BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Number Two May 2024 ISSN: 1350-7265

CONTENTS

(continued)

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

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Number Two May 2024 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 30 Number 2 May 2024 Pages 851–1693 ISI/BS

Volume 30 Number 2   May 2024   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 2024 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **30**(2), 2024, 851–877
https://doi.org/10.3150/22-BEJ1529

On Azadkia–Chatterjee's conditional dependence coefficient

HONGJIAN SHI[1,a], MATHIAS DRTON[1,b] and FANG HAN[2,c]

[1]*Department of Mathematics, Technical University of Munich, 85748 Garching bei München, Germany,* [a]*hongjian.shi@tum.de,* [b]*mathias.drton@tum.de*
[2]*Department of Statistics, University of Washington, Seattle, WA 98195, USA,* [c]*fanghan@uw.edu*

In recent work, Azadkia and Chatterjee (*Ann. Statist.* **49** (2021) 3070–3102) laid out an ingenious approach to defining consistent measures of conditional dependence. Their fully nonparametric approach forms statistics based on ranks and nearest neighbor graphs. The appealing nonparametric consistency of the resulting conditional dependence measure and the associated empirical conditional dependence coefficient has quickly prompted follow-up work that seeks to study its statistical efficiency. In this paper, we take up the framework of conditional randomization tests (CRT) for conditional independence and conduct a power analysis that considers two types of local alternatives, namely, parametric quadratic mean differentiable alternatives and nonparametric Hölder smooth alternatives. Our local power analysis shows that conditional independence tests using the Azadkia–Chatterjee coefficient remain inefficient even when aided with the CRT framework, and serves as motivation to develop variants of the approach; cf. Lin and Han (*Biometrika* **110** (2023) 283–299). As a byproduct, we resolve a conjecture of Azadkia and Chatterjee by proving central limit theorems for the considered conditional dependence coefficients, with explicit formulas for the asymptotic variances.

Keywords: Conditional independence; graph-based test; rank-based test; nearest neighbor graphs; local power analysis

References

Auddy, A., Deb, N. and Nandy, S. (2023). Exact detection thresholds for Chatterjee's correlation. *Bernoulli*. To appear.

Azadkia, M. and Chatterjee, S. (2021). A simple measure of conditional dependence. *Ann. Statist.* **49** 3070–3102. MR4352523 https://doi.org/10.1214/21-aos2073

Bergsma, W. (2004). Testing conditional independence for continuous random variables. Eurandom Report No. 2004-048. Available at https://www.eurandom.tue.nl/reports/2004/048-report.pdf.

Bergsma, W. (2011). Nonparametric testing of conditional independence by means of the partial copula. Available at arXiv:1101.4607v1.

Berrett, T.B., Samworth, R.J. and Yuan, M. (2019). Efficient multivariate entropy estimation via k-nearest neighbour distances. *Ann. Statist.* **47** 288–318. MR3909934 https://doi.org/10.1214/18-AOS1688

Berrett, T.B., Wang, Y., Barber, R.F. and Samworth, R.J. (2020). The conditional permutation test for independence while controlling for confounders. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 175–197. MR4060981

Bhattacharya, B.B. (2019). A general asymptotic framework for distribution-free graph-based two-sample tests. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 575–602. MR3961499

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

Bickel, P.J. and Breiman, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. MR0682809

Cai, Z., Li, R. and Zhang, Y. (2022). A distribution free conditional independence test with applications to causal discovery. *J. Mach. Learn. Res.* **23** Paper No. 85. MR4576670

1350-7265

Candès, E., Fan, Y., Janson, L. and Lv, J. (2018). Panning for gold: 'model-X' knockoffs for high dimensional controlled variable selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 551–577. MR3798878 https://doi.org/10.1111/rssb.12265

Canonne, C.L., Diakonikolas, I., Kane, D.M. and Stewart, A. (2018). Testing conditional independence of discrete distributions. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* 735–748. New York: ACM. MR3826290 https://doi.org/10.1145/3188745.3188756

Cao, S. and Bickel, P.J. (2020). Correlations with tailored extremal properties. Available at arXiv:2008.10177v2.

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, L.H.Y. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028. MR2073183 https://doi.org/10.1214/009117904000000450

Dawid, A.P. (1979). Conditional independence in statistical theory. *J. Roy. Statist. Soc. Ser. B* **41** 1–31. MR0535541

Dawid, A.P. (1980). Conditional independence for statistical operations. *Ann. Statist.* **8** 598–617. MR0568723

Deb, N., Ghosal, P. and Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. Available at arXiv:2010.01768v2.

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

Dette, H., Siburg, K.F. and Stoimenov, P.A. (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat.* **40** 21–41. MR3024030 https://doi.org/10.1111/j.1467-9469.2011.00767.x

Devroye, L. (1988). The expected size of some graphs in computational geometry. *Comput. Math. Appl.* **15** 53–64. MR0937563 https://doi.org/10.1016/0898-1221(88)90071-5

Devroye, L., Györfi, L., Lugosi, G. and Walk, H. (2018). A nearest neighbor estimate of the residual variance. *Electron. J. Stat.* **12** 1752–1778. MR3811758 https://doi.org/10.1214/18-EJS1438

Doran, G., Muandet, K., Zhang, K. and Schölkopf, B. (2014). A permutation-based kernel conditional independence test. In *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*. *UAI'14* 132–141. Arlington, Virginia, USA: AUAI Press.

Friedman, J.H. and Rafsky, L.C. (1979). Multivariate generalizations of the Wald-Wolfowitz and Smirnov two-sample tests. *Ann. Statist.* **7** 697–717. MR0532236

Fukumizu, K., Gretton, A., Sun, X. and Schölkopf, B. (2008). Kernel measures of conditional dependence. In *Advances in Neural Information Processing Systems 20* (J.C. Platt, D. Koller, Y. Singer and S.T. Roweis, eds.) 673–680. Curran Associates, Red Hook, NY.

Gamboa, F., Gremaud, P., Klein, T. and Lagnoux, A. (2022). Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics. *Bernoulli* **28** 2345–2374. MR4474546 https://doi.org/10.3150/21-bej1421

Hájek, J., Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests*, 2nd ed. *Probability and Mathematical Statistics*. San Diego, CA: Academic Press. MR1680991

Hallin, M., del Barrio, E., Cuesta-Albertos, J. and Matrán, C. (2021). Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *Ann. Statist.* **49** 1139–1165. MR4255122 https://doi.org/10.1214/20-aos1996

Henze, N. (1987). On the fraction of random points with specified nearest-neighbour interrelations and degree of attraction. *Adv. in Appl. Probab.* **19** 873–895. MR0914597 https://doi.org/10.2307/1427106

Henze, N. (1988). A multivariate two-sample test based on the number of nearest neighbor type coincidences. *Ann. Statist.* **16** 772–783. MR0947577 https://doi.org/10.1214/aos/1176350835

Henze, N. and Penrose, M.D. (1999). On the multivariate runs test. *Ann. Statist.* **27** 290–298. MR1701112 https://doi.org/10.1214/aos/1018031112

Hoeffding, W. (1952). The large-sample power of tests based on permutations of observations. *Ann. Math. Stat.* **23** 169–192. MR0057521 https://doi.org/10.1214/aoms/1177729436

Hoyer, P., Janzing, D., Mooij, J.M., Peters, J. and Schölkopf, B. (2009). Nonlinear causal discovery with additive noise models. In *Advances in Neural Information Processing Systems* (D. Koller, D. Schuurmans, Y. Bengio and L. Bottou, eds.) **21** 692–699. Curran Associates, Inc.

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

Huang, Z., Deb, N. and Sen, B. (2022). Kernel partial correlation coefficient—a measure of conditional dependence. *J. Mach. Learn. Res.* **23** Paper No. 216. MR4577169 https://doi.org/10.1086/287487

Koller, D. and Sahami, M. (1996). Toward Optimal Feature Selection. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*. *ICML'96* 284–292. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Kössler, W. and Rödel, E. (2007). The asymptotic efficacies and relative efficiencies of various linear rank tests for independence. *Metrika* **65** 3–28. MR2288045 https://doi.org/10.1007/s00184-006-0055-x

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **23**. Berlin: Springer. MR1102015 https://doi.org/10.1007/978-3-642-20212-4

Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR2135927

Li, S. (2011). Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Stat.* **4** 66–70. MR2813331 https://doi.org/10.3923/ajms.2011.66.70

Lin, Z. and Han, F. (2022). Limit theorems of Chatterjee's rank correlation. Available at arXiv:2204.08031v2.

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

Linton, O. and Gozalo, P. (1996). Conditional independence restrictions: testing and estimation. Cowles Foundation Discussion Paper No. 1140. Available at https://cowles.yale.edu/publications/cfdp/cfdp-1140.

Liu, R.Y. and Singh, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* **88** 252–260. MR1212489

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

Maathuis, M., Drton, M., Lauritzen, S. and Wainwright, M., eds. (2019). *Handbook of Graphical Models*. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. Boca Raton, FL: CRC Press.

Neykov, M., Balakrishnan, S. and Wasserman, L. (2021). Minimax optimal conditional independence testing. *Ann. Statist.* **49** 2151–2177. MR4319245 https://doi.org/10.1214/20-aos2030

Peters, J., Janzing, D. and Schölkopf, B. (2011). Causal inference on discrete data using additive noise models. *IEEE Trans. Pattern Anal. Mach. Intell.* **33** 2436–2450. https://doi.org/10.1109/TPAMI.2011.71

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

Petersen, L. and Hansen, N.R. (2021). Testing conditional independence via quantile regression based partial copulas. *J. Mach. Learn. Res.* **22** Paper No. 70. MR4253763

Póczos, B. and Schneider, J. (2012). Nonparametric estimation of conditional information and divergences. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics* (N.D. Lawrence and M. Girolami, eds.). *Proceedings of Machine Learning Research* **22** 914–923. La Palma, Canary Islands: PMLR.

Runge, J. (2018). Conditional independence testing based on a nearest-neighbor estimator of conditional mutual information. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics* (A. Storkey and F. Perez-Cruz, eds.). *Proceedings of Machine Learning Research* **84** 938–947. PMLR.

Shah, R.D. and Peters, J. (2020). The hardness of conditional independence testing and the generalised covariance measure. *Ann. Statist.* **48** 1514–1538. MR4124333 https://doi.org/10.1214/19-AOS1857

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

Shi, H., Drton, M. and Han, F. (2022b). Supplement to "On the power of Chatterjee's rank correlation". *Biometrika* **109**. https://doi.org/10.1093/biomet/asab028#supplementary-data

Shi, H., Drton, M. and Han, F. (2022c). 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. (2024). Supplement to "On Azadkia–Chatterjee's conditional dependence coefficient." https://doi.org/10.3150/22-BEJ1529SUPP

Shi, H., Hallin, M., Drton, M. and Han, F. (2022). On universally consistent and fully distribution-free rank tests of vector independence. *Ann. Statist.* **50** 1933–1959. MR4474478 https://doi.org/10.1214/21-aos2151

Song, K. (2009). Testing conditional independence via Rosenblatt transforms. *Ann. Statist.* **37** 4011–4045. MR2572451 https://doi.org/10.1214/09-AOS704

Stone, C.J. (1977). Consistent nonparametric regression. *Ann. Statist.* **5** 595–645. MR0443204

Strobl, E.V., Zhang, K. and Visweswaran, S. (2019). Approximate kernel-based conditional independence tests for fast non-parametric causal discovery. *J. Causal Inference* **7** Art. No. 20180017. MR4350065 https://doi.org/10.1515/jci-2018-0017

Su, L. and White, H. (2007). A consistent characteristic function-based test for conditional independence. *J. Econometrics* **141** 807–834. MR2413488 https://doi.org/10.1016/j.jeconom.2006.11.006

Su, L. and White, H. (2008). A nonparametric Hellinger metric test for conditional independence. *Econometric Theory* **24** 829–864. MR2428851 https://doi.org/10.1017/S0266466608080341

Su, L. and White, H. (2014). Testing conditional independence via empirical likelihood. *J. Econometrics* **182** 27–44. MR3212759 https://doi.org/10.1016/j.jeconom.2014.04.006

Székely, G.J. and Rizzo, M.L. (2013). Energy statistics: A class of statistics based on distances. *J. Statist. Plann. Inference* **143** 1249–1272. MR3055745 https://doi.org/10.1016/j.jspi.2013.03.018

Székely, G.J. and Rizzo, M.L. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.* **42** 2382–2412. MR3269983 https://doi.org/10.1214/14-AOS1255

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794

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

Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.* **38** 766–780. MR2859749 https://doi.org/10.1111/j.1467-9469.2011.00744.x

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. *Ann. Math. Stat.* **11** 147–162. MR0002083 https://doi.org/10.1214/aoms/1177731909

Wang, X., Pan, W., Hu, W., Tian, Y. and Zhang, H. (2015). Conditional distance correlation. *J. Amer. Statist. Assoc.* **110** 1726–1734. MR3449068 https://doi.org/10.1080/01621459.2014.993081

Zhang, K., Peters, J., Janzing, D. and Schölkopf, B. (2011). Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*. *UAI'11* 804–813. Arlington, Virginia, USA: AUAI Press.

Bernoulli **30**(2), 2024, 878–910
https://doi.org/10.3150/22-BEJ1563

Laplace priors and spatial inhomogeneity in Bayesian inverse problems

SERGIOS AGAPIOU[1,a] and SVEN WANG[2,b]

[1]*Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus,* [a]*agapiou.sergios@ucy.ac.cy*
[2]*Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, USA,*
[b]*svenwang@mit.edu*

Spatially inhomogeneous functions, which may be smooth in some regions and rough in other regions, are modelled naturally in a Bayesian manner using so-called *Besov priors* which are given by random wavelet expansions with Laplace-distributed coefficients. This paper studies theoretical guarantees for such prior measures – specifically, we examine their frequentist posterior contraction rates in the setting of non-linear inverse problems with Gaussian white noise. Our results are first derived under a general local Lipschitz assumption on the forward map. We then verify the assumption for two non-linear inverse problems arising from elliptic partial differential equations, the *Darcy flow* model from geophysics as well as a model for the *Schrödinger equation* appearing in tomography. In the course of the proofs, we also obtain novel concentration inequalities for penalized least squares estimators with ℓ^1 wavelet penalty, which have a natural interpretation as maximum a posteriori (MAP) estimators. The true parameter is assumed to belong to some spatially inhomogeneous Besov class B^{α}_{11}, with $\alpha > 0$ sufficiently large. In a setting with direct observations, we complement these upper bounds with a lower bound on the rate of contraction for *arbitrary* Gaussian priors. An immediate consequence of our results is that while Laplace priors can achieve minimax-optimal rates over B^{α}_{11}-classes, Gaussian priors are limited to a (by a polynomial factor) slower contraction rate. This gives information-theoretical justification for the intuition that Laplace priors are more compatible with ℓ^1 regularity structure in the underlying parameter.

Keywords: Bayesian nonparametric inference; frequentist consistency; inverse problems; Laplace prior; spatially inhomogeneous functions

References

[1] Abraham, K. and Nickl, R. (2019). On statistical Calderón problems. *Math. Stat. Learn.* **2** 165–216. MR4130599
[2] Agapiou, S., Burger, M., Dashti, M. and Helin, T. (2018). Sparsity-promoting and edge-preserving maximum *a posteriori* estimators in non-parametric Bayesian inverse problems. *Inverse Probl.* **34** 045002. MR3774703 https://doi.org/10.1088/1361-6420/aaacac
[3] Agapiou, S., Dashti, M. and Helin, T. (2021). Rates of contraction of posterior distributions based on p-exponential priors. *Bernoulli* **27** 1616–1642. MR4278794 https://doi.org/10.3150/20-bej1285
[4] Agapiou, S., Larsson, S. and Stuart, A.M. (2013). Posterior contraction rates for the Bayesian approach to linear ill-posed inverse problems. *Stochastic Process. Appl.* **123** 3828–3860. MR3084161 https://doi.org/10.1016/j.spa.2013.05.001
[5] Agapiou, S. and Wang, S. (2024). Supplement to "Laplace priors and spatial inhomogeneity in Bayesian inverse problems." https://doi.org/10.3150/22-BEJ1563SUPP
[6] Alt, H.W. (2016). *Linear Functional Analysis: An Application-Oriented Introduction*. *Universitext*. London: Springer London, Ltd. MR3497775 https://doi.org/10.1007/978-1-4471-7280-2
[7] Aurzada, F. (2007). On the lower tail probabilities of some random sequences in l_p. *J. Theoret. Probab.* **20** 843–858. MR2359058 https://doi.org/10.1007/s10959-007-0095-9

1350-7265

[8] 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
[9] Bonito, A., Cohen, A., DeVore, R., Petrova, G. and Welper, G. (2017). Diffusion coefficients estimation for elliptic partial differential equations. *SIAM J. Math. Anal.* **49** 1570–1592. MR3639575 https://doi.org/10.1137/16M1094476
[10] Borell, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239–252. MR0388475 https://doi.org/10.1007/BF02384761
[11] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958 https://doi.org/10.1214/aos/1032181159
[12] Castillo, I. (2008). Lower bounds for posterior rates with Gaussian process priors. *Electron. J. Stat.* **2** 1281–1299. MR2471287 https://doi.org/10.1214/08-EJS273
[13] 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
[14] Chen, V., Dunlop, M.M., Papaspiliopoulos, O. and Stuart, A.M. (2018). Dimension-robust MCMC in Bayesian inverse problems. Preprint. Available at arXiv:1803.03344.
[15] 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
[16] 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
[17] 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
[18] Dalalyan, A.S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183
[19] Dashti, M., Harris, S. and Stuart, A. (2012). Besov priors for Bayesian inverse problems. *Inverse Probl. Imaging* **6** 183–200. MR2942737 https://doi.org/10.3934/ipi.2012.6.183
[20] Dashti, M. and Stuart, A.M. (2017). The Bayesian approach to inverse problems. In *Handbook of Uncertainty Quantification. Vol. 1, 2, 3* 311–428. Cham: Springer. MR3839555
[21] Daubechies, I. (1992). *Ten Lectures on Wavelets*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **61**. Philadelphia, PA: SIAM. MR1162107 https://doi.org/10.1137/1.9781611970104
[22] Donoho, D.L. (1995). Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.* **2** 101–126. MR1325535 https://doi.org/10.1006/acha.1995.1008
[23] Donoho, D.L. and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. MR1635414 https://doi.org/10.1214/aos/1024691081
[24] Durmus, A. and Moulines, É. (2019). High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli* **25** 2854–2882. MR4003567 https://doi.org/10.3150/18-BEJ1073
[25] Edmunds, D.E. and Triebel, H. (1996). *Function Spaces, Entropy Numbers, Differential Operators*. *Cambridge Tracts in Mathematics* **120**. Cambridge: Cambridge Univ. Press. MR1410258 https://doi.org/10.1017/CBO9780511662201
[26] Engl, H.W., Hanke, M. and Neubauer, A. (1996). *Regularization of Inverse Problems*. *Mathematics and Its Applications* **375**. Dordrecht: Kluwer Academic. MR1408680
[27] 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
[28] Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist.* **35** 192–223. MR2332274 https://doi.org/10.1214/009053606000001172
[29] 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
[30] Gilbarg, D. and Trudinger, N.S. (2001). *Elliptic Partial Differential Equations of Second Order*. *Classics in Mathematics*. Berlin: Springer. MR1814364

