026, 2660–2685
https://doi.org/10.3150/25-BEJ1923

Energy distances for statistical inference on infinite dimensional Hilbert spaces without moment conditions

JIAJUN TANG[1,a] and HOLGER DETTE[2,b]

[1] *Department of Statistics, Harvard University, Cambridge, USA,* [a]*jiajuntang@fas.harvard.edu*
[2] *Fakultät für Mathematik, Ruhr-Universität Bochum, Bochum, Germany,* [b]*holger.dette@rub.de*

For statistical inference on an infinite-dimensional Hilbert space $\mathcal{H}$ with no moment conditions we introduce a new class of energy distances on the space of probability measures on $\mathcal{H}$. The proposed distances consist of the integrated squared modulus of the corresponding difference of the characteristic functionals with respect to a reference probability measure on the Hilbert space. Necessary and sufficient conditions are established for the reference probability measure to be *characteristic*, the property that guarantees that the distance defines a metric on the space of probability measures on $\mathcal{H}$. We also use these results to define new distance covariances, which can be used to measure the dependence between the marginals of a two dimensional distribution of $\mathcal{H}^2$ without existing moments. On the basis of the new distances we develop statistical inference for Hilbert space valued data, which does require any moment assumptions. As a consequence, our methods are robust with respect to heavy tails in finite dimensional data. In particular, we consider the problem of comparing the distributions of two samples and the problem of testing for independence and construct new minimax optimal tests for the corresponding hypotheses. We also develop aggregated (with respect to the reference measure) procedures for power enhancement and investigate the finite-sample properties by means of a simulation study.

Keywords: Characteristic functional; functional data; Gaussian measure; Hilbert-Schmidt independence criterion; Laplace measure; permutation test; probability measure on Hilbert space; minimax optimal testing; reproducing kernel Hilbert space

References

Albert, M., Laurent, B., Marrel, A. and Meynaoui, A. (2022). Adaptive test of independence based on HSIC measures. *Ann. Statist.* **50** 858–879.
Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.
Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606.
Berg, C., Christensen, J.P.R. and Ressel, P. (1984). *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*. New York: Springer.
Berger, M., Hermann, P. and Holzmann, H. (2025). From dense to sparse design: Optimal rates under the supremum norm for estimating the mean function in functional data analysis. *Bernoulli*. **31** 1858 – 1888.
Betken, A., Dehling, H. and Kroll, M. (2021). A bootstrap test for independence of time series based on the distance covariance. arXiv:2112.14091.
Bhar, S. and Dhar, S.S. (2023). Testing independence of infinite dimensional random elements: a sup-norm approach. arXiv:2301.00375.
Bhattacharya, R. and Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. *Ann. Statist.* **31** 1–29.
Bhattacharya, R. and Patrangenaru, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds II. *Ann. Statist.* **33** 1225–1259.
Bobrowski, O., Mukherjee, S. and Taylor, J.E. (2017). Topological consistency via kernel estimation. *Bernoulli* **23** 288–328.

Boente, G. and Salibian-Barrera, M. (2021). Robust functional principal components for sparse longitudinal data. *Metron* **79** 159–188.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Springer, Science & Business Media.

Buchet, M., Chazal, F., Oudot, S.Y. and Sheehy, D.R. (2016). Efficient and robust persistent homology for measures. *Comput. Geom.* **58** 70–96.

Bugni, F.A. and Horowitz, J.L. (2021). Permutation tests for equality of distributions of functional data. *J. Appl. Econometrics* **36** 861–877.

Carroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83** 1184–1186.

Chen, H., Jiang, G. and Yoshihira, K. (2006). Robust nonlinear dimensionality reduction for manifold learning. In *18th Int. Conf. Pattern Recognit* **2** 447–450.

Da Prato, G. (2006). *An Introduction to Infinite-Dimensional Analysis*. Springer, Science & Business Media.

Dehling, H., Matsui, M., Mikosch, T., Samorodnitsky, G. and Tafakori, L. (2020). Distance covariance for discretized stochastic processes. *Bernoulli* **26** 2758–2789.

Delaigle, A. (2008). An alternative view of the deconvolution problem. *Statist. Sinica* **18** 1025–1045.

Dette, H. and Tang, J. (2026). Supplement to "Energy distances for statistical inference on infinite dimensional Hilbert spaces without moment conditions." https://doi.org/10.3150/25-BEJ1923SUPP

Dryden, I.L. and Mardia, K.V. (2016). *Statistical Shape Analysis: With Applications in R*. Wiley.

Fan, Y., Lafaye de Micheaux, P., Penev, S. and Salopek, D. (2017). Multivariate nonparametric test of independence. *J. Multivariate Anal.* **153** 189–210.

Fasy, B., Lecci, F. and Wasserman, L. (2018). Robust topological inference: Distance to a measure and kernel distance. *J. Mach. Learn. Res.* **18** 1–40.

Feng, S., Ren, W., Han, M. and Chen, Y.W. (2019). Robust manifold broad learning system for large-scale noisy chaotic time series prediction: A perturbation perspective. *Neural Netw.* **117** 179–190.

Fernández, V.A., Jiménez Gamero, M.D. and Muñoz García, J. (2008). A test for the two-sample problem based on empirical characteristic functions. *Comput. Statist. Data Anal.* **52** 3730–3748.

Fromont, M., Laurent, B. and Reynaud-Bouret, P. (2013). The two-sample problem for Poisson processes: Adaptive tests with a nonasymptotic wild bootstrap approach. *Ann. Statist.* **41** 1431–1461.

Fukumizu, K., Gretton, A., Sun, X. and Schölkopf, B. (2007). Kernel measures of conditional dependence. *Adv. Neural Inf. Process. Syst.* **20**.

Gardner, R. (2002). The Brunn–Minkowski inequality. *Bull. Amer. Math. Soc.* **39** 355–405.

Gervini, D. (2009). Detecting and handling outlying trajectories in irregularly sampled functional datasets. *Ann. Appl. Stat.* **3** 1758–1775.

Gretton, A., Bousquet, O., Smola, A. and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Int. Conf. Algorithmic Learn. Theory* 63–77.

Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2007). A kernel statistical test of independence. *Adv. Neural Inf. Process. Syst.* **20**.

Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773.

Henze, N. and Jiménez-Gamero, M.D. (2021). A test for Gaussianity in Hilbert spaces via the empirical characteristic functional. *Scand. J. Stat.* **48** 406–428.

Henze, N. and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *J. Multivariate Anal.* **62** 1–23.

Hlavka, Z., Hlubinka, D. and Konasova, K. (2022). Functional ANOVA based on empirical characteristic functionals. *J. Multivariate Anal.* **189** 8.

Horvath, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer, Science & Business Media.

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley.

Kotz, S., Kozubowski, T. and Podgorski, K. (2001). *The Laplace Distribution and Generalizations*. Springer, Science & Business Media.

Krzysko, M., Smaga, L. and Kokoszka, P. (2022). Marginal distance and Hilbert-Schmidt covariances-based independence tests for multivariate functional data. *J. Artificial Intelligence Res.* **73** 1355–1384.
Kukush, A. (2020). *Gaussian Measures in Hilbert Space: Construction and Properties*. Wiley.
Kuo, H.H. (1975). *Gaussian Measures in Banach Spaces*. *Lecture Notes in Math.* **463**. Berlin: Springer.
Lee, C.E., Zhang, X. and Shao, X. (2020). Testing conditional mean independence for functional data. *Biometrika* **107** 331–346.
Lyons, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305.
Ma, Y. and Fu, Y.E. (2011). *Manifold Learning Theory and Applications*. CRC Press.
Meintanis, S.G., Huskova, M. and Hlavka, Z. (2022). Fourier-type tests of mutual independence between functional time series. *J. Multivariate Anal.* **189** 3.
Meintanis, S. and Iliopoulos, G. (2008). Fourier methods for testing multivariate independence. *Comput. Statist. Data Anal.* **52** 1884–1895.
Meintanis, S.G., Swanepoel, J. and Allison, J. (2014). The probability weighted characteristic function and goodness-of-fit testing. *J. Statist. Plann. Inference* **146** 122–132.
Meister, A. (2009). *Deconvolution Problems in Nonparametric Statistics*. *Lecture Notes in Stat.* Springer.
Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Philos. Trans. R. Soc. Lond.* **209** 415–446.
Miao, R., Zhang, X. and Wong, R.K. (2023). A wavelet-based independence test for functional data with an application to MEG functional connectivity. *J. Amer. Statist. Assoc.* **118** 1876–1889.
Patilea, V., Sanchez-Sellero, C. and Saumard, M. (2016). Testing the predictor effect on a functional response. *J. Amer. Statist. Assoc.* **111** 1684–1695.
Prokhorov, Y.V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157–214.
Rachev, S.T., Klebanov, L.B., Stoyanov, S.V. and Fabozzi, F. (2013). *The Methods of Distances in the Theory of Probability and Statistics*. Heidelberg: Springer.
Schmid, J., Kim, J. and Magnenat-Thalmann, N. (2011). Robust statistical shape models for MRI bone segmentation in presence of small field of view. *Med. Image Anal.* **15** 155–168.
Schrab, A., Kim, I., Guedj, B. and Gretton, A. (2022). Efficient aggregated kernel tests using incomplete U-statistics. *NeurIPS* **35**.
Shao, L. and Yao, F. (2023). Robust functional data analysis for discretely observed data. arXiv:2305.16236.
Smola, A., Gretton, A., Song, L. and Schölkopf, B. (2007). A Hilbert space embedding for distributions. In *Int. Conf. Algorithmic Learn. Theory* 13–31.
Sriperumbudur, B.K., Gretton, A., Fukumizu, K., Schölkopf, B. and Lanckriet, G.R. (2010). Hilbert space embeddings and metrics on probability measures. *J. Mach. Learn. Res.* **11** 1517–1561.
Szekely, G.J. and Rizzo, M.L. (2013). Energy statistics: A class of statistics based on distances. *J. Statist. Plann. Inference* **143** 1249–1272.
Szekely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794.
Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9** 2579–2605.
Wang, J.L., Chiou, J.M. and Müller, H.G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.
Wasserman, L. (2018). Topological data analysis. *Annu. Rev. Stat. Appl.* **5** 501–532.
Wynne, G. and Duncan, A.B. (2022). A kernel two-sample test for functional data. *J. Mach. Learn. Res.* **23** 3159–3209.
Zhu, C. and Wang, J.-L. (2023). Testing homogeneity: The trouble with sparse functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 705–731.

Bernoulli **32**(4), 2026, 2686–2710
https://doi.org/10.3150/25-BEJ1942

Statistical inference for Bures-Wasserstein flows

LEONARDO V. SANTORO[a] and VICTOR M. PANARETOS[b]

Institut de Mathématiques, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland,
[a]*leonardo.santoro@epfl.ch,* [b]*victor.panaretos@epfl.ch*

We develop a statistical framework for conducting inference on collections of time-varying covariance operators (covariance flows) over a general, possibly infinite dimensional, Hilbert space, and fully develop asymptotic theory based on interpretable and verifiable assumptions. We model the intrinsically non-linear structure of covariances by means of the Bures-Wasserstein metric geometry. We make use of the Riemannian-like structure induced by this metric to define a notion of mean and covariance of a random flow, and develop an associated Karhunen-Loève expansion. We then treat theproblem of estimation and construction of functional principal components from a finite collection of covariance flows, observed fully or irregularly. Our theoretical results are motivated by modern problems in functional data analysis, where one observes operator-valued random processes – for instance when analysing dynamic functional connectivity and fMRI data, or when analysing multiple functional time series in the frequency domain. Nevertheless, our framework is also novel in the finite-dimensions (matrix case), and we demonstrate what simplifications can be afforded then. We illustrate our methodology by means of simulations and data analyses.

Keywords: Bures-Wasserstein space; functional data analysis; Karhunen-Loève expansion

References

Ambrosio, L. and Gigli, N. (2008). Construction of the parallel transport in the Wasserstein space. *Methods Appl. Anal.* **15** 1–30.

Ambrosio, L., Gigli, N. and Savaré, G. (2004). Gradient flows with metric and differentiable structures, and applications to the Wasserstein space. *Rend. Accad. Naz. Lincei Mat. Appl.* **15** 327–343.

Arsigny, V., Fillard, P., Pennec, X. and Ayache, N. (2006). Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **56** 411–421.

Arsigny, V., Fillard, P., Pennec, X. and Ayache, N. (2007). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **29** 328–347.

Bhatia, R., Jain, T. and Lim, Y. (2019). On the Bures–Wasserstein distance between positive definite matrices. *Exp. Math.* **37** 165–191.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications 149*. Springer Science & Business Media.

Chau, J. and von Sachs, R. (2021). Intrinsic wavelet regression for curves of Hermitian positive definite matrices. *J. Amer. Statist. Assoc.* **116** 819–832.

Cuesta-Albertos, J., Matrán-Bea, C. and Tuero-Diaz, A. (1996). On lower bounds for the L 2-Wasserstein metric in a Hilbert space. *J. Theor. Probab.* **9** 263–283.

Dai, X., Lin, Z. and Müller, H.-G. (2021). Modeling sparse longitudinal data on Riemannian manifolds. *Biometrics* **77** 1328–1341.

Dai, X. and Müller, H.-G. (2018). Principal component analysis for functional data on Riemannian manifolds and spheres. *Ann. Statist.* **46** 3334–3361.

Dai, M., Zhang, Z. and Srivastava, A. (2019). Analyzing dynamical brain functional connectivity as trajectories on space of covariance matrices. *IEEE Trans. Med. Imag.* **39** 611–620.

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.

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

1350-7265

Gouic, T.L., Paris, Q., Rigollet, P. and Stromme, A.J. (2023). Fast convergence of empirical barycenters in Alexandrov spaces and the Wasserstein space. *J. Eur. Math. Soc. (JEMS)* **25** 2229–2250.

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators* **997**. John Wiley & Sons.

Huang, S.-G., Samdin, S.B., Ting, C.-M., Ombao, H. and Chung, M.K. (2020). Statistical model for dynamically-changing correlation matrices with application to brain connectivity. *J. Neurosci. Methods* **331** 108480.

Kantorovich, L.V. (1958). On the translocation of masses. *Manag. Sci.* **5** 1–4.

Kim, D., Lee, Y.K. and Park, B.U. (2020). Principal component analysis for Hilbertian functional data. *Commun. Stat. Appl. Methods* **27** 149–161.

Krebs, J., Rademacher, D. and von Sachs, R. (2022). Statistical inference for intrinsic wavelet estimators of SPD matrices in a log-Euclidean manifold. Preprint.

Kroshnin, A., Spokoiny, V. and Suvorikova, A. (2021). Statistical inference for Bures–Wasserstein barycenters. *Ann. Appl. Probab.* **31** 1264–1298.

Lin, Z. and Yao, F. (2019). Intrinsic Riemannian functional data analysis. *Ann. Statist.* **47** 3533–3577.

Masarotto, V., Panaretos, V.M. and Zemel, Y. (2019). Procrustes metrics on covariance operators and optimal transportation of Gaussian processes. *Sankhya A* **81** 172–213.

Masarotto, V., Panaretos, V.M. and Zemel, Y. (2024). Transportation-based functional ANOVA and PCA for covariance operators. *Electron. J. Stat.* **18** 1887–1916.

McCann, R.J. (1997). A convexity principle for interacting gases. *Adv. Math.* **128** 153–179.

Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Hist. Acad. R. Sci. Mém. Math. Phys.* 666–704.

Ocaña, F.A., Aguilera, A.M. and Valderrama, M.J. (1999). Functional principal components analysis by choice of norm. *J. Multivariate Anal.* **71** 262–276.

Panaretos, V.M., Kraus, D. and Maddocks, J.H. (2010). Second-order comparison of Gaussian random functions and the geometry of DNA minicircles. *J. Amer. Statist. Assoc.* **105** 670–682.

Panaretos, V.M. and Tavakoli, S. (2013a). Fourier analysis of stationary time series in function space. *Ann. Statist.* **41** 568–603.

Panaretos, V.M. and Tavakoli, S. (2013b). Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series. *Stoch. Process. Appl.* **123** 2779–2807.

Panaretos, V.M. and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** 405–431.

Petersen, A. and Müller, H.-G. (2019). Fréchet regression for random objects with Euclidean predictors. *Ann. Statist.* **47** 691–719.

Pigoli, D., Aston, J.A., Dryden, I.L. and Secchi, P. (2014). Distances and inference for covariance operators. *Biometrika* **101** 409–422.

Santoro, L.V. and Panaretos, V.M. (2025). Large sample theory for Bures–Wasserstein barycenters. *Ann. Appl. Probab.* To appear.

Santoro, L.V. and Panaretos, V.M. (2026). Supplement to "Statistical inference for Bures-Wasserstein flows." https://doi.org/10.3150/25-BEJ1942SUPP

Santoro, L.V., Waghmare, K.G. and Panaretos, V.M. (2024). A Karhunen–Loève theorem for random flows in Hilbert spaces. *Electron. Commun. Probab.* **29** 1–12.

Schwartzman, A. (2006). Random Ellipsoids and False Discovery Rates: Statistics for Diffusion Tensor Imaging Data. Stanford University.

Shao, L., Lin, Z. and Yao, F. (2022). Intrinsic Riemannian functional data analysis for sparse longitudinal observations. *Ann. Statist.* **50** 1696–1721.

Takatsu, A. (2011). Wasserstein geometry of Gaussian measures. *Osaka J. Math.* **48** 1005–1026.

Takesaki, M. et al. (2003). *Theory of Operator Algebras I* **125**. Springer.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer.

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.

Waghmare, K.G. and Panaretos, V.M. (2024). Extreme Points of Spectrahedra. arXiv preprint. Available at arXiv: 2410.14889.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590.

Yaqing Chen, Z.L. and Müller, H.-G. (2023). Wasserstein regression. *J. Amer. Statist. Assoc.* **118** 869–882.

Zhang, Z., Su, J., Klassen, E., Le, H. and Srivastava, A. (2018). Rate-invariant analysis of covariance trajectories. *J. Math. Imaging Vision* **60** 1306–1323.

Zhou, H., Lin, Z. and Yao, F. (2023). Intrinsic correlation analysis for Wasserstein functional data. *Statist. Sinica*. Preprint No. SS-2023-0147.

Bernoulli **32**(4), 2026, 2711–2740
https://doi.org/10.3150/25-BEJ1963

Discrete Feynman-Kac approximation for parabolic Anderson model using random walks

PANQIU XIA[1,a] and JIAYU ZHENG[2,b]

[1]*School of Mathematics, Cardiff University, Abacws, Senghennydd Road, Cathays, Cardiff, Wales, UK, CF24 4AG,* [a]*xiap@cardiff.ac.uk*
[2]*Faculty of Computational Mathematics and Cybernetics, Shenzhen MSU-BIT University, Shenzhen 518172, China,* [b]*jyzheng@smbu.edu.cn*

In this paper, we introduce a natively positive approximation method based on the Feynman-Kac representation using random walks, to approximate the solution to the one-dimensional parabolic Anderson model of Skorokhod type, with either a flat or a Dirac delta initial condition. Assuming the driving noise is a fractional Brownian sheet with Hurst parameters $H \geq \frac{1}{2}$ and $H_* \geq \frac{1}{2}$ in time and space, respectively, we also provide an error analysis of the proposed method. The error in $L^p(\Omega)$ norm is of order

$$O\left(h^{\frac{1}{2}[(2H+H_*-1)\wedge 1]-\epsilon}\right),$$

where $h > 0$ is the step size in time (resp. $\sqrt{h}$ in space), and $\epsilon > 0$ can be chosen arbitrarily small. This error order matches the Hölder continuity of the solution in time with a correction order ϵ, making it 'almost' optimal. Furthermore, these results provide a quantitative framework for convergence of the partition function of directed polymers in Gaussian environments to the parabolic Anderson model.

Keywords: Directed polymers in random environments; Feynman-Kac formula; fractional Brownian sheet; parabolic Anderson model; Skorokhod integral

References

Alberts, T., Khanin, K. and Jeremy Quastel, J. (2014). The intermediate disorder regime for directed polymers in dimension $1+1$. *Ann. Probab.* **42** 1212–1256.

Balan, R.M. and Chen, L. (2018). Parabolic Anderson model with space-time homogeneous Gaussian noise and rough initial condition. *J. Theoret. Probab.* **31** 2216–2265.

Balan, R.M., Quer-Sardanyons, L. and Song, J. (2019). Hölder continuity for the parabolic Anderson model with space-time homogeneous Gaussian noise. *Acta Math. Sci. Ser. B Engl. Ed.* **39** 717–730.

Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.

Bertini, L. and Cancrini, N. (1995). The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Stat. Phys.* **78** 1377–1401.

Bréhier, C.E., Cohen, D. and Ulander, J. (2024a). Positivity-preserving schemes for some nonlinear stochastic PDEs. In *Sixteenth International Conference Zaragoza-Pau on Mathematics and Its Applications*. *Monogr. Mat. García Galdeano* **43** 31–40. Zaragoza: Prensas Univ. Zaragoza.

Bréhier, C.E., Cohen, D. and Ulander, J. (2024b). Analysis of a positivity-preserving splitting scheme for some semilinear stochastic heat equations. *ESAIM: Math. Model. Numer. Anal.* **58** 1317–1346.

Caravenna, F., Sun, R. and Zygouras, N. (2016). Polynomial chaos and scaling limits of disordered systems. *J. Eur. Math. Soc.* **19** 1–65.

Chen, X. (2015). Precise intermittency for the parabolic Anderson equation with an $(1+1)$-dimensional time-space white noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1486–1499.

1350-7265

Chen, X. (2019). Parabolic Anderson model with rough or critical Gaussian noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 941–976.

Chen, L. and Dalang, R.C. (2015). Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions. *Ann. Probab.* **43** 3006–3051.

Chen, Y. and Gao, F. (2023). Scaling limits of directed polymers in spatial-correlated environment. *Electron. J. Probab.* **28** 1–57.

Chen, L. and Huang, J. (2019). Comparison principle for stochastic heat equation on $\mathbb{R}$. *Ann. Probab.* **47** 989–1035.

Comets, F. (2017). *Directed Polymers in Random Environments*. *Lecture Notes in Mathematics* **2175**. Cham: Springer.

Gordon, R.D. (1941). Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Ann. Math. Stat.* **12** 364–366.

Gubernatis, J., Kawashima, N. and Werner, P. (2016). *Quantum Monte Carlo Methods: Algorithms for Lattice Models*. Cambridge: Cambridge University Press.

Gubinelli, M. and Perkowski, N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269.

Guo, Y., Song, J. and Song, X. (2024). Stochastic fractional diffusion equations with Gaussian noise rough in space. *Bernoulli* **30** 1774–1799.

Gyöngy, I. (1998). Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise I. *Potential Anal.* **9** 1–25.

Gyöngy, I. (1999). Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise II. *Potential Anal.* **11** 1–37.

Herrell, R., Song, R., Wu, D. and Xiao, Y. (2020). Sharp space-time regularity of the solution to stochastic heat equation driven by fractional-colored noise. *Stoch. Anal. Appl.* **38** 747–768.

Hu, Y. (2017). *Analysis on Gaussian Spaces*. Hackensack, NJ: World Scientific.

Hu, Y., Lu, F. and Nualart, D. (2012). Feynman-Kac formula for the heat equation driven by fractional noise with Hurst parameter $H < \frac{1}{2}$. *Ann. Probab.* **40** 1041–1068.

Hu, Y., Nualart, D. and Song, J. (2011). Feynman-Kac formula for heat equation driven by fractional white noise. *Ann. Probab.* **39** 291–326.

Hu, Y., Huang, J., Nualart, D. and Tindel, S. (2015). Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.* **20** Paper No. 55, 50.

Hu, Y., Huang, J., Lê, K., Nualart, D. and Tindel, S. (2018). Parabolic Anderson model with rough dependence in space, computation and combinatorics in dynamics, stochastics and control. *Abel Symp.* **13** 477–498.

Huang, J., Lê, K. and Nualart, D. (2017). Large time asymptotics for the parabolic Anderson model driven by spatially correlated noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1305–1340.

Joseph, M. (2018). An invariance principle for the stochastic heat equation. *Stoch. Partial Differ. Equ. Anal. Comput.* **6** 690–745.

Kilbas, A.A., Srivastava, H.M. and Trujillo, J. (2006). *Theory and Applications of Fractional Differential Equations*. *North-Holland Mathematics Studies* **204**. Amsterdam: Elsevier.

Liggett, T.M. (1968). An invariance principle for conditioned sums of independent random variables. *J. Math. Mech.* **18** 559–570.

Luo, W. (2006). Wiener chaos expansion and numerical solutions of stochastic partial differential equations. Ann Arbor: ProQuest LLC, Thesis (Ph. D.)–California Institute of Technology.

Mémin, J., Mishura, Y. and Valkeila, E. (2001). Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion. *Statist. Probab. Lett.* **51** 197–206.

Mossel, E., O'Donnell, R. and Oleszkiewicz, K. (2010). Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.* **171** 295–341.

Mueller, C. (1991). On the support of solutions to the heat equation with noise. *Stoch. Stoch. Rep.* **37** 225–245.

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, second ed. *Probability and Its Applications (New York)*. Berlin: Springer-Verlag.

Petrov, V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]* **82**. New York: Springer-Verlag. Translated from the Russian by A. A. Brown.

Rang, G. (2020). From directed polymers in spatial-correlated environment to stochastic heat equations driven by fractional noise in 1 + 1 dimensions. *Stoch. Process. Appl.* **130** 3408–3444.

Rang, G., Song, J. and Wang, M. (2024). Scaling limit of a long-range random walk in time-correlated random environment. *Electron. J. Probab.* **29** 1–39.

Sanz-Solé, M. and Sarrà, M. (2000). Path properties of a class of Gaussian processes with applications to spde's. In *Stochastic Processes, Physics and Geometry: New Interplays, I (Leipzig, 1999)*. *CMS Conf. Proc.* **28** 303–316. Providence, RI: Amer. Math. Soc.

Villani, C. (2009). *Optimal Transport: Old and New*. *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer-Verlag.

Bernoulli **32**(4), 2026, 2741–2767
https://doi.org/10.3150/25-BEJ1964

A closed-form transition density expansion for elliptic and hypo-elliptic SDEs

YUGA IGUCHI[1,a] and ALEXANDROS BESKOS[2,b]

[1]*School of Mathematical Sciences, Lancaster University, Lancaster, UK,* [a]*y.iguchi@lancaster.ac.uk*
[2]*Department of Statistical Science, University College London, London, UK,* [b]*a.beskos@ucl.ac.uk*

We introduce a closed-form expansion for the transition density of elliptic and hypo-elliptic multivariate Stochastic Differential Equations (SDEs), over a period $\Delta \in (0,1)$, in terms of powers of $\Delta^{j/2}$, $j \geq 0$. Our methodology provides approximations of the transition density, easily evaluated via any software that performs symbolic calculations. A major part of the paper is devoted to an analytical control of the remainder in our expansion for fixed $\Delta \in (0,1)$. The obtained error bounds validate theoretically the methodology, by characterising the size of the distance from the true value. For elliptic SDEs, closed-form expansions are available, with some works identifying the size of the error for fixed Δ, as per our contribution. Our methodology allows for a uniform treatment of elliptic and hypo-elliptic SDEs, when earlier works are intrinsically restricted to an elliptic setting. We show numerical applications highlighting the effectiveness of our method, by carrying out parameter inference for hypo-elliptic SDEs that do not satisfy stated conditions. The latter are sufficient for controlling the remainder terms, but the closed-form expansion itself is applicable in general settings.

Keywords: Data augmentation; hypo-elliptic diffusion; transition density; stochastic differential equation

References

Ableidinger, M., Buckwar, E. and Hinterleitner, H. (2017). A stochastic version of the Jansen and Rit neural mass model: Analysis and numerics. *J. Math. Neurosci.* **7** 1–35.

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica* **70** 223–262.

Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.* **36** 906–937.

Andrieu, C., Doucet, A. and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 269–342.

Barilari, D. and Paoli, E. (2017). Curvature terms in small time heat kernel expansion for a model class of hypoelliptic Hörmander operators. *Nonlinear Anal.* **164** 118–134.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. Available at arXiv:1701.02434.

Chang, J. and Chen, S.X. (2011). On the approximate maximum likelihood estimation for diffusion processes. *Ann. Statist.* **39** 2820–2851.

Choi, S. (2015). Explicit form of approximate transition probability density functions of diffusion processes. *J. Econometrics* **187** 57–73.

Davie, A.M. (2022). Approximation of SDE solutions using local asymptotic expansions. *Stoch. Partial Differ. Equ.* **10** 720–734.

DeVille, L., Vanden-Eijnden, E. and Muratov, C. (2005). Two distinct mechanisms of coherence in randomly perturbed dynamical systems. *Phys. Rev. E* **72** 031105.

Ditlevsen, S. and Samson, A. (2019). Hypoelliptic diffusions: Filtering and inference from complete and partial observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 361–384.

Gloter, A. and Yoshida, N. (2021). Adaptive estimation for degenerate diffusion processes. *Electron. J. Stat.* **15** 1424–1472.

Habermann, K. (2019). Small-time fluctuations for the bridge in a model class of hypoelliptic diffusions of weak Hörmander type. *Electron. J. Probab.* **24**. https://doi.org/10.1214/19-ejp274

1350-7265

Hoffman, M., Gelman, A. et al. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15** 1593–1623.

Iguchi, Y. and Beskos, A. (2025). Parameter inference for hypo-elliptic diffusions under a weak design condition. *Electron. J. Stat.* **19** 1337–1369.

Iguchi, Y. and Beskos, A. (2026). Supplement to "A closed-form transition density expansion for elliptic and hypo-elliptic SDEs." https://doi.org/10.3150/25-BEJ1964SUPP

Iguchi, Y., Beskos, A. and Graham, M. (2024). Parameter inference for degenerate diffusion processes. *Stoch. Process. Appl.* **104384**.

Iguchi, Y., Beskos, A. and Graham, M. (2025). Parameter estimation with increased precision for elliptic and hypo-elliptic diffusions. *Bernoulli* **31** 333–358.

