BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty One Number Two May 2025 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty One Number Two May 2025 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 Volume 31 Number 2 May 2025 Pages 843–1691 ISI/BS

Volume 31 Number 2   May 2025   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 2025 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **31**(2), 2025, 843–867
https://doi.org/10.3150/21-BEJ1412

Accuracy of Gaussian approximation for high-dimensional posterior distributions

VLADIMIR SPOKOINY[1,a] and MAXIM PANOV[2,b]

[1] *Weierstrass Institute and Humboldt University Berlin, Germany, IITP RAS, and HSE University Moscow, Russia,* [a]*spokoiny@wias-berlin.de*
[2] *Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), Abu Dhabi, UAE,* [b]*maxim.panov@mbzuai.ac.ae*

The prominent Bernstein – von Mises (BvM) result claims that the posterior distribution after centering by the efficient estimator and standardizing by the square root of the total Fisher information is nearly standard normal. In particular, the prior completely washes out from the asymptotic posterior distribution. This fact is fundamental and justifies the Bayes approach from the frequentist viewpoint. In the nonparametric setup the situation changes dramatically and the impact of prior becomes essential even for the contraction of the posterior; see (*Ann. Statist.* **36** (2008) 1435–1463, *Ann. Statist.* **39** (2011) 2557–2584, *Ann. Statist.* **41** (2013) 1999–2028, *Ann. Statist.* **42** (2014) 1941–1969) for different models like Gaussian regression or i.i.d. model in different weak topologies. This paper offers another non-asymptotic approach to studying the behavior of the posterior for a special but rather popular and useful class of statistical models and for Gaussian priors. Our main results describe the accuracy of Gaussian approximation of the posterior. In particular, we show that restricting to the class of all centrally symmetric credible sets around the penalized maximum likelihood estimator (pMLE) allows to get Gaussian approximation up to order n^{-1}. We also derive tight finite sample bounds on posterior contraction in terms of the so-called effective dimension of the parameter space and address the question of frequentist reliability of Bayesian credible sets. The obtained results are specified for nonparametric log-density estimation and generalized regression.

Keywords: Posterior; concentration; contraction Gaussian approximation

References

Bontemps, D. (2011). Bernstein-von Mises theorems for Gaussian regression with increasing number of regressors. *Ann. Statist.* **39** 2557–2584.

Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028.

Castillo, I. and Nickl, R. (2014). On the Bernstein–von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969.

Castillo, I. and Rousseau, J. (2015). A Bernstein–von Mises theorem for smooth functionals in semiparametric models. *Ann. Statist.* **43** 2353–2383.

Cox, D.D. (1993). An analysis of Bayesian inference for nonparametric regression. *Ann. Statist.* **21** 903–923.

Freedman, D. (1999). On the Bernstein-von Mises theorem with infinite-dimensional parameters. *Ann. Statist.* **27** 1119–1140.

Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press.

Kleijn, B.J.K. and van der Vaart, A.W. (2006). Misspecification in infinite-dimensional Bayesian statistics. *Ann. Statist.* **34** 837–877.

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

Knapik, B.T., Szabó, B.T., van der Vaart, A.W. and van Zanten, J.H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probab. Theory Related Fields* **164** 771–813.

1350-7265

Leahu, H. (2011). On the Bernstein-von Mises phenomenon in the Gaussian white noise model. *Electron. J. Stat.* **5** 373–404.

Nickl, R. and Szabó, B. (2016). A sharp adaptive confidence ball for self-similar functions. *Stochastic Process. Appl.* **126** 3913–3934. In Memoriam: Evarist Gine.

Sniekers, S. and van der Vaart, A. (2015). Adaptive Bayesian credible sets in regression with a Gaussian process prior. *Electron. J. Stat.* **9** 2475–2527.

Spokoiny, V. (2017). Penalized maximum likelihood estimation and effective dimension. *AIHP* **53** 389–429.

Panov, M. and Spokoiny, V. (2015). Finite sample Bernstein - von Mises theorem for semiparametric problems. *Bayesian Anal.* **10** 665–710.

Spokoiny, V. (2019). Bayesian inference for nonlinear inverse problems. arXiv:1912.12694.

van der Vaart, A. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge Univ. Press.

van der Vaart, A.W. and van Zanten, J.H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463.

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

Yano, K. and Kato, K. (2020). On frequentist coverage errors of Bayesian credible sets in moderately high dimensions. *Bernoulli* **26** 616–641.

Spokoiny, V. and Panov, M. (2025). Supplement to “Accuracy of Gaussian approximation for high-dimensional posterior distributions.” https://doi.org/10.3150/21-BEJ1412SUPP

Bernoulli **31**(2), 2025, 868–893
https://doi.org/10.3150/23-BEJ1662

A Fourier representation of kernel Stein discrepancy with application to Goodness-of-Fit tests for measures on infinite dimensional Hilbert spaces

GEORGE WYNNE[1,a], MIKOŁAJ J. KASPRZAK[2,b] and ANDREW B. DUNCAN[3,4,c]

[1] *University of Bristol, Bristol, United Kingdom,* [a]*g.wynne@bristol.ac.uk*
[2] *ESSEC Business School, Cergy-Pontoise, France,* [b]*kasprzak@essec.edu*
[3] *Imperial College London, London, United Kingdom,* [c]*a.duncan@imperial.ac.uk*
[4] *The Alan Turing Institute, London, United Kingdom*

Kernel Stein discrepancy (KSD) is a widely used kernel-based measure of discrepancy between probability measures. It is often employed in the scenario where a user has a collection of samples from a candidate probability measure and wishes to compare them against a specified target probability measure. KSD has been employed in a range of settings including goodness-of-fit testing, parametric inference, MCMC output assessment and generative modelling. However, so far the method has been restricted to finite-dimensional data. We provide the first analysis of KSD in the generality of data lying in a separable Hilbert space, for example functional data. The main result is a novel Fourier representation of KSD obtained by combining the theory of measure equations with kernel methods. This allows us to prove that KSD can separate measures and thus is valid to use in practice. Additionally, our results improve the interpretability of KSD by decoupling the effect of the kernel and Stein operator. We demonstrate the efficacy of the proposed methodology by performing goodness-of-fit tests for various Gaussian and non-Gaussian functional models in a number of synthetic data experiments.

Keywords: Functional data analysis; Stein's method; testing

References

Albeverio, S., Bogachev, V. and Röckner, M. (1999). On uniqueness of invariant measures for finite- and infinite-dimensional diffusions. *Comm. Pure Appl. Math.* **52** 325–362. MR1656067 https://doi.org/10.1002/(SICI)1097-0312(199903)52:3<325::AID-CPA2>3.0.CO;2-V

Albeverio, S. and Röckner, M. (1991). Stochastic differential equations in infinite dimensions: Solutions via Dirichlet forms. *Probab. Theory Related Fields* **89** 347–386. MR1113223 https://doi.org/10.1007/BF01198791

Anastasiou, A., Barp, A., Briol, F.-X., Ebner, B., Gaunt, R.E., Ghaderinezhad, F., Gorham, J., Gretton, A., Ley, C., Liu, Q., Mackey, L., Oates, C.J., Reinert, G. and Swan, Y. (2023). Stein's method meets computational statistics: A review of some recent developments. *Statist. Sci.* **38** 120–139. MR4534646 https://doi.org/10.1214/22-sts863

Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404. MR0051437 https://doi.org/10.2307/1990404

Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2016). A new approach to the Stein-Tikhomirov method: With applications to the second Wiener chaos and Dickman convergence. Available at arXiv:1605.06819.

Barbour, A.D. (1988). Stein's method and Poisson process convergence **Special Vol. 25A** 175–184. A celebration of applied probability.

Barbour, A.D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659 https://doi.org/10.1007/BF01197887

Barp, A., Oates, C.J., Porcu, E. and Girolami, M. (2022a). A Riemann-Stein kernel method. *Bernoulli* **28** 2181–2208. MR4474540 https://doi.org/10.3150/21-bej1415

Barp, A., Simon-Gabriel, C.-J., Girolami, M. and Mackey, L. (2022b). Targeted separation and convergence with kernel discrepancies. Available at arXiv:2209.12835.

Beskos, A., Papaspiliopoulos, O., Roberts, G.O. and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 333–382. With discussions and a reply by the authors. MR2278331 https://doi.org/10.1111/j.1467-9868.2006.00552.x

Beskos, A., Roberts, G., Stuart, A. and Voss, J. (2008). MCMC methods for diffusion bridges. *Stoch. Dyn.* **8** 319–350. MR2444507 https://doi.org/10.1142/S0219493708002378

Bierkens, J., Grazzi, S., van der Meulen, F. and Schauer, M. (2021). A piecewise deterministic Monte Carlo method for diffusion bridges. *Stat. Comput.* **31** 37. MR4249043 https://doi.org/10.1007/s11222-021-10008-8

Bogachev, V.I. (1998). *Gaussian Measures*. *Mathematical Surveys and Monographs* **62**. Providence, RI: Amer. Math. Soc. MR1642391 https://doi.org/10.1090/surv/062

Bogachev, V.I. (2010). *Differentiable Measures and the Malliavin Calculus*. *Mathematical Surveys and Monographs* **164**. Providence, RI: Amer. Math. Soc. MR2663405 https://doi.org/10.1090/surv/164

Bogachev, V.I., Krylov, N.V. and Röckner, M. (2009). Elliptic and parabolic equations for measures. *Russian Math. Surveys* **64** 973–1078. MR2640966 https://doi.org/10.1070/RM2009v064n06ABEH004652

Bogachev, V.I. and Röckner, M. (1995). Regularity of invariant measures on finite- and infinite-dimensional spaces and applications. *J. Funct. Anal.* **133** 168–223. MR1351647 https://doi.org/10.1006/jfan.1995.1123

Bongiorno, E.G., Goia, A. and Vieu, P. (2019). Modeling functional data: A test procedure. *Comput. Statist.* **34** 451–468. MR3953667 https://doi.org/10.1007/s00180-018-0816-9

Bourguin, S. and Campese, S. (2020). Approximation of Hilbert-valued Gaussians on Dirichlet structures. *Electron. J. Probab.* **25** 150. MR4193891 https://doi.org/10.1214/20-ejp551

Bugni, F.A., Hall, P., Horowitz, J.L. and Neumann, G.R. (2009). Goodness-of-fit tests for functional data. *Econom. J.* **12** S1–S18. MR2543479 https://doi.org/10.1111/j.1368-423X.2008.00266.x

Carmeli, C., De Vito, E. and Toigo, A. (2006). Vector valued reproducing kernel Hilbert spaces of integrable functions and Mercer theorem. *Anal. Appl. (Singap.)* **4** 377–408. MR2265340 https://doi.org/10.1142/S0219530506000838

Carmeli, C., de Vito, E., Toigo, A. and Umanità, V. (2010). Vector valued reproducing kernel Hilbert spaces and universality. *Anal. Appl. (Singap.)* **8** 19–61. MR2603770 https://doi.org/10.1142/S0219530510001503

Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications (New York)*. Heidelberg: Springer. MR2732624 https://doi.org/10.1007/978-3-642-15007-4

Chevyrev, I. and Oberhauser, H. (2022). Signature moments to characterize laws of stochastic processes. *J. Mach. Learn. Res.* **23** 176. MR4577129 https://doi.org/10.5486/pmd.1976.23.1-2.24

Chwialkowski, K., Strathmann, H. and Gretton, A. (2016). A kernel test of goodness-of-fit. In *Proceedings of the 33rd International Conference on Machine Learning* **48** 2606–2615.

Cuesta-Albertos, J.A., del Barrio, E., Fraiman, R. and Matrán, C. (2007). The random projection method in goodness of fit for functional data. *Comput. Statist. Data Anal.* **51** 4814–4831. MR2364542 https://doi.org/10.1016/j.csda.2006.09.007

Da Prato, G. (2006). *An Introduction to Infinite-Dimensional Analysis*. *Universitext*. Berlin: Springer. MR2244975 https://doi.org/10.1007/3-540-29021-4

Da Prato, G. and Zabczyk, J. (2002). *Second Order Partial Differential Equations in Hilbert Spaces*. *London Mathematical Society Lecture Note Series* **293**. Cambridge: Cambridge Univ. Press. MR1985790 https://doi.org/10.1017/CBO9780511543210

Ditzhaus, M. and Gaigall, D. (2018). A consistent goodness-of-fit test for huge dimensional and functional data. *J. Nonparametr. Stat.* **30** 834–859. MR3871114 https://doi.org/10.1080/10485252.2018.1486402

Ebner, B. and Henze, N. (2020). Rejoinder on: Tests for multivariate normality—a critical review with emphasis on weighted L^2-statistics [MR4182842; MR4182843; MR4183844; MR4183845; MR4183841]. *TEST* **29** 911–913. MR4182846 https://doi.org/10.1007/s11749-020-00744-w

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. *Springer Series in Statistics*. New York: Springer. MR2229687

Fisher, M., Nolan, T., Graham, M., Prangle, D. and Oates, C. (2021). Measure transport with kernel Stein discrepancy. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics* **130** 1054–1062.

Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR3587782 https://doi.org/10.1017/9781139029834

Gorham, J. and Mackey, L. (2015). Measuring sample quality with Stein's method. In *Advances in Neural Information Processing Systems* **28**. Curran Associates.

Gorham, J. and Mackey, L. (2017). Measuring sample quality with kernels. In *Proceedings of the 34th International Conference on Machine Learning* **70** 1292–1301.

Götze, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724–739. MR1106283

Grathwohl, W., Wang, K.-C., Jacobsen, J.-H., Duvenaud, D. and Zemel, R. (2020). Learning the Stein discrepancy for training and evaluating energy-based models without sampling. In *Proceedings of the 37th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* **119** 3732–3747. PMLR.

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. MR2913716

Hairer, M., Stuart, A.M. and Voss, J. (2007). Analysis of SPDEs arising in path sampling. II. The nonlinear case. *Ann. Appl. Probab.* **17** 1657–1706. MR2358638 https://doi.org/10.1214/07-AAP441

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. MR4275152 https://doi.org/10.1111/sjos.12470

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

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR3379106 https://doi.org/10.1002/9781118762547

Hunt, B.R., Sauer, T. and Yorke, J.A. (1992). Prevalence: A translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)* **27** 217–238. MR1161274 https://doi.org/10.1090/S0273-0979-1992-00328-2

Hušková, M. and Janssen, P. (1993). Consistency of the generalized bootstrap for degenerate U-statistics. *Ann. Statist.* **21** 1811–1823. MR1245770 https://doi.org/10.1214/aos/1176349399

Jia, J. and Meng, D. (2021). Stein variational gradient descent on infinite-dimensional space and applications to statistical inverse problems. Available at arXiv:2102.09741.

Jitkrittum, W., Xu, W., Szabo, Z., Fukumizu, K. and Gretton, A. (2017). A linear-time kernel goodness-of-fit test. *Adv. Neural Inf. Process. Syst.* **30** 262–271.

Kadri, H., Duflos, E., Preux, P., Canu, S., Rakotomamonjy, A. and Audiffren, J. (2016). Operator-valued kernels for learning from functional response data. *J. Mach. Learn. Res.* **17** 20. MR3491114

Kukush, A. (2019). *Gaussian Measures in Hilbert Space*. New York: Wiley.

Ley, C., Reinert, G. and Swan, Y. (2017). Stein's method for comparison of univariate distributions. *Probab. Surv.* **14** 1–52. MR3595350 https://doi.org/10.1214/16-PS278

Liu, Q., Lee, J. and Jordan, M. (2016). A kernelized Stein discrepancy for goodness-of-fit tests. In *Proceedings of the 33rd International Conference on Machine Learning* **48** 276–284.

Liu, Q. and Wang, D. (2016). Stein variational gradient descent: A general purpose Bayesian inference algorithm. In *Advances in Neural Information Processing Systems* **29**.

Maniglia, S. and Rhandi, A. (2004). Gaussian Measures on Separable Hilbert Spaces and Applications.

Matsubara, T., Knoblauch, J., Briol, F.-X. and Oates, C.J. (2022). Robust generalised Bayesian inference for intractable likelihoods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 997–1022. MR4460583 https://doi.org/10.1111/rssb.12500

Micchelli, C.A. and Pontil, M. (2005). On learning vector-valued functions. *Neural Comput.* **17** 177–204. MR2175914 https://doi.org/10.1162/0899766052530802

Mijoule, G., Reinert, G. and Swan, Y. (2018). Stein operators, kernels and discrepancies for multivariate continuous distributions. Available at arXiv:1806.03478.

Müller, A. (1997). Integral probability metrics and their generating classes of functions. *Adv. in Appl. Probab.* **29** 429–443. MR1450938 https://doi.org/10.2307/1428011

Nourdin, I. and Peccati, G. (2009). *Normal Approximations with Malliavin Calculus*. Cambridge: Cambridge Univ. Press.

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. MR2724851 https://doi.org/10.1198/jasa.2010.tm09239

Paulsen, V.I. and Raghupathi, M. (2016). *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. *Cambridge Studies in Advanced Mathematics* **152**. Cambridge: Cambridge Univ. Press. MR3526117 https://doi.org/10.1017/CBO9781316219232

Pinski, F.J. and Stuart, A.M. (2010). Transition paths in molecules at finite temperature. *J. Chem. Phys.* **132** 184104.

Pomann, G.-M., Staicu, A.-M. and Ghosh, S. (2016). A two-sample distribution-free test for functional data with application to a diffusion tensor imaging study of multiple sclerosis. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **65** 395–414. MR3470583 https://doi.org/10.1111/rssc.12130

Rajput, B.S. (1972). Gaussian measures on L_p spaces, $1 \le p < \infty$. *J. Multivariate Anal.* **2** 382–403. MR0345157 https://doi.org/10.1016/0047-259X(72)90034-6

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR2514435

Roberts, G.O. and Stramer, O. (2001). On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm. *Biometrika* **88** 603–621. MR1859397 https://doi.org/10.1093/biomet/88.3.603

Rossi, F. and Villa, N. (2006). Support vector machine for functional data classification. *Neurocomputing* **69** 730–742.

Salvi, C., Lemercier, M., Liu, C., Horvath, B., Damoulas, T. and Lyons, T. (2021). Higher order kernel mean embeddings to capture filtrations of stochastic processes. In *Advances in Neural Information Processing Systems* **34** 16635–16647.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR0595165

Shih, H.-H. (2011). On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces. *J. Funct. Anal.* **261** 1236–1283. MR2807099 https://doi.org/10.1016/j.jfa.2011.04.016

South, L.F., Riabiz, M., Teymur, O. and Oates, C.J. (2022). Postprocessing of MCMC. *Annu. Rev. Stat. Appl.* **9** 529–555. MR4394919 https://doi.org/10.1146/annurev-statistics-040220-091727

Sriperumbudur, B.K., Gretton, A., Fukumizu, K., Schölkopf, B. and Lanckriet, G.R.G. (2010). Hilbert space embeddings and metrics on probability measures. *J. Mach. Learn. Res.* **11** 1517–1561. MR2645460

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 (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory* 583–602. Berkeley, CA: Univ. California Press. MR0402873

Stuart, A.M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. MR2652785 https://doi.org/10.1017/S0962492910000061

Tikhomirov, A.N. (1981). On the convergence rate in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25** 790–809.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Wynne, G. and Duncan, A.B. (2022). A kernel two-sample test for functional data. *J. Mach. Learn. Res.* **23** 73. MR4576658

Wynne, G., Kasprzak, M. and Duncan, A.B. (2025). Supplement to "A Fourier representation of kernel Stein discrepancy with application to Goodness-of-Fit tests for measures on infinite dimensional Hilbert spaces." https://doi.org/10.3150/23-BEJ1662SUPPA, https://doi.org/10.3150/23-BEJ1662SUPPB

Xu, W. (2022). Standardisation-function kernel Stein discrepancy: A unifying view on kernel Stein discrepancy tests for goodness-of-fit. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*. *Proceedings of Machine Learning Research* **151** 1575–1597. PMLR.

Yang, J., Liu, Q., Rao, V. and Neville, J. (2018). Goodness-of-fit testing for discrete distributions via Stein discrepancy. In *Proceedings of the 35th International Conference on Machine Learning* **80** 5561–5570.

Bernoulli **31**(2), 2025, 894–921
https://doi.org/10.3150/24-BEJ1752

An adaptive model checking test for the functional linear model

ENZE SHI[1,a], YI LIU[1,b], KE SUN[1,c], LINGZHU LI[2,e] and LINGLONG KONG[1,d]

[1]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada,*
[a]*eshi@ualberta.ca,* [b]*yliu16@ualberta.ca,* [c]*ksun6@ualberta.ca,* [d]*lkong@ualberta.ca*
[2]*School of Mathematics, Statistics and Mechanics, Beijing University of Technology, Beijing, China,*
[e]*lzli1@bjut.edu.cn*

Numerous studies have been devoted to the estimation and inference problems for functional linear models (FLM). However, few works focus on model checking problem that ensures the reliability of results. Limited tests in this area do not have tractable null distributions or asymptotic analysis under alternatives. Also, the functional predictor is usually assumed to be fully observed, which is impractical. To address these problems, we propose an adaptive model checking test for FLM. It combines regular moment-based and conditional moment-based tests, and achieves model adaptivity via the dimension of a residual-based subspace. The advantages of our test are manifold. First, it has a tractable chi-squared null distribution and higher powers under the alternatives than its components. Second, asymptotic properties under different underlying models are developed, including the unvisited local alternatives. Third, the test statistic is constructed upon finite grid points, which incorporates the discrete nature of collected data. We develop the desirable relationship between sample size and number of grid points to maintain the asymptotic properties. Besides, we provide a data-driven approach to estimate the dimension leading to model adaptivity, which is promising in sufficient dimension reduction. We conduct comprehensive numerical experiments to demonstrate the advantages the test inherits from its two simple components.

Keywords: Adaptive-to-model test; functional linear model; reproducing kernel Hilbert space; sufficient dimension reduction

References

Cai, T.T. and Hall, P. (2006). Prediction in functional linear regression. *Ann. Statist.* **34** 2159–2179. MR2291496 https://doi.org/10.1214/009053606000000830

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. MR2906870 https://doi.org/10.1214/11-AOS898

Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *J. Amer. Statist. Assoc.* **107** 1201–1216. MR3010906 https://doi.org/10.1080/01621459.2012.716337

Cardot, H., Ferraty, F., Mas, A. and Sarda, P. (2003). Testing hypotheses in the functional linear model. *Scand. J. Stat.* **30** 241–255. MR1965105 https://doi.org/10.1111/1467-9469.00329

Chen, F., Jiang, Q., Feng, Z. and Zhu, L. (2020). Model checks for functional linear regression models based on projected empirical processes. *Comput. Statist. Data Anal.* **144** 106897, 22. MR4044785 https://doi.org/10.1016/j.csda.2019.106897

Cook, R.D. and Li, B. (2002). Dimension reduction for conditional mean in regression. *Ann. Statist.* **30** 455–474. MR1902895 https://doi.org/10.1214/aos/1021379861

Cuesta-Albertos, J.A., García-Portugués, E., Febrero-Bande, M. and González-Manteiga, W. (2019). Goodness-of-fit tests for the functional linear model based on randomly projected empirical processes. *Ann. Statist.* **47** 439–467. MR3909938 https://doi.org/10.1214/18-AOS1693

Dette, H., Neumeyer, N. and Van Keilegom, I. (2007). A new test for the parametric form of the variance function in non-parametric regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 903–917. MR2368576 https://doi.org/10.1111/j.1467-9868.2007.00616.x

1350-7265

DeVore, R.A. and Lorentz, G.G. (1993). *Constructive Approximation*. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-662-02888-9

Escanciano, J.C. (2006). A consistent diagnostic test for regression models using projections. *Econometric Theory* **22** 1030–1051. MR2328527 https://doi.org/10.1017/S0266466606060506

Ferraty, F., Goia, A., Salinelli, E. and Vieu, P. (2013). Functional projection pursuit regression. *TEST* **22** 293–320. MR3062259 https://doi.org/10.1007/s11749-012-0306-2

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. *Springer Series in Statistics*. New York: Springer. MR2229687

Ferré, L. and Yao, A.-F. (2005). Smoothed functional inverse regression. *Statist. Sinica* **15** 665–683. MR2233905

Ferré, L. and Yao, A.F. (2003). Functional sliced inverse regression analysis. *Statistics* **37** 475–488. MR2022235 https://doi.org/10.1080/0233188031000112845

García-Portugués, E., González-Manteiga, W. and Febrero-Bande, M. (2014). A goodness-of-fit test for the functional linear model with scalar response. *J. Comput. Graph. Statist.* **23** 761–778. MR3224655 https://doi.org/10.1080/10618600.2013.812519

Goldsmith, J. and Kitago, T. (2016). Assessing systematic effects of stroke on motor control by using hierarchical function-on-scalar regression. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **65** 215–236. MR3456686 https://doi.org/10.1111/rssc.12115

Guo, X., Wang, T. and Zhu, L. (2016). Model checking for parametric single-index models: A dimension reduction model-adaptive approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 1013–1035. MR3557187 https://doi.org/10.1111/rssb.12147

Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. MR2332269 https://doi.org/10.1214/009053606000000957

Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. MR2278365 https://doi.org/10.1214/009053606000000272

Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21** 1926–1947. MR1245774 https://doi.org/10.1214/aos/1176349403

Lei, J. (2014). Adaptive global testing for functional linear models. *J. Amer. Statist. Assoc.* **109** 624–634. MR3223738 https://doi.org/10.1080/01621459.2013.856794

Li, B. (2018). *Sufficient Dimension Reduction: Methods and Applications*. London: Chapman & Hall. https://doi.org/10.1201/9781315119427.

Li, L., Zhu, X. and Zhu, L.-X. (2023). Adaptive-to-model hybrid of tests for regressions. *J. Amer. Statist. Assoc.* **118** 514–523. MR4571138 https://doi.org/10.1080/01621459.2021.1941052

Lian, H. (2015). Functional sufficient dimension reduction: Convergence rates and multiple functional case. *J. Statist. Plann. Inference* **167** 58–68. MR3383236 https://doi.org/10.1016/j.jspi.2015.05.008

Müller, H.-G. and Stadtmüller, U. (2005). Generalized functional linear models. *Ann. Statist.* **33** 774–805. MR2163159 https://doi.org/10.1214/009053604000001156

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Reiss, P.T., Huang, L. and Mennes, M. (2010). Fast function-on-scalar regression with penalized basis expansions. *Int. J. Biostat.* **6** Art. 28, 30. MR2683940 https://doi.org/10.2202/1557-4679.1246

Ritter, K., Wasilkowski, G.W. and Woźniakowski, H. (1995). Multivariate integration and approximation for random fields satisfying Sacks-Ylvisaker conditions. *Ann. Appl. Probab.* **5** 518–540. MR1336881

Shang, Z. and Cheng, G. (2015). Nonparametric inference in generalized functional linear models. *Ann. Statist.* **43** 1742–1773. MR3357877 https://doi.org/10.1214/15-AOS1322

Shi, E., Liu, Y., Sun, K., Li, L. and Kong, L. (2025). Supplement to "An adaptive model checking test for the functional linear model." https://doi.org/10.3150/24-BEJ1752SUPP

Song, J. (2019). On sufficient dimension reduction for functional data: Inverse moment-based methods. *Wiley Interdiscip. Rev.: Comput. Stat.* **11** e1459, 13. MR3999524 https://doi.org/10.1002/wics.1459

Song, J. and Li, B. (2021). Nonlinear and additive principal component analysis for functional data. *J. Multivariate Anal.* **181** Paper No. 104675, 21. MR4157686 https://doi.org/10.1016/j.jmva.2020.104675

Stewart, G.W. (1990). Stochastic perturbation theory. *SIAM Rev.* **32** 579–610. MR1084571 https://doi.org/10.1137/1032121

Stute, W., González Manteiga, W. and Presedo Quindimil, M. (1998). Bootstrap approximations in model checks for regression. *J. Amer. Statist. Assoc.* **93** 141–149. MR1614600 https://doi.org/10.2307/2669611

Sun, X., Du, P., Wang, X. and Ma, P. (2018). Optimal penalized function-on-function regression under a reproducing kernel Hilbert space framework. *J. Amer. Statist. Assoc.* **113** 1601–1611. MR3902232 https://doi.org/10.1080/01621459.2017.1356320

Tan, F., Zhu, X. and Zhu, L. (2018). A projection-based adaptive-to-model test for regressions. *Statist. Sinica* **28** 157–188. MR3752256

Wahba, G. (1990). *Spline Models for Observational Data*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. Philadelphia, PA: SIAM. MR1045442 https://doi.org/10.1137/1.9781611970128

Xue, K. and Yao, F. (2021). Hypothesis testing in large-scale functional linear regression. *Statist. Sinica* **31** 1101–1123. MR4286208 https://doi.org/10.5705/ss.20

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. MR2253106 https://doi.org/10.1214/009053605000000660

Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. MR2766857 https://doi.org/10.1214/09-AOS772

Zhang, Z., Wang, X., Kong, L. and Zhu, H. (2022). High-dimensional spatial quantile function-on-scalar regression. *J. Amer. Statist. Assoc.* **117** 1563–1578. MR4480732 https://doi.org/10.1080/01621459.2020.1870984

Zheng, J.X. (1996). A consistent test of functional form via nonparametric estimation techniques. *J. Econometrics* **75** 263–289. MR1413644 https://doi.org/10.1016/0304-4076(95)01760-7

Zhu, H., Li, R. and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *Ann. Statist.* **40** 2634–2666. MR3097615 https://doi.org/10.1214/12-AOS1045

Zhu, L., Wang, T., Zhu, L. and Ferré, L. (2010). Sufficient dimension reduction through discretization-expectation estimation. *Biometrika* **97** 295–304. MR2650739 https://doi.org/10.1093/biomet/asq018

Zhu, X., Guo, X., Wang, T. and Zhu, L. (2020). Dimensionality determination: A thresholding double ridge ratio approach. *Comput. Statist. Data Anal.* **146** 106910, 18. MR4065384 https://doi.org/10.1016/j.csda.2020.106910

Bernoulli **31**(2), 2025, 922–947
https://doi.org/10.3150/24-BEJ1753

Functional diffusion driven stochastic volatility model

PIOTR KOKOSZKA[1,a], NEDA MOHAMMADI[1,b], HAONAN WANG[1,c] and SHIXUAN WANG[2,d]

[1]*Department of Statistics, Colorado State University, Fort Collins, USA,* [a]*piotr.kokoszka@colostate.edu,* [b]*neda.mohammadi@colostate.edu,* [c]*haonan.wang@colostate.edu*
[2]*Department of Economics, University of Reading, Reading, UK,* [d]*shixuan.wang@reading.ac.uk*

We propose a stochastic volatility model for time series of curves. It is motivated by dynamics of intraday price curves that exhibit both between days dependence and intraday price evolution. The curves are suitably normalized to stationary in a function space and are functional analogs of point-to-point daily returns. The between curves dependence is modeled by a latent autoregression. The within curves behavior is modeled by a diffusion process. We establish the properties of the model and propose several approaches to its estimation. These approaches are justified by asymptotic arguments that involve an interplay between the latent autoregression and the intraday diffusions. The asymptotic framework combines the increasing number of daily curves and the refinement of the discrete grid on which each daily curve is observed. Consistency rates for the estimators of the intraday volatility curves are derived as well as the asymptotic normality of the estimators of the latent autoregression. The estimation approaches are further explored and compared by an application to intraday price curves of over seven thousand U.S. stocks and an informative simulation study.

Keywords: Functional time series; intraday price curves; Itô diffusion process; stochastic volatility

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*, Illustrated ed. Princeton, NJ: Princeton Univ. Press. https://doi.org/10.1515/9781400850327

Asai, M., McAleer, M. and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Rev.* **25** 145–175. MR2256285 https://doi.org/10.1080/07474930600713564

Aue, A., Horváth, L. and Pellatt, D.F. (2017). Functional generalized autoregressive conditional heteroskedasticity. *J. Time Series Anal.* **38** 3–21. MR3601312 https://doi.org/10.1111/jtsa.12192

Bachelier, L. (1900). Théorie de la spéculation. *Ann. Sci. Éc. Norm. Supér. (3)* **17** 21–86. MR1508978

Barndorff-Nielsen, O.E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 253–280. MR1904704 https://doi.org/10.1111/1467-9868.00336

Bauwens, L., Laurent, S. and Rombouts, J.V.K. (2006). Multivariate GARCH models: A survey. *J. Appl. Econometrics* **21** 79–109. MR2225523 https://doi.org/10.1002/jae.842

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Rev. Econ. Stat.* **72** 498–505. https://doi.org/10.2307/2109358

Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988). A capital asset pricing model with time-varying covariances. *J. Polit. Econ.* **96** 116–131.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. *Lecture Notes in Statistics* **149**. New York: Springer. MR1783138 https://doi.org/10.1007/978-1-4612-1154-9

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4

Cerovecki, C., Francq, C., Hörmann, S. and Zakoïan, J.-M. (2019). Functional GARCH models: The quasi-likelihood approach and its applications. *J. Econometrics* **209** 353–375. MR3944755 https://doi.org/10.1016/j.jeconom.2019.01.006

Chong, C., Hoffmann, M., Liu, Y., Rosenbaum, M. and Szymanski, G. (2022). Statistical inference for rough volatility: minimax theory. Available at arXiv:2210.01214 [math.ST].

Christensen, K., Hounyo, U. and Podolskij, M. (2018). Is the diurnal pattern sufficient to explain intraday variation in volatility? A nonparametric assessment. *J. Econometrics* **205** 336–362. MR3813521 https://doi.org/10.1016/j.jeconom.2018.03.016

Danielsson, J. (1998). Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models. *J. Empir. Finance* **5** 155–173. https://doi.org/10.1016/S0927-5398(97)00016-9

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* **20** 339–350. MR1939905 https://doi.org/10.1198/073500102288618487

Engle, R.F. and Kroner, K.F. (1995). Multivariate simultaneous generalized arch. *Econometric Theory* **11** 122–150. MR1325104 https://doi.org/10.1017/S0266466600009063

Galbraith, J.W., Zinde-Walsh, V. and Zhu, J. (2015). GARCH model estimation using estimated quadratic variation. *Econometric Rev.* **34** 1171–1191. MR3292154 https://doi.org/10.1080/07474938.2014.956629

Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. MR2278365 https://doi.org/10.1214/009053606000000272

Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *Rev. Econ. Stud.* **61** 247–264. https://doi.org/10.2307/2297980

Harvey, A. and Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *J. Bus. Econom. Statist.* **14** 429–434. https://doi.org/10.2307/1392251

Hörmann, S., Horváth, L. and Reeder, R. (2013). A functional version of the ARCH model. *Econometric Theory* **29** 267–288. MR3042756 https://doi.org/10.1017/S0266466612000345

Horváth, L., Kokoszka, P. and Rice, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82. MR3153649 https://doi.org/10.1016/j.jeconom.2013.11.002

Jang, P.A., Jauch, M. and Matteson, D.S. (2022). Functional stochastic volatility in financial option surfaces. *Data Sci. J.* **1** 6–19. https://doi.org/10.1080/26941899.2022.2152764

Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

Kokoszka, P. and Reimherr, M. (2013). Determining the order of the functional autoregressive model. *J. Time Series Anal.* **34** 116–129. MR3008019 https://doi.org/10.1111/j.1467-9892.2012.00816.x

Kokoszka, P., Mohammadi, N., Wang, H. and Wang, S. (2024). Supplement to "Functional diffusion driven stochastic volatility model." https://doi.org/10.3150/24-BEJ1753SUPP

Kuenzer, T. (2024). Estimation of functional ARMA models. *Bernoulli* **30** 117–142. MR4665572 https://doi.org/10.3150/23-bej1591

Lavancier, F. and Rochet, P. (2016). A general procedure to combine estimators. *Comput. Statist. Data Anal.* **94** 175–192. MR3412818 https://doi.org/10.1016/j.csda.2015.08.001

Le Gall, J.-F. (2016). *Brownian Motion, Martingales, and Stochastic Calculus*, French ed. *Graduate Texts in Mathematics* **274**. Cham: Springer. MR3497465 https://doi.org/10.1007/978-3-319-31089-3

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. MR2766854 https://doi.org/10.1214/10-AOS813

Mörters, P. and Peres, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge: Cambridge Univ. Press. With an appendix by Oded Schramm and Wendelin Werner. MR2604525 https://doi.org/10.1017/CBO9780511750489

Müller, H.-G., Sen, R. and Stadtmüller, U. (2011). Functional data analysis for volatility. *J. Econometrics* **165** 233–245. MR2846647 https://doi.org/10.1016/j.jeconom.2011.08.002

Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. *Universitext*. Berlin: Springer. MR2001996 https://doi.org/10.1007/978-3-642-14394-6

Paparoditis, E. and Shang, H.L. (2023). Bootstrap prediction bands for functional time series. *J. Amer. Statist. Assoc.* **118** 972–986. MR4595470 https://doi.org/10.1080/01621459.2021.1963262

Rosenbaum, M. (2008). Estimation of the volatility persistence in a discretely observed diffusion model. *Stochastic Process. Appl.* **118** 1434–1462. MR2427046 https://doi.org/10.1016/j.spa.2007.09.004

Shaman, P. and Stine, R.A. (1988). The bias of autoregressive coefficient estimators. *J. Amer. Statist. Assoc.* **83** 842–848. MR0963814

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In *Time Series Models* CRC Press/CRC.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models. In *Handbook of Financial Time Series* (T. Mikosch, J.-P. Kreiß, R.A. Davis and T.G. Andersen, eds.) 201–229. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-71297-8_9

Taylor, S.J. (1982). Financial returns modelled by the product of two stochastic processes: A study of daily sugar prices, 1961–79. In *Time Series Analysis: Theory and Practice* **1**.

Taylor, S.J. (1994). Modeling stochastic volatility: A review and comparative study. *Math. Finance* **4** 183–204. https://doi.org/10.1111/j.1467-9965.1994.tb00057.x

Taylor, S.J. (2008). *Modelling Financial Time Series*. Singapore: World Scientific. https://doi.org/10.1142/6578

Visser, M. (2011). GARCH parameter estimation using high-frequency data. *J. Financ. Econ.* **9** 162–197. https://doi.org/10.1093/jjfinec/nbq017

Vrontos, I.D., Dellaportas, P. and Politis, D.N. (2003). A full-factor multivariate GARCH model. *Econom. J.* **6** 312–334. MR2028238 https://doi.org/10.1111/1368-423X.t01-1-00111

Wilms, I., Basu, S., Bien, J. and Matteson, D.S. (2023). Sparse identification and estimation of large-scale vector autoregressive moving averages. *J. Amer. Statist. Assoc.* **118** 571–582. MR4571142 https://doi.org/10.1080/01621459.2021.1942013

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

Yu, J. and Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Rev.* **25** 361–384. MR2256291 https://doi.org/10.1080/07474930600713465

Bernoulli **31**(2), 2025, 948–972
https://doi.org/10.3150/24-BEJ1754

Blessing of dependence: Identifiability and geometry of discrete models with multiple binary latent variables

YUQI GU[a]

Department of Statistics, Columbia University, New York, NY, USA, [a]*yuqi.gu@columbia.edu*

Identifiability of discrete statistical models with latent variables is known to be challenging to study, yet crucial to a model's interpretability and reliability. This work presents a general algebraic technique to investigate identifiability of discrete models with latent and graphical components. Specifically, motivated by diagnostic tests collecting multivariate categorical data, we focus on discrete models with multiple binary latent variables. We consider the BLESS model, in which the latent variables can have arbitrary dependencies among themselves while the latent-to-observed measurement graph takes a "star-forest" shape. We establish necessary and sufficient graphical criteria for identifiability, and reveal an interesting and perhaps surprising geometry of blessing-of-dependence: under the minimal conditions for generic identifiability, the parameters are identifiable if and only if the latent variables are not statistically independent. Thanks to this theory, we can perform formal hypothesis tests of identifiability in the boundary case by testing marginal independence of the observed variables. In addition to the BLESS model, we also use the technique to show identifiability and the blessing-of-dependence geometry for a more flexible model, which has a general measurement graph beyond a start forest. Our results give new understanding of statistical properties of graphical models with latent variables. They also entail useful implications for designing diagnostic tests or surveys that measure binary latent traits.