[31] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics, [40]*. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862
[32] Giordano, M. and Nickl, R. (2020). Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. *Inverse Probl.* **36** 085001. MR4151406 https://doi.org/10.1088/1361-6420/ab7d2a
[33] Giordano, M. and Ray, K. (2022). Nonparametric Bayesian inference for reversible multidimensional diffusions. *Ann. Statist.* **50** 2872–2898. MR4500628 https://doi.org/10.1214/22-aos2213
[34] Hairer, M., Stuart, A.M. and Vollmer, S.J. (2014). Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions. *Ann. Appl. Probab.* **24** 2455–2490. MR3262508 https://doi.org/10.1214/13-AAP982
[35] Iglesias, M.A., Lin, K. and Stuart, A.M. (2014). Well-posed Bayesian geometric inverse problems arising in subsurface flow. *Inverse Probl.* **30** 114001. MR3274585 https://doi.org/10.1088/0266-5611/30/11/114001
[36] Isakov, V. (2006). *Inverse Problems for Partial Differential Equations*, 2nd ed. *Applied Mathematical Sciences* **127**. New York: Springer. MR2193218
[37] Ito, K. and Kunisch, K. (1994). On the injectivity and linearization of the coefficient-to-solution mapping for elliptic boundary value problems. *J. Math. Anal. Appl.* **188** 1040–1066. MR1305502 https://doi.org/10.1006/jmaa.1994.1479
[38] Kaipio, J. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems*. *Applied Mathematical Sciences* **160**. New York: Springer. MR2102218
[39] Kekkonen, H. (2022). Consistency of Bayesian inference with Gaussian process priors for a parabolic inverse problem. *Inverse Probl.* **38** Paper No. 035002. MR4385425 https://doi.org/10.1088/1361-6420/ac4839
[40] Kekkonen, H., Lassas, M., Saksman, E. and Siltanen, S. (2023). Random tree Besov priors—towards fractal imaging. *Inverse Probl. Imaging* **17** 507–531. MR4546960
[41] Knapik, B.T., van der Vaart, A.W. and van Zanten, J.H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.* **39** 2626–2657. MR2906881 https://doi.org/10.1214/11-AOS920
[42] Lassas, M., Saksman, E. and Siltanen, S. (2009). Discretization-invariant Bayesian inversion and Besov space priors. *Inverse Probl. Imaging* **3** 87–122. MR2558305 https://doi.org/10.3934/ipi.2009.3.87
[43] Lassas, M. and Siltanen, S. (2004). Can one use total variation prior for edge-preserving Bayesian inversion? *Inverse Probl.* **20** 1537–1563. MR2109134 https://doi.org/10.1088/0266-5611/20/5/013
[44] Lions, J.-L. and Magenes, E. (1972). *Non-homogeneous Boundary Value Problems and Applications. Vol. I*. *Die Grundlehren der Mathematischen Wissenschaften, Band 181*. New York: Springer. MR0350177
[45] Martin, J., Wilcox, L.C., Burstedde, C. and Ghattas, O. (2012). A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion. *SIAM J. Sci. Comput.* **34** A1460–A1487. MR2970260 https://doi.org/10.1137/110845598
[46] Meyer, Y. (1992). *Wavelets and Operators*. *Cambridge Studies in Advanced Mathematics* **37**. Cambridge: Cambridge Univ. Press. MR1228209
[47] Monard, F., Nickl, R. and Paternain, G.P. (2021). Consistent inversion of noisy non-Abelian X-ray transforms. *Comm. Pure Appl. Math.* **74** 1045–1099. MR4230066 https://doi.org/10.1002/cpa.21942
[48] Nickl, R. (2020). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc. (JEMS)* **22** 2697–2750. MR4118619 https://doi.org/10.4171/JEMS/975
[49] Nickl, R., van de Geer, S. and Wang, S. (2020). Convergence rates for penalized least squares estimators in PDE constrained regression problems. *SIAM/ASA J. Uncertain. Quantificat.* **8** 374–413. MR4074017 https://doi.org/10.1137/18M1236137
[50] Nickl, R. and Wang, S. (2023). On polynomial-time computation of high-dimensional posterior measures by Langevin-type algorithms. *J. Eur. Math. Soc.* To appear (available online). https://doi.org/10.4171/JEMS/1304
[51] Ray, K. (2013). Bayesian inverse problems with non-conjugate priors. *Electron. J. Stat.* **7** 2516–2549. MR3117105 https://doi.org/10.1214/13-EJS851
[52] Reed, M. and Simon, B. (1980). *Methods of Modern Mathematical Physics. I: Functional Analysis*, 2nd ed. New York: Academic Press. MR0751959
[53] Reiß, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.* **36** 1957–1982. MR2435461 https://doi.org/10.1214/07-AOS525
[54] Richter, G.R. (1981). An inverse problem for the steady state diffusion equation. *SIAM J. Appl. Math.* **41** 210–221. MR0628945 https://doi.org/10.1137/0141016

[55] Ročková, V. and Rousseau, J. (2021). Ideal Bayesian Spatial Adaptation. Preprint. Available at arXiv:2105.12793.

[56] 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

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

[58] Talagrand, M. (1994). The supremum of some canonical processes. *Amer. J. Math.* **116** 283–325. MR1269606 https://doi.org/10.2307/2374931

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

[60] Triebel, H. (2010). *Theory of Function Spaces*. *Modern Birkhäuser Classics*. Basel: Birkhäuser/Springer. MR3024598

[61] van de Geer, S. (2001). Least squares estimation with complexity penalties. *Math. Methods Statist.* **10** 355–374.

[62] van de Geer, S.A. (2000). *Applications of Empirical Process Theory*. *Cambridge Series in Statistical and Probabilistic Mathematics* **6**. Cambridge: Cambridge Univ. Press. MR1739079

[63] 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. MR2418663 https://doi.org/10.1214/009053607000000613

[64] Yeh, W.W.-G. (1986). Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.* **22** 95–108.

[65] Zhao, L.H. (2000). Bayesian aspects of some nonparametric problems. *Ann. Statist.* **28** 532–552. MR1790008 https://doi.org/10.1214/aos/1016218229

Bernoulli **30**(2), 2024, 911–935
https://doi.org/10.3150/23-BEJ1586

Variance estimation for sequential Monte Carlo algorithms: A backward sampling approach

YAZID JANATI EL IDRISSI[1,a], SYLVAIN LE CORFF[2,c] and YOHAN PETETIN[1,b]

[1]*SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France,*
[a]*yazid.janati_elidrissi@telecom-sudparis.eu,* [b]*yohan.petetin@telecom-sudparis.eu*
[2]*LPSM, Sorbonne Université, UMR CNRS 8001, 4 Place Jussieu, 75005 Paris, France,*
[c]*sylvain.le_corff@sorbonne-universite.fr*

In this paper, we consider the problem of online asymptotic variance estimation for particle filtering and smoothing. Current solutions for the particle filter rely on the particle genealogy and are either unstable or hard to tune in practice. We propose to mitigate these limitations by introducing a new estimator of the asymptotic variance based on the so called backward weights. The resulting estimator is weakly consistent and trades computational cost for more stability and reduced variance. We also propose a more computationally efficient estimator inspired by the *PaRIS* algorithm of (*Bernoulli* **23** (2017) 1951–1996). As an application, particle smoothing is considered and an estimator of the asymptotic variance of the Forward Filtering Backward Smoothing estimator applied to additive functionals is provided.

Keywords: Asymptotic variance; central limit theorem; particle filtering; particle smoothing; sequential Monte Carlo methods

References

Andrieu, C., Doucet, A. and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 269–342. MR2758115 https://doi.org/10.1111/j.1467-9868.2009.00736.x

Andrieu, C., Lee, A. and Vihola, M. (2018). Uniform ergodicity of the iterated conditional SMC and geometric ergodicity of particle Gibbs samplers. *Bernoulli* **24** 842–872. MR3706778 https://doi.org/10.3150/15-BEJ785

Cappé, O., Moulines, E. and Rydén, T. (2005). *Inference in Hidden Markov Models*. *Springer Series in Statistics*. New York: Springer. MR2159833

Cérou, F., Del Moral, P. and Guyader, A. (2011). A nonasymptotic theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 629–649. MR2841068 https://doi.org/10.1214/10-AIHP358

Chan, H.P. and Lai, T.L. (2013). A general theory of particle filters in hidden Markov models and some applications. *Ann. Statist.* **41** 2877–2904. MR3161451 https://doi.org/10.1214/13-AOS1172

Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.* **32** 2385–2411. MR2153989 https://doi.org/10.1214/009053604000000698

Chopin, N. and Papaspiliopoulos, O. (2020). *An Introduction to Sequential Monte Carlo*. *Springer Series in Statistics*. Cham: Springer. MR4215639 https://doi.org/10.1007/978-3-030-47845-2

Chopin, N. and Singh, S.S. (2015). On particle Gibbs sampling. *Bernoulli* **21** 1855–1883. MR3352064 https://doi.org/10.3150/14-BEJ629

Del Moral, P. (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. *Probability and Its Applications (New York)*. New York: Springer. MR2044973 https://doi.org/10.1007/978-1-4684-9393-1

Del Moral, P., Doucet, A. and Singh, S.S. (2010). A backward particle interpretation of Feynman-Kac formulae. *ESAIM Math. Model. Numer. Anal.* **44** 947–975. MR2731399 https://doi.org/10.1051/m2an/2010048

Del Moral, P. and Guionnet, A. (1999). Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9** 275–297. MR1687359 https://doi.org/10.1214/aoap/1029962742

Del Moral, P., Doucet, A. and Sumeetpal, S. (2010). Forward Smoothing using Sequential Monte Carlo. Preprint. Available at arXiv:1012.5390.

Douc, R. and Cappé, O. (2005). Comparison of resampling schemes for particle filtering. In *ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005* 64–69. IEEE.

Douc, R., Moulines, E. and Stoffer, D.S. (2014). *Nonlinear Time Series: Theory, Methods, and Applications with R Examples*. *Chapman & Hall/CRC Texts in Statistical Science Series*. Boca Raton, FL: CRC Press/CRC. MR3289095

Douc, R., Garivier, A., Moulines, E. and Olsson, J. (2011). Sequential Monte Carlo smoothing for general state space hidden Markov models. *Ann. Appl. Probab.* **21** 2109–2145. MR2895411 https://doi.org/10.1214/10-AAP735

Doucet, A., Godsill, S. and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **10** 197–208.

Du, Q. and Guyader, A. (2021). Variance estimation in adaptive sequential Monte Carlo. *Ann. Appl. Probab.* **31** 1021–1060. MR4278778 https://doi.org/10.1214/20-aap1611

Dubarry, C. and Le Corff, S. (2013). Non-asymptotic deviation inequalities for smoothed additive functionals in nonlinear state-space models. *Bernoulli* **19** 2222–2249. MR3160552 https://doi.org/10.3150/12-BEJ450

Fearnhead, P., Wyncoll, D. and Tawn, J. (2010). A sequential smoothing algorithm with linear computational cost. *Biometrika* **97** 447–464. MR2650750 https://doi.org/10.1093/biomet/asq013

Gloaguen, P., Le Corff, S. and Olsson, J. (2022). A pseudo-marginal sequential Monte Carlo online smoothing algorithm. *Bernoulli* **28** 2606–2633. MR4474556 https://doi.org/10.3150/21-bej1431

Godsill, S.J., Doucet, A. and West, M. (2004). Monte Carlo smoothing for nonlinear times series. *J. Amer. Statist. Assoc.* **99** 156–168. MR2054295 https://doi.org/10.1198/016214504000000151

Gordon, N.J., Salmond, D.J. and Smith, A.F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F* **140** 107–113. IET.

Janati El Idrissi, Y., Le Corff, S. and Petetin, Y. (2024). Supplement to "Variance estimation for sequential Monte Carlo algorithms: A backward sampling approach." https://doi.org/10.3150/23-BEJ1586SUPP

Kantas, N., Doucet, A., Singh, S.S., Maciejowski, J. and Chopin, N. (2015). On particle methods for parameter estimation in state-space models. *Statist. Sci.* **30** 328–351. MR3383884 https://doi.org/10.1214/14-STS511

Koskela, J., Jenkins, P.A., Johansen, A.M. and Spanò, D. (2020). Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo. *Ann. Statist.* **48** 560–583. MR4065174 https://doi.org/10.1214/19-AOS1823

Künsch, H.R. (2005). Recursive Monte Carlo filters: Algorithms and theoretical analysis. *Ann. Statist.* **33** 1983–2021. MR2211077 https://doi.org/10.1214/009053605000000426

Lee, A., Singh, S.S. and Vihola, M. (2020). Coupled conditional backward sampling particle filter. *Ann. Statist.* **48** 3066–3089. MR4152635 https://doi.org/10.1214/19-AOS1922

Lee, A. and Whiteley, N. (2018). Variance estimation in the particle filter. *Biometrika* **105** 609–625. MR3842888 https://doi.org/10.1093/biomet/asy028

Lindsten, F. and Schön, T.B. (2012). On the use of backward simulation in the particle Gibbs sampler. In *2012 IEEE ICASSP* 3845–3848. https://doi.org/10.1109/ICASSP.2012.6288756

Liu, J. and West, M. (2001). Sequential Monte Carlo methods in practice. In *Statistics for Engineering and Information Science* 225–246. New York: Springer.

Olsson, J. and Douc, R. (2019). Numerically stable online estimation of variance in particle filters. *Bernoulli* **25** 1504–1535. MR3920380 https://doi.org/10.3150/18-bej1028

Olsson, J. and Westerborn, J. (2017). Efficient particle-based online smoothing in general hidden Markov models: The PaRIS algorithm. *Bernoulli* **23** 1951–1996. MR3624883 https://doi.org/10.3150/16-BEJ801

Olsson, J., Cappé, O., Douc, R. and Moulines, E. (2008). Sequential Monte Carlo smoothing with application to parameter estimation in nonlinear state space models. *Bernoulli* **14** 155–179. MR2401658 https://doi.org/10.3150/07-BEJ6150

Pitt, M.K. and Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters. *J. Amer. Statist. Assoc.* **94** 590–599. MR1702328 https://doi.org/10.2307/2670179

Poyiadjis, G., Doucet, A. and Singh, S.S. (2011). Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. *Biometrika* **98** 65–80. MR2804210 https://doi.org/10.1093/biomet/asq062

Tanizaki, H. and Mariano, R. (1994). Prediction, filtering and smoothing in non-linear and non-normal cases using Monte Carlo integration. *J. Appl. Econometrics* **9** 163–79.

Bernoulli **30**(2), 2024, 936–960
https://doi.org/10.3150/23-BEJ1619

Reproduction of initial distributions from the first hitting time distribution for birth-and-death processes

KOSUKE YAMATO[1,a] and KOUJI YANO[2,b]

[1]*Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan,* [a]*kyamato.math@gmail.com*
[2]*Graduate School of Science, Osaka University, Osaka, Japan,* [b]*kyanomath@gmail.com*

For birth-and-death processes, we show that every initial distribution is reproduced from the first hitting time distribution. The reproduction is done by applying to the distribution function a differential operator defined through the eigenfunction of the generator. Using the spectral theory for generalized second-order differential operators, we study asymmetric random walks and binary branching processes.

Keywords: Birth-and-death processes; first hitting time

References

Anderson, W.J. (1991). *Continuous-Time Markov Chains*. *Springer Series in Statistics: Probability and Its Applications*. New York: Springer. An applications-oriented approach. MR1118840 https://doi.org/10.1007/978-1-4612-3038-0

Coddington, E.A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations*. New York-Toronto-London: McGraw-Hill, Inc. MR0069338

Diaconis, P. and Miclo, L. (2009). On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.* **22** 558–586. MR2530103 https://doi.org/10.1007/s10959-009-0234-6

Itô, K. (2006). *Essentials of Stochastic Processes*. *Translations of Mathematical Monographs* **231**. Providence, RI: Amer. Math. Soc. Translated from the 1957 Japanese original by Yuji Ito. MR2239081 https://doi.org/10.1090/mmono/231

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

Karlin, S. and McGregor, J.L. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.* **85** 489–546. MR0091566 https://doi.org/10.2307/1992942

Kotani, S. and Watanabe, S. (1982). Kreĭn's spectral theory of strings and generalized diffusion processes. In *Functional Analysis in Markov Processes (Katata/Kyoto, 1981)*. *Lecture Notes in Math.* **923** 235–259. Berlin: Springer. MR0661628

Magnus, W., Oberhettinger, F. and Soni, R.P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics*, enlarged ed. *Die Grundlehren der Mathematischen Wissenschaften, Band 52*. New York: Springer New York, Inc. MR0232968

McKean, H.P. Jr. (1956). Elementary solutions for certain parabolic partial differential equations. *Trans. Amer. Math. Soc.* **82** 519–548. MR0087012 https://doi.org/10.2307/1993060

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

Rogers, L.C.G. (1984). A diffusion first passage problem. In *Seminar on Stochastic Processes, 1983 (Gainesville, Fla., 1983)*. *Progr. Probab. Statist.* **7** 151–160. Boston, MA: Birkhäuser. MR0902417 https://doi.org/10.1007/978-1-4684-9169-2_10

1350-7265

Schilling, R.L., Song, R. and Vondraček, Z. (2012). *Bernstein Functions*, 2nd ed. *De Gruyter Studies in Mathematics* **37**. Berlin: de Gruyter. Theory and applications. MR2978140 https://doi.org/10.1515/9783110269338

Takemura, T. and Tomisaki, M. (2012). *h* transform of one-dimensional generalized diffusion operators. *Kyushu J. Math.* **66** 171–191. MR2962397 https://doi.org/10.2206/kyushujm.66.171

Yamato, K. (2021). Existence of Laplace transforms of the spectral measures for one-dimensional diffusions with an exit boundary. *Infinitely Divisible Processes and Related Topics (25), The Institute of Statistical Mathematics Cooperative Research Report* **446** 55–59.