Iguchi, Y. and Yamada, T. (2021). Operator splitting around Euler-Maruyama scheme and high order discretization of heat kernels. *ESAIM: Math. Model. Numer. Anal.* **55** S323–S367.

Jimenez, J.C., Mora, C. and Selva, M. (2017). A weak local linearization scheme for stochastic differential equations with multiplicative noise. *J. Comput. Appl. Math.* **313** 202–217. https://doi.org/10.1016/j.cam.2016.09.013

Kloeden, P. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer.

Lee, Y.D., Song, S. and Lee, E.K. (2014). The delta expansion for the transition density of diffusion models. *J. Econometrics* **178** 694–705.

Li, C. (2013). Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.* **41** 1350–1380.

Li, C. and Chen, D. (2016). Estimating jump–diffusions using closed-form likelihood expansions. *J. Econometrics* **195** 51–70.

Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stoch. Process. Appl.* **101** 185–232. https://doi.org/10.1016/s0304-4149(02)00150-3

Melnykova, A. (2020). Parametric inference for hypoelliptic ergodic diffusions with full observations. *Stat. Inference Stoch. Process.* **23** 595–635.

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*. Berlin: Springer.

Papaspiliopoulos, O., Roberts, G. and Sköld, M. (2007). A general framework for the parametrization of hierarchical models. *Statist. Sci.* **22** 59–73.

Papaspiliopoulos, O., Roberts, G.O. and Stramer, O. (2013). Data augmentation for diffusions. *J. Comput. Graph. Statist.* **22** 665–688. https://doi.org/10.1080/10618600.2013.783484

Pavliotis, G. (2014). *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations 60*. Springer.

Pigato, P. (2022). Density estimates and short-time asymptotics for a hypoelliptic diffusion process. *Stoch. Process. Appl.* **145** 117–142.

Pilipovic, P., Samson, A. and Ditlevsen, S. (2025). Strang splitting for parametric inference in second-order stochastic differential equations. *Stoch. Process. Appl.* **187** 104650.

Roberts, G. and Stramer, O. (2001). On inference for partially observed nonlinear diffusion models using the Metropolis–Hastings algorithm. *Biometrika* **88** 603–621.

Samson, A., Tamborrino, M. and Tubikanec, I. (2025). Inference for the stochastic FitzHugh-Nagumo model from real action potential data via approximate Bayesian computation. *Comput. Statist. Data Anal.* **204** 108095. https://doi.org/10.1016/j.csda.2024.108095

Stramer, O., Bognar, M. and Schneider, P. (2010). Bayesian inference for discretely sampled Markov processes with closed-form likelihood expansions. *J. Financ. Econom.* **8** 450–480.

Watanabe, S. (1987). Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Ann. Probab.* **15** 1–39.

Yang, N., Chen, N. and Wan, X. (2019). A new delta expansion for multivariate diffusions via the Itô-Taylor expansion. *J. Econometrics* **209** 256–288.

Yoshida, N. (1992). Asymptotic expansions of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe. *Probab. Theory Related Fields* **92** 275–311.

Bernoulli **32**(4), 2026, 2768–2793
https://doi.org/10.3150/25-BEJ1967

Cramér-type moderate deviations of drift estimation in the stochastic heat equation

HUI JIANG[a] and JINGYING ZHOU[b]

School of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, China,
[a]*huijiang@nuaa.edu.cn,* [b]*zhoujy0106@nuaa.edu.cn*

For the stochastic fractional heat equation driven by additive noise, we study the asymptotic properties of the maximum likelihood estimator (MLE) of the drift coefficient. Using parameter-dependent change of measure methods along with asymptotic analysis techniques, we establish (self-normalized) Cramér-type moderate deviations, (non-)uniform Berry-Esseen bounds, moderate deviations and precise large deviations. The study covers three asymptotic regimes: large time asymptotics, increasing number of Fourier modes, large time asymptotics and increasing number of Fourier modes. Our main results reveal how the asymptotic behaviors depend on the time horizon, the number of Fourier modes, the spatial dimension, and the fractional order.

Keywords: Berry-Esseen bound; change of measure method; Cramér-type moderate deviations; maximum likelihood estimator; stochastic heat equation

References

Bercu, B., Coutin, L. and Savy, N. (2012). Sharp large deviations for the non-stationary Ornstein-Uhlenbeck process. *Stoch. Process. Appl.* **122** 3393–3424.

Bercu, B. and Richou, A. (2015). Large deviations for the Ornstein-Uhlenbeck process with shift. *Adv. Appl. Probab.* **47** 880–901.

Bercu, B. and Rouault, A. (2002). Sharp large deviations for the Ornstein-Uhlenbeck process. *Theory Probab. Appl.* **46** 1–19.

Bishwal, J.P.N. (2000). Sharp Berry-Esseen bound for the maximum likelihood estimator in the Ornstein-Uhlenbeck process. *Sankhyā* **62** 1–10.

Blade, M.F., Es-Sebaiy, K. and Tudor, C.A. (2020). Ergodicity and drift parameter estimation for infinite-dimensional fractional Ornstein-Uhlenbeck process of the second kind. *Appl. Math. Optim.* **81** 785–814.

Cialenco, I. (2018). Statistical inference for SPDEs: An overview. *Stat. Inference Stoch. Process.* **21** 309–329.

Cialenco, I., Delgado-Vences, F. and Kim, H.J. (2020). Drift estimation for discretely sampled SPDEs. *Stoch. Partial Differ. Equ. Anal. Comput.* **8** 895–920.

Cialenco, I., Gong, R. and Huang, Y. (2018). Trajectory fitting estimators for SPDEs driven by additive noise. *Stat. Inference Stoch. Process.* **21** 1–19.

Cialenco, I. and Xu, L. (2015). Hypothesis testing for stochastic PDEs driven by additive noise. *Stoch. Process. Appl.* **125** 819–866.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Berlin: Springer-Verlag.

Douissi, S. and Alshahrani, F. (2025). Parameter estimation for fractional stochastic heat equations: Berry-Esseen bounds in CLTs. *Theory Probab. Appl.* **70** 257–281.

Douissi, S., Es-Sebaiy, K. and Viens, F. (2022). Asymptotics of Yule's nonsense correlation for Ornstein-Uhlenbeck paths: A Wiener chaos approach. *Electron. J. Stat.* **16** 3176–3211.

Douissi, S., Es-Sebaiy, K., Alshahrani, F. and Viens, F.G. (2022a). AR(1) processes driven by second-chaos white noise: Berry-Esseen bounds for quadratic variation and parameter estimation. *Stoch. Process. Appl.* **150** 886–918.

Douissi, S., Es-Sebaiy, K., Kerchev, G. and Nourdin, I. (2022b). Berry-Esseen bounds of second moment estimators for Gaussian processes observed at high frequency. *Electron. J. Stat.* **16** 636–670.

1350-7265

Es-Sebaiy, K. and Viens, F.G. (2019). Optimal rates for parameter estimation of stationary Gaussian processes. *Stoch. Process. Appl.* **129** 3018–3054.

Fan, X.Q. and Shao, Q.M. (2024). Cramér's moderate deviations for martingales with applications. *Ann. Inst. Henri Poincaré Probab. Stat.* **60** 2046–2074.

Fan, X.Q. and Shao, Q.M. (2025). Self-normalized Cramér type moderate deviations for martingales and applications. *Bernoulli* **31** 130–161.

Fan, X.Q., Grama, I., Liu, Q.S. and Shao, Q.M. (2019). Self-normalized Cramér type moderate deviations for martingales. *Bernoulli* **25** 2793–2823.

Florens-Landais, D. and Pham, H. (1999). Large deviations in estimation of an Ornstein-Uhlenbeck model. *J. Appl. Probab.* **36** 60–77.

Gao, F.Q. and Jiang, H. (2017). Deviation inequalities for quadratic Wiener functionals and moderate deviations for parameter estimators. *Sci. China Math.* **60** 1181–1196.

Gao, F.Q., Xiong, J. and Zhao, X.Q. (2018). Moderate deviations and nonparametric inference for monotone functions. *Ann. Statist.* **46** 1225–1254.

Hu, Y. and Nualart, D. (2010). Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Statist. Probab. Lett.* **80** 1030–1038.

Hu, Y., Nualart, D. and Zhou, H. (2019). Parameter estimation for fractional Ornstein-Uhlenbeck processes of general Hurst parameter. *Stat. Inference Stoch. Process.* **22** 111–142.

Hu, Y. and Song, J. (2013). Parameter estimation for fractional Ornstein-Uhlenbeck processes with discrete observations. In *Malliavin Calculus and Stochastic Analysis* (F. Viens, J. Feng, Y. Hu and E. Nualart, eds.). *Springer Proceedings in Mathematics and Statistics* **34** 427–442. New York: Springer.

Huebner, M., Lototsky, S.V. and Rozovskiĭ, B.L. (1997). Asymptotic properties of an approximate maximum likelihood estimator for stochastic PDEs. In *Statistics and Control of Stochastic Processes (Moscow, 1995/1996)* (Y.M. Kabanov, B.L. Rozovskiĭ and A.N. Shiryaev, eds.) 139–155. River Edge, NJ: World Science Publishing.

Huebner, M. and Rozovskiĭ, B.L. (1995). On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDEs. *Probab. Theory Related Fields* **103** 143–163.

Janák, J. (2020). Parameter estimation for stochastic partial differential equations of second order. *Appl. Math. Optim.* **82** 353–397.

Jiang, H. (2014). Moderate deviation for parameter estimator in the stochastic parabolic equations with additive fractional Brownian motion. *Stoch. Dyn.* **14** 1450002.

Jiang, H. and Zhou, J.Y. (2023). An exponential nonuniform Berry-Esseen bound of the fractional Ornstein-Uhlenbeck process. *J. Theoret. Probab.* **36** 1037–1058.

Jiang, H. and Zhou, J. (2026). Supplement to "Cramér-type moderate deviations of drift estimation in the stochastic heat equation." https://doi.org/10.3150/25-BEJ1967SUPP

Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. London: Springer-Verlag.

Lototsky, S.V. (2003). Parameter estimation for stochastic parabolic equations: Asymptotic properties of a two-dimensional projection-based estimator. *Stat. Inference Stoch. Process.* **6** 65–87.

Lototsky, S.V. (2009). Statistical inference for stochastic parabolic equations: A spectral approach. *Publ. Mat.* **53** 3–45.

Lototsky, S.V. and Rozovskiĭ, B.L. (1999). Spectral asymptotics of some functionals arising in statistical inference for SPDEs. *Stoch. Process. Appl.* **79** 69–94.

Lototsky, S.V. and Rozovskiĭ, B.L. (2017). *Stochastic Partial Differential Equations*. Cham: Springer.

Maslowski, B. and Pospíšil, J. (2008). Ergodicity and parameter estimates for infinite-dimensional fractional Ornstein-Uhlenbeck process. *Appl. Math. Optim.* **57** 401–429.

Maslowski, B. and Tudor, C.A. (2013). Drift parameter estimation for infinite-dimensional fractional Ornstein-Uhlenbeck process. *Bull. Sci. Math.* **137** 880–901.

Masuda, H. and Yoshida, N. (2013). Edgeworth expansion for the integrated Lévy driven Ornstein-Uhlenbeck process. *Electron. Commun. Probab.* **18** 10.

Petrov, V. (1975). *Sums of Independent Random Variables*. New York-Heidelberg: Springer-Verlag.

Phillips, P.C.B., Magdalinos, T. and Giraitis, L. (2010). Smoothing local-to-moderate unit root theory. *J. Econometrics* **158** 274–279.

Piterbarg, L.I. and Rozovskiĭ, B.L. (1997). On asymptotic problems of parameter estimation in stochastic PDE's: Discrete time sampling. *Math. Methods Statist.* **6** 200–223.

Pospíšil, J. and Tribe, R. (2007). Parameter estimates and exact variations for stochastic heat equations driven by space-time white noise. *Stoch. Anal. Appl.* **25** 593–611.

Shao, Q.M. (1999). A Cramér-type large deviation result for student's t-statistics. *J. Theoret. Probab.* **12** 385–398.

Shao, Q.M. and Zhang, Z.S. (2022). Berry-Esseen bounds for multivariate nonlinear statistics with applications to M-estimators and stochastic gradient descent algorithms. *Bernoulli* **28** 1548–1576.

Shubin, M.A. (2001). *Pseudodifferential Operators and Spectral Theory*. Berlin: Springer-Verlag.

Wang, Q.Y. and Jing, B.Y. (1999). An exponential nonuniform Berry-Esseen bound for self-normalized sums. *Ann. Probab.* **27** 2068–2088.

Xu, J., Zheng, Q. and Mu, J.Y. (2025). Maximum likelihood estimation for small noise multi-scale McKean-Vlasov stochastic differential equations. *Bernoulli* **31** 783–815.

Yoshihide, K. (2002). Edgeworth approximation in the O-U process. *Sankhya, Ser. A* **64** 16–41.

Zani, M. (2002). Large deviations for squared radial Ornstein-Uhlenbeck processes. *Stoch. Process. Appl.* **102** 25–42.

Zhang, Z.S. (2023). Cramér-type moderate deviation of normal approximation for unbounded exchangeable pairs. *Bernoulli* **29** 274–299.

Bernoulli **32**(4), 2026, 2794–2825
https://doi.org/10.3150/26-BEJ1968

On a pinning model in correlated Gaussian random environments

ZI'AN LI[1,2,a], JIAN SONG[1,2,b], RAN WEI[3,d] and HANG ZHANG[1,2,c]

[1]*Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao 266237, China,* [a]*202117106@mail.sdu.edu.cn,* [b]*txjsong@sdu.edu.cn,* [c]*202217118@mail.sdu.edu.cn*
[2]*Frontiers Science Center for Nonlinear Expectations, Ministry of Education, Shandong University, Qingdao 266237, China*
[3]*Department of Financial and Actuarial Mathematics, Xi'an Jiaotong-Liverpool University, 111 Ren'ai Road, Suzhou 215123, China,* [d]*ran.wei@xjtlu.edu.cn*

We consider a pinning model in correlated Gaussian random environments. For the model that is disorder relevant, we study its intermediate disorder regime and show that the rescaled partition functions converge to a non-trivial continuum limit in the Skorohod setting and in the Stratonovich setting, respectively. Our results partially confirm the Weinrib–Halperin prediction for disorder relevance/irrelevance.

Keywords: Correlated Gaussian random environments; pinning model; scaling limit; Weinrib–Halperin prediction

References

Alberts, T., Khanin, K. and Quastel, J. (2014). The intermediate disorder regime for directed polymers in dimension 1 + 1. *Ann. Probab.* **42** 1212–1256.
Berger, Q. (2013). Comments on the influence of disorder for pinning model in correlated Gaussian environment. *ALEA Lat. Amer. J. Probab. Math. Stat.* **10** 953–977.
Berger, Q. (2014). Pinning model in random correlated environment: Appearance of an infinite disorder regime. *J. Stat. Phys.* **155** 544–570.
Berger, Q. and Lacoin, H. (2018). Pinning on a defect line: Characterization of marginal disorder relevance and sharp asymptotics for the critical point shift. *J. Inst. Math. Jussieu* **17** 305–346.
Berger, Q. and Poisat, J. (2015). On the critical curves of the pinning and copolymer models in correlated Gaussian environment. *Electron. J. Probab.* **20** 1–35.
Billingsley, P. (1999). *Convergence of Probability Measures*. John Wiley & Sons.
Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. Cambridge university press.
Caravenna, F., Sun, R. and Zygouras, N. (2016). The continuum disordered pinning model. *Probab. Theory Related Fields* **164** 17–59.
Caravenna, F., Sun, R. and Zygouras, N. (2017a). Polynomial chaos and scaling limits of disordered systems. *J. Eur. Math. Soc.* **19** 1–65.
Caravenna, F., Sun, R. and Zygouras, N. (2017b). Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* **27** 3050–3112.
Caravenna, F., Toninelli, F.L. and Torri, N. (2017). Universality for the pinning model in the weak coupling regime. *Ann. Probab.* 2154–2209.
Cheliotis, D., Chino, Y. and Poisat, J. (2019). The random pinning model with correlated disorder given by a renewal set. *Ann. Henri Lebesgue* **2** 281–329.
Chen, Y. and Gao, F. (2023). Scaling limits of directed polymers in spatial-correlated environment. *Electron. J. Probab.* **28** 1–57.
Doney, R.A. (1997). One-sided local large deviation and renewal theorems in the case of infinite mean. *Probab. Theory Related Fields* **107** 451–465.
Giacomin, G. (2007). *Random Polymer Models*. Imperial College Press.

Giacomin, G. (2011). *Disorder and Critical Phenomena Through Basic Probability Models: École d'Été de Probabilités de Saint-Flour XL–2010 2025*. Springer Science & Business Media.

Harris, A.B. (1974). Effect of random defects on the critical behaviour of Ising models. *J. Phys. C, Solid State Phys.* **7** 1671.

Hu, Y. (2016). *Analysis on Gaussian Spaces*. World Scientific.

Janson, S. (1997). *Gaussian Hilbert Spaces 129*. Cambridge university press.

Junk, S. (2025). Local limit theorem for directed polymers beyond the L^2-phase. *J. Eur. Math. Soc.*

Li, L., Margarint, V. and Sun, R. (2025). Scaling limits of disorder relevant non-binary spin systems. *Stoch. Partial Differ. Equ. Anal. Comput.* 1–23.

Li, Z., Song, J., Wei, R. and Zhang, H. (2026). On a continuum pinning model with colored Gaussian disorder. In preparation.

Mémin, J., Mishura, Y. and Valkeila, E. (2001). Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion. *Statist. Probab. Lett.* **51** 197–206.

Nualart, D. (2006). *The Malliavin Calculus and Related Topics 1995*. Springer.

Pipiras, V. and Taqqu, M.S. (2000). Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields* **118** 251–291.

Poisat, J. (2012). Random pinning model with finite range correlations: Disorder relevant regime. *Stoch. Process. Appl.* **122** 3560–3579.

Poisat, J. (2013). Ruelle-Perron-Frobenius operator approach to the annealed pinning model with Gaussian long-range correlated disorder. *Markov Process. Related Fields* **19** 577–606.

Quastel, J., Ramirez, A. and Virág, B. (2025). KPZ fluctuations in the planar stochastic heat equation. *Duke Math. J.* **174** 1261–1340.

Rang, G. (2020). From directed polymers in spatial-correlated environment to stochastic heat equations driven by fractional noise in 1+ 1 dimensions. *Stoch. Process. Appl.* **130** 3408–3444.

Rang, G., Song, J. and Wang, M. (2024). Scaling limit of a long-range random walk in time-correlated random environment. *Electron. J. Probab.* **29** 1–39.

Shen, H., Song, J., Sun, R. and Xu, L. (2021). Scaling limit of a directed polymer among a Poisson field of independent walks. *J. Funct. Anal.* **281** 109066.

Wei, R. and Yu, J. (2024). The critical disordered pinning measure. arXiv preprint, arXiv:2402.17642.

Weinrib, A. and Halperin, B.I. (1983). Critical phenomena in systems with long-range-correlated quenched disorder. *Phys. Rev. B* **27** 413.

Bernoulli **32**(4), 2026, 2826–2839
https://doi.org/10.3150/26-BEJ1969

On a population model with memory

JEAN BERTOIN[a]

Institute of Mathematics, University of Zurich, Zürich, Switzerland, [a]*jean.bertoin@math.uzh.ch*

Consider first a memoryless population model described by the usual branching process with a given mean reproduction matrix on a finite space of types. Motivated by the consequences of atavism in Evolutionary Biology, we are interested in a modification of the dynamics where individuals keep full memory of their forebears and procreation involves the reactivation of a gene picked at random on the ancestral lineage. By comparing the spectral radii of the two mean reproduction matrices (with and without memory), we observe that, on average, the model with memory always grows at least as fast as the model without memory. The proof relies on analyzing a biased Markov chain on the space of memories. The existence of a unique ergodic law for the latter is demonstrated through asymptotic coupling.

Keywords: Markov chains with memory; multitype branching process; spectral radius

References

Athreya, K.B. and Ney, P.E. (2004). *Branching Processes*. Mineola, NY: Dover Publications, Inc. Reprint of the 1972 original, [Springer, New York, MR0373040]. MR2047480

Benaïm, M. and Hurth, T. (2022). *Markov Chains on Metric Spaces—a Short Course*. *Universitext*. Cham: Springer. MR4559704 https://doi.org/10.1007/978-3-031-11822-7

Bertoin, J. and Mallein, B. (2024). Reinforced Galton-Watson processes. I: Malthusian exponents. *Random Structures Algorithms* **65** 387–410. https://doi.org/10.1002/rsa.21219

Bertoin, J. and Mallein, B. (2025a). Reinforced Galton-Watson processes. II: Large time behaviors. *Ann. Appl. Probab.* **35** 1353–1372. https://doi.org/10.1214/24-AAP2144

Bertoin, J. and Mallein, B. (2025b). Reinforced Galton-Watson processes. III: Empirical offspring distributions. *Probab. Theory Related Fields*.

Bhattacharya, R. and Waymire, E.C. (2021). *Random Walk, Brownian Motion, and Martingales*. *Graduate Texts in Mathematics* **292**. Cham: Springer. MR4331427 https://doi.org/10.1007/978-3-030-78939-8

Bhattacharya, R. and Waymire, E.C. (2022). *Stationary Processes and Discrete Parameter Markov Processes*. *Graduate Texts in Mathematics* **293**. Cham: Springer. MR4744443 https://doi.org/10.1007/978-3-031-00943-3

Blath, J., Hermann, F. and Slowik, M. (2021). A branching process model for dormancy and seed banks in randomly fluctuating environments. *J. Math. Biol.* **83**. Paper No. 17, 40. MR4289473 https://doi.org/10.1007/s00285-021-01639-6

Blath, J., González Casanova, A., Kurt, N. and Wilke-Berenguer, M. (2016). A new coalescent for seed-bank models. *Ann. Appl. Probab.* **26** 857–891. MR3476627 https://doi.org/10.1214/15-AAP1106

Champagnat, N., Strickler, É. and Villemonais, D. (2025). Uniform Wasserstein convergence of penalized Markov processes. *Probab. Theory Related Fields* **192** 1031–1069. https://doi.org/10.1007/s00440-025-01385-6

Del Moral, P., Horton, E. and Jasra, A. (2023). On the stability of positive semigroups. *Ann. Appl. Probab.* **33** 4424–4490. MR4674056 https://doi.org/10.1214/22-aap1923

Del Moral, P. and Miclo, L. (2000). Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV*. *Lecture Notes in Math.* **1729** 1–145. Berlin: Springer. MR1768060 https://doi.org/10.1007/BFb0103798

Greven, A., den Hollander, F. and Oomen, M. (2025). Spatial populations with seed-bank: Renormalisation on the hierarchical group. *Mem. Amer. Math. Soc.* **307** ix+207. MR4899541 https://doi.org/10.1090/memo/1553

Hairer, M. and Mattingly, J.C. (2006). Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)* **164** 993–1032. MR2259251 https://doi.org/10.4007/annals.2006.164.993

Hairer, M., Mattingly, J.C. and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** 223–259. https://doi.org/10.1007/s00440-009-0250-6

Harris, T.E. (1963). *The Theory of Branching Processes*. *Die Grundlehren der mathematischen Wissenschaften* **119**. Berlin: Springer-Verlag. MR0163361

Kaj, I., Krone, S.M. and Lascoux, M. (2001). Coalescent theory for seed bank models. *J. Appl. Probab.* **38** 285–300. MR1834743 https://doi.org/10.1017/s0021900200019860

Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, Second ed. Cambridge: Cambridge University Press. With a prologue by Peter W. Glynn. MR2509253 https://doi.org/10.1017/CBO9780511626630

Newburgh, J.D. (1951). The variation of spectra. *Duke Math. J.* **18** 165–176. https://doi.org/10.1215/S0012-7094-51-01813-3

Wagner, G.P. (2014). *Homology, Genes, and Evolutionary Innovation*. Princeton: Princeton University Press.

Bernoulli **32**(4), 2026, 2840–2864
https://doi.org/10.3150/26-BEJ1970

Maximum likelihood estimation for Brownian motion tree models based on one sample

CHANDLER SQUIRES[1,a], MICHAEL TRUELL[1,b], JAN-CHRISTIAN HÜTTER[2,d], PIOTR ZWIERNIK[3,4,e] and CAROLINE UHLER[1,c]

[1] *Laboratory for Information and Decision Systems, and Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA,* [a]*csquires@mit.edu,* [b]*truellm@mit.edu,* [c]*cuhler@mit.edu*
[2] *Eric and Wendy Schmidt Center, Broad Institute of MIT and Harvard, Cambridge, MA, USA,* [d]*jchuetter.web@gmail.com*
[3] *Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain,* [e]*piotr.zwiernik@upf.edu*
[4] *Barcelona School of Economics, Barcelona, Spain*

We study the problem of maximum likelihood estimation given one sample ($n = 1$) over Brownian Motion Tree Models (BMTMs), a class of Gaussian models on trees. BMTMs are often used as a null model in phylogenetics, where the one-sample regime is common. Specifically, we show that, for any fixed tree, almost surely, the one-sample BMTM maximum likelihood estimator (MLE) exists, is unique, and corresponds to a fully observed tree. Moreover, we provide a polynomial time algorithm for its exact computation. We also consider the one-sample MLE over all possible BMTM tree structures and show that it exists almost surely, that it coincides with the MLE over diagonally dominant M-matrices, and that it admits a unique closed-form solution that corresponds to a path graph. Finally, we explore statistical properties of the one-sample BMTM MLE through numerical experiments.

Keywords: Brownian motion tree model; diagonally dominant Gaussian model; Laplacian-structured Gaussian Markov random field; linear Gaussian covariance model; maximum likelihood estimation; MTP_2 matrices; phylogenetic tree; tree-structured latent random variable model

References

[1] Anderson, T.W. (1970). Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. In *Essays in Probability and Statistics*. *University of North Carolina Monograph Series in Probability and Statistics* **3** 1–24. Chapel Hill, NC: Univ. North Carolina Press. MR0277057
[2] 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
[3] Brown, L.D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **9**. Hayward, CA: Institute of Mathematical Statistics. https://doi.org/10.1214/lnms/1215466757 MR882001
[4] Campbell, P.J., Pleasance, E.D., Stephens, P.J., Dicks, E., Rance, R., Goodhead, I., Follows, G.A., Green, A.R., Futreal, P.A. and Stratton, M.R. (2008). Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc. Natl. Acad. Sci. USA* **105** 13081–13086. https://doi.org/10.1073/pnas.0801523105
[5] Cooper, N. and Purvis, A. (2010). Body size evolution in mammals: Complexity in tempo and mode. *Amer. Nat.* **175** 727–738. https://doi.org/10.1086/652466
[6] Dellacherie, C., Martinez, S. and San Martin, J. (2014). *Inverse M-Matrices and Ultrametric Matrices*. *Lecture Notes in Mathematics* **2118**. Cham: Springer. MR3289211 https://doi.org/10.1007/978-3-319-10298-6
[7] Dey, D.K. and Srinivasan, C. (1985). Estimation of a covariance matrix under Stein's loss. *Ann. Statist.* **13** 1581–1591. https://doi.org/10.1214/aos/1176349756 MR811511
[8] Duval, A.M., Klivans, C.J. and Martin, J.L. (2009). Simplicial matrix-tree theorems. *Trans. Amer. Math. Soc.* **361** 6073–6114. MR2529925 https://doi.org/10.1090/S0002-9947-09-04898-3

1350-7265

[9] Felsenstein, J. (1973). Maximum-likelihood estimation of evolutionary trees from continuous characters. *Amer. J. Hum. Genet.* **25** 471–492. https://doi.org/10.2307/2408134
[10] Felsenstein, J. (1981). Evolutionary trees from gene frequencies and quantitative characters: Finding maximum likelihood estimates. *Evolution* **35** 1229–1242. https://doi.org/10.2307/2408134
[11] Garba, M.K., Nye, T.M.W., Lueg, J. and Huckemann, S.F. (2021). Information geometry for phylogenetic trees. *J. Math. Biol.* **82** Paper No. 19, 39. MR4218000 https://doi.org/10.1007/s00285-021-01553-x
[12] Gillman, M.P. (2007). Evolutionary dynamics of vertebrate body mass range. *Evolution* **61** 685–693. https://doi.org/10.1111/j.1558-5646.2007.00060.x
[13] Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D.A., Finstad, K.M., Amundson, R., Thomas, B.C. and Banfield, J.F. (2016). A new view of the tree of life. *Nat. Microbiol.* **1** 16048. https://doi.org/10.1038/nmicrobiol.2016.48
[14] Jäger, G. (2018). Global-scale phylogenetic linguistic inference from lexical resources. *Sci. Data* **5** 1–16. https://doi.org/10.1038/sdata.2018.189
[15] Kozlov, A., Alves, J.M., Stamatakis, A. and Posada, D. (2022). CellPhy: Accurate and fast probabilistic inference of single-cell phylogenies from scDNA-seq data. *Genome Biol.* **23** 1–30. https://doi.org/10.1186/s13059-021-02583-w
[16] Lauritzen, S., Uhler, C. and Zwiernik, P. (2019). Maximum likelihood estimation in Gaussian models under total positivity. *Ann. Statist.* **47** 1835–1863. MR3953437 https://doi.org/10.1214/17-AOS1668
[17] Lauritzen, S., Uhler, C. and Zwiernik, P. (2021). Total positivity in exponential families with application to binary variables. *Ann. Statist.* **49** 1436–1459. MR4298870 https://doi.org/10.1214/20-aos2007
[18] Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* **88** 365–411. MR2026339 https://doi.org/10.1016/S0047-259X(03)00096-4
[19] Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Statist.* **40** 1024–1060. MR2985942 https://doi.org/10.1214/12-AOS989
[20] Makam, V., Reichenbach, P. and Seigal, A. (2023). Symmetries in directed Gaussian graphical models. *Electron. J. Stat.* **17** 3969–4010. MR4677191 https://doi.org/10.1214/23-ejs2192
[21] Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A. and Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32** 268–274. https://doi.org/10.1093/molbev/msu300
[22] Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *Ann. Statist.* **1** 763–765. https://doi.org/10.1214/aos/1176342472 MR331643
[23] Roch, S. (2006). A short proof that phylogenetic tree reconstruction by maximum likelihood is hard. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **3** 92–94. https://doi.org/10.1109/tcbb.2006.4
[24] Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4** 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454
[25] Slawski, M. and Hein, M. (2015). Estimation of positive definite M-matrices and structure learning for attractive Gaussian Markov random fields. *Linear Algebra Appl.* **473** 145–179. MR3338330 https://doi.org/10.1016/j.laa.2014.04.020
[26] Squires, C., Truell, M., Hütter, J.-C., Zwiernik, P. and Uhler, C. (2026). Supplement to "Maximum likelihood estimation for Brownian motion tree models based on one sample." https://doi.org/10.3150/26-BEJ1970SUPPA, https://doi.org/10.3150/26-BEJ1970SUPPB
[27] Sturmfels, B., Timme, S. and Zwiernik, P. (2020). Estimating linear covariance models with numerical nonlinear algebra. *J. Algebr. Stat.* **11** 31–52. MR4196404 https://doi.org/10.2140/astat.2020.11.31
[28] Sturmfels, B., Uhler, C. and Zwiernik, P. (2020). Brownian motion tree models are toric. *Kybernetika (Prague)* **56** 1154–1175. MR4199908 https://doi.org/10.14736/kyb-2020-6-1154
[29] Tuffley, C. and Steel, M. (1997). Links between maximum likelihood and maximum parsimony under a simple model of site substitution. *Bull. Math. Biol.* **59** 581–607. https://doi.org/10.1016/s0092-8240(97)00001-3
[30] Ying, J., Cardoso, J.V.M. and Palomar, D.P. (2021). Minimax estimation of Laplacian constrained precision matrices. In *In International Conference on Artificial Intelligence and Statistics* **130** 3736–3744. PMLR.
[31] Zwiernik, P., Uhler, C. and Richards, D. (2017). Maximum likelihood estimation for linear Gaussian covariance models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1269–1292. MR3689318 https://doi.org/10.1111/rssb.12217

Bernoulli **32**(4), 2026, 2865–2893
https://doi.org/10.3150/26-BEJ1971

Bootstrapping LASSO estimators under variable selection consistency in high dimensions and some higher order refinements

DEBRAJ DAS[1,a], ARINDAM CHATTERJEE[2,b] and SOUMENDRA N. LAHIRI[3,c]

[1]*Department of Mathematics, Indian Institute of Technology, Bombay, Mumbai 400076, India,*
[a]*debrajdas@math.iitb.ac.in*
[2]*Statistical Sciences Unit, Indian Statistical Institute Delhi, 7, S. J. S. Sansanwal Marg, New Delhi 110016, India,*
[b]*cha@isid.ac.in*
[3]*Department of Mathematics and Statistics, Washington University in St. Louis, MO 63130, United States,*
[c]*s.lahiri@wustl.edu*

We consider statistical inference based on the LASSO (cf. Tibshirani (*J. Roy. Statist. Soc. Ser. B, Methodol.* **58** (1996) 267–288)) in high dimensional regression problems. It is well known that the LASSO produces biased estimator of the regression parameter. The bias problem is further exacerbated when the LASSO has the variable selection consistency property (cf. Lahiri (*Ann. Statist.* **49** (2021) 820–844)). The centered and scaled point estimators are dominated by the bias term and fail to converge to *any* non-degenerate limit distribution. In contrast, we show that the LASSO interval estimators based on the Bootstrap are surprisingly accurate. Here, we consider two variants of the Bootstrap, namely, the Residual Bootstrap and the Perturbation Bootstrap, and show that under some regularity conditions, the Bootstrap approximations generated by both variants automatically adjust for the effects of the bias and are second order correct. This is an important finding as centered and scaled LASSO estimators under variable selection consistency fail to have a non-degenerate limit distribution and the Bootstrap approximations provide a viable way of constructing valid confidence intervals. Building on these second order results, we next consider construction of two-sided symmetric Bootstrap confidence intervals and show that with suitable choices of the studentized pivotal quantities, Bootstrap based two-sided symmetric confidence intervals attain a level of accuracy $O(n^{-2})$ where n is the sample size. Thus, even with penalization and with diverging model parameter dimension, the performance of the Bootstrap here remains comparable to its finite dimensional counterpart in the traditional Smooth Function model (cf. Hall (*The Bootstrap and Edgeworth Expansion* (1992) Springer-Verlag)). We also establish similar results for the heteroscedastic case for the Perturbation Bootstrap and report results from a moderate simulation study in support of the theoretical findings.