Keywords: Algebraic statistics; contingency table; diagnostic test; generic identifiability; graphical model; hypothesis testing; latent class model; multivariate categorical data

References

Allman, E.S., Matias, C. and Rhodes, J.A. (2009). Identifiability of parameters in latent structure models with many observed variables. *Ann. Statist.* **37** 3099–3132. MR2549554 https://doi.org/10.1214/09-AOS689

Allman, E.S., Matias, C. and Rhodes, J.A. (2011). Parameter identifiability in a class of random graph mixture models. *J. Statist. Plann. Inference* **141** 1719–1736. MR2763202 https://doi.org/10.1016/j.jspi.2010.11.022

Allman, E.S. and Rhodes, J.A. (2008). The identifiability of covarion models in phylogenetics. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **6** 76–88. https://doi.org/10.1109/TCBB.2008.52

Allman, E.S., Rhodes, J.A., Stanghellini, E. and Valtorta, M. (2015). Parameter identifiability of discrete Bayesian networks with hidden variables. *J. Causal Inference* **3** 189–205. MR4289435 https://doi.org/10.1515/jci-2014-0021

Anderson, T.W. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. **5** 111–150. Berkeley-Los Angeles, Calif.: Univ. California Press. MR0084943

Bing, X., Bunea, F. and Wegkamp, M. (2023). Detecting approximate replicate components of a high-dimensional random vector with latent structure. *Bernoulli* **29** 1368–1391. MR4550227 https://doi.org/10.3150/22-bej1502

Bing, X., Bunea, F., Ning, Y. and Wegkamp, M. (2020). Adaptive estimation in structured factor models with applications to overlapping clustering. *Ann. Statist.* **48** 2055–2081. MR4134786 https://doi.org/10.1214/19-AOS1877

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR0996025 https://doi.org/10.1002/9781118619179

Carreira-Perpinán, M.A. and Renals, S. (2000). Practical identifiability of finite mixtures of multivariate Bernoulli distributions. *Neural Comput.* **12** 141–152. https://doi.org/10.1162/089976600300015925
Chen, Y., Culpepper, S. and Liang, F. (2020). A sparse latent class model for cognitive diagnosis. *Psychometrika* **85** 121–153. MR4092105 https://doi.org/10.1007/s11336-019-09693-2
Chen, Y., Liu, J., Xu, G. and Ying, Z. (2015). Statistical analysis of Q-matrix based diagnostic classification models. *J. Amer. Statist. Assoc.* **110** 850–866. MR3367269 https://doi.org/10.1080/01621459.2014.934827
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. MR2813153
Culpepper, S.A. (2015). Bayesian estimation of the DINA model with Gibbs sampling. *J. Educ. Behav. Stat.* **40** 454–476. https://doi.org/10.3102/1076998615595403
Culpepper, S.A. (2019). An exploratory diagnostic model for ordinal responses with binary attributes: Identifiability and estimation. *Psychometrika* **84** 921–940. MR4023486 https://doi.org/10.1007/s11336-019-09683-4
Drton, M. (2009). Likelihood ratio tests and singularities. *Ann. Statist.* **37** 979–1012. MR2502658 https://doi.org/10.1214/07-AOS571
Erosheva, E.A., Fienberg, S.E. and Joutard, C. (2007). Describing disability through individual-level mixture models for multivariate binary data. *Ann. Appl. Stat.* **1** 502–537. MR2415745 https://doi.org/10.1214/07-AOAS126
Fang, G., Liu, J. and Ying, Z. (2019). On the identifiability of diagnostic classification models. *Psychometrika* **84** 19–40. MR3910906 https://doi.org/10.1007/s11336-018-09658-x
Fang, G., Guo, J., Xu, X., Ying, Z. and Zhang, S. (2021). Identifiability of bifactor models. *Statist. Sinica* **31** 2309–2330. MR4338090 https://doi.org/10.5705/ss.20
Fienberg, S.E., Hersh, P., Rinaldo, A. and Zhou, Y. (2010). Maximum likelihood estimation in latent class models for contingency table data. In *Algebraic and Geometric Methods in Statistics* 27–62. Cambridge: Cambridge Univ. Press. MR2642657
Gassiat, E., Cleynen, A. and Robin, S. (2016). Inference in finite state space non parametric hidden Markov models and applications. *Stat. Comput.* **26** 61–71. MR3439359 https://doi.org/10.1007/s11222-014-9523-8
Gassiat, E. and Rousseau, J. (2016). Nonparametric finite translation hidden Markov models and extensions. *Bernoulli* **22** 193–212. MR3449780 https://doi.org/10.3150/14-BEJ631
George, A.C. and Robitzsch, A. (2015). Cognitive diagnosis models in R: A didactic. *Quant. Meth. Psych.* **11** 189–205. https://doi.org/10.20982/tqmp.11.3.p189
Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR3617773
Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61** 215–231. MR0370936 https://doi.org/10.1093/biomet/61.2.215
Gu, Y. (2024). Supplement to "Blessing of dependence: identifiability and geometry of discrete models with multiple binary latent variables." https://doi.org/10.3150/24-BEJ1754SUPP
Gu, Y. and Dunson, D.B. (2023). Bayesian pyramids: Identifiable multilayer discrete latent structure models for discrete data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 399–426. MR4726971 https://doi.org/10.1093/jrsssb/qkad010
Gu, Y. and Xu, G. (2019). The sufficient and necessary condition for the identifiability and estimability of the DINA model. *Psychometrika* **84** 468–483. MR3947373 https://doi.org/10.1007/s11336-018-9619-8
Gu, Y. and Xu, G. (2021). Sufficient and necessary conditions for the identifiability of the Q-matrix. *Statist. Sinica* **31** 449–472. MR4270392 https://doi.org/10.5705/ss.20
Gu, Y. and Xu, G. (2023). A joint MLE approach to large-scale structured latent attribute analysis. *J. Amer. Statist. Assoc.* **118** 746–760. MR4571155 https://doi.org/10.1080/01621459.2021.1955689
Gyllenberg, M., Koski, T., Reilink, E. and Verlaan, M. (1994). Nonuniqueness in probabilistic numerical identification of bacteria. *J. Appl. Probab.* **31** 542–548. MR1274807 https://doi.org/10.2307/3215044
Halpern, Y. and Sontag, D. (2013). Unsupervised learning of noisy-or Bayesian networks. In *Uncertainty in Artificial Intelligence* 272–281. Citeseer.
Hinton, G.E., Osindero, S. and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Comput.* **18** 1527–1554. MR2224485 https://doi.org/10.1162/neco.2006.18.7.1527
Junker, B.W. and Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Appl. Psychol. Meas.* **25** 258–272. MR1842982 https://doi.org/10.1177/01466210122032064

Kolda, T.G. and Bader, B.W. (2009). Tensor decompositions and applications. *SIAM Rev.* **51** 455–500. MR2535056 https://doi.org/10.1137/07070111X

Kruskal, J.B. (1977). Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.* **18** 95–138. MR0444690 https://doi.org/10.1016/0024-3795(77)90069-6

Lauritzen, S.L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. Oxford University Press, New York: The Clarendon Press. Oxford Science Publications. MR1419991

Lazarsfeld, P.F. and Henry, N.W. (1968). *Latent Structure Analysis*. Boston, MA: Houghton Mifflin.

Mealli, F., Pacini, B. and Stanghellini, E. (2016). Identification of principal causal effects using additional outcomes in concentration graphs. *J. Educ. Behav. Stat.* **41** 463–480. https://doi.org/10.3102/1076998616646199

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. MR3072945 https://doi.org/10.1613/jair.3879

Rupp, A.A. and Templin, J.L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement* **6** 219–262. https://doi.org/10.1080/15366360802490866

Settimi, R. and Smith, J.Q. (2000). Geometry, moments and conditional independence trees with hidden variables. *Ann. Statist.* **28** 1179–1205. MR1811324 https://doi.org/10.1214/aos/1015956712

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. MR3551782 https://doi.org/10.1093/biomet/asw032

Shwe, M.A., Middleton, B., Heckerman, D.E., Henrion, M., Horvitz, E.J., Lehmann, H.P. and Cooper, G.F. (1991). Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. *Methods Inf. Med.* **30** 241–255.

Silva, R., Scheines, R., Glymour, C. and Spirtes, P. (2006). Learning the structure of linear latent variable models. *J. Mach. Learn. Res.* **7** 191–246. MR2274367

Stanghellini, E. and Vantaggi, B. (2013). Identification of discrete concentration graph models with one hidden binary variable. *Bernoulli* **19** 1920–1937. MR3129039 https://doi.org/10.3150/12-BEJ435

Tatsuoka, K.K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *J. Educ. Meas.* **20** 345–354. https://doi.org/10.1111/j.1745-3984.1983.tb00212.x

von Davier, M. and Lee, Y.-S. (2019). *Handbook of Diagnostic Classification Models*. Cham: Springer. https://doi.org/10.1007/978-3-030-05584-4

Wu, Z., Deloria-Knoll, M. and Zeger, S.L. (2017). Nested partially latent class models for dependent binary data; estimating disease etiology. *Biostatistics* **18** 200–213. MR3825117 https://doi.org/10.1093/biostatistics/kxw037

Xu, G. (2017). Identifiability of restricted latent class models with binary responses. *Ann. Statist.* **45** 675–707. MR3650397 https://doi.org/10.1214/16-AOS1464

Zwiernik, P. (2016). *Semialgebraic Statistics and Latent Tree Models*. *Monographs on Statistics and Applied Probability* **146**. Boca Raton, FL: CRC Press/CRC. MR3379921

Zwiernik, P. and Smith, J.Q. (2012). Tree cumulants and the geometry of binary tree models. *Bernoulli* **18** 290–321. MR2888708 https://doi.org/10.3150/10-BEJ338

Bernoulli **31**(2), 2025, 973–1006
https://doi.org/10.3150/24-BEJ1755

Functional linear and single-index models: A unified approach via Gaussian Stein identity

KRISHNAKUMAR BALASUBRAMANIAN[1,a], HANS-GEORG MÜLLER[1,b] and BHARATH K. SRIPERUMBUDUR[2,c]

[1]*Department of Statistics, University of California, Davis, Davis, USA,* [a]*kbala@ucdavis.edu,* [b]*hgmueller@ucdavis.edu*
[2]*Department of Statistics, Pennsylvania State University, University Park, USA,* [c]*bks18@psu.edu*

Functional linear and single-index models are core regression methods in functional data analysis and are widely used for performing regression in a wide range of applications when the covariates are random functions coupled with scalar responses. In the existing literature, however, the construction of associated estimators and the study of their theoretical properties is invariably carried out on a case-by-case basis for specific models under consideration. In this work, assuming the predictors are Gaussian processes, we provide a unified methodological and theoretical framework for estimating the index in functional linear, and its direction in single-index models. In the latter case, the proposed approach does not require the specification of the link function. In terms of methodology, we show that the reproducing kernel Hilbert space (RKHS) based functional linear least-squares estimator, when viewed through the lens of an *infinite-dimensional Gaussian Stein's identity*, also provides an estimator of the index of the single-index model. Theoretically, we characterize the convergence rates of the proposed estimators for both linear and single-index models. Our analysis has several key advantages: (i) it does not require restrictive commutativity assumptions for the covariance operator of the random covariates and the integral operator associated with the reproducing kernel; and (ii) the true index parameter can lie outside of the chosen RKHS, thereby allowing for index misspecification as well as for quantifying the degree of such index misspecification. Several existing results emerge as special cases of our analysis.

Keywords: Covariance operator; functional single-index models; functional regression; Gaussian Stein's identity; integral operator; operator commutativity; reproducing kernel Hilbert space

References

Babichev, D. and Bach, F. (2018). Slice inverse regression with score functions. *Electron. J. Stat.* **12** 1507–1543. MR3804844 https://doi.org/10.1214/18-EJS1428

Balasubramanian, K., Müller, H.-G. and Sriperumbudur, B.K. (2024). Supplement to “Functional linear and single-index models: A unified approach via Gaussian Stein identity.” https://doi.org/10.3150/24-BEJ1755SUPP

Blanchard, G. and Mücke, N. (2018). Optimal rates for regularization of statistical inverse learning problems. *Found. Comput. Math.* **18** 971–1013. MR3833647 https://doi.org/10.1007/s10208-017-9359-7

Brillinger, D.R. (1983). A generalized linear model with “Gaussian” regressor variables. In *A Festschrift for Erich L. Lehmann*. *Wadsworth Statist./Probab. Ser.* 97–114. Belmont, CA: Wadsworth. MR0689741

Brumback, B.A. and Rice, J.A. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves. *J. Amer. Statist. Assoc.* **93** 961–994. With comments and a rejoinder by the authors. MR1649194 https://doi.org/10.2307/2669837

Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *J. Amer. Statist. Assoc.* **107** 1201–1216. MR3010906 https://doi.org/10.1080/01621459.2012.716337

Caponnetto, A. and De Vito, E. (2007). Optimal rates for the regularized least-squares algorithm. *Found. Comput. Math.* **7** 331–368. MR2335249 https://doi.org/10.1007/s10208-006-0196-8

Cardot, H., Ferraty, F. and Sarda, P. (1999). Functional linear model. *Statist. Probab. Lett.* **45** 11–22. MR1718346 https://doi.org/10.1016/S0167-7152(99)00036-X

1350-7265

Cardot, H., Ferraty, F. and Sarda, P. (2003). Spline estimators for the functional linear model. *Statist. Sinica* **13** 571–591. MR1997162

Chen, D., Hall, P. and Müller, H.-G. (2011). Single and multiple index functional regression models with nonparametric link. *Ann. Statist.* **39** 1720–1747. MR2850218 https://doi.org/10.1214/11-AOS882

Chen, Y. and Müller, H.-G. (2022). Uniform convergence of local Fréchet regression with applications to locating extrema and time warping for metric space valued trajectories. *Ann. Statist.* **50** 1573–1592. MR4441132 https://doi.org/10.1214/21-aos2163

Cordes, H.O. (1987). *Spectral Theory of Linear Differential Operators and Comparison Algebras*. *London Mathematical Society Lecture Note Series* **76**. Cambridge: Cambridge Univ. Press. MR0890743 https://doi.org/10.1017/CBO9780511662836

Cucker, F. and Smale, S. (2002). On the mathematical foundations of learning. *Bull. Amer. Math. Soc. (N.S.)* **39** 1–49. MR1864085 https://doi.org/10.1090/S0273-0979-01-00923-5

Cuevas, A., Febrero, M. and Fraiman, R. (2002). Linear functional regression: The case of fixed design and functional response. *Canad. J. Statist.* **30** 285–300. MR1926066 https://doi.org/10.2307/3315952

De Vito, E., Rosasco, L., Caponnetto, A., De Giovannini, U. and Odone, F. (2005). Learning from examples as an inverse problem. *J. Mach. Learn. Res.* **6** 883–904. MR2249842

Dicker, L.H., Foster, D.P. and Hsu, D. (2017). Kernel ridge vs. principal component regression: Minimax bounds and the qualification of regularization operators. *Electron. J. Stat.* **11** 1022–1047. MR3629418 https://doi.org/10.1214/17-EJS1258

Engle, R.F., Granger, C.W., Rice, J. and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *J. Amer. Statist. Assoc.* **81** 310–320.

Goldstein, L., Minsker, S. and Wei, X. (2018). Structured signal recovery from non-linear and heavy-tailed measurements. *IEEE Trans. Inf. Theory* **64** 5513–5530. MR3832320 https://doi.org/10.1109/TIT.2018.2842216

Goldstein, L. and Wei, X. (2019). Non-Gaussian observations in nonlinear compressed sensing via Stein discrepancies. *Inf. Inference* **8** 125–159. MR3922405 https://doi.org/10.1093/imaiai/iay006

Grenander, U. (1950). Stochastic processes and statistical inference. *Ark. Mat.* **1** 195–277. MR0039202 https://doi.org/10.1007/BF02590638

Grenander, U. (2008). *Probabilities on Algebraic Structures*. Courier Corporation.

Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. MR2332269 https://doi.org/10.1214/009053606000000957

Hall, P. and Van Keilegom, I. (2007). Two-sample tests in functional data analysis starting from discrete data. *Statist. Sinica* **17** 1511–1531. MR2413533

He, G., Müller, H.-G. and Wang, J.-L. (2000). Extending correlation and regression from multivariate to functional data. In *Asymptotics in Statistics and Probability* (M.L. Puri, ed.) 301–315. VSP International Science Publishers.

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR3379106 https://doi.org/10.1002/9781118762547

Hsing, T. and Ren, H. (2009). An RKHS formulation of the inverse regression dimension-reduction problem. *Ann. Statist.* **37** 726–755. MR2502649 https://doi.org/10.1214/07-AOS589

Hsu, D., Kakade, S.M. and Zhang, T. (2014). Random design analysis of ridge regression. *Found. Comput. Math.* **14** 569–600. MR3201956 https://doi.org/10.1007/s10208-014-9192-1

James, G.M. (2002). Generalized linear models with functional predictors. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 411–432. MR1924298 https://doi.org/10.1111/1467-9868.00342

Jiang, C.-R. and Wang, J.-L. (2011). Functional single index models for longitudinal data. *Ann. Statist.* **39** 362–388. MR2797850 https://doi.org/10.1214/10-AOS845

Jiang, C.-R., Yu, W. and Wang, J.-L. (2014). Inverse regression for longitudinal data. *Ann. Statist.* **42** 563–591. MR3210979 https://doi.org/10.1214/13-AOS1193

Kimeldorf, G. and Wahba, G. (1971). Some results on Tchebycheffian spline functions. *J. Math. Anal. Appl.* **33** 82–95. MR0290013 https://doi.org/10.1016/0022-247X(71)90184-3

Kuo, H.-H. and Lee, Y.-J. (2011). Integration by parts formula and the Stein lemma on abstract Wiener space. *Commun. Stoch. Anal.* **5** 405–418. MR2814486 https://doi.org/10.31390/cosa.5.2.10

Li, K.-C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86** 316–342. With discussion and a rejoinder by the author. MR1137117

Li, K.-C. (1992). On principal Hessian directions for data visualization and dimension reduction: Another application of Stein's lemma. *J. Amer. Statist. Assoc.* **87** 1025–1039. MR1209564

Li, K.-C. and Duan, N. (1989). Regression analysis under link violation. *Ann. Statist.* **17** 1009–1052. MR1015136 https://doi.org/10.1214/aos/1176347254

Li, Y. and Hsing, T. (2010). Deciding the dimension of effective dimension reduction space for functional and high-dimensional data. *Ann. Statist.* **38** 3028–3062. MR2722463 https://doi.org/10.1214/10-AOS816

Li, B. and Song, J. (2017). Nonlinear sufficient dimension reduction for functional data. *Ann. Statist.* **45** 1059–1095. MR3662448 https://doi.org/10.1214/16-AOS1475

Lin, J., Rudi, A., Rosasco, L. and Cevher, V. (2020). Optimal rates for spectral algorithms with least-squares regression over Hilbert spaces. *Appl. Comput. Harmon. Anal.* **48** 868–890. MR4068943 https://doi.org/10.1016/j.acha.2018.09.009

Morris, J.S. (2015). Functional regression. *Annu. Rev. Stat. Appl.* **2** 321–359.

Müller, H.-G. and Stadtmüller, U. (2005). Generalized functional linear models. *Ann. Statist.* **33** 774–805. MR2163159 https://doi.org/10.1214/009053604000001156

Müller, H.-G., Stadtmüller, U. and Yao, F. (2006). Functional variance processes. *J. Amer. Statist. Assoc.* **101** 1007–1018. MR2324140 https://doi.org/10.1198/016214506000000186

Plan, Y. and Vershynin, R. (2016). The generalized Lasso with non-linear observations. *IEEE Trans. Inf. Theory* **62** 1528–1537. MR3472264 https://doi.org/10.1109/TIT.2016.2517008

Rajput, B.S. (1972). Gaussian measures on L_p spaces, $1 \leq p \leq \infty$. *J. Multivariate Anal.* **2** 382–403. MR0345157 https://doi.org/10.1016/0047-259X(72)90034-6

Rajput, B.S. and Cambanis, S. (1972). Gaussian processes and Gaussian measures. *Ann. Math. Stat.* **43** 1944–1952. MR0370728 https://doi.org/10.1214/aoms/1177690865

Ramsay, J.O. and Dalzell, C.J. (1991). Some tools for functional data analysis. *J. Roy. Statist. Soc. Ser. B* **53** 539–572. With discussion and a reply by the authors. MR1125714

Schölkopf, B., Herbrich, R. and Smola, A.J. (2001). A generalized representer theorem. In *Computational Learning Theory (Amsterdam, 2001). Lecture Notes in Computer Science* **2111** 416–426. Berlin: Springer. MR2042050 https://doi.org/10.1007/3-540-44581-1_27

Shang, Z. and Cheng, G. (2015). Nonparametric inference in generalized functional linear models. *Ann. Statist.* **43** 1742–1773. MR3357877 https://doi.org/10.1214/15-AOS1322

Shih, H.-H. (2011). On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces. *J. Funct. Anal.* **261** 1236–1283. MR2807099 https://doi.org/10.1016/j.jfa.2011.04.016

Smale, S. and Zhou, D.-X. (2005). Shannon sampling. II. Connections to learning theory. *Appl. Comput. Harmon. Anal.* **19** 285–302. MR2186447 https://doi.org/10.1016/j.acha.2005.03.001

Smale, S. and Zhou, D.-X. (2007). Learning theory estimates via integral operators and their approximations. *Constr. Approx.* **26** 153–172. MR2327597 https://doi.org/10.1007/s00365-006-0659-y

Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*. Springer.

Tong, H. and Ng, M. (2018). Analysis of regularized least squares for functional linear regression model. *J. Complexity* **49** 85–94. MR3852604 https://doi.org/10.1016/j.jco.2018.08.001

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Wang, C. and Zhou, D.-X. (2011). Optimal learning rates for least squares regularized regression with unbounded sampling. *J. Complexity* **27** 55–67. MR2745300 https://doi.org/10.1016/j.jco.2010.10.002

Wu, Q., Ying, Y. and Zhou, D.-X. (2006). Learning rates of least-square regularized regression. *Found. Comput. Math.* **6** 171–192. MR2228738 https://doi.org/10.1007/s10208-004-0155-9

Yang, Z., Balasubramanian, K. and Liu, H. (2017). High-dimensional non-Gaussian single index models via thresholded score function estimation. In *International Conference on Machine Learning* 3851–3860. PMLR.

Yang, Z., Balasubramanian, K., Wang, Z. and Liu, H. (2017). Estimating high-dimensional non-Gaussian multiple index models via Stein's lemma. In *Advances in Neural Information Processing Systems* **30**.

Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. MR2766857 https://doi.org/10.1214/09-AOS772

Zhu, H., Yao, F. and Zhang, H.H. (2014). Structured functional additive regression in reproducing kernel Hilbert spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 581–603. MR3210729 https://doi.org/10.1111/rssb.12036

Bernoulli **31**(2), 2025, 1007–1031
https://doi.org/10.3150/24-BEJ1756

Theoretical properties of angular halfspace depth

STANISLAV NAGY[a] and PETRA LAKETA[b]

Department of Probability and Mathematical Statistics, Charles University, Prague, Czech Republic,
[a]*nagy@karlin.mff.cuni.cz,* [b]*laketa@karlin.mff.cuni.cz*

The angular halfspace depth (*ahD*) is a natural modification of the celebrated halfspace (or Tukey) depth to the setup of directional data. It allows us to define elements of nonparametric inference, such as the median, the inter-quantile regions, or the rank statistics, for datasets supported on the unit sphere. Despite being introduced in 1987, *ahD* has never received ample recognition in the literature, mainly due to the lack of efficient algorithms for its computation. With the recent progress on the computational front, *ahD* exhibits the potential for developing viable nonparametric statistics techniques for directional datasets. In this paper, we thoroughly treat the theoretical properties of *ahD*. We show that similarly to the classical halfspace depth for multivariate data, also *ahD* satisfies many desirable properties of a statistical depth function. Further, we derive uniform continuity/consistency results for the associated set of directional medians, and the central regions of *ahD*, the latter representing a depth-based analogue of the quantiles for directional data.

Keywords: Angular halfspace depth; angular Tukey's depth; directional data; median; nonparametric methods

References

[1] Agostinelli, C. and Romanazzi, M. (2013). Nonparametric analysis of directional data based on data depth. *Environ. Ecol. Stat.* **20** 253–270. MR3068658 https://doi.org/10.1007/s10651-012-0218-z

[2] Berge, C. (1997). *Topological Spaces: Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity*. Mineola, NY: Dover. Translated from the French original by E. M. Patterson, Reprint of the 1963 translation. MR1464690

[3] Besau, F. and Werner, E.M. (2016). The spherical convex floating body. *Adv. Math.* **301** 867–901. MR3539392 https://doi.org/10.1016/j.aim.2016.07.001

[4] Buttarazzi, D., Pandolfo, G. and Porzio, G.C. (2018). A boxplot for circular data. *Biometrics* **74** 1492–1501. MR3908165 https://doi.org/10.1111/biom.12889

[5] Chernozhukov, V., Galichon, A., Hallin, M. and Henry, M. (2017). Monge–Kantorovich depth, quantiles, ranks and signs. *Ann. Statist.* **45** 223–256. MR3611491 https://doi.org/10.1214/16-AOS1450

[6] Donoho, D.L. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20** 1803–1827. MR1193313 https://doi.org/10.1214/aos/1176348890

[7] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. Revised reprint of the 1989 original. MR1932358 https://doi.org/10.1017/CBO9780511755347

[8] Dyckerhoff, R. (2017). Convergence of depths and depth-trimmed regions. arXiv preprint. Available at arXiv:1611.08721. https://doi.org/10.48550/arXiv.1611.08721

[9] Dyckerhoff, R. and Nagy, S. (2023). Exact computation of angular halfspace depth. Under review.

[10] Hallin, M., Liu, H. and Verdebout, T. (2024). Nonparametric measure-transportation-based methods for directional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* To appear. https://doi.org/10.48550/arXiv.2212.10345.

[11] Huber, P.J. and Ronchetti, E.M. (2009). *Robust Statistics*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR2488795 https://doi.org/10.1002/9780470434697

[12] Hubert, M., Rousseeuw, P. and Segaert, P. (2017). Multivariate and functional classification using depth and distance. *Adv. Data Anal. Classif.* **11** 445–466. MR3688976 https://doi.org/10.1007/s11634-016-0269-3

[13] Hubert, M., Rousseeuw, P.J. and Segaert, P. (2015). Multivariate functional outlier detection. *Stat. Methods Appl.* **24** 177–202. MR3376852 https://doi.org/10.1007/s10260-015-0297-8
[14] Jiang, J. (2010). *Large Sample Techniques for Statistics*. *Springer Texts in Statistics*. New York: Springer. MR2675055 https://doi.org/10.1007/978-1-4419-6827-2
[15] Konen, D. (2022). Topics in Multivariate Spatial Quantiles. Ph.D. thesis, Université Libre de Bruxelles.
[16] Konen, D. and Paindaveine, D. (2023). Spatial quantiles on the hypersphere. *Ann. Statist.* **51** 2221–2245. MR4678802 https://doi.org/10.1214/23-aos2332
[17] Laketa, P. and Nagy, S. (2022). Halfspace depth for general measures: The ray basis theorem and its consequences. *Statist. Papers* **63** 849–883. MR4426418 https://doi.org/10.1007/s00362-021-01259-8
[18] Laketa, P., Pokorný, D. and Nagy, S. (2022). Simple halfspace depth. *Electron. Commun. Probab.* **27** 1–12. MR4529627 https://doi.org/10.1214/22-ecp503
[19] Ley, C., Sabbah, C. and Verdebout, T. (2014). A new concept of quantiles for directional data and the angular Mahalanobis depth. *Electron. J. Stat.* **8** 795–816. MR3217789 https://doi.org/10.1214/14-EJS904
[20] Ley, C. and Verdebout, T. (2017). *Modern Directional Statistics*. *Chapman & Hall/CRC Interdisciplinary Statistics Series*. Boca Raton, FL: CRC Press. MR3752655
[21] Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414. MR1041400 https://doi.org/10.1214/aos/1176347507
[22] Liu, R.Y., Parelius, J.M. and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference. *Ann. Statist.* **27** 783–858. MR1724033 https://doi.org/10.1214/aos/1018031260
[23] Liu, R.Y. and Singh, K. (1992). Ordering directional data: Concepts of data depth on circles and spheres. *Ann. Statist.* **20** 1468–1484. MR1186260 https://doi.org/10.1214/aos/1176348779
[24] Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. Revised reprint of *Statistics of directional data* by Mardia [MR0336854 (49 #1627)]. MR1828667
[25] Mizera, I. and Volauf, M. (2002). Continuity of halfspace depth contours and maximum depth estimators: Diagnostics of depth-related methods. *J. Multivariate Anal.* **83** 365–388. MR1945959 https://doi.org/10.1006/jmva.2001.2053
[26] Mosler, K. (2013). Depth statistics. In *Robustness and Complex Data Structures* (C. Becker, R. Fried and S. Kuhnt, eds.). 17–34. Heidelberg: Springer. MR3135871 https://doi.org/10.1007/978-3-642-35494-6_2
[27] Mosler, K. and Mozharovskyi, P. (2022). Choosing among notions of multivariate depth statistics. *Statist. Sci.* **37** 348–368. MR4444371 https://doi.org/10.1214/21-sts827
[28] Nagy, S., Demni, H., Buttarazzi, D. and Porzio, G.C. (2023). Theory of angular depth for classification of directional data. *Adv. Data Anal. Classif.* To appear. https://doi.org/10.1007/s11634-023-00557-3.
[29] Nagy, S. and Dvořák, J. (2021). Illumination depth. *J. Comput. Graph. Statist.* **30** 78–90. MR4235966 https://doi.org/10.1080/10618600.2020.1776717
[30] Nagy, S., Gijbels, I., Omelka, M. and Hlubinka, D. (2016). Integrated depth for functional data: Statistical properties and consistency. *ESAIM Probab. Stat.* **20** 95–130. MR3528619 https://doi.org/10.1051/ps/2016005
[31] Nagy, S., Schütt, C. and Werner, E.M. (2019). Halfspace depth and floating body. *Stat. Surv.* **13** 52–118. MR3973130 https://doi.org/10.1214/19-ss123
[32] Pandolfo, G., Paindaveine, D. and Porzio, G.C. (2018). Distance-based depths for directional data. *Canad. J. Statist.* **46** 593–609. MR3902616 https://doi.org/10.1002/cjs.11479
[33] Pokorný, D., Laketa, P. and Nagy, S. (2024). Another look at halfspace depth: Flag halfspaces with applications. *J. Nonparametr. Stat.* **36** 165–181. MR4706178 https://doi.org/10.1080/10485252.2023.2236721
[34] Rockafellar, R.T. and Wets, R.J.-B. (1998). *Variational Analysis*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **317**. Berlin: Springer. MR1491362 https://doi.org/10.1007/978-3-642-02431-3
[35] Rousseeuw, P.J. and Ruts, I. (1999). The depth function of a population distribution. *Metrika* **49** 213–244. MR1731769
[36] Schneider, R. (2014). *Convex Bodies: The Brunn–Minkowski Theory*, expanded ed. *Encyclopedia of Mathematics and Its Applications* **151**. Cambridge: Cambridge Univ. Press. MR3155183
[37] Serfling, R. (2006). Depth functions in nonparametric multivariate inference. In *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **72** 1–16. Providence, RI: Amer. Math. Soc. MR2343109 https://doi.org/10.1090/dimacs/072/01

[38] Small, C.G. (1987). Measures of centrality for multivariate and directional distributions. *Canad. J. Statist.* **15** 31–39. MR0887986 https://doi.org/10.2307/3314859

[39] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

[40] Watson, G.S. (1983). *Statistics on Spheres*. *University of Arkansas Lecture Notes in the Mathematical Sciences* **6**. New York: Wiley. MR0709262

[41] Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. *Ann. Statist.* **28** 461–482. MR1790005 https://doi.org/10.1214/aos/1016218226

Bernoulli **31**(2), 2025, 1032–1057
https://doi.org/10.3150/24-BEJ1759

Adaptive estimation of irregular mean and covariance functions

STEVEN GOLOVKINE[1,a], NICOLAS KLUTCHNIKOFF[2,b] and VALENTIN PATILEA[3,c]

[1]*MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland,* [a]*steven.golovkine@ul.ie*
[2]*Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France,* [b]*nicolas.klutchnikoff@univ-rennes2.fr*
[3]*Univ Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France,* [c]*valentin.patilea@ensai.fr*

Nonparametric estimators for the mean and the covariance functions of functional data are proposed. The setup covers a wide range of practical situations. The random trajectories are not necessarily differentiable, have unknown regularity, and are measured with error at discrete design points. The measurement error could be heteroscedastic. The design points could be either randomly drawn or common for all curves. The estimators depend on the local regularity of the stochastic process generating the functional data. We consider a simple estimator of this local regularity which exploits the replication and regularization features of functional data. Next, we use the "smoothing first, then estimate" approach for the mean and the covariance functions. They can be applied with both sparsely or densely sampled curves, are easy to calculate and to update, and perform well in simulations. Simulations built upon an example of a real data set illustrate the effectiveness of the new approach.

Keywords: Functional data analysis; Hölder exponent; kernel smoothing; minimax optimality

References

Balança, P. (2015). Some sample path properties of multifractional Brownian motion. *Stochastic Process. Appl.* **125** 3823–3850. MR3373305 https://doi.org/10.1016/j.spa.2015.05.008

Belloni, A., Chernozhukov, V., Chetverikov, D. and Kato, K. (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *J. Econometrics* **186** 345–366. MR3343791 https://doi.org/10.1016/j.jeconom.2015.02.014

Bertin, K. (2004). Minimax exact constant in sup-norm for nonparametric regression with random design. *J. Statist. Plann. Inference* **123** 225–242. MR2062980 https://doi.org/10.1016/S0378-3758(03)00154-X

Bibinger, M., Jirak, M. and Vetter, M. (2017). Nonparametric change-point analysis of volatility. *Ann. Statist.* **45** 1542–1578. MR3670188 https://doi.org/10.1214/16-AOS1499

Blanke, D. and Vial, C. (2014). Global smoothness estimation of a Gaussian process from general sequence designs. *Electron. J. Stat.* **8** 1152–1187. MR3263116 https://doi.org/10.1214/14-EJS925

Cai, T. and Yuan, M. (2010). Nonparametric covariance function estimation for functional and longitudinal data. University of Pennsylvania and Georgia Institute 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. MR2906870 https://doi.org/10.1214/11-AOS898

Cai, T.T. and Yuan, M. (2016). Minimax and adaptive estimation of covariance operator for random variables observed on a lattice graph. *J. Amer. Statist. Assoc.* **111** 253–265. MR3494657 https://doi.org/10.1080/01621459.2014.1001067

Carroll, C., Gajardo, A., Chen, Y., Dai, X., Fan, J., Hadjipantelis, P.Z., Han, K., Ji, H., Mueller, H.-G. and Wang, J.-L. (2021). *Fdapace: Functional Data Analysis and Empirical Dynamics*. R package version 0.5.6..

Corcuera, J.M., Hedevang, E., Pakkanen, M.S. and Podolskij, M. (2013). Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.* **123** 2552–2574. MR3054536 https://doi.org/10.1016/j.spa.2013.03.011

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. London: CRC Press. MR1383587

Gaïffas, S. (2007). Sharp estimation in sup norm with random design. *Statist. Probab. Lett.* **77** 782–794. MR2369683 https://doi.org/10.1016/j.spl.2006.11.017

Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39** 1608–1632. MR2850214 https://doi.org/10.1214/11-AOS883

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. (2024). Supplement to “Adaptive estimation of irregular mean and covariance functions.” https://doi.org/10.3150/24-BEJ1759SUPP

Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. MR2278365 https://doi.org/10.1214/009053606000000272

Hsing, T., Brown, T. and Thelen, B. (2016). Local intrinsic stationarity and its inference. *Ann. Statist.* **44** 2058–2088. MR3546443 https://doi.org/10.1214/15-AOS1402

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. MR2766854 https://doi.org/10.1214/10-AOS813

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Berlin: Springer. MR1725357 https://doi.org/10.1007/978-3-662-06400-9

Shen, J. and Hsing, T. (2020). Hurst function estimation. *Ann. Statist.* **48** 838–862. MR4102678 https://doi.org/10.1214/19-AOS1825

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

Wong, R.K.W. and Zhang, X. (2019). Nonparametric operator-regularized covariance function estimation for functional data. *Comput. Statist. Data Anal.* **131** 131–144. MR3906800 https://doi.org/10.1016/j.csda.2018.05.013

Zhang, J.-T. and Chen, J. (2007). Statistical inferences for functional data. *Ann. Statist.* **35** 1052–1079. MR2341698 https://doi.org/10.1214/009053606000001505

Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321. MR3546451 https://doi.org/10.1214/16-AOS1446

Zhang, X. and Wang, J.-L. (2018). Optimal weighting schemes for longitudinal and functional data. *Statist. Probab. Lett.* **138** 165–170. MR3788733 https://doi.org/10.1016/j.spl.2018.03.007

Bernoulli **31**(2), 2025, 1058–1083
https://doi.org/10.3150/24-BEJ1760

Change-point detection in low-rank VAR processes

FARIDA ENIKEEVA[1,a], OLGA KLOPP[2,c] and MATHILDE ROUSSELOT[1,b]

[1]*Laboratoire de Mathématiques et Applications, UMR CNRS 7348, Université de Poitiers, Poitiers, France,* [a]*farda.enikeeva@math.univ-poitiers.fr,* [b]*mathilde.rousselot@math.univ-poitiers.fr*
[2]*ESSEC Business School and CREST, Paris, France,* [c]*olga.klopp@math.cnrs.fr*

Vector autoregressive (VAR) models are widely used in multivariate time series analysis for describing the short-time dynamics of the data. The reduced-rank VAR models are of particular interest when dealing with high-dimensional and highly correlated time series. Many results for these models are based on the stationarity assumption that does not hold in several applications when the data exhibits structural breaks. We consider a low-rank piecewise stationary VAR model with possible changes in the transition matrix of the observed process. We develop a new test of the presence of a change-point in the transition matrix and show its minimax optimality with respect to the dimension and the sample size. Our two-step change-point detection strategy is based on the construction of estimators for the transition matrices and using them in a penalized version of the likelihood ratio test statistic. The effectiveness of the proposed method is demonstrated using both synthetic and real data.