Yamato, K. (2022). A unifying approach to non-minimal quasi-stationary distributions for one-dimensional diffusions. *J. Appl. Probab.* **59** 1106–1128. MR4507684 https://doi.org/10.1017/jpr.2022.2

Yamato, K. and Yano, K. (2024). Supplement to "Reproduction of initial distributions from the first hitting time distribution for birth-and-death processes." https://doi.org/10.3150/23-BEJ1619SUPP

Yano, K. (2006). Excursion measure away from an exit boundary of one-dimensional diffusion processes. *Publ. Res. Inst. Math. Sci.* **42** 837–878. MR2266999

Bernoulli **30**(2), 2024, 961–982
https://doi.org/10.3150/23-BEJ1620

Maximal displacement of spectrally negative branching Lévy processes

CHRISTOPHE PROFETA[a]

Université Paris-Saclay, CNRS, Univ Evry, Laboratoire de Mathématiques et Modélisation d'Evry, 91037, Evry-Courcouronnes, France, [a]*christophe.profeta@univ-evry.fr*

We consider a branching Markov process in continuous time in which the particles evolve independently as spectrally negative Lévy processes. When the branching mechanism is critical or subcritical, the process will eventually die and we may define its overall maximum, i.e. the maximum location ever reached by a particle. The purpose of this paper is to give asymptotic estimates for the survival function of this maximum. In particular, we show that in the critical case the asymptotics is polynomial when the underlying Lévy process oscillates or drifts towards $+\infty$, and is exponential when it drifts towards $-\infty$.

Keywords: Branching process; extreme values; spectrally negative Lévy process

References

Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. *Die Grundlehren der Mathematischen Wissenschaften, Band 196*. New York: Springer. MR0373040

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

Bertoin, J. and Mallein, B. (2018). Biggins' martingale convergence for branching Lévy processes. *Electron. Commun. Probab.* **23** 83. MR3873790 https://doi.org/10.1214/18-ECP185

Bramson, M.D. (1978). Maximal displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31** 531–581. MR0494541 https://doi.org/10.1002/cpa.3160310502

Hubalek, F. and Kyprianou, E. (2011). Old and new examples of scale functions for spectrally negative Lévy processes. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*. *Progress in Probability* **63** 119–145. Basel: Birkhäuser/Springer Basel AG. MR2857022 https://doi.org/10.1007/978-3-0348-0021-1_8

Kuznetsov, A., Kyprianou, A.E. and Rivero, V. (2011). The theory of scale functions for spectrally negative Lévy processes. In *Lévy Matters II*. *Lecture Notes in Math.* **2061** 97–186.

Kyprianou, A.E. (1999). A note on branching Lévy processes. *Stochastic Process. Appl.* **82** 1–14. MR1695066 https://doi.org/10.1016/S0304-4149(99)00010-1

Kyprianou, A.E. (2014). *Fluctuations of Lévy Processes with Applications. Introductory Lectures*, 2nd ed. *Universitext*. Heidelberg: Springer. MR3155252 https://doi.org/10.1007/978-3-642-37632-0

Lalley, S.P. and Shao, Y. (2015). On the maximal displacement of critical branching random walk. *Probab. Theory Related Fields* **162** 71–96. MR3350041 https://doi.org/10.1007/s00440-014-0566-8

Lalley, S.P. and Shao, Y. (2016). Maximal displacement of critical branching symmetric stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1161–1177. MR3531704 https://doi.org/10.1214/15-AIHP677

Mallein, B. and Shi, Q. (2023). A necessary and sufficient condition for the convergence of the derivative martingale in a branching Lévy process. *Bernoulli* **29** 597–624. MR4497260 https://doi.org/10.3150/22-bej1470

Profeta, C. (2022). Extreme values of critical and subcritical branching stable processes with positive jumps. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 1421–1433. MR4517728

Sawyer, S. and Fleischman, J. (1979). Maximum geographic range of a mutant allele considered as a subtype of a Brownian branching random field. *Proc Natl Acad Sci U S A* **76** 872–875. https://doi.org/10.1073/pnas.76.2.872

Bernoulli **30**(2), 2024, 983–1006
https://doi.org/10.3150/23-BEJ1621

Malliavin calculus techniques for local asymptotic mixed normality and their application to hypoelliptic diffusions

MASAAKI FUKASAWA[1,a] and TEPPEI OGIHARA[2,3,b]

[1]*Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560–8531, Japan,* [a]*fukasawa@sigmath.es.osaka-u.ac.jp*
[2]*Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113–8656, Japan,* [b]*ogihara@mist.i.u-tokyo.ac.jp*
[3]*Risk Analysis Research Center, The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan*

We study sufficient conditions for a local asymptotic mixed normality property of statistical models. We accommodate the framework of Jeganathan [*Sankhyā Ser. A* **44** (1982) 173–212] to a triangular array of variable dimension to, in particular, treat high-frequency observations of stochastic processes. When observations are smooth in the Malliavin sense, with the aid of Malliavin calculus techniques by Gobet [*Bernoulli* **7** (2001) 899–912], we further give tractable sufficient conditions which do not require Aronson-type estimates of the transition density function. The transition density function is even allowed to have zeros. For an application, we prove the local asymptotic mixed normality property of hypoelliptic diffusion models under high-frequency observations, in both complete and partial observation frameworks. The former and the latter extend previous results for elliptic diffusions and for integrated diffusions, respectively.

Keywords: Hypoelliptic diffusion processes; integrated diffusion processes; local asymptotic mixed normality; L^2 regularity condition; Malliavin calculus; partial observations

References

[1] Fukasawa, M. and Ogihara, T. (2024). Supplement to "Malliavin calculus techniques for local asymptotic mixed normality and their application to hypoelliptic diffusions." https://doi.org/10.3150/23-BEJ1621SUPP
[2] Genon-Catalot, V. and Jacod, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **29** 119–151. MR1204521
[3] Gloter, A. and Gobet, E. (2008). LAMN property for hidden processes: The case of integrated diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 104–128. MR2451573 https://doi.org/10.1214/07-AIHP111
[4] Gloter, A. and Jacod, J. (2001). Diffusions with measurement errors. I. Local asymptotic normality. *ESAIM Probab. Stat.* **5** 225–242. MR1875672 https://doi.org/10.1051/ps:2001110
[5] Gloter, A. and Yoshida, N. (2021). Adaptive estimation for degenerate diffusion processes. *Electron. J. Stat.* **15** 1424–1472. MR4255288 https://doi.org/10.1214/20-ejs1777
[6] Gobet, E. (2001). Local asymptotic mixed normality property for elliptic diffusion: A Malliavin calculus approach. *Bernoulli* **7** 899–912. MR1873834 https://doi.org/10.2307/3318625
[7] Gobet, E. (2002). LAN property for ergodic diffusions with discrete observations. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 711–737. MR1931584 https://doi.org/10.1016/S0246-0203(02)01107-X
[8] Hájek, J. (1969/70). A characterization of limiting distributions of regular estimates. *Z. Wahrsch. Verw. Gebiete* **14** 323–330. MR0283911 https://doi.org/10.1007/BF00533669
[9] Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of Statistics* 175–194. Berkeley, CA: Univ. California Press. MR0400513

[10] Hall, P. (1977). Martingale invariance principles. *Ann. Probab.* **5** 875–887. MR0517471 https://doi.org/10.1214/aop/1176995657

[11] Jeganathan, P. (1982). On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44** 173–212. MR0688800

[12] Jeganathan, P. (1983). Some asymptotic properties of risk functions when the limit of the experiment is mixed normal. *Sankhyā Ser. A* **45** 66–87. MR0749355

[13] Kohatsu-Higa, A., Nualart, E. and Tran, N.K. (2017). LAN property for an ergodic diffusion with jumps. *Statistics* **51** 419–454. MR3609328 https://doi.org/10.1080/02331888.2016.1239727

[14] Melnykova, A. (2020). Parametric inference for hypoelliptic ergodic diffusions with full observations. *Stat. Inference Stoch. Process.* **23** 595–635. MR4136706 https://doi.org/10.1007/s11203-020-09222-4

[15] 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

[16] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Berlin: Springer. MR2200233

[17] Ogihara, T. (2015). Local asymptotic mixed normality property for nonsynchronously observed diffusion processes. *Bernoulli* **21** 2024–2072. MR3378458 https://doi.org/10.3150/14-BEJ634

[18] Ogihara, T. (2018). Parametric inference for nonsynchronously observed diffusion processes in the presence of market microstructure noise. *Bernoulli* **24** 3318–3383. MR3788175 https://doi.org/10.3150/17-BEJ962

[19] Sweeting, T.J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8** 1375–1381. MR0594652

[20] Tongcang, L., Simon, K., David, M. and Mark, G.R. (2010). Measurement of the instantaneous velocity of a Brownian particle. *Science* **328** 1673–1675.

Bernoulli **30**(2), 2024, 1007–1028
https://doi.org/10.3150/23-BEJ1622

On the separation cut-off phenomenon for Brownian motions on high dimensional spheres

MARC ARNAUDON[1,a], KOLÉHÈ COULIBALY-PASQUIER[2,b] and LAURENT MICLO[3,c]

[1]*Institut de Mathématiques de Bordeaux, Université de Bordeaux, CNRS and INP Bordeaux, Bordeaux, France,* [a]*marc.arnaudon@math.u-bordeaux.fr*
[2]*Institut Élie Cartan de Lorraine, Université de Lorraine and CNRS, Nancy, France,* [b]*kolehe.coulibaly@univ-lorraine.fr*
[3]*Toulouse School of Economics and Institut de Mathématiques de Toulouse, CNRS and Université de Toulouse, Toulouse, France,* [c]*laurent.miclo@math.cnrs.fr*

This paper proves that the separation convergence toward the uniform distribution abruptly occurs at times around $\ln(n)/n$ for the (time-accelerated by 2) Brownian motion on the sphere with a high dimension n. The arguments are based on a new and elementary perturbative approach for estimating hitting times in a small noise context. The quantitative estimates thus obtained are applied to the strong stationary times constructed in (Arnaudon, Coulibaly-Pasquier and Miclo (2020)) to deduce the wanted cut-off phenomenon.

Keywords: Hitting times; separation discrepancy; small noise one-dimensional diffusions; spherical Brownian motions; strong stationary times

References

Arnaudon, M., Coulibaly-Pasquier, K. and Miclo, L. (2020). Construction of set-valued dual processes on manifolds. HAL Preprint.

Arnaudon, M., Coulibaly-Pasquier, K. and Miclo, L. (2024). Supplement to "On the separation cut-off phenomenon for Brownian motions on high dimensional spheres." https://doi.org/10.3150/23-BEJ1622SUPP

Coulibaly-Pasquier, K. and Miclo, L. (2021). On the evolution by duality of domains on manifolds. *Mém. Soc. Math. Fr. (N.S.)* **171** 110. MR4372500 https://doi.org/10.24033/msmf.47

Diaconis, P. (1996). The cutoff phenomenon in finite Markov chains. *Proc. Natl. Acad. Sci. USA* **93** 1659–1664. MR1374011 https://doi.org/10.1073/pnas.93.4.1659

Diaconis, P. and Fill, J.A. (1990). Strong stationary times via a new form of duality. *Ann. Probab.* **18** 1483–1522. MR1071805

Diaconis, P. and Saloff-Coste, L. (2006). Separation cut-offs for birth and death chains. *Ann. Appl. Probab.* **16** 2098–2122. MR2288715 https://doi.org/10.1214/105051606000000501

Ding, J., Lubetzky, E. and Peres, Y. (2010). Total variation cutoff in birth-and-death chains. *Probab. Theory Related Fields* **146** 61–85. MR2550359 https://doi.org/10.1007/s00440-008-0185-3

Hermon, J., Lacoin, H. and Peres, Y. (2016). Total variation and separation cutoffs are not equivalent and neither one implies the other. *Electron. J. Probab.* **21** Paper No. 44, 36. MR3530321 https://doi.org/10.1214/16-EJP4687

Levin, D.A., Peres, Y. and Wilmer, E.L. (2009). *Markov Chains and Mixing Times*. Providence, RI: Amer. Math. Soc. MR2466937 https://doi.org/10.1090/mbk/058

Méliot, P.-L. (2014). The cut-off phenomenon for Brownian motions on compact symmetric spaces. *Potential Anal.* **40** 427–509. MR3201989 https://doi.org/10.1007/s11118-013-9356-7

Saloff-Coste, L. (1994). Precise estimates on the rate at which certain diffusions tend to equilibrium. *Math. Z.* **217** 641–677. MR1306030 https://doi.org/10.1007/BF02571965

Bernoulli **30**(2), 2024, 1029–1054
https://doi.org/10.3150/23-BEJ1623

A recursive distributional equation for the stable tree

NICHOLAS CHEE[1,a], FRANZ REMBART[2,b] and MATTHIAS WINKEL[3,c]

[1,a]*niccheekn@gmail.com*
[2]*Columbia University, Department of Statistics, 1255 Amsterdam Avenue, New York, NY 10027, USA,* [b]*franz.rembart@gmail.com*
[3]*University of Oxford, Department of Statistics, 24–29 St Giles', Oxford OX1 3LB, UK,* [c]*winkel@stats.ox.ac.uk*

We provide a new characterisation of Duquesne and Le Gall's α-stable tree, $\alpha \in (1,2]$, as the solution of a recursive distributional equation (RDE) of the form $\mathcal{T} \overset{d}{=} g(\xi, \mathcal{T}_i, i \geq 0)$, where g is a concatenation operator, $\xi = (\xi_i, i \geq 0)$ a sequence of scaling factors, $\mathcal{T}_i$, $i \geq 0$, and $\mathcal{T}$ are i.i.d. trees independent of ξ. This generalises the characterisation of the Brownian Continuum Random Tree proved by Albenque and Goldschmidt, based on self-similarity observed by Aldous. By relating to previous results on a rather different class of RDE, we explore the present RDE and obtain for a large class of similar RDEs that the fixpoint is unique (up to multiplication by a constant) and attractive.

Keywords: Gromov–Hausdorff distance; recursive distributional equation; $\mathbb{R}$-tree; stable tree

References

[1] Abraham, R., Delmas, J.-F. and Hoscheit, P. (2013). A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18** 14. MR3035742 https://doi.org/10.1214/EJP.v18-2116
[2] Addario-Berry, L., Dieuleveut, D. and Goldschmidt, C. (2019). Inverting the cut-tree transform. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 1349–1376. MR4010938 https://doi.org/10.1214/18-aihp921
[3] Albenque, M. and Goldschmidt, C. (2015). The Brownian continuum random tree as the unique solution to a fixed point equation. *Electron. Commun. Probab.* **20** 61. MR3399812 https://doi.org/10.1214/ECP.v20-4250
[4] Aldous, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226
[5] Aldous, D. (1994). Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.* **22** 527–545. MR1288122
[6] Aldous, D.J. and Bandyopadhyay, A. (2005). A survey of max-type recursive distributional equations. *Ann. Appl. Probab.* **15** 1047–1110. MR2134098 https://doi.org/10.1214/105051605000000142
[7] Berestycki, J., Berestycki, N. and Schweinsberg, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35** 1835–1887. MR2349577 https://doi.org/10.1214/009117906000001114
[8] Bertoin, J. (2002). Self-similar fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 319–340. MR1899456 https://doi.org/10.1016/S0246-0203(00)01073-6
[9] Blackwell, D. and MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614
[10] Broutin, N. and Sulzbach, H. (2021). Self-similar real trees defined as fixed points and their geometric properties. *Electron. J. Probab.* **26** 88. MR4278599 https://doi.org/10.1214/21-ejp647
[11] Chaumont, L. and Pardo, J.C. (2009). On the genealogy of conditioned stable Lévy forests. *ALEA Lat. Am. J. Probab. Math. Stat.* **6** 261–279. MR2534486
[12] Croydon, D. and Hambly, B. (2008). Self-similarity and spectral asymptotics for the continuum random tree. *Stochastic Process. Appl.* **118** 730–754. MR2411518 https://doi.org/10.1016/j.spa.2007.06.005
[13] Curien, N. and Haas, B. (2013). The stable trees are nested. *Probab. Theory Related Fields* **157** 847–883. MR3129805 https://doi.org/10.1007/s00440-012-0472-x

1350-7265 This research was funded, in whole or in part, by [UKRI].