Keywords: Bootstrap; confidence interval; Edgeworth expansion; LASSO; second order correct; variable selection consistency

References

Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19** 521–547. MR3037163 https://doi.org/10.3150/11-BEJ410

Bhattacharya, R.N. and Ghosh, J.K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. MR0471142

Camponovo, L. (2015). On the validity of the pairs bootstrap for lasso estimators. *Biometrika* **102** 981–987. MR3431568 https://doi.org/10.1093/biomet/asv039

Chatterjee, A. and Lahiri, S.N. (2010). Asymptotic properties of the residual bootstrap for Lasso estimators. *Proc. Amer. Math. Soc.* **138** 4497–4509. MR2680074 https://doi.org/10.1090/S0002-9939-2010-10474-4

Chatterjee, A. and Lahiri, S.N. (2011). Bootstrapping lasso estimators. *J. Amer. Statist. Assoc.* **106** 608–625. MR2847974 https://doi.org/10.1198/jasa.2011.tm10159

1350-7265

Chatterjee, A. and Lahiri, S.N. (2013). Rates of convergence of the adaptive LASSO estimators to the oracle distribution and higher order refinements by the bootstrap. *Ann. Statist.* **41** 1232–1259. MR3113809 https://doi.org/10.1214/13-AOS1106

Choudhury, M. and Das, D. (2025). Does K-fold CV based penalty perform variable selection or does it lead to $n^{1/2}$-consistency in Lasso? arXiv preprint, arXiv:2507.12457.

Das, D., Chatterjee, A. and Lahiri, S.N. (2026). Supplement to "Bootstrapping LASSO estimators under variable selection consistency in high dimensions and some higher order refinements." https://doi.org/10.3150/26-BEJ1971SUPP

Das, D., Gregory, K. and Lahiri, S.N. (2019). Perturbation bootstrap in Adaptive Lasso. *Ann. Statist.* **47** 2080–2116. MR3953445 https://doi.org/10.1214/18-AOS1741

Das, D. and Lahiri, S.N. (2019a). Distributional consistency of the lasso by perturbation bootstrap. *Biometrika* **106** 957–964. MR4031208 https://doi.org/10.1093/biomet/asz029

Das, D. and Lahiri, S.N. (2019b). Second order correctness of perturbation bootstrap M-estimator of multiple linear regression parameter. *Bernoulli* **25** 654–682. MR3892332 https://doi.org/10.3150/17-bej1001

Dezeure, R., Bühlmann, P., Meier, L. and Meinshausen, N. (2015). High-dimensional inference: Confidence intervals, p-values and R-software hdi. *Statist. Sci.* **30** 533–558. MR3432840 https://doi.org/10.1214/15-STS527

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

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. MR1946581

Friedman, J.H., Hastie, T. and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33** 1–22. https://doi.org/10.18637/jss.v033.i01

Friedman, J., Hastie, T., Höfling, H. and Tibshirani, R. (2007). Pathwise coordinate optimization. *Ann. Appl. Stat.* **1** 302–332. MR2415737 https://doi.org/10.1214/07-AOAS131

Fu, W.J. (1998). Penalized regressions: The bridge versus the lasso. *J. Comput. Graph. Statist.* **7** 397–416. MR1646710 https://doi.org/10.2307/1390712

Hall, P. (1988). On symmetric bootstrap confidence intervals. *J. Roy. Statist. Soc. Ser. B, Methodol.* **50** 35–45. MR0954730

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. *Springer Series in Statistics*. New York: Springer-Verlag. MR1145237

Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909.

Knight, K. and Fu, W. (2000). Asymptotics for lasso-type estimators. *Ann. Statist.* **28** 1356–1378. MR1805787

Kuchibhotla, A.K. and Chakrabortty, A. (2022). Moving beyond sub-Gaussianity in high-dimensional statistics: Applications in covariance estimation and linear regression. *Inf. Inference* **iaac012**. https://doi.org/10.1093/imaiai/iaac012

Lahiri, S.N. (2021). Necessary and sufficient conditions for variable selection consistency of the LASSO in high dimensions. *Ann. Statist.* **49** 820–844. MR4255109 https://doi.org/10.1214/20-aos1979

Lee, J.D., Sun, D.L., Sun, Y. and Taylor, J.E. (2016). Exact post-selection inference, with application to the lasso. *Ann. Statist.* **44** 907–927. MR3485948 https://doi.org/10.1214/15-AOS1371

Liu, R.Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Ann. Statist.* **16** 1696–1708. https://doi.org/10.1214/aos/1176351062 MR964947

Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. MR2278363

Minnier, J., Tian, L. and Cai, T. (2011). A perturbation method for inference on regularized regression estimates. *J. Amer. Statist. Assoc.* **106** 1371–1382. MR2896842 https://doi.org/10.1198/jasa.2011.tm10382

Osborne, M.R., Presnell, B. and Turlach, B.A. (2000). On the LASSO and its dual. *J. Comput. Graph. Statist.* **9** 319–337. MR1822089

Taylor, J. and Tibshirani, R. (2018). Post-selection inference for ℓ_1-penalized likelihood models. *Canad. J. Statist.* **46** 41–61. MR3767165 https://doi.org/10.1002/cjs.11313

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

Tibshirani, R.J. (2013). The lasso problem and uniqueness. *Electron. J. Stat.* **7** 1456–1490. MR3066375 https://doi.org/10.1214/13-EJS815

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

Vladimirova, M., Girard, S., Nguyen, H. and Arbel, J. (2020). Sub-Weibull distributions: Generalizing sub-Gaussian and sub-exponential properties to heavier tailed distributions. *Stat* **9** e318. https://doi.org/10.1002/sta4.318

Wagener, J. and Dette, H. (2013). The adaptive Lasso in high-dimensional sparse heteroscedastic models. *Math. Methods Statist.* **22** 137–154. MR3071959 https://doi.org/10.3103/S106653071302004X

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

Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* **14** 1261–1295. https://doi.org/10.1214/aos/1176350142

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

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

Zou, H. (2006). The Adaptive Lasso and its oracle properties. *J. Amer. Statist. Assoc.* **101** 1418–1429. MR2279469

Zou, H. and Li, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Ann. Statist.* **36** 1509–1533. MR2435443 https://doi.org/10.1214/009053607000000802

Bernoulli **32**(4), 2026, 2894–2915
https://doi.org/10.3150/26-BEJ1972

Voter model stability with respect to conservative noises

GIDEON AMIR[1,a], OMER ANGEL[2,b], RANGEL BALDASSO[3,d] and DANIEL DE LA RIVA[2,c]

[1] *Bar Ilan University, Ramat Gan, Israel,* [a]*gideon.amir@biu.ac.il*
[2] *University of British Columbia, Vancouver, Canada,* [b]*angel@math.ubc.ca,* [c]*delariva@math.ubc.ca*
[3] *Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil,* [d]*rangel@puc-rio.br*

The notions of noise sensitivity and stability were recently extended for the voter model. In this model, the vertices of a graph have opinions that are updated by uniformly selecting edges. We further extend stability results to different classes of perturbations. We consider two different types of noise: in the first one, an exclusion process is performed on the edge selections, while in the second, independent Brownian motions are applied to such a sequence. In both cases, we prove stability of the consensus opinion provided the noise is run for a short amount of time, depending on the underlying graph structure. This is done by analyzing the expected size of the pivotal set, whose definition differs from the usual one in order to reflect the change associated with these noises.

Keywords: Brownian motion; interchange process; noise sensitivity; voter model

References

[1] Ahlberg, D. and Baldasso, R. (2018). Noise sensitivity and Voronoi percolation. *Electron. J. Probab.* **23**.
[2] Ahlberg, D., de la Riva, D. and Griffiths, S. (2021). On the rate of convergence in quenched Voronoi percolation. *Electron. J. Probab.* **26** 1–26.
[3] Ahlberg, D., Broman, E., Griffiths, S. and Morris, R. (2014). Noise sensitivity in continuum percolation. *Israel J. Math.* **201** 847–899.
[4] Ahlberg, D., Griffiths, S., Morris, R. and Tassion, V. (2016). Quenched Voronoi percolation. *Adv. Math.* **286** 889–911.
[5] Aleliunas, R., Karp, R.M., Lipton, R.J., Lovász, L. and Rackoff, C. (1979). Random walks, universal traversal sequences, and the complexity of maze problems. In *20th Annual Symposium on Foundations of Computer Science (sfcs 1979)* 218–223. IEEE Computer Society.
[6] Amir, G., Angel, O., Baldasso, R. and Peretz, R. (2022). Dynamical noise sensitivity for the voter model. *Electron. Commun. Probab.* **27**.
[7] Benjamini, I., Kalai, G. and Schramm, O. (1999). Noise sensitivity of Boolean functions and applications to percolation. *Publ. Math. IHÉS* **90**.
[8] Bollobás, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics*. Cambridge University Press.
[9] Broman, E.I., Garban, C. and Steif, J.E. (2013). Exclusion sensitivity of Boolean functions. *Probab. Theory Related Fields* **155**.
[10] Forsström, M.P. (2016). Monotonicity properties of exclusion sensitivity. *Electron. J. Probab.* **21** 1–22.
[11] Garban, C., Pete, G. and Schramm, O. (2010). The Fourier spectrum of critical percolation. *Acta Math.* **205** 19–104.
[12] Garban, C. and Vanneuville, H. (2016). Exceptional times for percolation under exclusion dynamics. arXiv preprint. Available at arXiv:1605.04766.
[13] Garban, C. and Vanneuville, H. (2020). Bargmann-Fock percolation is noise sensitive. *Electron. J. Probab.* **25** 1–20.

1350-7265

[14] Levin, D.A., Peres, Y. and Wilmer, E.L. (2006). *Markov Chains and Mixing Times*. American Mathematical Society.
[15] Liggett, T.M. (1985). *Interacting Particle Systems*. Berlin Heidelberg: Springer.
[16] Mossel, E., O'Donnell, R. and Oleszkiewicz, K. (2005). Noise stability of functions with low influences: Invariance and optimality. *Ann. Math.* **171** 21–30.
[17] O'Donnell, R. (2014). *Analysis of Boolean Functions*. Cambridge University Press.
[18] Oliveira, R.I. (2012). On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.* **364**.
[19] Schramm, O. and Steif, J.E. (2010). Quantitative noise sensitivity and exceptional times for percolation. *Ann. Math.* **171** 619–672.
[20] Tassion, V. and Vanneuville, H. (2023). Noise sensitivity of percolation via differential inequalities. *Proc. Lond. Math. Soc.* **126** 1063–1091.
[21] Tassion, V. and Vanneuville, H. (2025). Noise sensitivity of crossings for high temperature Ising model. arXiv preprint. Available at arXiv:2505.11457.

Bernoulli **32**(4), 2026, 2916–2943
https://doi.org/10.3150/26-BEJ1973

Optimal confidence bands for shape-restricted regression in multidimensions

ASHLEY (PRATYAY) DATTA[1,a], SOMABHA MUKHERJEE[2,b] and BODHISATTVA SEN[3,c]

[1] *Department of Statistics, Columbia University, New York, USA,* [a] *pd2511@columbia.edu*
[2] *Department of Statistics and Data Science, National University of Singapore, Singapore, Singapore,* [b] *somabha@nus.edu.sg*
[3] *Department of Statistics, Columbia University, New York, USA,* [c] *bodhi@stat.columbia.edu*

In this paper, we propose and study construction of confidence bands for shape-constrained regression functions when the predictor is multivariate. In particular, we consider the continuous multidimensional white noise model given by $dY(\boldsymbol{t}) = n^{1/2} f(\boldsymbol{t})\,d\boldsymbol{t} + dW(\boldsymbol{t})$, where Y is the observed stochastic process on $[0,1]^d$ ($d \geq 1$), W is the standard Brownian sheet on $[0,1]^d$, and f is the unknown function of interest assumed to belong to a (shape-constrained) function class, e.g., coordinate-wise monotone functions or convex functions. The constructed confidence bands are based on local kernel averaging with bandwidth chosen automatically via a multivariate multiscale statistic. The confidence bands have guaranteed coverage for every n and for every member of the underlying function class. Under monotonicity/convexity constraints on f, the proposed confidence bands automatically adapt (in terms of width) to the global and local (Hölder) smoothness and intrinsic dimensionality of the unknown f; the bands are also shown to be optimal in a certain sense. These bands have (almost) parametric ($n^{-1/2}$) widths when the underlying function has "low-complexity" (e.g., piecewise constant/affine). We also develop an optimal testing methodology for checking if the underlying regression function is actually coordinate-wise monotone.

Keywords: Automatic adaptation; continuous multidimensional white-noise model; coordinate-wise nondecreasing function; Hölder smoothness; minimax optimal; multivariate convex function

References

Allon, G., Beenstock, M., Hackman, S., Passy, U. and Shapiro, A. (2007). Nonparametric estimation of concave production technologies by entropic methods. *J. Appl. Econometrics* **22** 795–816.

Arias-Castro, E., Donoho, D. and Huo, X. (2005). Near-optimal detection of geometric objects by fast multiscale methods. *IEEE Trans. Inf. Theory* **51** 2402–2425. https://doi.org/10.1109/TIT.2005.850056

Banerjee, M., Durot, C. and Sen, B. (2019). Divide and conquer in nonstandard problems and the super-efficiency phenomenon. *Ann. Statist.* **47** 720–757. MR3909948 https://doi.org/10.1214/17-AOS1633

Banerjee, M. and Wellner, J.A. (2001). Likelihood ratio tests for monotone functions. *Ann. Statist.* **29** 1699–1731. MR1891743 https://doi.org/10.1214/aos/1015345959

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference Under Order Restrictions. The Theory and Application of Isotonic Regression*. London–New York–Sydney: John Wiley & Sons. MR326887

Bellec, P.C. (2021). Adaptive confidence sets in shape restricted regression. *Bernoulli* **27** 66–92. https://doi.org/10.3150/20-BEJ1223

Brown, L.D., Cai, T.T. and Zhou, H.H. (2008). Robust nonparametric estimation via wavelet median regression. *Ann. Statist.* **36** 2055–2084. https://doi.org/10.1214/07-AOS513

Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958

Brunk, H.D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Statist.* **26** 607–616. MR73894 https://doi.org/10.1214/aoms/1177728420

1350-7265

Brunk, H.D. (1970). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)* 177–197. London: Cambridge Univ. Press. MR0277070

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

Carter, A.V. (2006). A continuous Gaussian approximation to a nonparametric regression in two dimensions. *Bernoulli* **12** 143–156. MR2202326

Chambers, R. (1982). Duality, the output effect, and applied comparative statics. *Amer. J. Agric. Econ.* **64** 152–156.

Chan, H.P. and Walther, G. (2013). Detection with the scan and the average likelihood ratio. *Statist. Sinica* **23** 409–428. MR3076173

Chatterjee, S., Guntuboyina, A. and Sen, B. (2015). On risk bounds in isotonic and other shape restricted regression problems. *Ann. Statist.* **43** 1774–1800. MR3357878 https://doi.org/10.1214/15-AOS1324

Chatterjee, S., Guntuboyina, A. and Sen, B. (2018). On matrix estimation under monotonicity constraints. *Bernoulli* **24** 1072–1100. MR3706788 https://doi.org/10.3150/16-BEJ865

Chatterjee, S. and Lafferty, J. (2019). Adaptive risk bounds in unimodal regression. *Bernoulli* **25** 1–25. MR3892309 https://doi.org/10.3150/16-bej922

Chen, X., Lin, Q. and Sen, B. (2020). On degrees of freedom of projection estimators with applications to multivariate nonparametric regression. *J. Amer. Statist. Assoc.* **115** 173–186. MR4078455 https://doi.org/10.1080/01621459.2018.1537917

Chen, X., Chernozhukov, V., Fernández-Val, I., Kostyshak, S. and Luo, Y. (2018). Shape-enforcing operators for point and interval estimators. Available at arXiv:1809.01038v3.

Datta, A., Mukherjee, S. and Sen, B. (2026). Supplement to "Optimal confidence bands for shape-restricted regression in multidimensions." https://doi.org/10.3150/26-BEJ1973SUPP

Datta, P. and Sen, B. (2021). Optimal inference with a multidimensional multiscale statistic. *Electron. J. Stat.* **15** 5203–5244.

Davies, P.L. (1995). Data features. *Stat. Neerl.* **49** 185–245.

Deng, H., Han, Q. and Sen, B. (2022). Inference for local parameters in convexity constrained models. *J. Amer. Statist. Assoc.* 1–15.

Deng, H., Han, Q. and Zhang, C.-H. (2021). Confidence intervals for multiple isotonic regression and other monotone models. *Ann. Statist.* **49** 2021–2052. MR4319240 https://doi.org/10.1214/20-aos2025

Deng, H. and Zhang, C.-H. (2020). Isotonic regression in multi-dimensional spaces and graphs. *Ann. Statist.* **48** 3672–3698. MR4185824 https://doi.org/10.1214/20-AOS1947

Donoho, D.L. (1988). One-sided inference about functionals of a density. *Ann. Statist.* **16** 1390–1420.

Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425–455.

Donoho, D.L. and Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.* **90** 1200–1224.

Donoho, D.L. and Low, M.G. (1992). Renormalization exponents and optimal pointwise rates of convergence. *Ann. Statist.* **20** 944–970. MR1165601 https://doi.org/10.1214/aos/1176348665

Dümbgen, L. (1998). New goodness-of-fit tests and their application to nonparametric confidence sets. *Ann. Statist.* **26** 288–314. MR1611768 https://doi.org/10.1214/aos/1030563987

Dümbgen, L. (2003). Optimal confidence bands for shape-restricted curves. *Bernoulli* **9** 423–449. MR1997491 https://doi.org/10.3150/bj/1065444812

Dümbgen, L. and Spokoiny, V.G. (2001). Multiscale testing of qualitative hypotheses. *Ann. Statist.* **29** 124–152. MR1833961 https://doi.org/10.1214/aos/996986504

Durot, C., Kulikov, V.N. and Lopuhaä, H.P. (2012). The limit distribution of the L_∞-error of Grenander-type estimators. *Ann. Statist.* **40** 1578–1608. MR3015036 https://doi.org/10.1214/12-AOS1015

Fan, J. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. Chapman and Hall. https://doi.org/10.1201/9780203748725

Freyberger, J. and Reeves, B. (2018). Inference under shape restrictions. Social Science Research Network.

Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics* **40**. New York: Cambridge University Press. MR3588285 https://doi.org/10.1017/CBO9781107337862

Groeneboom, P. and Jongbloed, G. (2014). Estimators, algorithms and asymptotics. In *Nonparametric Estimation Under Shape Constraints*. *Cambridge Series in Statistical and Probabilistic Mathematics* **38**. New York: Cambridge University Press. MR3445293 https://doi.org/10.1017/CBO9781139020893

Groeneboom, P., Jongbloed, G. and Wellner, J.A. (2001). Estimation of a convex function: Characterizations and asymptotic theory. *Ann. Statist.* **29** 1653–1698. MR1891742 https://doi.org/10.1214/aos/1015345958

Groeneboom, P. and Wellner, J.A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. *DMV Seminar* **19**. Basel: Birkhäuser Verlag. MR1180321 https://doi.org/10.1007/978-3-0348-8621-5

Gu, C. (2002). *Smoothing Spline ANOVA Models*. *Springer Series in Statistics*. New York: Springer-Verlag. MR1876599 https://doi.org/10.1007/978-1-4757-3683-0

Guntuboyina, A. and Sen, B. (2015). Global risk bounds and adaptation in univariate convex regression. *Probab. Theory Related Fields* **163** 379–411. MR3405621 https://doi.org/10.1007/s00440-014-0595-3

Guntuboyina, A. and Sen, B. (2018). Nonparametric shape-restricted regression. *Statist. Sci.* **33** 568–594. MR3881209 https://doi.org/10.1214/18-STS665

Han, Q. and Zhang, C.-H. (2020). Limit distribution theory for block estimators in multiple isotonic regression. *Ann. Statist.* **48** 3251–3282. MR4185808 https://doi.org/10.1214/19-AOS1928

Han, Q., Wang, T., Chatterjee, S. and Samworth, R.J. (2019). Isotonic regression in general dimensions. *Ann. Statist.* **47** 2440–2471. MR3988762 https://doi.org/10.1214/18-AOS1753

Hart, J.D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. New York: Springer.

Hengartner, N.W. and Stark, P.B. (1995). Finite-sample confidence envelopes for shape-restricted densities. *Ann. Statist.* **23** 525–550. MR1332580 https://doi.org/10.1214/aos/1176324534

Hildreth, C. (1954). Point estimates of ordinates of concave functions. *J. Amer. Statist. Assoc.* **49** 598–619. MR0065093

Johnstone, I.M. (1999). Wavelets and the theory of non-parametric function estimation. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **357** 2475–2493.

Keshavarz, A., Wang, Y. and Boyd, S. (2011). Imputing a convex objective function. In *2011 IEEE International Symposium on Intelligent Control* 613–619.

Khoshnevisan, D. (1996). Five lectures on Brownian sheet. Summer Internship Program University of Wisconsin–Madison.

König, C., Munk, A. and Werner, F. (2020). Multidimensional multiscale scanning in exponential families: Limit theory and statistical consequences. *Ann. Statist.* **48** 655–678. MR4102671 https://doi.org/10.1214/18-AOS1806

Kuosmanen, T. and Johnson, A.L. (2010). Data envelopment analysis as nonparametric least-squares regression. *Oper. Res.* **58** 149–160.

Kur, G., Gao, F., Guntuboyina, A. and Sen, B. (2020). Convex regression in multidimensions: Suboptimality of least squares estimators. arXiv preprint. Available at arXiv:2006.02044.

Lian, H., Zhao, K. and Lv, S. (2019). Projected spline estimation of the nonparametric function in high- dimensional partially linear models for massive data. *Ann. Statist.* **47** 2922–2949.

Luss, R., Rosset, S. and Shahar, M. (2012). Efficient regularized isotonic regression with application to gene-gene interaction search. *Ann. Appl. Stat.* **6** 253–283. MR2951537 https://doi.org/10.1214/11-AOAS504

McMahan, H.B., Holt, G., Sculley, D., Young, M., Ebner, D., Grady, J., Nie, L., Phillips, T., Davydov, E., Golovin, D. et al. (2013). Ad click prediction: A view from the trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1222–1230.

Meyer, M. and Woodroofe, M. (2000). On the degrees of freedom in shape-restricted regression. *Ann. Statist.* **28** 1083–1104. MR1810920 https://doi.org/10.1214/aos/1015956708

Mukherjee, S., Patra, R., Johnson, A. and Morita, H. (2023). Least squares estimation of a monotone quasiconvex regression function. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* https://doi.org/10.1093/jrsssb/qkad133

Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density. *Sankhyā Ser. A* **31** 23–36. MR0267677

Protter, P.E. (2005). *Stochastic Integration and Differential Equations*. *Stochastic Modelling and Applied Probability* **2**. Springer.

Reiß, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.* **36** 1957–1982. https://doi.org/10.1214/07-AOS525

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Chichester: John Wiley & Sons, Ltd. MR0961262

Seijo, E. and Sen, B. (2011). Nonparametric least squares estimation of a multivariate convex regression function. *Ann. Statist.* **39** 1633–1657. MR2850215 https://doi.org/10.1214/10-AOS852

Sen, B., Banerjee, M. and Woodroofe, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *Ann. Statist.* **38** 1953–1977. MR2676880 https://doi.org/10.1214/09-AOS777

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original. Translated by Vladimir Zaiats. MR2724359 https://doi.org/10.1007/b13794

Varian, H.R. (1992). *Microeconomic Analysis*. WW Norton.

Varian, H.R. (2010). *Intermediate Microeconomics, a Modern Approach*, 8th ed. Macmillan & Company.

Wahba, G. (1990). *Spline Models for Observational Data*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. Society for Industrial and Applied Mathematics.

Walther, G. (2010). Optimal and fast detection of spatial clusters with scan statistics. *Ann. Statist.* **38** 1010–1033. MR2604703 https://doi.org/10.1214/09-AOS732

Walther, G. and Perry, A. (2022). Calibrating the scan statistic: Finite sample performance versus asymptotics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1608–1639. MR4515552

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. *Monographs on Statistics and Applied Probability* **60**. London: Chapman and Hall, Ltd. MR1319818 https://doi.org/10.1007/978-1-4899-4493-1

Zhang, C.-H. (2002). Risk bounds in isotonic regression. *Ann. Statist.* **30** 528–555. MR1902898 https://doi.org/10.1214/aos/1021379864

Bernoulli **32**(4), 2026, 2944–2968
https://doi.org/10.3150/26-BEJ1974

Debiased estimation and variable selection under function-on-scalar linear regression models with ultrahigh-dimensional covariates subject to measurement error

YIFAN SUN[1,a] and GRACE Y. YI[1,2,b]

[1]*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Canada,* [a]*ysun2343@uwo.ca*
[2]*Department of Computer Science, University of Western Ontario, London, Canada,* [b]*gyi5@uwo.ca*

In real-world applications, data are often error-contaminated; naively applying conventional methods without accommodating the measurement error effects often yields inconsistent estimates. Biased results can be further exacerbated by the ultrahigh-dimensionality of covariates. Focusing on the widely used function-on-scalar linear regression model, this article develops new methods for simultaneous parameter estimation and variable selection with error-prone covariates that can be ultrahigh-dimensional. The proposed framework provides flexibility to handle different types of measurement error models. We rigorously establish asymptotic properties of the proposed estimators under mild conditions. Notably, the convergence rates and limiting distributions of the proposed estimators depend on the nature of measurement error. Our findings highlight the significant differences of settings with ultrahigh dimensions compared to scenarios with finite dimensions, as well as the drastically different influence of different measurement error processes. For efficient computation, we design algorithms with data-driven tuning. We evaluate the finite sample performance of the proposed method through simulation studies and a real data application, demonstrating its effectiveness in addressing the challenges posed by error-contaminated and ultrahigh-dimensional of covariates.