Keywords: High-dimensional; minimax optimality; nuclear norm penalization; statistical testing; structural breaks

References

Ahn, S.K. and Reinsel, G.C. (1988). Nested reduced-rank autoregressive models for multiple time series. *J. Amer. Statist. Assoc.* **83** 849–856. MR0963815

Alquier, P., Bertin, K., Doukhan, P. and Garnier, R. (2020). High-dimensional VAR with low-rank transition. *Stat. Comput.* **30** 1139–1153. MR4108695 https://doi.org/10.1007/s11222-020-09929-7

Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. MR2572452 https://doi.org/10.1214/09-AOS707

Bai, P., Safikhani, A. and Michailidis, G. (2020). Multiple change points detection in low rank and sparse high dimensional vector autoregressive models. *IEEE Trans. Signal Process.* **68** 3074–3089. MR4114742 https://doi.org/10.1109/TSP.2020.2993145

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. MR4681620 https://doi.org/10.1080/01621459.2022.2079514

Bańbura, M., Giannone, D. and Reichlin, L. (2010). Large Bayesian vector auto regressions. *J. Appl. Econometrics* **25** 71–92. MR2751790 https://doi.org/10.1002/jae.1137

Basseville, M. and Nikiforov, I.V. (1993). *Detection of Abrupt Changes: Theory and Application*. *Prentice Hall Information and System Sciences Series*. Englewood Cliffs, NJ: Prentice Hall, Inc. MR1210954

Basu, S., Li, X. and Michailidis, G. (2019). Low rank and structured modeling of high-dimensional vector autoregressions. *IEEE Trans. Signal Process.* **67** 1207–1222. MR3925874 https://doi.org/10.1109/TSP.2018.2887401

Basu, S. and Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567. MR3357870 https://doi.org/10.1214/15-AOS1315

Enikeeva, F., Klopp, O. and Rousselot, M. (2024). Supplement to "Change-point detection in low-rank VAR processes." https://doi.org/10.3150/24-BEJ1760SUPP

Fan, J., Lv, J. and Qi, L. (2011). Sparse high-dimensional models in economics. *Ann. Rev. Econ.* **3** 291–317.

Gorrostieta, C., Fiecas, M., Ombao, H., Burke, E. and Cramer, S. (2013). Hierarchical vector auto-regressive models and their applications to multi-subject effective connectivity. *Front. Comput. Neurosci.* **7**.

1350-7265

Ji, S. and Ye, J. (2009). An accelerated gradient method for trace norm minimization. In *Proceedings of the 26th Annual International Conference on Machine Learning*. *ACM International Conference Proceeding Series* **382** 457–464. ACM.

Khan, F., Saeed, A. and Ali, S. (2020). Modelling and forecasting of new cases, deaths and recover cases of COVID-19 by using vector autoregressive model in Pakistan. *Chaos Solitons Fractals* **140** 110189, 5. MR4135451 https://doi.org/10.1016/j.chaos.2020.110189

Liu, H., Gao, C. and Samworth, R.J. (2021). Minimax rates in sparse, high-dimensional change point detection. *Ann. Statist.* **49** 1081–1112. MR4255120 https://doi.org/10.1214/20-aos1994

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer. MR2172368 https://doi.org/10.1007/978-3-540-27752-1

McCracken, M.W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *J. Bus. Econom. Statist.* **34** 574–589. MR3547997 https://doi.org/10.1080/07350015.2015.1086655

Melnyk, I. and Banerjee, A. (2016). Estimating structured vector autoregressive models. In *Proceedings of the 33rd International Conference on Machine Learning* **48** 830–839. New York, USA: PMLR.

Michel, C.M. and Murray, M.M. (2012). Towards the utilization of EEG as a brain imaging tool. *NeuroImage* **61** 371–385. https://doi.org/10.1016/j.neuroimage.2011.12.039

Negahban, S. and Wainwright, M.J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.* **39** 1069–1097. MR2816348 https://doi.org/10.1214/10-AOS850

Reinsel, G.C., Velu, R.P. and Chen, K. (1998). *Multivariate Reduced-Rank Regression: Theory and Applications*. *Lecture Notes in Statistics* **136**. New York: Springer. MR1719704 https://doi.org/10.1007/978-1-4757-2853-8

Safikhani, A. and Shojaie, A. (2022). Joint structural break detection and parameter estimation in high-dimensional nonstationary VAR models. *J. Amer. Statist. Assoc.* **117** 251–264. MR4399083 https://doi.org/10.1080/01621459.2020.1770097

She, Y., He, Y., Li, S. and Wu, D. (2015). Joint association graph screening and decomposition for large-scale linear dynamical systems. *IEEE Trans. Signal Process.* **63** 389–401. MR3299471 https://doi.org/10.1109/TSP.2014.2373315

Stock, J.H. and Watson, M.W. (2001). Vector autoregressions. *J. Electron. Publ.* **15** 101–115.

Velu, R.P., Reinsel, G.C. and Wichern, D.W. (1986). Reduced rank models for multiple time series. *Biometrika* **73** 105–118. MR0836438 https://doi.org/10.1093/biomet/73.1.105

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. With a foreword by Sara van de Geer. MR3837109 https://doi.org/10.1017/9781108231596

Verzelen, N., Fromont, M., Lerasle, M. and Reynaud-Bouret, P. (2023). Optimal change-point detection and localization. *Ann. Statist.* **51** 1586–1610. MR4658569 https://doi.org/10.1214/23-aos2297

Wang, D. and Tsay, R.S. (2023). Rate-optimal robust estimation of high-dimensional vector autoregressive models. *Ann. Statist.* **51** 846–877. MR4601004 https://doi.org/10.1214/23-aos2278

Wang, D., Yu, Y., Rinaldo, A. and Willett, R. (2019). Localizing Changes in High-Dimensional Vector Autoregressive Processes. arXiv:1909.06359.

Wild, B., Eichler, M., Friederich, H.-C., Hartmann, M., Zipfel, S. and Herzog, W. (2010). A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC Med. Res. Methodol.* **10** 28.

Bernoulli **31**(2), 2025, 1084–1113
https://doi.org/10.3150/24-BEJ1761

Hölder regularity and roughness: Construction and examples

ERHAN BAYRAKTAR[1,a], PURBA DAS[2,b] and DONGHAN KIM[3,c]

[1] *Department of Mathematics, University of Michigan, Ann Arbor, USA,* [a] *erhan@umich.edu*
[2] *Department of Mathematics, King's College London, London, UK,* [b] *purba.das@kcl.ac.uk*
[3] *Department of Mathematical Sciences, KAIST, Daejeon, South Korea,* [c] *kimdonghan@kaist.ac.kr*

We study how to construct a stochastic process on a finite interval with given 'roughness' and its finite joint moments. We first extend Ciesielski's isomorphism (1960) along a general sequence of partitions, and provide a characterization of Hölder regularity of a function in terms of the coefficients along Schauder basis. Using this characterization we propose a better (pathwise) estimator of Hölder exponent. As an additional application, we construct fake (fractional) Brownian motions with some path properties and finite moments same as (fractional) Brownian motions. These belong to non-Gaussian families of stochastic processes which are statistically difficult to distinguish from real (fractional) Brownian motions.

Keywords: Ciesielski's isomorphism; fake fractional Brownian motion; fractional Brownian motion; generalized Faber-Schauder system; matching moments; normality tests; pathwise Hölder regularity estimator; roughness

References

[1] Albin, J.M.P. (2008). A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statist. Probab. Lett.* **78** 682–686. MR2409532 https://doi.org/10.1016/j.spl.2007.09.031
[2] Bayraktar, E., Das, P. and Kim, D. (2024). Supplement to "Hölder regularity and roughness: Construction and examples." https://doi.org/10.3150/24-BEJ1761SUPPA
[3] Bayraktar, E., Das, P. and Kim, D. (2024). Supplement to "Hölder regularity and roughness: Construction and examples." https://doi.org/10.3150/24-BEJ1761SUPPB
[4] Beiglböck, M., Lowther, G., Pammer, G. and Schachermayer, W. (2024). Faking Brownian motion with continuous Markov martingales. *Finance Stoch.* **28** 259–284. MR4682847 https://doi.org/10.1007/s00780-023-00526-w
[5] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. *Probability and Its Applications (New York)*. London: Springer London, Ltd. MR2387368 https://doi.org/10.1007/978-1-84628-797-8
[6] Ciesielski, Z. (1960). On the isomorphisms of the spaces H_α and m. *Bull. Acad. Pol. Sci., Sér. Sci. Math. Astron. Phys.* **8** 217–222. MR0132389
[7] Collins, J. and De Luca, C. (1995). Upright, correlated random walks: A statistical-biomechanics approach to the human postural control system. *Chaos* **5** 57–63.
[8] Cont, R. and Das, P. (2022). Quadratic variation along refining partitions: Constructions and examples. *J. Math. Anal. Appl.* **512** Paper No. 126173, 35. MR4401981 https://doi.org/10.1016/j.jmaa.2022.126173
[9] Cont, R. and Das, P. (2023). Quadratic variation and quadratic roughness. *Bernoulli* **29** 496–522. MR4497256 https://doi.org/10.3150/22-bej1466
[10] Cont, R. and Das, P. (2024). Measuring the roughness of a signal. Working Paper.
[11] Echelard, A. (2007). Analyse 2-microlocale et application au débruitage Thesis, Ecole Centrale de Nantes (ECN); Univ. Nantes.
[12] Filatova, D.V. and Grzywaczewski, M. (2008). Mathematical modeling in selected biological systems with fractional Brownian motion. In *2008 Conference on Human System Interactions* 909–914. IEEE.

1350-7265

[13] Gurbatov, S.N., Simdyankin, S.I., Aurell, E., Frisch, U. and Tóth, G. (1997). On the decay of Burgers turbulence. *J. Fluid Mech.* **344** 339–374. MR1471369 https://doi.org/10.1017/S0022112097006241

[14] Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Math. Ann.* **69** 331–371. MR1511592 https://doi.org/10.1007/BF01456326

[15] Hamdouche, M., Henry-Labordere, P. and Pham, H. (2023). Generative modeling for time series via Schrödinger bridge. Preprint. Available at arXiv:2304.05093.

[16] Hamza, K. and Klebaner, F.C. (2007). A family of non-Gaussian martingales with Gaussian marginals. *J. Appl. Math. Stoch. Anal.* Art. ID 92723, 19. MR2335977 https://doi.org/10.1155/2007/92723

[17] Han, X. and Schied, A. (2021). The Hurst roughness exponent and its model-free estimation. Preprint. Available at arXiv:2111.10301.

[18] Hobson, D. (2016). Mimicking martingales. *Ann. Appl. Probab.* **26** 2273–2303. MR3543897 https://doi.org/10.1214/15-AAP1147

[19] Hobson, D.G. (2013). Fake exponential Brownian motion. *Statist. Probab. Lett.* **83** 2386–2390. MR3093829 https://doi.org/10.1016/j.spl.2013.06.030

[20] Jaffard, S. (2004). Wavelet techniques in multifractal analysis. In *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Part 2*. *Proc. Sympos. Pure Math.* **72** 91–151. Providence, RI: Amer. Math. Soc. MR2112122 https://doi.org/10.1090/pspum/072.2/2112122

[21] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

[22] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V. (1994). On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Trans. Netw.* **2** 1–15.

[23] Lo, A.W. (1991). Long-term memory in stock market prices. *Econometrica* **59** 1279–1313.

[24] Madan, D.B. and Yor, M. (2002). Making Markov martingales meet marginals: With explicit constructions. *Bernoulli* **8** 509–536. MR1914701

[25] Mishura, Y. and Schied, A. (2016). Constructing functions with prescribed pathwise quadratic variation. *J. Math. Anal. Appl.* **442** 117–137. MR3498321 https://doi.org/10.1016/j.jmaa.2016.04.056

[26] Mishura, Y. and Valkeila, E. (2011). An extension of the Lévy characterization to fractional Brownian motion. *Ann. Probab.* **39** 439–470. MR2789503 https://doi.org/10.1214/10-AOP555

[27] Mishura, Y.S. (2008). *Stochastic Calculus for Fractional Brownian Motion and Related Processes*. *Lecture Notes in Math.* **1929**. Berlin: Springer. MR2378138 https://doi.org/10.1007/978-3-540-75873-0

[28] Oleszkiewicz, K. (2008). On fake Brownian motions. *Statist. Probab. Lett.* **78** 1251–1254. MR2444313 https://doi.org/10.1016/j.spl.2007.11.026

[29] Rogers, L.C.G. (1997). Arbitrage with fractional Brownian motion. *Math. Finance* **7** 95–105. MR1434408 https://doi.org/10.1111/1467-9965.00025

[30] Sadhu, T. and Wiese, K.J. (2021). Functionals of fractional Brownian motion and the three arcsine laws. *Phys. Rev. E* **104** Paper No. 054112, 47. MR4349821 https://doi.org/10.1103/physreve.104.054112

[31] Schauder, J. (1928). Eine Eigenschaft des Haarschen Orthogonalsystems. *Math. Z.* **28** 317–320. MR1544958 https://doi.org/10.1007/BF01181164

[32] Schied, A. (2016). On a class of generalized Takagi functions with linear pathwise quadratic variation. *J. Math. Anal. Appl.* **433** 974–990. MR3398747 https://doi.org/10.1016/j.jmaa.2015.08.022

[33] Semadeni, Z. (1982). *Schauder Bases in Banach Spaces of Continuous Functions*. *Lecture Notes in Math.* **918**. Berlin: Springer. MR0653986

[34] Taqqu, M.S., Teverovsky, V. and Willinger, W. (1997). Is network traffic self-similar or multifractal? *Fractals* **5** 63–73.

Bernoulli **31**(2), 2025, 1114–1139
https://doi.org/10.3150/24-BEJ1763

Optimal spectral recovery of a planted vector in a subspace

CHENG MAO[1,a] and ALEXANDER S. WEIN[2,b]

[1] *School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA,* [a]*cheng.mao@math.gatech.edu*
[2] *Department of Mathematics, University of California, Davis, Davis, CA, USA,* [b]*aswein@ucdavis.edu*

Recovering a planted vector v in an n-dimensional random subspace of $\mathbb{R}^N$ is a generic task related to many problems in machine learning and statistics, such as dictionary learning, subspace recovery, principal component analysis, and non-Gaussian component analysis. In this work, we study computationally efficient estimation and detection of a planted vector v whose ℓ_4 norm differs from that of a Gaussian vector with the same ℓ_2 norm. For instance, in the special case where v is an $N\rho$-sparse vector with Bernoulli–Gaussian or Bernoulli–Rademacher entries, our results include the following: (1) We give an improved analysis of a slight variant of the spectral method proposed by Hopkins, Schramm, Shi, and Steurer (2016), showing that it approximately recovers v with high probability in the regime $n\rho \ll \sqrt{N}$. This condition subsumes the conditions $\rho \ll 1/\sqrt{n}$ or $n\sqrt{\rho} \lesssim \sqrt{N}$ required by previous work up to polylogarithmic factors. We achieve ℓ_∞ error bounds for the spectral estimator via a leave-one-out analysis, from which it follows that a simple thresholding procedure exactly recovers v with Bernoulli–Rademacher entries, even in the dense case $\rho = 1$. (2) We study the associated detection problem and show that in the regime $n\rho \gg \sqrt{N}$, any spectral method from a large class (and more generally, any low-degree polynomial of the input) fails to detect the planted vector. This matches the condition for recovery and offers evidence that no polynomial-time algorithm can succeed in recovering a Bernoulli–Gaussian vector v when $n\rho \gg \sqrt{N}$.

Keywords: Computational lower bound; eigenvector perturbation; low-degree method; planted sparse vector; spectral method

References

Abbe, E., Fan, J., Wang, K. and Zhong, Y. (2020). Entrywise eigenvector analysis of random matrices with low expected rank. *Ann. Statist.* **48** 1452–1474. MR4124330 https://doi.org/10.1214/19-AOS1854

Bandeira, A.S., Kunisky, D. and Wein, A.S. (2020). Computational hardness of certifying bounds on constrained PCA problems. In *11th Innovations in Theoretical Computer Science Conference*. *LIPIcs. Leibniz Int. Proc. Inform.* **151** Art. No. 78, 29. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4048181

Barak, B., Kelner, J.A. and Steurer, D. (2014). Rounding sum-of-squares relaxations. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing* 31–40.

Barak, B., Hopkins, S., Kelner, J., Kothari, P.K., Moitra, A. and Potechin, A. (2019). A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM J. Comput.* **48** 687–735. MR3945259 https://doi.org/10.1137/17M1138236

Bean, D.M. (2014). Non-Gaussian component analysis. University of California, Berkeley. MR3295320

Blanchard, G., Kawanabe, M., Sugiyama, M., Spokoiny, V., Müller, K.-R. and Roweis, S. (2006). In search of non-Gaussian components of a high-dimensional distribution. *J. Mach. Learn. Res.* **7** 247–282. MR2274368

Candès, E.J., Li, X., Ma, Y. and Wright, J. (2011). Robust principal component analysis? *J. ACM* **58** Art. 11, 37. MR2811000 https://doi.org/10.1145/1970392.1970395

Cape, J., Tang, M. and Priebe, C.E. (2019). The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics. *Ann. Statist.* **47** 2405–2439. MR3988761 https://doi.org/10.1214/18-AOS1752

Chen, Y., Chi, Y., Fan, J., Ma, C. et al. (2021). Spectral methods for data science: A statistical perspective. *Found. Trends Mach. Learn.* **14** 566–806.

1350-7265

d'Orsi, T., Kothari, P.K., Novikov, G. and Steurer, D. (2020). Sparse PCA: Algorithms, adversarial perturbations and certificates. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science* 553–564. Los Alamitos, CA: IEEE Computer Soc. MR4232066 https://doi.org/10.1109/FOCS46700.2020.00058

Davis, D., Diaz, M. and Wang, K. (2021). Clustering a mixture of gaussians with unknown covariance. arXiv preprint, arXiv:2110.01602.

Demanet, L. and Hand, P. (2014). Scaling law for recovering the sparsest element in a subspace. *Inf. Inference* **3** 295–309. MR3311454 https://doi.org/10.1093/imaiai/iau007

Diakonikolas, I. and Kane, D. (2022). Non-Gaussian component analysis via lattice basis reduction. *Proc. Mach. Learn. Res.* **178** 4535–4547.

Diakonikolas, I., Kane, D.M. and Stewart, A. (2017). Statistical query lower bounds for robust estimation of high-dimensional Gaussians and Gaussian mixtures (extended abstract). In *58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017* 73–84. Los Alamitos, CA: IEEE Computer Soc. MR3734219 https://doi.org/10.1109/FOCS.2017.16

Diakonikolas, I., Kane, D., Pensia, A., Pittas, T. and Stewart, A. (2021). Statistical query lower bounds for list-decodable linear regression. *Adv. Neural Inf. Process. Syst.* **34** 3191–3204.

Ding, Y., Kunisky, D., Wein, A.S. and Bandeira, A.S. (2021). The average-case time complexity of certifying the restricted isometry property. *IEEE Trans. Inf. Theory* **67** 7355–7361. MR4345126 https://doi.org/10.1109/TIT.2021.3112823

Dudeja, R. and Hsu, D. (2024). Statistical-computational trade-offs in tensor PCA and related problems via communication complexity. *Ann. Statist.* **52** 131–156. MR4718410 https://doi.org/10.1214/23-aos2331

Fan, J., Wang, W. and Zhong, Y. (2017). An ℓ_∞ eigenvector perturbation bound and its application to robust covariance estimation. *J. Mach. Learn. Res.* **18** Paper No. 207, 42. MR3827095

Ghosh, M., Jeronimo, F.G., Jones, C., Potechin, A. and Rajendran, G. (2020). Sum-of-squares lower bounds for Sherrington-Kirkpatrick via planted affine planes. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science* 954–965. Los Alamitos, CA: IEEE Computer Soc. MR4232101

Goyal, N. and Shetty, A. (2019). Non-Gaussian component analysis using entropy methods. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* 840–851. New York: ACM. MR4003388

Hopkins, S. (2018). Statistical inference and the sum of squares method. PhD Thesis, Cornell University. MR3864930

Hopkins, S.B. and Steurer, D. (2017). Efficient Bayesian estimation from few samples: Community detection and related problems. In *58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017* 379–390. Los Alamitos, CA: IEEE Computer Soc. MR3734245 https://doi.org/10.1109/FOCS.2017.42

Hopkins, S.B., Shi, J., Schramm, T. and Steurer, D. (2016). Fast spectral algorithms from sum-of-squares proofs: Tensor decomposition and planted sparse vectors. In *STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing* 178–191. New York: ACM. MR3536564 https://doi.org/10.1145/2897518.2897529

Hopkins, S.B., Kothari, P.K., Potechin, A., Raghavendra, P., Schramm, T. and Steurer, D. (2017). The power of sum-of-squares for detecting hidden structures. In *58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017* 720–731. Los Alamitos, CA: IEEE Computer Soc. MR3734275 https://doi.org/10.1109/FOCS.2017.72

Koltchinskii, V. and Xia, D. (2016). Perturbation of linear forms of singular vectors under Gaussian noise. In *High Dimensional Probability VII. Progress in Probability* **71** 397–423. Cham: Springer. MR3565274 https://doi.org/10.1007/978-3-319-40519-3_18

Kunisky, D., Wein, A.S. and Bandeira, A.S. (2022). Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *Mathematical Analysis, Its Applications and Computation. Springer Proc. Math. Stat.* **385** 1–50. Cham: Springer. MR4461037 https://doi.org/10.1007/978-3-030-97127-4_1

Lei, L. (2019). Unified $\ell_{2\to\infty}$ eigenspace perturbation theory for symmetric random matrices. arXiv preprint, arXiv:1909.04798.

Mao, C. and Wein, A.S. (2021). Optimal spectral recovery of a planted vector in a subspace. arXiv preprint, arXiv:2105.15081v1.

Mao, C. and Wein, A.S. (2024). Supplement to "Optimal spectral recovery of a planted vector in a subspace." https://doi.org/10.3150/24-BEJ1763SUPP

Mohanty, S., Raghavendra, P. and Xu, J. (2020). Lifting sum-of-squares lower bounds: Degree-2 to degree-4. In *STOC '20—Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing* 840–853. New York: ACM. MR4141804 https://doi.org/10.1145/3357713.3384319

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

Qu, Q., Sun, J. and Wright, J. (2016). Finding a sparse vector in a subspace: Linear sparsity using alternating directions. *IEEE Trans. Inf. Theory* **62** 5855–5880. MR3552428 https://doi.org/10.1109/TIT.2016.2601599

Qu, Q., Zhu, Z., Li, X., Tsakiris, M.C., Wright, J. and Vidal, R. (2020). Finding the sparsest vectors in a subspace: Theory, algorithms, and applications. arXiv preprint, arXiv:2001.06970.

Schramm, T. and Wein, A.S. (2022). Computational barriers to estimation from low-degree polynomials. *Ann. Statist.* **50** 1833–1858. MR4441142 https://doi.org/10.1214/22-aos2179

Spielman, D.A., Wang, H. and Wright, J. (2012). Exact recovery of sparsely-used dictionaries. In *Conference on Learning Theory*. *Proc. Mach. Learn. Res.* **23** 37–1.

Sun, J., Qu, Q. and Wright, J. (2017). Complete dictionary recovery over the sphere II: Recovery by Riemannian trust-region method. *IEEE Trans. Inf. Theory* **63** 885–914. MR3604647 https://doi.org/10.1109/TIT.2016.2632149

Szegö, G. (1939). *Orthogonal Polynomials*. *American Mathematical Society Colloquium Publications* **23**. New York: Amer. Math. Soc. MR0000077

Tan, Y.S. and Vershynin, R. (2018). Polynomial time and sample complexity for non-Gaussian component analysis: Spectral methods. In *Conference on Learning Theory*. *Proc. Mach. Learn. Res.* **75** 498–534.

Vempala, S.S. and Xiao, Y. (2011). Structure from local optima: Learning subspace juntas via higher order PCA. arXiv preprint, arXiv:1108.3329.

Zadik, I., Song, M.J., Wein, A.S. and Bruna, J. (2022). Lattice-based methods surpass sum-of-squares in clustering. In *Conference on Learning Theory*. *Proc. Mach. Learn. Res.* **178** 1247–1248.

Zhai, Y., Mehta, H., Zhou, Z. and Ma, Y. (2020a). Understanding l4-based dictionary learning: Interpretation, stability, and robustness. In *International Conference on Learning Representations* 1–28.

Zhai, Y., Yang, Z., Liao, Z., Wright, J. and Ma, Y. (2020b). Complete dictionary learning via ℓ^4-norm maximization over the orthogonal group. *J. Mach. Learn. Res.* **21** Paper No. 165, 68. MR4209451

Bernoulli **31**(2), 2025, 1140–1159
https://doi.org/10.3150/24-BEJ1764

Isotonic conditional laws

SEBASTIAN ARNOLD[a] and JOHANNA ZIEGEL[b]

Seminar for Statistics, ETH Zurich, Switzerland, [a]*sebastian.arnold@stat.math.ethz.ch,* [b]*ziegel@stat.math.ethz.ch*

We introduce isotonic conditional laws (ICL) which extend the classical notion of conditional laws by the additional requirement that there exists an isotonic relationship between the random variable of interest and the conditioning random object. We show existence and uniqueness of ICL building on conditional expectations given σ-lattices. ICL corresponds to a classical conditional law if and only if the latter is already isotonic. ICL is motivated from a statistical point of view by showing that ICL emerges equivalently as the minimizer of an expected score where the scoring rule may be taken from a large class comprising the continuous ranked probability score (CRPS). Furthermore, ICL is calibrated in the sense that it is invariant to certain conditioning operations, and the corresponding event probabilities and quantiles are simultaneously optimal with respect to all relevant scoring functions. We develop a new notion of general conditional functionals given σ-lattices which is of independent interest.

Keywords: Calibration; conditional functional; conditional law; isotonicity; σ-lattice

References

Armerin, F. (2014). The conditional quantile as a minimizer. Working Paper.

Arnold, S. (2020). Isotonic distributional approximation. Master's thesis, Universität Bern.

Arnold, S., Henzi, A. and Ziegel, J.F. (2023). Sequentially valid tests for forecast calibration. *Ann. Appl. Stat.* **17** 1909–1935. MR4637650 https://doi.org/10.1214/22-aoas1697

Arnold, S. and Ziegel, J. (2024). Supplement to "Isotonic conditional laws." https://doi.org/10.3150/24-BEJ1764SUPP

Ayer, M., Brunk, H.D., Ewing, G.M., Reid, W.T. and Silverman, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Stat.* **26** 641–647. MR0073895 https://doi.org/10.1214/aoms/1177728423

Barlow, R.E., Brunk, H.D., Bartholomew, D.J. and Bremner, J.M. (1972). *Statistical Inference Under Order Restrictions*. London: Wiley.

Bauer, H. (1996). *Probability Theory*. *De Gruyter Studies in Mathematics* **23**. Berlin: de Gruyter. Translated from the fourth (1991) German edition by Robert B. Burckel and revised by the author. MR1385460 https://doi.org/10.1515/9783110814668

Bauer, H. (2001). *Measure and Integration Theory*. *De Gruyter Studies in Mathematics* **26**. Berlin: de Gruyter. Translated from the German by Robert B. Burckel. MR1897176 https://doi.org/10.1515/9783110866209

Brunk, H.D. (1961). Best fit to a random variable by a random variable measurable with respect to a σ-lattice. *Pacific J. Math.* **11** 785–802. MR0133844

Brunk, H.D. (1963). On an extension of the concept conditional expectation. *Proc. Amer. Math. Soc.* **14** 298–304. MR0148090 https://doi.org/10.2307/2034631

Brunk, H.D. (1965). Conditional expectation given a σ-lattice and applications. *Ann. Math. Stat.* **36** 1339–1350. MR0185629 https://doi.org/10.1214/aoms/1177699895

Brunk, H.D. and Johansen, S. (1970). A generalized Radon-Nikodym derivative. *Pacific J. Math.* **34** 585–617. MR0267066

Dawid, A.P. (1984). Statistical theory. The prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–292. MR0763811 https://doi.org/10.2307/2981683

Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998). Evaluating density forecasts with applications to financial risk management. *Internat. Econom. Rev.* **39** 863–883.

1350-7265

Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 505–562. MR3506792 https://doi.org/10.1111/rssb.12154

El Barmi, H. and Mukerjee, H. (2005). Inferences under a stochastic ordering constraint: The k-sample case. *J. Amer. Statist. Assoc.* **100** 252–261. MR2156835 https://doi.org/10.1198/016214504000000764

Embrechts, P. and Hofert, M. (2013). A note on generalized inverses. *Math. Methods Oper. Res.* **77** 423–432. MR3072795 https://doi.org/10.1007/s00186-013-0436-7

Fissler, T. and Holzmann, H. (2022). Measurability of functionals and of ideal point forecasts. *Electron. J. Stat.* **16** 5019–5034. MR4490414 https://doi.org/10.1214/22-ejs2062

Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. MR2847988 https://doi.org/10.1198/jasa.2011.r10138

Gneiting, T. and Katzfuss, M. (2014). Probabilistic forecasting. *Annu. Rev. Stat. Appl.* **1** 125–151.

Gneiting, T. and Raftery, A.E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548 https://doi.org/10.1198/016214506000001437

Gneiting, T. and Ranjan, R. (2013). Combining predictive distributions. *Electron. J. Stat.* **7** 1747–1782. MR3080409 https://doi.org/10.1214/13-EJS823

Gneiting, T. and Resin, J. (2023). Regression diagnostics meets forecast evaluation: Conditional calibration, reliability diagrams, and coefficient of determination. *Electron. J. Stat.* **17** 3226–3286. MR4669757 https://doi.org/10.1214/23-ejs2180

Guo, C., Pleiss, G., Sun, Y. and Weinberger, K.Q. (2017). On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning* (D. Precup and Y.W. Teh, eds.). *Proceedings of Machine Learning Research* **70** 1321–1330. PMLR.

Gupta, C., Podkopaev, A. and Ramdas, A. (2020). Distribution-free binary classification: Prediction sets, confidence intervals and calibration. In *Advances in Neural Information Processing Systems* (H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan and H. Lin, eds.) **33** 3711–3723. Curran Associates.

Henzi, A., Kleger, G.-R. and Ziegel, J.F. (2023). Distributional (single) index models. *J. Amer. Statist. Assoc.* **118** 489–503. MR4571136 https://doi.org/10.1080/01621459.2021.1938582

Henzi, A. and Ziegel, J.F. (2022). Valid sequential inference on probability forecast performance. *Biometrika* **109** 647–663. MR4472840 https://doi.org/10.1093/biomet/asab047

Henzi, A., Ziegel, J.F. and Gneiting, T. (2021). Isotonic distributional regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 963–993. MR4349124

Jordan, A.I., Mühlemann, A. and Ziegel, J.F. (2022). Characterizing the optimal solutions to the isotonic regression problem for identifiable functionals. *Ann. Inst. Statist. Math.* **74** 489–514. MR4417369 https://doi.org/10.1007/s10463-021-00808-0

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

Kuenzi, N.J. (1969). An investigation of some problems concerning the additivity of conditional expectation with respect to a sigma-lattice. PhD thesis, University of Iowa. MR2618670

Lee, C.-I.C. (1976). Conditional Expectation given a σ-lattice, Isotonic Regression, Univariate and Multivariate. PhD thesis, Oregon State University. MR2625527

Matheson, J.E. and Winkler, R.L. (1976). Scoring rules for continuous probability distributions. *Manage. Sci.* **22** 1087–1096.

Mösching, A. and Dümbgen, L. (2020). Monotone least squares and isotonic quantiles. *Electron. J. Stat.* **14** 24–49. MR4047593 https://doi.org/10.1214/19-EJS1659

Richmond, T. (2020). *General Topology—an Introduction*. *De Gruyter Textbook*. Berlin: de Gruyter. MR4202805 https://doi.org/10.1515/9783110686579-201

Schulz, B. and Lerch, S. (2022). Machine learning methods for postprocessing ensemble forecasts of wind gusts: A systematic comparison. *Mon. Weather Rev.* **150** 235–257.

Strähl, C. and Ziegel, J. (2017). Cross-calibration of probabilistic forecasts. *Electron. J. Stat.* **11** 608–639. MR3619318 https://doi.org/10.1214/17-EJS1244

Thorarinsdottir, T.L., Scheuerer, M. and Heinz, C. (2016). Assessing the calibration of high-dimensional ensemble forecasts using rank histograms. *J. Comput. Graph. Statist.* **25** 105–122. MR3474039 https://doi.org/10.1080/10618600.2014.977447

Tomkins, R.J. (1975). On conditional medians. *Ann. Probab.* **3** 375–379. MR0385947 https://doi.org/10.1214/aop/1176996411

Tsyplakov, A. (2013). Evaluation of probabilistic forecasts: Proper scoring rules and moments. Preprint.

Ziegel, J.F. and Gneiting, T. (2014). Copula calibration. *Electron. J. Stat.* **8** 2619–2638. MR3291527 https://doi.org/10.1214/14-EJS964

Bernoulli **31**(2), 2025, 1160–1180
https://doi.org/10.3150/24-BEJ1765

Speeding up Monte Carlo integration: Control neighbors for optimal convergence

RÉMI LELUC[1,a], FRANÇOIS PORTIER[2,b], JOHAN SEGERS[3,c] and AIGERIM ZHUMAN[3,d]

[1]*CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France,* [a]*remi.leluc@gmail.com*
[2]*CREST, ENSAI, Bruz, France,* [b]*francois.portier@gmail.com*
[3]*ISBA/LIDAM, UCLouvain, Louvain-la-Neuve, Belgium,* [c]*johan.segers@uclouvain.be,* [d]*aigerim.zhuman@uclouvain.be*

A novel linear integration rule called *control neighbors* is proposed in which nearest neighbor estimates act as control variates to speed up the convergence rate of the Monte Carlo procedure on metric spaces. The main result is the $O(n^{-1/2}n^{-s/d})$ convergence rate – where n stands for the number of evaluations of the integrand and d for the dimension of the domain – of this estimate for Hölder functions with regularity $s \in (0,1]$, a rate which, in some sense, is optimal. Several numerical experiments validate the complexity bound and highlight the good performance of the proposed estimator.

Keywords: Control variates; Hölder functions; metric space; Monte Carlo; nearest neighbor; optimal convergence rate; variance reduction; Voronoi cell

References

Achddou, J., Lam-Weil, J., Carpentier, A. and Blanchard, G. (2019). A minimax near-optimal algorithm for adaptive rejection sampling. In *Algorithmic Learning Theory* 94–126. PMLR.

Audibert, J.-Y. and Tsybakov, A.B. (2007). Fast learning rates for plug-in classifiers. *Ann. Statist.* **35** 608–633. MR2336861 https://doi.org/10.1214/009053606000001217

Bakhvalov, N.S. (2015). On the approximate calculation of multiple integrals [translation of 0115275]. *J. Complexity* **31** 502–516. MR3348079 https://doi.org/10.1016/j.jco.2014.12.003

Bardenet, R. and Hardy, A. (2020). Monte Carlo with determinantal point processes. *Ann. Appl. Probab.* **30** 368–417. MR4068314 https://doi.org/10.1214/19-AAP1504

Bentley, J.L. (1975). Multidimensional binary search trees used for associative searching. *Commun. ACM* **18** 509–517.

Biau, G. and Devroye, L. (2015). *Lectures on the Nearest Neighbor Method*. *Springer Series in the Data Sciences*. Cham: Springer. MR3445317 https://doi.org/10.1007/978-3-319-25388-6

Blanchet, J., Chen, H., Lu, Y. and Ying, L. (2023). When can regression-adjusted control variates help? In *Adv. Neural Inf. Process. Syst.* **36** 36566–36578.

Bonneel, N., Rabin, J., Peyré, G. and Pfister, H. (2015). Sliced and Radon Wasserstein barycenters of measures. *J. Math. Imaging Vision* **51** 22–45. MR3300482 https://doi.org/10.1007/s10851-014-0506-3

Caflisch, R.E. (1998). Monte Carlo and quasi-Monte Carlo methods. In *Acta Numerica, 1998*. *Acta Numer.* **7** 1–49. Cambridge: Cambridge Univ. Press. MR1689431 https://doi.org/10.1017/S0962492900002804

Chopin, N. and Gerber, M. (2024). Higher-order Monte Carlo through cubic stratification. *SIAM J. Numer. Anal.* **62** 229–247. MR4695762 https://doi.org/10.1137/22M1532287

Coeurjolly, J.-F., Mazoyer, A. and Amblard, P.-O. (2021). Monte Carlo integration of non-differentiable functions on $[0,1]^\iota$, $\iota = 1,\dots,d$, using a single determinantal point pattern defined on $[0,1]^d$. *Electron. J. Stat.* **15** 6228–6280. MR4355707 https://doi.org/10.1214/21-ejs1929

Combes, R. (2015). An extension of McDiarmid's inequality. ArXiv preprint. Available at arXiv:1511.05240.

Courty, N., Flamary, R., Habrard, A. and Rakotomamonjy, A. (2017). Joint distribution optimal transportation for domain adaptation. *Adv. Neural Inf. Process. Syst.* **30** 3733–3742.

Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions. *Numer. Math.* **31** 377–403.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Adv. Neural Inf. Process. Syst.* **26** 2292–2300.

Diaconis, P. and Evans, S.N. (2001). Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** 2615–2633. MR1828463 https://doi.org/10.1090/S0002-9947-01-02800-8

Doucet, A., De Freitas, N., Gordon, N.J. et al. (2001). *Sequential Monte Carlo Methods in Practice* **2**. New York: Springer.

Evans, M. and Swartz, T. (2000). *Approximating Integrals via Monte Carlo and Deterministic Methods*. *Oxford Statistical Science Series*. Oxford: Oxford Univ. Press. MR1859163

Friedman, J.H., Bentley, J.L. and Finkel, R.A. (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Software* **3** 209–226.

Gadat, S., Klein, T. and Marteau, C. (2016). Classification in general finite dimensional spaces with the k-nearest neighbor rule. *Ann. Statist.* **44** 982–1009. MR3485951 https://doi.org/10.1214/15-AOS1395

Genevay, A., Peyré, G. and Cuturi, M. (2018). Learning generative models with Sinkhorn divergences. In *Int. Conf. Artif. Intell. Stat.* 1608–1617. PMLR.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. *Applications of Mathematics (New York)* **53**. New York: Springer. MR1999614

Glynn, P.W. and Szechtman, R. (2002). Some new perspectives on the method of control variates. In *Monte Carlo and Quasi-Monte Carlo Methods, 2000 (Hong Kong)* 27–49. Berlin: Springer. MR1958845

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A.C. (2017). Improved training of Wasserstein GANs. In *Adv. Neural Inf. Process. Syst.* **30** 5769–5779.

Haber, S. (1966). A modified Monte-Carlo quadrature. *Math. Comp.* **20** 361–368. MR0210285 https://doi.org/10.2307/2003589

Haber, S. (1967). A modified Monte-Carlo quadrature. II. *Math. Comp.* **21** 388–397. MR0234606 https://doi.org/10.2307/2003241

Haber, S. (1969). Stochastic quadrature formulas. *Math. Comp.* **23** 751–764. MR0260139 https://doi.org/10.2307/2004961

Han, R. (2023). Sliced Wasserstein distance between probability measures on Hilbert spaces. ArXiv preprint. Available at arXiv:2307.05802.

Higdon, D., McDonnell, J.D., Schunck, N., Sarich, J. and Wild, S.M. (2015). A Bayesian approach for parameter estimation and prediction using a computationally intensive model. *J. Phys. G, Nucl. Part. Phys.* **42** 034009.

Hinrichs, A., Krieg, D., Novak, E., Prochno, J. and Ullrich, M. (2020). On the power of random information. In *Multivariate Algorithms and Information-Based Complexity*. *Radon Ser. Comput. Appl. Math.* **27** 43–64. Berlin: de Gruyter. MR4654505

Hoff III, K.E., Keyser, J., Lin, M., Manocha, D. and Culver, T. (1999). Fast computation of generalized Voronoi diagrams using graphics hardware. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques* 277–286.

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9** 90–95.

Ichida, R. (2000). The injectivity radius and the fundamental group of compact homogeneous Riemannian manifolds of positive curvature. *Tsukuba J. Math.* **24** 139–156. MR1791336 https://doi.org/10.21099/tkbjm/1496164051

Johnson, J., Douze, M. and Jégou, H. (2019). Billion-scale similarity search with GPUs. *IEEE Trans. Big Data* **7** 535–547.