[14] Dieuleveut, D. (2015). The vertex-cut-tree of Galton-Watson trees converging to a stable tree. *Ann. Appl. Probab.* **25** 2215–2262. MR3349006 https://doi.org/10.1214/14-AAP1047
[15] Duquesne, T. and Le Gall, J.-F. (2002). Random trees, Lévy processes and spatial branching processes. *Astérisque* **281**.
[16] Duquesne, T. and Wang, M. (2017). Decomposition of Lévy trees along their diameter. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 539–593. MR3634265 https://doi.org/10.1214/15-AIHP725
[17] Evans, S.N. (2008). *Probability and Real Trees*. *Lecture Notes in Math.* **1920**. Berlin: Springer. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587 https://doi.org/10.1007/978-3-540-74798-7
[18] Haas, B., Pitman, J. and Winkel, M. (2009). Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Probab.* **37** 1381–1411. MR2546748 https://doi.org/10.1214/08-AOP434
[19] Marchal, P. (2008). A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*. *Discrete Math. Theor. Comput. Sci. Proc., AI* 489–499. Nancy: Assoc. Discrete Math. Theor. Comput. Sci. MR2508809
[20] Miermont, G. (2005). Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields* **131** 341–375. MR2123249 https://doi.org/10.1007/s00440-004-0373-8
[21] Miermont, G. (2009). Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)* **42** 725–781. MR2571957 https://doi.org/10.24033/asens.2108
[22] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Berlin: Springer. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR2245368
[23] Pitman, J. and Winkel, M. (2015). Regenerative tree growth: Markovian embedding of fragmenters, bifurcators, and bead splitting processes. *Ann. Probab.* **43** 2611–2646. MR3395470 https://doi.org/10.1214/14-AOP945
[24] Rembart, F. and Winkel, M. (2018). Recursive construction of continuum random trees. *Ann. Probab.* **46** 2715–2748. MR3846837 https://doi.org/10.1214/17-AOP1237
[25] Rösler, U. and Rüschendorf, L. (2001). The contraction method for recursive algorithms. *Algorithmica* **29** 3–33.

Bernoulli **30**(2), 2024, 1055–1078
https://doi.org/10.3150/23-BEJ1624

Rearranged dependence measures

CHRISTOPHER STROTHMANN[1,a], HOLGER DETTE[2,c] and
KARL FRIEDRICH SIBURG[1,b]

[1]*Department of Mathematics, TU Dortmund University, Vogelpothsweg 87, 44221 Dortmund, Germany,*
[a]*christopher.strothmann@mathematik.tu-dortmund.de,* [b]*karl.f.siburg@mathematik.tu-dortmund.de*
[2]*Department of Mathematics, Ruhr-University Bochum, Universitätsstraße 150, 44780 Bochum, Germany,*
[c]*holger.dette@rub.de*

Most of the popular dependence measures for two random variables X and Y (such as Pearson's and Spearman's correlation, Kendall's τ and Gini's γ) vanish whenever X and Y are independent. However, neither does a vanishing dependence measure necessarily imply independence, nor does a measure equal to 1 imply that one variable is a measurable function of the other. Yet, both properties are natural properties for a convincing dependence measure. In this paper, we present a general approach to transforming a given dependence measure into a new one which exactly characterizes independence as well as functional dependence. Our approach uses the concept of monotone rearrangements as introduced by Hardy and Littlewood and is applicable to a broad class of measures. In particular, we are able to define a rearranged Spearman's ρ and a rearranged Kendall's τ which do attain the value 0 if and only if both variables are independent, and the value 1 if and only if one variable is a measurable function of the other. We also present simple estimators for the rearranged dependence measures, prove their consistency and illustrate their finite sample properties by means of a simulation study and a data example.

Keywords: Coefficient of correlation; copula; decreasing rearrangement; measure of dependence

References

Anevski, D. and Fougères, A.-L. (2019). Limit properties of the monotone rearrangement for density and regression function estimation. *Bernoulli* **25** 549–583. MR3892329 https://doi.org/10.3150/17-bej998

Ansari, J. and Rüschendorf, L. (2021). Sklar's theorem, copula products, and ordering results in factor models. *Depend. Model.* **9** 267–306. MR4327840 https://doi.org/10.1515/demo-2021-0113

Auddy, A., Deb, N. and Nandy, S. (2021). Exact detection thresholds for Chatterjee's correlation. Available at https://arxiv.org/abs/2104.15140.

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

Bennett, C. and Sharpley, R. (1988). *Interpolation of Operators*. *Pure and Applied Mathematics* **129**. Boston, MA: Academic Press. MR0928802

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

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

Camirand Lemyre, F., Carroll, R.J. and Delaigle, A. (2022). Semiparametric estimation of the distribution of episodically consumed foods measured with error. *J. Amer. Statist. Assoc.* **117** 469–481. MR4399099 https://doi.org/10.1080/01621459.2020.1787840

Cao, S. and Bickel, P.J. (2020). Correlations with tailored extremal properties. Available at http://arxiv.org/abs/2008.10177.

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, S.X. and Huang, T.-M. (2007). Nonparametric estimation of copula functions for dependence modelling. *Canad. J. Statist.* **35** 265–282. MR2393609 https://doi.org/10.1002/cjs.5550350205

Chernozhukov, V., Fernández-Val, I. and Galichon, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika* **96** 559–575. MR2538757 https://doi.org/10.1093/biomet/asp030
Chernozhukov, V., Fernández-Val, I. and Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica* **78** 1093–1125. MR2667913 https://doi.org/10.3982/ECTA7880
Chong, K.M. and Rice, N.M. (1971). *Equimeasurable Rearrangements of Functions*. *Queen's Papers in Pure and Applied Mathematics* **28**. Kingston, Ont.: Queen's Univ. MR0372140
Cover, T.M. and Thomas, J.A. (2006). *Elements of Information Theory*, 2nd ed. Hoboken, NJ: Wiley Interscience. MR2239987
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
Deb, N., Ghosal, P. and Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. Available at http://arxiv.org/abs/2010.01768.
Dette, H., Neumeyer, N. and Pilz, K.F. (2006). A simple nonparametric estimator of a strictly monotone regression function. *Bernoulli* **12** 469–490. MR2232727 https://doi.org/10.3150/bj/1151525131
Dette, H., Siburg, K.F. and Stoimenov, P.A. (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat.* **40** 21–41. MR3024030 https://doi.org/10.1111/j.1467-9469.2011.00767.x
Dette, H. and Volgushev, S. (2008). Non-crossing non-parametric estimates of quantile curves. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 609–627. MR2420417 https://doi.org/10.1111/j.1467-9868.2008.00651.x
Dette, H. and Wu, W. (2019). Detecting relevant changes in the mean of nonstationary processes—a mass excess approach. *Ann. Statist.* **47** 3578–3608. MR4025752 https://doi.org/10.1214/19-AOS1811
Durante, F. and Sempi, C. (2016). *Principles of Copula Theory*. Boca Raton, FL: CRC Press. MR3443023
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
Gamboa, F., Gremaud, P., Klein, T. and Lagnoux, A. (2022). Global sensitivity analysis: A novel generation of mighty estimators based on rank statistics. *Bernoulli* **28** 2345–2374. MR4474546 https://doi.org/10.3150/21-bej1421
Geenens, G. and Lafaye de Micheaux, P. (2022). The Hellinger correlation. *J. Amer. Statist. Assoc.* **117** 639–653. MR4436302 https://doi.org/10.1080/01621459.2020.1791132
Genest, C., Nešlehová, J.G. and Rémillard, B. (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions. *J. Multivariate Anal.* **159** 82–110. MR3668549 https://doi.org/10.1016/j.jmva.2017.04.002
Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2008). A kernel statistical test of independence. In *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer and S. Roweis, eds.) **20**. Curran Associates.
Griessenberger, F., Junker, R.R. and Trutschnig, W. (2022). On a multivariate copula-based dependence measure and its estimation. *Electron. J. Stat.* **16** 2206–2251. MR4401220 https://doi.org/10.1214/22-ejs2005
Hardy, G.H., Littlewood, J.E. and Pólya, G. (1988). *Inequalities*. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. Reprint of the 1952 edition. MR0944909
Hofert, M., Kojadinovic, I., Mächler, M. and Yan, J. (2020). copula: Multivariate dependence with copulas. R package version 1.0-1 available at https://CRAN.R-project.org/package=copula.
Junker, R.R., Griessenberger, F. and Trutschnig, W. (2021). Estimating scale-invariant directed dependence of bivariate distributions. *Comput. Statist. Data Anal.* **153** Paper No. 107058, 22. MR4141460 https://doi.org/10.1016/j.csda.2020.107058
Kasper, T., Griessenberger, F., Junker, R.R., Petzel, V. and Trutschnig, W. (2022). qad: Quantification of Asymmetric Dependence R package version 1.0.4 available at https://CRAN.R-project.org/package=qad.
Kinney, J.B. and Atwal, G.S. (2014). Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci. USA* **111** 3354–3359. MR3200177 https://doi.org/10.1073/pnas.1309933111
Lehmann, E.L. (1959). *Testing Statistical Hypotheses*. New York: Wiley; London: CRC Press. MR0107933
Li, X., Mikusiński, P. and Taylor, M.D. (1998). Strong approximation of copulas. *J. Math. Anal. Appl.* **225** 608–623. MR1644300 https://doi.org/10.1006/jmaa.1998.6056
Li, X., Mikusiński, P., Sherwood, H. and Taylor, M.D. (1997). On approximation of copulas. In *Distributions with Given Marginals and Moment Problems (Prague, 1996)* 107–116. Dordrecht: Kluwer Academic. MR1614663

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

Marshall, A.W., Olkin, I. and Arnold, B.C. (2011). *Inequalities: Theory of Majorization and Its Applications*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2759813 https://doi.org/10.1007/978-0-387-68276-1

Mikusiński, P., Sherwood, H. and Taylor, M.D. (1992). Shuffles of Min. *Stochastica* **13** 61–74. MR1197328

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

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

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., Lander, E. S., Mitzenmacher, M. and Sabeti, P. C. (2011). Detecting novel associations in large data sets. *Science* **334** 1518–1524.

Rosenblatt, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.* **3** 1–14. MR0428579

Ryff, J.V. (1965). Orbits of L^1-functions under doubly stochastic transformations. *Trans. Amer. Math. Soc.* **117** 92–100. MR0209866 https://doi.org/10.2307/1994198

Ryff, J.V. (1970). Measure preserving transformations and rearrangements. *J. Math. Anal. Appl.* **31** 449–458. MR0419734 https://doi.org/10.1016/0022-247X(70)90038-7

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

Shi, H., Drton, M. and Han, F. (2021). On Azadkia-Chatterjee's conditional dependence coefficient. Available at http://arxiv.org/abs/2108.06827.

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

Siburg, K.F. and Strothmann, C. (2021). Stochastic monotonicity and the Markov product for copulas. *J. Math. Anal. Appl.* **503** Paper No. 125348, 14. MR4263102 https://doi.org/10.1016/j.jmaa.2021.125348

Spellman, P.T., Gavin, S., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D. and Futcher, B. (1998). Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol. Biol. Cell* **9** 3273–3297.

Stone, C.J. (1984). An asymptotically optimal window selection rule for kernel density estimates. *Ann. Statist.* **12** 1285–1297. MR0760688 https://doi.org/10.1214/aos/1176346792

Strothmann, C., Dette, H. and Siburg, K.F. (2023). RDM: Quantify dependence using rearranged dependence measures. R package version 0.1.1 available at https://cran.r-project.org/package=RDM/.

Strothmann, C., Dette, H. and Siburg, K.F. (2024). Supplement to "Rearranged dependence measures." https://doi.org/10.3150/23-BEJ1624SUPP

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

Team, R.C. (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Trutschnig, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.* **384** 690–705. MR2825218 https://doi.org/10.1016/j.jmaa.2011.06.013

Zhang, K. (2019). BET on independence. *J. Amer. Statist. Assoc.* **114** 1620–1637. MR4047288 https://doi.org/10.1080/01621459.2018.1537921

Bernoulli **30**(2), 2024, 1079–1104
https://doi.org/10.3150/23-BEJ1625

A diffusion approach to Stein's method on Riemannian manifolds

HUILING LE[a], ALEXANDER LEWIS[b], KARTHIK BHARATH[c] and CHRISTOPHER FALLAIZE[d]

School of Mathematical Sciences, University Park, University of Nottingham, Nottingham, UK,
[a]*Huiling.Le@nottingham.ac.uk,* [b]*pmxal12@exmail.nottingham.ac.uk,* [c]*Karthik.Bharath@nottingham.ac.uk,*
[d]*Chris.Fallaize@nottingham.ac.uk*

We detail an approach to developing Stein's method for bounding integral metrics on probability measures defined on a Riemannian manifold $\boldsymbol{M}$. Our approach exploits the relationship between the generator of a diffusion on $\boldsymbol{M}$ having a target invariant measure and its characterising Stein operator. We consider a pair of such diffusions with different starting points, and through analysis of the distance process between the pair, derive Stein factors, which bound the solution to the Stein equation and its derivatives. The Stein factors contain curvature-dependent terms and reduce to those currently available for $\mathbb{R}^m$, and moreover imply that the bounds for $\mathbb{R}^m$ remain valid when $\boldsymbol{M}$ is a flat manifold.

Keywords: Coupling; integral metrics; Stein equation; stochastic flow; Wasserstein distance

References

Azagra, D., Ferrera, J., López-Mesas, F. and Rangel, Y. (2007). Smooth approximation of Lipschitz functions on Riemannian manifolds. *J. Math. Anal. Appl.* **326** 1370–1378. MR2280987 https://doi.org/10.1016/j.jmaa.2006.03.088

Bakry, D. (1986). Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *C. R. Acad. Sci. Paris Sér. I Math.* **303** 23–26. MR0849620

Barbour, A.D. (1988). Stein's method and Poisson process convergence. *J. Appl. Probab.* **25** 175–184. MR0974580 https://doi.org/10.1017/s0021900200040341

Barbour, A.D. and Chen, L.H.Y. (2014). Stein's (magic) method. Preprint. Available at arXiv:1411.1179.

Barden, D. and Le, H. (1997). Some consequences of the nature of the distance function on the cut locus in a Riemannian manifold. *J. Lond. Math. Soc. (2)* **56** 369–383. MR1489143 https://doi.org/10.1112/S002461079700553X

Chatterjee, S. and Meckes, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. MR2453473

Cheeger, J. and Ebin, D.G. (1975). *Comparison Theorems in Riemannian Geometry*. *North-Holland Mathematical Library* **9**. Amsterdam: North-Holland. MR0458335

Chen, P., Nourdin, I., Xu, L. and Yang, X. (2019). Multivariate stable approximation in Wasserstein distance by Stein's method. Preprint. Available at arXiv:1911.12917.

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

Fang, X., Shao, Q.-M. and Xu, L. (2019). Multivariate approximations in Wasserstein distance by Stein's method and Bismut's formula. *Probab. Theory Related Fields* **174** 945–979. MR3980309 https://doi.org/10.1007/s00440-018-0874-5

Greene, R.E. and Wu, H. (1979). *Function Theory on Manifolds Which Possess a Pole*. *Lecture Notes in Math.* **699**. Berlin: Springer. MR0521983

1350-7265 This research was funded, in whole or in part, by [Engineering and Physical Sciences Research Council (EPSRC), UK, EPSRC EP/V048104/1].

Kendall, W.S. (1986a). Stochastic differential geometry, a coupling property, and harmonic maps. *J. Lond. Math. Soc. (2)* **33** 554–566. MR0850971 https://doi.org/10.1112/jlms/s2-33.3.554

Kendall, W.S. (1986b). Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics* **19** 111–129. MR0864339 https://doi.org/10.1080/17442508608833419

Kent, J.T. (1994). The complex Bingham distribution and shape analysis. *J. Roy. Statist. Soc. Ser. B* **56** 285–299. MR1281934

Kobayashi, S. and Nomizu, K. (1963). *Foundations of Differential Geometry. Vol I*. New York: Wiley. MR0152974

Kuwada, K. (2010). Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.* **258** 3758–3774. MR2606871 https://doi.org/10.1016/j.jfa.2010.01.010

Le, H.L. and Barden, D. (1995). Itô correction terms for the radial parts of semimartingales on manifolds. *Probab. Theory Related Fields* **101** 133–146. MR1314177 https://doi.org/10.1007/BF01192198

Le, H., Lewis, A., Bharath, K. and Fallaize, C. (2024). Supplement to "A diffusion approach to Stein's method on Riemannian manifolds." https://doi.org/10.3150/23-BEJ1625SUPP

Lewis, A. (2021). Stein's Method for Probability Distributions on $\mathbb{S}^1$. Preprint. Available at arXiv:2105.13199.

Mackey, L. and Gorham, J. (2016). Multivariate Stein factors for a class of strongly log-concave distributions. *Electron. Commun. Probab.* **21** Paper No. 56, 14 pp. MR3548768 https://doi.org/10.1214/16-ecp15

Meckes, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V: The Luminy Volume*. *Inst. Math. Stat. (IMS) Collect.* **5** 153–178. Beachwood, OH: IMS. MR2797946 https://doi.org/10.1214/09-IMSCOLL511

Mijoule, G., Reinert, G. and Swan, Y. (2019). Stein operators, kernels and discrepancies for multivariate continuous distributions. Preprint. Available at arXiv:1806.03478.

O'Neill, B. (1983). *Semi-Riemannian Geometry: With Applications to Relativity*. *Pure and Applied Mathematics* **103**. New York: Academic Press. MR0719023

Röllin, A. (2012). On the optimality of Stein factors. In *Probability Approximations and Beyond*. *Lect. Notes Stat.* **205** 61–72. New York: Springer. MR3289376 https://doi.org/10.1007/978-1-4614-1966-2_5

Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293. MR2861132 https://doi.org/10.1214/11-PS182

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, CaA: Univ. California Press. MR0402873

Sturm, K.-T. (2006a). On the geometry of metric measure spaces. I. *Acta Math.* **196** 65–131. MR2237206 https://doi.org/10.1007/s11511-006-0002-8

Sturm, K.-T. (2006b). On the geometry of metric measure spaces. II. *Acta Math.* **196** 133–177. MR2237207 https://doi.org/10.1007/s11511-006-0003-7

Thalmaier, A. (1997). On the differentiation of heat semigroups and Poisson integrals. *Stoch. Stoch. Rep.* **61** 297–321. MR1488139 https://doi.org/10.1080/17442509708834123

Thompson, J. (2020). Approximation of Riemannian measures by Stein's method. Preprint. Available at arXiv:2001.009.

von Renesse, M.-K. and Sturm, K.-T. (2005). Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.* **58** 923–940. MR2142879 https://doi.org/10.1002/cpa.20060

Bernoulli **30**(2), 2024, 1105–1153
https://doi.org/10.3150/23-BEJ1626

A new shape of extremal clusters for certain stationary semi-exponential processes with moderate long range dependence

ZAO-LI CHEN[1,a] and GENNADY SAMRODNITSKY[2,b]

[1]*Department of Mathematics, Cornell University, Ithaca, New York 14850, USA,* [a]*zc288@cornell.edu*
[2]*School of Operation Research and Information Engineering, Cornell University, Ithaca, New York 14850, USA,* [b]*gs18@cornell.edu*

Extremal clusters of stationary processes with long memory can be quite intricate. For certain stationary infinitely divisible processes with subexponential tails an extremal cluster may consist of a single extreme value distributed over a stable regenerative set. This happens both in the case of power-like tails and in the case of certain lighter tails, e.g. lognormal-like tails, In this paper we show that in the case of semi-exponential tails, a new shape of extremal clusters arises. In this case each stable regenerative set supports a random panoply of varying extremes.