Keywords: Bias; convergence rate; functional data analysis; function-on-scalar regression; measurement error; ultrahigh dimension; variable selection

References

[1] Barber, R.F., Reimherr, M. and Schill, T. (2017). The function-on-scalar LASSO with applications to longitudinal GWAS. *Electron. J. Stat.* **11** 1351–1389.
[2] Berkson, J. (1950). Are there two regressions? *J. Amer. Statist. Assoc.* **45** 164–180.
[3] Breheny, P. and Huang, J. (2015). Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. *Stat. Comput.* **25** 173–187.
[4] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Berlin: Springer Science & Business Media.
[5] Cai, X., Xue, L. and Cao, J. (2022). Robust estimation and variable selection for function-on-scalar regression. *Canad. J. Statist.* **50** 162–179.
[6] Cai, T.T. and Yuan, M. (2011). Optimal estimation of the mean function based on discretely sampled functional data: Phase transition. *Ann. Statist.* **39** 2330–2355.
[7] Cai, T.T., Zhang, C.-H. and Zhou, H.H. (2010). Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.* **38** 2118–2144.
[8] Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd ed. ed. New York: Chapman and Hall/CRC.

[9] Chen, Y., Goldsmith, J. and Ogden, R.T. (2016). Variable selection in function-on-scalar regression. *Stat* **5** 88–101.

[10] Chiang, C.-T., Rice, J.A. and Wu, C.O. (2001). Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. *J. Amer. Statist. Assoc.* **96** 605–619.

[11] Datta, A. and Zou, H. (2017). Cocolasso for high-dimensional error-in-variables regression. *Ann. Statist.* **45** 2400–2426.

[12] De Boor, C. (2001). *A Practical Guide to Splines*, 1st, revised ed. ed. New York: Springer.

[13] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360.

[14] Fan, J. and Lv, J. (2010). A selective overview of variable selection in high dimensional feature space. *Statist. Sinica* **20** 101–148.

[15] Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* **32** 928–961.

[16] Fan, Z. and Reimherr, M. (2017). High-dimensional adaptive function-on-scalar regression. *Econom. Stat.* **1** 167–183.

[17] Fan, J., Xue, L. and Zou, H. (2016). Multitask quantile regression under the transnormal model. *J. Amer. Statist. Assoc.* **111** 1726–1735.

[18] Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. New York: Springer.

[19] Kim, Y., Choi, H. and Oh, H.-S. (2008). Smoothly clipped absolute deviation on high dimensions. *J. Amer. Statist. Assoc.* **103** 1665–1673.

[20] Kong, D., Xue, K., Yao, F. and Zhang, H.H. (2016). Partially functional linear regression in high dimensions. *Biometrika* **103** 147–159.

[21] Küchenhoff, H., Bender, R. and Langner, I. (2007). Effect of Berkson measurement error on parameter estimates in Cox regression models. *Lifetime Data Anal.* **13** 261–272.

[22] Li, X., Wang, L., Wang, H.J. and Initiative, A.D.N. (2021). Sparse learning and structure identification for ultrahigh-dimensional image-on-scalar regression. *J. Amer. Statist. Assoc.* **116** 1994–2008.

[23] Loh, P.-L. and Wainwright, M.J. (2011). High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. In *Adv. Neural. Inf. Process. Syst* **24**.

[24] Masiuk, S., Shklyar, S., Kukush, A., Carroll, R., Kovgan, L. and Likhtarov, I. (2016). Estimation of radiation risk in presence of classical additive and Berkson multiplicative errors in exposure doses. *Biostatistics* **17** 422–436.

[25] Meng, S., Huang, Z., Zhang, J. and Jiang, Z. (2021). Estimation on functional partially linear single index measurement error model. *Comm. Statist. Theory Methods* **52** 1–23.

[26] Parodi, A. and Reimherr, M. (2018). Simultaneous variable selection and smoothing for high-dimensional function-on-scalar regression. *Electron. J. Stat.* **12** 4602–4639.

[27] Phillips, G.M. and Taylor, P.J. (1996). *Theory and Applications of Numerical Analysis*, 2nd ed. ed. London: Elsevier.

[28] Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. ed. New York: Springer.

[29] Sun, Y. and Yi, G.Y. (2026a). Supplement to "Debiased estimation and variable selection under function-on-scalar linear regression models with ultrahigh-dimensional covariates subject to measurement error." https://doi.org/10.3150/26-BEJ1974SUPP

[30] Sun, Y. and Yi, G.Y. (2026b). Estimation and variable selection under the function-on-scalar linear model with covariate measurement error. *Statist. Sinica* **36** 525–543.

[31] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **58** 267–288.

[32] 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.

[33] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge: Cambridge University Press.

[34] Wainwright, M.J. (2019). *High-Dimensional Statistics: A Non-asymptotic Viewpoint*. Cambridge: Cambridge University Press.

[35] Wang, L., Chen, G. and Li, H. (2007). Group SCAD regression analysis for microarray time course gene expression data. *Bioinformatics* **23** 1486–1494.

[36] Wang, L., Kim, Y. and Li, R. (2013). Calibrating non-convex penalized regression in ultra-high dimension. *Ann. Statist.* **41** 2505–2536.
[37] Xiao, L. (2019). Asymptotic theory of penalized splines. *Electron. J. Stat.* **13** 747–794.
[38] Xiao, L. (2020). Asymptotic properties of penalized splines for functional data. *Bernoulli* **26** 2847–2875.
[39] Yi, G.Y. (2017). *Statistical Analysis with Measurement Error or Misclassification: Strategy, Method and Application*. New York: Springer.
[40] Yi, G.Y., Delaigle, A. and Gustafson, P. (2021). *Handbook of Measurement Error Models*. New York: Chapman and Hall/CRC.
[41] Yueh, W.-C. (2005). Eigenvalues of several tridiagonal matrices. *Appl. Math. E-Notes* **5** 66–74.
[42] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942.
[43] Zhang, J.-T. and Chen, J. (2007). Statistical inferences for functional data. *Ann. Statist.* **35** 1052–1079.
[44] Zhao, M., Gao, Y. and Cui, Y. (2022). Variable selection for longitudinal varying coefficient errors-in-variables models. *Comm. Statist. Theory Methods* **51** 3713–3738.
[45] Zhou, S., Shen, X. and Wolfe, D.A. (1998). Local asymptotics for regression splines and confidence regions. *Ann. Statist.* **26** 1760–1782.
[46] Zhu, H., Li, R. and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *Ann. Statist.* **40** 2634–2666.
[47] Zhu, H., Zhang, R. and Zhu, G. (2020). Estimation and inference in semi-functional partially linear measurement error models. *J. Syst. Sci. Complex.* **33** 1179–1199.
[48] Zhu, H., Zhang, R., Yu, Z., Lian, H. and Liu, Y. (2019). Estimation and testing for partially functional linear errors-in-variables models. *J. Multivariate Anal.* **170** 296–314.

Bernoulli **32**(4), 2026, 2969–2998
https://doi.org/10.3150/26-BEJ1975

Large deviations for Independent Metropolis Hastings and Metropolis-adjusted Langevin algorithm

FEDERICA MILINANNI[1,a] and PIERRE NYQUIST[2,b]

[1] *Institute for Computational and Experimental Research in Mathematics (ICERM), Brown University, Providence, RI 02903, USA,* [a] *federica_milinanni@brown.edu*
[2] *Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden,* [b] *pnyquist@chalmers.se*

In this paper, we prove large deviation principles for the empirical measures associated with the Independent Metropolis Hastings (IMH) sampler and the Metropolis-adjusted Langevin Algorithm (MALA). These are the first large deviation results for empirical measures of Markov chains arising from specific Metropolis-Hastings methods on a continuous state space. Moreover, we show that the existing large deviation framework, that we developed in a previous work (Milinanni and Nyquist, 2024) does not cover the Random Walk Metropolis sampler, even in cases when the underlying Markov chain is geometrically ergodic.

Keywords: Empirical measure; Large deviations; Lyapunov function; Markov chain Monte Carlo; Metropolis-Hastings

References

[1] Andrieu, C., De Freitas, N., Doucet, A. and Jordan, M.I. (2003). An introduction to MCMC for machine learning. *Machine Learning* **50** 5–43.
[2] 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.
[3] Andrieu, C., Lee, A., Power, S. and Wang, A.Q. (2023). Weak Poincaré Inequalities for Markov chains: theory and applications.
[4] Atchadé, Y.F. and Perron, F. (2007). On the geometric ergodicity of Metropolis-Hastings algorithms. *Statistics* **41** 77–84.
[5] Bédard, M. and Rosenthal, J.S. (2008). Optimal scaling of Metropolis algorithms: Heading toward general target distributions. *Canad. J. Statist.* **36** 483–503.
[6] Besag, J. (1994). Comments on "Representations of knowledge in complex systems" by U. Grenander and M. I. Miller. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **56** 4.
[7] Bierkens, J. (2016). Non-reversible Metropolis-Hastings. *Stat. Comput.* **26** 1213–1228.
[8] Bierkens, J., Nyquist, P. and Schlottke, M.C. (2021). Large deviations for the empirical measure of the zig-zag process. *Ann. Appl. Probab.* **31** 2811–2843.
[9] Brown, A. and Jones, G.L. (2024). Exact convergence analysis for Metropolis–Hastings independence samplers in Wasserstein distances. *J. Appl. Probab.* **61** 33–54.
[10] Budhiraja, A. and Dupuis, P. (2003). Large deviations for the emprirical measures of reflecting Brownian motion and related constrained processes in R_+. *Electron. J. Probab.* **8** 1–46.
[11] Budhiraja, A. and Dupuis, P. (2019). Analysis and approximation of rare events. Representations and weak convergence methods. *Series Prob. Theory and Stoch. Modelling* **94**.
[12] Chatterjee, S. and Diaconis, P. (2018). The sample size required in importance sampling. *Ann. Appl. Probab.* **28** 1099–1135.

1350-7265

[13] Christensen, O.F., Roberts, G.O. and Rosenthal, J.S. (2005). Scaling limits for the transient phase of local Metropolis–Hastings algorithms. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 253–268.
[14] Dembo, A. and Zeitouni, O. (2009). *Large Deviations Techniques and Applications 38*. Springer.
[15] Diaconis, P., Holmes, S. and Neal, R.M. (2000). Analysis of a nonreversible Markov chain sampler. *Ann. Appl. Probab.* **10** 726–752. https://doi.org/10.1214/aoap/1019487508
[16] Doll, J.D., Dupuis, P. and Nyquist, P. (2018). A large deviations analysis of certain qualitative properties of parallel tempering and infinite swapping algorithms. *Appl. Math. Optim.* **78** 103–144.
[17] Dupuis, P. and Ellis, R.E. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley.
[18] Dupuis, P. and Wu, G.-J. (2022). Analysis and optimization of certain parallel Monte Carlo methods in the low temperature limit. *Multiscale Model. Simul.* **10** 220–249.
[19] Dupuis, P., Liu, Y., Plattner, N. and Doll, J.D. (2012). On the infinite swapping limit for parallel tempering. *Multiscale Model. Simul.* **10** 986–1022.
[20] Feng, J. and Kurtz, T.G. (2006). *Large Deviations for Stochastic Processes 131*. American Mathematical Soc.
[21] Franke, B., Hwang, C.R., Pai, H.M. and Sheu, S.J. (2010). The behavior of the spectral gap under growing drift. *Multiscale Model. Simul.* **362** 1325–1350.
[22] Frigessi, A., Stefano, P., Hwang, C.R. and Sheu, S.J. (1993). Convergence rates of the Gibbs sampler, the Metropolis algorithm and other single-site updating dynamics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **55** 205–219.
[23] Gelman, A., Gilks, W.R. and Roberts, G.O. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7** 110–120.
[24] Glatt-Holtz, N., Krometis, J. and Mondaini, C. (2023). On the accept–reject mechanism for Metropolis–Hastings algorithms. *Ann. Appl. Probab.* **33** 5279–5333.
[25] Hairer, M., Stuart, A.M. and Vollmer, S.J. (2014). Spectral gaps for a Metropolis–Hastings algorithm in infinite dimensions. *Ann. Appl. Probab.* **24** 2455–2490.
[26] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications.
[27] Hult, H. and Nyquist, P. (2016). Large deviations for weighted empirical measures arising in importance sampling. *Stochastic Process. Appl.* **126** 138–170.
[28] Jarner, S.F. and Hansen, E. (2000). Geometric ergodicity of Metropolis algorithms. *Stochastic Process. Appl.* **85** 341–361.
[29] Jourdain, B., Leliévre, T. and Miasojedow, B. (2014). Optimal scaling for the transient phase of Metropolis Hastings algorithms: The longtime behavior. *Bernoulli* **20** 1930–1978.
[30] Kontoyiannis, I. and Meyn, S.P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.* **13** 304–362.
[31] Kontoyiannis, I. and Meyn, S.P. (2005). Large deviations asymptotics and the spectral theory of multiplicatively regular Markov processes. *Electron. J. Probab.* **10** 61–123.
[32] Mattingly, J.C., Pillai, N.S. and Stuart, A.M. (2012). Diffusion limits of the random walk Metropolis algorithm in high dimensions. *Ann. Appl. Probab.* **22** 881–930.
[33] Mengersen, K.L. and Tweedie, R.L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.* **24** 101–121.
[34] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21** 1087–1092.
[35] Meyn, S.P. and Tweedie, R.L. (2009). Markov Chains and Stochastic Stability.
[36] Milinanni, F. (2024). Large deviation-based tuning schemes for Metropolis-Hastings algorithms. Preprint, arXiv:2409.20337.
[37] Milinanni, F. and Nyquist, P. (2024). A large deviation principle for the empirical measures of Metropolis–Hastings chains. *Stochastic Process. Appl.* **170**.
[38] Plattner, N., Doll, J.D., Dupuis, P., Wang, H., Liu, Y. and Gubernatis, J.E. (2011). An infinite swapping approach to the rare-event sampling problem. *J. Chem. Phys.* **135** 134111.
[39] Power, S., Rudolf, D., Sprungk, B. and Wang, A.Q. (2024). Weak Poincaré inequality comparisons for ideal and hybrid slice sampling.
[40] Rey-Bellet, L. and Spiliopoulos, K. (2015). Irreversible Langevin samplers and variance reduction: A large deviations approach. *Nonlinearity* **28**. 2081.

[41] Rey-Bellet, L. and Spiliopoulos, K. (2016). Improving the convergence of reversible samplers. *J. Stat. Phys.* **164** 472–494.
[42] Robert, C.P. and Casella, G. (2004). *Monte Carlo Statistical Methods*, 2nd ed. *Springer Texts in Statistics*. New York, NY: Springer.
[43] Roberts, G.O. (2003). Linking theory and practice of MCMC. Oxford Statistical Science Series 145–166.
[44] Roberts, G.O. and Rosenthal, J.S. (1997). Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.* **2** 13–25.
[45] 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.
[46] Roberts, G.O. and Rosenthal, J.S. (2001). Optimal scaling for various Metropolis-Hastings algorithms. *Statist. Sci.* **16** 351–367.
[47] Roberts, G.O. and Tweedie, R.L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* 341–363.
[48] Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika* **83** 95–110.
[49] Rosenthal, J.S. (2003). Asymptotic variance and convergence rates of nearly-periodic Markov chain Monte Carlo algorithms. *J. Amer. Statist. Assoc.* **98** 169–177.
[50] Roy, V. and Zhang, L. (2023). Convergence of position-dependent MALA with application to conditional simulation in GLMMs. *J. Comput. Graph. Statist.* **32** 501–512.
[51] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1728.
[52] Tierney, L. (1998). A note on Metropolis-Hastings kernels for general state spaces. *Ann. Appl. Probab.* **8** 1–9.
[53] Wang, G. (2022). Exact convergence analysis of the independent Metropolis-Hastings algorithms. *Bernoulli* **28** 2012–2033.

Bernoulli **32**(4), 2026, 2999–3023
https://doi.org/10.3150/26-BEJ1976

Spectral independence of almost fully correlated random matrices

OLEKSII KOLUPAIEV[a]

Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria,
[a]*Oleksii.Kolupaiev@ist.ac.at*

We study the joint spectral properties of two coupled random matrices $H^{(1)}$ and $H^{(2)}$, which are either real symmetric or complex Hermitian. The entries of these matrices exhibit polynomially decaying correlations, both within each matrix and between them. Surprisingly, we find that under extremely weak decorrelation condition, permitting $H^{(1)}$ and $H^{(2)}$ to be almost fully correlated, the fluctuations of their individual eigenvalues in the bulk of the spectrum are still asymptotically independent. Furthermore, we demonstrate that this decorrelation condition is optimal.

Keywords: Correlated random matrix; eigenvalue perturbation theory; local law; Wigner-type matrix

References

Ajanki, O., Erdős, L. and Krüger, T. (2015). Universality for general Wigner-type matrices. *Probab. Theory Related Fields* **169** 667–727.

Ajanki, O.H., Erdős, L. and Krüger, T. (2016). Local spectral statistics of Gaussian matrices with correlated entries. *J. Stat. Phys.* **163** 280–302.

Alt, J., Erdős, L. and Krüger, T. (2020). The Dyson equation with linear self-energy: Spectral bands, edges and cusps. *Doc. Math.* **25** 1421–1539.

Bao, Z., Cipolloni, G., Erdős, L., Henheik, J. and Kolupaiev, O. (2025a). Decorrelation transition in the Wigner minor process. *Probab. Theory Related Fields* 1–52. https://doi.org/10.1007/s00440-025-01422-4

Bao, Z., Cipolloni, G., Erdős, L., Henheik, J. and Kolupaiev, O. (2025b). Law of fractional logarithm for random matrices. Preprint. Available at arXiv:2503.18922.

Biane, P. (1997). On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46** 705–718.

Bourgade, P. and Falconet, H. (2022). Liouville quantum gravity from random matrix dynamics. Preprint. Available at arXiv:2206.03029.

Cipolloni, G. and Erdős, L. (2020). Fluctuations for differences of linear eigenvalue statistics for sample covariance matrices. *Random Matrices Theory Appl.* **9**.

Cipolloni, G., Erdős, L. and Schröder, D. (2023). Central limit theorem for linear eigenvalue statistics of non-Hermitian random matrices. *Comm. Pure Appl. Math.* **76** 899–1136.

Cipolloni, G., Erdős, L. and Schröder, D. (2024). Mesoscopic central limit theorem for non-Hermitian random matrices. *Probab. Theory Related Fields* **188** 1131–1182. MR4716346 https://doi.org/10.1007/s00440-023-01229-1

Cipolloni, G., Erdős, L., Henheik, J. and Kolupaiev, O. (2024). Eigenvector decorrelation for random matrices. Available at arXiv:2410.10718.

Erdős, L. (2019). *The Matrix Dyson Equation and Its Applications for Random Matrices*. *IAS/Park City Mathematics Series* **26** 75–158.

Erdős, L., Henheik, J. and Riabov, V. (2025). Cusp universality for correlated random matrices. *Comm. Math. Phys.* **406**.

Erdős, L., Krüger, T. and Schröder, D. (2019). Random matrices with slow correlation decay. *Forum Math. Sigma* **7**.

1350-7265

Erdős, L. and Yau, H.T. (2017). *A Dynamical Approach to Random Matrix Theory* **28**. American Mathematical Soc.

Erdős, L., Knowles, A., Yau, H.T. and Yin, J. (2013). The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** 1–58.

Eynard, B. (1997). Eigenvalue distribution of large random matrices, from one matrix to several coupled matrices. *Nuclear Phys. B* **506** 633–664.

Eynard, B. (1998). Correlation functions of eigenvalues of multi-matrix models, and the limit of a time-dependent matrix. *J. Phys. A: Math. Gen.* **31** 8081–8102.

Figalli, A. and Guionnet, A. (2016). Universality in several-matrix models via approximate transport maps. *Acta Math.* **217** 81–176.

Filev, V.G. and O'Connor, D. (2013). Multi-matrix models at general coupling. *J. Phys. A: Math. Theor.* **46**.

Forrester, P.J. and Nagao, T. (2011). Determinantal correlations for classical projection processes. *J. Stat. Mech. Theory Exp.* **2011** P08011.

Gorin, V. and Shkolnikov, M. (2015). Multilevel Dyson Brownian motions via Jack polynomials. *Probab. Theory Related Fields* **163** 413–463. MR3418747 https://doi.org/10.1007/s00440-014-0596-2

Gorin, V. and Shkolnikov, M. (2017). Interacting particle systems at the edge of multilevel Dyson Brownian motions. *Adv. Math.* **304** 90–130. MR3558206 https://doi.org/10.1016/j.aim.2016.08.034

Jana, I. and Rani, S. (2025). CLT for LES of correlated Non-Hermitian Random Matrices. Preprint. Available at arXiv:2503.22542.

Landon, B., Sosoe, P. and Yau, H.-T. (2019). Fixed energy universality of Dyson Brownian motion. *Adv. Math.* **346** 1137–1332.

Schnelli, K. and Xu, Y. (2023). Quantitative Tracy–Widom laws for the largest eigenvalue of generalized Wigner matrices. *Electron. J. Probab.* **28** 1–38.

Bernoulli **32**(4), 2026, 3024–3041
https://doi.org/10.3150/26-BEJ1977

Central limit theorems for random boundary polytopes

MATTHIAS REITZNER[1,a] and MATHIAS SONNLEITNER[2,b]

[1]*Institute of Mathematics, University of Osnabrück, Osnabrück, Germany,*
[a]*matthias.reitzner@uni-osnabrueck.de*
[2]*Institute of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz, Linz, Austria,*
[b]*mathias.sonnleitner@jku.at*

The number of faces of the convex hull of n independent and identically distributed random points chosen on the boundary of a smooth convex body in $\mathbb{R}^d$ is investigated. In dimensions two and three the number of k-faces is known to be constant almost surely and in dimension four and higher the variance is known to be non-zero if $k \geq 1$. We show that it is of order n. This is complemented by a central limit theorem with a Berry-Esseen bound which is of optimal order $n^{-1/2}$. We derive similar results for the Poissonized model, where additionally the number of random points is Poisson distributed. As a main tool, we develop a representation of the number of faces as a sum of exponentially stabilizing score functions.

Keywords: Berry-Esseen bound; central limit theorem; random approximation; random polytope; stochastic geometry; variance expansion

References

Bárány, I. (1992). Random polytopes in smooth convex bodies. *Mathematika* **39** 81–92.
Bárány, I. (2004). Corrigendum: Random polytopes in smooth convex bodies. *Mathematika* **51** 31.
Bárány, I. and Buchta, C. (1993). Random polytopes in a convex polytope, independence of shape, and concentration of vertices. *Math. Ann.* **297** 467–497.
Bárány, I. and Reitzner, M. (2010a). On the variance of random polytopes. *Adv. Math.* **225** 1986–2001.
Bárány, I. and Reitzner, M. (2010b). Poisson polytopes. *Ann. Probab.* **38** 1507–1531.
Bárány, I. and Steiger, W. (2013). On the variance of random polygons. *Comput. Geom.* **46** 173–180.
Besau, F., Rosen, D. and Thäle, C. (2021). Random inscribed polytopes in projective geometries. *Math. Ann.* **381** 1345–1372.
Böröczky, K. Jr., Fodor, F. and Hug, D. (2013). Intrinsic volumes of random polytopes with vertices on the boundary of a convex body. *Trans. Amer. Math. Soc.* **365** 785–809.
Böröczky, K. Jr., Hoffmann, L.M. and Hug, D. (2008). Expectation of intrinsic volumes of random polytopes. *Period. Math. Hungar.* **57** 143–164.
Buchta, C., Müller, J.S. and Tichy, R.F. (1985). Stochastical approximation of convex bodies. *Math. Ann.* **271** 225–235.
Calka, P., Schreiber, T. and Yukich, J.E. (2013). Brownian limits, local limits and variance asymptotics for convex hulls in the ball. *Ann. Probab.* **41** 50–108.
Calka, P. and Yukich, J.E. (2014). Variance asymptotics for random polytopes in smooth convex bodies. *Probab. Theory Related Fields* **158** 435–463.
Calka, P. and Yukich, J.E. (2017). Variance asymptotics and scaling limits for random polytopes. *Adv. Math.* **304** 1–55.
Efron, B. (1965). The convex hull of a random set of points. *Biometrika* **52** 331–343.
Englund, G. (1981). A remainder term estimate for the normal approximation in classical occupancy. *Ann. Probab.* **9** 684–692.
Groeneboom, P. (1988). Limit theorems for convex hulls. *Probab. Theory Related Fields* **79** 327–368.

1350-7265

Grünbaum, B. (2003). *Convex Polytopes*, 2nd ed. *Graduate Texts in Mathematics*. New York, NY: Springer.

Grygierek, J. (2021). Multivariate normal approximation for functionals of random polytopes. *J. Theoret. Probab.* **34** 897–922.

Gusakova, A., Reitzner, M. and Thäle, C. (2023). Variance expansion and Berry-Esseen bound for the number of vertices of a random polygon in a polygon. *Ann. Henri Lebesgue* **6** 875–906.

Herold, F., Hug, D. and Thäle, C. (2021). Does a central limit theorem hold for the k-skeleton of Poisson hyperplanes in hyperbolic space? *Probab. Theory Related Fields* **179** 889–968.

Kabluchko, Z. (2021). Angles of random simplices and face numbers of random polytopes. *Adv. Math.* **380** 107612.

Kabluchko, Z. (2023). On expected face numbers of random beta and beta′ polytopes. *Beitr. Algebra Geom.* **64** 155–174.

Kabluchko, Z., Rosen, D. and Thäle, C. (2025). Fluctuations of λ-geodesic Poisson hyperplanes in hyperbolic space. *Israel J. Math.* 1–42.

Kabluchko, Z., Thäle, C. and Zaporozhets, D. (2020). Beta polytopes and Poisson polyhedra: f-vectors and angles. *Adv. Math.* **374** 107333.

Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993.

Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks 7*. Cambridge: Cambridge University Press.

Leroux, B., Rademacher, L., Schütt, C. and Werner, E.M. (2025). Expected extremal area of facets of random polytopes. *Math. Ann.* **392** 2657–2708.

Pardon, J. (2011). Central limit theorems for random polygons in an arbitrary convex set. *Ann. Probab.* **39** 881–903.

Pardon, J. (2012). Central limit theorems for uniform model random polygons. *J. Theoret. Probab.* **25** 823–833.

Prochno, J., Schütt, C., Sonnleitner, M. and Werner, E.M. (2025). Random approximation of convex bodies in Hausdorff metric. *Math. Ann.* **392** 4525–4542. https://doi.org/10.1007/s00208-025-03186-7

Reitzner, M. (2002). Random points on the boundary of smooth convex bodies. *Trans. Amer. Math. Soc.* **354** 2243–2278.

Reitzner, M. (2003). Random polytopes and the Efron-Stein jackknife inequality. *Ann. Probab.* **31** 2136–2166.

Reitzner, M. (2004). Stochastical approximation of smooth convex bodies. *Mathematika* **51** 11–29.

Reitzner, M. (2005a). The combinatorial structure of random polytopes. *Adv. Math.* **191** 178–208.

Reitzner, M. (2005b). Central limit theorems for random polytopes. *Probab. Theory Related Fields* **133** 483–507.

Reitzner, M., Schütt, C. and Werner, E.M. (2023). The convex hull of random points on the boundary of a simple polytope. *Discrete Comput. Geom.* **69** 453–504.

Rényi, A. and Sulanke, R. (1963). Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrsch. Verw. Gebiete* **2** 75–84.

Rényi, A. and Sulanke, R. (1964). Über die konvexe Hülle von n zufällig gewählten Punkten II. *Z. Wahrsch. Verw. Gebiete* **3** 138–147.

Richardson, R.M., Vu, V.H. and Wu, L. (2008). An inscribing model for random polytopes. *Discrete Comput. Geom.* **39** 469–499.

Schneider, R. (1988). Random approximation of convex sets. *J. Microsc.* **151** 211–227.

Schneider, R. (2014). *Convex Bodies: The Brunn-Minkowski Theory*, 2nd ed. Cambridge: Cambridge University Press.

Schneider, R. and Wieacker, J.A. (1980). Random polytopes in a convex body. *Z. Wahrsch. Verw. Gebiete* **52** 69–73.

Schütt, C. (1994). Random polytopes and affine surface area. *Math. Nachr.* **170** 227–249.

Schütt, C. and Werner, E. (2003). Polytopes with vertices chosen randomly from the boundary of a convex body. In *Geometric Aspects of Functional Analysis (2001–2002)* (V.D. Milman and G. Schechtman, eds.). *Lecture Notes in Mathematics* **1807** 241–422. Springer.

Stemeseder, J. (2014). Random polytopes with vertices on the sphere. PhD thesis, Universität Salzburg.

Thäle, C. (2018). Central limit theorem for the volume of random polytopes with vertices on the boundary. *Discrete Comput. Geom.* **59** 990–1000.

Thäle, C., Turchi, N. and Wespi, F. (2018). Random polytopes: Central limit theorems for intrinsic volumes. *Proc. Amer. Math. Soc.* **146** 3063–3071.

Turchi, N. and Wespi, F. (2018). Limit theorems for random polytopes with vertices on convex surfaces. *Adv. Appl. Probab.* **50** 1227–1245.

Wieacker, J.A. (1978). Einige Probleme der polyedrischen Approximation, Diplomarbeit, Universität Freiburg.

Bernoulli **32**(4), 2026, 3042–3062
https://doi.org/10.3150/26-BEJ1978

Finite-dimensional Gaussian approximation for deep neural networks: Universality in random weights

KRISHNAKUMAR BALASUBRAMANIAN[1,a] and NATHAN ROSS[2,b]

[1]*Department of Statistics, University of California, Davis, United States of America,* [a]*kbala@ucdavis.edu*
[2]*School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia,*
[b]*nathan.ross@unimelb.edu.au*

We study the Finite-Dimensional Distributions (FDDs) of deep neural networks with randomly initialized weights that have finite-order moments. Specifically, we establish Gaussian approximation bounds in the Wasserstein-1 norm between the FDDs and their Gaussian limit assuming a Lipschitz activation function and allowing the layer widths to grow to infinity at arbitrary relative rates. In the special case where all widths are proportional to a common scale parameter n and there are $L-1$ hidden layers, we obtain convergence rates of order $n^{-(1/6)^{L-1}+\epsilon}$, for any $\epsilon > 0$.