Kohler, M. and Krzyżak, A. (2013). Optimal global rates of convergence for interpolation problems with random design. *Statist. Probab. Lett.* **83** 1871–1879. MR3069891 https://doi.org/10.1016/j.spl.2013.04.018

Kohler, M., Krzyżak, A. and Walk, H. (2006). Rates of convergence for partitioning and nearest neighbor regression estimates with unbounded data. *J. Multivariate Anal.* **97** 311–323. MR2234025 https://doi.org/10.1016/j.jmva.2005.03.006

Kokarev, G. (2021). Berger's inequality in the presence of upper sectional curvature bound. *Int. Math. Res. Not.* **2022** 11262–11303.

Kolouri, S., Nadjahi, K., Simsekli, U., Badeau, R. and Rohde, G. (2019). Generalized sliced Wasserstein distances. *Adv. Neural Inf. Process. Syst.* **32** 261–272.

Krieg, D. and Novak, E. (2017). A universal algorithm for multivariate integration. *Found. Comput. Math.* **17** 895–916. MR3682216 https://doi.org/10.1007/s10208-016-9307-y

Krieg, D., Novak, E. and Sonnleitner, M. (2022). Recovery of Sobolev functions restricted to iid sampling. *Math. Comp.* **91** 2715–2738. MR4473101 https://doi.org/10.1090/mcom/3763

Lee, J.M. (2019). *Introduction to Riemannian Manifolds*. *Graduate Texts in Mathematics*. Berlin: Springer.

Leluc, R., Dieuleveut, A., Portier, F., Segers, J. and Zhuman, A. (2024a). Sliced-Wasserstein estimation with spherical harmonics as control variates. ArXiv preprint. Available at arXiv:2402.01493.

Leluc, R., Portier, F. and Segers, J. (2021). Control variate selection for Monte Carlo integration. *Stat. Comput.* **31** 50. MR4277333 https://doi.org/10.1007/s11222-021-10011-z

Leluc, R., Portier, F., Segers, J. and Zhuman, A. (2024b). Supplement to "Speeding up Monte Carlo integration: Control neighbors for optimal convergence." https://doi.org/10.3150/24-BEJ1765SUPP

Mezzadri, F. (2007). How to generate random matrices from the classical compact groups. *Notices Amer. Math. Soc.* **54** 592–604. MR2311982

Nadjahi, K., Durmus, A., Jacob, P.E., Badeau, R. and Simsekli, U. (2021). Fast approximation of the sliced-Wasserstein distance using concentration of random projections. *Adv. Neural Inf. Process. Syst.* **34** 12411–12424.

Newton, N.J. (1994). Variance reduction for simulated diffusions. *SIAM J. Appl. Math.* **54** 1780–1805. MR1301282 https://doi.org/10.1137/S0036139992236220

Nguyen, K. and Ho, N. (2024). Sliced Wasserstein Estimation with Control Variates. ArXiv preprint. Available at arXiv:2305.00402.

Novak, E. (1988). *Deterministic and Stochastic Error Bounds in Numerical Analysis*. *Lecture Notes in Math.* **1349**. Berlin: Springer. MR0971255 https://doi.org/10.1007/BFb0079792

Novak, E. (2016). Some results on the complexity of numerical integration. In *Monte Carlo and Quasi-Monte Carlo Methods*. *Springer Proc. Math. Stat.* **163** 161–183. Springer. MR3536658 https://doi.org/10.1007/978-3-319-33507-0_6

Oates, C.J., Cockayne, J., Briol, F.-X. and Girolami, M. (2019). Convergence rates for a class of estimators based on Stein's method. *Bernoulli* **25** 1141–1159. MR3920368 https://doi.org/10.3150/17-bej1016

Oates, C.J., Girolami, M. and Chopin, N. (2017). Control functionals for Monte Carlo integration. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 695–718. MR3641403 https://doi.org/10.1111/rssb.12185

Omohundro, S.M. (1989). *Five Balltree Construction Algorithms*. International Computer Science Institute Berkeley.

Pastur, L. and Vasilchuk, V. (2004). On the moments of traces of matrices of classical groups. *Comm. Math. Phys.* **252** 149–166. MR2104877 https://doi.org/10.1007/s00220-004-1231-3

Pedregosa, F., Varoquaux, G., Gramfort, A. et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12** 2825–2830. MR2854348

Peyré, G., Cuturi, M. et al. (2019). Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** 355–607.

Portier, F. (2023). Nearest neighbor process: Weak convergence and non-asymptotic bound. ArXiv preprint. Available at arXiv:2110.15083.

Portier, F. and Segers, J. (2018). On the weak convergence of the empirical conditional copula under a simplifying assumption. *J. Multivariate Anal.* **166** 160–181. MR3799641 https://doi.org/10.1016/j.jmva.2018.03.002

Portier, F. and Segers, J. (2019). Monte Carlo integration with a growing number of control variates. *J. Appl. Probab.* **56** 1168–1186. MR4041454 https://doi.org/10.1017/jpr.2019.78

Rabin, J., Peyré, G., Delon, J. and Bernot, M. (2012). Wasserstein barycenter and its application to texture mixing. In *Scale Space and Variational Methods in Computer Vision: Third International Conference, SSVM 2011, Ein-Gedi, Israel, May 29–June 2, 2011, Revised Selected Papers 3* 435–446. Springer.

Richards, F.M. (1974). The interpretation of protein structures: Total volume, group volume distributions and packing density. *J. Mol. Biol.* **82** 1–14.

Rubinstein, R.Y. (1981). *Simulation and the Monte Carlo Method*. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR0624270

Rycroft, C. (2009). Voro++: A three-dimensional Voronoi cell library in C++ Technical Report Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989). Design and analysis of computer experiments. *Statist. Sci.* **4** 409–435. MR1041765

South, L.F., Oates, C.J., Mira, A. and Drovandi, C. (2023). Regularized zero-variance control variates. *Bayesian Anal.* **18** 865–888. MR4626360 https://doi.org/10.1214/22-ba1328

Stein, E.M. and Shakarchi, R. (2005). *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. *Princeton Lectures in Analysis* **3**. Princeton, NJ: Princeton Univ. Press. MR2129625

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *J. Roy. Statist. Soc. Ser. B* **36** 111–147. MR0356377

Toscano-Palmerin, S. and Frazier, P.I. (2022). Bayesian optimization with expensive integrands. *SIAM J. Optim.* **32** 417–444. MR4408265 https://doi.org/10.1137/19M1303125

Villani, C. (2009). *Optimal Transport: Old and New*. Berlin, Heidelberg: Springer.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nat. Methods* **17** 261–272.

Zimmermann, R. and Stoye, J. (2024). High curvature means low-rank: On the sectional curvature of Grassmann and Stiefel manifolds and the underlying matrix trace inequalities. ArXiv preprint. Available at arXiv:2403.01879.

Bernoulli **31**(2), 2025, 1181–1205
https://doi.org/10.3150/24-BEJ1766

Bayesian uncertainty quantification and structure detection for multiple change points models

EDUARD BELITSER[1,a] and SUBHASHIS GHOSAL[2,b]

[1]*Department of Mathematics, VU Amsterdam, Amsterdam, The Netherlands,* [a]*e.n.belitser@vu.nl*
[2]*Department of Statistics, North Carolina State University, Raleigh, USA,* [b]*sghosal@ncsu.edu*

Data observed over a long period of time may contain several change points, where the distribution of a variable changes but remains the same over the blocks in between. This useful qualitative structure allows precise estimation and uncertainty quantification for a long vector of parameters. Detecting these change points is another important objective. In this paper, we derive a concentration inequality for an empirical Bayes procedure, obtain the frequentist coverage of a suitable confidence ball of the optimal size constructed from the posterior distribution and study the problem of change point detection. We adopt an oracle approach to quantify the estimation error locally and show that the estimation error of the proposed procedure matches with the oracle rate, thus automatically implying minimax optimality, adaptively over all change point structures. Under a condition on the minimum magnitude of the changes, we show that precisely all change points are detected with high probability, and accompany this with a lower bound result asserting the minimality of that condition. Our results are non-asymptotic and robust in that normality is used only as a working model in the procedure, but the true distribution may not be normal. We discuss important extensions of our results to Hilbert space-valued parameters to address the multiple change point problem for multivariate and functional data. Finally, we describe a possible computational procedure using the simulated annealing method.

Keywords: Change points; confidence ball; empirical Bayes; oracle; posterior contraction

References

[1] Arias-Castro, E., Donoho, D.L. and Huo, X. (2005). Near-optimal detection of geometric objects by fast multiscale methods. *IEEE Trans. Inf. Theory* **51** 2402–2425.
[2] Aston, J. and Kirch, C. (2012). Evaluating stationarity via change-point alternatives with applications to fMRI data. *Ann. Appl. Stat.* **6** 1906–1948.
[3] Belitser, E. (2017). On coverage and local radial rates of credible sets. *Ann. Statist.* **45** 1124–1151.
[4] Belitser, E. and Ghosal, S. (2020). Empirical Bayes oracle uncertainty quantification for regression. *Ann. Statist.* **31** 536–559.
[5] Belitser, E. and Ghosal, S. (2024). Supplement to "Bayesian uncertainty quantification and structure detection for multiple change points models." https://doi.org/10.3150/24-BEJ1766SUPP
[6] Belitser, E. and Nurushev, N. (2019). General framework for projection structures. https://doi.org/10.48550/arXiv.1904.01003
[7] Belitser, E. and Nurushev, N. (2020). Needles and straw in a haystack: Robust confidence for possibly sparse sequences. *Bernoulli* **26** 191–225.
[8] Braun, J.V., Braun, R.K. and Müller, H.-G. (2000). Multiple changepoint fitting via quasilikelihood, with application to DNA sequence segmentation. *Biometrika* **87** 301–314.
[9] Carlin, B.P., Gelfand, A.E. and Smith, A.F. (1992). Hierarchical Bayesian analysis of changepoint problems. *Appl. Stat.* **41** 389–405. https://doi.org/10.2307/2347570
[10] Carlstein, E., Müller, H.-G. and Siegmund, D., eds. (1994). *Change-Point Problems*. *Institute of Mathematical Statistics Lecture Notes–Monograph Series* **23**. Hayward, CA: IMS.

1350-7265

[11] Chan, H.P. and Walther, G. (2013). Detection with the scan and the average likelihood ratio. *Statist. Sinica* **1** 409–428.
[12] Chan, N.H., Yau, C.Y. and Zhang, R.-M. (2014). Group LASSO for structural break time series. *J. Amer. Statist. Assoc.* **109** 590–599.
[13] Chen, J. and Gupta, A.K. (1997). Testing and locating variance changepoints with application to stock prices. *J. Amer. Statist. Assoc.* **92** 739–747.
[14] Chen, J. and Gupta, A.K. (2000). *Parametric Statistical Change-Point Analysis*. Boston, MA: Birkhäuser, Inc.
[15] Cho, H. and Kirch, C. (2022). Two-stage data segmentation permitting multiscale change points, heavy tails and dependence. *Ann. Inst. Statist. Math.* **74** 653–684.
[16] Cobb, G.W. (1978). The problem of the Nile: Conditional solution to a changepoint problem. *Biometrika* **65** 243–251.
[17] Conci, A. and Kubrusly, C. (2017). Distances between sets—a survey. *Adv. Math. Sci. Appl.* **26** 1–18.
[18] Eichinger, B. and Kirch, C. (2018). A mosum procedure for the estimation of multiple random change points. *Bernoulli* **24** 526–564.
[19] Fang, X., Li, J. and Siegmund, D. (2020). Segmentation and estimation of change-point models: False positive control and confidence regions. *Ann. Statist.* **48** 1615–1647.
[20] Frick, K., Munk, A. and Sieling, H. (2014). Multiscale change-point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 495–580.
[21] Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.* **42** 2243–2281.
[22] Gao, C., Han, F. and Zhang, C.-H. (2020). On estimation of isotonic piecewise constant signals. *Ann. Statist.* **48** 629–654.
[23] Hao, N., Niu, Y.S. and Zhang, H. (2013). Multiple change-point detection via a screening and ranking algorithm. *Statist. Sinica* **23** 1553–1572.
[24] Harchaoui, Z. and Lévy-Leduc, C. (2010). Multiple change-point estimation with a total variation penalty. *J. Amer. Statist. Assoc.* **105** 1480–1493.
[25] Huang, T., Wu, B., Lizardi, P. and Zhao, H. (2005). Detection of DNA copy number alterations using penalized least squares regression. *Bioinformatics* **21** 3811–3817. https://doi.org/10.1093/bioinformatics/bti646
[26] Lee, T.-S. (2010). Change-point problems: Bibliography and review. *J. Stat. Theory Pract.* **4** 643–662.
[27] Lin, K., Rinaldo, A., Sharpnack, J. and Tibshirani, R. (2017). *A sharp error analysis for the fused lasso, with application to approximate changepoint screening*. *Adv. Neurol. Inform. Process. Syst.* **30**.
[28] Niu, Y.S., Hao, N. and Zhang, H. (2016). Multiple change-point detection: A selective overview. *Statist. Sci.* **31** 611–623.
[29] Niu, Y.S. and Zhang, H. (2012). The screening and ranking algorithm to detect DNA copy number variations. *Ann. Appl. Stat.* **6** 1306–1326.
[30] Olshen, A.B., Venkatraman, E.S., Lucito, R. and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5** 557–572. https://doi.org/10.1093/biostatistics/kxh008
[31] Page, E. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42** 523–527.
[32] Rinaldo, A. (2009). Properties and refinements of the fused lasso. *Ann. Statist.* **37** 2922–2952.
[33] Tibshirani, R. and Wang, P. (2008). Spatial smoothing and hot spot detection for cgh data using the fused lasso. *Biostatistics* **9** 18–29. https://doi.org/10.1093/biostatistics/kxm013
[34] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series on Statistics*. New York, NY: Springer.
[35] Verzelen, N., Fromont, M., Lerasle, M. and Reynaud-Bouret, P. (2023). Optimal change-point detection and localization. *Ann. Statist.* **51** 1386–1610.
[36] Wang, D., Yu, Y. and Rinaldo, A. (2020). Univariate mean change-point detection: Penalization, CUSUM and optimality. *Electron. J. Stat.* **14** 1917–1961.
[37] Wang, T. and Samworth, R. (2018). High dimensional change-point estimation via sparse projection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 57–83.
[38] Yu, Y., Chatterjee, S. and Xu, H. (2022). Localising change points in piecewise polynomials of general degrees. *Electron. J. Stat.* **16** 1855–1890.

[39] Zhang, N.R. and Siegmund, D.O. (2007). A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data. *Biometrics* **63** 22–32.

Bernoulli **31**(2), 2025, 1206–1229
https://doi.org/10.3150/24-BEJ1767

High-dimensional variable selection with heterogeneous signals: A precise asymptotic perspective

SAPTARSHI ROY[1,a], AMBUJ TEWARI[1,b] and ZIWEI ZHU[2,c]

[1]*Department of Statistics, University of Michigan, Ann Arbor, USA,* [a]*roysapta@umich.edu,* [b]*tewaria@umich.edu*
[2]*Radix Trading, Chicago, USA,* [c]*zzw9348ustc@gmail.com*

We study the problem of exact support recovery for high-dimensional sparse linear regression under independent Gaussian design when the signals are weak, rare, and possibly heterogeneous. Under a suitable scaling of the sample size and signal sparsity, we fix the minimum signal magnitude at the information-theoretic optimal rate and investigate the asymptotic selection accuracy of best subset selection (BSS) and marginal screening (MS) procedures. We show that despite the ideal setup, somewhat surprisingly, marginal screening can fail to achieve exact recovery with probability converging to one in the presence of heterogeneous signals, whereas BSS enjoys model consistency whenever the minimum signal strength is above the information-theoretic threshold. To mitigate the computational intractability of BSS, we also propose an efficient two-stage algorithmic framework called ETS (Estimate Then Screen) comprised of an estimation step and gradient coordinate screening step, and under the same scaling assumption on sample size and sparsity, we show that ETS achieves model consistency under the same information-theoretic optimal requirement on the minimum signal strength as BSS. Finally, we present a simulation study comparing ETS with LASSO and marginal screening. The numerical results agree with our asymptotic theory even for realistic values of the sample size, dimension and sparsity.

Keywords: Heterogeneous signals; high-dimensional statistics; iterative hard thresholding; marginal screening; model consistency; variable selection

References

Aeron, S., Saligrama, V. and Zhao, M. (2010). Information theoretic bounds for compressed sensing. *IEEE Trans. Inf. Theory* **56** 5111–5130. MR2808668 https://doi.org/10.1109/TIT.2010.2059891

Agarwal, A., Negahban, S. and Wainwright, M.J. (2012). Fast global convergence of gradient methods for high-dimensional statistical recovery. *Ann. Statist.* **40** 2452–2482. MR3097609 https://doi.org/10.1214/12-AOS1032

Akçakaya, M. and Tarokh, V. (2010). Shannon-theoretic limits on noisy compressive sampling. *IEEE Trans. Inf. Theory* **56** 492–504. MR2589459 https://doi.org/10.1109/TIT.2009.2034796

Arias-Castro, E., Candès, E.J. and Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *Ann. Statist.* **39** 2533–2556. MR2906877 https://doi.org/10.1214/11-AOS910

Bertsimas, D., King, A. and Mazumder, R. (2016). Best subset selection via a modern optimization lens. *Ann. Statist.* **44** 813–852. MR3476618 https://doi.org/10.1214/15-AOS1388

Bertsimas, D. and Parys, B. (2020). Sparse high-dimensional regression: Exact scalable algorithms and phase transitions. *Ann. Statist.* **48** 300–323. MR4065163 https://doi.org/10.1214/18-AOS1804

Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. MR2533469 https://doi.org/10.1214/08-AOS620

Blumensath, T. and Davies, M.E. (2009). Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* **27** 265–274. MR2559726 https://doi.org/10.1016/j.acha.2009.04.002

Bogdan, M., van den Berg, E., Sabatti, C., Su, W. and Candès, E.J. (2015). SLOPE—adaptive variable selection via convex optimization. *Ann. Appl. Stat.* **9** 1103–1140. MR3418717 https://doi.org/10.1214/15-AOAS842

1350-7265

Candès, E.J., Romberg, J. and Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **52** 489–509. MR2236170 https://doi.org/10.1109/TIT.2005.862083

Candes, E.J. and Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inf. Theory* **52** 5406–5425. MR2300700 https://doi.org/10.1109/TIT.2006.885507

Consortium, W.T.C.C. et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447** 661–678. https://doi.org/10.1038/nature05911

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

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

Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 849–911. MR2530322 https://doi.org/10.1111/j.1467-9868.2008.00674.x

Fan, J., Shao, Q.-M. and Zhou, W.-X. (2018). Are discoveries spurious? Distributions of maximum spurious correlations and their applications. *Ann. Statist.* **46** 989–1017. MR3797994 https://doi.org/10.1214/17-AOS1575

Fletcher, A.K., Rangan, S. and Goyal, V.K. (2009). Necessary and sufficient conditions for sparsity pattern recovery. *IEEE Trans. Inf. Theory* **55** 5758–5772. MR2597192 https://doi.org/10.1109/TIT.2009.2032726

Genovese, C.R., Jin, J., Wasserman, L. and Yao, Z. (2012). A comparison of the lasso and marginal regression. *J. Mach. Learn. Res.* **13** 2107–2143. MR2956354

Guo, Y., Zhu, Z. and Fan, J. (2020). Best subset selection is robust against design dependence. arXiv preprint. Available at arXiv:2007.01478. https://doi.org/10.48550/arXiv.2007.01478

Guo, X., Zhang, H., Wang, Y. and Wu, J.-L. (2015). Model selection and estimation in high dimensional regression models with group SCAD. *Statist. Probab. Lett.* **103** 86–92. MR3350866 https://doi.org/10.1016/j.spl.2015.04.017

Hastie, T., Tibshirani, R. and Tibshirani, R. (2020). Best subset, forward stepwise or Lasso? Analysis and recommendations based on extensive comparisons. *Statist. Sci.* **35** 579–592. MR4175382 https://doi.org/10.1214/19-STS733

Hazimeh, H., Mazumder, R. and Saab, A. (2022). Sparse regression at scale: Branch-and-bound rooted in first-order optimization. *Math. Program.* **196** 347–388. MR4506808 https://doi.org/10.1007/s10107-021-01712-4

Hong, H.G., Kang, J. and Li, Y. (2018). Conditional screening for ultra-high dimensional covariates with survival outcomes. *Lifetime Data Anal.* **24** 45–71. MR3742906 https://doi.org/10.1007/s10985-016-9387-7

Huang, D., Li, R. and Wang, H. (2014). Feature screening for ultrahigh dimensional categorical data with applications. *J. Bus. Econom. Statist.* **32** 237–244. MR3207836 https://doi.org/10.1080/07350015.2013.863158

Huang, J., Ma, S. and Zhang, C.-H. (2008). Adaptive Lasso for sparse high-dimensional regression models. *Statist. Sinica* **18** 1603–1618. MR2469326

Huang, T.-J., McKeague, I.W. and Qian, M. (2019). Marginal screening for high-dimensional predictors of survival outcomes. *Statist. Sinica* **29** 2105–2139. MR3970349

Jain, P., Tewari, A. and Kar, P. (2014). On iterative hard thresholding methods for high-dimensional M-estimation. *Adv. Neural Inf. Process. Syst.* **27**.

Ji, P., Jin, J. et al. (2012). UPS delivers optimal phase diagram in high-dimensional variable selection. *Ann. Statist.* **40** 73–103. MR3013180 https://doi.org/10.1214/11-AOS947

Jin, Y., Kim, Y.-H. and Rao, B.D. (2011). Limits on support recovery of sparse signals via multiple-access communication techniques. *IEEE Trans. Inf. Theory* **57** 7877–7892. MR2895366 https://doi.org/10.1109/TIT.2011.2170116

Jin, J., Zhang, C.-H. and Zhang, Q. (2014). Optimality of graphlet screening in high dimensional variable selection. *J. Mach. Learn. Res.* **15** 2723–2772. MR3270749 https://doi.org/10.1631/jzus.a1400233

Kowshik, S.S. and Polyanskiy, Y. (2021). Fundamental limits of many-user MAC with finite payloads and fading. *IEEE Trans. Inf. Theory* **67** 5853–5884. MR4345042 https://doi.org/10.1109/tit.2021.3091423

Leisenring, W., Pepe, M.S. and Longton, G. (1997). A marginal regression modelling framework for evaluating medical diagnostic tests. *Stat. Med.* **16** 1263–1281. https://doi.org/10.1002/(SICI)1097-0258(19970615)16:11<1263::AID-SIM550>3.0.CO;2-M

Li, J., Zheng, Q., Peng, L. and Huang, Z. (2016). Survival impact index and ultrahigh-dimensional model-free screening with survival outcomes. *Biometrics* **72** 1145–1154. MR3591599 https://doi.org/10.1111/biom.12499

Li, Y., Hong, H.G., Ahmed, S.E. and Li, Y. (2019). Weak signals in high-dimensional regression: Detection, estimation and prediction. *Appl. Stoch. Models Bus. Ind.* **35** 283–298. MR3946127 https://doi.org/10.1002/asmb.2340

Li, X., Jiang, H., Haupt, J., Arora, R., Liu, H., Hong, M. and Zhao, T. (2020). On fast convergence of proximal algorithms for SQRT-Lasso optimization: Don't worry about its nonsmooth loss function. In *Proceedings of the 35th Uncertainty in Artificial Intelligence Conference*. *Proceedings of Machine Learning Research* **115** 49–59.

Liu, H. and Foygel Barber, R. (2020). Between hard and soft thresholding: Optimal iterative thresholding algorithms. *Inf. Inference* **9** 899–933. MR4188230 https://doi.org/10.1093/imaiai/iaz027

Lu, W. (2005). Marginal regression of multivariate event times based on linear transformation models. *Lifetime Data Anal.* **11** 389–404. MR2186652 https://doi.org/10.1007/s10985-005-2969-4

Marttinen, P., Gillberg, J., Havulinna, A., Corander, J. and Kaski, S. (2013). Genome-wide association studies with high-dimensional phenotypes. *Stat. Appl. Genet. Mol. Biol.* **12** 413–431. MR3101038 https://doi.org/10.1515/sagmb-2012-0032

Ndaoud, M. and Tsybakov, A.B. (2020). Optimal variable selection and adaptive noisy compressed sensing. *IEEE Trans. Inf. Theory* **66** 2517–2532. MR4087700 https://doi.org/10.1109/TIT.2020.2965738

Rad, K.R. (2011). Nearly sharp sufficient conditions on exact sparsity pattern recovery. *IEEE Trans. Inf. Theory* **57** 4672–4679. MR2840483 https://doi.org/10.1109/TIT.2011.2145670

Robins, J.M., Scheines, R., Spirtes, P. and Wasserman, L. (2003). Uniform consistency in causal inference. *Biometrika* **90** 491–515. MR2006831 https://doi.org/10.1093/biomet/90.3.491

Roy, S., Tewari, A. and Zhu, Z. (2024). Supplement to "High-dimensional variable selection with heterogeneous signals: A precise asymptotic perspective." https://doi.org/10.3150/24-BEJ1767SUPP

Scarlett, J. and Cevher, V. (2017). Limits on support recovery with probabilistic models: An information-theoretic framework. *IEEE Trans. Inf. Theory* **63** 593–620. MR3599962 https://doi.org/10.1109/TIT.2016.2606605

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

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

Tikhonov, A. (1943). On the stability of inverse problems. *Proc. USSR Acad. Sci.* **39** 195–198.

Van De Geer, S.A. and Bühlmann, P. (2009). On the conditions used to prove oracle results for the Lasso. *Electron. J. Stat.* **3** 1360–1392. MR2576316 https://doi.org/10.1214/09-EJS506

Wainwright, M.J. (2009a). 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

Wainwright, M.J. (2009b). Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting. *IEEE Trans. Inf. Theory* **55** 5728–5741. MR2597190 https://doi.org/10.1109/TIT.2009.2032816

Wang, W., Wainwright, M.J. and Ramchandran, K. (2010). Information-theoretic limits on sparse signal recovery: Dense versus sparse measurement matrices. *IEEE Trans. Inf. Theory* **56** 2967–2979. MR2683451 https://doi.org/10.1109/TIT.2010.2046199

Wang, H., Yang, Y. and Su, W.J. (2022). The price of competition: Effect size heterogeneity matters in high dimensions. *IEEE Trans. Inf. Theory* **68** 5268–5294. MR4476397 https://doi.org/10.1109/tit.2022.3166720

Wang, C.D., Chen, Z., Lian, Y. and Chen, M. (2022). Asset selection based on high frequency Sharpe ratio. *J. Econometrics* **227** 168–188. MR4377179 https://doi.org/10.1016/j.jeconom.2020.05.007

Wasserman, L. and Roeder, K. (2009). High-dimensional variable selection. *Ann. Statist.* **37** 2178–2201. MR2543689 https://doi.org/10.1214/08-AOS646

Xiao, L. and Zhang, T. (2013). A proximal-gradient homotopy method for the sparse least-squares problem. *SIAM J. Optim.* **23** 1062–1091. MR3057332 https://doi.org/10.1137/120869997

Xie, W. and Deng, X. (2020). Scalable algorithms for the sparse ridge regression. *SIAM J. Optim.* **30** 3359–3386. MR4188848 https://doi.org/10.1137/19M1245414

Yang, S., Wen, J., Eckert, S.T., Wang, Y., Liu, D.J., Wu, R., Li, R. and Zhan, X. (2020). Prioritizing genetic variants in GWAS with lasso using permutation-assisted tuning. *Bioinformatics* **36** 3811–3817. https://doi.org/10.1093/bioinformatics/btaa229

Zhang, C.-H. and Huang, J. (2008). The sparsity and bias of the LASSO selection in high-dimensional linear regression. *Ann. Statist.* **36** 1567–1594. MR2435448 https://doi.org/10.1214/07-AOS520

Zhang, Y., Wainwright, M.J. and Jordan, M.I. (2014). Lower bounds on the performance of polynomial-time algorithms for sparse linear regression. In *Conference on Learning Theory* 921–948.

Zhao, T., Liu, H. and Zhang, T. (2018). Pathwise coordinate optimization for sparse learning: Algorithm and theory. *Ann. Statist.* **46** 180–218. MR3766950 https://doi.org/10.1214/17-AOS1547

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

Zheng, Z., Fan, Y. and Lv, J. (2014). High dimensional thresholded regression and shrinkage effect. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 627–649. MR3210731 https://doi.org/10.1111/rssb.12037

Zhu, Z. and Wu, S. (2021). On the early solution path of best subset selection. arXiv preprint. Available at arXiv:2107.06939.

Zhu, J., Wen, C., Zhu, J., Zhang, H. and Wang, X. (2020). A polynomial algorithm for best-subset selection problem. *Proc. Natl. Acad. Sci. USA* **117** 33117–33123. MR4269239 https://doi.org/10.1073/pnas.2014241117

Bernoulli **31**(2), 2025, 1230–1255
https://doi.org/10.3150/24-BEJ1768

Multiple testing under negative dependence

ZIYU CHI[1,a], AADITYA RAMDAS[2,b] and RUODU WANG[3,c]

[1]*Department of IEOR, University of California, Berkeley, Berkeley CA, USA,* [a]*ziyu.chi@berkeley.edu*
[2]*Department of Statistics and Machine Learning, Carnegie Mellon University, Pittsburgh PA, USA,*
[b]*aramdas@cmu.edu*
[3]*Department of Statistics and Actuarial Science, University of Waterloo, Waterloo ON, Canada,*
[c]*wang@uwaterloo.ca*

The multiple testing literature has primarily dealt with three types of dependence assumptions between p-values: independence, positive regression dependence, and arbitrary dependence. In this paper, we provide what we believe are the first theoretical results under various notions of negative dependence (negative Gaussian dependence, negative regression dependence, negative association, negative orthant dependence and weak negative dependence). These include the Simes global null test and the Benjamini-Hochberg procedure, which are known experimentally to be anti-conservative under negative dependence. The anti-conservativeness of these procedures is bounded by factors smaller than that under arbitrary dependence (in particular, by factors independent of the number of hypotheses). We also provide new results about negatively dependent e-values, and provide several examples as to when negative dependence may arise. Our proofs are elementary and short, thus amenable to extensions.

Keywords: E-values; false discovery rate; negative association; negative orthant dependence; simes test

References

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* **57** 289–300. MR1325392

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

Blanchard, G., Neuvial, P. and Roquain, E. (2020). Post hoc confidence bounds on false positives using reference families. *Ann. Statist.* **48** 1281–1303. MR4124323 https://doi.org/10.1214/19-AOS1847

Block, H.W., Savits, T.H. and Shaked, M. (1982). Some concepts of negative dependence. *Ann. Probab.* **10** 765–772. MR0659545

Block, H.W., Savits, T.H. and Shaked, M. (1985). A concept of negative dependence using stochastic ordering. *Statist. Probab. Lett.* **3** 81–86. MR0792794 https://doi.org/10.1016/0167-7152(85)90029-X

Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

Chen, Y., Liu, P., Tan, K.S. and Wang, R. (2023). Trade-off between validity and efficiency of merging p-values under arbitrary dependence. *Statist. Sinica* **33** 851–872. MR4575325

Chi, Z., Ramdas, A. and Wang, R. (2024). Supplement to "Multiple testing under negative dependence." https://doi.org/10.3150/24-BEJ1768SUPP

Cui, X., Dickhaus, T., Ding, Y. and Hsu, J.C. (2021). *Handbook of Multiple Comparisons*. Boca Raton: CRC Press.

Delattre, S. and Roquain, E. (2016). On empirical distribution function of high-dimensional Gaussian vector components with an application to multiple testing. *Bernoulli* **22** 302–324. MR3449784 https://doi.org/10.3150/14-BEJ659

Finner, H., Roters, M. and Strassburger, K. (2017). On the Simes test under dependence. *Statist. Papers* **58** 775–789. MR3686850 https://doi.org/10.1007/s00362-015-0725-8

Gou, J. and Tamhane, A.C. (2018). Hochberg procedure under negative dependence. *Statist. Sinica* **28** 339–362. MR3752264

Hochberg, Y. and Rom, D. (1995). Extensions of multiple testing procedures based on Simes' test. *J. Statist. Plann. Inference* **48** 141–152. MR1366786 https://doi.org/10.1016/0378-3758(95)00005-T

Hommel, G. (1983). Tests of the overall hypothesis for arbitrary dependence structures. *Biom. J.* **25** 423–430. MR0735888

Howard, S.R., Ramdas, A., McAuliffe, J. and Sekhon, J. (2020). Time-uniform Chernoff bounds via nonnegative supermartingales. *Probab. Surv.* **17** 257–317. MR4100718 https://doi.org/10.1214/18-PS321

Huber, P.J. (1963). A remark on a paper of Trawinski and David entitled: "Selection of the best treatment in a paired-comparison experiment". *Ann. Math. Stat.* **34** 92–94. MR0144430 https://doi.org/10.1214/aoms/1177704244

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

Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11** 286–295. MR0684886 https://doi.org/10.1214/aos/1176346079

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467–498. MR0599685 https://doi.org/10.1016/0047-259X(80)90065-2

Kluger, D.M. and Owen, A.B. (2024). A central limit theorem for the Benjamini-Hochberg false discovery proportion under a factor model. *Bernoulli* **30** 743–769. MR4665596 https://doi.org/10.3150/23-bej1615

Lauzier, J.-G., Lin, L. and Wang, R. (2023). Pairwise counter-monotonicity. *Insurance Math. Econom.* **111** 279–287. MR4598917 https://doi.org/10.1016/j.insmatheco.2023.05.006

Lehmann, E.L. (1966). Some concepts of dependence. *Ann. Math. Stat.* **37** 1137–1153. MR0202228 https://doi.org/10.1214/aoms/1177699260

Malinovsky, Y. and Rinott, Y. (2023). On tournaments and negative dependence. *J. Appl. Probab.* **60** 945–954. MR4624050 https://doi.org/10.1017/jpr.2022.104

Müller, A. (1997). Stochastic orders generated by integrals: A unified study. *Adv. in Appl. Probab.* **29** 414–428. MR1450937 https://doi.org/10.2307/1428010

Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR1889865

Nelsen, R.B. (2007). *An Introduction to Copulas*. Berlin: Springer.

Puccetti, G. and Wang, R. (2015). Extremal dependence concepts. *Statist. Sci.* **30** 485–517. MR3432838 https://doi.org/10.1214/15-STS525

Ramdas, A.K., Barber, R.F., Wainwright, M.J. and Jordan, M.I. (2019). A unified treatment of multiple testing with prior knowledge using the p-filter. *Ann. Statist.* **47** 2790–2821. MR3988773 https://doi.org/10.1214/18-AOS1765

Reiner-Benaim, A. (2007). FDR control by the BH procedure for two-sided correlated tests with implications to gene expression data analysis. *Biom. J.* **49** 107–126. MR2339220 https://doi.org/10.1002/bimj.200510313

Roux, M. (2018). Inférence de graphes par une procédure de test multiple avec application en Neuroimagerie Ph.D. thesis Université Grenoble Alpes (ComUE).

Rüger, B. (1978). Das maximale Signifikanzniveau des Tests: "Lehne H_0 ab, wenn k unter n gegebenen Tests zur Ablehnung führen". *Metrika* **25** 171–178. MR0526476 https://doi.org/10.1007/BF02204362

Sarkar, T.K. (1969). *Some Lower Bounds of Reliability*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–Stanford University. MR2618992

Sarkar, S.K. (1998). Some probability inequalities for ordered MTP_2 random variables: A proof of the Simes conjecture. *Ann. Statist.* **26** 494–504. MR1626047 https://doi.org/10.1214/aos/1028144846

Shafer, G., Shen, A., Vereshchagin, N. and Vovk, V. (2011). Test martingales, Bayes factors and p-values. *Statist. Sci.* **26** 84–101. MR2849911 https://doi.org/10.1214/10-STS347

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

Shao, Q.-M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoret. Probab.* **13** 343–356. MR1777538 https://doi.org/10.1023/A:1007849609234

Simes, R.J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* **73** 751–754. MR0897872 https://doi.org/10.1093/biomet/73.3.751

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Syst. Tech. J.* **41** 463–501. MR0133183 https://doi.org/10.1002/j.1538-7305.1962.tb02419.x

Su, W.J. (2018). The FDR-linking theorem. arXiv preprint arXiv:1812.08965.

Tukey, J.W. (1953). The problem of multiple comparisons. Technical Report, Princeton University.

Vovk, V. and Wang, R. (2020a). Combining *p*-values via averaging. *Biometrika* **107** 791–808. MR4186488 https://doi.org/10.1093/biomet/asaa027

Vovk, V. and Wang, R. (2020b). True and false discoveries with independent e-values. arXiv preprint arXiv:2003.00593.

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

Vovk, V., Wang, B. and Wang, R. (2022). Admissible ways of merging p-values under arbitrary dependence. *Ann. Statist.* **50** 351–375. MR4382020 https://doi.org/10.1214/21-aos2109

Wang, R. and Ramdas, A. (2022). False discovery rate control with e-values. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 822–852. MR4460577

Bernoulli **31**(2), 2025, 1256–1276
https://doi.org/10.3150/24-BEJ1769

Stochastic selection problem for a Stratonovich SDE with power non-linearity

ILYA PAVLYUKEVICH[1,a] and GEORGIY SHEVCHENKO[2,b]

[1]*Institute of Mathematics, Friedrich Schiller University Jena, Jena, Germany,* [a]*ilya.pavlyukevich@uni-jena.de*
[2]*Kyiv School of Economics, Kyiv, Ukraine,* [b]*gshevchenko@kse.org.ua*

In our paper (*Bernoulli* **26** (2020) 1381–1409), we found all strong Markov solutions that spend zero time at 0 of the Stratonovich stochastic differential equation $\mathrm{d}X = |X|^\alpha \circ \mathrm{d}B$, $\alpha \in (0,1)$. These solutions have the form $X_t^\theta = F(B_t^\theta)$, where $F(x) = \frac{1}{1-\alpha}|x|^{1/(1-\alpha)} \operatorname{sign} x$ and B^θ is the skew Brownian motion with skewness parameter $\theta \in [-1,1]$ starting at $F^{-1}(X_0)$. In this paper we show how an addition of small external additive noise εW restores uniqueness. In the limit as $\varepsilon \to 0$, we recover heterogeneous diffusion corresponding to the physically symmetric case $\theta = 0$.

Keywords: Generalized Itô's formula; heterogeneous diffusion process; local time; non-uniqueness; selection problem; singular stochastic differential equation; skew Brownian motion; Stratonovich integral

References

Aït-Sahalia, Y. and Jacod, J. (2018). Semimartingale: Itô or not? *Stochastic Process. Appl.* **128** 233–254. MR3729537 https://doi.org/10.1016/j.spa.2017.04.006

Bafico, R. and Baldi, P. (1982). Small random perturbations of Peano phenomena. *Stochastics* **6** 279–292. MR0665404 https://doi.org/10.1080/17442508208833208

Cherny, A.S. (2001). Principal values of the integral functionals of Brownian motion: Existence, continuity and an extension of Itô's formula. In *Séminaire de Probabilités, XXXV*. *Lecture Notes in Math.* **1755** 348–370. Berlin: Springer. MR1837297 https://doi.org/10.1007/978-3-540-44671-2_24

Cherny, A.S. and Engelbert, H.-J. (2005). *Singular Stochastic Differential Equations*. *Lecture Notes in Math.* **1858**. Berlin: Springer. MR2112227 https://doi.org/10.1007/b104187

Cherstvy, A.G., Chechkin, A.V. and Metzler, R. (2013). Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes. *New J. Phys.* **15** 083039. MR3104237 https://doi.org/10.1088/1367-2630/15/8/083039

Delarue, F. and Flandoli, F. (2014). The transition point in the zero noise limit for a 1D Peano example. *Discrete Contin. Dyn. Syst.* **34** 4071–4083. MR3195359 https://doi.org/10.3934/dcds.2014.34.4071

Föllmer, H., Protter, P. and Shiryayev, A.N. (1995). Quadratic covariation and an extension of Itô's formula. *Bernoulli* **1** 149–169. MR1354459 https://doi.org/10.2307/3318684

Girsanov, I.V. (1962). An example of non-uniqueness of the solution of the stochastic equation of K. Ito. *Theory Probab. Appl.* **7** 325–331.