Keywords: Extreme value theory; long range dependence; random sup-measure; stable regenerative set; subexponential distributions; semi-exponential distributions; Gumbel maximum domain of attraction

References

Aaronson, J. (1997). *An Introduction to Infinite Ergodic Theory*. *Mathematical Surveys and Monographs* **50**. Providence, RI: Amer. Math. Soc. MR1450400 https://doi.org/10.1090/surv/050

Chen, Z. and Samorodnitsky, G. (2020). Extreme value theory for long-range-dependent stable random fields. *J. Theoret. Probab.* **33** 1894–1918. MR4166186 https://doi.org/10.1007/s10959-019-00951-8

Chen, Z. and Samorodnitsky, G. (2022). Extremal clustering under moderate long range dependence and moderately heavy tails. *Stochastic Process. Appl.* **145** 86–116. MR4354404 https://doi.org/10.1016/j.spa.2021.12.001

Čistjakov, V.P. (1964). A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory Probab. Appl.* **9** 640–648.

Doney, R.A. (1997). One-sided local large deviation and renewal theorems in the case of infinite mean. *Probab. Theory Related Fields* **107** 451–465. MR1440141 https://doi.org/10.1007/s004400050093

Durieu, O. and Wang, Y. (2018). A family of random sup-measures with long-range dependence. *Electron. J. Probab.* **23** Paper No. 107. MR3870450 https://doi.org/10.1214/18-ejp235

Embrechts, P., Goldie, C.M. and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahrsch. Verw. Gebiete* **49** 335–347. MR0547833 https://doi.org/10.1007/BF00535504

Foss, S., Korshunov, D. and Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd ed. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR3097424 https://doi.org/10.1007/978-1-4614-7101-1

Goldie, C.M. and Resnick, S. (1988). Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Adv. in Appl. Probab.* **20** 706–718. MR0967994 https://doi.org/10.2307/1427356

Harris, T.E. and Robbins, H. (1953). Ergodic theory of Markov chains admitting an infinite invariant measure. *Proc. Natl. Acad. Sci. USA* **39** 860–864. MR0056873 https://doi.org/10.1073/pnas.39.8.860

Kingman, J.F.C. (1968). The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B* **30** 499–510. MR0254907

Lacaux, C. and Samorodnitsky, G. (2016). Time-changed extremal process as a random sup measure. *Bernoulli* **22** 1979–2000. MR3498020 https://doi.org/10.3150/15-BEJ717

1350-7265

Molchanov, I. (2017). *Theory of Random Sets*, 2nd ed. *Probability Theory and Stochastic Modelling* **87**. London: Springer. MR3751326

Molchanov, I. and Strokorb, K. (2016). Max-stable random sup-measures with comonotonic tail dependence. *Stochastic Process. Appl.* **126** 2835–2859. MR3522303 https://doi.org/10.1016/j.spa.2016.03.004

O'Brien, G.L., Torfs, P.J.J.F. and Vervaat, W. (1990). Stationary self-similar extremal processes. *Probab. Theory Related Fields* **87** 97–119. MR1076958 https://doi.org/10.1007/BF01217748

Pitman, E.J.G. (1980). Subexponential distribution functions. *J. Aust. Math. Soc. A* **29** 337–347. MR0569522

Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes*. *Applied Probability. A Series of the Applied Probability Trust* **4**. New York: Springer. MR0900810 https://doi.org/10.1007/978-0-387-75953-1

Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424

Resnick, S., Samorodnitsky, G. and Xue, F. (2000). Growth rates of sample covariances of stationary symmetric α-stable processes associated with null recurrent Markov chains. *Stochastic Process. Appl.* **85** 321–339. MR1731029 https://doi.org/10.1016/S0304-4149(99)00081-2

Samorodnitsky, G. (2004). Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes. *Ann. Probab.* **32** 1438–1468. MR2060304 https://doi.org/10.1214/009117904000000261

Samorodnitsky, G. (2016). *Stochastic Processes and Long Range Dependence*. *Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3561100 https://doi.org/10.1007/978-3-319-45575-4

Samorodnitsky, G. and Wang, Y. (2019). Extremal theory for long range dependent infinitely divisible processes. *Ann. Probab.* **47** 2529–2562. MR3980927 https://doi.org/10.1214/18-AOP1318

Sato, K. (2013). *Lévy Processes and Infinitely Divisible Distributions*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR3185174

Spitzer, F. (1964). *Principles of Random Walk*. *The University Series in Higher Mathematics*. Princeton: D. Van Nostrand Co., Inc. MR0171290

Taylor, S.J. and Wendel, J.G. (1966). The exact Hausdorff measure of the zero set of a stable process. *Z. Wahrsch. Verw. Gebiete* **6** 170–180. MR0210196 https://doi.org/10.1007/BF00537139

Wang, Y. (2022). Choquet random sup-measures with aggregations. *Extremes* **25** 25–54. MR4376583 https://doi.org/10.1007/s10687-021-00425-3

Whitt, W. (1971). Weak convergence of first passage time processes. *J. Appl. Probab.* **8** 417–422. MR0307335 https://doi.org/10.1017/s0021900200035440

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 **30**(2), 2024, 1154–1176
https://doi.org/10.3150/23-BEJ1627

Asymptotic normality for a modified quadratic variation of the Hermite process

ANTOINE AYACHE[a] and CIPRIAN A. TUDOR[b]

CNRS, Université de Lille, Laboratoire Paul Painlevé UMR 8524, F-59655 Villeneuve d'Ascq, France,
[a]*antoine.ayache@univ-lille.fr,* [b]*ciprian.tudor@univ-lille.fr*

We consider a modified quadratic variation of the Hermite process based on some well-chosen increments of this process. These special increments have the very useful property to be independent and identically distributed up to asymptotically negligible remainders. We prove that this modified quadratic variation satisfies a Central Limit Theorem and we derive its rate of convergence under the Wasserstein distance via Stein-Malliavin calculus. As a consequence, we construct, for the first time in the literature related to Hermite processes, a strongly consistent and asymptotically normal estimator for the Hurst parameter.

Keywords: Asymptotic normality; fractional Brownian motion; Hermite process; Hurst index estimation; multiple Wiener-Itô integrals; Ornstein-Uhlenbeck process; Stein-Malliavin calculus; strong consistency

References

Assaad, O. and Tudor, C.A. (2020). Parameter identification for the Hermite Ornstein-Uhlenbeck process. *Stat. Inference Stoch. Process.* **23** 251–270. MR4123924 https://doi.org/10.1007/s11203-020-09219-z

Ayache, A. (2020). Lower bound for local oscillations of Hermite processes. *Stochastic Process. Appl.* **130** 4593–4607. MR4108464 https://doi.org/10.1016/j.spa.2020.01.009

Bardet, J.-M. and Tudor, C.A. (2010). A wavelet analysis of the Rosenblatt process: Chaos expansion and estimation of the self-similarity parameter. *Stochastic Process. Appl.* **120** 2331–2362. MR2728168 https://doi.org/10.1016/j.spa.2010.08.003

Bourguin, S., Diez, C.-P. and Tudor, C.A. (2021). Limiting behavior of large correlated Wishart matrices with chaotic entries. *Bernoulli* **27** 1077–1102. MR4255227 https://doi.org/10.3150/20-bej1266

Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.* **8** 3. MR1961165 https://doi.org/10.1214/EJP.v8-125

Chronopoulou, A., Tudor, C.A. and Viens, F.G. (2011). Self-similarity parameter estimation and reproduction property for non-Gaussian Hermite processes. *Commun. Stoch. Anal.* **5** 161–185. MR2808541 https://doi.org/10.31390/cosa.5.1.10

Chronopoulou, A., Viens, F.G. and Tudor, C.A. (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electron. J. Stat.* **3** 1393–1435. MR2578831 https://doi.org/10.1214/09-EJS423

Clausel, M., Roueff, F., Taqqu, M.S. and Tudor, C. (2013). High order chaotic limits of wavelet scalograms under long-range dependence. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 979–1011. MR3151747

Clausel, M., Roueff, F., Taqqu, M.S. and Tudor, C. (2014). Wavelet estimation of the long memory parameter for Hermite polynomial of Gaussian processes. *ESAIM Probab. Stat.* **18** 42–76. MR3143733 https://doi.org/10.1051/ps/2012026

Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process.* **4** 199–227. MR1856174 https://doi.org/10.1023/A:1017507306245

Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52. MR0550122 https://doi.org/10.1007/BF00535673

Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 407–436. MR1465796 https://doi.org/10.1016/S0246-0203(97)80099-4

1350-7265

Janson, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge: Cambridge Univ. Press. MR1474726 https://doi.org/10.1017/CBO9780511526169

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

Nourdin, I. and Tran, T.T.D. (2019). Statistical inference for Vasicek-type model driven by Hermite processes. *Stochastic Process. Appl.* **129** 3774–3791. MR3997661 https://doi.org/10.1016/j.spa.2018.10.005

Nualart, D. (2006). *Malliavin Calculus and Related Topics*, Second Edition ed. New York: Springer.

Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. Cambridge: Cambridge Univ. Press.

Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302.

Tudor, C.A. (2013). *Analysis of Variations for Self-Similar Processes*. Cham: Springer.

Tudor, C.A. and Viens, F.G. (2009). Variations and estimators for self-similarity parameters via Malliavin calculus. *Ann. Probab.* **37** 2093–2134. MR2573552 https://doi.org/10.1214/09-AOP459

Bernoulli **30**(2), 2024, 1177–1196
https://doi.org/10.3150/23-BEJ1628

Inverse covariance operators of multivariate nonstationary time series

JONAS KRAMPE[1,2,a] and SUHASINI SUBBA RAO[3,b]

[1]*Department of Economics, University of Mannheim, Mannheim, Germany*
[2]*Department of Statistics and Data Science, Cornell University, Ithaca, USA,* [a]*jonas.krampe@cornell.edu*
[3]*Department of Statistics, Texas A&M University, College Station, USA,* [b]*suhasini@stat.tamu.edu*

For multivariate stationary time series many important properties, such as partial correlation, graphical models and autoregressive representations are encoded in the inverse of its spectral density matrix. This is not true for nonstationary time series, where the pertinent information lies in the inverse infinite dimensional covariance matrix operator associated with the multivariate time series. This necessitates the study of the covariance of a multivariate nonstationary time series and its relationship to its inverse. We show that if the rows/columns of the infinite dimensional covariance matrix decay at a certain rate then the rate (up to a factor) transfers to the rows/columns of the inverse covariance matrix. This is used to obtain a nonstationary autoregressive representation of the time series and a Baxter-type bound between the parameters of the autoregressive infinite representation and the corresponding finite autoregressive projection. The aforementioned results lay the foundation for the subsequent analysis of locally stationary time series. In particular, we show that smoothness properties on the covariance matrix transfer to (i) the inverse covariance (ii) the parameters of the vector autoregressive representation and (iii) the partial covariances. All results are set up in such a way that the constants involved depend only on the eigenvalue of the covariance matrix and can be applied in the high-dimensional settings with non-diverging eigenvalues.

Keywords: Autoregressive parameters; Baxter's inequality; high dimensional time series; local stationarity; partial covariance

References

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

Basu, S. and Subba Rao, S. (2022). Graphical models for nonstationary time series. *Ann. Statist.* To appear.

Brillinger, D.R. (2001). *Time Series: Data Analysis and Theory*. *Classics in Applied Mathematics* **36**. Philadelphia, PA: SIAM. MR1853554 https://doi.org/10.1137/1.9780898719246

Cheng, R. and Pourahmadi, M. (1993). Baxter's inequality and convergence of finite predictors of multivariate stochastic processes. *Probab. Theory Related Fields* **95** 115–124. MR1207310 https://doi.org/10.1007/BF01197341

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. (2000a). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762–1794. MR1835040 https://doi.org/10.1214/aos/1015957480

Dahlhaus, R. (2000b). Graphical interaction models for multivariate time series. *Metrika* **51** 157–172. MR1790930 https://doi.org/10.1007/s001840000055

Dahlhaus, R. and Polonik, W. (2006). Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary processes. *Ann. Statist.* **34** 2790–2824. MR2329468 https://doi.org/10.1214/009053606000000867

Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli* **25** 1013–1044. MR3920364 https://doi.org/10.3150/17-bej1011

Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying ARCH processes. *Ann. Statist.* **34** 1075–1114. MR2278352 https://doi.org/10.1214/009053606000000227

1350-7265

Demko, S., Moss, W.F. and Smith, P.W. (1984). Decay rates for inverses of band matrices. *Math. Comp.* **43** 491–499. MR758197 https://doi.org/10.2307/2008290

Ding, X., Qiu, Z. and Chen, X. (2017). Sparse transition matrix estimation for high-dimensional and locally stationary vector autoregressive models. *Electron. J. Stat.* **11** 3871–3902. MR3714301 https://doi.org/10.1214/17-EJS1325

Ding, X. and Zhou, Z. (2020). Estimation and inference for precision matrices of nonstationary time series. *Ann. Statist.* **48** 2455–2477. MR4134802 https://doi.org/10.1214/19-AOS1894

Ding, X. and Zhou, Z. (2023). Auto-regressive approximations to nonstationary time series with inference and applications. *Ann. Statist.* **51** 1207–1231. MR4630946 https://doi.org/10.1214/23-aos2288

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Rev. Econ. Stat.* **82** 540–554. MR4455743

Hannan, E.J. and Deistler, M. (2012). *The Statistical Theory of Linear Systems*. *Classics in Applied Mathematics* **70**. Philadelphia, PA: SIAM. MR3397291 https://doi.org/10.1137/1.9781611972191.ch1

Karmakar, S., Richter, S. and Wu, W.B. (2022). Simultaneous inference for time-varying models. *J. Econometrics* **227** 408–428. MR4384679 https://doi.org/10.1016/j.jeconom.2021.03.002

Krampe, J., Kreiss, J.-P. and Paparoditis, E. (2021). Bootstrap based inference for sparse high-dimensional time series models. *Bernoulli* **27** 1441–1466. MR4278792 https://doi.org/10.3150/20-bej1239

Krampe, J. and McMurry, T.L. (2021). Estimating Wold matrices and vector moving average processes. *J. Time Series Anal.* **42** 201–221. MR4303045 https://doi.org/10.1111/jtsa.12562

Krampe, J. and Paparoditis, E. (2021). Sparsity concepts and estimation procedures for high-dimensional vector autoregressive models. *J. Time Series Anal.* **42** 554–579. MR4325665 https://doi.org/10.1111/jtsa.12586

Krampe, J. and Paparoditis, E. (2022). Frequency Domain Statistical Inference for High-Dimensional Time Series. Preprint. Available at arXiv:2206.02250.

Krampe, J. and Subba Rao, S. (2024). Supplement to “Inverse covariance operators of multivariate nonstationary time series.” https://doi.org/10.3150/23-BEJ1628SUPP

Künsch, H.R. (1995). A note on causal solutions for locally stationary AR-processes. Technical report.

Liu, Y., Taniguchi, M. and Ombao, H. (2021). Statistical inference for local Granger causality. Preprint. Available at arXiv:2103.00209.

Meyer, M. and Kreiss, J.-P. (2015). On the vector autoregressive sieve bootstrap. *J. Time Series Anal.* **36** 377–397. MR3343007 https://doi.org/10.1111/jtsa.12090

Park, T., Eckley, I.A. and Ombao, H.C. (2014). Estimating time-evolving partial coherence between signals via multivariate locally stationary wavelet processes. *IEEE Trans. Signal Process.* **62** 5240–5250. MR3268108 https://doi.org/10.1109/TSP.2014.2343937

Priestley, M.B. (1981). *Spectral Analysis and Time Series. Vol. 2: Multivariate Series, Prediction and Control*. *Probability and Mathematical Statistics*. London–New York: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0628736

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

Subba Rao, S. (2006). On some nonstationary, nonlinear random processes and their stationary approximations. *Adv. in Appl. Probab.* **38** 1155–1172. MR2285698 https://doi.org/10.1239/aap/1165414596

Truquet, L. (2019). Local stationarity and time-inhomogeneous Markov chains. *Ann. Statist.* **47** 2023–2050. MR3953443 https://doi.org/10.1214/18-AOS1739

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614 https://doi.org/10.1214/12-AOS1043

Wiener, N. and Masani, P. (1958). The prediction theory of multivariate stochastic processes. II. The linear predictor. *Acta Math.* **99** 93–137. MR0097859 https://doi.org/10.1007/BF02392423

Xu, H., Wang, D., Zhao, Z. and Yu, Y. (2022). Change point inference in high-dimensional regression models under temporal dependence. Preprint. Available at arXiv:2207.12453.

Yan, Y., Gao, J. and Peng, B. (2021). On Time-Varying VAR Models: Estimation, Testing and Impulse Response Analysis. Preprint. Available at arXiv:2111.00450.

Zhang, D. and Wu, W.B. (2021). Second-order asymptotics for high dimensional locally stationary processes. *Ann. Statist.* **49** 233–254.

Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. MR2541444 https://doi.org/10.1214/08-AOS636

Bernoulli **30**(2), 2024, 1197–1230
https://doi.org/10.3150/23-BEJ1629

Rough paths and symmetric-Stratonovich integrals driven by singular covariance Gaussian processes

ALBERTO OHASHI[1,a] and FRANCESCO RUSSO[2,b]

[1]*Department of Mathematics, Universidade de Brasília, Brasília, Brazil,* [a]*ohashi@mat.unb.br*
[2]*ENSTA Paris, Unité de Mathématiques appliquées, OC, Institut Polytechnique de Paris, Paris, France,* [b]*francesco.russo@ensta-paris.fr*

We examine the relation between a stochastic version of the rough integral with the symmetric-Stratonovich integral in the sense of regularization. Under mild regularity conditions in the sense of Malliavin calculus, we establish equality between stochastic rough and symmetric-Stratonovich integrals driven by a class of Gaussian processes. As a by-product, we show that solutions of multi-dimensional rough differential equations driven by a large class of Gaussian rough paths they are actually solutions to Stratonovich stochastic differential equations. We obtain almost sure convergence rates of the first-order Stratonovich scheme to rough integrals in the sense of Gubinelli. In case the time-increment of the Malliavin derivative of the integrands is regular enough, the rates are essentially sharp. The framework applies to a large class of Gaussian processes whose the second-order derivative of the covariance function is a sigma-finite non-positive measure on $\mathbb{R}^2_+$ off diagonal.

Keywords: Rough paths; Stratonovich integrals

References

Alòs, E., León, J.A. and Nualart, D. (2001). Stochastic Stratonovich calculus fBm for fractional Brownian motion with Hurst parameter less than $1/2$. *Taiwanese J. Math.* **5** 609–632. MR1849782 https://doi.org/10.11650/twjm/1500574954

Alòs, E., Mazet, O. and Nualart, D. (2001). Stochastic calculus with respect to Gaussian processes. *Ann. Probab.* **29** 766–801. MR1849177 https://doi.org/10.1214/aop/1008956692

Alòs, E. and Nualart, D. (2003). Stochastic integration with respect to the fractional Brownian motion. *Stoch. Stoch. Rep.* **75** 129–152. MR1978896 https://doi.org/10.1080/1045112031000078917

Butkovsky, O., Dareiotis, K. and Gerencsér, M. (2021). Approximation of SDEs: A stochastic sewing approach. *Probab. Theory Related Fields* **181** 975–1034. MR4344136 https://doi.org/10.1007/s00440-021-01080-2

Cass, T., Friz, P. and Victoir, N. (2009). Non-degeneracy of Wiener functionals arising from rough differential equations. *Trans. Amer. Math. Soc.* **361** 3359–3371. MR2485431 https://doi.org/10.1090/S0002-9947-09-04677-7

Cass, T. and Lim, N. (2019). A Stratonovich-Skorohod integral formula for Gaussian rough paths. *Ann. Probab.* **47** 1–60. MR3909965 https://doi.org/10.1214/18-AOP1254

Cass, T. and Lim, N. (2021). Skorohod and rough integration for stochastic differential equations driven by Volterra processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 132–168. MR4255171 https://doi.org/10.1214/20-aihp1074

Cass, T., Litterer, C. and Lyons, T. (2013). Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **41** 3026–3050. MR3112937 https://doi.org/10.1214/12-AOP821

Cass, T., Hairer, M., Litterer, C. and Tindel, S. (2015). Smoothness of the density for solutions to Gaussian rough differential equations. *Ann. Probab.* **43** 188–239. MR3298472 https://doi.org/10.1214/13-AOP896

Cheridito, P. and Nualart, D. (2005). Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter $H \in (0, \frac{1}{2})$. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 1049–1081. MR2172209 https://doi.org/10.1016/j.anihpb.2004.09.004

1350-7265

Coutin, L., Friz, P. and Victoir, N. (2007). Good rough path sequences and applications to anticipating stochastic calculus. *Ann. Probab.* **35** 1172–1193. MR2319719 https://doi.org/10.1214/009117906000000827

Coutin, L. and Qian, Z. (2002). Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields* **122** 108–140. MR1883719 https://doi.org/10.1007/s004400100158

Coviello, R., di Girolami, C. and Russo, F. (2011). On stochastic calculus related to financial assets without semimartingales. *Bull. Sci. Math.* **135** 733–774. MR2838099 https://doi.org/10.1016/j.bulsci.2011.06.008

Coviello, R. and Russo, F. (2007). Nonsemimartingales: Stochastic differential equations and weak Dirichlet processes. *Ann. Probab.* **35** 255–308. MR2303950 https://doi.org/10.1214/009117906000000566

Deya, A., Neuenkirch, A. and Tindel, S. (2012). A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 518–550. MR2954265 https://doi.org/10.1214/10-AIHP392

Errami, M. and Russo, F. (2003). n-covariation, generalized Dirichlet processes and calculus with respect to finite cubic variation processes. *Stochastic Process. Appl.* **104** 259–299. MR1961622 https://doi.org/10.1016/S0304-4149(02)00238-7

Feyel, D. and de La Pradelle, A. (2006). Curvilinear integrals along enriched paths. *Electron. J. Probab.* **11** 860–892. MR2261056 https://doi.org/10.1214/EJP.v11-356

Friz, P.K. and Hairer, M. (2020). *A Course on Rough Paths*. *Universitext*. Cham: Springer. With an introduction to regularity structures. MR4174393 https://doi.org/10.1007/978-3-030-41556-3

Friz, P.K., Hocquet, A. and Lê, K. (2021). Rough stochastic differential equations. Preprint. Available at arXiv:2106.10340.

Friz, P. and Riedel, S. (2014). Convergence rates for the full Gaussian rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 154–194. MR3161527 https://doi.org/10.1214/12-AIHP507

Friz, P.K. and Victoir, N.B. (2010a). *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge: Cambridge Univ. Press. MR2604669 https://doi.org/10.1017/CBO9780511845079

Friz, P. and Victoir, N. (2010b). Differential equations driven by Gaussian signals. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 369–413. MR2667703 https://doi.org/10.1214/09-AIHP202

Friz, P.K., Gess, B., Gulisashvili, A. and Riedel, S. (2016). The Jain-Monrad criterion for rough paths and applications to random Fourier series and non-Markovian Hörmander theory. *Ann. Probab.* **44** 684–738. MR3456349 https://doi.org/10.1214/14-AOP986

Garzón, J., León, J.A. and Torres, S. (2023). Forward integration of bounded variation coefficients with respect to Hölder continuous processes. *Bernoulli* **29** 1877–1904. MR4580900 https://doi.org/10.3150/22-bej1524

Gomes, A.O., Ohashi, A., Russo, F. and Teixeira, A. (2021). Rough paths and regularization. *J. Stoch. Anal.* **2** Art. 1, 20 pp. MR4331260

Gradinaru, M. and Nourdin, I. (2003). Approximation at first and second order of m-order integrals of the fractional Brownian motion and of certain semimartingales. *Electron. J. Probab.* **8** no. 18, 26 pp. MR2041819 https://doi.org/10.1214/EJP.v8-166

Gradinaru, M., Russo, F. and Vallois, P. (2003). Generalized covariations, local time and Stratonovich Itô's formula for fractional Brownian motion with Hurst index $H \geq \frac{1}{4}$. *Ann. Probab.* **31** 1772–1820. MR2016600 https://doi.org/10.1214/aop/1068646366

Gradinaru, M., Nourdin, I., Russo, F. and Vallois, P. (2005). m-order integrals and generalized Itô's formula: The case of a fractional Brownian motion with any Hurst index. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 781–806. MR2144234 https://doi.org/10.1016/j.anihpb.2004.06.002

Gubinelli, M. (2004). Controlling rough paths. *J. Funct. Anal.* **216** 86–140. MR2091358 https://doi.org/10.1016/j.jfa.2004.01.002

Hu, Y., Jolis, M. and Tindel, S. (2013). On Stratonovich and Skorohod stochastic calculus for Gaussian processes. *Ann. Probab.* **41** 1656–1693. MR3098687 https://doi.org/10.1214/12-AOP751

Hu, Y., Liu, Y. and Nualart, D. (2016). Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions. *Ann. Appl. Probab.* **26** 1147–1207. MR3476635 https://doi.org/10.1214/15-AAP1114

Kruk, I. and Russo, F. (2010). Malliavin-Skorohod calculus and Paley-Wiener integral for covariance singular processes. Preprint. Available at arXiv:1011.6478v1.

Kruk, I., Russo, F. and Tudor, C.A. (2007). Wiener integrals, Malliavin calculus and covariance measure structure. *J. Funct. Anal.* **249** 92–142. MR2338856 https://doi.org/10.1016/j.jfa.2007.03.031

Lê, K. (2020). A stochastic sewing lemma and applications. *Electron. J. Probab.* **25** Paper No. 38, 55 pp. MR4089788 https://doi.org/10.1214/20-ejp442

Liu, Y., Selk, Z. and Tindel, S. (2020). Convergence of trapezoid rule to rough integrals. Preprint. Available at arXiv:2005.06500.

Liu, Y. and Tindel, S. (2019). First-order Euler scheme for SDEs driven by fractional Brownian motions: The rough case. *Ann. Appl. Probab.* **29** 758–826. MR3910017 https://doi.org/10.1214/17-AAP1374

Lyons, T.J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. MR1654527 https://doi.org/10.4171/RMI/240

Matsuda, T. and Perkowski, N. (2022). An extension of the stochastic sewing lemma and applications to fractional stochastic calculus. Preprint. Available at arXiv:2206.01686.

Neuenkirch, A. and Nourdin, I. (2007). Exact rate of convergence of some approximation schemes associated to SDEs driven by a fractional Brownian motion. *J. Theoret. Probab.* **20** 871–899. MR2359060 https://doi.org/10.1007/s10959-007-0083-0

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Berlin: Springer. MR2200233

Nualart, D. and Pardoux, É. (1988). Stochastic calculus with anticipating integrands. *Probab. Theory Related Fields* **78** 535–581. MR0950346 https://doi.org/10.1007/BF00353876

Ohashi, A. and Russo, F. (2024). Supplement to "Rough paths and symmetric-Stratonovich integrals driven by singular covariance Gaussian processes." https://doi.org/10.3150/23-BEJ1629SUPP

Perkowski, N. and Prömel, D.J. (2016). Pathwise stochastic integrals for model free finance. *Bernoulli* **22** 2486–2520. MR3498035 https://doi.org/10.3150/15-BEJ735

Perlman, M.D. (1974). Jensen's inequality for a convex vector-valued function on an infinite-dimensional space. *J. Multivariate Anal.* **4** 52–65. MR0362421 https://doi.org/10.1016/0047-259X(74)90005-0

Russo, F. and Tudor, C.A. (2006). On bifractional Brownian motion. *Stochastic Process. Appl.* **116** 830–856. MR2218338 https://doi.org/10.1016/j.spa.2005.11.013

Russo, F. and Vallois, P. (1993). Forward, backward and symmetric stochastic integration. *Probab. Theory Related Fields* **97** 403–421. MR1245252 https://doi.org/10.1007/BF01195073

Russo, F. and Vallois, P. (2007). Elements of stochastic calculus via regularization. In *Séminaire de Probabilités XL*. *Lecture Notes in Math.* **1899** 147–185. Berlin: Springer. MR2409004 https://doi.org/10.1007/978-3-540-71189-6_7

Russo, F. and Vallois, P. (2022). *Stochastic Calculus via Regularizations*. *Bocconi & Springer Series* **11**. Cham: Springer. MR4559653 https://doi.org/10.1007/978-3-031-09446-0

Song, J. and Tindel, S. (2022). Skorohod and Stratonovich integrals for controlled processes. *Stochastic Process. Appl.* **150** 569–595. MR4426165 https://doi.org/10.1016/j.spa.2022.05.002

Wong, E. and Zakai, M. (1965). On the relation between ordinary and stochastic differential equations. *Internat. J. Engrg. Sci.* **3** 213–229. MR0183023 https://doi.org/10.1016/0020-7225(65)90045-5

Bernoulli **30**(2), 2024, 1231–1256
https://doi.org/10.3150/23-BEJ1630

Mean stationarity test in time series: A signal variance-based approach

HON KIU TO[a] and KIN WAI CHAN[b]

Department of Statistics, The Chinese University of Hong Kong, Shatin, Hong Kong, [a]*tojames@link.cuhk.edu.hk*, [b]*kinwaichan@cuhk.edu.hk*

Inference of mean structure is an important problem in time series analysis. Various tests have been developed to test for different mean structures, for example, the presence of structural breaks, and parametric mean structures. However, many of them are designed for handling specific mean structures, and may lose power upon violation of such structural assumptions. In this paper, we propose a new mean stationarity test built around the signal variance. The proposed test is based on a super-efficient estimator which could achieve a convergence rate faster than $\sqrt{n}$. It can detect non-constancy of the mean function under serial dependence. It is shown to have promising power, especially in detecting hardly noticeable oscillating structures. The proposal is further generalized to test for smooth trend structures and relative signal variability.

Keywords: Difference variate; mean stationarity; non-linear time series; relative variability; signal variance; super-efficiency

References

Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59** 817–858. MR1106513 https://doi.org/10.2307/2938229

Andrews, D.W.K. (1995). Nonparametric kernel estimation for semiparametric models. *Econometric Theory* **11** 560–596. MR1349935 https://doi.org/10.1017/S0266466600009427

Antoch, J. and Jarušková, D. (2013). Testing for multiple change points. *Comput. Statist.* **28** 2161–2183. MR3107296 https://doi.org/10.1007/s00180-013-0401-1

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66** 47–78. MR1616121 https://doi.org/10.2307/2998540

Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** 107–144. Update of, and a supplement to, the 1986 original. MR2178042 https://doi.org/10.1214/154957805100000104

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

Casini, A. (2023). Theory of evolutionary spectra for heteroskedasticity and autocorrelation robust inference in possibly misspecified and nonstationary models. *J. Econometrics* **235** 372–392. MR4602874 https://doi.org/10.1016/j.jeconom.2022.05.001

Casini, A. and Perron, P. (2021). Prewhitened long-run variance estimation robust to nonstationarity. Manuscript.

Chan, K.W. (2022a). Optimal difference-based variance estimators in time series: A general framework. *Ann. Statist.* **50** 1376–1400. MR4441124 https://doi.org/10.1214/21-aos2154

Chan, K.W. (2022b). Mean-structure and autocorrelation consistent covariance matrix estimation. *J. Bus. Econom. Statist.* **40** 201–215. MR4356567 https://doi.org/10.1080/07350015.2020.1796397

Chan, K.W. and Yau, C.Y. (2017). High-order corrected estimator of asymptotic variance with optimal bandwidth. *Scand. J. Stat.* **44** 866–898. MR3730019 https://doi.org/10.1111/sjos.12279

Chan, K.W. and Yau, C.Y. (2023). Asymptotically constant risk estimator of time-average variance constant. Manuscript.

1350-7265

Chen, L., Wang, W. and Wu, W.B. (2022). Inference of breakpoints in high-dimensional time series. *J. Amer. Statist. Assoc.* **117** 1951–1963. MR4528482 https://doi.org/10.1080/01621459.2021.1893178
Chen, L. and Wu, W.B. (2019). Testing for trends in high-dimensional time series. *J. Amer. Statist. Assoc.* **114** 869–881. MR3963187 https://doi.org/10.1080/01621459.2018.1456935
Cheng, C.H. and Chan, K.W. (2023). A general framework for constructing locally self-normalized multiple-change-point tests. Manuscript.
Crainiceanu, C.M. and Vogelsang, T.J. (2007). Nonmonotonic power for tests of a mean shift in a time series. *J. Stat. Comput. Simul.* **77** 457–476. MR2405424 https://doi.org/10.1080/10629360600569394
Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. With a foreword by David Kendall. MR2743035
Dalla, V., Giraitis, L. and Phillips, P.C.B. (2015). Testing mean stability of heteroskedastic time series. Cowles Foundation Discussion Papers 2441.
Dedecker, J. and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields* **132** 203–236. MR2199291 https://doi.org/10.1007/s00440-004-0394-3
Dette, H. and Wied, D. (2016). Detecting relevant changes in time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 371–394. MR3454201 https://doi.org/10.1111/rssb.12121
Dette, H. and Wu, W. (2019). Detecting relevant changes in the mean of nonstationary processes—a mass excess approach. *Ann. Statist.* **47** 3578–3608. MR4025752 https://doi.org/10.1214/19-AOS1811
Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.* **42** 2243–2281. MR3269979 https://doi.org/10.1214/14-AOS1245
Gallant, A.R. (1987). *Nonlinear Statistical Models*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR0921029 https://doi.org/10.1002/9780470316719
Górecki, T., Horváth, L. and Kokoszka, P. (2018). Change point detection in heteroscedastic time series. *Econom. Stat.* **7** 63–88. MR3824127 https://doi.org/10.1016/j.ecosta.2017.07.005
Hall, P., Kay, J.W. and Titterington, D.M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika* **77** 521–528. MR1087842 https://doi.org/10.1093/biomet/77.3.521
Hobijn, B., Franses, P.H. and Ooms, M. (2004). Generalizations of the KPSS-test for stationarity. *Stat. Neerl.* **58** 483–502. MR2106350 https://doi.org/10.1111/j.1467-9574.2004.00272.x
Horváth, L., Kokoszka, P. and Steinebach, J. (1999). Testing for changes in multivariate dependent observations with an application to temperature changes. *J. Multivariate Anal.* **68** 96–119. MR1668911 https://doi.org/10.1006/jmva.1998.1780
Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Teor. Veroyatn. Primen.* **7** 361–392. MR0148125
Jiang, F., Zhao, Z. and Shao, X. (2022). Modelling the COVID-19 infection trajectory: A piecewise linear quantile trend model. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1589–1607. MR4515551
Juhl, T. and Xiao, Z. (2009). Tests for changing mean with monotonic power. *J. Econometrics* **148** 14–24. MR2494814 https://doi.org/10.1016/j.jeconom.2008.08.020
Karmakar, S. and Wu, W.B. (2020). Optimal Gaussian approximation for multiple time series. *Statist. Sinica* **30** 1399–1417. MR4257539 https://doi.org/10.5705/ss.202017.0303
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econometrics* **54** 159–178.
Liu, X. and Chan, K.W. (2023). No-lose converging kernel estimation of long-run variance. Manuscript.
Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Stat.* **28** 329–348. MR0088833 https://doi.org/10.1214/aoms/1177706962
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. and Zhang, X. (2010). Testing for change points in time series. *J. Amer. Statist. Assoc.* **105** 1228–1240. MR2752617 https://doi.org/10.1198/jasa.2010.tm10103
To, H.K. and Chan, K.W. (2024). Supplement to "Mean stationarity test in time series: A signal variance-based approach." https://doi.org/10.3150/23-BEJ1630SUPP

Vats, D. and Flegal, J.M. (2022). Lugsail lag windows for estimating time-average covariance matrices. *Biometrika* **109** 735–750. MR4472845 https://doi.org/10.1093/biomet/asab049

Wu, W.B. (2004). A test for detecting changes in mean. In *Time Series Analysis and Applications to Geophysical Systems* (D.R. Brillinger, E.A. Robinson and F. Schoenberg, eds.) **45** 105–121. New York: Springer-Verlag.