Keywords: Deep neural networks; non-Gaussian weights; Stein's Method

References

Apollonio, N., Franzina, G. and Torrisi, G.L. (2025). Simulating Posterior Bayesian Neural Networks with Dependent Weights. arXiv preprint. Available at arXiv:2507.22095.

Apollonio, N., De Canditiis, D., Franzina, G., Stolfi, P. and Torrisi, G.L. (2025). Normal approximation of random Gaussian neural networks. *Stoch. Syst.* **15** 88–110.

Balasubramanian, K., Goldstein, L., Ross, N. and Salim, A. (2024). Gaussian random field approximation via Stein's method with applications to wide random neural networks. *Appl. Comput. Harmon. Anal.* **72** 27. Paper No. 101668.

Basteri, A. and Trevisan, D. (2024). Quantitative Gaussian approximation of randomly initialized deep neural networks. *Mach. Learn.* **113** 6373–6393.

Bordino, A., Favaro, S. and Fortini, S. (2023). Infinitely wide limits for deep Stable neural networks: Sub-linear, linear and super-linear activation functions. arXiv preprint. Available at arXiv:2304.04008.

Bordino, A., Favaro, S. and Fortini, S. (2024). Non-asymptotic approximations of Gaussian neural networks via second-order Poincaré inequalities. In *Proceedings of the 6th Symposium on Advances in Approximate Bayesian Inference* **1** 34.

Cammarota, V., Marinucci, D., Salvi, M. and Vigogna, S. (2024). A quantitative functional central limit theorem for shallow neural networks. *Mod. Stoch. Theory Appl.* **11** 85–108.

Carvalho, L., Costa, J.L., Mourão, J. and Oliveira, G. (2023). Wide neural networks: From non-Gaussian random fields at initialization to the NTK geometry of training. arXiv preprint. Available at arXiv:2304.03385.

Celli, L. and Peccati, G. (2025). Entropic bounds for conditionally Gaussian vectors and applications to neural networks. arXiv preprint. Available at arXiv:2504.08335.

Chatterjee, S. and Meckes, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Amer. J. Probab. Math. Stat.* **4** 257–283.

Chen, L.H., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. Springer.

de Matthews, A.G., Hron, J., Rowland, M., Turner, R.E. and Ghahramani, Z. (2018). Gaussian process behaviour in wide deep neural networks. In *International Conference on Learning Representations*.

1350-7265

Der, R. and Lee, D. (2005). Beyond Gaussian processes: On the distributions of infinite networks. In *Advances in Neural Information Processing Systems* **18**.

Eldan, R., Mikulincer, D. and Schramm, T. (2021). Non-asymptotic approximations of neural networks by Gaussian processes. In *Conference on Learning Theory* 1754–1775. PMLR.

Favaro, S., Fortini, S. and Peluchetti, S. (2023). Deep stable neural networks: Large-width asymptotics and convergence rates. *Bernoulli* **29** 2574–2597.

Favaro, S., Hanin, B., Marinucci, D., Nourdin, I. and Peccati, G. (2025). Quantitative CLTs in deep neural networks. *Probab. Theory Related Fields* **191** 933–977.

Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics* 249–256.

Goldstein, L. and Rinott, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17.

Golikov, E. and Yang, G. (2022). Non-Gaussian tensor programs. In *Advances in Neural Information Processing Systems* **35**.

Götze, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724–739.

Hanin, B. (2023). Random neural networks in the infinite width limit as Gaussian processes. *Ann. Appl. Probab.* **33** 4798–4819.

Jacot, A., Gabriel, F. and Hongler, C. (2018). Neural tangent kernel: Convergence and generalization in neural networks. In *Advances in Neural Information Processing Systems* **31**.

Jung, P., Lee, H., Lee, J. and Yang, H. (2023). α-Stable convergence of heavy/light-tailed infinitely wide neural networks. *Adv. Appl. Probab.* **55** 1415–1441.

Klukowski, A. (2022). Rate of convergence of polynomial networks to Gaussian processes. In *Conference on Learning Theory* 701–722. PMLR.

Lee, J., Sohl Dickstein, J., Pennington, J., Novak, R., Schoenholz, S. and Bahri, Y. (2018). Deep neural networks as Gaussian processes. In *International Conference on Learning Representations*.

Lee, H., Ayed, F., Jung, P., Lee, J., Yang, H. and Caron, F. (2023). Deep neural networks with dependent weights: Gaussian process mixture limit, heavy tails, sparsity and compressibility. *J. Mach. Learn. Res.* **24** 1–78.

Li, H., De, S., Xu, Z., Studer, C., Samet, H. and Goldstein, T. (2017). Training quantized nets: A deeper understanding. In *Advances in Neural Information Processing Systems* **30**.

Loría, J. and Bhadra, A. (2023). Posterior inference on shallow infinitely wide Bayesian neural networks under weights with unbounded variance. arXiv preprint. Available at arXiv:2305.10664.

Meckes, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V: The Luminy Volume*. *Inst. Math. Stat. (IMS) Collect.* **5** 153–178. Beachwood, OH: Inst. Math. Statist.

Neal, R.M. (1996). Priors for infinite networks. In *Bayesian Learning for Neural Networks* 29–53. Springer.

Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality 192*. Cambridge University Press.

Raič, M. (2018). A multivariate central limit theorem for Lipschitz and smooth test functions. Preprint. Available at arXiv:1812.08268.

Reinert, G. and Röllin, A. (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** 2150–2173.

Soto, T. (2024). Wide stable neural networks: Sample regularity, functional convergence and Bayesian inverse problems. arXiv preprint. Available at arXiv:2407.03909.

Trevisan, D. (2023). Wide deep neural networks with Gaussian weights are very close to Gaussian processes. arXiv preprint. Available at arXiv:2312.11737.

Bernoulli **32**(4), 2026, 3063–3087
https://doi.org/10.3150/26-BEJ1979

Asymptotics of resampling without replacement in robust and logistic regression

PIERRE C. BELLEC[1,a] and TAKUYA KORIYAMA[2,b]

[1]*Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA,* [a]*pierre.bellec@rutgers.edu*
[2]*Booth School of Business, University of Chicago, Chicago, IL 60637, USA,* [b]*tkoriyam@uchicago.edu*

This paper studies the asymptotics of resampling without replacement in the proportional regime where dimension p and sample size n are of the same order. For a given dataset $(\boldsymbol{X},\boldsymbol{y}) \in \mathbb{R}^{n\times p} \times \mathbb{R}^n$ and fixed subsample ratio $q \in (0,1)$, the practitioner samples independently of $(\boldsymbol{X},\boldsymbol{y})$ iid subsets $I_1,\dots,I_M$ of $\{1,\dots,n\}$ of size qn and trains estimators $\hat{\boldsymbol{b}}(I_1),\dots,\hat{\boldsymbol{b}}(I_M)$ on the corresponding subsets of rows of $(\boldsymbol{X},\boldsymbol{y})$. Understanding the performance of the bagged estimate $\bar{\boldsymbol{b}} = M^{-1}\sum_{m=1}^M \hat{\boldsymbol{b}}(I_m)$, for instance its squared error, requires us to understand correlations between two distinct $\hat{\boldsymbol{b}}(I_m)$ and $\hat{\boldsymbol{b}}(I_{m'})$ trained on different subsets I_m and $I_{m'}$. In robust linear regression and logistic regression, we characterize the limit in probability of the correlation between two estimates trained on different subsets of the data. The limit is characterized as the unique solution of a simple nonlinear equation. We further provide data-driven estimators that are consistent for estimating this limit. These estimators of the limiting correlation allow us to estimate the squared error of the bagged estimate $\bar{\boldsymbol{b}}$, and for instance perform parameter tuning to choose the optimal subsample ratio q. As a by-product of the proof argument, we obtain the limiting distribution of the bivariate pair $(\boldsymbol{x}_i^T\hat{\boldsymbol{b}}(I_m), \boldsymbol{x}_i^T\hat{\boldsymbol{b}}(I_{m'}))$ for observations $i \in I_m \cap I_{m'}$, i.e., for observations used to train both estimates.

Keywords: Bootstrap aggregating; high dimensional asymptotics; proportional regime

References

Araki, H. and Yamagami, S. (1981). An inequality for Hilbert-Schmidt norm. *Comm. Math. Phys.* **81** 89–96.

Bayati, M. and Montanari, A. (2011). The LASSO risk for Gaussian matrices. *IEEE Trans. Inf. Theory* **58** 1997–2017.

Bellec, P.C. (2023). Out-of-sample error estimation for M-estimators with convex penalty. *Inf. Inference* **12** 2782–2817.

Bellec, P.C. (2025). Observable adjustments in single-index models for regularized M-estimators with bounded p/n. *Ann. Statist.* **53** 531–560.

Bellec, P.C. and Koriyama, T. (2023). Existence of solutions to the nonlinear equations characterizing the precise error of M-estimators. arXiv preprint arXiv:2312.13254.

Bellec, P.C. and Koriyama, T. (2025). Error estimation and adaptive tuning for unregularized robust M-estimator. *J. Mach. Learn. Res.* **26** 1–40.

Bellec, P.C. and Koriyama, T. (2026). Supplement to "Asymptotics of resampling without replacement in robust and logistic regression." https://doi.org/10.3150/26-BEJ1979SUPP

Bellec, P.C. and Shen, Y. (2022). Derivatives and residual distribution of regularized M-estimators with application to adaptive tuning. In *Proceedings of the Conference on Learning Theory*. *Proceedings of Machine Learning Research* **178** 1912–1947. PMLR.

Bellec, P.C. and Zhang, C.-H. (2021). Second-order Stein: SURE for SURE and other applications in high-dimensional inference. *Ann. Statist.* **49** 1864–1903.

Bellec, P.C. and Zhang, C.-H. (2023). Debiasing convex regularized estimators and interval estimation in linear models. *Ann. Statist.* **51** 391–436.

Bellec, P.C., Du, J.-H., Koriyama, T., Patil, P. and Tan, K. (2025). Corrected generalized cross-validation for finite ensembles of penalized estimators. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **87** 289–318.

Bhatia, R. (2013). *Matrix Analysis 169*. Springer Science & Business Media.

Bogachev, V.I. (1998). *Gaussian Measures* **62**. American Mathematical Soc.

Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford university press.

Breiman, L. (1996). Bagging predictors. *Mach. Learn.* **24** 123–140.

Breiman, L. (2001). Using iterated bagging to debias regressions. *Mach. Learn.* **45** 261–277.

Bühlmann, P. and Yu, B. (2002). Analyzing bagging. *Ann. Statist.* **30** 927–961.

Candès, E.J. and Sur, P. (2020). The phase transition for the existence of the maximum likelihood estimate in high-dimensional logistic regression. *Ann. Statist.* **48** 27–42.

Chaudhuri, A. (2014). *Modern Survey Sampling*. CRC Press.

Chun-hui, C. and Ji-guang, S. (1989). Perturbation bounds for the polar factors. *J. Comput. Math.* 397–401.

Clarté, L., Vandenbroucque, A., Dalle, G., Loureiro, B., Krzakala, F. and Zdeborová, L. (2024). Analysis of bootstrap and subsampling in high-dimensional regularized regression. arXiv preprint arXiv:2402.13622.

Davidson, K.R. and Szarek, S.J. (2001). Local operator theory, random matrices and Banach spaces. *Handb. Geom. Banach Spaces* **1** 131.

Donoho, D. and Montanari, A. (2016). High dimensional robust M-estimation: Asymptotic variance via approximate message passing. *Probab. Theory Related Fields* **166** 935–969.

Du, J.-H., Patil, P. and Kuchibhotla, A.K. (2023). Subsample ridge ensembles: Equivalences and generalized cross-validation. In *International Conference on Machine Learning*.

Edelman, A. (1988). Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.* **9** 543–560.

El Karoui, N. (2013). Asymptotic behavior of unregularized and ridge-regularized high-dimensional robust regression estimators: rigorous results. arXiv preprint 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.

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** 1–66.

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.

Koriyama, T., Patil, P., Du, J.-H., Tan, K. and Bellec, P.C. (2026). Precise asymptotics of bagging regularized M-estimators. *Ann. Statist.* forthcoming.

Krogh, A. and Sollich, P. (1997). Statistical mechanics of ensemble learning. *Phys. Rev. E* **55** 811–825.

LeJeune, D., Javadi, H. and Baraniuk, R. (2020). The implicit regularization of ordinary least squares ensembles. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* 3525–3535. PMLR.

Loureiro, B., Gerbelot, C., Cui, H., Goldt, S., Krzakala, F., Mezard, M. and Zdeborová, L. (2021). Learning curves of generic features maps for realistic datasets with a teacher-student model. In *In Advances in Neural Information Processing Systems* **34** 18137–18151. Curran Associates, Inc.

Loureiro, B., Gerbelot, C., Refinetti, M., Sicuro, G. and Krzakala, F. (2022). Fluctuations, bias, variance & ensemble of learners: Exact asymptotics for convex losses in high-dimension. In *Proceedings of the 39th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* **162** 14283–14314. PMLR.

Patil, P., Du, J.-H. and Kuchibhotla, A.K. (2023). Bagging in overparameterized learning: Risk characterization and risk monotonization. *J. Mach. Learn. Res.* **24** 1–113.

Salehi, F., Abbasi, E. and Hassibi, B. (2019). *The Impact of Regularization on High-Dimensional Logistic Regression. In Advances in Neural Information Processing Systems* **32**. Curran Associates, Inc.

Sollich, P. and Krogh, A. (1995). Learning with ensembles: How overfitting can be useful. In *Advances in Neural Information Processing Systems* (D. Touretzky, M.C. Mozer and M. Hasselmo, eds.) **8**. MIT Press.

Sur, P. and Candès, E.J. (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proc. Natl. Acad. Sci. USA* **116** 14516–14525.

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.

van Hemmen, J.L. and Ando, T. (1980). An inequality for trace ideals. *Comm. Math. Phys.* **76** 143–148.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science 47*. Cambridge university press.

Zhao, Q., Sur, P. and Candes, E.J. (2022). The asymptotic distribution of the MLE in high-dimensional logistic models: Arbitrary covariance. *Bernoulli* **28** 1835–1861.

Bernoulli **32**(4), 2026, 3088–3112
https://doi.org/10.3150/26-BEJ1980

Theoretical analysis of transfer learning for temporally dependent observations

ABOLFAZL SAFIKHANI[1,a] and MINGLIANG MA[2,b]

[1]*Department of Statistics, George Mason University, Fairfax, US,* [a]*asafikha@gmu.edu*
[2]*Department of Statistics, University of Florida, Gainesville, US,* [b]*maminglian@ufl.edu*

Transfer learning has become increasingly important in recent years as it enables models to quickly adapt to new tasks, environments and data sets. It can improve the accuracy of machine learning models and reduce the training time required for them. However, theoretical foundation of transfer learning algorithms is rather scarce, particularly for time series models. In this work, we focus on high-dimensional vector autoregressive models and provide a two-step procedure to conduct transfer learning utilizing auxiliary data sets followed by constructing confidence intervals for model parameters in the high-dimensional regime. Further, a new method to select informative sets from auxiliary sets is introduced. Finally, two debiasing techniques are developed to perform inference for model parameters. Theoretical properties of all proposed algorithms are established under mild conditions which allow for heavy-tailed distributions and dependence among auxiliary and target data sets. Given certain level of similarity between the informative models and the target model, it is shown that the proposed algorithm achieves the minimax rate. Lastly, the empirical performance of proposed methods is tested through analyzing both simulated data as well as an EEG data set.

Keywords: High-dimensional data; inference; temporal dependence; transfer learning; vector autoregressive model

References

Bai, Y. and Safikhani, A. (2023). A unified framework for change point detection in high-dimensional linear models. *Statist. Sinica* **33** 1721–1748.
Bai, P., Safikhani, A. and Michailidis, G. (2023). Multiple change point detection in reduced rank high dimensional vector autoregressive models. *J. Amer. Statist. Assoc.* **118** 2776–2792.
Bastani, H. (2021). Predicting with proxies: Transfer learning in high dimension. *Manag. Sci.* **67** 2964–2984.
Basu, S. and Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567.
Cai, T.T. and Wei, H. (2021). Transfer learning for nonparametric classification: Minimax rate and adaptive classifier. *Ann. Statist.* **49** 100–128.
Cai, C., Wang, S., Xu, Y., Zhang, W., Tang, K., Ouyang, Q., Lai, L. and Pei, J. (2020). Transfer learning for drug discovery. *J. Med. Chem.* **63** 8683–8694.
de Cheveigné, A. and Arzounian, D. (2018). Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. *NeuroImage* **172** 903–912.
Dedecker, J. and Prieur, C. (2004). Coupling for τ-dependent sequences and applications. *J. Theor. Probab.* **17** 861–885.
Deshpande, Y., Javanmard, A. and Mehrabi, M. (2021). Online debiasing for adaptively collected high-dimensional data with applications to time series analysis. *J. Amer. Statist. Assoc.* 1–14.
Ding, X. and Zhou, Z. (2020). Estimation and inference for precision matrices of nonstationary time series. *Ann. Statist.*
Ding, X. and Zhou, Z. (2023). AutoRegressive approximations to nonstationary time series with inference and applications. *Ann. Statist.* **51** 1207–1231.
Farhadi, A., Forsyth, D. and White, R. (2007). Transfer learning in sign language. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit* 1–8. https://doi.org/10.1109/CVPR.2007.383346

Gong, B., Shi, Y., Sha, F. and Grauman, K. (2012). Geodesic flow kernel for unsupervised domain adaptation. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit* 2066–2073. IEEE.

Hall, P. and Heyde, C.C. (2014). *Martingale Limit Theory and Its Application*. Academic press.

Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909.

Kulis, B., Saenko, K. and Darrell, T. (2011). What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In *CVPR 2011* 1785–1792. IEEE.

Li, S., Cai, T.T. and Li, H. (2022a). Transfer learning for high-dimensional linear regression: Prediction, estimation and minimax optimality. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 149–173.

Li, S., Cai, T.T. and Li, H. (2022b). Transfer learning in large-scale Gaussian graphical models with false discovery rate control. *J. Amer. Statist. Assoc.* 1–13.

Li, Z., Liu, D., He, Y. and Zhang, X. (2023). The Role of Fine-tuning: Transfer Learning for High-dimensional M-estimators with Decomposable Regularizers. arXiv e-prints. arXiv-2306.

Loh, P.-L. and Wainwright, M.J. (2011). High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. *Adv. Neural Inf. Process. Syst.* **24**.

Lounici, K., Pontil, M., Tsybakov, A.B. and Van De Geer, S. (2009). Taking advantage of sparsity in multi-task learning. arXiv preprint. Available at arXiv:0903.1468.

Ma, M. and Safikhani, A. (2022). Theoretical analysis of deep neural networks for temporally dependent observations. *Adv. Neural Inf. Process. Syst.*

Nezamfar, H., Orhan, U., Purwar, S., Hild, K., Oken, B. and Erdogmus, D. (2011). Decoding of multichannel EEG activity from the visual cortex in response to pseudorandom binary sequences of visual stimuli. *Int. J. Imaging Syst. Technol.* **21** 139–147.

Olivas, E.S., Guerrero, J.D.M., Martinez-Sober, M., Magdalena-Benedito, J.R., Serrano, L. et al. (2009). *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques: Algorithms, Methods, and Techniques*. IGI global.

Safikhani, A. and Ma, M. (2026). Supplement to "Theoretical analysis of transfer learning for temporally dependent observations." https://doi.org/10.3150/26-BEJ1980SUPP

Tian, Y. and Feng, Y. (2023). Transfer learning under high-dimensional generalized linear models. *J. Amer. Statist. Assoc.* **118** 2684–2697.

Trujillo, L.T., Stanfield, C.T. and Vela, R.D. (2017). The effect of electroencephalogram (EEG) reference choice on information-theoretic measures of the complexity and integration of EEG signals. *Front. Neurosci.* **11** 425.

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.*

Wan, Z., Yang, R., Huang, M., Zeng, N. and Liu, X. (2021). A review on transfer learning in EEG signal analysis. *Neurocomputing* **421** 1–14.

Wang, J. and Politis, D.N. (2021). Consistent autoregressive spectral estimates: Nonlinear time series and large autocovariance matrices. *J. Time Series Anal.* **42** 580–596.

Weiss, K., Khoshgoftaar, T.M. and Wang, D. (2016). A survey of transfer learning. *J. Big Data* **3** 1–40.

Wong, K.C., Li, Z. and Tewari, A. (2020). Lasso guarantees for β-mixing heavy-tailed time series. *Ann. Statist.* **48** 1124–1142. https://doi.org/10.1214/19-AOS1840

Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154.

Yousuf, K. (2018). Variable screening for high dimensional time series. *Electron. J. Stat.* **12** 667–702.

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.

Zhang, W., Deng, L., Zhang, L. and Wu, D. (2022). A survey on negative transfer. *IEEE/CAA J. Autom. Sin.*

Zheng, C., Dasgupta, S., Xie, Y., Haris, A. and Chen, Y.Q. (2019). On data enriched logistic regression. arXiv preprint. Available at arXiv:1911.06380.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H. and He, Q. (2020). A comprehensive survey on transfer learning. *Proc. IEEE* **109** 43–76.

Bernoulli **32**(4), 2026, 3113–3142
https://doi.org/10.3150/26-BEJ1981

E-values for exponential families: The general case

YUNDA HAO[1,a] and PETER GRÜNWALD[2,3,b]

[1] *Department of Statistics and Data Science, the Chinese University of Hong Kong, Hong Kong, China,* [a] *yundahao@cuhk.edu.hk*
[2] *CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands,* [b] *pdg@cwi.nl*
[3] *Mathematical Institute, Leiden University, Leiden, The Netherlands*

We analyze common types of e-variables and e-processes for composite exponential family nulls: the optimal e-variable based on the reverse information projection (RIPr), a conditional (COND) e-variable, and the universal inference (UI) and sequentialized RIPr e-processes. Whereas earlier derivations of the RIPr e-variable, for parametric and nonparametric nulls alike, were restricted to cases in which it reduces to a simple-vs.-simple likelihood, we manage to derive it also in 'anti-simple' cases in which it cannot be so reduced. We characterize the RIPr for simple and Bayes-mixture based alternatives, either precisely (for Gaussian nulls and alternatives) or in an approximate sense (general exponential family nulls). We also provide conditions under which the RIPr e-variable is (again exactly vs. asymptotically) equal to the COND e-variable, and we determine, up to $o(1)$, the e-power of the four e-statistics as a function of sample size. For d-dimensional null and alternative, the e-power of UI tends to be smaller by a term of $(d/2)\log n + O(1)$ than that of the COND e-variable, which is the clear winner.

Keywords: Exponential family; e-value; growth rate; hypothesis testing

References

[1] Bar-Lev, S.K. (2023). The exponential dispersion model generated by the Landau distribution —a comprehensive review and further developments. *Mathematics* **11** 4343. https://doi.org/10.3390/math11204343
[2] Barndorff-Nielsen, O. (2014). *Information and Exponential Families: In Statistical Theory*. Wiley. https://doi.org/10.1002/9781118857281
[3] Bhattacharya, R.N. and Rao, R.R. (1976). *Normal Approximation and Asymptotic Expansions*. Wiley. https://doi.org/10.1137/1.9780898719895.ch1
[4] Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer. https://doi.org/10.5555/1162264
[5] Brown, L.D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. *Lecture Notes-Monograph Series* **9**. https://doi.org/10.1214/lnms/1215466757
[6] Dawid, A.P. (1984). Present position and potential developments: Some personal views: Statistical theory, the prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–290. https://doi.org/10.2307/2981683
[7] Efron, B. (2022). *Exponential Families in Theory and Practice*. *IMS Textbooks*. Cambridge University Press. https://doi.org/10.1017/9781108773157
[8] Grünwald, P., De Heide, R. and Koolen, W. (2024). Safe testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 1091–1128. With Discussion. https://doi.org/10.1093/jrsssb/qkae011
[9] Grünwald, P. and De Rooij, S. (2005). Asymptotic log-loss of prequential maximum likelihood codes. In *COLT* 652–667. Springer. https://doi.org/10.1007/11503415_44
[10] Grünwald, P., Lardy, T., Hao, Y., Lev, S.K.B. and de Jong, M. (2024). Optimal E-values for exponential families: The simple case. In *Information Theory, Probability and Statistical Learning - a Festschrift in Honor of Andrew Barron* (J. Klusowski, I. Kontoyiannis and C. Rush, eds.). Springer. arXiv preprint, To appear, 2025, arXiv:2404.19465. https://doi.org/10.48550/arXiv.2404.19465
[11] Hao, Y. and Grünwald, P. (2025a). E-values for exponential families: the general case. arXiv preprint, version 2 (extended version of the present article including results under misspecification), arXiv:2409.11134. https://doi.org/10.48550/arXiv.2409.11134

1350-7265

[12] Hao, Y. and Grünwald, P. (2026). Supplement to “E-values for exponential families: The general case.” https://doi.org/10.3150/26-BEJ1981SUPP

[13] Hao, Y., Grünwald, P., Lardy, T., Long, L. and Adams, R. (2024). E-values for k-sample tests with exponential families. *Sankhya A* **86** 596–636. https://doi.org/10.1007/s13171-024-00339-9

[14] Hendriksen, A., de Heide, R. and Grünwald, P. (2021). Optional stopping with Bayes factors: A categorization and extension of folklore results, with an application to invariant situations. *Bayesian Anal.* **16** 961–989. https://doi.org/10.1214/20-BA1234

[15] Kotłowski, W. and Grünwald, P. (2011). Maximum likelihood vs. Sequential normalized maximum likelihood in on-line density estimation. In *COLT. Proc. Mach. Learn. Res.* **19** 457–476. PMLR. https://proceedings.mlr.press/v19/kotlowski11a.html.

[16] Lardy, T., Grünwald, P. and Harremoës, P. (2024). Reverse information projections and optimal E-statistics. *IEEE Trans. Inf. Theory* **70** 7616–7631. https://doi.org/10.1109/TIT.2024.3444458

[17] Larsson, M., Ramdas, A. and Ruf, J. (2025). The numeraire e-variable and reverse information projection. *Ann. Statist.* **53** 1015–1–1015–43.

[18] Lindon, M., Ham, D.W., Tingley, M. and Bojinov, I. (2024). Anytime-Valid Linear Models and Regression Adjusted Causal Inference in Randomized Experiments. arXiv preprint, arXiv:2210.08589.

[19] Pérez-Ortiz, M.F., Lardy, T., De Heide, R. and Grünwald, P. (2024). E-statistics, group invariance and anytime valid testing. *Ann. Statist.* **52** 1410–1432. https://doi.org/10.1214/24-AOS2394

[20] Ramdas, A., Grünwald, P., Vovk, V. and Shafer, G. (2023). Game-theoretic statistics and safe anytime-valid inference. *Statist. Sci.* **38** 576–601. https://doi.org/10.1214/23-STS894

[21] Shafer, G. (2021). Testing by betting: A strategy for statistical and scientific communication (with discussion and response). *J. Roy. Statist. Soc. Ser. A* **184** 407–478. https://doi.org/10.1111/rssa.12647

[22] Tse, T. and Davison, A.C. (2022). A note on universal inference. *Stat* **11** e501. https://doi.org/10.1002/sta4.501

[23] Turner, R.J., Ly, A. and Grünwald, P.D. (2024). Generic e-variables for exact sequential k-sample tests that allow for optional stopping. *J. Statist. Plann. Inference* **230** 106116. https://doi.org/10.1016/j.jspi.2023.106116

[24] Vovk, V. and Wang, R. (2021). E-values: Calibration, combination and applications. *Ann. Statist.* **49** 1736–1754. https://doi.org/10.1214/20-AOS2020

[25] Wang, Q., Wang, R. and Ziegel, J. (2022). E-backtesting. *Management Science*. arXiv preprint, To appear. arXiv:2209.00991. https://dx.doi.org/10.2139/ssrn.4206997

[26] Wasserman, L., Ramdas, A. and Balakrishnan, S. (2020). Universal inference. *Proc. Natl. Acad. Sci. USA* **117** 16880–16890. https://doi.org/10.1073/pnas.1922664117

[27] 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. With Discussion. https://doi.org/10.1093/jrsssb/qkad009

Bernoulli **32**(4), 2026, 3143–3167
https://doi.org/10.3150/26-BEJ1982

Mixing phases and metastability for the Glauber dynamics on the p-spin Curie-Weiss model

RAMKRISHNA JYOTI SAMANTA[1,a], SOMABHA MUKHERJEE[2,b] and JIANG ZHANG[2,c]

[1]*Department of Statistical Science, University College London, London, UK,* [a]*ramkrishna.samanta.24@ucl.ac.uk*
[2]*Department of Statistics and Data Science, National University of Singapore, Singapore, Singapore,* [b]*somabha@nus.edu.sg,* [c]*zhangj@nus.edu.sg*

The Glauber dynamics for the classical 2-spin Curie-Weiss model on N nodes with inverse temperature β and zero external field is known to mix in time $\Theta(N \log N)$ for $\beta < \frac{1}{2}$, in time $\Theta(N^{3/2})$ at $\beta = \frac{1}{2}$, and in time $\exp(\Omega(N))$ for $\beta > \frac{1}{2}$. In this paper, we consider the p-spin generalization of the Curie-Weiss model with an external field h, and identify three disjoint regions partitioning the parameter space, with the corresponding Glauber dynamics exhibiting three different orders of mixing times in these regions. The construction of these disjoint regions depends on the number of local maximizers of a certain *negative free-energy* function $H_{\beta,h,p}$, and the behavior of the second derivative of $H_{\beta,h,p}$ at such a local maximizer. Specifically, we show that if $H_{\beta,h,p}$ has a unique local maximizer m_* with $H''_{\beta,h,p}(m_*) < 0$ and no other stationary point, then the Glauber dynamics mixes in time $\Theta(N \log N)$; if $H_{\beta,h,p}$ has multiple local maximizers, then the mixing time is $\exp(\Omega(N))$; and if $H_{\beta,h,p}$ has a unique local maximizer m_* with $H''_{\beta,h,p}(m_*) = 0$, then the mixing time is $\Theta(N^{3/2})$. We provide an explicit description of the geometry of these three different phases within the parameter space. Finally, we show that if $H_{\beta,h,p}$ has multiple local maximizers (*metastable states*), then one can create a restricted version of the original Glauber dynamics, which still mixes in time $\Theta(N \log N)$.