Jacod, J. and Protter, P. (2012). *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Heidelberg: Springer. MR2859096 https://doi.org/10.1007/978-3-642-24127-7

Krylov, N.V. (2009). *Controlled Diffusion Processes*. *Stochastic Modelling and Applied Probability* **14**. Berlin: Springer. MR2723141

Le Gall, J.-F. (1984). One-dimensional stochastic differential equations involving the local times of the unknown process. In *Stochastic Analysis and Applications (Swansea, 1983)*. *Lecture Notes in Math.* **1095** 51–82. Berlin: Springer. MR0777514 https://doi.org/10.1007/BFb0099122

Pavlyukevich, I. and Pilipenko, A. (2020). Generalized Peano problem with Lévy noise. *Electron. Commun. Probab.* **25** Paper No. 85. MR4195179 https://doi.org/10.1214/20-ecp365

1350-7265

Pavlyukevich, I. and Shevchenko, G. (2020). Stratonovich stochastic differential equation with irregular coefficients: Girsanov's example revisited. *Bernoulli* **26** 1381–1409. MR4058371 https://doi.org/10.3150/19-BEJ1161

Pilipenko, A. and Proske, F.N. (2018). On a selection problem for small noise perturbation in the multidimensional case. *Stoch. Dyn.* **18** 1850045. MR3869883 https://doi.org/10.1142/S0219493718500454

Protter, P.E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. *Applications of Mathematics (New York)* **21**. Berlin: Springer. MR2020294

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Berlin: Springer. MR1725357 https://doi.org/10.1007/978-3-662-06400-9

Sandev, T., Domazetoski, V., Kocarev, L., Metzler, R. and Chechkin, A. (2022). Heterogeneous diffusion with stochastic resetting. *J. Phys. A* **55** Paper No. 074003. MR4378847 https://doi.org/10.1088/1751-8121/ac491c

Trevisan, D. (2013). Zero noise limits using local times. *Electron. Commun. Probab.* **18** no. 31. MR3056068 https://doi.org/10.1214/ECP.v18-2587

Veretennikov, A.Y. (1983). Approximation of ordinary differential equations by stochastic differential equations. *Mat. Zametki* **33** 476–477.

Bernoulli **31**(2), 2025, 1277–1299
https://doi.org/10.3150/24-BEJ1770

Extremal shot noise processes and random cutout sets

CLÉMENT FOUCART[1,a] and LINGLONG YUAN[2,b]

[1]*CMAP, Ecole Polytechnique, Palaiseau, France and LAGA, Institut Galilée, Université Sorbonne Paris Nord, Villetaneuse, France,* [a]*foucart@math.univ-paris13.fr*
[2]*Department of Mathematical Sciences, University of Liverpool, Liverpool, United Kingdom,* [b]*linglong.yuan@liverpool.ac.uk*

We study some fundamental properties, such as the transience, the recurrence, the first passage times and the zero set of a certain type of sawtooth Markov processes, called extremal shot noise processes. The sets of zeros of the latter are Mandelbrot's random cutout sets, i.e. the sets obtained after placing Poisson random covering intervals on the positive half-line. Based on this connection, we provide a new proof of Fitzsimmons-Fristedt-Shepp Theorem which characterizes the random cutout sets.

Keywords: Extremal process; first passage time; invariant function; random covering; sawtooth process; shot noise process; subordinator

References

Barral, J. and Fan, A. (2004). Densities of some Poisson T-martingales and random covering numbers. *C. R. Math. Acad. Sci. Paris* **338** 571–574. MR2057032 https://doi.org/10.1016/j.crma.2004.01.027

Barral, J. and Fan, A. (2005). Covering numbers of different points in Dvoretzky covering. *Bull. Sci. Math.* **129** 275–317. MR2134123 https://doi.org/10.1016/j.bulsci.2004.05.007

Bertoin, J. (1991). Increase of a Lévy process with no positive jumps. *Stoch. Stoch. Rep.* **37** 247–251. MR1149349 https://doi.org/10.1080/17442509108833739

Bertoin, J. (1994). On nowhere differentiability for Lévy processes. *Stoch. Stoch. Rep.* **50** 205–210. MR1786116 https://doi.org/10.1080/17442509408833936

Bertoin, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press. MR1406564

Bertoin, J. (1999). Subordinators: Examples and applications. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*. (P. Bernard, ed.). *Lecture Notes in Math.* **1717** 1–91. Berlin: Springer. MR1746300 https://doi.org/10.1007/978-3-540-48115-7_1

Bi, H. and Delmas, J.-F. (2014). A population model with non-neutral mutations using branching processes with immigration. *Electron. J. Probab.* **19** 1–23. MR3238782 https://doi.org/10.1214/EJP.v19-2939

Blumenthal, R.M. (1992). *Excursions of Markov Processes*. *Probability and Its Applications*. Boston, MA: Birkhäuser, Inc. MR1138461 https://doi.org/10.1007/978-1-4684-9412-9

Blumenthal, R.M. and Getoor, R.K. (1964). Local times for Markov processes. *Z. Wahrsch. Verw. Gebiete* **3** 50–74. MR0165569 https://doi.org/10.1007/BF00531683

Böttcher, B., Schilling, R. and Wang, J. (2013). *Lévy Matters. III: Lévy-Type Processes: Construction, Approximation and Sample Path Properties*. *Lecture Notes in Math.* **2099**. Cham: Springer. MR3156646 https://doi.org/10.1007/978-3-319-02684-8

Boxma, O., Perry, D., Stadje, W. and Zacks, S. (2006). A Markovian growth-collapse model. *Adv. in Appl. Probab.* **38** 221–243. MR2213972 https://doi.org/10.1239/aap/1143936148

Davis, M.H.A. (1984). Piecewise-deterministic Markov processes: A general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B* **46** 353–388. MR0790622

1350-7265

Davis, M.H.A. (1993). *Markov Models and Optimization*. *Monographs on Statistics and Applied Probability* **49**. London: CRC Press. MR1283589 https://doi.org/10.1007/978-1-4899-4483-2

Dombry, C. (2012). Extremal shot noises, heavy tails and max-stable random fields. *Extremes* **15** 129–158. MR2915578 https://doi.org/10.1007/s10687-011-0131-0

Dwass, M. (1966). Extremal processes. II. *Illinois J. Math.* **10** 381–391. MR0193661

Evans, S.N. and Ralph, P.L. (2010). Dynamics of the time to the most recent common ancestor in a large branching population. *Ann. Appl. Probab.* **20** 1–25. MR2582640 https://doi.org/10.1214/09-AAP616

Fitzsimmons, P.J. (1981). Characterization and convergence of Markov sets. Ph.D. thesis, Ann Arbor, MI: ProQuest LLC. MR2631534

Fitzsimmons, P.J. (2022). Monotonicity properties of regenerative sets and Lorden's inequality. In *Dirichlet Forms and Related Topics*. *Springer Proc. Math. Stat.* **394** 109–117. Singapore: Springer. MR4508842 https://doi.org/10.1007/978-981-19-4672-1_7

Fitzsimmons, P.J., Fristedt, B. and Maisonneuve, B. (1985). Intersections and limits of regenerative sets. *Z. Wahrsch. Verw. Gebiete* **70** 157–173. MR0799144 https://doi.org/10.1007/BF02451426

Fitzsimmons, P.J., Fristedt, B. and Shepp, L.A. (1985). The set of real numbers left uncovered by random covering intervals. *Z. Wahrsch. Verw. Gebiete* **70** 175–189. MR0799145 https://doi.org/10.1007/BF02451427

Foucart, C. and Yuan, L. (2023). Weak convergence of continuous-state branching processes with large immigration. Preprint. Available at arXiv:2311.04045.

Foucart, C. and Uribe Bravo, G. (2014). Local extinction in continuous-state branching processes with immigration. *Bernoulli* **20** 1819–1844. MR3263091 https://doi.org/10.3150/13-BEJ543

Heinrich, L. and Molchanov, I.S. (1994). Some limit theorems for extremal and union shot-noise processes. *Math. Nachr.* **168** 139–159. MR1282636 https://doi.org/10.1002/mana.19941680109

Iksanov, A. and Kabluchko, Z. (2018). Functional limit theorems for Galton-Watson processes with very active immigration. *Stochastic Process. Appl.* **128** 291–305. MR3729539 https://doi.org/10.1016/j.spa.2017.04.012

Kahane, J.-P. (1990). Recouvrements aléatoires et théorie du potentiel. *Colloq. Math.* **60/61** 387–411. MR1096386 https://doi.org/10.4064/cm-60-61-2-287-411

Kahane, J.-P. (2000). Random coverings and multiplicative processes. In *Fractal Geometry and Stochastics, II. Proceedings of the 2nd Conference, Greifswald/ Koserow, Germany, August 28–September 2, 1998*. *Progress in Probability* **46** 125–146. Basel: Birkhäuser. MR1785624

Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications (New York)*. New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

Kella, O. and Stadje, W. (2001). On hitting times for compound Poisson dams with exponential jumps and linear release rate. *J. Appl. Probab.* **38** 781–786. MR1860216 https://doi.org/10.1017/s0021900200018945

Löpker, A. and Stadje, W. (2011). Hitting times and the running maximum of Markovian growth-collapse processes. *J. Appl. Probab.* **48** 295–312. MR2840300 https://doi.org/10.1239/jap/1308662628

Maisonneuve, B. (1971). Ensembles régénératifs, temps locaux et subordinateurs. In *Séminaire de Probabilités, V (Univ. Strasbourg, Année Universitaire 1969–1970)*. *Lecture Notes in Math., Vol. 191* 147–169. Berlin: Springer. MR0448586

Mandelbrot, B.B. (1972). Renewal sets and random cutouts. *Z. Wahrsch. Verw. Gebiete* **22** 145–157. MR0309162 https://doi.org/10.1007/BF00532733

Marchal, P. (2015). A class of special subordinators with nested ranges. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 533–544. MR3335014 https://doi.org/10.1214/13-AIHP595

Resnick, S.I. (1987). *Extreme Values, Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2364939

Resnick, S.I. and Rubinovitch, M. (1973). The structure of extremal processes. *Adv. in Appl. Probab.* **5** 287–307. MR0350867 https://doi.org/10.2307/1426037

Rivero, V.M. (2003). On random sets connected to the partial records of Poisson point process. *J. Theoret. Probab.* **16** 277–307. MR1956832 https://doi.org/10.1023/A:1022247025107

Roch, S. (2024). *Modern Discrete Probability. An Essential Toolkit*. Cambridge: Cambridge Univ. Press.

Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 1: Foundations*, 2nd ed. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press.

Serra, J. (1984). *Image Analysis and Mathematical Morphology*. London: Academic Press [Harcourt Brace Jovanovich, Publishers]. English version revised by Noel Cressie. MR0753649

Shepp, L.A. (1972). Covering the line with random intervals. *Z. Wahrsch. Verw. Gebiete* **23** 163–170. MR0322923 https://doi.org/10.1007/BF00536556

Bernoulli **31**(2), 2025, 1300–1324
https://doi.org/10.3150/24-BEJ1771

Siblings in d-dimensional nearest neighbour trees

JÉRÔME CASSE[a]

Laboratoire de mathématiques d'Orsay, Université Paris-Saclay, CNRS, Orsay, France,
[a]*jerome.casse.math@gmail.com*

Pick n ordered and independent uniform points on the d-dimensional flat torus. Then, link each point to its closest one that arrived before. This constructs random labelled trees called nearest neighbour trees of size n on the d-dimensional flat torus. These trees share some properties with the random recursive tree: the height of the last arrival node, the mean degree of the root, etc. On the contrary, the number of leaves seems to depend on dimension d, but no such properties have been proved yet. In this article, we prove that the mean number of siblings depends on d. In particular, we give explicit calculations of this number when n goes to infinity. In dimension 1, it is $1 + \ln 2$ and, in any dimension d, it has an explicit integral form, but unfortunately, it does not give an explicit number. Nevertheless, we show that it converges to 2 when $d \to \infty$ at a speed of order $(\sqrt{3}/2)^d$. To prove these results, we look at the local limit of those trees and we do some fine computations about the intersection of two balls in dimension d. In particular, we obtain a non-trivial upper bound for those intersections in some precise cases.

Keywords: High dimension; intersection of balls; local limit; nearest neighbour tree

References

[1] Abraham, R. and Delmas, J.-F. (2014). Local limits of conditioned Galton-Watson trees: The condensation case. *Electron. J. Probab.* **19** no. 56, 29. MR3227065 https://doi.org/10.1214/ejp.v19-3164

[2] Aldous, D. (2018). Random partitions of the plane via Poissonian coloring and a self-similar process of coalescing planar partitions. *Ann. Probab.* **46** 2000–2037. MR3813984 https://doi.org/10.1214/17-AOP1218

[3] Basdevant, A.-L., Blanc, G., Curien, N. and Singh, A. (2023). Fractal properties of the frontier in Poissonian coloring. *ALEA Lat. Am. J. Probab. Math. Stat.* **20** 1565–1586. MR4683386 https://doi.org/10.30757/alea.v20-59

[4] Baur, E. and Bertoin, J. (2014). Cutting edges at random in large recursive trees. In *Stochastic Analysis and Applications 2014*. *Springer Proc. Math. Stat.* **100** 51–76. Cham: Springer. MR3332709 https://doi.org/10.1007/978-3-319-11292-3_3

[5] Berger, N., Bollobás, B., Borgs, C., Chayes, J. and Riordan, O. (2003). Degree distribution of the FKP network model. In *Automata, Languages and Programming*. *Lecture Notes in Computer Science* **2719** 725–738. Berlin: Springer. MR2080740 https://doi.org/10.1007/3-540-45061-0_57

[6] Casse, J. (2024). Supplement to "Siblings in d-dimensional nearest neighbour trees." https://doi.org/10.3150/24-BEJ1771SUPP

[7] Curien, N. (2023). Random graphs: quodlibet. https://www.imo.universite-paris-saclay.fr/~nicolas.curien/enseignement.html.

[8] Devroye, L. (1987). Branching processes in the analysis of the heights of trees. *Acta Inform.* **24** 277–298. MR0894557 https://doi.org/10.1007/BF00265991

[9] Devroye, L. (1988). Applications of the theory of records in the study of random trees. *Acta Inform.* **26** 123–130. MR0969872 https://doi.org/10.1007/BF02915448

[10] Devroye, L., Fawzi, O. and Fraiman, N. (2010). The height of scaled attachment random recursive trees. In *21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10)*. *Discrete Math. Theor. Comput. Sci. Proc., AM* 129–141. Nancy: Assoc. Discrete Math. Theor. Comput. Sci. MR2735337 https://doi.org/10.46298/dmtcs.2788

1350-7265

[11] Devroye, L. and Lu, J. (1995). The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures Algorithms* **7** 1–14. MR1346281 https://doi.org/10.1002/rsa.3240070102
[12] Dobrow, R.P. (1996). On the distribution of distances in recursive trees. *J. Appl. Probab.* **33** 749–757. MR1401472 https://doi.org/10.1017/s002190020010018x
[13] Drmota, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Vienna: Springer. MR2484382 https://doi.org/10.1007/978-3-211-75357-6
[14] Fabrikant, A., Koutsoupias, E. and Papadimitriou, C.H. (2002). Heuristically optimized trade-offs: A new paradigm for power laws in the Internet. In *Automata, Languages and Programming*. *Lecture Notes in Computer Science* **2380** 110–122. Berlin: Springer. MR2062451 https://doi.org/10.1007/3-540-45465-9_11
[15] Frieze, A. and Karoński, M. (2023). *Random Graphs and Networks: A First Course*. Cambridge: Cambridge Univ. Press. MR4567944 https://doi.org/10.1017/9781009260268
[16] Goh, W. and Schmutz, E. (2002). Limit distribution for the maximum degree of a random recursive tree. *J. Comput. Appl. Math.* **142** 61–82. MR1910519 https://doi.org/10.1016/S0377-0427(01)00460-5
[17] Goldschmidt, C. and Martin, J.B. (2005). Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **10** 718–745. MR2164028 https://doi.org/10.1214/EJP.v10-265
[18] Jordan, J. (2010). Degree sequences of geometric preferential attachment graphs. *Adv. in Appl. Probab.* **42** 319–330. MR2675104 https://doi.org/10.1239/aap/1275055230
[19] Jordan, J. and Wade, A.R. (2015). Phase transitions for random geometric preferential attachment graphs. *Adv. in Appl. Probab.* **47** 565–588. MR3360390 https://doi.org/10.1239/aap/1435236988
[20] Lichev, L. and Mitsche, D. (2024). New results for the random nearest neighbor tree. *Probab. Theory Related Fields* **189** 229–279. MR4750559 https://doi.org/10.1007/s00440-024-01268-2
[21] Mahmoud, H.M. (1991). Limiting distributions for path lengths in recursive trees. *Probab. Engrg. Inform. Sci.* **5** 53–59. MR1183165 https://doi.org/10.1017/S0269964800001881
[22] Manna, S.S. and Sen, P. (2002). Modulated scale-free network in Euclidean space. *Phys. Rev. E* **66** 066114. https://doi.org/10.1103/PhysRevE.66.066114
[23] Pain, M. and Sénizergues, D. (2022). Correction terms for the height of weighted recursive trees. *Ann. Appl. Probab.* **32** 3027–3059. MR4474526 https://doi.org/10.1214/21-aap1756
[24] Penrose, M.D. (2005). Multivariate spatial central limit theorems with applications to percolation and spatial graphs. *Ann. Probab.* **33** 1945–1991. MR2165584 https://doi.org/10.1214/009117905000000206
[25] Penrose, M.D. and Wade, A.R. (2008). Limit theory for the random on-line nearest-neighbor graph. *Random Structures Algorithms* **32** 125–156. MR2387554 https://doi.org/10.1002/rsa.20185
[26] Pittel, B. (1994). Note on the heights of random recursive trees and random m-ary search trees. *Random Structures Algorithms* **5** 337–347. MR1262983 https://doi.org/10.1002/rsa.3240050207
[27] Preater, J. (2009). A species of voter model driven by immigration. *Statist. Probab. Lett.* **79** 2131–2137. MR2572044 https://doi.org/10.1016/j.spl.2009.07.003
[28] Smythe, R.T. and Mahmoud, H.M. (1994). A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.* **51** 1–29. MR1445048
[29] Steele, J.M. (1989). Cost of sequential connection for points in space. *Oper. Res. Lett.* **8** 137–142. MR1007269 https://doi.org/10.1016/0167-6377(89)90039-4
[30] Trauthwein, T. (2022). Quantitative CLTs on the Poisson space via Skorohod estimates and p-Poincaré inequalities. Preprint. Available at arXiv:2212.03782. https://doi.org/10.48550/arXiv.2212.03782
[31] Wade, A.R. (2009). Asymptotic theory for the multidimensional random on-line nearest-neighbour graph. *Stochastic Process. Appl.* **119** 1889–1911. MR2519349 https://doi.org/10.1016/j.spa.2008.09.006

Bernoulli **31**(2), 2025, 1325–1350
https://doi.org/10.3150/24-BEJ1772

A nonparametric distribution-free test of independence among continuous random vectors based on L_1-norm

NOUR-EDDINE BERRAHOU[1,a], SALIM BOUZEBDA[2,c] and LAHCEN DOUGE[1,b]

[1]*Laboratoire de Modélisation des Systèmes Complexes, University Cadi Ayyad, Marrakesh, Morocco,* [a]*n.berrahou@uca.ac.ma,* [b]*lahcen.douge@uca.ac.ma*
[2]*Université de technologie de Compiègne, LMAC (Laboratory of Applied Mathematics of Compiègne), France,* [c]*salim.bouzebda@utc.fr*

We propose a novel statistical test to assess the mutual independence of multidimensional random vectors. Our approach is based on the L_1-distance between the joint density function and the product of the marginal densities associated with the presumed independent vectors. Under the null hypothesis, we employ Poissonization techniques to establish the asymptotic normal approximation of the corresponding test statistic, without imposing any regularity assumptions on the underlying Lebesgue density function, denoted as $f(\cdot)$. Remarkably, we observe that the limiting distribution of the L_1-based statistics remains unaffected by the specific form of $f(\cdot)$. This unexpected result contributes to the robustness and versatility of our method. Moreover, our tests exhibit nontrivial local power against a subset of local alternatives, which converge to the null hypothesis at a rate of $n^{-1/2}h_n^{-d/4}$, $d \geq 2$, where n represents the sample size and h_n denotes the bandwidth. Finally, the theory is supported by a comprehensive simulation study to investigate the finite-sample performance of our proposed test. The results demonstrate that our testing procedure generally outperforms existing approaches across various examined scenarios.

Keywords: Asymptotic normality; distribution-free tests; independence test; kernel density function estimator; L_1-distance

References

Ahmad, I.A. and Li, Q. (1997). Testing independence by nonparametric kernel method. *Statist. Probab. Lett.* **34** 201–210. MR1457514 https://doi.org/10.1016/S0167-7152(96)00183-6

Akaike, H. (1954). An approximation to the density function. *Ann. Inst. Statist. Math., Tokyo* **6** 127–132. MR0067412 https://doi.org/10.1007/bf02900741

Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. *Wiley Series in Probability and Mathematical Statistics*. New York-Chichester-Brisbane: Wiley. MR0576407

Auddy, A., Deb, N. and Nandy, S. (2024). Exact detection thresholds and minimax optimality of Chatterjee's correlation coefficient. *Bernoulli* **30** 1640–1668. MR4699567 https://doi.org/10.3150/23-bej1648

Bach, F.R. and Jordan, M.I. (2003). Kernel independent component analysis. *J. Mach. Learn. Res.* **3** 1–48. MR1966051 https://doi.org/10.1162/153244303768966085

Beirlant, J. and Mason, D.M. (1995). On the asymptotic normality of L_p-norms of empirical functionals. *Math. Methods Statist.* **4** 1–19. MR1324687

Bergsma, W. and Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20** 1006–1028. MR3178526 https://doi.org/10.3150/13-BEJ514

Berrahou, N.-E., Bouzebda, S. and Douge, L. (2024). Supplement to "A nonparametric distribution-free test of independence among continuous random vectors based on L_1-norm." https://doi.org/10.3150/24-BEJ1772SUPP

Berrett, T.B. (2022). Discussion of 'Multi-scale Fisher's independence test for multivariate dependence' [4472834]. *Biometrika* **109** 589–592. MR4472835 https://doi.org/10.1093/biomet/asac023

Berrett, T.B., Grose, D.J. and Samworth, R.J. (2018). IndepTest: *Nonparametric Independence Tests Based on Entropy Estimation*. R Package Version 0.2.0. Available at https://cran.r-project.org/web/Packages/IndepTest/index.html.

Berrett, T.B., Kontoyiannis, I. and Samworth, R.J. (2021). Optimal rates for independence testing via U-statistic permutation tests. *Ann. Statist.* **49** 2457–2490. MR4338371 https://doi.org/10.1214/20-aos2041

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

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

Binet, A. and Vaschide, N. (1897). Corrélation des épreuves physiques. *L'Année Psychologique* **4** 142–172. https://doi.org/10.3406/psy.1897.2892

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

Böttcher, B., Keller-Ressel, M. and Schilling, R.L. (2019). Distance multivariance: New dependence measures for random vectors. *Ann. Statist.* **47** 2757–2789. MR3988772 https://doi.org/10.1214/18-AOS1764

Bouzebda, S. (2011). Some new multivariate tests of independence. *Math. Methods Statist.* **20** 192–205. MR2908757 https://doi.org/10.3103/S1066530711030021

Bouzebda, S. (2014). General tests of independence based on empirical processes indexed by functions. *Stat. Methodol.* **21** 59–87. MR3229031 https://doi.org/10.1016/j.stamet.2014.03.001

Bouzebda, S. and Nemouchi, B. (2023). Weak-convergence of empirical conditional processes and conditional U-processes involving functional mixing data. *Stat. Inference Stoch. Process.* **26** 33–88. MR4562252 https://doi.org/10.1007/s11203-022-09276-6

Bouzebda, S. and Taachouche, N. (2023). On the variable bandwidth kernel estimation of conditional U-statistics at optimal rates in sup-norm. *Phys. A* **625** Paper No. 129000, 72. MR4612299 https://doi.org/10.1016/j.physa.2023.129000

Chakraborty, S. and Zhang, X. (2019). Distance metrics for measuring joint dependence with application to causal inference. *J. Amer. Statist. Assoc.* **114** 1638–1650. MR4047289 https://doi.org/10.1080/01621459.2018.1513364

Chan, N.-H., Lee, T.C.M. and Peng, L. (2010). On nonparametric local inference for density estimation. *Comput. Statist. Data Anal.* **54** 509–515. MR2756443 https://doi.org/10.1016/j.csda.2009.09.021

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

Chen, A. and Bickel, P.J. (2006). Efficient independent component analysis. *Ann. Statist.* **34** 2825–2855. MR2329469 https://doi.org/10.1214/009053606000000939

Cox, T.F. and Czanner, G. (2016). A practical divergence measure for survival distributions that can be estimated from Kaplan-Meier curves. *Stat. Med.* **35** 2406–2421. MR3513521 https://doi.org/10.1002/sim.6868

Csörgő, S. (1985). Testing for independence by the empirical characteristic function. *J. Multivariate Anal.* **16** 290–299. MR0793494 https://doi.org/10.1016/0047-259X(85)90022-3

Csörgő, M. and Horváth, L. (1988). Central limit theorems for L_p-norms of density estimators. *Probab. Theory Related Fields* **80** 269–291. MR0968821 https://doi.org/10.1007/BF00356106

Deb, N., Ghosal, P. and Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. Available at arXiv:2010.01768.

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

Deheuvels, P. (1977). Estimation non paramétrique de la densité par histogrammes généralisés. II. *Publ. Inst. Statist. Univ. Paris* **22** 1–23. MR0539513

Deheuvels, P. (1981). A Kolmogorov-Smirnov type test for independence and multivariate samples. *Rev. Roumaine Math. Pures Appl.* **26** 213–226. MR0616038

Devroye, L. (1987). *A Course in Density Estimation*. *Progress in Probability and Statistics* **14**. Boston, MA: Birkhäuser, Inc. MR0891874

Devroye, L. (1989). The double kernel method in density estimation. *Ann. Inst. Henri Poincaré Probab. Stat.* **25** 533–580. MR1045250

Devroye, L. (1997). Universal smoothing factor selection in density estimation: Theory and practice. *Test* **6** 223–320. MR1616896 https://doi.org/10.1007/BF02564701

Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View*. *Wiley Series in Probability and Mathematical Statistics: Tracts on Probability and Statistics*. New York: Wiley. MR0780746

Devroye, L., Györfi, L. and Lugosi, G. (2002). A note on robust hypothesis testing. *IEEE Trans. Inf. Theory* **48** 2111–2114. MR1930019 https://doi.org/10.1109/TIT.2002.1013154

Devroye, L. and Lugosi, G. (2001). *Combinatorial Methods in Density Estimation*. *Springer Series in Statistics*. New York: Springer. MR1843146 https://doi.org/10.1007/978-1-4613-0125-7

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

Dugué, D. (1975). Sur des tests d'indépendance "indépendants de la loi". *C. R. Acad. Sci. Paris Sér. A-B* **281** A1103–A1104. MR0395029

Duong, T. and Hazelton, M.L. (2005). Cross-validation bandwidth matrices for multivariate kernel density estimation. *Scand. J. Stat.* **32** 485–506. MR2204631 https://doi.org/10.1111/j.1467-9469.2005.00445.x

Eggermont, P.P.B. and LaRiccia, V.N. (2001). *Maximum Penalized Likelihood Estimation. Vol. I: Density Estimation*. *Springer Series in Statistics*. New York: Springer. MR1837879

Epsnečnikov, V.A. (1969). Nonparametric estimation of a multidimensional probability density. *Teor. Veroyatn. Primen.* **14** 156–162. MR0250422

Fan, Y., Lafaye de Micheaux, P., Penev, S. and Salopek, D. (2017). Multivariate nonparametric test of independence. *J. Multivariate Anal.* **153** 189–210. MR3578846 https://doi.org/10.1016/j.jmva.2016.09.014

Felber, T., Kohler, M. and Krzyżak, A. (2015). Adaptive density estimation from data with small measurement errors. *IEEE Trans. Inf. Theory* **61** 3446–3456. MR3352508 https://doi.org/10.1109/TIT.2015.2421297

Feuerverger, A. (1993). A consistent test for bivariate dependence. *Int. Stat. Rev.* **61** 419–433. MR4205649

Fukumizu, K., Bach, F.R. and Jordan, M.I. (2004). Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces. *J. Mach. Learn. Res.* **5** 73–99. MR2247974 https://doi.org/10.1162/153244303768966111

Gan, L., Narisetty, N.N. and Liang, F. (2019). Bayesian regularization for graphical models with unequal shrinkage. *J. Amer. Statist. Assoc.* **114** 1218–1231. MR4011774 https://doi.org/10.1080/01621459.2018.1482755

Gao, J. and Gijbels, I. (2008). Bandwidth selection in nonparametric kernel testing. *J. Amer. Statist. Assoc.* **103** 1584–1594. MR2504206 https://doi.org/10.1198/016214508000000968

Giné, E., Mason, D.M. and Zaitsev, A.Yu. (2003). The L_1-norm density estimator process. *Ann. Probab.* **31** 719–768. MR1964947 https://doi.org/10.1214/aop/1048516534

Gorsky, S. and Ma, L. (2022). Multi-scale Fisher's independence test for multivariate dependence. *Biometrika* **109** 569–587. MR4472834 https://doi.org/10.1093/biomet/asac013

Gretton, A. and Györfi, L. (2010). Consistent nonparametric tests of independence. *J. Mach. Learn. Res.* **11** 1391–1423. MR2645456

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

Hahn, T. (2005). Cuba—a library for multidimensional numerical integration. *Comput. Phys. Commun.* **168** 78–95. MR2136794 https://doi.org/10.1016/j.cpc.2005.01.010

Hall, P. and Wand, M.P. (1988). Minimizing L_1 distance in nonparametric density estimation. *J. Multivariate Anal.* **26** 59–88. MR0955204 https://doi.org/10.1016/0047-259X(88)90073-5

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

Härdle, W. (1990). *Applied Nonparametric Regression*. *Econometric Society Monographs* **19**. Cambridge: Cambridge Univ. Press. MR1161622 https://doi.org/10.1017/CCOL0521382483

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

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

Horowitz, J.L. and Spokoiny, V.G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica* **69** 599–631. MR1828537 https://doi.org/10.1111/1468-0262.00207

Horváth, L. (1991). On L_p-norms of multivariate density estimators. *Ann. Statist.* **19** 1933–1949. MR1135157 https://doi.org/10.1214/aos/1176348379

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

Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93. https://doi.org/10.1093/biomet/30.1-2.81

Lauritzen, S.L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. New York: Oxford University Press. MR1419991 https://doi.org/10.1093/oso/9780198522195.001.0001

Le Cam, L. (1970). Remarques sur le théorème limite central dans les espaces localement convexes. In *Les Probabilités sur les Structures Algébriques (Actes Colloq. Internat. CNRS, No. 186, Clermont-Ferrand, 1969)* 233–249. Paris: Éditions du Centre National de la Recherche Scientifique. MR0410832

Lee, S., Song, K. and Whang, Y.-J. (2013). Testing functional inequalities. *J. Econometrics* **172** 14–32. MR2997127 https://doi.org/10.1016/j.jeconom.2012.08.006

Leung, D. and Drton, M. (2018). Testing independence in high dimensions with sums of rank correlations. *Ann. Statist.* **46** 280–307. MR3766953 https://doi.org/10.1214/17-AOS1550

Matteson, D.S. and Tsay, R.S. (2017). Independent component analysis via distance covariance. *J. Amer. Statist. Assoc.* **112** 623–637. MR3671757 https://doi.org/10.1080/01621459.2016.1150851

Moon, K.R. (2016). Nonparametric Estimation of Distributional Functionals and Applications. Thesis (Ph.D.)—University of Michigan.

Nabeya, S. (1951). Absolute moments in 2-dimensional normal distribution. *Ann. Inst. Statist. Math.* **3** 2–6. MR0045347 https://doi.org/10.1007/bf02949770

Nadaraya, È.A. (1989). *Nonparametric Estimation of Probability Densities and Regression Curves*. *Mathematics and Its Applications (Soviet Series)* **20**. Dordrecht: Kluwer Academic. MR1093466 https://doi.org/10.1007/978-94-009-2583-0

Nguyen, D. and Eisenstein, J. (2017). A kernel independence test for geographical language variation. *Comput. Linguist.* **43** 567–592. MR3708620 https://doi.org/10.1162/COLI_a_00293

Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Stat.* **33** 1065–1076. MR0143282 https://doi.org/10.1214/aoms/1177704472

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2548166 https://doi.org/10.1017/CBO9780511803161

Pearson, K. (1920). Notes on the history of correlation. *Biometrika* **13** 25–45. https://doi.org/10.2307/2331722

Peters, J., Mooij, J.M., Janzing, D. and Schölkopf, B. (2014). Causal discovery with continuous additive noise models. *J. Mach. Learn. Res.* **15** 2009–2053. MR3231600

Pfister, N. and Peters, J. (2017). dHSIC: *Independence Testing via Hilbert-Schmidt Independence Criterion*. R Package Version 2.1. Available at https://cran.r-project.org/web/packages/dHSIC/index.html.

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

Pinelis, I. (1994). On a majorization inequality for sums of independent random vectors. *Statist. Probab. Lett.* **19** 97–99. MR1256696 https://doi.org/10.1016/0167-7152(94)90139-2

Qiu, T., Xu, W. and Zhu, L. (2023). Independence tests with random subspace of two random vectors in high dimension. *J. Multivariate Anal.* **195** Paper No. 105160, 22. MR4544605 https://doi.org/10.1016/j.jmva.2023.105160

Rizzo, M.L. (2009). New goodness-of-fit tests for Pareto distributions. *Astin Bull.* **39** 691–715. MR2751846 https://doi.org/10.2143/AST.39.2.2044654

Rizzo, M. and Székely, G. (2016). E-Statistics: Multivariate inference via the energy of data. R package version 1.7-0. Available at https://cran.r-project.org/web/packages/energy/index.html.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* **27** 832–837. MR0079873 https://doi.org/10.1214/aoms/1177728190

Rosenblatt, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.* **3** 1–14. MR0428579

Rosenblatt, M. and Wahlen, B.E. (1992). A nonparametric measure of independence under a hypothesis of independent components. *Statist. Probab. Lett.* **15** 245–252. MR1190260 https://doi.org/10.1016/0167-7152(92)90197-D

Sain, S.R., Baggerly, K.A. and Scott, D.W. (1994). Cross-validation of multivariate densities. *J. Amer. Statist. Assoc.* **89** 807–817. MR1294726

Samworth, R.J. and Yuan, M. (2012). Independent component analysis via nonparametric maximum likelihood estimation. *Ann. Statist.* **40** 2973–3002. MR3097966 https://doi.org/10.1214/12-AOS1060

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1191168 https://doi.org/10.1002/9780470316849

Scott, D.W. and Wand, M.P. (1991). Feasibility of multivariate density estimates. *Biometrika* **78** 197–205. MR1118245 https://doi.org/10.1093/biomet/78.1.197

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

Sheng, W. and Yin, X. (2016). Sufficient dimension reduction via distance covariance. *J. Comput. Graph. Statist.* **25** 91–104. MR3474038 https://doi.org/10.1080/10618600.2015.1026601

Shergin, V.V. (1990). The central limit theorem for finitely dependent random variables. In *Probability Theory and Mathematical Statistics, Vol. II (Vilnius, 1989)* 424–431. Vilnius: "Mokslas". MR1153895

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

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

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9

Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15** 72–101. https://doi.org/10.2307/1422689

Sriperumbudur, B.K., Fukumizu, K., Gretton, A., Schölkopf, B. and Lanckriet, G.R.G. (2012). On the empirical estimation of integral probability metrics. *Electron. J. Stat.* **6** 1550–1599. MR2988458 https://doi.org/10.1214/12-EJS722

Stein, E.M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series, No. 30*. Princeton, NJ: Princeton Univ. Press. MR0290095

Sweeting, T.J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** 28–41. MR0428400 https://doi.org/10.1214/aop/1176995888

Székely, G.J. and Rizzo, M.L. (2005). Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method. *J. Classification* **22** 151–183. MR2231170 https://doi.org/10.1007/s00357-005-0012-9

Székely, G.J. and Rizzo, M.L. (2009). Brownian distance covariance. *Ann. Appl. Stat.* **3** 1236–1265. https://doi.org/10.1214/09-AOAS312

Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. MR2382665 https://doi.org/10.1214/009053607000000505

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

Watson, G.S. and Leadbetter, M.R. (1963). On the estimation of the probability density. I. *Ann. Math. Stat.* **34** 480–491. https://doi.org/10.1214/aoms/1177704159

Yang, G. (2012). *The Energy Goodness-of-Fit Test for Univariate Stable Distributions*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)—Bowling Green State University.

Yatracos, Y.G. (1985). Rates of convergence of minimum distance estimators and Kolmogorov's entropy. *Ann. Statist.* **13** 768–774. https://doi.org/10.1214/aos/1176349553

Zhang, W., Gao, W. and Ng, H.K.T. (2023). Multivariate tests of independence based on a new class of measures of independence in reproducing kernel Hilbert space. *J. Multivariate Anal.* **195** Paper No. 105144. https://doi.org/10.1016/j.jmva.2022.105144

Zhang, Q., Filippi, S., Gretton, A. and Sejdinovic, D. (2018). Large-scale kernel methods for independence testing. *Stat. Comput.* **28** 113–130. https://doi.org/10.1007/s11222-016-9721-7

Bernoulli **31**(2), 2025, 1351–1376
https://doi.org/10.3150/24-BEJ1773

Optimal transport and Wasserstein distances for causal models

PATRICK CHERIDITO[1,a] and STEPHAN ECKSTEIN[2,b]

[1]*Department of Mathematics, ETH Zurich, Zurich, Switzerland,* [a]*patrick.cheridito@math.ethz.ch*
[2]*Department of Mathematics, University of Tübingen, Tübingen, Germany,* [b]*stephan.eckstein@uni-tuebingen.de*

In this paper, we introduce a variant of optimal transport adapted to the causal structure given by an underlying directed graph G. Different graph structures lead to different specifications of the optimal transport problem. For instance, a fully connected graph yields standard optimal transport, a linear graph structure corresponds to causal optimal transport between the distributions of two discrete-time stochastic processes, and an empty graph leads to a notion of optimal transport related to CO-OT, Gromov–Wasserstein distances and factored OT. We derive different characterizations of G-causal transport plans and introduce Wasserstein distances between causal models that respect the underlying graph structure. We show that average treatment effects are continuous with respect to G-causal Wasserstein distances and small perturbations of structural causal models lead to small deviations in G-causal Wasserstein distance. We also introduce an interpolation between causal models based on G-causal Wasserstein distance and compare it to standard Wasserstein interpolation.