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

Wu, W.B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320. MR2353389 https://doi.org/10.1214/009117907000000060

Wu, W.B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4** 207–226. MR2812816 https://doi.org/10.4310/SII.2011.v4.n2.a15

Wu, W.B., Woodroofe, M. and Mentz, G. (2001). Isotonic regression: Another look at the changepoint problem. *Biometrika* **88** 793–804. MR1859410 https://doi.org/10.1093/biomet/88.3.793

Wu, W.B. and Zhao, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. MR2323759 https://doi.org/10.1111/j.1467-9868.2007.00594.x

Zhang, T. and Lavitas, L. (2018). Unsupervised self-normalized change-point testing for time series. *J. Amer. Statist. Assoc.* **113** 637–648. MR3832215 https://doi.org/10.1080/01621459.2016.1270214

Zhang, T. and Wu, W.B. (2011). Testing parametric assumptions of trends of a nonstationary time series. *Biometrika* **98** 599–614. MR2836409 https://doi.org/10.1093/biomet/asr017

Bernoulli **30**(2), 2024, 1257–1286
https://doi.org/10.3150/23-BEJ1631

Bayesian estimation of nonlinear Hawkes processes

DÉBORAH SULEM[1,a], VINCENT RIVOIRARD[2,b] and JUDITH ROUSSEAU[3,c]

[1]*Barcelona School of Economics, Barcelona, Spain,* [a]*deborah.sulem@bse.eu*
[2]*Ceremade, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, Paris, France,*
[b]*Vincent.Rivoirard@dauphine.fr*
[3]*University of Oxford, Department of Statistics, Oxford, United Kingdom,* [c]*judith.rousseau@stats.ox.ac.uk*

Multivariate point processes (MPPs) are widely applied to model the occurrences of events, e.g., natural disasters, online message exchanges, financial transactions or neuronal spike trains. In the Hawkes process model, the probability of occurrences of future events depend on the past of the process. This model is particularly popular for modelling interactive phenomena such as disease expansion. In this work we consider the nonlinear multivariate Hawkes model, which allows to account for *excitation* and *inhibition* between interacting entities. We provide theoretical guarantees for applying nonparametric Bayesian estimation methods in this context. In particular, we obtain concentration rates of the posterior distribution on the parameters, under mild assumptions on the prior distribution and the model. These results also lead to convergence rates of Bayesian estimators. Another object of interest in event-data modelling is to infer the *graph of interaction* - or Granger causal graph. In this case, we provide consistency guarantees; in particular, we prove that the posterior distribution is consistent on the graph adjacency matrix of the process, as well as a Bayesian estimator based on an adequate loss function.

Keywords: Nonlinear Hawkes processes; nonparametric Bayesian inference; Granger-causal Graph

References

Apostolopoulou, I., Linderman, S., Miller, K. and Dubrawski, A. (2019). Mutually regressive point processes. *Adv. Neural Inf. Process. Syst.* **32**.

Arbel, J., Gayraud, G. and Rousseau, J. (2013). Bayesian optimal adaptive estimation using a sieve prior. *Scand. J. Stat.* **40** 549–570. MR3091697 https://doi.org/10.1002/sjos.12002

Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.* **123** 2475–2499. MR3054533 https://doi.org/10.1016/j.spa.2013.04.007

Bacry, E., Bompaire, M., Gaïffas, S. and Muzy, J.-F. (2020). Sparse and low-rank multivariate Hawkes processes. *J. Mach. Learn. Res.* **21** Paper No. 50. MR4095329

Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24** 1563–1588. MR1411506 https://doi.org/10.1214/aop/1065725193

Brémaud, P., Nappo, G. and Torrisi, G.L. (2002). Rate of convergence to equilibrium of marked Hawkes processes. *J. Appl. Probab.* **39** 123–136. MR1895148 https://doi.org/10.1017/s0021900200021562

Carstensen, L., Sandelin, A., Winther, O. and Hansen, N.R. (2010). Multivariate Hawkes process models of the occurrence of regulatory elements. *BMC Bioinform.* **11** 456.

Chen, S., Witten, D. and Shojaie, A. (2017). Nearly assumptionless screening for the mutually-exciting multivariate Hawkes process. *Electron. J. Stat.* **11** 1207–1234. MR3634334 https://doi.org/10.1214/17-EJS1251

Chen, S., Shojaie, A., Shea-Brown, E. and Witten, D. (2017). The multivariate Hawkes process in high dimensions: Beyond mutual excitation. Available at arXiv:1707.04928v2.

Chornoboy, E.S., Schramm, L.P. and Karr, A.F. (1988). Maximum likelihood identification of neural point process systems. *Biol. Cybernet.* **59** 265–275. MR0961117 https://doi.org/10.1007/BF00332915

1350-7265

Costa, M., Graham, C., Marsalle, L. and Tran, V.C. (2020). Renewal in Hawkes processes with self-excitation and inhibition. *Adv. in Appl. Probab.* **52** 879–915. MR4153592 https://doi.org/10.1017/apr.2020.19

Dassios, A. and Zhao, H. (2011). A dynamic contagion process. *Adv. in Appl. Probab.* **43** 814–846. MR2858222 https://doi.org/10.1239/aap/1316792671

Delattre, S. and Fournier, N. (2016). Statistical inference versus mean field limit for Hawkes processes. *Electron. J. Stat.* **10** 1223–1295. MR3499526 https://doi.org/10.1214/16-EJS1142

Delattre, S., Fournier, N. and Hoffmann, M. (2016). Hawkes processes on large networks. *Ann. Appl. Probab.* **26** 216–261. MR3449317 https://doi.org/10.1214/14-AAP1089

Deutsch, I. and Ross, G.J. (2022). Bayesian estimation of multivariate Hawkes processes with inhibition and sparsity. ArXiv preprint. Available at arXiv:2201.05009.

Didelez, V. (2008). Graphical models for marked point processes based on local independence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 245–264. MR2412641 https://doi.org/10.1111/j.1467-9868.2007.00634.x

Donnet, S., Rivoirard, V. and Rousseau, J. (2020). Nonparametric Bayesian estimation for multivariate Hawkes processes. *Ann. Statist.* **48** 2698–2727. MR4152118 https://doi.org/10.1214/19-AOS1903

Du, N., Farajtabar, M., Ahmed, A., Smola, A.J. and Song, L. (2015). Dirichlet-Hawkes processes with applications to vlustering vontinuous-time document streams. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. *KDD'15* 219–228. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2783258.2783411

Du, N., Dai, H., Trivedi, R., Upadhyay, U., Gomez-Rodriguez, M. and Song, L. (2016). Recurrent marked temporal point processes: Embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1555–1564.

Eichler, M., Dahlhaus, R. and Dueck, J. (2017). Graphical modeling for multivariate Hawkes processes with nonparametric link functions. *J. Time Series Anal.* **38** 225–242. MR3611742 https://doi.org/10.1111/jtsa.12213

Embrechts, P., Liniger, T. and Lin, L. (2011). Multivariate Hawkes processes: An application to financial data. *J. Appl. Probab.* **48A** 367–378. MR2865638 https://doi.org/10.1239/jap/1318940477

Ertekin, Ş., Rudin, C. and McCormick, T.H. (2015). Reactive point processes: A new approach to predicting power failures in underground electrical systems. *Ann. Appl. Stat.* **9** 122–144. MR3341110 https://doi.org/10.1214/14-AOAS789

Farajtabar, M., Wang, Y., Gomez Rodriguez, M., Li, S., Zha, H. and Song, L. (2015). Coevolve: A joint point process model for information diffusion and network co-evolution. *Adv. Neural Inf. Process. Syst.* **28**.

Gao, F. and Zhu, L. (2018a). Some asymptotic results for nonlinear Hawkes processes. *Stochastic Process. Appl.* **128** 4051–4077. MR3906978 https://doi.org/10.1016/j.spa.2018.01.007

Gao, X. and Zhu, L. (2018b). Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues. *Queueing Syst.* **90** 161–206. MR3850052 https://doi.org/10.1007/s11134-018-9570-5

Gerhard, F., Deger, M. and Truccolo, W. (2017). On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs. *PLoS Comput. Biol.* **13** e1005390. https://doi.org/10.1371/journal.pcbi.1005390

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

Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist.* **35** 192–223. MR2332274 https://doi.org/10.1214/009053606000001172

Graham, C. (2021). Regenerative properties of the linear Hawkes process with unbounded memory. *Ann. Appl. Probab.* **31** 2844–2863. MR4350975 https://doi.org/10.1214/21-aap1664

Granger, C.W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 424–438.

Gusto, G. and Schbath, S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes' model. *Stat. Appl. Genet. Mol. Biol.* **4** Art. 24. MR2170440 https://doi.org/10.2202/1544-6115.1119

Hansen, N.R., Reynaud-Bouret, P. and Rivoirard, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21** 83–143. MR3322314 https://doi.org/10.3150/13-BEJ562

Hawkes, A.G. (1971). Point spectra of some mutually exciting point processes. *J. Roy. Statist. Soc. Ser. B* **33** 438–443. MR0358976

Hillairet, C., Huang, L., Khabou, M. and Réveillac, A. (2022). The Malliavin-Stein method for Hawkes functionals. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 1293–1328. MR4512148 https://doi.org/10.30757/alea.v19-52

Isham, V. and Westcott, M. (1979). A self-correcting point process. *Stochastic Process. Appl.* **8** 335–347. MR0535308 https://doi.org/10.1016/0304-4149(79)90008-5

Karabash, D. (2012). On stability of Hawkes process. ArXiv preprint. Available at arXiv:1201.1573.

Karabash, D. and Zhu, L. (2015). Limit theorems for marked Hawkes processes with application to a risk model. *Stoch. Models* **31** 433–451. MR3395721 https://doi.org/10.1080/15326349.2015.1024868

Lambert, R., Tuleau-Malot, C., Bessaih, T., Rivoirard, V., Bouret, Y., Leresche, N. and Reynaud-Bouret, P. (2017). Reconstructing the functional connectivity of multiple spike trains using Hawkes models. *J. Neurosci. Methods* **297**. https://doi.org/10.1016/j.jneumeth.2017.12.026

Lewis, E. and Mohler, G. (2011). A nonparametric EM algorithm for multiscale Hawkes processes. *J. Nonparametr. Stat.* **1** 1–20.

Malem-Shinitski, N., Ojeda, C. and Opper, M. (2022). Variational Bayesian inference for nonlinear Hawkes process with Gaussian process self-effects. *Entropy* **24** Paper No. 356. MR4406556 https://doi.org/10.3390/e24030356

Massoulié, L. (1998). Stability results for a general class of interacting point processes dynamics, and applications. *Stochastic Process. Appl.* **75** 1–30. MR1629010 https://doi.org/10.1016/S0304-4149(98)00006-4

Mei, H. and Eisner, J.M. (2017). The neural Hawkes process: A neurally self-modulating multivariate point process. *Adv. Neural Inf. Process. Syst.* **30**.

Menon, A. and Lee, Y. (2018). Proper loss functions for nonlinear Hawkes processes. In *Proceedings of the AAAI Conference on Artificial Intelligence* **32**.

Miscouridou, X., Caron, F. and Teh, Y.W. (2018). Modelling sparsity, heterogeneity, reciprocity and community structure in temporal interaction data. *Adv. Neural Inf. Process. Syst.* **31**.

Møller, J. and Rasmussen, J.G. (2005). Perfect simulation of Hawkes processes. *Adv. in Appl. Probab.* **37** 629–646. MR2156552 https://doi.org/10.1239/aap/1127483739

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.* **83** 9–27.

Raad, M.B. (2019). Renewal time points for Hawkes processes. ArXiv preprint. Available at arXiv:1906.02036.

Raad, M.B., Ditlevsen, S. and Löcherbach, E. (2020). Stability and mean-field limits of age dependent Hawkes processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1958–1990. MR4116713 https://doi.org/10.1214/19-AIHP1023

Rasmussen, J.G. (2013). Bayesian inference for Hawkes processes. *Methodol. Comput. Appl. Probab.* **15** 623–642. MR3085883 https://doi.org/10.1007/s11009-011-9272-5

Reynaud-Bouret, P. and Roy, E. (2006). Some non asymptotic tail estimates for Hawkes processes. *Bull. Belg. Math. Soc. Simon Stevin* **13** 883–896. MR2293215

Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.* **38** 2781–2822. MR2722456 https://doi.org/10.1214/10-AOS806

Reynaud-Bouret, P., Rivoirard, V., Grammont, F. and Tuleau-Malot, C. (2014). Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *J. Math. Neurosci.* **4** Art. 3. MR3197017 https://doi.org/10.1186/2190-8567-4-3

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

Stone, C.J. (1994). The use of polynomial splines and their tensor products in multivariate function estimation. *Ann. Statist.* **22** 118–184. With discussion by Andreas Buja and Trevor Hastie and a rejoinder by the author. MR1272079 https://doi.org/10.1214/aos/1176325361

Sulem, D., Rivoirard, V. and Rousseau, J. (2024). Supplement to "Bayesian estimation of nonlinear Hawkes processes." https://doi.org/10.3150/23-BEJ1631SUPP

Torrisi, G.L. (2016). Gaussian approximation of nonlinear Hawkes processes. *Ann. Appl. Probab.* **26** 2106–2140. MR3543891 https://doi.org/10.1214/15-AAP1141

Torrisi, G.L. (2017). Poisson approximation of point processes with stochastic intensity, and application to nonlinear Hawkes processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 679–700. MR3634270 https://doi.org/10.1214/15-AIHP730

Truccolo, W., Eden, U.T., Fellows, M.R., Donoghue, J.P. and Brown, E.N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J. Neurophysiol.* **93** 1074–1089. https://doi.org/10.1152/jn.00697.2004

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. MR2418663 https://doi.org/10.1214/009053607000000613

van der Vaart, A.W. and van Zanten, J.H. (2009). Adaptive Bayesian estimation using a Gaussian random field with inverse gamma bandwidth. *Ann. Statist.* **37** 2655–2675. MR2541442 https://doi.org/10.1214/08-AOS678

Veen, A. and Schoenberg, F.P. (2008). Estimation of space-time branching process models in seismology using an EM-type algorithm. *J. Amer. Statist. Assoc.* **103** 614–624. MR2523998 https://doi.org/10.1198/016214508000000148

Wang, Y., Xie, B., Du, N. and Song, L. (2016). Isotonic Hawkes processes. In *International Conference on Machine Learning* 2226–2234.

Xu, H., Farajtabar, M. and Zha, H. (2016). Learning granger causality for Hawkes processes. In *33rd International Conference on Machine Learning, ICML 2016* **4** 2576–2588.

Zhou, F., Luo, S., Li, Z., Fan, X., Wang, Y., Sowmya, A. and Chen, F. (2021). Efficient EM-variational inference for nonparametric Hawkes process. *Stat. Comput.* **31** 1–11.

Zhou, F., Kong, Q., Zhang, Y., Feng, C. and Zhu, J. (2021). Nonlinear Hawkes processes in time-varying system. ArXiv preprint. Available at arXiv:2106.04844.

Zhou, F., Kong, Q., Deng, Z., Kan, J., Zhang, Y., Feng, C. and Zhu, J. (2022). Efficient inference for dynamic flexible interactions of neural populations. *J. Mach. Learn. Res.* **23** Paper No. 211. MR4577164

Bernoulli **30**(2), 2024, 1287–1312
https://doi.org/10.3150/23-BEJ1632

Optimal weighted pooling for inference about the tail index and extreme quantiles

ABDELAATI DAOUIA[1,a], SIMONE A. PADOAN[2,b] and GILLES STUPFLER[3,c]

[1] *Toulouse School of Economics, University of Toulouse Capitole, France,* [a] *abdelaati.daouia@tse-fr.eu*
[2] *Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milano, Italy,* [b] *simone.padoan@unibocconi.it*
[3] *Univ Angers, CNRS, LAREMA, SFR MATHSTIC, F-49000 Angers, France,* [c] *gilles.stupfler@univ-angers.fr*

This paper investigates pooling strategies for tail index and extreme quantile estimation from heavy-tailed data. To fully exploit the information contained in several samples, we present general weighted pooled Hill estimators of the tail index and weighted pooled Weissman estimators of extreme quantiles calculated through a nonstandard geometric averaging scheme. We develop their large-sample asymptotic theory across a fixed number of samples, covering the general framework of heterogeneous sample sizes with different and asymptotically dependent distributions. Our results include optimal choices of pooling weights based on asymptotic variance and MSE minimization. In the important application of distributed inference, we prove that the variance-optimal distributed estimators are asymptotically equivalent to the benchmark Hill and Weissman estimators based on the unfeasible combination of subsamples, while the AMSE-optimal distributed estimators enjoy a smaller AMSE than the benchmarks in the case of large bias. We consider additional scenarios where the number of subsamples grows with the total sample size and effective subsample sizes can be low. We extend our methodology to handle serial dependence and the presence of covariates. Simulations confirm the statistical inferential theory of our pooled estimators. Two applications to real weather and insurance data are showcased.