Keywords: Curie-Weiss; Glauber dynamics; metastability; mixing time; phase transition

References

Aizenman, M. and Holley, R. (1987). Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime. In *Spatial Stochastic Processes* (R. Durrett and H. Kesten, eds.). *IMA Volumes in Mathematics and Its Applications* **8** 1–11. New York: Springer. https://doi.org/10.1007/978-1-4613-8734-3_1

Ananikyan, L.N. (2007). The hexagonal recursive approximation with multisite-interaction Ising model for the solid and fluid ^{3}He system. *Int. J. Mod. Phys. B* **21** 755–772. https://doi.org/10.1142/S0217979207036667

Aslanov, L.A. (1988). Crystal symmetry and atomic interactions. *Comput. Math. Appl.* **16** 443–451. https://doi.org/10.1016/0898-1221(88)90234-9

Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*, 2nd ed. Chapman and Hall/CRC. https://doi.org/10.1201/9780203487808

Barra, A. (2009). Notes on ferromagnetic p-spin and REM. *Math. Methods Appl. Sci.* **32** 783–797. https://doi.org/10.1002/mma.1065

Bergamaschi, T.R., Menke, T., Banner, W.P., Di Paolo, A., Weber, S.J., Hirjibehedin, C.F., Kerman, A.J. and Oliver, W.D. (2022). Distinguishing multi-spin interactions from lower-order effects. *Phys. Rev. Appl.* **18** 044018. https://doi.org/10.1103/PhysRevApplied.18.044018

Bhamidi, S., Bresler, G. and Sly, A. (2008). Mixing time of exponential random graphs. In *49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)* 803–812. https://doi.org/10.1109/FOCS.2008.75

Bhowal, S. and Mukherjee, S. (2023). Limit Theorems and Phase Transitions in the Tensor Curie-Weiss Potts Model. arXiv preprint. Available at arXiv:2307.01052. https://doi.org/10.48550/arXiv.2307.01052

Bhowal, S. and Mukherjee, S. (2025). Rates of convergence of the magnetization in the tensor Curie–Weiss Potts model. *J. Stat. Phys.* **192** 1–25. https://doi.org/10.1007/s10955-024-03382-w

Bianchi, A. (2008). Glauber dynamics on nonamenable graphs: Boundary conditions and mixing time. *Electron. J. Probab.* **13** 1980–2012. https://doi.org/10.1214/EJP.v13-574

Bresler, G., Nagaraj, D.M. and Nichani, E. (2022). Metastable mixing of Markov chains: Efficiently sampling low temperature exponential random graphs. *Ann. Appl. Probab.* **32** 517–554. https://doi.org/10.1214/23-AAP1971

Bubley, R. and Dyer, M. (1997). Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS)* 223–231. IEEE. https://doi.org/10.1109/SFCS.1997.646111

Can, V.H., van der Hofstad, R. and Kumagai, T. (2021). Glauber dynamics for Ising models on random regular graphs: Cut-off and metastability. *ALEA Lat. Amer. J. Probab. Math. Stat.* **18** 1441–1482. https://doi.org/10.30757/ALEA.v18-52

Chen, Z., Liu, K. and Vigoda, E. (2020). Rapid mixing of Glauber dynamics up to uniqueness via contraction. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)* 1307–1318. https://doi.org/10.1137/20M136685X

Contucci, P., Mingione, E. and Osabutey, G. (2024). Limit theorems for the cubic mean-field Ising model. *Ann. Henri Poincaré* **25** 5019–5044. https://doi.org/10.1007/s00023-024-01420-7

Cuff, P.J., Ding, J.K., Louidor, O., Lubetzky, E., Peres, Y. and Sly, A. (2012). Glauber dynamics for the mean-field Potts model. *J. Stat. Phys.* **149** 432–477. https://doi.org/10.1007/s10955-012-0599-2

Daskalakis, C., Dikkala, N. and Panageas, I. (2020). Logistic regression with peer-group effects via inference in higher-order Ising models. In *International Conference on Artificial Intelligence and Statistics (AISTATS). Proceedings of Machine Learning Research (PMLR)* 3653–3663.

DeMuse, R., Easlick, T. and Yin, M. (2019). Mixing time of vertex-weighted exponential random graphs. *J. Comput. Appl. Math.* **362** 443–459. https://doi.org/10.1016/j.cam.2018.07.038

Derrida, B. (1980). Random-energy model: Limit of a family of disordered models. *Phys. Rev. Lett.* **45** 79–82. https://doi.org/10.1103/PhysRevLett.45.79

Ding, J., Lubetzky, E. and Peres, Y. (2014). The mixing time evolution of Glauber dynamics for the mean-field Ising model. *Ann. Probab.* **42** 1042–1071. https://doi.org/10.1214/13-AOP845

Ding, J. and Peres, Y. (2009). Mixing time for the Ising model: A uniform lower bound for all graphs. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 1020–1028. https://doi.org/10.1214/10-AIHP402

Eichelsbacher, P. (2024). Stein's method and a cubic mean-field model. *J. Stat. Phys.* **191**. https://doi.org/10.1007/s10955-024-03373-x

Gardner, E. (1985). Spin glasses with p-spin interactions. *Nuclear Phys. B* **257** 747–765. https://doi.org/10.1016/0550-3213(85)90374-9

Geman, S. and Graffigne, C. (1986). Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians* 1496–1517. https://doi.org/10.1007/bfb0028368

Gheissari, R. and Sinclair, A. (2024). Spatial mixing and the random-cluster dynamics on lattices. *Random Structures Algorithms* **64** 490–534. https://doi.org/10.1002/rsa.21191

Green, P.J. and Richardson, S. (2002). Hidden Markov models and disease mapping. *J. Amer. Statist. Assoc.* **97** 1055–1070. https://doi.org/10.1198/016214502388618870

Griffiths, R.B., Weng, C.-Y. and Langer, J.S. (1966). Relaxation times for metastable states in the mean-field model of a ferromagnet. *Phys. Rev.* **149** 301–305. https://doi.org/10.1103/PhysRev.149.301

Hayes, T.P. and Sinclair, A. (2005). A general lower bound for mixing of single-site dynamics on graphs. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)* 511–520. https://doi.org/10.1214/105051607000000104

He, R. and Lok, J. (2024). On Approximating the Potts Model with Contracting Glauber Dynamics. Available at arXiv:2404.18778. https://doi.org/10.48550/arXiv.2404.18778

Heringa, J.R., Blöte, H.W.J. and Hoogland, A. (1989). Phase transitions in self-dual Ising models with multispin interactions and a field. *Phys. Rev. Lett.* **63** 1546–1549. https://doi.org/10.1103/PhysRevLett.63.1546

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79** 2554–2558. https://doi.org/10.1073/pnas.79.8.2554

Ising, E. (1925). Beitrag zur Theorie der Ferromagnetismus. *Z. Phys.* **31** 253–258. https://doi.org/10.1007/BF02980577

Jörg, T., Krzakala, F., Kurchan, J., Maggs, A.C. and Pujos, J. (2010). Energy gaps in quantum first-order mean-field–like transitions: The problems that quantum annealing cannot solve. *Europhys. Lett.* **89** 40004. https://doi.org/10.1209/0295-5075/89/40004

Levin, D.A., Luczak, M.J. and Peres, Y. (2007). Glauber dynamics for the mean-field Ising model: Cut-off, critical power law, and metastability. *Probab. Theory Related Fields* **146** 223–265. https://doi.org/10.1007/s00440-008-0189-z

Levin, D.A., Peres, Y. and Wilmer, E.L. (2006). *Markov Chains and Mixing Times*. American Mathematical Society.

Mikulincer, D. and Piana, A. (2024). Stochastic Localization with Non-Gaussian Tilts and Applications to Tensor Ising Models. Available at arXiv:2412.12720. https://doi.org/10.48550/arXiv.2412.12720

Montanari, A. and Saberi, A. (2010). The spread of innovations in social networks. *Proc. Natl. Acad. Sci. USA* **107** 20196–20201. https://doi.org/10.1073/pnas.1004098107

Mossel, E. and Sly, A. (2009). Rapid mixing of Gibbs sampling on graphs that are sparse on average. *Random Structures Algorithms* **35** 250–270. https://doi.org/10.1002/rsa.20276

Mukherjee, S., Son, J. and Bhattacharya, B.B. (2021). Fluctuations of the magnetization in the p-spin Curie–Weiss model. *Comm. Math. Phys.* **387** 681–728. https://doi.org/10.1007/s00220-021-04182-z

Mukherjee, S., Son, J. and Bhattacharya, B.B. (2022). Estimation in tensor Ising models. *Inf. Inference* **11** 1457–1500. https://doi.org/10.1093/imaiai/iaac007

Mukherjee, S., Son, J. and Bhattacharya, B.B. (2025). Phase transitions of the maximum likelihood estimators in the p-spin Curie-Weiss model. *Bernoulli* **31** 1502–1526. https://doi.org/10.3150/24-BEJ1779

Ohkuwa, M., Nishimori, H. and Lidar, D.A. (2018). Reverse annealing for the fully connected p-spin model. *Phys. Rev. A* **98** 022314. https://doi.org/10.1103/PhysRevA.98.022314

Osabutey, G. (2024). Phase properties of the mean-field Ising model with three-spin interaction. *J. Math. Phys.* **65** 061901. https://doi.org/10.1063/5.0183805

Samanta, R.J., Mukherjee, S. and Zhang, J. (2026). Supplement to "Mixing Phases and Metastability for the Glauber Dynamics on the p-Spin Curie-Weiss Model." https://doi.org/10.3150/26-BEJ1982SUPP

Suzuki, M. (1972). Solution and critical behavior of some "three-dimensional" Ising models with a four-spin interaction. *Phys. Rev. Lett.* **28** 507–510. https://doi.org/10.1103/PhysRevLett.28.507

Suzuki, M. and Fisher, M.E. (1971). Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models. *J. Math. Phys.* **12** 235–246. https://doi.org/10.1063/1.1665583

Turban, L. (2016). One-dimensional Ising model with multispin interactions. *J. Phys. A: Math. Theor.* **49**. https://doi.org/10.1088/1751-8113/49/35/355002

Yamashiro, Y., Ohkuwa, M., Nishimori, H. and Lidar, D.A. (2019). Dynamics of reverse annealing for the fully-connected p-spin model. *Phys. Rev. A* **100** 052321. https://doi.org/10.1103/PhysRevA.100.052321

Bernoulli **32**(4), 2026, 3168–3194
https://doi.org/10.3150/26-BEJ1983

Asymptotic analysis of the finite predictor for fractional Gaussian noise

PAVEL CHIGANSKY[1,a] and MARINA KLEPTSYNA[2,b]

[1]*Department of Statistics, The Hebrew University of Jerusalem, Jerusalem, Israel,*
[a]*Pavel.Chigansky@mail.huji.ac.il*
[2]*Laboratoire Manceau de Mathématiques, Le Mans Université, Le Mans, France,*
[b]*Marina.Kleptsyna@univ-lemans.fr*

This paper proposes a new approach to the asymptotic analysis of the finite predictor for stationary sequences. Our method yields the exact asymptotics of both the relative prediction error and the partial correlation coefficients. The underlying assumptions are analytic in nature, making the approach applicable to processes with long-range dependence. The ARMA-type process driven by fractional Gaussian noise (fGn), which had previously remained elusive, is used as a case study.

Keywords: Fractional Gaussian noise; long-range dependence; prediction

References

Afterman, D., Chigansky, P., Kleptsyna, M. and Marushkevych, D. (2022). Linear filtering with fractional noises: Large time and small noise asymptotics. *SIAM J. Control Optim.* **60** 1463–1487.
Babayan, N.M. and Ginovyan, M.S. (2023). Asymptotic behavior of the prediction error for stationary sequences. *Probab. Surv.* **20** 664–721. MR4607530 https://doi.org/10.1214/23-ps21
Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer-Verlag. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4
Chigansky, P. and Kleptsyna, M. (2018). Exact asymptotics in eigenproblems for fractional Brownian covariance operators. *Stoch. Process. Appl.* **128** 2007–2059. MR3797652 https://doi.org/10.1016/j.spa.2017.08.019
Chigansky, P. and Kleptsyna, M. (2026). Supplement to "Asymptotic analysis of the finite predictor for fractional Gaussian noise." https://doi.org/10.3150/26-BEJ1983SUPP
Embrechts, P. and Maejima, M. (2002). *Selfsimilar Processes*. *Princeton Series in Applied Mathematics*. Princeton, NJ: Princeton University Press. MR1920153
Gakhov, F.D. (1990). *Boundary Value Problems*. New York: Dover Publications, Inc. Translated from the Russian, Reprint of the 1966 translation. MR1106850
Grenander, U. and Szegő, G. (1984). *Toeplitz Forms and Their Applications*, 2nd ed. New York: Chelsea Publishing Co. MR0890515
Ibragimov, I.A. and Solev, V.N. (1968). Asymptotic behavoir of the prediction error of a stationary sequence with a spectral density of special form. *Teor. Veroyatn. Primen.* **13** 746–751. MR0242242
Inoue, A. (2000). Asymptotics for the partial autocorrelation function of a stationary process. *J. Anal. Math.* **81** 65–109. MR1785278 https://doi.org/10.1007/BF02788986
Inoue, A. (2008). AR and MA representation of partial autocorrelation functions, with applications. *Probab. Theory Related Fields* **140** 523–551.
Inoue, A. (2023). Representation theorems in finite prediction, with applications. *Sugaku Expositions* **36** 173–197. MR4652498
Lewin, L. (1981). *Polylogarithms and Associated Functions*. New York-Amsterdam: North-Holland Publishing Co. With a foreword by A. J. Van der Poorten. MR618278
Pal′cev, B.V. (1974). Asymptotic behavior of the spectrum and eigenfunctions of convolution operators on a finite interval with the kernel having a homogeneous Fourier transform. *Dokl. Akad. Nauk SSSR* **218** 28–31. MR0355697

Pal′tsev, B.V. (2003). Asymptotics of the spectrum of integral convolution operators on a finite interval with homogeneous polar kernels. *Izv. Ross. Akad. Nauk Ser. Mat.* **67** 67–154. MR2001765 https://doi.org/10.1070/IM2003v067n04ABEH000443

Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics* **45**. Cambridge: Cambridge University Press. MR3729426

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (2007). *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge: Cambridge University Press. MR2371990

Stein, E.M. and Shakarchi, R. (2003). *Complex Analysis*. *Princeton Lectures in Analysis 2*. Princeton, NJ: Princeton University Press. MR1976398

Ukai, S. (1971). Asymptotic distribution of eigenvalues of the kernel in the Kirkwood-Riseman integral equation. *J. Math. Phys.* **12** 83–92. MR0275084

Bernoulli **32**(4), 2026, 3195–3221
https://doi.org/10.3150/26-BEJ1984

Parameter estimation from local measurements for a class of stochastic Burgers equations

JOSEF JANÁK[a] and ENRICO PRIOLA[b]

University of Pavia, Department of Mathematics, Corso Strada Nuova 65, 27100 Pavia, Italy,
[a]*josef.janak@seznam.cz,* [b]*enrico.priola@unipv.it*

We deal with a class of semilinear SPDEs driven by space–time white noise that includes the one dimensional stochastic Burgers equation. Such equations can have nonlocal and quadratic nonlinearities. We consider the problem of estimation of the diffusivity parameter in front of the second-order spatial derivative. Based on local observations in space, we study the estimator derived in (*Ann. Appl. Probab.* **31** (2021) 1–38) for linear stochastic heat equation that has also been used in (*Bernoulli* **29** (2023) 2035–2061) to cover certain class of semilinear SPDEs including stochastic Burgers equations driven by trace class noise. The space-time white noise case we consider has also relevant physical motivations. After we establish new regularity results for the solution, we are able to show that our proposed estimator is strongly consistent and asymptotically normal.

Keywords: Augmented MLE; central limit theorem; diffusivity estimation; generalized stochastic Burgers equation; local measurements; space-time white noise

References

[1] Albeveiro, S. and Rüdiger, B. (2005). Subordination of symmetric quasi-regular Dirichlet forms. *Random Oper. Stoch. Equ.* **13** 17–38.
[2] Altmeyer, R. and Reiß, M. (2021). Nonparametric estimation for linear SPDEs from local measurements. *Ann. Appl. Probab.* **31** 1–38.
[3] Altmeyer, R., Bretschneider, T., Janák, J. and Reiß, M. (2022). Parameter estimation in an SPDE model for cell repolarisation. *SIAM/ASA J. Uncertain. Quantificat.* **10** 179–199.
[4] Altmeyer, R., Cialenco, I. and Pasemann, G. (2023). Parameter estimation for semilinear SPDEs from local measurements. *Bernoulli* **29** 2035–2061.
[5] Ball, J.M. (1977). Strongly continuous semigroups, weak solutions, and variation of constants formula. *Proc. Amer. Math. Soc.* **63** 370–373.
[6] Bertini, L., Cancrini, N. and Jona-Lasinio, G. (1994). The stochastic Burgers equation. *Comm. Math. Phys.* **165** 211–232.
[7] Bibinger, M. and Trabs, M. (2019). On central limit theorems for power variations of the solution to the stochastic heat equation. In *Stochastic Models. Statistics and Their Applications. Springer Proc. Math. Stat.* **294** 69–84. Cham: Springer.
[8] Bibinger, M. and Trabs, M. (2020). Volatility estimation in stochastic PDEs using high-frequency observations. *Stoch. Process. Appl.* **130** 3005–3052.
[9] Chambers, D.H., Adrian, R.J., Moin, P., Stewart, D.S. and Sung, H.J. (1988). Karhunen-Loeve expansion of Burgers' model of turbulence. *Phys. Fluids* **31** 2573–2582.
[10] Cialenco, I. and Glatt-Holtz, N. (2011). Parameter estimation for the stochastically perturbed Navier-Stokes equations. *Stoch. Process. Appl.* **121** 701–724.
[11] Cohn, L.D. (2013). *Measure Theory*, 2nd ed. New York: Birkhäuser.
[12] Da Prato, G. and Debussche, A. (2007). m-dissipativity of Kolmogorov operators corresponding to Burgers equations with space-time white noise. *Potential Anal.* **26** 31–55.
[13] Da Prato, G. and Gatarek, D. (1995). Stochastic Burgers equation with correlated noise. *Stoch. Stoch. Rep.* **52** 29–41.

[14] Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite Dimensional Systems* **229**. Cambridge University Press.
[15] Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge University Press.
[16] Da Prato, G., Debussche, A. and Temam, R. (1994). Stochastic Burgers' equation. *NoDEA Nonlinear Differential Equations Appl.* **1** 389–402.
[17] Fukushima, M., Oshima, Y. and Takeda, M. (2010). *Dirichlet Norms and Symmetric Markov Processes, Vol. 19*. Walter de Gruyter.
[18] Gyöngy, I. (1998). Existence and uniqueness results for semilinear stochastic partial differential equations. *Stoch. Process. Appl.* **73** 271–299.
[19] Gyöngy, I. and Pardoux, E. (1993). On quasi-linear stochastic partial differential equations. *Probab. Theory Related Fields* **94** 413–425.
[20] Gyöngy, I. and Rovira, C. (1999). On stochastic partial differential equations with polynomial nonlinearities. *Stoch. Stoch. Rep.* **67** 123–146.
[21] Hairer, M. (2009). An Introduction to Stochastic PDEs. Available at arXiv:0907.4178v1.
[22] Hairer, M. and Voss, J. (2011). Approximations to the stochastic Burgers equation. *J. Nonlinear Sci.* **21** 897–920.
[23] Henry, D. (1981). *Geomeric Theory of Semilinear Parabolic Equations*. *Lecture Notes in Math.* **840**. Berlin-New York: Springer-Verlag.
[24] Hildebrandt, F. and Trabs, M. (2023). Nonparametric calibration for stochastic reaction-diffusion equations based on discrete observations. *Stoch. Process. Appl.* **162** 171–217.
[25] Hirsch, F. (1999). Lipschitz functions and fractional Sobolev spaces. *Potential Anal.* **11** 415–429.
[26] Huebner, M. and Rozovskii, B. (1995). On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDEs. *Probab. Theory Related Fields* **103** 143–163.
[27] Iscoe, I., Marcus, M.B., McDonald, D., Talagrand, M. and Zinn, J. (1990). Continuity of l^2-valued Ornstein-Uhlenbeck processes. *Ann. Probab.* 68–84.
[28] Itô, K. and Nisio, M. (1968). On the convergence of sums of independent Banach space valued random variables. *Osaka J. Math.* **5** 35–48.
[29] Janák, J. and Priola, E. (2026). Supplement to "Parameter estimation from local measurements for a class of stochastic Burgers equations." https://doi.org/10.3150/26-BEJ1984SUPP
[30] Janák, J. and Reiß, M. (2024). Estimation of the diffusivity parameter in the stochastic heat equation with multiplicative noise. *Stoch. Process. Appl.* **175**.
[31] Janson, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge University Press.
[32] Kardar, M., Parisi, G. and Zhang, Y.C. (1986). Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** 889–892.
[33] Kim, J.U. (2006). On the stochastic Burgers equation with a polynomial nonlinearity in the real line. *Discrete Contin. Dyn. Syst.* **6** 835–866.
[34] Kumar, A. and Mohan, M.T. (2024). Absolute continuity of the solution to stochastic generalized Burgers-Huxley equation. *Stoch. Partial Differ. Equ. Anal. Comput.* **12** 1983–2043.
[35] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. *Springer Series in Statistics*.
[36] Liptser, R.S. and Shiryayev, A.N. (1989). *Theory of Martingales*. *Mathematics and Its Applications (Soviet Series)* **49**. Dordrecht: Kluwer Academic.
[37] Loève, M. (1978). *Probability Theory II*, 4th ed. *Graduate Texts in Mathematics* **46**. Springer-Verlag.
[38] Lunardi, A. (2012). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser. Reprint of the 1995 Edition.
[39] Pasemann, G. and Stannat, W. (2020). Drift estimation for stochastic reaction-diffusion systems. *Electron. J. Stat.* **14** 547–579.
[40] Priola, E. (2021). An optimal regularity result for Kolmogorov equations and weak uniqueness for some critical SPDEs. *Ann. Probab.* **49** 1310–1346. Correction: ibid. (2023). **51** 2387–2395.
[41] Strichartz, R.S. (1967). Multipliers on fractional Sobolev spaces. *J. Math. Mech.* **16** 1031–1060.
[42] Tersenov, A.S. (2010). On the generalized Burgers equation. *NoDEA Nonlinear Differential Equations Appl.* **17** 437–452.

Bernoulli **32**(4), 2026, 3222–3248
https://doi.org/10.3150/26-BEJ1985

Bicausal optimal transport for SDEs with irregular coefficients

MICHAELA HITZ[a] and BENJAMIN A. ROBINSON[b]

Department of Statistics, University of Klagenfurt, Klagenfurt, Austria, [a]*michaela.hitz@aau.at,* [b]*benjamin.robinson@aau.at*

We solve constrained optimal transport problems in which the marginal laws are given by the laws of solutions of stochastic differential equations (SDEs). We consider SDEs with irregular coefficients, making only minimal regularity assumptions. We show that the so-called synchronous coupling is optimal among bicausal couplings, that is couplings that respect the flow of information encoded in the stochastic processes. Our results provide a method to numerically compute the adapted Wasserstein distance between laws of SDEs with irregular coefficients. We show that this can be applied to quantifying model uncertainty in stochastic optimisation problems. Moreover, we introduce a transformation-based semi-implicit numerical scheme and establish the first strong convergence result for SDEs with exponentially growing and discontinuous drift.

Keywords: Adapted Wasserstein distance; bicausal couplings; numerical methods for stochastic differential equations; optimal transport

References

[1] Acciaio, B., Backhoff-Veraguas, J. and Zalashko, A. (2020). Causal optimal transport and its links to enlargement of filtrations and continuous-time stochastic optimization. *Stoch. Process. Appl.* **130** 2918–2953. MR4080734 https://doi.org/10.1016/j.spa.2019.08.009

[2] Aldous, D.J. (1981). Weak Convergence and General Theory of Processes. Unpublished monograph. Department of Statistics, University of California. Berkeley.

[3] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, second ed. *Lectures in Mathematics ETH Zürich.* Basel: Birkhäuser Verlag. MR2401600 https://doi.org/10.1007/978-3-7643-8722-8

[4] Analui, B. and Pflug, G.C. (2014). On distributionally robust multiperiod stochastic optimization. *Comput. Manag. Sci.* **11** 197–220. MR3232022 https://doi.org/10.1007/s10287-014-0213-y

[5] Backhoff-Veraguas, J., Källblad, S. and Robinson, B.A. (2025). Adapted Wasserstein distance between the laws of SDEs. *Stoch. Process. Appl.* **189** 104689. MR4916099 https://doi.org/10.1016/j.spa.2025.104689

[6] Backhoff-Veraguas, J., Beiglböck, M., Lin, Y. and Zalashko, A. (2017). Causal transport in discrete time and applications. *SIAM J. Optim.* **27** 2528–2562. MR3738324 https://doi.org/10.1137/16M1080197

[7] Backhoff-Veraguas, J., Bartl, D., Beiglböck, M. and Eder, M. (2020a). Adapted Wasserstein distances and stability in mathematical finance. *Finance Stoch.* **24** 601–632. MR4118990 https://doi.org/10.1007/s00780-020-00426-3

[8] Backhoff-Veraguas, J., Bartl, D., Beiglböck, M. and Eder, M. (2020b). All adapted topologies are equal. *Probab. Theory Related Fields* **178** 1125–1172. MR4168395 https://doi.org/10.1007/s00440-020-00993-8

[9] Bartl, D. and Wiesel, J. (2023). Sensitivity of multiperiod optimization problems with respect to the adapted Wasserstein distance. *SIAM J. Financial Math.* **14** 704–720. MR4599319 https://doi.org/10.1137/22M1537746

[10] Becker, S., Cheridito, P., Jentzen, A. and Welti, T. (2021). Solving high-dimensional optimal stopping problems using deep learning. *European J. Appl. Math.* **32** 470–514. MR4253974 https://doi.org/10.1017/S0956792521000073

1350-7265

[11] Bion-Nadal, J. and Talay, D. (2019). On a Wasserstein-type distance between solutions to stochastic differential equations. *Ann. Appl. Probab.* **29** 1609–1639. MR3914552 https://doi.org/10.1214/18-AAP1423
[12] Block, H.W., Griffith, W.S. and Savits, T.H. (1989). L-Superadditive structure functions. *Adv. Appl. Probab.* **21** 919–929. MR1039635 https://doi.org/10.2307/1427774
[13] Bonnier, P., Liu, C. and Oberhauser, H. (2023). Adapted topologies and higher rank signatures. *Ann. Appl. Probab.* **33** 2136–2175. MR4583667 https://doi.org/10.1214/22-aap1862
[14] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402
[15] Cheridito, P. and Eckstein, S. (2025). Optimal transport and Wasserstein distances for causal models. *Bernoulli* **31** 1351–1376. MR4863079 https://doi.org/10.3150/24-bej1773
[16] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems* (C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani and K.Q. Weinberger, eds.) **26** 2292–2300.
[17] Daley, D.J. (1968). Stochastically monotone Markov chains. *Z. Wahrsch. Verw. Gebiete* **10** 305–317. MR242270 https://doi.org/10.1007/BF00531852
[18] Dareiotis, K. and Gerencsér, M. (2020). On the regularisation of the noise for the Euler–Maruyama scheme with irregular drift. *Electron. J. Probab.* **25** 1–18. MR4125787 https://doi.org/10.1214/20-ejp479
[19] Dareiotis, K., Gerencsér, M. and Lê, K. (2023). Quantifying a convergence theorem of Gyöngy and Krylov. *Ann. Appl. Probab.* **33** 2291–2323. MR4583671 https://doi.org/10.1214/22-aap1867
[20] Davie, A.M. (2022). Approximation of SDE solutions using local asymptotic expansions. *Stoch. Partial Differ. Equ. Anal. Comput.* **10** 720–734. MR4491501 https://doi.org/10.1007/s40072-021-00232-8
[21] Eckstein, S. and Pammer, G. (2024). Computational methods for adapted optimal transport. *Ann. Appl. Probab.* **34** 675–713. MR4696288 https://doi.org/10.1214/23-aap1975
[22] Gerencsér, M., Lampl, G. and Ling, C. (2025). The Milstein scheme for singular SDEs with Hölder continuous drift. *IMA J. Numer. Anal.* **45** 3077–3108. MR4966453 https://doi.org/10.1093/imanum/drae083
[23] Giles, M. and Szpruch, L. (2013). Multilevel Monte Carlo methods for applications in finance. In *Recent Developments in Computational Finance*. *Interdiscip. Math. Sci.* **14** 3–47. Hackensack, NJ: World Sci. Publ. MR3288633 https://doi.org/10.1142/9789814436434_0001
[24] Gyöngy, I. (1998). A note on Euler's approximations. *Potential Anal.* **8** 205–216. MR1625576 https://doi.org/10.1023/A:1008605221617
[25] Gyöngy, I. and Krylov, N. (1996). Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields* **105** 143–158. MR1392450 https://doi.org/10.1007/BF01203833
[26] Gyöngy, I. and Rásonyi, M. (2011). A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stoch. Process. Appl.* **121** 2189–2200. MR2822773 https://doi.org/10.1016/j.spa.2011.06.008
[27] Higham, D.J., Mao, X. and Stuart, A.M. (2002). Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40** 1041–1063. MR1949404 https://doi.org/10.1137/S0036142901389530
[28] Hitz, M. and Robinson, B.A. (2026). Supplement to "Bicausal optimal transport for SDEs with irregular coefficients." https://doi.org/10.3150/26-BEJ1985SUPP
[29] Hu, Y. (1996). Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. In *Stochastic Analysis and Related Topics, V (Silivri, 1994)* (H. Körezlioglu, B. Øksendal and A.S. Üstünel, eds.). *Progr. Probab.* **38** 183–202. Boston, MA: Birkhäuser Boston. MR1396331 https://doi.org/10.1007/978-1-4612-2450-1_9
[30] Hu, H. and Gan, S. (2022). Strong convergence of the tamed Euler scheme for scalar SDEs with superlinearly growing and discontinuous drift coefficient. Available at arXiv:2206.00088. https://doi.org/10.48550/arXiv2206.00088
[31] Hutzenthaler, M. and Jentzen, A. (2015). Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *Mem. Amer. Math. Soc.* **236**. MR3364862 https://doi.org/10.1090/memo/1112
[32] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E. (2011). Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** 1563–1576. MR2795791 https://doi.org/10.1098/rspa.2010.0348

[33] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E. (2012). Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** 1611–1641. MR2985171 https://doi.org/10.1214/11-AAP803

[34] Jiang, Y. and Obłój, J. (2024). Sensitivity of causal distributionally robust optimization. Available at arXiv:2408.17109. https://doi.org/10.48550/arXiv2408.17109

[35] Jourdain, B. and Pagès, G. (2023). Convex ordering of solutions to one-dimensional SDEs. Available at arXiv:2312.09779. https://doi.org/10.48550/arXiv.2312.09779.