Keywords: Average treatment effect; causality; directed graphs; optimal transport; Wasserstein distance

References

[1] Acciaio, B., Backhoff-Veraguas, J. and Jia, J. (2021). Cournot-Nash equilibrium and optimal transport in a dynamic setting. *SIAM J. Control Optim.* **59** 2273–2300. MR4274834 https://doi.org/10.1137/20M1321462
[2] Agueh, M. and Carlier, G. (2011). Barycenters in the Wasserstein space. *SIAM J. Math. Anal.* **43** 904–924. MR2801182 https://doi.org/10.1137/100805741
[3] Aldous, D. (1981). Weak convergence and the general theory of processes. Incomplete draft.
[4] Arjovsky, M., Chintala, S. and Bottou, L. (2017). Wasserstein generative adversarial networks. In *International Conference on Machine Learning* 214–223. PMLR.
[5] Backhoff-Veraguas, J., Bartl, D., Beiglböck, M. and Eder, M. (2020). All adapted topologies are equal. *Probab. Theory Related Fields* **178** 1125–1172. MR4168395 https://doi.org/10.1007/s00440-020-00993-8
[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. and Pammer, G. (2022). Stability of martingale optimal transport and weak optimal transport. *Ann. Appl. Probab.* **32** 721–752. MR4386541 https://doi.org/10.1214/21-aap1694
[8] Bartl, D., Beiglböck, M. and Pammer, G. (2021). The Wasserstein space of stochastic processes. arXiv Preprint arXiv:2104.14245.
[9] Beiglböck, M. and Juillet, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. MR3456332 https://doi.org/10.1214/14-AOP966
[10] Blanchet, J. and Murthy, K. (2019). Quantifying distributional model risk via optimal transport. *Math. Oper. Res.* **44** 565–600. MR3959085 https://doi.org/10.1287/moor.2018.0936
[11] Bongers, S., Forré, P., Peters, J. and Mooij, J.M. (2021). Foundations of structural causal models with cycles and latent variables. *Ann. Statist.* **49** 2885–2915. MR4338897 https://doi.org/10.1214/21-aos2064
[12] Bonneel, N. and Coeurjolly, D. (2019). Spot: Sliced partial optimal transport. *ACM Trans. Graph.* **38** 1–13.
[13] Bourbaki, N. (1966). *Elements of Mathematics: General Topology, Part 2*. *Actualités Scientifiques et Industrielles*. Hermann.

1350-7265

[14] Cheridito, P., Kiiski, M., Prömel, D.J. and Soner, H.M. (2021). Martingale optimal transport duality. *Math. Ann.* **379** 1685–1712. MR4238277 https://doi.org/10.1007/s00208-019-01952-y
[15] Cheridito, P., Kupper, M. and Tangpi, L. (2017). Duality formulas for robust pricing and hedging in discrete time. *SIAM J. Financial Math.* **8** 738–765. MR3705784 https://doi.org/10.1137/16M1064088
[16] Chiappori, P.-A., McCann, R.J. and Nesheim, L.P. (2010). Hedonic price equilibria, stable matching, and optimal transport: Equivalence, topology, and uniqueness. *Econom. Theory* **42** 317–354. MR2564439 https://doi.org/10.1007/s00199-009-0455-z
[17] Chizat, L., Peyré, G., Schmitzer, B. and Vialard, F.-X. (2018). Unbalanced optimal transport: Dynamic and Kantorovich formulations. *J. Funct. Anal.* **274** 3090–3123. MR3782988 https://doi.org/10.1016/j.jfa.2018.03.008
[18] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. *Adv. Neural Inf. Process. Syst.* **26**.
[19] Cuturi, M. and Doucet, A. (2014). Fast computation of Wasserstein barycenters. In *International Conference on Machine Learning* 685–693. PMLR.
[20] De Gennaro Aquino, L. and Eckstein, S. (2020). Minmax methods for optimal transport and beyond: Regularization, approximation and numerics. *Adv. Neural Inf. Process. Syst.* **33** 13818–13830.
[21] De Lara, L., González-Sanz, A., Asher, N. and Loubes, J.-M. (2021). Transport-based Counterfactual Models. arXiv Preprint arXiv:2108.13025.
[22] De Philippis, G. and Figalli, A. (2014). The Monge–Ampère equation and its link to optimal transportation. *Bull. Amer. Math. Soc. (N.S.)* **51** 527–580. MR3237759 https://doi.org/10.1090/S0273-0979-2014-01459-4
[23] Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *J. Bank. Finance* **37** 2750–2764.
[24] Figalli, A., Maggi, F. and Pratelli, A. (2010). A mass transportation approach to quantitative isoperimetric inequalities. *Invent. Math.* **182** 167–211. MR2672283 https://doi.org/10.1007/s00222-010-0261-z
[25] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7
[26] Galichon, A. (2016). *Optimal Transport Methods in Economics*. Princeton, NJ: Princeton Univ. Press. MR3586373 https://doi.org/10.1515/9781400883592
[27] Ghosal, P. and Sen, B. (2022). Multivariate ranks and quantiles using optimal transport: Consistency, rates and nonparametric testing. *Ann. Statist.* **50** 1012–1037. MR4404927 https://doi.org/10.1214/21-aos2136
[28] Gurobi Optimization, LLC (2022). Gurobi Optimizer Reference Manual.
[29] Hellwig, M.F. (1996). Sequential decisions under uncertainty and the maximum theorem. *J. Math. Econom.* **25** 443–464. MR1394333 https://doi.org/10.1016/0304-4068(95)00739-3
[30] Holland, P.W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc.* **81** 945–970. With discussion and a reply by the author. MR0867618
[31] Huang, Y., Leung, C.H., Wu, Q., Yan, X., Ma, S., Yuan, Z., Wang, D. and Huang, Z. (2022). Robust causal learning for the estimation of average treatment effects. In *2022 International Joint Conference on Neural Networks (IJCNN 2022)*. IEEE.
[32] Kallenberg, O. (2021). *Foundations of Modern Probability*. *Probability Theory and Stochastic Modelling* **99**. Cham: Springer. Third edition [of 1464694]. MR4226142 https://doi.org/10.1007/978-3-030-61871-1
[33] Kantorovich, L.V. (1942). On the translocation of masses. *C. R. (Dokl.) Acad. Sci. URSS* **37** 199–201. MR0009619
[34] Kantorovich, L.V. (1959/60). Mathematical methods of organizing and planning production. *Manage. Sci.* **6** 366–422. MR0129016 https://doi.org/10.1287/mnsc.6.4.366
[35] Kocaoglu, M., Snyder, C., Dimakis, A.G. and Vishwanath, S. (2018). CausalGAN: Learning causal implicit generative models with adversarial training. In *International Conference on Learning Representations*.
[36] Korman, J. and McCann, R.J. (2015). Optimal transportation with capacity constraints. *Trans. Amer. Math. Soc.* **367** 1501–1521. MR3286490 https://doi.org/10.1090/S0002-9947-2014-06032-7
[37] Lassalle, R. (2018). Causal transport plans and their Monge-Kantorovich problems. *Stoch. Anal. Appl.* **36** 452–484. MR3784142 https://doi.org/10.1080/07362994.2017.1422747
[38] Lauritzen, S.L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. Oxford University Press, New York: The Clarendon Press. Oxford Science Publications. MR1419991

[39] Li, J. and Lin, L. (2021). Optimal transport with relaxed marginal constraints. *IEEE Access* **9** 58142–58160.
[40] Lorenz, D.A., Manns, P. and Meyer, C. (2021). Quadratically regularized optimal transport. *Appl. Math. Optim.* **83** 1919–1949. MR4261277 https://doi.org/10.1007/s00245-019-09614-w
[41] Mémoli, F. (2011). Gromov–Wasserstein distances and the metric approach to object matching. *Found. Comput. Math.* **11** 417–487. MR2811584 https://doi.org/10.1007/s10208-011-9093-5
[42] Mohajerin Esfahani, P. and Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Math. Program.* **171** 115–166. MR3844536 https://doi.org/10.1007/s10107-017-1172-1
[43] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Mem. Math. Phys. Acad. R. Sci.* 666–704.
[44] Niles-Weed, J. and Berthet, Q. (2022). Minimax estimation of smooth densities in Wasserstein distance. *Ann. Statist.* **50** 1519–1540. MR4441130 https://doi.org/10.1214/21-aos2161
[45] Nutz, M. and Wang, R. (2022). The directional optimal transport. *Ann. Appl. Probab.* **32** 1400–1420. MR4414708 https://doi.org/10.1214/21-aap1712
[46] Pammer, G. (2024). A note on the adapted weak topology in discrete time. *Electron. Commun. Probab.* **29** Paper No. 1, 13. MR4696436
[47] Pass, B. (2015). Multi-marginal optimal transport: Theory and applications. *ESAIM Math. Model. Numer. Anal.* **49** 1771–1790. MR3423275 https://doi.org/10.1051/m2an/2015020
[48] Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2548166 https://doi.org/10.1017/CBO9780511803161
[49] Peters, J., Bühlmann, P. and Meinshausen, N. (2016). Causal inference by using invariant prediction: Identification and confidence intervals. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 947–1012. With comments and a rejoinder. MR3557186 https://doi.org/10.1111/rssb.12167
[50] Peyré, G. and Cuturi, M. (2019). *Computational Optimal Transport: With Applications to Data Science*. Now Publishers and Trends.
[51] 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
[52] Pflug, G.C. and Pichler, A. (2015). Dynamic generation of scenario trees. *Comput. Optim. Appl.* **62** 641–668. MR3426120 https://doi.org/10.1007/s10589-015-9758-0
[53] Pflug, G.C., Pichler, A. and Wozabal, D. (2012). The 1/N investment strategy is optimal under high model ambiguity. *J. Bank. Financ.* **36** 410–417.
[54] Plečko, D. and Meinshausen, N. (2020). Fair data adaptation with quantile preservation. *J. Mach. Learn. Res.* **21** Paper No. 242, 44. MR4209528
[55] Rachev, S.T. (1984). The Monge-Kantorovich problem on mass transfer and its applications in stochastics. *Teor. Veroyatn. Primen.* **29** 625–653. MR0773434
[56] Rachev, S.T. and Rüschendorf, L. (1998). *Mass Transportation Problems. Vol. II Applications*. *Probability and Its Applications (New York)*. New York: Springer. MR1619171
[57] Rosenbaum, P.R. and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70** 41–55. MR0742974 https://doi.org/10.1093/biomet/70.1.41
[58] Sävje, F., Aronow, P.M. and Hudgens, M.G. (2021). Average treatment effects in the presence of unknown interference. *Ann. Statist.* **49** 673–701. MR4255103 https://doi.org/10.1214/20-aos1973
[59] Schölkopf, B. (2022). Causality for machine learning. In *Probabilistic and Causal Inference: The Works of Judea Pearl* 765–804.
[60] Séjourné, T., Peyré, G. and Vialard, F.-X. (2022). Unbalanced optimal transport, from theory to numerics. arxiv Preprint arXiv:2211.08775.
[61] Sriperumbudur, B.K., Fukumizu, K., Gretton, A., Schölkopf, B. and Lanckriet, G.R.G. (2012). On the empirical estimation of integral probability metrics. *Electron. J. Stat.* **6** 1550–1599. MR2988458 https://doi.org/10.1214/12-EJS722
[62] Sudakov, V.N. (1979). Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.* **2** i–v, 1–178. MR0530375
[63] Titouan, V., Redko, I., Flamary, R. and Courty, N. (2020). Co-optimal transport. *Adv. Neural Inf. Process. Syst.* **33** 17559–17570.

[64] Tran, Q.H., Janati, H., Redko, I., Rémi, F. and Courty, N. (2021). Factored couplings in multi-marginal optimal transport via difference of convex programming. In *NeurIPS 2021 Optimal Transport and Machine Learning Workshop (OTML)*.

[65] Vallander, S.S. (1973). Calculations of the Vasseršteĭn distance between probability distributions on the line. *Teor. Veroyatn. Primen.* **18** 824–827. MR0328982

[66] Vaserstein, L.N. (1969). Markov processes over denumerable products of spaces describing large system of automata. *Problemy Peredachi Informatsii* **5** 64–72. MR0314115 https://doi.org/10.1016/s0016-0032(33)90010-1

[67] Villani, C. (2009). *Optimal Transport: Old and New*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9

[68] Xu, T. and Acciaio, B. (2021). Quantized Conditional COT-GAN for Video Prediction. arXiv Preprint arXiv:2106.05658.

[69] Xu, T., Wenliang, L.K., Munn, M. and Acciaio, B. (2020). Cot-gan: Generating sequential data via causal optimal transport. *Adv. Neural Inf. Process. Syst.* **33** 8798–8809.

[70] Yang, I. (2017). A convex optimization approach to distributionally robust Markov decision processes with Wasserstein distance. *IEEE Control Syst. Lett.* **1** 164–169. MR4208527

Bernoulli **31**(2), 2025, 1377–1401
https://doi.org/10.3150/24-BEJ1774

Reversibility of elliptical slice sampling revisited

MAREIKE HASENPFLUG[1,a], VIACHESLAV TELEZHNIKOV[2,c] and DANIEL RUDOLF[1,b]

[1]*Faculty of Computer Science and Mathematics, Universität Passau, Innstraße 33, 94032 Passau, Germany,* [a]*mareike.hasenplfug@uni-passau.de,* [b]*daniel.rudolf@uni-passau.de*
[2]*previously known as Viacheslav Natarovskii, Expert Analytics GmbH, Hubertusstraße 83, 82131 Gauting, Germany,* [c]*viacheslav.telezhnikov@expertanalytics.de*

We extend elliptical slice sampling, a Markov chain transition kernel suggested in Murray, Adams and MacKay (In *The Proceedings of the 13th International Conference on Artificial Intelligence and Statistics* (2010) 541–548), to infinite-dimensional separable Hilbert spaces and discuss its well-definedness. We point to a regularity requirement, provide an alternative proof of the desirable reversibility property and show that it induces a positive semi-definite Markov operator. Crucial within the proof of the formerly mentioned results is the analysis of a shrinkage Markov chain that may be interesting on its own.

Keywords: Elliptical slice sampling; reversibility; shrinkage procedure

References

Beskos, A., Pinski, F.J., Sanz-Serna, J.M. and Stuart, A.M. (2011). Hybrid Monte Carlo on Hilbert spaces. *Stochastic Process. Appl.* **121** 2201–2230. MR2822774 https://doi.org/10.1016/j.spa.2011.06.003

Beskos, A., Girolami, M., Lan, S., Farrell, P.E. and Stuart, A.M. (2017). Geometric MCMC for infinite-dimensional inverse problems. *J. Comput. Phys.* **335** 327–351. MR3612501 https://doi.org/10.1016/j.jcp.2016.12.041

Cotter, S.L., Roberts, G.O., Stuart, A.M. and White, D. (2013). MCMC methods for functions: Modifying old algorithms to make them faster. *Statist. Sci.* **28** 424–446. MR3135540 https://doi.org/10.1214/13-STS421

Cui, T., Law, K.J.H. and Marzouk, Y.M. (2016). Dimension-independent likelihood-informed MCMC. *J. Comput. Phys.* **304** 109–137. MR3422405 https://doi.org/10.1016/j.jcp.2015.10.008

Da Prato, G. and Zabczyk, J. (2002). *Second Order Partial Differential Equations in Hilbert Spaces*. *London Mathematical Society Lecture Note Series* **293**. Cambridge: Cambridge Univ. Press. MR1985790 https://doi.org/10.1017/CBO9780511543210

Douc, R., Moulines, E., Priouret, P. and Soulier, P. (2018). *Markov Chains*. *Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3889011 https://doi.org/10.1007/978-3-319-97704-1

Gessner, A., Kanjilal, O. and Hennig, P. (2020). Integrals over Gaussians under linear domain constraints. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics* 2764–2774.

Grenioux, L., O. Durmus, A., Moulines, E. and Gabrié, M. (2023). On sampling with approximate transport maps. In *Proceedings of the 40th International Conference on Machine Learning, PMLR* **202** 11698–11733.

Habeck, M., Rudolf, D. and Sprungk, B. (2020). Stability of doubly-intractable distributions. *Electron. Commun. Probab.* **25** Paper No. 62, 13. MR4151879 https://doi.org/10.1214/20-ECP341

Habeck, M., Hasenpflug, M., Kodgirwar, S. and Rudolf, D. (2023). Geodesic slice sampling on the sphere. arXiv:2301.08056. Preprint. https://doi.org/10.48550/arXiv.2301.08056

Jones, G.L. (2004). On the Markov chain central limit theorem. *Probab. Surv.* **1** 299–320. MR2068475 https://doi.org/10.1214/154957804100000051

Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications (New York)*. New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

1350-7265

Łatuszyński, K. and Rudolf, D. (2024). Convergence of hybrid slice sampling via spectral gap. *J. Appl. Probab.* Accepted. https://doi.org/10.48550/arXiv.1409.2709

Law, K.J.H. (2014). Proposals which speed up function-space MCMC. *J. Comput. Appl. Math.* **262** 127–138. MR3162309 https://doi.org/10.1016/j.cam.2013.07.026

Lie, H.C., Rudolf, D., Sprungk, B. and Sullivan, T.J. (2023). Dimension-independent Markov chain Monte Carlo on the sphere. *Scand. J. Stat.* **50** 1818–1858. MR4677377 https://doi.org/10.1111/sjos.12653

Medina-Aguayo, F., Rudolf, D. and Schweizer, N. (2020). Perturbation bounds for Monte Carlo within Metropolis via restricted approximations. *Stochastic Process. Appl.* **130** 2200–2227. MR4074698 https://doi.org/10.1016/j.spa.2019.06.015

Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. With a prologue by Peter W. Glynn. MR2509253 https://doi.org/10.1017/CBO9780511626630

Murray, I., Adams, R.P. and MacKay, D.J.C. (2010). Elliptical slice sampling. In *The Proceedings of the 13th International Conference on Artificial Intelligence and Statistics* 541–548.

Murray, I. and Graham, M. (2016). Pseudo-marginal slice sampling. In *The Proceedings of the 19th International Conference on Artificial Intelligence and Statistics* 911–919.

Natarovskii, V., Rudolf, D. and Sprungk, B. (2021a). Geometric convergence of elliptical slice sampling. In *Proceedings of the 38th International Conference on Machine Learning, PMLR* 7969–7978.

Natarovskii, V., Rudolf, D. and Sprungk, B. (2021b). Quantitative spectral gap estimate and Wasserstein contraction of simple slice sampling. *Ann. Appl. Probab.* **31** 806–825. MR4254496 https://doi.org/10.1214/20-aap1605

Neal, R.M. (1999). Regression and classification using Gaussian process priors. In *Bayesian Statistics, 6 (Alcoceber, 1998)* 475–501. New York: Oxford Univ. Press. MR1723510

Neal, R.M. (2003). Slice sampling. *Ann. Statist.* **31** 705–767. With discussions and a rejoinder by the author. MR1994729 https://doi.org/10.1214/aos/1056562461

Nishihara, R., Murray, I. and Adams, R.P. (2014). Parallel MCMC with generalized elliptical slice sampling. *J. Mach. Learn. Res.* **15** 2087–2112. MR3231603

Ottobre, M., Pillai, N.S., Pinski, F.J. and Stuart, A.M. (2016). A function space HMC algorithm with second order Langevin diffusion limit. *Bernoulli* **22** 60–106. MR3449777 https://doi.org/10.3150/14-BEJ621

Pinski, F.J., Simpson, G., Stuart, A.M. and Weber, H. (2015). Algorithms for Kullback-Leibler approximation of probability measures in infinite dimensions. *SIAM J. Sci. Comput.* **37** A2733–A2757. MR3424069 https://doi.org/10.1137/14098171X

Ray, P., Pati, D. and Bhattacharya, A. (2020). Efficient Bayesian shape-restricted function estimation with constrained Gaussian process priors. *Stat. Comput.* **30** 839–853. MR4108680 https://doi.org/10.1007/s11222-020-09922-0

Rudolf, D. (2012). Explicit error bounds for Markov chain Monte Carlo. *Dissertationes Math.* **485** 1–93. MR2977521 https://doi.org/10.4064/dm485-0-1

Rudolf, D. and Schweizer, N. (2018). Perturbation theory for Markov chains via Wasserstein distance. *Bernoulli* **24** 2610–2639. MR3779696 https://doi.org/10.3150/17-BEJ938

Rudolf, D. and Sprungk, B. (2018). On a generalization of the preconditioned Crank-Nicolson Metropolis algorithm. *Found. Comput. Math.* **18** 309–343. MR3777781 https://doi.org/10.1007/s10208-016-9340-x

Rudolf, D. and Sprungk, B. (2022). Robust random walk-like Metropolis-Hastings algorithms for concentrating posteriors. arXiv:2202.12127. Preprint. https://doi.org/10.48550/arXiv.2202.12127

Schär, P., Habeck, M. and Rudolf, D. (2023). Gibbsian polar slice sampling. In *Proceedings of the 40th International Conference on Machine Learning* 30204–30223.

Sprungk, B. (2020). On the local Lipschitz stability of Bayesian inverse problems. *Inverse Probl.* **36** 055015, 31. MR4105328 https://doi.org/10.1088/1361-6420/ab6f43

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1762. With discussion and a rejoinder by the author. MR1329166 https://doi.org/10.1214/aos/1176325750

Bernoulli **31**(2), 2025, 1402–1427
https://doi.org/10.3150/24-BEJ1775

Heat kernel estimates for kinetic SDEs with drifts being unbounded and in Kato's class

CHONGYANG REN[1,a] and XICHENG ZHANG[2,b]

[1]*School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China,* [a]*rcy.math@whu.edu.cn*
[2]*School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China,*
[b]*XichengZhang@gmail.com*

In this paper we investigate the existence and uniqueness of weak solutions for kinetic stochastic differential equations with Hölder diffusion and unbounded singular drifts in Kato's class. Moreover, we also establish sharp two-sided estimates for the density of the solution. In particular, the drift b can be in the mixed $L^q_t L^{p_1}_{x_1} L^{p_2}_{x_2}$ space with $2/q + d/p_1 + 3d/p_2 < 1$. As an application, we show the existence and uniqueness of weak solution to a second order singular interacting particle system in $\mathbb{R}^{dN}$.

Keywords: Heat kernel estimates; Kato's class; kinetic SDEs; Krylov's estimate

References

[1] Aizenman, M. and Simon, B. (1982). Brownian motion and Harnack inequality for Schrödinger operators. *Comm. Pure Appl. Math.* **35** 209–273. MR0644024 https://doi.org/10.1002/cpa.3160350206

[2] Boué, M. and Dupuis, P. (1998). A variational representation for certain functionals of Brownian motion. *Ann. Probab.* **26** 1641–1659. MR1675051 https://doi.org/10.1214/aop/1022855876

[3] Chaudru de Raynal, P., Honoré, I. and Menozzi, S. (2021). Sharp Schauder estimates for some degenerate Kolmogorov equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **22** 989–1089. MR4334312

[4] Chaudru de Raynal, P. and Menozzi, S. (2022). Regularization effects of a noise propagating through a chain of differential equations: An almost sharp result. *Trans. Amer. Math. Soc.* **375** 1–45. MR4358660 https://doi.org/10.1090/tran/7947

[5] Chaudru de Raynal, P.E. (2017). Strong existence and uniqueness for degenerate SDE with Hölder drift. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 259–286. MR3606742 https://doi.org/10.1214/15-AIHP716

[6] Chaudru de Raynal, P.E., Menozzi, S., Pesce, A. and Zhang, X. (2023). Heat kernel and gradient estimates for kinetic SDEs with low regularity coefficients. *Bull. Sci. Math.* **183** 103229. MR4554678 https://doi.org/10.1016/j.bulsci.2023.103229

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

[8] Delarue, F. and Menozzi, S. (2010). Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259** 1577–1630. MR2659772 https://doi.org/10.1016/j.jfa.2010.05.002

[9] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes: Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658

[10] Fedrizzi, E., Flandoli, F., Priola, E. and Vovelle, J. (2017). Regularity of stochastic kinetic equations. *Electron. J. Probab.* **22** 48. MR3661662 https://doi.org/10.1214/17-EJP65

[11] Gadat, S. and Panloup, F. (2014). Long time behaviour and stationary regime of memory gradient diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 564–601. MR3189085 https://doi.org/10.1214/12-AIHP536

[12] Garrido, L., Seglar, P. and Shepherd, P.J., eds. (1978). *Stochastic Processes in Nonequilibrium Systems*. *Lecture Notes in Physics* **84**. Berlin: Springer.

[13] Hao, Z., Röckner, M. and Zhang, X. (2024). Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions. *SIAM J. Math. Anal.* **56** 2661–2713. MR4722362 https://doi.org/10.1137/23M1556666

[14] Hao, Z., Wu, M. and Zhang, X. (2020). Schauder estimates for nonlocal kinetic equations and applications. *J. Math. Pures Appl. (9)* **140** 139–184. MR4124429 https://doi.org/10.1016/j.matpur.2020.06.003

[15] Hérau, F. and Nier, F. (2004). Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.* **171** 151–218. MR2034753 https://doi.org/10.1007/s00205-003-0276-3

[16] Kato, T. (1972). Schrödinger operators with singular potentials. *Israel J. Math.* **13** 135–148 (1973). MR0333833 https://doi.org/10.1007/BF02760233

[17] Kolmogoroff, A. (1934). Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. (2)* **35** 116–117. MR1503147 https://doi.org/10.2307/1968123

[18] Konakov, V., Menozzi, S. and Molchanov, S. (2010). Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 908–923. MR2744877 https://doi.org/10.1214/09-AIHP207

[19] Krylov, N.V. (1996). *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. *Graduate Studies in Mathematics* **12**. Providence, RI: Amer. Math. Soc. MR1406091 https://doi.org/10.1090/gsm/012

[20] Ling, C. and Xie, L. (2022). Strong solutions of stochastic differential equations with coefficients in mixed-norm spaces. *Potential Anal.* **57** 227–241. MR4443124 https://doi.org/10.1007/s11118-021-09913-4

[21] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266 https://doi.org/10.1016/S0304-4149(02)00150-3

[22] Menozzi, S. (2011). Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16** 234–250. MR2802040 https://doi.org/10.1214/ECP.v16-1619

[23] Skorokhod, A.V. (1965). *Studies in the Theory of Random Processes*. Reading, MA: Addison-Wesley. Translated from the Russian by Scripta Technica, Inc. MR0185620

[24] Soize, C. (1994). *The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions*. *Series on Advances in Mathematics for Applied Sciences* **17**. River Edge, NJ: World Scientific. MR1287386 https://doi.org/10.1142/9789814354110

[25] Stroock, D.W. and Varadhan, S.R.S. (1979). *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften* **233**. Berlin: Springer. MR0532498

[26] Talay, D. (2002). Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** 163–198.

[27] van Kampen, N.G. (1981). *Stochastic Processes in Physics and Chemistry*. *Lecture Notes in Math.* **888**. Amsterdam: North-Holland. MR0648937

[28] Villani, C. (2009). Hypocoercivity. *Mem. Amer. Math. Soc.* **202** iv+141. MR2562709 https://doi.org/10.1090/S0065-9266-09-00567-5

[29] Wang, F.-Y. and Zhang, X. (2016). Degenerate SDE with Hölder-Dini drift and non-Lipschitz noise coefficient. *SIAM J. Math. Anal.* **48** 2189–2226. MR3511355 https://doi.org/10.1137/15M1023671

[30] Xia, P., Xie, L., Zhang, X. and Zhao, G. (2020). $L^q(L^p)$-theory of stochastic differential equations. *Stochastic Process. Appl.* **130** 5188–5211. MR4108486 https://doi.org/10.1016/j.spa.2020.03.004

[31] Zhang, Q. (1996). A Harnack inequality for the equation $\nabla(a\nabla u) + b\nabla u = 0$, when $|b| \in K_{n+1}$. *Manuscripta Math.* **89** 61–77. MR1368536 https://doi.org/10.1007/BF02567505

[32] Zhang, Q.S. (1997). Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$. *Manuscripta Math.* **93** 381–390. MR1457736 https://doi.org/10.1007/BF02677479

[33] Zhang, X. (2009). A variational representation for random functionals on abstract Wiener spaces. *J. Math. Kyoto Univ.* **49** 475–490. MR2583599 https://doi.org/10.1215/kjm/1260975036

[34] Zhang, X. (2018). Stochastic Hamiltonian flows with singular coefficients. *Sci. China Math.* **61** 1353–1384. MR3833741 https://doi.org/10.1007/s11425-017-9127-0

[35] Zhang, X. and Zhao, G. (2018). Singular Brownian diffusion processes. *Commun. Math. Stat.* **6** 533–581.

Bernoulli **31**(2), 2025, 1428–1450
https://doi.org/10.3150/24-BEJ1776

Brownian motion conditioned to spend limited time outside a bounded interval – an extreme example of entropic repulsion

FRANK AURZADA[1,a], MARTIN KOLB[2,b] and DOMINIC T. SCHICKENTANZ[1,2,c]

[1]*Department of Mathematics, Technical University of Darmstadt, Darmstadt, Germany,*
[a]*aurzada@mathematik.tu-darmstadt.de*
[2]*Institute of Mathematics, Paderborn University, Paderborn, Germany,* [b]*kolb@math.uni-paderborn.de,*
[c]*info@dominic-schickentanz.com*

We show that a Brownian motion on $\mathbb{R}_{\geq 0}$ which is allowed to spend a total of $s > 0$ time units outside a bounded interval does not leave the interval at all. This can be seen as an extreme example of entropic repulsion. Moreover, we explicitly determine the exact asymptotic behavior of the probability that a Brownian motion on $[0,T]$ spends limited time outside a bounded interval, as $T \to \infty$.

Keywords: Brownian motion; conditioned process; extreme entropic repulsion; occupation time

References

[1] Aurzada, F. and Schickentanz, D.T. (2022). Brownian motion conditioned to spend limited time below a barrier. *Stochastic Process. Appl.* **146** 360–381. MR4374940 https://doi.org/10.1016/j.spa.2022.01.007
[2] Barker, A. (2020). Transience and recurrence of Markov processes with constrained local time. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** 993–1045. MR4182158 https://doi.org/10.30757/alea.v17-36
[3] Benjamini, I. and Berestycki, N. (2010). Random paths with bounded local time. *J. Eur. Math. Soc. (JEMS)* **12** 819–854. MR2654081 https://doi.org/10.4171/JEMS/216
[4] Benjamini, I. and Berestycki, N. (2011). An integral test for the transience of a Brownian path with limited local time. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 539–558. MR2814422 https://doi.org/10.1214/10-AIHP371
[5] Borodin, A.N. and Salminen, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Basel: Birkhäuser. MR1912205 https://doi.org/10.1007/978-3-0348-8163-0
[6] Champagnat, N. and Villemonais, D. (2016). Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields* **164** 243–283. MR3449390 https://doi.org/10.1007/s00440-014-0611-7
[7] Darling, D.A. and Siegert, A.J.F. (1953). The first passage problem for a continuous Markov process. *Ann. Math. Stat.* **24** 624–639. MR0058908 https://doi.org/10.1214/aoms/1177728918
[8] Doob, J.L. (1957). Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France* **85** 431–458. MR0109961
[9] Ingham, A.E. (1941). A Tauberian theorem for partitions. *Ann. of Math. (2)* **42** 1075–1090. MR0005522 https://doi.org/10.2307/1970462
[10] Kac, M. (1949). On distributions of certain Wiener functionals. *Trans. Amer. Math. Soc.* **65** 1–13. MR0027960 https://doi.org/10.2307/1990512
[11] Kac, M. (1951). On some connections between probability theory and differential and integral equations. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950* 189–215. Berkeley–Los Angeles, CA: Univ. California Press. MR0045333
[12] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

1350-7265

[13] Knight, F.B. (1969). Brownian local times and taboo processes. *Trans. Amer. Math. Soc.* **143** 173–185. MR0253424 https://doi.org/10.2307/1995241
[14] Kolb, M. and Savov, M. (2016). Transience and recurrence of a Brownian path with limited local time. *Ann. Probab.* **44** 4083–4132. MR3572332 https://doi.org/10.1214/15-AOP1069
[15] Kolb, M. and Steinsaltz, D. (2012). Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.* **40** 162–212. MR2917771 https://doi.org/10.1214/10-AOP623
[16] Korevaar, J. (2004). *Tauberian Theory: A Century of Developments*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **329**. Berlin: Springer. MR2073637 https://doi.org/10.1007/978-3-662-10225-1
[17] Pinsky, R.G. (1985). On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Ann. Probab.* **13** 363–378. MR0781410
[18] Profeta, C. (2015). Some limiting laws associated with the integrated Brownian motion. *ESAIM Probab. Stat.* **19** 148–171. MR3386368 https://doi.org/10.1051/ps/2014018
[19] Steinsaltz, D. and Evans, S.N. (2007). Quasistationary distributions for one-dimensional diffusions with killing. *Trans. Amer. Math. Soc.* **359** 1285–1324. MR2262851 https://doi.org/10.1090/S0002-9947-06-03980-8
[20] Stroock, D.W. (1993). *Probability Theory, an Analytic View*. Cambridge: Cambridge Univ. Press. MR1267569
[21] Whitt, W. (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Stat.* **41** 939–944. MR0261646 https://doi.org/10.1214/aoms/1177696970

Bernoulli **31**(2), 2025, 1451–1474
https://doi.org/10.3150/24-BEJ1777

Generalized Hadamard differentiability of the copula mapping and its applications

NATALIE NEUMEYER[1,a] and MAREK OMELKA[2,b]

[1]*Department of Mathematics, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany,* [a]*natalie.neumeyer@uni-hamburg.de*
[2]*Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic,* [b]*omelka@karlin.mff.cuni.cz*

We consider the copula mapping, which maps a joint cumulative distribution function to the corresponding copula. Its Hadamard differentiability was shown in van der Vaart and Wellner (1996), Fermanian, Radulović and Wegkamp (2004) and (under less strict assumptions) in Bücher and Volgushev (2013). This differentiability result has proved to be a powerful tool to show the weak convergence of empirical copula processes in various settings using the functional delta method. We state a generalization of the Hadamard differentiability results and illustrate how it can be used for the derivations of asymptotic expansions and the weak convergence of empirical copula processes in the presence of covariates. The usefulness of this result is illustrated on several applications which include a multidimensional functional linear model, where the copula of the error vector describes the dependency between the components of the vector of observations, given the functional covariate.

Keywords: Empirical copula; functional delta method; Hadamard derivative; pseudo-observations; regression residuals; weak convergence

References

Berghaus, B., Bücher, A. and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli* **23** 743–772. MR3556791 https://doi.org/10.3150/15-BEJ751

Bücher, A. and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.* **119** 61–70. MR3061415 https://doi.org/10.1016/j.jmva.2013.04.003

Cardot, H., Mas, A. and Sarda, P. (2007). CLT in functional linear regression models. *Probab. Theory Related Fields* **138** 325–361. MR2299711 https://doi.org/10.1007/s00440-006-0025-2

Côté, M.-P., Genest, C. and Omelka, M. (2019). Rank-based inference tools for copula regression, with property and casualty insurance applications. *Insurance Math. Econom.* **89** 1–15. MR4007254 https://doi.org/10.1016/j.insmatheco.2019.08.001

Dedecker, J. and Louhichi, S. (2002). Maximal inequalities and empirical central limit theorems. In *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sorensen, eds.). 137–159. Boston, MA: Birkhäuser. MR1958779

Dette, H., Pardo-Fernández, J.C. and Van Keilegom, I. (2009). Goodness-of-fit tests for multiplicative models with dependent data. *Scand. J. Stat.* **36** 782–799. MR2573308 https://doi.org/10.1111/j.1467-9469.2009.00648.x

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. *Springer Series in Statistics*. New York: Springer. MR1964455 https://doi.org/10.1007/b97702

Fermanian, J.-D., Radulović, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. MR2093613 https://doi.org/10.3150/bj/1099579158

Gänssler, P. and Stute, W. (1987). *Seminar on Empirical Processes*. *DMV Seminar* **9**. Basel: Birkhäuser. MR0902803 https://doi.org/10.1007/978-3-0348-6269-1

Ghoudi, K. and Rémillard, B. (1998). Empirical processes based on pseudo-observations. In *Asymptotic Methods in Probability and Statistics, a Volume in Honour of Miklós Csörgő* (B. Szyskowitz, ed.). 171–197. Amsterdam: Elsevier. MR1661523

1350-7265

Gijbels, I., Omelka, M. and Veraverbeke, N. (2015). Estimation of a copula when a covariate affects only marginal distributions. *Scand. J. Stat.* **42** 1109–1126. MR3426313 https://doi.org/10.1111/sjos.12154

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

Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. MR2332269 https://doi.org/10.1214/009053606000000957

Masry, E. (1996). Multivariate local polynomial regression for time series: Uniform strong consistency and rates. *J. Time Series Anal.* **17** 571–599. MR1424907 https://doi.org/10.1111/j.1467-9892.1996.tb00294.x

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

Neumeyer, N. and Omelka, M. (2024). Supplement to "Generalized Hadamard differentiability of the copula mapping and its applications." https://doi.org/10.3150/24-BEJ1777SUPP

Neumeyer, N., Omelka, M. and Hudecová, Š. (2019). A copula approach for dependence modeling in multivariate nonparametric time series. *J. Multivariate Anal.* **171** 139–162. MR3892031 https://doi.org/10.1016/j.jmva.2018.11.016

Nickl, R. and Pötscher, B.M. (2007). Bracketing metric entropy rates and empirical central limit theorems for function classes of Besov- and Sobolev-type. *J. Theoret. Probab.* **20** 177–199. MR2324525 https://doi.org/10.1007/s10959-007-0058-1

Omelka, M., Gijbels, I. and Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *Ann. Statist.* **37** 3023–3058. MR2541454 https://doi.org/10.1214/08-AOS666

Portier, F. and Segers, J. (2018). On the weak convergence of the empirical conditional copula under a simplifying assumption. *J. Multivariate Anal.* **166** 160–181. MR3799641 https://doi.org/10.1016/j.jmva.2018.03.002

Rémillard, B. (2017). Goodness-of-fit tests for copulas of multivariate time series. *Econom. J.* **5**. https://doi.org/10.3390/econometrics5010013

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. MR2948900 https://doi.org/10.3150/11-BEJ387

Shang, Z. and Cheng, G. (2015). Nonparametric inference in generalized functional linear models. *Ann. Statist.* **43** 1742–1773. MR3357877 https://doi.org/10.1214/15-AOS1322

Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33** 357–375. MR2193980 https://doi.org/10.1002/cjs.5540330304

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

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942 https://doi.org/10.1214/074921707000000382

Yeon, H., Dai, X. and Nordman, D.J. (2023). Bootstrap inference in functional linear regression models with scalar response. *Bernoulli* **29** 2599–2626. MR4631413 https://doi.org/10.3150/22-BEJ1554

Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. MR2766857 https://doi.org/10.1214/09-AOS772

Bernoulli **31**(2), 2025, 1475–1501
https://doi.org/10.3150/24-BEJ1778

Bayesian inference for k-monotone densities with applications to multiple testing

KANG WANG[1,a] and SUBHASHIS GHOSAL[2,b]

[1]*Department of Biostatistics, M. D. Anderson Cancer Center, Houston, TX 77030, USA,* [a]*kwang8@mdanderson.org*
[2]*Department of Statistics, North Carolina State University, Raleigh, NC 27606, USA,* [b]*sghosal@stat.ncsu.edu*

Shape restriction imposed on a function of interest, such as a regression or density function, allows for its estimation without smoothness assumptions. The concept of k-monotonicity encompasses a family of shape restrictions, including decreasing and convex decreasing as special cases corresponding to $k = 1$ and $k = 2$. We consider Bayesian approaches to estimate a k-monotone density. By utilizing a kernel mixture representation and putting a Dirichlet process or a finite mixture prior on the mixing distribution, we show that the posterior contraction rate in the Hellinger distance is $(n/\log n)^{-k/(2k+1)}$ for a k-monotone density, which is minimax optimal up to a polylogarithmic factor. When $k_0 = \infty$ or the true k-monotone density is a finite J_0-component mixture of the kernel, the contraction rate improves to the nearly parametric rate. Moreover, by putting a prior on k, we show that almost the same rates hold even when the best value of k is unknown. A specific application in modeling the density of p-values in a large-scale multiple testing problem is considered. Simulation studies are conducted to evaluate the performance of the proposed method.