[36] Kantorovich, L. (1942). On the translocation of masses. *C. R. (Dokl.) Acad. Sci. URSS* **37** 199–201. MR0009619

[37] Kloeden, P.E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. *Applications of Mathematics* **23**. Berlin: Springer. MR1214374 https://doi.org/10.1007/978-3-662-12616-5

[38] Knothe, H. (1957). Contributions to the theory of convex bodies. *Michigan Math. J.* **4** 39–52. MR83759 https://doi.org/10.1307/mmj/1028990175

[39] Lassalle, R. (2018). Causal transference plans and their Monge–Kantorovich problems. *Stoch. Anal. Appl.* **36** 452–484. MR3784142 https://doi.org/10.1080/07362994.2017.1422747

[40] Leobacher, G. and Szölgyenyi, M. (2016). A numerical method for SDEs with discontinuous drift. *BIT* **56** 151–162. MR3486457 https://doi.org/10.1007/s10543-015-0549-x

[41] Leobacher, G. and Szölgyenyi, M. (2017a). A strong order $1/2$ method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Probab.* **27** 2383–2418. MR3693529 https://doi.org/10.1214/16-AAP1262

[42] Leobacher, G. and Szölgyenyi, M. (2017b). Numerical methods for SDEs with drift discontinuous on a set of positive reach. *Internat. Math. News* **235** 1–16. https://doi.org/10.48550/arXiv1708.06188

[43] Leobacher, G. and Szölgyenyi, M. (2018). Convergence of the Euler–Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numer. Math.* **138** 219–239. MR3745015 https://doi.org/10.1007/s00211-017-0903-9

[44] Leobacher, G., Szölgyenyi, M. and Thonhauser, S. (2015). On the existence of solutions of a class of SDEs with discontinuous drift and singular diffusion. *Electron. Commun. Probab.* **20** 14. MR3304412 https://doi.org/10.1214/ECP.v20-3149

[45] Liu, Y. and Pagès, G. (2022). Monotone convex order for the McKean–Vlasov processes. *Stoch. Process. Appl.* **152** 312–338. MR4453161 https://doi.org/10.1016/j.spa.2022.06.003

[46] Mao, X. and Szpruch, L. (2013a). Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Comput. Appl. Math.* **238** 14–28. MR2972586 https://doi.org/10.1016/j.cam.2012.08.015

[47] Mao, X. and Szpruch, L. (2013b). Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. *Stochastics* **85** 144–171. MR3011916 https://doi.org/10.1080/17442508.2011.651213

[48] Milstein, G.N. (1975). Approximate integration of stochastic differential equations. *Theory Probab. Appl.* **19** 557–562. MR356225 https://doi.org/10.1137/1119062

[49] Milstein, G.N., Repin, Y.M. and Tretyakov, M.V. (2002). Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.* **40** 1583–1604. MR1951908 https://doi.org/10.1137/S0036142901395588

[50] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Hist. Acad. R. Sci. Mém. Math. Phys.* 666–704.

[51] Müller-Gronbach, T., Sabanis, S. and Yaroslavtseva, L. (2025). Existence, uniqueness and approximation of solutions of SDEs with superlinear coefficients in the presence of discontinuities of the drift coefficient. *Commun. Math. Sci.* **23** 389–429. MR4846712 https://doi.org/10.4310/cms.241217213438

[52] Müller-Gronbach, T. and Yaroslavtseva, L. (2020). On the performance of the Euler–Maruyama scheme for SDEs with discontinuous drift coefficient. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1162–1178. MR4076779 https://doi.org/10.1214/19-AIHP997

[53] Müller-Gronbach, T. and Yaroslavtseva, L. (2022). A strong order $3/4$ method for SDEs with discontinuous drift coefficient. *IMA J. Numer. Anal.* **42** 229–259. MR4367654 https://doi.org/10.1093/imanum/draa078

[54] Neuenkirch, A. and Szölgyenyi, M. (2021). The Euler–Maruyama scheme for SDEs with irregular drift: Convergence rates via reduction to a quadrature problem. *IMA J. Numer. Anal.* **41** 1164–1196. MR4246871 https://doi.org/10.1093/imanum/draa007

[55] Neuenkirch, A., Szölgyenyi, M. and Szpruch, L. (2019). An adaptive Euler–Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis. *SIAM J. Numer. Anal.* **57** 378–403. MR3914573 https://doi.org/10.1137/18M1170017
[56] Ngo, H.-L. and Taguchi, D. (2017a). On the Euler–Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients. *IMA J. Numer. Anal.* **37** 1864–1883. MR3712177 https://doi.org/10.1093/imanum/drw058
[57] Ngo, H.-L. and Taguchi, D. (2017b). Strong convergence for the Euler–Maruyama approximation of stochastic differential equations with discontinuous coefficients. *Statist. Probab. Lett.* **125** 55–63. MR3626068 https://doi.org/10.1016/j.spl.2017.01.027
[58] Ngo, H.-L. and Taguchi, D. (2019). On the Euler–Maruyama scheme for SDEs with bounded variation and Hölder continuous coefficients. *Math. Comput. Simul.* **161** 102–112. MR3926802 https://doi.org/10.1016/j.matcom.2019.01.012
[59] Peyré, G. and Cuturi, M. (2019). Computational optimal transport. *Found. Trends Mach. Learn.* **11** 355–607. https://doi.org/10.1561/2200000073
[60] Pflug, G.C. and Pichler, A. (2012). A distance for multistage stochastic optimization models. *SIAM J. Optim.* **22** 1–23. MR2902682 https://doi.org/10.1137/110825054
[61] Pflug, G.C. and Pichler, A. (2014). *Multistage Stochastic Optimization*. *Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3288310 https://doi.org/10.1007/978-3-319-08843-3
[62] Pichler, A. and Weinhardt, M. (2021). The nested Sinkhorn divergence to learn the nested distance. *Comput. Manag. Sci.* **19** 269–293. MR4423922 https://doi.org/10.1007/s10287-021-00415-7
[63] Przybyłowicz, P., Schwarz, V. and Szölgyenyi, M. (2024). A higher order approximation method for jump-diffusion SDEs with discontinuous drift coefficient. *J. Math. Anal. Appl.* **538** 128319. MR4728130 https://doi.org/10.1016/j.jmaa.2024.128319
[64] Przybyłowicz, P. and Szölgyenyi, M. (2021). Existence, uniqueness, and approximation of solutions of jump-diffusion SDEs with discontinuous drift. *Appl. Math. Comput.* **403** 126191. MR4235987 https://doi.org/10.1016/j.amc.2021.126191
[65] Przybyłowicz, P., Szölgyenyi, M. and Xu, F. (2021). Existence and uniqueness of solutions of SDEs with discontinuous drift and finite activity jumps. *Statist. Probab. Lett.* **174** 109072. MR4241638 https://doi.org/10.1016/j.spl.2021.109072
[66] Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.* **23** 470–472. MR49525 https://doi.org/10.1214/aoms/1177729394
[67] Rüschendorf, L. (1985). The Wasserstein distance and approximation theorems. *Z. Wahrsch. Verw. Gebiete* **70** 117–129. MR795791 https://doi.org/10.1007/BF00532240
[68] Sabanis, S. (2013). A note on tamed Euler approximations. *Electron. Commun. Probab.* **18** 1–10. MR3070913 https://doi.org/10.1214/ECP.v18-2824
[69] Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians*. *Progress in Nonlinear Differential Equations and Their Applications* **87**. Cham: Birkhäuser. MR3409718 https://doi.org/10.1007/978-3-319-20828-2
[70] Spendier, K. and Szölgyenyi, M. (2024). Convergence of the tamed-Euler–Maruyama method for SDEs with discontinuous and polynomially growing drift. In *Monte Carlo and Quasi-Monte Carlo Methods. MCQMC 2022* (A. Hinrichs, P. Kritzer and F. Pillichshammer, eds.). *Springer Proc. Math. Stat.* **460** 601–623. Springer. MR4774942 https://doi.org/10.1007/978-3-031-59762-6_30
[71] Szölgyenyi, M. (2021). Stochastic differential equations with irregular coefficients: Mind the gap! *Internat. Math. News* **246** 43–56. https://doi.org/10.48550/arXiv2104.11505
[72] Talay, D. (1983). Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. *Stochastics* **9** 275–306. MR0707643 https://doi.org/10.1080/17442508308833257
[73] Villani, C. (2003). *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Providence, RI: American Mathematical Society. MR1964483 https://doi.org/10.1090/gsm/058
[74] Watanabe, S. and Yamada, T. (1971). On the uniqueness of solutions of stochastic differential equations II. *J. Math. Kyoto Univ.* **11** 553–563. MR288876 https://doi.org/10.1215/kjm/1250523620
[75] Yaroslavtseva, L. (2022). An adaptive strong order 1 method for SDEs with discontinuous drift coefficient. *J. Math. Anal. Appl.* **513** 126180. MR4404565 https://doi.org/10.1016/j.jmaa.2022.126180
[76] Zvonkin, A.K. (1974). A transformation of the phase space of a diffusion process that removes the drift. *Math. USSR, Sb.* **22** 129–149. MR336813 https://doi.org/10.1070/SM1974v022n01ABEH001689

Bernoulli **32**(4), 2026, 3249–3276
https://doi.org/10.3150/26-BEJ1986

Optimal rates for estimating the covariance kernel from synchronously sampled functional data

MAX BERGER[a] and HAJO HOLZMANN[b]

Philipps-Universität, Marburg, Germany, [a]*mberger@mathematik.uni-marburg.de,*
[b]*holzmann@mathematik.uni-marburg.de*

We obtain minimax-optimal convergence rates in the supremum norm, including information-theoretic lower bounds, for estimating the covariance kernel of a stochastic process which is repeatedly observed at discrete, synchronous design points. We focus on the supremum norm instead of the simpler L_2 norm, since it forms the basis for the construction of uniform confidence bands and is a stronger norm than the L_2-norm, thus arguably corresponding to a better visualization of the estimation error. For dense design, assuming Hölder-smooth sample paths we obtain the $\sqrt{n}$-rate of convergence in the supremum norm without additional logarithmic factors which typically occur in the results in the literature. Surprisingly, in the transition from dense to sparse design the rates do not reflect the two-dimensional nature of the covariance kernel but correspond to those for univariate mean function estimation. Our estimation method can make use of higher-order smoothness of the covariance kernel away from the diagonal, and does not require the same smoothness on the diagonal itself. Hence, our results cover covariance kernels of processes with rough, non-differentiable sample paths. Moreover, the estimator does not use mean function estimation to form residuals, and no smoothness assumptions on the mean have to be imposed. In the dense case we also obtain a central limit theorem in the supremum norm, which can be used as the basis for the construction of uniform confidence sets. Extensions to estimating partial derivatives as well as to asynchronous designs are also discussed. Simulations and real-data applications illustrate the practical usefulness of the methods.

Keywords: Covariance kernel; functional data; optimal rates of convergence; supremum norm; synchronously sample data

References

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. John Wiley & Sons.

Azmoodeh, E., Sottinen, T., Viitasaari, L. and Yazigi, A. (2014). Necessary and sufficient conditions for Hölder continuity of Gaussian processes. *Statist. Probab. Lett.* **94** 230–235.

Belhakem, R., Picard, F., Rivoirard, V. and Roche, A. (2025). Minimax estimation of functional principal components from noisy discretized functional data. *Scand. J. Stat.* **52** 38–80.

Berger, M. (2025). From Dense to Sparse Design: Optimal Rates under the Supremum Norm in Functional Data Analysis. Dissertation, Philipps-Universität Marburg.

Berger, M., Hermann, P. and Holzmann, H. (2025). From dense to sparse design: Optimal rates under the supremum norm for estimating the mean function in functional data analysis. *Bernoulli* **31** 1858–1888. https://doi.org/10.3150/24-BEJ1789

Berger, M. and Holzmann, H. (2025). Smooth and rough paths in mean derivative estimation for functional data. arXiv preprint. Available at arXiv:2503.24066.

Berger, M. and Holzmann, H. (2026). Supplement to "Optimal rates for estimating the covariance kernel from synchronously sampled functional data." https://doi.org/10.3150/26-BEJ1986SUPP

Cai, T. and Yuan, M. (2010). Nonparametric covariance function estimation for functional and longitudinal data. University of Pennsylvania and Georgia Inistitute of Technology.

Cai, T.T. and Yuan, M. (2011). Optimal estimation of the mean function based on discretely sampled functional data: Phase transition. *Ann. Statist.* **39** 2330–2355.

Cardot, H., Degras, D. and Josserand, E. (2013). Confidence bands for Horvitz-Thompson estimators using sampled noisy functional data. *Bernoulli* **19** 2067–2097. https://doi.org/10.3150/12-BEJ443

Dai, X., Müller, H.-G. and Tao, W. (2018). Derivative principal component analysis for representing the time dynamics of longitudinal and functional data. *Statist. Sinica* **28** 1583–1609. https://doi.org/10.5705/ss.202016.0427

Dette, H. and Kokot, K. (2022). Detecting relevant differences in the covariance operators of functional time series: A sup-norm approach. *Ann. Inst. Statist. Math.* **74** 195–231.

Dette, H., Kokot, K. and Aue, A. (2020). Functional data analysis in the Banach space of continuous functions. *Ann. Statist.* **48** 1168–1192.

Dette, H. and Wu, W. (2021). Confidence surfaces for the mean of locally stationary functional time series. Preprint. Available at arXiv:2109.03641 [math.ST].

Golovkine, S., Klutchnikoff, N. and Patilea, V. (2025). Adaptive estimation of irregular mean and covariance functions. *Bernoulli* **31** 1032–1057.

Hall, P. and Hosseini-Nasab, M. (2009). Theory for high-order bounds in functional principal components analysis. In *Mathematical Proceedings of the Cambridge Philosophical Society* **146** 225–256. Cambridge University Press.

Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* 1493–1517.

Li, Y. and Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351.

Liebl, D. and Reimherr, M. (2023). Fast and fair simultaneous confidence bands for functional parameters. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 842–868. https://doi.org/10.1093/jrsssb/qkad026

Lin, Z. and Wang, J.-L. (2022). Mean and covariance estimation for functional snippets. *J. Amer. Statist. Assoc.* **117** 348–360.

Mohammadi, N. and Panaretos, V.M. (2024). Functional data analysis with rough sample paths? *J. Nonparametr. Stat.* **36** 4–22.

Pollard, D. (1990). Empirical processes: Theory and applications. In *NSF-CBMS Regional Conference Series in Probability and Statistics* i–86. JSTOR.

Ramsay, J. and Silvermann, B. (1998). *Functional Data Analysis*. *Springer Series in Statistics*.

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Ser. Stat.* New York, NY: Springer.

Turkmen, R. and Civciv, H. (2007). Some bounds for the singular values of matrices. *Appl. Math. Sci.* **1** 2443–2449.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Science & Business Media.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science 47*. Cambridge university press.

Wang, L., Brown, L.D. and Cai, T. (2008). Effect of mean on variance function estimation in nonparametric regression. *Ann. Statist.* **36** 646–664.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Xiao, L. (2020). Asymptotic properties of penalized splines for functional data. *Bernoulli* **26** 2847–2875.

Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321.

Zhou, H., Wei, D. and Yao, F. (2022). Theory of functional principal component analysis for discretely observed data. arXiv preprint. Available at arXiv:2209.08768.

Bernoulli **32**(4), 2026, 3277–3299
https://doi.org/10.3150/26-BEJ1988

Stratified permutational Berry–Esseen bounds and their applications to statistics

PENGFEI TIAN[1,a], FAN YANG[2,b] and PENG DING[3,c]

[1]*Qiuzhen College, Tsinghua University, Beijing 100084, China,* [a]*tpf24@mails.tsinghua.edu.cn*
[2]*Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China,*
[b]*yangfan1987@tsinghua.edu.cn*
[3]*Department of Statistics, University of California, Berkeley, Berkeley, California 94720, USA,*
[c]*pengdingpku@berkeley.edu*

The stratified linear permutation statistic arises in various statistics problems, including stratified and post-stratified survey sampling, stratified and post-stratified experiments, conditional permutation tests, etc. Although we can derive the Berry–Esseen bounds for the stratified linear permutation statistic based on existing bounds for the non-stratified statistics, those bounds are not sharp, and moreover, this strategy does not work in general settings with heterogeneous strata with varying sizes. We first use Stein's method to obtain a unified stratified permutational Berry–Esseen bound that can accommodate heterogeneous strata. We then apply the bound to various statistics problems, leading to stronger theoretical quantifications and thereby facilitating statistical inference in those problems.

Keywords: Causal inference; design-based inference; experimental design; post-stratification; Stein's method; survey sampling; zero-bias transformation

References

Abadie, A. and Imbens, G.W. (2008). Estimation of the conditional variance in paired experiments. *Ann. Écon. Stat.* **91/92** 175–87.
Berry, A.C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** 122–136.
Bickel, P.J. and Freedman, D.A. (1984). Asymptotic normality and the bootstrap in stratified sampling. *Ann. Statist.* **12** 470–482.
Bolthausen, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** 379–386.
Chen, L.H. (2021). Stein's method of normal approximation: Some recollections and reflections. *Ann. Statist.* **49** 1850–1863.
Chen, L.H. and Fang, X. (2015). On the error bound in a combinatorial central limit theorem. *Bernoulli* **21** 335–359.
Chen, L.H., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. New York: Springer.
Cochran, W.G. (1977). *Sampling Techniques*, 3rd ed. New York: Wiley.
de Chaisemartin, C. and Deeb, A. (2024). Estimating treatment-effect heterogeneity across sites in multi-site randomized experiments with imperfect compliance. arXiv preprint. Available at arXiv:2405.17254.
D'Haultfœuille, X. and Tuvaandorj, P. (2024). A robust permutation test for subvector inference in linear regressions. *Quant. Econ.* **15** 27–87.
Ding, P. (2024). *A First Course in Causal Inference*. Boca Raton: CRC Press.
Esseen, C.-G. (1942). On the Liapunov limit error in the theory of probability. *Ark. Mat. Astron. Fys.* **28** 1–19.
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.
Fogarty, C.B. (2018a). Regression-assisted inference for the average treatment effect in paired experiments. *Biometrika* **105** 994–1000.

Fogarty, C.B. (2018b). On mitigating the analytical limitations of finely stratified experiments. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 1035–1056.

Fraser, D.A.S. (1956). A vector form of the Wald–Wolfowitz–Hoeffding theorem. *Ann. Math. Stat.* **27** 540–543.

Fuller, W.A. (1966). Estimation employing post strata. *J. Amer. Statist. Assoc.* **61** 1172–1183.

Goldstein, L. (2005). Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **42** 661–683.

Goldstein, L. (2007). L^1 bounds in normal approximation. *Ann. Probab.* **35** 1888–1930.

Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952.

Hájek, J. (1961). Some extensions of the Wald–Wolfowitz–Noether theorem. *Ann. Math. Stat.* **32** 506–523.

Higgins, M.J., Sävje, F. and Sekhon, J.S. (2016). Improving massive experiments with threshold blocking. *Proc. Natl. Acad. Sci. USA* **113** 7369–7376.

Ho, S.-T. and Chen, L.H. (1978). An L_p bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.* **6** 231–249.

Hoeffding, W. (1951). A combinatorial central limit theorem. *Ann. Math. Stat.* **22** 558–566.

Holt, D. and Smith, T. (1979). Post stratification. *J. R. Stat. Soc. A, General* **142** 33–46.

Imai, K. (2008). Variance identification and efficiency analysis in randomized experiments under the matched-pair design. *Stat. Med.* **27** 4857–4873.

Imai, K., King, G. and Stuart, E.A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *J. Roy. Statist. Soc., Ser. A, Statist. Soc.* **171** 481–502.

Imai, K., King, G., Nall, C. et al. (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.

Imbens, G.W. and Rosenbaum, P.R. (2005). Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *J. Roy. Statist. Soc., Ser. A, Statist. Soc.* **168** 109–126.

Imbens, G.W. and Rubin, D.B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York: Cambridge University Press.

Lahiri, S.N., Chatterjee, A. and Maiti, T. (2007). Normal approximation to the hypergeometric distribution in nonstandard cases and a sub-Gaussian Berry–Esseen theorem. *J. Statist. Plann. Inference* **137** 3570–3590.

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.

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.

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *Ann. Appl. Stat.* **7** 295–318.

Little, R.J. (1993). Post-stratification: A modeler's perspective. *J. Amer. Statist. Assoc.* **88** 1001–1012.

Liu, H., Ren, J. and Yang, Y. (2024). Randomization-based joint central limit theorem and efficient covariate adjustment in randomized block 2^K factorial experiments. *J. Amer. Statist. Assoc.* **119** 136–150.

Liu, H. and Yang, Y. (2020). Regression-adjusted average treatment effect estimators in stratified randomized experiments. *Biometrika* **107** 935–948.

McHugh, R. and Matts, J. (1983). Post-stratification in the randomized clinical trial. *Biometrics* **39** 217–225.

Miratrix, L.W., Sekhon, J.S. and Yu, B. (2013). Adjusting treatment effect estimates by post-stratification in randomized experiments. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 369–396.

Móri, T.F. (2009). Sharp inequalities between centered moments. *J. Inequal. Pure Appl. Math.* **10** 1–19.

Motoo, M. (1956). On the Hoeffding's combinatrial central limit theorem. *Ann. Inst. Statist. Math.* **8** 145–154.

Noether, G.E. (1949). On a theorem by Wald and Wolfowitz. *Ann. Math. Stat.* **20** 455–458.

Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality 192*. Cambridge University Press.

Panaretos, V.M. and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** 405–431.

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.

Rosenbaum, P.R. (2002). *Obs. Stud.*, 2nd ed. New York: Springer.

Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293.

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.

Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory* **6** 583–603. University of California Press.

Tian, P., Ren, J. and Ma, Y. (2026). Balancing Covariates in Survey Experiments. arXiv preprint. Available at arXiv:2602.07390.

Tian, P., Yang, F. and Ding, P. (2026). Supplement to "Stratified Permutational Berry–Esseen Bounds and Their Applications to Statistics." https://doi.org/10.3150/26-BEJ1988SUPP

Tuvaandorj, P. (2024). A combinatorial central limit theorem for stratified randomization. arXiv preprint. Available at arXiv:2402.14764.

von Bahr, B. (1976). Remainder term estimate in a combinatorial limit theorem. *Z. Wahrsch. Verw. Gebiete* **35** 131–139.

Wald, A. and Wolfowitz, J. (1944). Statistical tests based on permutations of the observations. *Ann. Math. Stat.* **15** 358–372.

Wang, Y. and Li, X. (2022). Rerandomization with diminishing covariate imbalance and diverging number of covariates. *Ann. Statist.* **50** 3439–3465.

Bernoulli **32**(4), 2026, 3300–3326
https://doi.org/10.3150/26-BEJ1989

Estimating the hyperuniformity exponent of spatial point processes

GABRIEL MASTRILLI[1,2,a], BARTŁOMIEJ BŁASZCZYSZYN[2,c] and FRÉDÉRIC LAVANCIER[1,b]

[1] *Univ Rennes, Ensai, CNRS, CREST, Rennes, France,* [a]*gabriel.mastrilli@ensai.fr,* [b]*frederic.lavancier@ensai.fr*
[2] *Inria, ENS/PSL Univ, Paris, France,* [c]*bartek.blaszczyszyn@ens.fr*

We address the challenge of estimating the hyperuniformity exponent α of a spatial point process, given only one realization of it. Assuming that the structure factor S of the point process follows a vanishing power law at the origin (the typical case of a hyperuniform point process), this exponent is defined as the slope near the origin of $\log S$. Our estimator is built upon the (expanding window) asymptotic variance of some shot-noise wavelet transforms, of the point process. By combining several scales and several wavelets, we develop a multi-scale, multi-taper estimator $\widehat{\alpha}$. We analyze its asymptotic behavior, proving its consistency under various settings, and enabling the construction of asymptotic confidence intervals for α when $\alpha < d$ and under Brillinger mixing. This construction is derived from a multivariate central limit theorem where the normalisations are non-standard and vary among the components. We also present a non-asymptotic deviation inequality providing insights into the influence of tapers on the bias-variance trade-off of $\widehat{\alpha}$. Finally, we investigate the performance of $\widehat{\alpha}$ through simulations, and we apply our method to the analysis of hyperuniformity in a real dataset of marine algae.

Keywords: Hyperuniformity; scattering intensity; spatial point process; structure factor; wavelets

References

Abramowitz, M. and Stegun, I.A. (1968). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables 55*. US Government printing office.

Baccelli, F., Blaszczyszyn, B. and Karray, M. (2020). *Random Measures, Point Processes, and Stochastic Geometry*. Inria. hal-02460214.

Bartlett, M.S. (1963). The spectral analysis of point processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **25** 264–281.

Bartlett, M.S. (1964). The spectral analysis of two-dimensional point processes. *Biometrika* **51** 299–311.

Biscio, C.A.N. and Lavancier, F. (2016a). Quantifying repulsiveness of determinantal point processes. *Bernoulli* **22** 2001–2028.

Biscio, C.A.N. and Lavancier, F. (2016b). Brillinger mixing of determinantal point processes and statistical applications. *Electron. J. Stat.* **10** 582–607.

Biscio, C.A.N. and Lavancier, F. (2017). Contrast estimation for parametric stationary determinantal point processes. *Scand. J. Stat.* **44** 204–229.

Björklund, M. and Hartnick, T. (2024). Hyperuniformity and non-hyperuniformity of quasicrystals. *Math. Ann.* **389** 365–426.

Boursier, J. (2022). Decay of correlations and thermodynamic limit for the circular Riesz gas. Available at arXiv:2209.00396.

Brémaud, P. (2014). *Fourier Analysis of Stochastic Processes*. Cham: Springer.

Brillinger, D.R. (1972). The spectral analysis of stationary interval functions. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics* **6** 483–514. University of California Press.

Brillinger, D.R. (1997). Some wavelet analyses of point process data. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. No. 97CB36136)* **2** 1087–1091. IEEE.

1350-7265

Brochard, A. (2022). Wavelet-based representations of point processes for modelling and statistical learning. PhD thesis, Ecole Normale Supérieure (Paris).
Chatterjee, S. (2019). Rigidity of the three-dimensional hierarchical Coulomb gas. *Probab. Theory Related Fields* **175** 1123–1176.
Chiu, S.N., Stoyan, D., Kendall, W.S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*. John Wiley & Sons.
Cohen, E.A.K. and Gibberd, A.J. (2021). Wavelet spectra for multivariate point processes. *Biometrika* **109** 837–851.
Coste, S. (2021). Order, fluctuations, rigidities. Unpublished manuscript.
Daley, D.J. and Vere-Jones, D. (2007). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. New York, NY: Springer.
Daley, D.J., Vere-Jones, D. et al. (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. New York, NY: Springer.
Degl'Innocenti, R., Shah, Y., Masini, L., Ronzani, A., Pitanti, A., Ren, Y., Jessop, D., Tredicucci, A., Beere, H. and Ritchie, D. (2015). THz quantum cascade lasers based on a hyperuniform design. In *Quantum Sensing and Nanophotonic Devices XII* **9370** 27–32. SPIE.
Dereudre, D. and Flimmel, D. (2024). Non-hyperuniformity of Gibbs point processes with short-range interactions. *J. Appl. Probab.* **61** 1380–1406.
Dereudre, D., Hardy, A., Leblé, T. and Maïda, M. (2021). DLR equations and rigidity for the sine-beta process. *Comm. Pure Appl. Math.* **74** 172–222.
Dong, L. (2023). Hyperuniform organization in human settlements. Available at arXiv:2306.04149.
Evans, J.W. (1993). Random and cooperative sequential adsorption. *Rev. Modern Phys.* **65** 1281.
Faÿ, G., Moulines, E., Roueff, F. and Taqqu, M.S. (2009). Estimators of long-memory: Fourier versus wavelets. *J. Econometrics* **151** 159–177.
Florescu, M., Torquato, S. and Steinhardt, P.J. (2009). Designer disordered materials with large, complete photonic band gaps. *Proc. Natl. Acad. Sci. USA* **106** 20658–20663.
Folland, G.B. (2009). *Fourier Analysis and Its Applications 4*. American Mathematical Soc.
Froufe-Pérez, L.S., Engel, M., Damasceno, P.F., Muller, N., Haberko, J., Glotzer, S.C. and Scheffold, F. (2016). Role of short-range order and hyperuniformity in the formation of band gaps in disordered photonic materials. *Phys. Rev. Lett.* **117** 053902.
Gabrielli, A. (2004). Point processes and stochastic displacement fields. *Phys. Rev. E* **70** 066131.
Gaultier, L. (2016). Fluctuations of smooth linear statistics of determinantal point processes. PhD thesis, KTH Royal Institute of Technology.
Ghosh, S. and Lebowitz, J.L. (2017). Fluctuations, large deviations and rigidity in hyperuniform systems: A brief survey. *Indian J. Pure Appl. Math.* **48** 609–631.
Ginibre, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449.
Gorsky, S., Britton, W., Chen, Y., Montaner, J., Lenef, A., Raukas, M. and Dal Negro, L. (2019). Engineered hyperuniformity for directional light extraction. *APL Photon.* **4**.
Grainger, J.P., Rajala, T.A., Murrell, D.J. and Olhede, S.C. (2025). Spectral estimation for point processes and random fields. *Biometrika*.
Haimi, A. and Romero, J.L. (2024). Normality of smooth statistics for planar determinantal point processes. *Bernoulli* **30** 666–682.
Hawat, D., Gautier, G., Bardenet, R. and Lachièze-Rey, R. (2023). On estimating the structure factor of a point process, with applications to hyperuniformity. *Stat. Comput.* **61**.
Heinrich, L. (2012). Asymptotic methods in statistics of random point processes. In *Stochastic Geometry, Spatial Statistics and Random Fields: Asymptotic Methods* 115–150. Springer.
Heinrich, L. (2016). On the strong Brillinger-mixing property of α-determinantal point processes and some applications. *Appl. Math.* **61** 443–461.
Heinrich, L. and Klein, S. (2011). Central limit theorem for the integrated squared error of the empirical second-order product density and goodness-of-fit tests for stationary point processes. *Stat. Risk Model.* **28** 359–387.
Heinrich, L. and Pawlas, Z. (2013). Absolute regularity and Brillinger-mixing of stationary point processes. *Lith. Math. J.* **53** 293–310.
Heinrich, L. and Schmidt, V. (1985). Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. Appl. Probab.* **17** 709–730.

Hexner, D. and Levine, D. (2017). Noise, diffusion, and hyperuniformity. *Phys. Rev. Lett.* **118** 020601.

Huang, M., Hu, W., Yang, S., Liu, Q.-X. and Zhang, H. (2021). Circular swimming motility and disordered hyperuniform state in an algae system. *Proc. Natl. Acad. Sci. USA* **118** e2100493118.

Jiao, Y., Lau, T., Hatzikirou, H., Meyer-Hermann, M., Corbo, J.C. and Torquato, S. (2014). Avian photoreceptor patterns represent a disordered hyperuniform solution to a multiscale packing problem. *Phys. Rev. E* **89** 022721.

Jolivet, E. (1981). Central limit theorem and convergence of empirical processes for stationary point processes. In *Point Processes and Queuing Problems (Colloqium, Keszthely, 1978)*. *Colloq. Math. Soc. János Bolyai* **24** 117–161. Amsterdam-New York: North-Holland. (82m:60059). MR617406

Kanter, M. (1975). Stable densities under change of scale and total variation inequalities. *Ann. Probab.* 697–707.

Kim, J. and Torquato, S. (2018). Effect of imperfections on the hyperuniformity of many-body systems. *Phys. Rev. B* **5** 054105.

Klatt, M.A., Kim, J. and Torquato, S. (2020). Cloaking the underlying long-range order of randomly perturbed lattices. *Phys. Rev. E* **101** 032118.

Klatt, M.A., Last, G. and Henze, N. (2022). A genuine test for hyperuniformity. Available at arXiv:2210.12790.

Klatt, M.A., Last, G. and Yogeshwaran, D. (2020). Hyperuniform and rigid stable matchings. *Random Structures Algorithms* **57** 439–473.

Kunz, H. (1974). The one-dimensional classical electron gas. *Ann. Phys.* **85** 303–335.

Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process 7*. Cambridge University Press.

Lavancier, F., Møller, J. and Rubak, E. (2015). Determinantal point process models and statistical inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 853–877.

Leblé, T. (2021). The two-dimensional one-component plasma is hyperuniform. Available at arXiv:2104.05109.

Lei, Q.-L. and Ni, R. (2019). Hydrodynamics of random-organizing hyperuniform fluids. *Proc. Natl. Acad. Sci. USA* **116** 22983–22989.

Leseur, O., Pierrat, R. and Carminati, R. (2016). High-density hyperuniform materials can be transparent. *Optica* **3** 763–767.

Lin, C., Steinhardt, P.J. and Torquato, S. (2017). Hyperuniformity variation with quasicrystal local isomorphism class. *J. Phys., Condens. Matter* **29** 204003.

Macchi, O. (1975). The coincidence approach to stochastic point processes. *Adv. Appl. Probab.* **7** 83–122.

Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Elsevier.

Martelli, F., Torquato, S., Giovambattista, N. and Car, R. (2017). Large-scale structure and hyperuniformity of amorphous ices. *Phys. Rev. Lett.* **119** 136002.

Mastrilli, G. (2026). Asymptotic fluctuations of smooth linear statistics of independently perturbed lattices. *J. Stat. Phys.* **193** 24.

Mastrilli, G., Błaszczyszyn, B. and Lavancier, F. (2026). Supplement to "Estimating the hyperuniformity exponent of spatial point processes." https://doi.org/10.3150/26-BEJ1989SUPPA, https://doi.org/10.3150/26-BEJ1989SUPPB

Mayer, A., Balasubramanian, V., Mora, T. and Walczak, A.M. (2015). How a well-adapted immune system is organized. *Proc. Natl. Acad. Sci. USA* **112** 5950–5955.

Møller, J. and O'Reilly, E. (2021). Couplings for determinantal point processes and their reduced Palm distributions with a view to quantifying repulsiveness. *J. Appl. Probab.* **58** 469–483.

Møller, J. and Waagepetersen, R.P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. CRC press.

Mugglestone, M.A. and Renshaw, E. (1996). A practical guide to the spectral analysis of spatial point processes. *Comput. Statist. Data Anal.* **21** 43–65.

Muller, N., Haberko, J., Marichy, C. and Scheffold, F. (2013). Silicon hyperuniform disordered photonic materials with a pronounced gap in the shortwave infrared. *Adv. Opt. Mater.* **2** 115–119.

Nazarov, F. and Sodin, M. (2011). Fluctuations in random complex zeroes: Asymptotic normality revisited. *Int. Math. Res. Not.* **2011** 5720–5759.

Oğuz, E.C., Socolar, J.E., Steinhardt, P.J. and Torquato, S. (2017). Hyperuniformity of quasicrystals. *Phys. Rev. B* **95** 054119.

Percival, D.B. and Walden, A.T. (2000). *Wavelet Methods for Time Series Analysis 4*. Cambridge University Press.

Percival, D.B. and Walden, A.T. (2020). *Spectral Analysis for. Univariate Time Series* **51**. Cambridge University Press.

Philcox, O.H. and Torquato, S. (2023). Disordered heterogeneous universe: Galaxy distribution and clustering across length scales. *Phys. Rev. X* **13** 011038.

Rajala, T.A., Olhede, S.C., Grainger, J.P. and Murrell, D.J. (2023). What is the Fourier transform of a spatial point process? *IEEE Trans. Inf. Theory*.

Ramaswamy, S. (2010). The mechanics and statistics of active matter. *Annu. Rev. Condens. Matter Phys.* **1** 323–345.

Rider, B. and Virag, B. (2007). Complex determinantal processes and $H1$ noise. *Electron. J. Probab.* **12** 1238–1257.

Riedel, K.S. and Sidorenko, A. (1995). Minimum bias multiple taper spectral estimation. *IEEE Trans. Signal Process.* **43** 188–195.

Sodin, M. and Tsirelson, B. (2004). Random complex zeroes, I. Asymptotic normality. *Israel J. Math.* **144** 125–149.

Soshnikov, A. (2000). Determinantal random point fields. *Russian Math. Surveys* **55** 923.

Soshnikov, A. (2002). Gaussian limit for determinantal random point fields. *Ann. Probab.* 171–187.

Taleb, Y. and Cohen, E.A.K. (2021). Multiresolution analysis of point processes and statistical thresholding for Haar wavelet-based intensity estimation. *Ann. Inst. Statist. Math.* **73** 395–423.

Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. *Proc. IEEE* **70** 1055.

Torquato, S. (2016). Hyperuniformity and its generalizations. *Phys. Rev. E* **94** 022122.

Torquato, S. (2018). Hyperuniform states of matter. *Phys. Rep.* **745** 1–95.

Torquato, S. and Stillinger, F.H. (2003). Local density fluctuations, hyperuniformity, and order metrics. *Phys. Rev. E* **68** 041113.

Uchaikin, V. and Saenko, V. (2002). Simulation of random vectors with isotropic fractional stable distributions and calculation of their probability density functions. *J. Math. Sci.* **112** 4211–4228.

Valkó, B. and Virág, B. (2009). Continuum limits of random matrices and the Brownian carousel. *Invent. Math.* **177** 463–508.

Walden, A.T. (2000). A unified view of multitaper multivariate spectral estimation. *Biometrika* **87** 767–788.

Weijs, J.H., Jeanneret, R., Dreyfus, R. and Bartolo, D. (2015). Emergent hyperuniformity in periodically driven emulsions. *Phys. Rev. Lett.* **115** 108301.

Wilken, S., Chaderjian, A. and Saleh, O.A. (2023). Spatial organization of phase-separated DNA droplets. *Phys. Rev. X* **13** 031014.

Yang, J. and Guan, Y. (2026). Fourier analysis of spatial point processes. *Bernoulli* **32** 370–392.

Yu, S. (2023). Evolving scattering networks for engineering disorder. *Nat. Comput. Sci.* **3** 128–138.

Zachary, C.E. and Torquato, S. (2009). Hyperuniformity in point patterns and two-phase random heterogeneous media. *J. Stat. Mech. Theory Exp.* **2009** P12015.

Zhu, Z., Hallet, B., Sipos, A.A., Domokos, G. and Liu, Q.-X. (2023). Dispersed pebbles on Mars: An unexpected setting for hyperuniformity. Available at arXiv:2312.13818.

Bernoulli **32**(4), 2026, 3327–3353
https://doi.org/10.3150/26-BEJ1990

Detecting practically significant dependencies in metric spaces via distance correlations

HOLGER DETTE[a] and MARIUS KROLL[b]

Faculty of Mathematics, Ruhr-Universität Bochum, Bochum, Germany, [a]*holger.dette@rub.de,*
[b]*marius.kroll@rub.de*

We take a different look at the problem of testing the independence of two metric-space-valued random variables using the distance correlation. Instead of testing if the distance correlation vanishes exactly, we are interested in the hypothesis that it does not exceed a certain threshold. Our testing problem is motivated by the observation that in many cases it is more reasonable to test for a practically significant dependency since it is rare that a hypothesis of perfect independence is exactly satisfied. This point of view also reflects statistical practice, where one often classifies the strength of the association in categories such as 'small', 'medium' and 'large' and the precise definitions depend on the specific application. To address these problems we develop a pivotal test for the hypothesis that the distance correlation between two random variables does not exceed a pre-specified threshold Δ. We also determine a minimum value $\hat{\Delta}_\alpha$ from the data such that the hypothesis is rejected for all $\Delta \leq \hat{\Delta}_\alpha$ at controlled type I error α. This quantity can be interpreted as a measure of evidence against the hypothesis that the distance correlation is less or equal than' Δ. The new test is applicable to processes taking values in separable metric spaces of strong negative type, covering Euclidean as well as functional data. We do not assume independent observations, and instead prove our results for absolutely regular sample generating processes, which includes many time series such as ARMA and GARCH models. Our approach is based on a new functional limit theorem for the sequential distance correlation process, and can also be used to construct confidence intervals for the distance correlation without the need for resampling.

Keywords: Distance correlation; distance covariance; relevant hypotheses; self-normalisation; tests for independence

References

Albert, M., Bouret, Y., Fromont, M. and Reynaud-Bouret, P. (2015). Bootstrap and permutation tests of independence for point processes. *Ann. Statist.* **43** 2537–2564.
Arcones, M.A. (1998). The law of large numbers for U-statistics under absolute regularity. *Electron. Commun. Probab.* **3** 13–19.
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.
Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29** 1165–1188.
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.
Betken, A. and Dehling, H. (2024). Distance Correlation for Long-Range Dependent Time Series. Preprint. Available at arXiv:2107.03041.
Betken, A., Dehling, H. and Kroll, M. (2024). A Bootstrap Test for Independence of Time Series Based on the Distance Covariance. Preprint. Available at arXiv:2112.14091.
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.
Bosco, F.A., Aguinis, H., Singh, K., Field, J.G. and Pierce, C.A. (2015). Correlational effect size benchmarks. *J. Appl. Psychol.* **100** 431–449.

Bradley, R.C. (2007). *Introduction to Strong Mixing Conditions 1*. Heber City: Kendrick Press.

Brydges, C.R. (2019). Effect size guidelines, sample size calculations, and statistical power in gerontology. *Innov. Aging* **3** igz036.

Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18** 17–39.

Chatterjee, S. (2021). A new coefficient of correlation. *J. Amer. Statist. Assoc.* **116** 2009–2022.

Chu, B. (2023). A distance-based test of independence between two multivariate time series. *J. Multivariate Anal.* **195** 105151.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Davis, R.A., Matsui, M., Mikosch, T. and Wan, P. (2018). Applications of distance correlation to time series. *Bernoulli* **24** 3087–3116.

Dehling, H., Matsui, M., Mikosch, T., Samorodnitsky, G. and Tafakori, L. (2020). Distance covariance for discretized stochastic processes. *Bernoulli* **26** 2758–2789.

Denker, M. and Keller, G. (1983). On U-statistics and v. Mises' statistics for weakly dependent processes. *Z. Wahrsch. Verw. Gebiete* **64** 505–522.

Dette, H. and Kroll, M. (2026). Supplement to "Detecting practically significant dependencies in metric spaces via distance correlations." https://doi.org/10.3150/26-BEJ1990SUPP

Dette, H., Siburg, K.F. and Stoimenov, P.A. (2012). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat.* **40** 21–41.

Dette, H. and Tang, J. (2024). New energy distances for statistical inference on infinite dimensional Hilbert spaces without moment conditions.

Doukhan, P. (1994). *Mixing: Properties and Examples*. Berlin – New York – Heidelberg: Springer Verlag.

Edelmann, D., Fokianos, K. and Pitsillou, M. (2019). An updated literature review of distance correlation and its applications to time series. *Int. Stat. Rev.* **87** 237–262.

Fokianos, K. and Pitsillou, M. (2017). Consistent testing for pairwise dependence in time series. *Technometrics* **59** 262–270.

Geenens, G. and Lafaye de Micheaux, P. (2020). The Hellinger correlation. *J. Amer. Statist. Assoc.* 1–15.

Gretton, A. and Györfi, L. (2010). Consistent nonparametric tests of independence. *J. Mach. Learn. Res.* **11** 1391–1423.

Gretton, A., Bousquet, O., Smola, A. and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory* (S. Jain, H.U. Simon and E. Tomita, eds.) 63–77. Berlin, Heidelberg: Springer Berlin Heidelberg.

Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2008). A kernel statistical test of independence. In *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer and S. Roweis, eds.) **20** 585–592. Curran Associates, Inc.

Heller, R., Heller, Y. and Gorfine, M. (2012). A consistent multivariate test of association based on ranks of distances. *Biometrika* **100** 503–510.

Hemphill, J. (2003). Interpreting the magnitude of correlation coefficients. *Amer. Psychol.* **58** 78–9.

Hoeffding, W. (1948a). A non-parametric test of independence. *Ann. Math. Stat.* **19** 546–557.

Hoeffding, W. (1948b). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325.

Hosking, J.R.M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.* **75** 602–608.

Huang, W., Tang, M., Chen, Y.-L., Zhang, T.-L., Hong, T., Li, J., Lv, G.-H., Yan, Y., Ouyang, Z.-H. and Zou, M.-X. (2022). Incidence and risk factors for cerebrovascular-specific mortality in patients with colorectal cancer: A registry-based cohort study involving 563298 patients. *Cancers* **14** 2053.

Janson, S. (2021). On distance covariance in metric and Hilbert spaces. *ALEA Lat. Amer. J. Probab. Math. Stat.* **18** 1353–1393.

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

Kroll, M. (2022). Asymptotic behaviour of the empirical distance covariance for dependent data. *J. Theor. Probab.* **35** 1226–1246.

Kroll, M. (2023). The Empirical Distance Covariance of Weakly Dependent Data PhD thesis, Ruhr-Universität, Bochum. https://doi.org/10.13154/294-9840

Kroll, M. (2026). relevant-dependence. GitHub Code. Available at https://github.com/marius-kroll/relevant-dependence.

Kuelbs, J. and Philipp, W. (1980). Almost sure invariance principles for partial sums of mixing B-valued random variables. *Ann. Probab.* **8** 1003–1036.

Lovakov, A. and Agadullina, E. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *Eur. J. Soc. Psychol.* **51** 485–504.

Lyons, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305.

Lyons, R. (2014). Hyperbolic space has strong negative type. *Illinois J. Math.* **58** 1009–1013.

McDonald, D.J., Shalizi, C.R. and Schervish, M. (2015). Estimating beta-mixing coefficients via histograms. *Electron. J. Stat.* **9** 2855–2883.

Mokkadem, A. (1988). Mixing properties of ARMA processes. *Stoch. Process. Appl.* **29** 309–315.

Pearson, K. (1920). Notes on the history of correlation. *Biometrika* **13** 25–45.

Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press.

Quintana, D. (2016). Statistical considerations for reporting and planning heart rate variability case-control studies. *Psychophysiology* **54**.

Rizzo, M. and Székely, G. (2024). energy: E-Statistics: Multivariate Inference via the Energy of Data R package version 1.7-12. https://doi.org/10.32614/CRAN.package.energy

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.

Sen, A. and Sen, B. (2014). Testing independence and goodness-of-fit in linear models. *Biometrika* **101** 927–942.

Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. New York – Chichester – Brisbane – Toronto – Singapore: John Wiley & Sons.

Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15** 72–101.

Székely, G.J. and Rizzo, M.L. (2009). Brownian distance covariance. *Ann. Appl. Stat.* **3** 1236–1265.

Székely, G.J. and Rizzo, M.L. (2012). On the uniqueness of distance covariance. *Statist. Probab. Lett.* **82** 2278–2282.

Székely, G.J. and Rizzo, M.L. (2013). The distance correlation t-test of independence in high dimension. *J. Multivariate Anal.* **117** 193–213.

Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794.

Timan, A.F. and Vestfrid, I.A. (1983). Any separable ultrametric space can be isomterically imbedded in l_2. *Funct. Anal. Appl.* **17** 70–71.

Tsaparas, P., Mariño-Ramírez, L., Bodenreider, O., Koonin, E.V. and Jordan, I.K. (2006). Global similarity and local divergence in human and mouse gene co-expression networks. *BMC Evol. Biol.* **6** 70.

Tse, Y.K. (2002). Residual-based diagnostics for conditional heteroscedasticity models. *J. Econometrics* **108** 191–212.

Tukey, J.W. (1991). The philosophy of multiple comparisons. *Statist. Sci.* **6** 100–116.

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. New York – Berlin – Heidelberg: Springer-Verlag.

Volkonskii, V.A. and Rozanov, Y.A. (1959). Some limit theorems for random functions. I. *Theory Probab. Appl.* **4** 178–197.

Volkonskii, V.A. and Rozanov, Y.A. (1961). Some limit theorems for random functions. II. *Theory Probab. Appl.* **6** 186–198.

Wang, G., Li, W.K. and Zhu, K. (2021). New HSIC-based tests for independence between two stationary multivariate time series. *Statist. Sinica* **31** 269–300.

Wellek, S. (2010). *Testing Statistical Hypotheses of Equivalence and Noninferiority*. CRC Press.

Yao, S., Zhang, X. and Shao, X. (2018). Testing mutual independence in high dimension via distance covariance. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 455–480.

Zhang, X., Song, L., Gretton, A. and Smola, A. (2008). Kernel measures of independence for non-iid data. In *Advances in Neural Information Processing Systems* (D. Koller, D. Schuurmans, Y. Bengio and L. Bottou, eds.) **21**. Curran Associates, Inc.

Zhou, Z. (2012). Measuring nonlinear dependence in time-series, a distance correlation approach. *J. Time Series Anal.* **33** 438–457.

Bernoulli **32**(4), 2026, 3354–3381
https://doi.org/10.3150/26-BEJ1991

Structural adaptation and rate accelerated estimation in bivariate functional data

OMAR KASSI[a] and SUNNY G.W. WANG[b]

Univ Rennes, Ensai, CNRS, CREST-UMR 9194, F-35000 Rennes, France, [a]*omar.kassi@ensai.fr,* [b]*sunny.wang@ensai.fr*

We introduce directional regularity, a new definition of anisotropy for multivariate functional data. Instead of taking the conventional view, which determines anisotropy as a notion of smoothness along a dimension, directional regularity additionally views anisotropy through the lens of directions. We show that faster rates of convergence for smoothing can be obtained through a change-of-basis by adapting to the anisotropy of a bivariate process. An algorithm for the estimation and identification of the change-of-basis matrix is constructed, made possible due to the replication structure of functional data. Non-asymptotic bounds are provided for our algorithm, supplemented by numerical evidence from an extensive simulation study. Finally, a real-world rainfall measurement dataset is analyzed with our methods.

Keywords: Anisotropy; adaptive algorithms; effective smoothness; multivariate functional data

References

Ammous, S., Dedecker, J. and Duval, C. (2024). Adaptive directional estimator of the density in $\mathbb{R}^d$ for independent and mixing sequences. *J. Multivariate Anal.* **203** 105332. https://doi.org/10.1016/j.jmva.2024.105332

Amorino, C. (2021). Rate of estimation for the stationary distribution of jump-processes over anisotropic Holder classes. *Electron. J. Stat.* **15** 5067–5116. MR4347384 https://doi.org/10.1214/21-ejs1913

Azzimonti, L., Sangalli, L.M., Secchi, P., Domanin, M. and Nobile, F. (2015). Blood flow velocity field estimation via spatial regression with PDE penalization. *J. Amer. Statist. Assoc.* **110** 1057–1071. MR3420684 https://doi.org/10.1080/01621459.2014.946036

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.

Bernardi, M.S., Carey, M., Ramsay, J.O. and Sangalli, L.M. (2018). Modeling spatial anisotropy via regression with partial differential regularization. *J. Multivariate Anal.* **167** 15–30.

Bertin, K. (2004). Asymptotically exact minimax estimation in sup-norm for anisotropic Hölder classes. *Bernoulli* **10** 873–888. MR2093615 https://doi.org/10.3150/bj/1099579160

Biermé, H., Moisan, L. and Richard, F. (2015). A turning-band method for the simulation of anisotropic fractional Brownian fields. *J. Comput. Graph. Statist.* **24** 885–904. MR3397238 https://doi.org/10.1080/10618600.2014.946603

Coeurjolly, J.-F. and Porcu, E. (2018). Fast and exact simulation of complex-valued stationary Gaussian processes through embedding circulant matrix. *J. Comput. Graph. Statist.* **27** 278–290. MR3816264 https://doi.org/10.1080/10618600.2017.1385468

Daly, C., Bryant, K., University, O.S. and Service, U.N.R.C. (2013). The PRISM climate and weather system – an introduction. Technical report.

Davies, S. and Hall, P. (1999). Fractal analysis of surface roughness by using spatial data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 3–37. With discussion and a reply by the authors. MR1664088 https://doi.org/10.1111/1467-9868.00160

Dieker, T. (2004). Simulation of fractional brownian motion. Master's thesis, University of Twente.

Fan, Y. and Guerre, E. (2016). Multivariate local polynomial estimators: Uniform boundary properties and asymptotic linear representation. In *Essays in Honor of Aman Ullah* **36** 489–537. Emerald Group Publishing Limited.

1350-7265

Friedland, C.J., Joyner, T.A., Massarra, C., Rohli, R.V. et al. (2016). Isotropic and anisotropic Kriging approaches for interpolating surface-level wind speeds across large, geographically diverse regions. *Geomat. Nat. Hazards Risk* **8** 1–18. https://doi.org/10.1080/19475705.2016.1185749

Germain, C., Da Costa, J.P., Lavialle, O. and Baylou, P. (2003). Multiscale estimation of vector field anisotropy application to texture characterization. *Signal Process.* **83** 1487–1503. https://doi.org/10.1016/S0165-1684(03)00064-1

Golovkine, S., Klutchnikoff, N. and Patilea, V. (2022). Learning the smoothness of noisy curves with application to online curve estimation. *Electron. J. Stat.* **16** 1485–1560. MR4390502 https://doi.org/10.1214/22-EJS1997

Golovkine, S., Klutchnikoff, N. and Patilea, V. (2025). Adaptive estimation of irregular mean and covariance functions. *Bernoulli* **31** 1032–1057. MR4863066 https://doi.org/10.3150/24-bej1759

Greenberg, S., Aladjem, M. and Kogan, D. (2002). Fingerprint image enhancement using filtering techniques. *Real-Time Imaging* **8** 227–236. https://doi.org/10.1006/rtim.2002.0279

Hart, E.M. and Bell, K. (2015). PRISM: download data from the Oregon state prism climate project R package version 0.2.3. https://doi.org/10.32614/CRAN.package.prism

Herbin, E. (2006). From *N* parameter fractional Brownian motions to *N* parameter multifractional Brownian motions. *Rocky Mountain J. Math.* **36** 1249–1284. MR2274895 https://doi.org/10.1216/rmjm/1181069415

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. *Springer Series in Statistics*. New York: Springer. MR2920735 https://doi.org/10.1007/978-1-4614-3655-3

Jain, A., Hong, L. and Bolle, R. (1997). On-line fingerprint verification. *IEEE Trans. Pattern Anal. Mach. Intell.* **19** 302–314. https://doi.org/10.1109/34.587996

Jiang, X. (2005). On orientation and anisotropy estimation for online fingerprint authentication. *IEEE Trans. Signal Process.* **53** 4038–4049. MR2246664 https://doi.org/10.1109/TSP.2005.855417

Kassi, O., Klutchnikoff, N. and Patilea, V. (2025). Learning the regularity of multivariate functional data. *Electron. J. Stat.* **19** 4174–4229. MR4958400 https://doi.org/10.1214/25-ejs2433

Kassi, O. and Wang, S.G.W. (2026). Supplement to "Structural adaptation and rate accelerated estimation in bivariate functional data." https://doi.org/10.3150/26-BEJ1991SUPPA, https://doi.org/10.3150/26-BEJ1991SUPPB

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. *Texts in Statistical Science Series*. Boca Raton, FL: CRC Press. MR3793167

Lepski, O. (2015). Adaptive estimation over anisotropic functional classes via oracle approach. *Ann. Statist.* **43** 1178–1242. MR3346701 https://doi.org/10.1214/14-AOS1306

Lepski, O.V. and Rebelles, G. (2020). Structural adaptation in the density model. *Math. Stat. Learn.* **3** 345–386. MR4362042 https://doi.org/10.4171/msl/24

Maissoro, H., Patilea, V. and Vimond, M. (2025). Adaptive estimation for weakly dependent functional times series. *J. Time Series Anal.* https://doi.org/10.1111/jtsa.70006

Matheron, G. (1973). The intrinsic random functions and their applications. *Adv. Appl. Probab.* **5** 439–468. https://doi.org/10.2307/1425829 MR356209

Niemi, T.J., Guillaume, J.H.A., Kokkonen, T., Hoang, T.M.T. and Seed, A.W. (2015). Role of spatial anisotropy in design storm generation: Experiment and interpretation. *Water Resour. Res.* **51** 1–21. https://doi.org/10.1002/2015WR017521

Oliveira, A.L.G., Lima, J.P., Brasco, T.L. and Amaral, L.R. (2022). The importance of modeling the effects of trend and anisotropy on soil fertility maps. *Comput. Electron. Agric.* **196** 106877. https://doi.org/10.1016/j.compag.2022.106877

PRISM Climate Group, O.S.U. (2025). PRISM climate data. Available at https://prism.oregonstate.edu/. Accessed: 2025-06-19.

Ramsay, T. (2002). Spline smoothing over difficult regions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 307–319. MR1904707 https://doi.org/10.1111/1467-9868.00339

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Rice, J.A. and Wu, C.O. (2001). Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics* **57** 253–259. MR1833314 https://doi.org/10.1111/j.0006-341X.2001.00253.x

Richard, F.J.P. (2016). Tests of isotropy for rough textures of trended images. *Statist. Sinica* **26** 1279–1304. MR3559953

Samarov, A. and Tsybakov, A. (2004). Nonparametric independent component analysis. *Bernoulli* **10** 565–582. MR2076063 https://doi.org/10.3150/bj/1093265630

Sangalli, L.M., Ramsay, J.O. and Ramsay, T.O. (2013). Spatial spline regression models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 681–703. MR3091654 https://doi.org/10.1111/rssb.12009

Shen, J. and Hsing, T. (2020). Hurst function estimation. *Ann. Statist.* **48** 838–862. https://doi.org/10.1214/19-AOS1825

Somvanshi, P. and Rane, M. (2012). Survey of palmprint recognition. *Int. J. Sci. Eng. Res.* **3** 1–5.

Stein, M.L. (2002). Fast and exact simulation of fractional Brownian surfaces. *J. Comput. Graph. Statist.* **11** 587–599. MR1938447 https://doi.org/10.1198/106186002466

Wang, Y., Hu, J. and Yan, H. (2005). A gradient-based weighted averaging method for estimation of fingerprint orientation fields. In *Proceedings of the Digital Image Computing: Techniques and Applications (DICTA)* 191–198.

Wang, S.G.W., Patilea, V. and Klutchnikoff, N. (2025). Adaptive functional principal components analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **87** 603–631. https://doi.org/10.1093/jrsssb/qkae106

Wood, S.N., Bravington, M.V. and Hedley, S.L. (2008). Soap film smoothing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 931–955. MR2530324 https://doi.org/10.1111/j.1467-9868.2008.00665.x

Wood, A.T.A. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0,1]^d$. *J. Comput. Graph. Statist.* **3** 409–432. MR1323050 https://doi.org/10.2307/1390903

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. MR2160561 https://doi.org/10.1198/016214504000001745