Keywords: Adaptation; contraction rates; Dirichlet process mixture; mixture representation; shape restrictions

References

[1] Balabdaoui, F. and Wellner, J.A. (2007). Estimation of a k-monotone density: Limit distribution theory and the spline connection. *Ann. Statist.* **35** 2536–2564. MR2382657 https://doi.org/10.1214/009053607000000262
[2] Balabdaoui, F. and Wellner, J.A. (2010). Estimation of a k-monotone density: Characterizations, consistency and minimax lower bounds. *Stat. Neerl.* **64** 45–70. MR2830965 https://doi.org/10.1111/j.1467-9574.2009.00438.x
[3] Bayarri, M.J. and Berger, J.O. (2000). p values for composite null models. *J. Amer. Statist. Assoc.* **95** 1127–1142, 1157–1170. With comments and a rejoinder by the authors. MR1804239 https://doi.org/10.2307/2669749
[4] Bean, G.J., Dimarco, E.A., Mercer, L.D., Thayer, L.K., Roy, A. and Ghosal, S. (2013). Finite skew-mixture models for estimation of positive false discovery rates. *Stat. Methodol.* **10** 46–57. MR2974809 https://doi.org/10.1016/j.stamet.2012.05.005
[5] Birgé, L. (1987). Estimating a density under order restrictions: Nonasymptotic minimax risk. *Ann. Statist.* **15** 995–1012. MR0902241 https://doi.org/10.1214/aos/1176350488
[6] Blei, D.M. and Jordan, M.I. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Anal.* **1** 121–143. MR2227367 https://doi.org/10.1214/06-BA104
[7] Brunner, L.J. (1992). Bayesian nonparametric methods for data from a unimodal density. *Statist. Probab. Lett.* **14** 195–199. MR1173617 https://doi.org/10.1016/0167-7152(92)90021-V
[8] Brunner, L.J. and Lo, A.Y. (1989). Bayes methods for a symmetric unimodal density and its mode. *Ann. Statist.* **17** 1550–1566. MR1026299 https://doi.org/10.1214/aos/1176347381
[9] Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *Ann. Statist.* **43** 1986–2018. MR3375874 https://doi.org/10.1214/15-AOS1334

[10] DeVore, R.A. and Lorentz, G.G. (1993). *Constructive Approximation*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **303**. Berlin: Springer. MR1261635

[11] Escobar, M.D. and West, M. (1995). Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* **90** 577–588. MR1340510

[12] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. MR0270403

[13] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

[14] Ferguson, T.S. (1983). Bayesian density estimation by mixtures of normal distributions. In *Recent Advances in Statistics* 287–302. New York: Academic Press. MR0736538

[15] Gao, F. (2008). Entropy estimate for k-monotone functions via small ball probability of integrated Brownian motion. *Electron. Commun. Probab.* **13** 121–130. MR2386068 https://doi.org/10.1214/ECP.v13-1357

[16] Gao, F. and Wellner, J.A. (2009). On the rate of convergence of the maximum likelihood estimator of a k-monotone density. *Sci. China Ser. A* **52** 1525–1538. MR2520591 https://doi.org/10.1007/s11425-009-0102-y

[17] Ghosal, S. (2001). Convergence rates for density estimation with Bernstein polynomials. *Ann. Statist.* **29** 1264–1280. MR1873330 https://doi.org/10.1214/aos/1013203453

[18] Ghosal, S., Ghosh, J.K. and Ramamoorthi, R.V. (1999). Posterior consistency of Dirichlet mixtures in density estimation. *Ann. Statist.* **27** 143–158. MR1701105 https://doi.org/10.1214/aos/1018031105

[19] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228

[20] Ghosal, S., Lember, J. and van der Vaart, A. (2008). Nonparametric Bayesian model selection and averaging. *Electron. J. Stat.* **2** 63–89. MR2386086 https://doi.org/10.1214/07-EJS090

[21] Ghosal, S. and Roy, A. (2009). Bayesian nonparametric approach to multiple testing. In *Perspectives in Mathematical Sciences. I. Stat. Sci. Interdiscip. Res.* **7** 139–164. Hackensack, NJ: World Sci. Publ. MR2581743 https://doi.org/10.1142/9789814273633_0007

[22] Ghosal, S. and Roy, A. (2011). Predicting false discovery proportion under dependence. *J. Amer. Statist. Assoc.* **106** 1208–1218. MR2894775 https://doi.org/10.1198/jasa.2011.tm10488

[23] Ghosal, S. and Roy, A. (2011). Identifiability of the proportion of null hypotheses in skew-mixture models for the p-value distribution. *Electron. J. Stat.* **5** 329–341. MR2802046 https://doi.org/10.1214/11-EJS609

[24] Ghosal, S., Roy, A. and Tang, Y. (2008). Posterior consistency of Dirichlet mixtures of beta densities in estimating positive false discovery rates. In *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen. Inst. Math. Stat. (IMS) Collect.* **1** 105–115. Beachwood, OH: IMS. MR2462198 https://doi.org/10.1214/193940307000000077

[25] Ghosal, S. and van der Vaart, A. (2007). Posterior convergence rates of Dirichlet mixtures at smooth densities. *Ann. Statist.* **35** 697–723. MR2336864 https://doi.org/10.1214/009053606000001271

[26] Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR3587782 https://doi.org/10.1017/9781139029834

[27] Ghosal, S. and van der Vaart, A.W. (2001). Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. *Ann. Statist.* **29** 1233–1263. MR1873329 https://doi.org/10.1214/aos/1013203453

[28] Grenander, U. (1956). On the theory of mortality measurement. II. *Skand. Aktuarietidskr.* **39** 125–153. MR0093415 https://doi.org/10.1080/03461238.1956.10414944

[29] Groeneboom, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983). Wadsworth Statist./Probab. Ser.* 539–555. Belmont, CA: Wadsworth. MR0822052

[30] Groeneboom, P. and Jongbloed, G. (2014). *Nonparametric Estimation Under Shape Constraints: Estimators, Algorithms and Asymptotics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **38**. New York: Cambridge Univ. Press. MR3445293 https://doi.org/10.1017/CBO9781139020893

[31] 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

[32] Huang, J. and Wellner, J.A. (1995). Estimation of a monotone density or monotone hazard under random censoring. *Scand. J. Stat.* **22** 3–33. MR1334065

[33] Huang, Y. and Zhang, C.-H. (1994). Estimating a monotone density from censored observations. *Ann. Statist.* **22** 1256–1274. MR1311975 https://doi.org/10.1214/aos/1176325628
[34] Jewell, N.P. (1982). Mixtures of exponential distributions. *Ann. Statist.* **10** 479–484. MR0653523
[35] Kalli, M., Griffin, J.E. and Walker, S.G. (2011). Slice sampling mixture models. *Stat. Comput.* **21** 93–105. MR2746606 https://doi.org/10.1007/s11222-009-9150-y
[36] Kopotun, K. and Shadrin, A. (2003). On k-monotone approximation by free knot splines. *SIAM J. Math. Anal.* **34** 901–924. MR1969607 https://doi.org/10.1137/S0036141002358514
[37] Kruijer, W. and van der Vaart, A. (2008). Posterior convergence rates for Dirichlet mixtures of beta densities. *J. Statist. Plann. Inference* **138** 1981–1992. MR2406419 https://doi.org/10.1016/j.jspi.2007.07.012
[38] Langaas, M., Lindqvist, B.H. and Ferkingstad, E. (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 555–572. MR2168204 https://doi.org/10.1111/j.1467-9868.2005.00515.x
[39] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. MR0733519 https://doi.org/10.1214/aos/1176346412
[40] MacEachern, S.N. and Müller, P. (1998). Estimating mixture of Dirichlet process models. *J. Comput. Graph. Statist.* **7** 223–238.
[41] Mammen, E. (1991). Nonparametric regression under qualitative smoothness assumptions. *Ann. Statist.* **19** 741–759. MR1105842 https://doi.org/10.1214/aos/1176348118
[42] Mariucci, E., Ray, K. and Szabó, B. (2020). A Bayesian nonparametric approach to log-concave density estimation. *Bernoulli* **26** 1070–1097. MR4058361 https://doi.org/10.3150/19-BEJ1139
[43] Martin, R. (2019). Empirical priors and posterior concentration rates for a monotone density. *Sankhya A* **81** 493–509. MR4043484 https://doi.org/10.1007/s13171-018-0147-5
[44] Martin, R. and Walker, S.G. (2019). Data-driven priors and their posterior concentration rates. *Electron. J. Stat.* **13** 3049–3081. MR4010592 https://doi.org/10.1214/19-ejs1600
[45] Petrone, S. (1999). Bayesian density estimation using Bernstein polynomials. *Canad. J. Statist.* **27** 105–126. MR1703623 https://doi.org/10.2307/3315494
[46] Petrone, S. and Wasserman, L. (2002). Consistency of Bernstein polynomial posteriors. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 79–100. MR1881846 https://doi.org/10.1111/1467-9868.00326
[47] Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density. *Sankhyā Ser. A* **31** 23–36. MR0267677
[48] Robins, J.M., van der Vaart, A. and Ventura, V. (2000). Asymptotic distribution of p values in composite null models. *J. Amer. Statist. Assoc.* **95** 1143–1167, 1171–1172. With comments and a rejoinder by the authors. MR1804240 https://doi.org/10.2307/2669750
[49] Rousseau, J. (2010). Rates of convergence for the posterior distributions of mixtures of betas and adaptive nonparametric estimation of the density. *Ann. Statist.* **38** 146–180. MR2589319 https://doi.org/10.1214/09-AOS703
[50] Salomond, J.-B. (2014). Concentration rate and consistency of the posterior distribution for selected priors under monotonicity constraints. *Electron. J. Stat.* **8** 1380–1404. MR3263126 https://doi.org/10.1214/14-EJS929
[51] Shen, W., Tokdar, S.T. and Ghosal, S. (2013). Adaptive Bayesian multivariate density estimation with Dirichlet mixtures. *Biometrika* **100** 623–640. MR3094441 https://doi.org/10.1093/biomet/ast015
[52] Storey, J.D. (2002). A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 479–498. MR1924302 https://doi.org/10.1111/1467-9868.00346
[53] Storey, J.D. and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100** 9440–9445. MR1994856 https://doi.org/10.1073/pnas.1530509100
[54] Tang, Y., Ghosal, S. and Roy, A. (2007). Nonparametric Bayesian estimation of positive false discovery rates. *Biometrics* **63** 1126–1134, 1312. MR2414590 https://doi.org/10.1111/j.1541-0420.2007.00819.x
[55] Tokdar, S.T. (2006). Posterior consistency of Dirichlet location-scale mixture of normals in density estimation and regression. *Sankhyā* **68** 90–110. MR2301566
[56] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2
[57] Walker, S.G. (2007). Sampling the Dirichlet mixture model with slices. *Comm. Statist. Simulation Comput.* **36** 45–54. MR2370888 https://doi.org/10.1080/03610910601096262

[58] Wang, K. and Ghosal, S. (2024). Supplement to "Bayesian inference for k-monotone densities with applications to multiple testing." https://doi.org/10.3150/24-BEJ1778SUPP

[59] Williamson, R.E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.* **23** 189–207. MR0077581

[60] Wu, Y. and Ghosal, S. (2008). Kullback Leibler property of kernel mixture priors in Bayesian density estimation. *Electron. J. Stat.* **2** 298–331. MR2399197 https://doi.org/10.1214/07-EJS130

[61] Wu, Y. and Ghosal, S. (2010). The L_1-consistency of Dirichlet mixtures in multivariate Bayesian density estimation. *J. Multivariate Anal.* **101** 2411–2419. MR2719871 https://doi.org/10.1016/j.jmva.2010.06.012

Bernoulli **31**(2), 2025, 1502–1526
https://doi.org/10.3150/24-BEJ1779

Phase transitions of the maximum likelihood estimators in the *p*-spin Curie-Weiss model

SOMABHA MUKHERJEE[1,a], JAESUNG SON[2,b] and BHASWAR B. BHATTACHARYA[3,c]

[1]*Deapartment of Statistics and Data Science, National University of Singapore, Singapore,* [a]*somabha@nus.edu.sg*
[2]*Department of Statistics, Columbia University, New York, USA,* [b]*js4638@columbia.edu*
[3]*Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, USA,* [c]*bhaswar@wharton.upenn.edu*

In this paper we consider the problem of parameter estimation in the p-spin Curie-Weiss model, for $p \geq 3$. We provide a complete description of the limiting properties of the maximum likelihood (ML) estimators of the inverse temperature and the magnetic field given a single realization from the p-spin Curie-Weiss model, complementing the well-known results in the 2-spin case by Comets and Gidas (1991). Our results unearth various new phase transitions and surprising limit theorems, such as the existence of a 'critical' curve in the parameter space, where the limiting distribution of the ML estimators is a mixture with both continuous and discrete components. The number of mixture components is either two or three, depending on, among other things, the sign of one of the parameters and the parity of p. Another interesting revelation is the existence of certain 'special' points in the parameter space where the ML estimators exhibit a superefficiency phenomenon, converging to a non-Gaussian limiting distribution at rate $N^{3/4}$. Using these results we can obtain asymptotically valid confidence intervals for the inverse temperature and the magnetic field at all points in the parameter space where consistent estimation is possible.

Keywords: Central limit theorems; estimation; Ising models; magnetization; phase transitions; spin-systems; superefficiency

References

Anandkumar, A., Tan, V.Y.F., Huang, F. and Willsky, A.S. (2012). High-dimensional structure estimation in Ising models: Local separation criterion. *Ann. Statist.* **40** 1346–1375. MR3015028 https://doi.org/10.1214/12-AOS1009

Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2015). *Hierarchical Modeling and Analysis for Spatial Data*, 2nd ed. *Monographs on Statistics and Applied Probability* **135**. Boca Raton, FL: CRC Press. MR3362184

Barra, A. (2009). Notes on ferromagnetic p-spin and REM. *Math. Methods Appl. Sci.* **32** 783–797. MR2507932 https://doi.org/10.1002/mma.1065

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. MR0373208

Besag, J. (1975). Statistical analysis of non-lattice data. *J. R. Stat. Soc., Ser. D, Stat.* **24** 179–195. https://doi.org/10.2307/2987782

Bhattacharya, B.B. and Mukherjee, S. (2018). Inference in Ising models. *Bernoulli* **24** 493–525. MR3706767 https://doi.org/10.3150/16-BEJ886

Bresler, G. (2015). Efficiently learning Ising models on arbitrary graphs [extended abstract]. In *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing* 771–782. New York: ACM. MR3388257

Bresler, G. and Nagaraj, D. (2018). Optimal single sample tests for structured versus unstructured network data. In *Conference on Learning Theory (COLT)* 1657–1690.

1350-7265

Cao, Y., Neykov, M. and Liu, H. (2022). High-temperature structure detection in ferromagnets. *Inf. Inference* **11** 55–102. MR4403982 https://doi.org/10.1093/imaiai/iaaa032

Chao, C. and Liu, C. (1968). Approximating discrete probability distributions with dependence trees. *IEEE Trans. Inf. Theory* **14** 462–467. https://doi.org/10.1109/TIT.1968.1054142

Chatterjee, S. (2007). Estimation in spin glasses: A first step. *Ann. Statist.* **35** 1931–1946. MR2363958 https://doi.org/10.1214/009053607000000109

Chatterjee, S. and Shao, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964 https://doi.org/10.1214/10-AAP712

Comets, F. (1992). On consistency of a class of estimators for exponential families of Markov random fields on the lattice. *Ann. Statist.* **20** 455–468. MR1150354 https://doi.org/10.1214/aos/1176348532

Comets, F. and Gidas, B. (1991). Asymptotics of maximum likelihood estimators for the Curie-Weiss model. *Ann. Statist.* **19** 557–578. MR1105836 https://doi.org/10.1214/aos/1176348111

Dagan, Y., Daskalakis, C., Dikkala, N. and Kandiros, A.V. (2021). Learning Ising models from one or multiple samples. In *STOC '21—Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing* 161–168. New York: ACM. MR4398832 https://doi.org/10.1145/3406325.3451074

Daskalakis, C., Dikkala, N. and Kamath, G. (2019). Testing Ising models. *IEEE Trans. Inf. Theory* **65** 6829–6852. MR4030862 https://doi.org/10.1109/TIT.2019.2932255

Daskalakis, C., Dikkala, N. and Panageas, I. (2019). Regression from dependent observations. In *STOC'19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* 881–889. New York: ACM. MR4003392

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)* 3653–3663.

Daskalakis, C., Mossel, E. and Roch, S. (2011). Evolutionary trees and the Ising model on the Bethe lattice: A proof of Steel's conjecture. *Probab. Theory Related Fields* **149** 149–189. MR2773028 https://doi.org/10.1007/s00440-009-0246-2

Dembo, A. and Montanari, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563 https://doi.org/10.1214/09-BJPS027

Ellis, R.S. (2006). *Entropy, Large Deviations, and Statistical Mechanics*. *Classics in Mathematics*. Berlin: Springer. Reprint of the 1985 original. MR2189669 https://doi.org/10.1007/3-540-29060-5

Ellis, R.S. and Newman, C.M. (1978). The statistics of Curie-Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332 https://doi.org/10.1007/BF01012508

Geman, S. and Graffigne, C. (1987). Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)* 1496–1517. Providence, RI: Amer. Math. Soc. MR0934354

Ghosal, P. and Mukherjee, S. (2020). Joint estimation of parameters in Ising model. *Ann. Statist.* **48** 785–810. MR4102676 https://doi.org/10.1214/19-AOS1822

Gidas, B. (1988). Consistency of maximum likelihood and pseudolikelihood estimators for Gibbs distributions. In *Stochastic Differential Systems, Stochastic Control Theory and Applications (Minneapolis, Minn., 1986)*. *IMA Vol. Math. Appl.* **10** 129–145. New York: Springer. MR0934721 https://doi.org/10.1007/978-1-4613-8762-6_10

Green, P.J. and Richardson, S. (2002). Hidden Markov models and disease mapping. *J. Amer. Statist. Assoc.* **97** 1055–1070. MR1951259 https://doi.org/10.1198/016214502388618870

Guyon, X. and Künsch, H.R. (1992). Asymptotic comparison of estimators in the Ising model. In *Stochastic Models, Statistical Methods, and Algorithms in Image Analysis (Rome, 1990)*. *Lect. Notes Stat.* **74** 177–198. Berlin: Springer. MR1188486 https://doi.org/10.1007/978-1-4612-2920-9_12

Hamilton, L., Koehler, F. and Ankur, M. (2017). Information theoretic properties of Markov random fields, and their algorithmic applications. In *Advances in Neural Information Processing Systems (NIPS)* 2463–2472.

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. MR0652033 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., Krchan, 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**. https://doi.org/10.1209/0295-5075/89/40004

Klivans, A.R. and Meka, R. (2017). Learning graphical models using multiplicative weights. In *58th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2017* 343–354. Los Alamitos, CA: IEEE Computer Soc. MR3734242 https://doi.org/10.1109/FOCS.2017.39

Levin, D.A., Luczak, M.J. and Peres, Y. (2010). Glauber dynamics for the mean-field Ising model: Cut-off, critical power law, and metastability. *Probab. Theory Related Fields* **146** 223–265. MR2550363 https://doi.org/10.1007/s00440-008-0189-z

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

Mukherjee, R. and Ray, G. (2022). On testing for parameters in Ising models. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 164–187. MR4375621 https://doi.org/10.1214/21-aihp1157

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. MR4315659 https://doi.org/10.1007/s00220-021-04182-z

Mukherjee, S., Son, J. and Bhattacharya, B.B. (2024). Supplement to "Phase transitions of the maximum likelihood estimators in the p-spin Curie-Weiss model." https://doi.org/10.3150/24-BEJ1779SUPP

Mukherjee, S., Niu, Z., Halder, S., Bhattacharya, B.B. and Michailidis, G. (2022). High dimensional logistic regression under network dependence. arXiv:2110.03200. https://doi.org/10.48550/arXiv.2110.03200

Neykov, M. and Liu, H. (2019). Property testing in high-dimensional Ising models. *Ann. Statist.* **47** 2472–2503. MR3988763 https://doi.org/10.1214/18-AOS1754

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

Pickard, D.K. (1987). Inference for discrete Markov fields: The simplest nontrivial case. *J. Amer. Statist. Assoc.* **82** 90–96. MR0883337

Ravikumar, P., Wainwright, M.J. and Lafferty, J.D. (2010). High-dimensional Ising model selection using ℓ_1-regularized logistic regression. *Ann. Statist.* **38** 1287–1319. MR2662343 https://doi.org/10.1214/09-AOS691

Santhanam, N.P. and Wainwright, M.J. (2012). Information-theoretic limits of selecting binary graphical models in high dimensions. *IEEE Trans. Inf. Theory* **58** 4117–4134. MR2943079 https://doi.org/10.1109/TIT.2012.2191659

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. MR0275824 https://doi.org/10.1063/1.1665583

Turban, L. (2016). One-dimensional Ising model with multispin interactions. *J. Phys. A* **49** 355002, 16. MR3543453 https://doi.org/10.1088/1751-8113/49/35/355002

Xue, L., Zou, H. and Cai, T. (2012). Nonconcave penalized composite conditional likelihood estimation of sparse Ising models. *Ann. Statist.* **40** 1403–1429. MR3015030 https://doi.org/10.1214/12-AOS1017

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 **31**(2), 2025, 1527–1551
https://doi.org/10.3150/24-BEJ1780

A bootstrapped test of covariance stationarity based on orthonormal transformations

JONATHAN B. HILL[a] and TIANQI LI

Dept. of Economics, University of North Carolina, Chapel Hill, USA, [a]*jbhill@email.unc.edu*

We propose a covariance stationarity test for an otherwise dependent and possibly globally non-stationary time series. We work in a generalized version of the new setting in Jin, Wang and Wang (*J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** (2015) 893–922), who exploit Walsh (*Amer. J. Math.* **45** (1923) 5–24) functions in order to compare sub-sample covariances with the full sample counterpart. They impose strict stationarity under the null, only consider linear processes under either hypothesis in order to achieve a parametric estimator for an inverted high dimensional asymptotic covariance matrix, and do not consider any other orthonormal basis. Conversely, we work with a general orthonormal basis under mild conditions that include Haar wavelet and Walsh functions, and we allow for linear or nonlinear processes with possibly non-iid innovations. This is important in macroeconomics and finance where nonlinear feedback and random volatility occur in many settings. We completely sidestep asymptotic covariance matrix estimation and inversion by bootstrapping a max-correlation difference statistic, where the maximum is taken over the correlation lag h and basis generated sub-sample counter k (the number of systematic samples). We achieve a higher feasible rate of increase for the maximum lag and counter $\mathcal{H}_T$ and $\mathcal{K}_T$. Of particular note, our test is capable of detecting breaks in variance, and distant, or very mild, deviations from stationarity.

Keywords: Covariance stationarity; max-correlation test; multiplier bootstrap; orthonormal basis; Walsh functions

References

Ahmed, N. and Rao, K.R. (1975). *Orthogonal Transformations for Digital Signal Processing*. New York: Springer.

Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59** 817–858. MR1106513 https://doi.org/10.2307/2938229

Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. MR2572452 https://doi.org/10.1214/09-AOS707

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

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4

Busetti, F. and Taylor, A.M.R. (2003). Variance shifts, structural breaks, and stationarity tests. *J. Bus. Econom. Statist.* **21** 510–531. MR2008043 https://doi.org/10.1198/073500103288619269

Chang, J., Chen, X. and Wu, M. (2024). Central limit theorems for high dimensional dependent data. *Bernoulli* **30** 712–742. MR4665595 https://doi.org/10.3150/23-bej1614

Chang, J., Jiang, Q. and Shao, X. (2023). Testing the martingale difference hypothesis in high dimension. *J. Econometrics* **235** 972–1000. MR4602899 https://doi.org/10.1016/j.jeconom.2022.09.001

Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. MR3161448 https://doi.org/10.1214/13-AOS1161

Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Testing many moment inequalities. Preprint. Available at arXiv:1312.7614.

Chernozhukov, V., Chetverikov, D. and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields* **162** 47–70. MR3350040 https://doi.org/10.1007/s00440-014-0565-9

Chernozhukov, V., Chetverikov, D. and Kato, K. (2019). Inference on causal and structural parameters using many moment inequalities. *Rev. Econ. Stud.* **86** 1867–1900. MR4009488 https://doi.org/10.1093/restud/rdy065

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620

Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *J. Econometrics* **151** 101–112. MR2559818 https://doi.org/10.1016/j.jeconom.2009.03.002

de Jong, R.M. (1997). Central limit theorems for dependent heterogeneous random variables. *Econometric Theory* **13** 353–367. MR1455176 https://doi.org/10.1017/S0266466600005843

de Jong, R.M. and Davidson, J. (2000). Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica* **68** 407–423. MR1748008 https://doi.org/10.1111/1468-0262.00115

Dedecker, J., Doukhan, P., Lang, G., León R., J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence: With Examples and Applications*. *Lecture Notes in Statistics* **190**. New York: Springer. MR2338725

Dette, H., Preuss, P. and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *J. Amer. Statist. Assoc.* **106** 1113–1124. MR2894768 https://doi.org/10.1198/jasa.2011.tm10811

Doukhan, P. (1994). *Mixing: Properties and Examples*. *Lecture Notes in Statistics* **85**. New York: Springer. MR1312160 https://doi.org/10.1007/978-1-4612-2642-0

Dwivedi, Y. and Subba Rao, S. (2011). A test for second-order stationarity of a time series based on the discrete Fourier transform. *J. Time Series Anal.* **32** 68–91. MR2790673 https://doi.org/10.1111/j.1467-9892.2010.00685.x

Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: Wiley. MR3186556 https://doi.org/10.1002/9780470670057

Giai Gianetto, Q. and Raïssi, H. (2015). Testing instantaneous causality in presence of nonconstant unconditional covariance. *J. Bus. Econom. Statist.* **33** 46–53. MR3303740 https://doi.org/10.1080/07350015.2014.920614

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869. MR1055437

Gray, S.F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *J. Financ. Econ.* **42** 27–62.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Math. Ann.* **69** 331–371. MR1511592 https://doi.org/10.1007/BF01456326

Hendry, D.F. and Massmann, M. (2007). Co-breaking: Recent advances and a synopsis of the literature. *J. Bus. Econom. Statist.* **25** 33–51. MR2338869 https://doi.org/10.1198/073500106000000422

Hill, J.B. and Li, T. (2024). Supplement to "A bootstrapped test of covariance stationarity based on orthonormal transformations." https://doi.org/10.3150/24-BEJ1780SUPP

Hill, J.B. and Motegi, K. (2020). A max-correlation white noise test for weakly dependent time series. *Econometric Theory* **36** 907–960. MR4165431 https://doi.org/10.1017/s0266466619000367

Inglot, T. and Ledwina, T. (2006). Towards data driven selection of a penalty function for data driven Neyman tests. *Linear Algebra Appl.* **417** 124–133. MR2238600 https://doi.org/10.1016/j.laa.2005.10.023

Jentsch, C. and Subba Rao, S. (2015). A test for second order stationarity of a multivariate time series. *J. Econometrics* **185** 124–161. MR3300340 https://doi.org/10.1016/j.jeconom.2014.09.010

Jin, L., Wang, S. and Wang, H. (2015). A new non-parametric stationarity test of time series in the time domain. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 893–922. MR3414133 https://doi.org/10.1111/rssb.12091

Kreiss, J.-P. and Paparoditis, E. (2015). Bootstrapping locally stationary processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 267–290. MR3299408 https://doi.org/10.1111/rssb.12068

Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241. MR1015147 https://doi.org/10.1214/aos/1176347265

Liu, R.Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Ann. Statist.* **16** 1696–1708. MR0964947 https://doi.org/10.1214/aos/1176351062

McLeish, D.L. (1975). A maximal inequality and dependent strong laws. *Ann. Probab.* **3** 829–839. MR0400382 https://doi.org/10.1214/aop/1176996269

Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 879–904. MR3124795 https://doi.org/10.1111/rssb.12015

Nason, G.P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 271–292. MR1749539 https://doi.org/10.1111/1467-9868.00231

Neumann, M.H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25** 38–76. MR1429917 https://doi.org/10.1214/aos/1034276621

Paparoditis, E. (2010). Validating stationarity assumptions in time series analysis by rolling local periodograms. *J. Amer. Statist. Assoc.* **105** 839–851. MR2724865 https://doi.org/10.1198/jasa.2010.tm08243

Perron, P. (2006). Dealing with structural breaks. In *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory* (K. Patterson and T.C. Mills, eds.) **1** 278–352. Basingstoke: Palgrave Macmillan.

Preuß, P., Vetter, M. and Dette, H. (2013). A test for stationarity based on empirical processes. *Bernoulli* **19** 2715–2749. MR3160569 https://doi.org/10.3150/12-BEJ472

Priestley, M.B. and Subba Rao, T. (1969). A test for non-stationarity of time-series. *J. Roy. Statist. Soc. Ser. B* **31** 140–149. MR0269062

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** 43–47. MR0074711 https://doi.org/10.1073/pnas.42.1.43

Shao, X. (2010). The dependent wild bootstrap. *J. Amer. Statist. Assoc.* **105** 218–235. MR2656050 https://doi.org/10.1198/jasa.2009.tm08744

Shao, X. (2011). A bootstrap-assisted spectral test of white noise under unknown dependence. *J. Econometrics* **162** 213–224. MR2795613 https://doi.org/10.1016/j.jeconom.2011.01.001

Stoffer, D.S. (1987). Walsh–Fourier analysis of discrete-valued time series. *J. Time Series Anal.* **8** 449–467. MR0917797 https://doi.org/10.1111/j.1467-9892.1987.tb00008.x

Stoffer, D.S. (1991). Walsh–Fourier analysis and its statistical applications. *J. Amer. Statist. Assoc.* **86** 461–485. MR1137131

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *J. Amer. Statist. Assoc.* **89** 208–218.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. MR3837109 https://doi.org/10.1017/9781108231596

von Sachs, R. and Neumann, M.H. (2000). A wavelet-based test for stationarity. *J. Time Series Anal.* **21** 597–613. MR1794489 https://doi.org/10.1111/1467-9892.00200

Walsh, J.L. (1923). A Closed Set of Normal Orthogonal Functions. *Amer. J. Math.* **45** 5–24. MR1506485 https://doi.org/10.2307/2387224

Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. MR2172215 https://doi.org/10.1073/pnas.0506715102

Zhang, X. and Cheng, G. (2014). Bootstrapping high dimensional time series. Preprint. Available at arXiv:1406.1037.

Zhang, D. and Wu, W.B. (2017). Gaussian approximation for high dimensional time series. *Ann. Statist.* **45** 1895–1919. MR3718156 https://doi.org/10.1214/16-AOS1512

Bernoulli **31**(2), 2025, 1552–1578
https://doi.org/10.3150/24-BEJ1783

Averaging symmetric positive-definite matrices on the space of eigen-decompositions

SUNGKYU JUNG[1,a], BRIAN ROOKS[2,b], DAVID GROISSER[3,c] and ARMIN SCHWARTZMAN[4,d]

[1]*Department of Statistics and Institute for Data Innovation in Science, Seoul National University, Gwanak-gu, Seoul 08826, South Korea,* [a]*sungkyu@snu.ac.kr*
[2]*Statistics & Data Corporation, Tempe, AZ 85288, USA,* [b]*btrooks88@gmail.com*
[3]*Department of Mathematics, University of Florida, Gainesville, FL 32611, USA,* [c]*groisser@ufl.edu*
[4]*Halıcıoğlu Data Science Institute and Division of Biostatistics, University of California, San Diego, La Jolla, CA 92093, USA,* [d]*armins@ucsd.edu*

We study extensions of Fréchet means for random objects in the space $\mathrm{Sym}^+(p)$ of $p \times p$ symmetric positive-definite matrices using the scaling-rotation geometric framework introduced by Jung et al. [*SIAM J. Matrix. Anal. Appl.* **36** (2015) 1180-1201]. The scaling-rotation framework is designed to enjoy a clearer interpretation of the changes in random ellipsoids in terms of scaling and rotation. In this work, we formally define the *scaling-rotation (SR) mean set* to be the set of Fréchet means in $\mathrm{Sym}^+(p)$ with respect to the scaling-rotation distance. Since computing such means requires a difficult optimization, we also define the *partial scaling-rotation (PSR) mean set* lying in the space of eigen-decompositions as a proxy for the SR mean set. The PSR mean set is easier to compute and its projection to $\mathrm{Sym}^+(p)$ often coincides with SR mean set. Minimal conditions are required to ensure that the mean sets are non-empty. Because eigen-decompositions are never unique, neither are PSR means, but we give sufficient conditions for the sample PSR mean to be unique up to the action of a certain finite group. We also establish strong consistency of the sample PSR means as estimators of the population PSR mean set, and a central limit theorem. In an application to multivariate tensor-based morphometry, we demonstrate that a two-group test using the proposed PSR means can have greater power than the two-group test using the usual affine-invariant geometric framework for symmetric positive-definite matrices.

Keywords: Central limit theorem; manifolds; scaling-rotation distance; strong consistency

References

Afsari, B. (2011). Riemannian L^p center of mass: Existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.* **139** 655–673. MR2736346 https://doi.org/10.1090/S0002-9939-2010-10541-5

Afsari, B., Tron, R. and Vidal, R. (2013). On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM J. Control Optim.* **51** 2230–2260. MR3057324 https://doi.org/10.1137/12086282X

Alexander, D.C. (2005). Multiple-fiber reconstruction algorithms for diffusion MRI. *Ann. N.Y. Acad. Sci.* **1064** 113–133. https://doi.org/10.1196/annals.1340.018

Arsigny, V., Fillard, P., Pennec, X. and Ayache, N. (2006/07). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **29** 328–347. MR2288028 https://doi.org/10.1137/050637996

Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F. and Connelly, A. (2005). A rigorous framework for diffusion tensor calculus. *Magn. Reson. Med.* **53** 221–225. https://doi.org/10.1002/mrm.20334

Bhattacharya, A. and Bhattacharya, R. (2012). *Nonparametric Inference on Manifolds. Institute of Mathematical Statistics (IMS) Monographs* **2**. Cambridge: Cambridge Univ. Press. With applications to shape spaces. MR2934285 https://doi.org/10.1017/CBO9781139094764

Bhattacharya, R. and Lin, L. (2017). Omnibus CLTs for Fréchet means and nonparametric inference on non-Euclidean spaces. *Proc. Amer. Math. Soc.* **145** 413–428. MR3565392 https://doi.org/10.1090/proc/13216

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

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

Carmichael, O., Chen, J., Paul, D. and Peng, J. (2013). Diffusion tensor smoothing through weighted Karcher means. *Electron. J. Stat.* **7** 1913–1956. MR3084676 https://doi.org/10.1214/13-EJS825

Collard, A., Bonnabel, S., Phillips, C. and Sepulchre, R. (2014). Anisotropy preserving DTI processing. *Int. J. Comput. Vis.* **107** 58–74.

Dryden, I.L., Koloydenko, A. and Zhou, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3** 1102–1123. MR2750388 https://doi.org/10.1214/09-AOAS249

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

Eltzner, B. and Huckemann, S.F. (2019). A smeary central limit theorem for manifolds with application to high-dimensional spheres. *Ann. Statist.* **47** 3360–3381. MR4025745 https://doi.org/10.1214/18-AOS1781

Eltzner, B., Galaz-García, F., Huckemann, S. and Tuschmann, W. (2021). Stability of the cut locus and a central limit theorem for Fréchet means of Riemannian manifolds. *Proc. Amer. Math. Soc.* **149** 3947–3963. MR4291592 https://doi.org/10.1090/proc/15429

Feragen, A. and Fuster, A. (2017). Geometries and interpolations for symmetric positive definite matrices. In *Modeling, Analysis, and Visualization of Anisotropy*(T. Schultz, E. Özarslan and I. Hotz, eds.). *Math. Vis.* 85–113. Cham: Springer. MR3822438

Feragen, A. and Nye, T. (2020). Statistics on stratified spaces. In *Riemannian Geometric Statistics in Medical Image Analysis* (X. Pennec, S. Sommer and T. Fletcher, eds.) 299–342. London: Elsevier/Academic Press. MR4254219 https://doi.org/10.1016/B978-0-12-814725-2.00016-9

Fletcher, P.T., Lu, C., Pizer, S.M. and Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imag.* **23** 995–1005.

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

Gower, J.C. (1975). Generalized Procrustes analysis. *Psychometrika* **40** 33–51. MR0405725 https://doi.org/10.1007/BF02291478

Groisser, D., Jung, S. and Schwartzman, A. (2017). Geometric foundations for scaling-rotation statistics on symmetric positive definite matrices: Minimal smooth scaling-rotation curves in low dimensions. *Electron. J. Stat.* **11** 1092–1159. MR3631822 https://doi.org/10.1214/17-EJS1250

Groisser, D., Jung, S. and Schwartzman, A. (2021). Uniqueness questions in a scaling-rotation geometry on the space of symmetric positive-definite matrices. *Differential Geom. Appl.* **79** Paper No. 101798, 40. MR4296196 https://doi.org/10.1016/j.difgeo.2021.101798

Groisser, D., Jung, S. and Schwartzman, A. (2023). A genericity property of Fréchet sample means on Riemannian manifolds. Preprint. Available at arXiv:2309.13823.

Hotz, T. and Huckemann, S. (2015). Intrinsic means on the circle: Uniqueness, locus and asymptotics. *Ann. Inst. Statist. Math.* **67** 177–193. MR3297863 https://doi.org/10.1007/s10463-013-0444-7

Huckemann, S. (2011a). Inference on 3D Procrustes means: Tree bole growth, rank deficient diffusion tensors and perturbation models. *Scand. J. Stat.* **38** 424–446. MR2833839 https://doi.org/10.1111/j.1467-9469.2010.00724.x

Huckemann, S.F. (2011b). Intrinsic inference on the mean geodesic of planar shapes and tree discrimination by leaf growth. *Ann. Statist.* **39** 1098–1124. MR2816349 https://doi.org/10.1214/10-AOS862

Huckemann, S.F. and Eltzner, B. (2021). Data analysis on nonstandard spaces. *Wiley Interdiscip. Rev.: Comput. Stat.* **13** Paper No. e1526, 19. MR4242812 https://doi.org/10.1002/wics.1526

Jung, S., Schwartzman, A. and Groisser, D. (2015). Scaling-rotation distance and interpolation of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **36** 1180–1201. MR3379023 https://doi.org/10.1137/140967040

Jung, S., Rooks, B., Groisser, D. and Schwartzman, A. (2024). Supplement to “Averaging symmetric positive-definite matrices on the space of eigen-decompositions.” https://doi.org/10.3150/24-BEJ1783SUPP

Krishnamoorthy, K. and Yu, J. (2004). Modified Nel and van der Merwe test for the multivariate Behrens-Fisher problem. *Statist. Probab. Lett.* **66** 161–169. MR2029731 https://doi.org/10.1016/j.spl.2003.10.012

Lepore, F., Brun, C., Chou, Y.-Y., Chiang, M.-C., Dutton, R., Hayashi, K., Luders, E., Lopez, O., Aizenstein, H., Toga, A.W., Becker, J. and Thompson, P. (2008). Generalized tensor-based morphometry of HIV/AIDS using multivariate statistics on deformation tensors. *IEEE Trans. Med. Imag.* **27** 129–141.

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

Liu, H., Wang, X. and Zhu, J. (2022). Quantiles, ranks and signs in metric spaces. Preprint. Available at arXiv:2209.04090.

Manton, J.H. (2004). A globally convergent numerical algorithm for computing the centre of mass on compact Lie groups. In *Int. Conf. Control Autom. Robot. Vis.* **3** 2211–2216.

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. MR3982195 https://doi.org/10.1007/s13171-018-0130-1

Paquette, N., Shi, J., Wang, Y., Lao, Y., Ceschin, R., Nelson, M., Panigrahy, A. and Lepore, N. (2017). Ventricular shape and relative position abnormalities in preterm neonates. *NeuroImage Clin.* **15** 483–493.

Pennec, X., Fillard, P. and Ayache, N. (2006). A Riemannian framework for tensor computing. *Int. J. Comput. Vis.* **66** 41–66.

Schötz, C. (2019). Convergence rates for the generalized Fréchet mean via the quadruple inequality. *Electron. J. Stat.* **13** 4280–4345. MR4023955 https://doi.org/10.1214/19-EJS1618

Schötz, C. (2022). Strong laws of large numbers for generalizations of Fréchet mean sets. *Statistics* **56** 34–52. MR4392146 https://doi.org/10.1080/02331888.2022.2032063

Schwartzman, A. (2016). Lognormal distributions and geometric averages of symmetric positive definite matrices. *Int. Stat. Rev.* **84** 456–486. MR3580425 https://doi.org/10.1111/insr.12113

Schwartzman, A., Dougherty, R.F. and Taylor, J.E. (2010). Group comparison of eigenvalues and eigenvectors of diffusion tensors. *J. Amer. Statist. Assoc.* **105** 588–599. MR2724844 https://doi.org/10.1198/jasa.2010.ap07291

Wang, Y., Salehian, H., Cheng, G. and Vemuri, B.C. (2014). Tracking on the product manifold of shape and orientation for tractography from diffusion MRI. *Conf Comput Vis Pattern Recognit Workshops* **2014** 3051–3056. https://doi.org/10.1109/CVPR.2014.390

Ziezold, H. (1977). On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. In *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Tech. Univ. Prague, Prague, 1974), Vol. A* 591–602. Dordrecht-Boston, Mass.: Reidel. MR0501230

Bernoulli **31**(2), 2025, 1579–1605
https://doi.org/10.3150/24-BEJ1784

On a near-optimal and efficient algorithm for the sparse pooled data problem

MAX HAHN-KLIMROTH[1,a], REMCO VAN DER HOFSTAD[2,b], NOELA MÜLLER[2,c] and CONNOR RIDDLESDEN[2,d]

[1]*Goethe University Frankfurt, Frankfurt am Main, Germany,* [a]*hahnklim@mathematik.uni-frankfurt.de*
[2]*Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands,* [b]*r.w.v.d.hofstad@tue.nl,* [c]*n.s.muller@tue.nl,* [d]*c.d.riddlesden@tue.nl*

The pooled data problem asks to identify the unknown labels of a set of items from condensed measurements. More precisely, given n items, assume that each item has a label in $\{0,1,\ldots,d\}$, encoded via the ground-truth $\boldsymbol{\sigma}$. We call the pooled data problem sparse if the number of non-zero entries of $\boldsymbol{\sigma}$ scales as $k \sim n^{\theta}$ for $\theta \in (0,1)$. The information that is revealed about $\boldsymbol{\sigma}$ comes from pooled measurements, each indicating how many items of each label are contained in the pool. The most basic question is to design a pooling scheme that uses as few pools as possible, while reconstructing $\boldsymbol{\sigma}$ with high probability. Variants of the problem and its combinatorial ramifications have been studied for at least 35 years. However, the study of the modern question of *efficient* inference of the labels has suggested a statistical-to-computational gap of order $\ln n$ in the minimum number of pools needed for theoretically possible versus efficient inference. In this article, we resolve the question whether this $\ln n$-gap is artificial or of a fundamental nature by the design of an efficient algorithm, called SCIENT, based upon a novel pooling scheme on a number of pools very close to the information-theoretic threshold.

Keywords: Phase transition; pooled data; quantitative group testing; spatial coupling; statistical inference

References

Aldridge, M., Johnson, O. and Scarlett, J. (2019). Group testing: An information theory perspective. *Found. Trends Commun. Inf. Theory* **15** 196–392. https://doi.org/10.1561/0100000099

Bandeira, A.S., El Alaoui, A., Hopkins, S., Schramm, T., Wein, A.S. and Zadik, I. (2022). The Franz-Parisi criterion and computational trade-offs in high dimensional statistics. *Adv. Neural Inf. Process. Syst.* **35** 33831–33844.

Candès, E.J., Romberg, J. and Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **52** 489–509. MR2236170 https://doi.org/10.1109/TIT.2005.862083

Chan, C.L., Che, P.H., Jaggi, S. and Saligrama, V. (2011). Non-adaptive probabilistic group testing with noisy measurements: Near-optimal bounds with efficient algorithms. In *Conf. Commun. Contr. Comp. (Allerton)* 1832–1839. IEEE. https://doi.org/10.1109/Allerton.2011.6120391

Chen, X., Ma, T., Zhang, J. and Zhou, Y. (2019). Optimal design of process flexibility for general production systems. *Oper. Res.* **67** 516–531. MR3943938 https://doi.org/10.1287/opre.2018.1780

Coja-Oghlan, A., Gebhard, O., Hahn-Klimroth, M. and Loick, P. (2021). Optimal group testing. *Combin. Probab. Comput.* **30** 811–848. MR4328351 https://doi.org/10.1017/s096354832100002x

Coja-Oghlan, A., Gebhard, O., Hahn-Klimroth, M., Wein, A.S. and Zadik, I. (2022). Statistical and computational phase transitions in group testing. In *Conf. Learn. Theory (COLT)* **178** 4764–4781.

Djačkov, A.G. (1977). On a search model of false coins. In *Topics in Information Theory (Second Colloq., Keszthely, 1975). Colloq. Math. Soc. János Bolyai* **16** 163–170. Budapest: János Bolyai Math. Soc. MR0505668

Donoho, D.L. and Tanner, J. (2006). Thresholds for the recovery of sparse solutions via l1 minimization. In *Conf. Inform. Sci. and Syst. (CISS)* 202–206. IEEE. https://doi.org/10.1109/CISS.2006.286462

1350-7265

El Alaoui, A., Ramdas, A., Krzakala, F., Zdeborová, L. and Jordan, M.I. (2019a). Decoding from pooled data: Phase transitions of message passing. *IEEE Trans. Inf. Theory* **65** 572–585. MR3901010 https://doi.org/10.1109/TIT.2018.2855698

El Alaoui, A., Ramdas, A., Krzakala, F., Zdeborová, L. and Jordan, M.I. (2019b). Decoding from pooled data: Sharp information-theoretic bounds. *SIAM J. Math. Data Sci.* **1** 161–188. MR3949705 https://doi.org/10.1137/18M1183339

Feige, U. and Lellouche, A. (2020). Quantitative group testing and the rank of random matrices. CoRR. Available at arXiv:2006.09074.

Gebhard, O., Hahn-Klimroth, M., Kaaser, D. and Loick, P. (2023). Information-theoretic and algorithmic aspects of parallel and distributed reconstruction from pooled data. *Parallel Distrib. Comput. Ser.* **180** 104718. https://doi.org/10.1016/j.jpdc.2023.104718

Grebinski, V. and Kucherov, G. (2000). Optimal reconstruction of graphs under the additive model. *Algorithmica* **28** 104–124. MR1776240 https://doi.org/10.1007/s004530010033

Hahn-Klimroth, M. and Müller, N. (2022). Near optimal efficient decoding from pooled data. In *Conf. Learn. Theory (COLT)* **178** 3395–3409.

Janson, S. (2002). On concentration of probability. In *Contemporary Combinatorics*. *Bolyai Soc. Math. Stud.* **10** 289–301. Budapest: János Bolyai Math. Soc. MR1919574

Janson, S., Łuczak, T. and Rucinski, A. (2000). *Random Graphs*. *Wiley-Interscience Series in Discrete Mathematics and Optimization*. New York: Wiley Interscience. MR1782847 https://doi.org/10.1002/9781118032718

Jiménez Felström, A. and Zigangirov, K.S. (1999). Time-varying periodic convolutional codes with low-density parity-check matrix. *IEEE Trans. Inf. Theory* **45** 2181–2191. MR1720679 https://doi.org/10.1109/18.782171

Johnson, O., Aldridge, M. and Scarlett, J. (2019). Performance of group testing algorithms with near-constant tests per item. *IEEE Trans. Inf. Theory* **65** 707–723. MR3904909 https://doi.org/10.1109/TIT.2018.2861772

Karimi, E., Kazemi, F., Heidarzadeh, A., Narayanan, K.R. and Sprintson, A. (2019a). Non-adaptive quantitative group testing using irregular sparse graph codes. In *Conf. Commun. Contr. Comp. (Allerton)*. 608–614. IEEE. https://doi.org/10.1109/ALLERTON.2019.8919896

Karimi, E., Kazemi, F., Heidarzadeh, A., Narayanan, K.R. and Sprintson, A. (2019b). Sparse graph codes for non-adaptive quantitative group testing. In *IEEE Inform. Theory Workshop (ITW)* 1–5. https://doi.org/10.1109/ITW44776.2019.8989403

Kudekar, S. and Pfister, H.D. (2010). The effect of spatial coupling on compressive sensing. In *Conf. Commun. Contr. Comp. (Allerton)*. 347–353. IEEE. https://doi.org/10.1109/ALLERTON.2010.5706927

Kudekar, S., Richardson, T. and Urbanke, R.L. (2013). Spatially coupled ensembles universally achieve capacity under belief propagation. *IEEE Trans. Inf. Theory* **59** 7761–7813. MR3142263 https://doi.org/10.1109/TIT.2013.2280915

Liang, W. and Zou, J. (2021). Neural group testing to accelerate deep learning. In *IEEE Int. Symp. Inform. Theory (ISIT)* 958–963. https://doi.org/10.1109/ISIT45174.2021.9518038

Martins, J.P., Santos, R. and Sousa, R. (2014). Testing the maximum by the mean in quantitative group tests. In *New Advances in Statistical Modeling and Applications*. *Stud. Theor. Appl. Stat. Sel. Papers Stat. Soc.* 55–63. Cham: Springer. MR3234684

Mazumdar, A. and Pal, S. (2022). Support recovery in universal one-bit compressed sensing. In *13th Innovations in Theoretical Computer Science Conference*. *LIPIcs. Leibniz Int. Proc. Inform.* **215** Art. No. 106, 20 pp. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4459477

Scarlett, J. and Cevher, V. (2017). Phase transitions in the pooled data problem. *Adv. Neural Inf. Process. Syst.* **30** 376–384.

Sham, P., Bader, J.S., Craig, I., O'Donovan, M. and Owen, M. (2002). DNA pooling: A tool for large-scale association studies. *Nat. Rev. Genet.* **3** 862–871. https://doi.org/10.1038/nrg930

Shapiro, H.S. (1960). Problem e 1399. *Amer. Math. Monthly* **67** 82.

Soleymani, M., Mahdavifar, H. and Javidi, T. (2023). Non-adaptive quantitative group testing via Plotkin-type constructions. In *IEEE Int. Symp. Inform. Theory (ISIT)* 1854–1859. https://doi.org/10.1109/ISIT54713.2023.10206896

Tan, N., Cobo, P.P., Scarlett, J. and Venkataramanan, R. (2024). Approximate message passing with rigorous guarantees for pooled data and quantitative group testing. Preprint. Available at arXiv:2309.15507.

Wang, C., Zhao, Q. and Chuah, C.N. (2015). Group testing under sum observations for heavy hitter detection. In *Inform. Theory Appl. Workshop (ITA)* 149–153. IEEE. https://doi.org/10.1109/ITA.2015.7308980

Wein, A.S. (2021). Optimal low-degree hardness of maximum independent set. *Math. Stat. Learn.* **4** 221–251. MR4383734 https://doi.org/10.4171/MSL/25

Zhou, Y., Porwal, U., Zhang, C., Ngo, H.Q., Nguyen, X., Ré, C. and Govindaraju, V. (2014). Parallel feature selection inspired by group testing. *Adv. Neural Inf. Process. Syst.* **27** 3554–3562.

Bernoulli **31**(2), 2025, 1606–1632
https://doi.org/10.3150/24-BEJ1785

Heat content for Gaussian processes: Small-time asymptotic analysis

KEI KOBAYASHI[1,a] and HYUNCHUL PARK[2,3,b]

[1] *Department of Mathematics, Fordham University, New York, NY, USA,* [a] *kkobayashi5@fordham.edu*
[2] *Department of Mathematics, State University of New York at New Paltz, New Paltz, NY, USA*
[3] *School of Mathematics, Korea Institute for Advanced Study, Seoul, South Korea,* [b] *parkh@newpaltz.edu*

This paper establishes the small-time asymptotic behaviors of the regular heat content and spectral heat content for general Gaussian processes in both one-dimensional and multi-dimensional settings, where the boundary of the underlying domain satisfies some smoothness condition. For the amount of heat loss associated with the spectral heat content, the exact asymptotic behavior with the rate function being the expected supremum process is obtained, whereas for the regular heat content, the exact asymptotic behavior is described in terms of the standard deviation function.

Keywords: Asymptotic behavior; Gaussian process; regular heat content; spectral heat content

References

Acuña Valverde, L. (2016). On the one dimensional spectral heat content for stable processes. *J. Math. Anal. Appl.* **441** 11–24. MR3488045 https://doi.org/10.1016/j.jmaa.2016.03.086

Acuña Valverde, L. (2017). Heat content for stable processes in domains of $\mathbb{R}^d$. *J. Geom. Anal.* **27** 492–524. MR3606559 https://doi.org/10.1007/s12220-016-9688-9

Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. *Springer Monographs in Mathematics*. New York: Springer. MR2319516

Aikawa, H., Kilpeläinen, T., Shanmugalingam, N. and Zhong, X. (2007). Boundary Harnack principle for p-harmonic functions in smooth Euclidean domains. *Potential Anal.* **26** 281–301. MR2286038 https://doi.org/10.1007/s11118-006-9036-y

Ascione, G., Mishura, Y. and Pirozzi, E. (2020). Time-changed fractional Ornstein-Uhlenbeck process. *Fract. Calc. Appl. Anal.* **23** 450–483. MR4098657 https://doi.org/10.1515/fca-2020-0022

Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1700749 https://doi.org/10.1002/9780470316962

Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR1015093

Borovkov, K., Mishura, Y., Novikov, A. and Zhitlukhin, M. (2017). Bounds for expected maxima of Gaussian processes and their discrete approximations. *Stochastics* **89** 21–37. MR3574693 https://doi.org/10.1080/17442508.2015.1126282

Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.* **8** no. 3, 14. MR1961165 https://doi.org/10.1214/EJP.v8-125

Durrett, R. (2019). *Probability—Theory and Examples*. *Cambridge Series in Statistical and Probabilistic Mathematics* **49**. Cambridge: Cambridge Univ. Press. Fifth edition of [MR1068527]. MR3930614 https://doi.org/10.1017/9781108591034

Grzywny, T., Park, H. and Song, R. (2019). Spectral heat content for Lévy processes. *Math. Nachr.* **292** 805–825. MR3937619 https://doi.org/10.1002/mana.201800035

Houdré, C. and Villa, J. (2003). An example of infinite dimensional quasi-helix. In *Stochastic Models (Mexico City, 2002)*. *Contemp. Math.* **336** 195–201. Providence, RI: Amer. Math. Soc. MR2037165 https://doi.org/10.1090/conm/336/06034

1350-7265

Jansen, S. (1997). *Gaussian Hilbert Spaces*. Cambridge: Cambridge Univ. Press. https://doi.org/10.1017/CBO9780511526169

Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

Kobayashi, K. and Park, H. (2022). Large-time and small-time behaviors of the spectral heat content for time-changed stable processes. *Electron. Commun. Probab.* **27** Paper No. 39, 11. MR4478125 https://doi.org/10.1086/287711

Kobayashi, K. and Park, H. (2023). Spectral heat content for time-changed killed Brownian motions. *J. Theoret. Probab.* **36** 1148–1180. MR4591869 https://doi.org/10.1007/s10959-022-01188-8

Kobayashi, K. and Park, H. (2024). A unified approach to the small-time behavior of the spectral heat content for isotropic Lévy processes. *Stochastic Process. Appl.* **171** Paper No. 104331, 11. MR4712387 https://doi.org/10.1016/j.spa.2024.104331

Lifshits, M. (2012). *Lectures on Gaussian Processes*. *SpringerBriefs in Mathematics*. Heidelberg: Springer. MR3024389 https://doi.org/10.1007/978-3-642-24939-6

Marcus, M.B. and Shepp, L.A. (1972). Sample behavior of Gaussian processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory* 423–441. Berkeley, CA: Univ. California Press. MR0402896

Miranda, M. Jr., Pallara, D., Paronetto, F. and Preunkert, M. (2007). Short-time heat flow and functions of bounded variation in $\mathbf{R}^N$. *Ann. Fac. Sci. Toulouse Math. (6)* **16** 125–145. MR2325595

Park, H. (2022). Higher order terms of the spectral heat content for killed subordinate and subordinate killed Brownian motions related to symmetric α-stable processes in $\mathbb{R}$. *Potential Anal.* **57** 283–303. MR4443127 https://doi.org/10.1007/s11118-021-09916-1

Park, H. and Song, R. (2019). Small time asymptotics of spectral heat contents for subordinate killed Brownian motions related to isotropic α-stable processes. *Bull. Lond. Math. Soc.* **51** 371–384. MR3937594 https://doi.org/10.1112/blms.12235

Park, H. and Song, R. (2022). Spectral heat content for α-stable processes in $C^{1,1}$ open sets. *Electron. J. Probab.* **27** Paper No. 22, 19. MR4379201 https://doi.org/10.1214/22-ejp752

Samorodnitsky, G. (1991). Probability tails of Gaussian extrema. *Stochastic Process. Appl.* **38** 55–84. MR1116304 https://doi.org/10.1016/0304-4149(91)90072-K

van den Berg, M. and Davies, E.B. (1989). Heat flow out of regions in $\mathbf{R}^m$. *Math. Z.* **202** 463–482. MR1022816 https://doi.org/10.1007/BF01221585

van den Berg, M. and Gilkey, P.B. (1994). Heat content asymptotics of a Riemannian manifold with boundary. *J. Funct. Anal.* **120** 48–71. MR1262245 https://doi.org/10.1006/jfan.1994.1022

van den Berg, M. and Gittins, K. (2015). Uniform bounds for the heat content of open sets in Euclidean space. *Differential Geom. Appl.* **40** 67–85. MR3333096 https://doi.org/10.1016/j.difgeo.2015.01.010

van den Berg, M. and Le Gall, J.-F. (1994). Mean curvature and the heat equation. *Math. Z.* **215** 437–464. MR1262526 https://doi.org/10.1007/BF02571723

van den Berg, M. and Srisatkunarajah, S. (1990). Heat flow and Brownian motion for a region in $\mathbf{R}^2$ with a polygonal boundary. *Probab. Theory Related Fields* **86** 41–52. MR1061947 https://doi.org/10.1007/BF01207512

Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. *Springer Series in Operations Research*. New York: Springer. MR1876437

Bernoulli **31**(2), 2025, 1633–1663
https://doi.org/10.3150/24-BEJ1786

Random line graphs and edge-attributed network inference

ZACHARY LUBBERTS[1,a], AVANTI ATHREYA[2,b], YOUNGSER PARK[3,d] and CAREY E. PRIEBE[2,c]

[1]*Department of Statistics, University of Virginia, Charlottesville, VA, USA,* [a]*zlubberts@virginia.edu*
[2]*Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, USA,* [b]*dathrey1@jhu.edu,* [c]*cep@jhu.edu*
[3]*Center for Imaging Science, Johns Hopkins University, Baltimore, MD, USA,* [d]*youngser@jhu.edu*

We extend the latent position random graph model to the *line graph of a random graph*, which is formed by creating a vertex for each edge in the original random graph, and connecting each pair of edges incident to a common vertex in the original graph. We prove concentration inequalities for the spectrum of general line graphs, as well as limiting distribution results for the largest eigenvalue and the empirical spectral distribution for the line graph of an Erdős-Rényi random graph. For the stochastic blockmodel, we establish that although naive spectral decompositions can fail to extract necessary signal for edge clustering, there exist signal-preserving singular subspaces of the line graph that can be recovered through a carefully-chosen projection. Moreover, we can consistently estimate edge latent positions in a random line graph, even though such graphs are of a random size, typically have high rank, and possess no spectral gap. Our results demonstrate that the line graph of a stochastic block model exhibits underlying block structure, and in simulations, we synthesize and test our methods against several commonly-used techniques, including tensor decompositions, for cluster recovery and edge covariate inference. By incorporating information encoded in both vertices and edges, the random line graph improves network inference.

Keywords: Edge attributes; edge covariates; network inference; random line graphs; spectral decomposition

References

[1] Airoldi, E.M., Costa, T.B. and Chan, S.H. (2013). Stochastic blockmodel approximation of a graphon: Theory and consistent estimation. *Adv. Neural Inf. Process. Syst.* **26** 692–700.

[2] Arroyo, J., Athreya, A., Cape, J., Chen, G., Priebe, C.E. and Vogelstein, J.T. (2021). Inference for multiple heterogeneous networks with a common invariant subspace. *J. Mach. Learn. Res.* **22** Paper No. 142, 49. MR4318498

[3] Athreya, A., Priebe, C.E., Tang, M., Lyzinski, V., Marchette, D.J. and Sussman, D.L. (2016). A limit theorem for scaled eigenvectors of random dot product graphs. *Sankhya A* **78** 1–18. MR3494576 https://doi.org/10.1007/s13171-015-0071-x

[4] Binkiewicz, N., Vogelstein, J.T. and Rohe, K. (2017). Covariate-assisted spectral clustering. *Biometrika* **104** 361–377. MR3698259 https://doi.org/10.1093/biomet/asx008

[5] Bullmore, E. and Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10** 186–198.

[6] Diaconis, P. and Janson, S. (2008). Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)* **28** 33–61. MR2463439

[7] Ding, X. and Jiang, T. (2010). Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.* **20** 2086–2117. MR2759729 https://doi.org/10.1214/10-AAP677

[8] Durante, D. and Dunson, D.B. (2018). Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* **13** 29–58. MR3737942 https://doi.org/10.1214/16-BA1030

[9] Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **5** 17–61. MR0125031

1350-7265

[10] Evans, T. and Lambiotte, R. (2009). Line graphs, link partitions, and overlapping communities. *Phys. Rev. E* **80** 016105.

[11] Füredi, Z. and Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241.

[12] Gross, J.L. and Yellen, J. (2006). *Graph Theory and Its Applications*, 2nd ed. *Discrete Mathematics and Its Applications (Boca Raton)*. Boca Raton, FL: CRC Press/CRC. MR2181153

[13] Hoff, P.D., Raftery, A.E. and Handcock, M.S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. MR1951262 https://doi.org/10.1198/016214502388618906

[14] Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183 https://doi.org/10.1017/CBO9780511810817

[15] Karoński, M., Scheinerman, E.R. and Singer-Cohen, K.B. (1999). On random intersection graphs: The subgraph problem. *Combin. Probab. Comput.* **8** 131–159.

[16] Lehot, P.G.H. (1974). An optimal algorithm to detect a line graph and output its root graph. *J. Assoc. Comput. Mach.* **21** 569–575. MR0347690 https://doi.org/10.1145/321850.321853

[17] Lelarge, M., Massoulié, L. and Xu, J. (2015). Reconstruction in the labelled stochastic block model. *IEEE Trans. Netw. Sci. Eng.* **2** 152–163. MR3453283 https://doi.org/10.1109/TNSE.2015.2490580

[18] Liu, D., Trajanovski, S. and Van Mieghem, P. (2013). Random line graphs and a linear law for assortativity. *Phys. Rev. E* **87** 012816.

[19] Lubberts, Z., Athreya, A., Park, Y. and Priebe, C.E. (2024). Supplement to "Random line graphs and edge-attributed network inference." https://doi.org/10.3150/24-BEJ1786SUPP

[20] Lyzinski, V., Sussman, D.L., Tang, M., Athreya, A. and Priebe, C.E. (2014). Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electron. J. Stat.* **8** 2905–2922. MR3299126 https://doi.org/10.1214/14-EJS978

[21] Mańka-Krasoń, A., Mwijage, A. and Kułakowski, K. (2010). Clustering in random line graphs. *Comput. Phys. Commun.* **181** 118–121. MR2575193 https://doi.org/10.1016/j.cpc.2009.09.010

[22] Paul, S. and Chen, Y. (2020). Spectral and matrix factorization methods for consistent community detection in multi-layer networks. *Ann. Statist.* **48** 230–250. MR4065160 https://doi.org/10.1214/18-AOS1800

[23] Priebe, C.E., Park, Y., Tang, M., Athreya, A., Lyzinski, V., Vogelstein, J.T., Qin, Y., Cocanougher, B., Eichler, K., Zlatic, M. and Cardona, A. (2017). Semiparametric spectral modeling of the Drosophila connectome. arXiv preprint, arXiv:1705.03297.

[24] Skiena, S. (1990). *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*. Redwood City, CA: Addison-Wesley Company. With programs by the author and Anil Bhansali. MR1061378

[25] Smith, A.L., Asta, D.M. and Calder, C.A. (2019). The geometry of continuous latent space models for network data. *Statist. Sci.* **34** 428–453. MR4017522 https://doi.org/10.1214/19-STS702

[26] Sussman, D.L., Tang, M., Fishkind, D.E. and Priebe, C.E. (2012). A consistent adjacency spectral embedding for stochastic blockmodel graphs. *J. Amer. Statist. Assoc.* **107** 1119–1128. MR3010899 https://doi.org/10.1080/01621459.2012.699795

[27] Sussman, D.L., Tang, M. and Priebe, C.E. (2014). Consistent latent position estimation and vertex classification for random dot product graphs. *IEEE Trans. Pattern Anal. Mach. Intell.* **36** 48–57. https://doi.org/10.1109/TPAMI.2013.135

[28] Takemura, S.Y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K. et al. (2013). A visual motion detection circuit suggested by Drosophila connectomics. *Nature* **500** 175–181.

[29] Tang, M., Athreya, A., Sussman, D.L., Lyzinski, V., Park, Y. and Priebe, C.E. (2017). A semiparametric two-sample hypothesis testing problem for random graphs. *J. Comput. Graph. Statist.* **26** 344–354. MR3640191 https://doi.org/10.1080/10618600.2016.1193505

[30] Tang, M., Athreya, A., Sussman, D.L., Lyzinski, V. and Priebe, C.E. (2017). A nonparametric two-sample hypothesis testing problem for random graphs. *Bernoulli* **23** 1599–1630. MR3624872 https://doi.org/10.3150/15-BEJ789

[31] Tropp, J.A. (2015). An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.* **8** 1–230.

[32] Whitney, H. (1932). Congruent Graphs and the Connectivity of Graphs. *Amer. J. Math.* **54** 150–168. MR1506881 https://doi.org/10.2307/2371086

[33] Wolfe, P.J. and Olhede, S.C. Nonparametric graphon estimation. arXiv preprint, arXiv:1309.5936.
[34] Yu, Y., Wang, T. and Samworth, R.J. (2015). A useful variant of the Davis-Kahan theorem for statisticians. *Biometrika* **102** 315–323. MR3371006 https://doi.org/10.1093/biomet/asv008
[35] Zhang, A. and Xia, D. (2018). Tensor SVD: Statistical and computational limits. *IEEE Trans. Inf. Theory* **64** 7311–7338. MR3876445 https://doi.org/10.1109/TIT.2018.2841377

Bernoulli **31**(2), 2025, 1664–1691
https://doi.org/10.3150/24-BEJ1787

Sharper dimension-free bounds on the Frobenius distance between sample covariance and its expectation

NIKITA PUCHKIN[1,a], FEDOR NOSKOV[1,b] and VLADIMIR SPOKOINY[2,1,3,c]

[1]*HSE University, Moscow, Russian Federation,* [a]*npuchkin@hse.ru,* [b]*fnoskov@hse.ru*
[2]*Weierstrass Institute and Humboldt University, Berlin, Germany,* [c]*spokoiny@wias-berlin.de*
[3]*Institute for Information Transmission Problems, Moscow, Russian Federation*

We study properties of a sample covariance estimate $\widehat{\Sigma}$ given a finite sample of n i.i.d. centered random elements in $\mathbb{R}^d$ with the covariance matrix Σ. We derive dimension-free bounds on the squared Frobenius norm of $(\widehat{\Sigma} - \Sigma)$ under reasonable assumptions. For instance, we show that $\|\widehat{\Sigma} - \Sigma\|_{\mathrm{F}}^2$ differs from its expectation by at most $\mathcal{O}(\mathrm{Tr}(\Sigma^2)/n)$ with overwhelming probability, which is a significant improvement over the existing results. This allows us to establish the concentration phenomenon for the squared Frobenius distance between the covariance and its empirical counterpart in the case of moderately large effective rank of Σ.

Keywords: Covariance estimation; cumulant generating function; effective rank; Hanson-Wright inequality; Orlicz norm

References

Abdalla, P. (2024). Covariance estimation under missing observations and $L_4 - L_2$ moment equivalence. *Electron. J. Stat.* **18** 2665–2686. MR4767501 https://doi.org/10.1214/24-ejs2264

Abdalla, P. and Zhivotovskiy, N. (2022). Covariance estimation: Optimal dimension-free guarantees for adversarial corruption and heavy tails. Preprint, arXiv:2205.08494.

Adamczak, R. (2015). A note on the Hanson-Wright inequality for random vectors with dependencies. *Electron. Commun. Probab.* **20** 72. MR3407216 https://doi.org/10.1214/ECP.v20-3829

Adamczak, R. and Wolff, P. (2015). Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. *Probab. Theory Related Fields* **162** 531–586. MR3383337 https://doi.org/10.1007/s00440-014-0579-3

Adamczak, R., Litvak, A.E., Pajor, A. and Tomczak-Jaegermann, N. (2010). Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles. *J. Amer. Math. Soc.* **23** 535–561. MR2601042 https://doi.org/10.1090/S0894-0347-09-00650-X

Adamczak, R., Litvak, A.E., Pajor, A. and Tomczak-Jaegermann, N. (2011). Sharp bounds on the rate of convergence of the empirical covariance matrix. *C. R. Math. Acad. Sci. Paris* **349** 195–200. MR2769907 https://doi.org/10.1016/j.crma.2010.12.014

Bai, J. and Shi, S. (2011). Estimating high dimensional covariance matrices and its applications. *Ann. Econ. Financ.* **12** 199–215.

Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9

Bickel, P.J. and Levina, E. (2008a). Covariance regularization by thresholding. *Ann. Statist.* **36** 2577–2604. MR2485008 https://doi.org/10.1214/08-AOS600

Bickel, P.J. and Levina, E. (2008b). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227. MR2387969 https://doi.org/10.1214/009053607000000758

Bourgain, J. (1999). Random points in isotropic convex sets. In *Convex Geometric Analysis (Berkeley, CA, 1996). Math. Sci. Res. Inst. Publ.* **34** 53–58. Cambridge: Cambridge Univ. Press. MR1665576

Bunea, F. and Xiao, L. (2015). On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA. *Bernoulli* **21** 1200–1230. MR3338661 https://doi.org/10.3150/14-BEJ602

Cai, T.T., Ren, Z. and Zhou, H.H. (2013). Optimal rates of convergence for estimating Toeplitz covariance matrices. *Probab. Theory Related Fields* **156** 101–143. MR3055254 https://doi.org/10.1007/s00440-012-0422-7

Cai, T.T., Ren, Z. and Zhou, H.H. (2016). Estimating structured high-dimensional covariance and precision matrices: Optimal rates and adaptive estimation. *Electron. J. Stat.* **10** 1–59. MR3466172 https://doi.org/10.1214/15-EJS1081

Cai, T.T., Zhang, C.-H. and Zhou, H.H. (2010). Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.* **38** 2118–2144. MR2676885 https://doi.org/10.1214/09-AOS752

Cai, T.T. and Zhou, H.H. (2012). Optimal rates of convergence for sparse covariance matrix estimation. *Ann. Statist.* **40** 2389–2420. MR3097607 https://doi.org/10.1214/12-AOS998

Chen, S.X., Zhang, L.-X. and Zhong, P.-S. (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105** 810–819. MR2724863 https://doi.org/10.1198/jasa.2010.tm09560

Dahmen, J., Keysers, D., Pitz, M. and Ney, H. (2000). Structured covariance matrices for statistical image object recognition. In *Mustererkennung 2000, 22. DAGM-Symposium* 99–106. Springer, Berlin.

Fan, J., Rigollet, P. and Wang, W. (2015). Estimation of functionals of sparse covariance matrices. *Ann. Statist.* **43** 2706–2737. MR3405609 https://doi.org/10.1214/15-AOS1357

Giannopoulos, A., Hartzoulaki, M. and Tsolomitis, A. (2005). Random points in isotropic unconditional convex bodies. *J. Lond. Math. Soc. (2)* **72** 779–798. MR2190337 https://doi.org/10.1112/S0024610705006897

Götze, F., Sambale, H. and Sinulis, A. (2021). Concentration inequalities for polynomials in α-sub-exponential random variables. *Electron. J. Probab.* **26** 48. MR4247973 https://doi.org/10.1214/21-ejp606

Haghighatshoar, S. and Caire, G. (2018). Low-complexity massive MIMO subspace estimation and tracking from low-dimensional projections. *IEEE Trans. Signal Process.* **66** 1832–1844. MR3797838 https://doi.org/10.1109/TSP.2018.2795560

Han, Q. (2022). Exact spectral norm error of sample covariance. Preprint, arXiv:2207.13594.

Han, Y. and Wu, W.B. (2020). Test for high dimensional covariance matrices. *Ann. Statist.* **48** 3565–3588. MR4185820 https://doi.org/10.1214/20-AOS1943

Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864 https://doi.org/10.1214/aoms/1177693335

Hero, A. and Rajaratnam, B. (2012). Hub discovery in partial correlation graphs. *IEEE Trans. Inf. Theory* **58** 6064–6078. MR2966078 https://doi.org/10.1109/TIT.2012.2200825

Holtz, M. (2011). *Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance*. *Lecture Notes in Computational Science and Engineering* **77**. Berlin: Springer. MR2743492 https://doi.org/10.1007/978-3-642-16004-2

Hsu, D., Kakade, S.M. and Zhang, T. (2012). A tail inequality for quadratic forms of subgaussian random vectors. *Electron. Commun. Probab.* **17** 52. MR2994877 https://doi.org/10.1214/ECP.v17-2079

Kannan, R., Lovász, L. and Simonovits, M. (1997). Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures Algorithms* **11** 1–50. MR1608200 https://doi.org/10.1002/(SICI)1098-2418(199708)11:1<1::AID-RSA1>3.0.CO;2-X

Klochkov, Y. and Zhivotovskiy, N. (2020). Uniform Hanson-Wright type concentration inequalities for unbounded entries via the entropy method. *Electron. J. Probab.* **25** 22. MR4073683 https://doi.org/10.1214/20-ejp422

Koltchinskii, V. and Lounici, K. (2017). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23** 110–133. MR3556768 https://doi.org/10.3150/15-BEJ730

Krim, H. and Viberg, M. (1996). Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* **13** 67–94.

Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** 1302–1338. MR1805785 https://doi.org/10.1214/aos/1015957395

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. MR1926169 https://doi.org/10.1214/aos/1031689018

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Finance* **10** 603–621.

Leng, C. and Pan, G. (2018). Covariance estimation via sparse Kronecker structures. *Bernoulli* **24** 3833–3863. MR3788190 https://doi.org/10.3150/17-BEJ980

Li, J. and Chen, S.X. (2012). Two sample tests for high-dimensional covariance matrices. *Ann. Statist.* **40** 908–940. MR2985938 https://doi.org/10.1214/12-AOS993

Lounici, K. (2014). High-dimensional covariance matrix estimation with missing observations. *Bernoulli* **20** 1029–1058. MR3217437 https://doi.org/10.3150/12-BEJ487

Mendelson, S. and Zhivotovskiy, N. (2020). Robust covariance estimation under L_4-L_2 norm equivalence. *Ann. Statist.* **48** 1648–1664. MR4124338 https://doi.org/10.1214/19-AOS1862

Minasyan, A. and Zhivotovskiy, N. (2023). Statistically optimal robust mean and covariance estimation for anisotropic Gaussians. Preprint, arXiv:2301.09024.

Oliveira, R.I. (2016). The lower tail of random quadratic forms with applications to ordinary least squares. *Probab. Theory Related Fields* **166** 1175–1194. MR3568047 https://doi.org/10.1007/s00440-016-0738-9

Pacreau, G. and Lounici, K. (2023). Robust covariance estimation with missing values and cell-wise contamination. In *Advances in Neural Information Processing Systems* **36** 72124–72136. Curran Associates.

Paouris, G. (2006). Concentration of mass on convex bodies. *Geom. Funct. Anal.* **16** 1021–1049. MR2276533 https://doi.org/10.1007/s00039-006-0584-5

Puchkin, N., Noskov, F. and Spokoiny, V. (2024). Supplement to "Sharper dimension-free bounds on the Frobenius distance between sample covariance and its expectation." https://doi.org/10.3150/24-BEJ1787SUPP

Rudelson, M. (1999). Random vectors in the isotropic position. *J. Funct. Anal.* **164** 60–72. MR1694526 https://doi.org/10.1006/jfan.1998.3384

Rudelson, M. and Vershynin, R. (2013). Hanson-Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.* **18** 82. MR3125258 https://doi.org/10.1214/ECP.v18-2865

Sambale, H. (2023). Some notes on concentration for α-subexponential random variables. In *High Dimensional Probability IX—the Ethereal Volume*. *Progress in Probability* **80** 167–192. Cham: Birkhäuser/Springer. MR4633263 https://doi.org/10.1007/978-3-031-26979-0_7

Schäfer, J. and Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* **4** 32. MR2183942 https://doi.org/10.2202/1544-6115.1175

Schudy, W. and Sviridenko, M. (2012). Concentration and moment inequalities for polynomials of independent random variables. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* 437–446. New York: ACM. MR3205227

Spokoiny, V. (2023). Concentration of a high dimensional sub-gaussian vector. Preprint, arXiv:2305.07885.

Srivastava, N. and Vershynin, R. (2013). Covariance estimation for distributions with $2+\varepsilon$ moments. *Ann. Probab.* **41** 3081–3111. MR3127875 https://doi.org/10.1214/12-AOP760

Tikhomirov, K. (2018). Sample covariance matrices of heavy-tailed distributions. *Int. Math. Res. Not. IMRN* **20** 6254–6289. MR3872323 https://doi.org/10.1093/imrn/rnx067

Tropp, J.A. (2012). User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12** 389–434. MR2946459 https://doi.org/10.1007/s10208-011-9099-z

Tsiligkaridis, T. and Hero, A.O. (2013). Covariance estimation in high dimensions via Kronecker product expansions. *IEEE Trans. Signal Process.* **61** 5347–5360.

Vershynin, R. (2012a). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170

Vershynin, R. (2012b). How close is the sample covariance matrix to the actual covariance matrix? *J. Theoret. Probab.* **25** 655–686. MR2956207 https://doi.org/10.1007/s10959-010-0338-z

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. With a foreword by Sara van de Geer. MR3837109 https://doi.org/10.1017/9781108231596

Xiao, H. and Wu, W.B. (2012). Covariance matrix estimation for stationary time series. *Ann. Statist.* **40** 466–493. MR3014314 https://doi.org/10.1214/11-AOS967

Xie, J. and Bentler, P.M. (2003). Covariance structure models for gene expression microarray data. *Struct. Equ. Model.* **10** 566–582. MR2011195 https://doi.org/10.1207/S15328007SEM1004_5

Youssef, P. (2013). Estimating the covariance of random matrices. *Electron. J. Probab.* **18** 107. MR3151727 https://doi.org/10.1214/EJP.v18-2579

Zhang, Y. and Schneider, J. (2010). Learning multiple tasks with a sparse matrix-normal penalty. In *Advances in Neural Information Processing Systems* **23**. Curran Associates.

Zhivotovskiy, N. (2024). Dimension-free bounds for sums of independent matrices and simple tensors via the variational principle. *Electron. J. Probab.* **29** 13. MR4693860 https://doi.org/10.1214/23-ejp1021