Keywords: Extreme values; heavy tails; inference; pooling; testing

References

[1] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR2108013 https://doi.org/10.1002/0470012382
[2] Chen, L., Li, D. and Zhou, C. (2021). Distributed inference for tail empirical and quantile processes. Available at arXiv:2108.01327v1.
[3] Chen, L., Li, D. and Zhou, C. (2022). Distributed inference for the extreme value index. *Biometrika* **109** 257–264. MR4374653 https://doi.org/10.1093/biomet/asab001
[4] Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. *Suppl. J. R. Stat. Soc.* **4** 102–118.
[5] Daouia, A., Padoan, S.A. and Stupfler, G. (2024). Supplement to "Optimal weighted pooling for inference about the tail index and extreme quantiles." https://doi.org/10.3150/23-BEJ1632SUPP
[6] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3
[7] Dematteo, A. and Clémençon, S. (2016). On tail index estimation based on multivariate data. *J. Nonparametr. Stat.* **28** 152–176. MR3463555 https://doi.org/10.1080/10485252.2015.1124105
[8] Drees, H. (1998). Optimal rates of convergence for estimates of the extreme value index. *Ann. Statist.* **26** 434–448. MR1608148 https://doi.org/10.1214/aos/1030563992
[9] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9** 617–657. MR1996273 https://doi.org/10.3150/bj/1066223272

1350-7265

[10] Einmahl, J.H.J., de Haan, L. and Zhou, C. (2016). Statistics of heteroscedastic extremes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 31–51. MR3453645 https://doi.org/10.1111/rssb.12099

[11] Einmahl, J.H.J., Yang, F. and Zhou, C. (2021). Testing the multivariate regular variation model. *J. Bus. Econom. Statist.* **39** 907–919. MR4319681 https://doi.org/10.1080/07350015.2020.1737533

[12] Girard, S., Stupfler, G. and Usseglio-Carleve, A. (2021). Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models. *Ann. Statist.* **49** 3358–3382. MR4352533 https://doi.org/10.1214/21-aos2087

[13] Girard, S., Stupfler, G. and Usseglio-Carleve, A. (2022). On automatic bias reduction for extreme expectile estimation. *Stat. Comput.* **32** 64. MR4466976 https://doi.org/10.1007/s11222-022-10118-x

[14] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174. MR0378204

[15] Kinsvater, P., Fried, R. and Lilienthal, J. (2016). Regional extreme value index estimation and a test of tail homogeneity. *Environmetrics* **27** 103–115. MR3481322 https://doi.org/10.1002/env.2376

[16] Knight, K. and Bassett, G.W. (2003). Second order improvements of sample quantiles using subsamples. Unpublished manuscript.

[17] Padoan, S.A. and Stupfler, G. (2022). Joint inference on extreme expectiles for multivariate heavy-tailed distributions. *Bernoulli* **28** 1021–1048. MR4388928 https://doi.org/10.3150/21-bej1375

[18] Qi, Y. (2010). On the tail index of a heavy tailed distribution. *Ann. Inst. Statist. Math.* **62** 277–298. MR2592099 https://doi.org/10.1007/s10463-008-0176-2

[19] Stupfler, G. (2019). On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes* **22** 749–769. MR4031856 https://doi.org/10.1007/s10687-019-00351-5

[20] Velthoen, J., Cai, J.-J., Jongbloed, G. and Schmeits, M. (2019). Improving precipitation forecasts using extreme quantile regression. *Extremes* **22** 599–622. MR4031851 https://doi.org/10.1007/s10687-019-00355-1

[21] Wang, H.J., Li, D. and He, X. (2012). Estimation of high conditional quantiles for heavy-tailed distributions. *J. Amer. Statist. Assoc.* **107** 1453–1464. MR3036407 https://doi.org/10.1080/01621459.2012.716382

[22] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73** 812–815. MR0521329

Bernoulli **30**(2), 2024, 1313–1346
https://doi.org/10.3150/23-BEJ1633

Testing with p*-values: Between p-values, mid p-values, and e-values

RUODU WANG[a]

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada,
[a]*wang@uwaterloo.ca*

We introduce the notion of p*-values (p*-variables), which generalizes p-values (p-variables) in several senses. The new notion has four natural interpretations: operational, probabilistic, Bayesian, and frequentist. A main example of a p*-value is a mid p-value, which arises in the presence of discrete test statistics. A unified stochastic representation for p-values, mid p-values, and p*-values is obtained to illustrate the relationship between the three objects. We study several ways of merging arbitrarily dependent or independent p*-values into one p-value or p*-value. Admissible calibrators of p*-values to and from p-values and e-values are obtained with nice mathematical forms, revealing the role of p*-values as a bridge between p-values and e-values. The notion of p*-values becomes useful in many situations even if one is only interested in p-values, mid p-values, or e-values. In particular, deterministic tests based on p*-values can be applied to improve some classic methods for p-values and e-values.

Keywords: Arbitrary dependence; average of p-values; mid p-values; posterior predictive p-values; test martingale

References

[1] Bates, S., Candès, E., Lei, L., Romano, Y. and Sesia, M. (2023). Testing for outliers with conformal *p*-values. *Ann. Statist.* **51** 149–178. MR4564852 https://doi.org/10.1214/22-aos2244

[2] Benjamini, Y. and Hochberg, Y. (1997). Multiple hypotheses testing with weights. *Scand. J. Stat.* **24** 407–418. MR1481424 https://doi.org/10.1111/1467-9469.00072

[3] 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

[4] 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

[5] Döhler, S., Durand, G. and Roquain, E. (2018). New FDR bounds for discrete and heterogeneous tests. *Electron. J. Stat.* **12** 1867–1900. MR3813600 https://doi.org/10.1214/18-EJS1441

[6] Duan, B., Ramdas, A., Balakrishnan, S. and Wasserman, L. (2020). Interactive martingale tests for the global null. *Electron. J. Stat.* **14** 4489–4551. MR4194269 https://doi.org/10.1214/20-EJS1790

[7] Efron, B. (2010). *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction. Institute of Mathematical Statistics (IMS) Monographs* **1**. Cambridge: Cambridge Univ. Press. MR2724758 https://doi.org/10.1017/CBO9780511761362

[8] Genovese, C. and Wasserman, L. (2004). A stochastic process approach to false discovery control. *Ann. Statist.* **32** 1035–1061. MR2065197 https://doi.org/10.1214/009053604000000283

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

[10] Grünwald, P., de Heide, R. and Koolen, W.M. (2020). Safe testing. Available at arXiv:1906.07801v2.

[11] Habiger, J.D. (2015). Multiple test functions and adjusted *p*-values for test statistics with discrete distributions. *J. Statist. Plann. Inference* **167** 1–13. MR3383232 https://doi.org/10.1016/j.jspi.2015.06.003

[12] Howard, S.R., Ramdas, A., McAuliffe, J. and Sekhon, J. (2021). Time-uniform, nonparametric, nonasymptotic confidence sequences. *Ann. Statist.* **49** 1055–1080. MR4255119 https://doi.org/10.1214/20-aos1991

[13] Huber, M. (2019). Halving the bounds for the Markov, Chebyshev, and Chernoff inequalities using smoothing. *Amer. Math. Monthly* **126** 915–927. MR4033226 https://doi.org/10.1080/00029890.2019.1656484

1350-7265

[14] Lancaster, H.O. (1952). Statistical control of counting experiments. *Biometrika* **39** 419–422.

[15] Liu, F. and Wang, R. (2021). A theory for measures of tail risk. *Math. Oper. Res.* **46** 1109–1128. MR4312589 https://doi.org/10.1287/moor.2020.1072

[16] Liu, Y. and Xie, J. (2020). Cauchy combination test: A powerful test with analytic p-value calculation under arbitrary dependency structures. *J. Amer. Statist. Assoc.* **115** 393–402. MR4078471 https://doi.org/10.1080/01621459.2018.1554485

[17] Mao, T., Wang, B. and Wang, R. (2019). Sums of standard uniform random variables. *J. Appl. Probab.* **56** 918–936. MR4015643 https://doi.org/10.1017/jpr.2019.52

[18] Meng, X.-L. (1994). Posterior predictive p-values. *Ann. Statist.* **22** 1142–1160. MR1311969 https://doi.org/10.1214/aos/1176325622

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

[20] Nutz, M., Wang, R. and Zhang, Z. (2022). Martingale transports and Monge maps. Available at arXiv:2209.14432.

[21] Ramdas, A., Grünwald, P., Vovk, V. and Shafer, G. (2022). Game-theoretic statistics and safe anytime-valid inference. Available at arXiv:2210.01948.

[22] 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

[23] Rubin-Delanchy, P., Heard, N.A. and Lawson, D.J. (2019). Meta-analysis of mid-p-values: Some new results based on the convex order. *J. Amer. Statist. Assoc.* **114** 1105–1112. MR4011765 https://doi.org/10.1080/01621459.2018.1469994

[24] Rüschendorf, L. (1982). Random variables with maximum sums. *Adv. in Appl. Probab.* **14** 623–632. MR0665297 https://doi.org/10.2307/1426677

[25] Rüschendorf, L. (2013). *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*. *Springer Series in Operations Research and Financial Engineering*. Heidelberg: Springer. MR3051756 https://doi.org/10.1007/978-3-642-33590-7

[26] 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

[27] Shafer, G. (2021). Testing by betting: A strategy for statistical and scientific communication. *J. Roy. Statist. Soc. Ser. A* **184** 407–431. MR4255905 https://doi.org/10.1111/rssa.12647

[28] 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

[29] 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

[30] 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

[31] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36** 423–439. MR0177430 https://doi.org/10.1214/aoms/1177700153

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

[33] Vovk, V., Gammerman, A. and Shafer, G. (2005). *Algorithmic Learning in a Random World*. New York: Springer. MR2161220

[34] 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

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

[36] 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

[37] Vovk, V. and Wang, R. (2023). Confidence and discoveries with E-values. *Statist. Sci.* **38** 329–354. MR4596762 https://doi.org/10.1214/22-sts874

[38] Wang, R. (2014). Sum of arbitrarily dependent random variables. *Electron. J. Probab.* **19** no. 84, 18. MR3263641 https://doi.org/10.1214/EJP.v19-3373

[39] 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

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

[41] Wilson, D.J. (2019). The harmonic mean p-value for combining dependent tests. *Proc. Natl. Acad. Sci. USA* **116** 1195–1200. MR3904688 https://doi.org/10.1073/pnas.1814092116

Bernoulli **30**(2), 2024, 1347–1374
https://doi.org/10.3150/23-BEJ1634

Sequential testing for elicitable functionals via supermartingales

PHILIPPE CASGRAIN[1,a], MARTIN LARSSON[2,b] and JOHANNA ZIEGEL[3,c]

[1]*ETH Zürich, Zürich, Switzerland,* [a]*philippe.casgrain@math.ethz.ch*
[2]*Carnegie Mellon University, Pittsburgh, USA,* [b]*larsson@cmu.edu*
[3]*University of Bern, Bern, Switzerland,* [c]*johanna.ziegel@unibe.ch*

We design sequential tests for a large class of nonparametric null hypotheses based on elicitable and identifiable functionals. Such functionals are defined in terms of scoring functions and identification functions, which are ideal building blocks for constructing nonnegative supermartingales under the null. This in turn yields sequential tests via Ville's inequality. Using regret bounds from Online Convex Optimization, we obtain rigorous guarantees on the asymptotic power of the tests for a wide range of alternative hypotheses. Our results allow for bounded and unbounded data distributions, assuming that a sub-ψ tail bound is satisfied.

Keywords: Anytime valid testing; elicitable functionals; identifiable functionals; online optimization; sequential statistics

References

Albers, C. (2019). The problem with unadjusted multiple and sequential statistical testing. *Nat. Commun.* **10** 1–4.
Caponnetto, A. (2005). A note on the role of squared loss in regression. CBCL Paper, Massachusetts Institute of Technology, Cambridge, MA.
Choe, Y.J. and Ramdas, A. (2021). Comparing sequential forecasters. Preprint. Available at arXiv:2110.00115.
Cornfeld, I.P., Fomin, S.V. and Sinai, Y.G. (2012). *Ergodic Theory*. New York: Springer.
de la Peña, V.H., Klass, M.J. and Lai, T.L. (2004). Self-normalized processes: Exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.* **32** 1902–1933. MR2073181 https://doi.org/10.1214/009117904000000397
Dimitriadis, T., Fissler, T. and Ziegel, J.F. (2020). The efficiency gap. Preprint. Available at arXiv:2010.14146.
Dimitriadis, T., Fissler, T. and Ziegel, J. (2023). Characterizing M-estimators. *Biometrika* asad026. https://doi.org/10.1093/biomet/asad026
Fissler, T. (2017). On higher order elicitability and some limit theorems on the Poisson and Wiener space. Ph.D. thesis, Univ. Bern.
Fissler, T. and Ziegel, J.F. (2016). Higher order elicitability and Osband's principle. *Ann. Statist.* **44** 1680–1707. MR3519937 https://doi.org/10.1214/16-AOS1439
Freedman, D.A. (1975). On tail probabilities for martingales. *Ann. Probab.* **3** 100–118. MR0380971 https://doi.org/10.1214/aop/1176996452
Frongillo, R. and Kash, I. (2015). Vector-valued property elicitation. In *Proceedings of the 28th Conference on Learning Theory*. *PMLR* **40** 710–727.
Frongillo, R.M. and Kash, I.A. (2021). Elicitation complexity of statistical properties. *Biometrika* **108** 857–879. MR4341356 https://doi.org/10.1093/biomet/asaa093
Frongillo, R. and Nobel, A. (2020). Memoryless sequences for general losses. *J. Mach. Learn. Res.* **21** Paper No. 80, 28. MR4119148
Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. MR2847988 https://doi.org/10.1198/jasa.2011.r10138
Grünwald, P., de Heide, R. and Koolen, W. (2019). Safe testing. Preprint. Available at arXiv:1906.07801.

1350-7265

Hagerup, T. and Rüb, C. (1990). A guided tour of Chernoff bounds. *Inform. Process. Lett.* **33** 305–308. MR1045520 https://doi.org/10.1016/0020-0190(90)90214-I

Hazan, E. (2016). Introduction to online convex optimization. *Found. Trends Optim.* **2** 157–325.

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

Hertz, D. (2020). Improved Hoeffding's lemma and Hoeffding's tail bounds. Preprint. Available at arXiv:2012.03535.

Hoeffding, W. (1994). Probability inequalities for sums of bounded random variables. In *The Collected Works of Wassily Hoeffding* 409–426. New York, NY: Springer.

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

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

Hsu, J.C. (1996). *Multiple Comparisons: Theory and Methods*. London: CRC Press. MR1629127 https://doi.org/10.1007/978-1-4899-7180-7

Jun, K.-S. and Orabona, F. (2019). Parameter-free online convex optimization with sub-exponential noise. In *Conference on Learning Theory* 1802–1823. PMLR.

Komunjer, I. and Vuong, Q. (2010a). Semiparametric efficiency bound in time-series models for conditional quantiles. *Econometric Theory* **26** 383–405. MR2600568 https://doi.org/10.1017/S0266466609100038

Komunjer, I. and Vuong, Q. (2010b). Efficient estimation in dynamic conditional quantile models. *J. Econometrics* **157** 272–285. MR2661600 https://doi.org/10.1016/j.jeconom.2010.01.001

Lambert, N., Pennock, D.M. and Shoham, Y. (2008). Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce* 129–138. Chicago, Il, USA. Extended abstract.

McMahan, H.B. (2017). A survey of algorithms and analysis for adaptive online learning. *J. Mach. Learn. Res.* **18** Paper No. 90, 50. MR3714253

O'Neill, R. and Wetherill, G.B. (1971). The present state of multiple comparison methods. *J. Roy. Statist. Soc. Ser. B* **33** 218–241, 241–250. MR0321252

Orabona, F. and Jun, K.-S. (2021). Tight concentrations and confidence sequences from the regret of universal portfolio. Preprint. Available at arXiv:2110.14099.

Osband, K.H. (1985). Providing incentives for better cost forecasting. Ph.D. thesis, Univ. California, Berkeley.

Ramdas, A., Ruf, J., Larsson, M. and Koolen, W. (2020). Admissible anytime-valid sequential inference must rely on nonnegative martingales. Preprint. Available at arXiv:2009.03167.

Ramdas, A., Ruf, J., Larsson, M. and Koolen, W.M. (2022a). Testing exchangeability: Fork-convexity, supermartingales and e-processes. *Internat. J. Approx. Reason.* **141** 83–109. MR4364897 https://doi.org/10.1016/j.ijar.2021.06.017

Ramdas, A., Grünwald, P., Vovk, V. and Shafer, G. (2022b). Game-theoretic statistics and safe anytime-valid inference. Preprint. Available at arXiv:2210.01948.

Robbins, H. and Siegmund, D. (1972). A class of stopping rules for testing parametric hypotheses. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. IV: Biology and Health* 37–41. Berkeley, Calif.: Univ. California Press. MR0403111

Rockafellar, R.T. and Wets, R.J.-B. (2009). *Variational Analysis*. Berlin, Heidelberg: Springer.

Savage, L.J. (1971). Elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc.* **66** 783–801. MR0331571

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

Shalev-Shwartz, S. (2011). Online learning and online convex optimization. *Found. Trends Mach. Learn.* **4** 107–194.

Shekhar, S. and Ramdas, A. (2021). Nonparametric two-sample testing by betting. Preprint. Available at arXiv:2112.09162.

Steinwart, I., Pasin, C., Williamson, R. and Zhang, S. (2014). Elicitation and identification of properties. In *Conference on Learning Theory* 482–526. PMLR.

Ville, J. (1939). *Étude Critique de la Notion de Collectif*. NUMDAM [place of publication not identified]. MR3533075

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

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Stat.* **16** 117–186. MR0013275 https://doi.org/10.1214/aoms/1177731118

Waudby-Smith, I. and Ramdas, A. (2023). Estimating means of bounded random variables by betting. *J. Roy. Statist. Soc. Ser. B* qkad009. https://doi.org/10.1093/jrsssb/qkad009

Xu, Z., Wang, R. and Ramdas, A. (2021). A unified framework for bandit multiple testing. Preprint. Available at arXiv:2107.07322.

Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning* 928–936.

Bernoulli **30**(2), 2024, 1375–1400
https://doi.org/10.3150/23-BEJ1635

Comparison principle for stochastic heat equations driven by α-stable white noises

YONGJIN WANG[1,2,a], CHENGXIN YAN[1,b] and XIAOWEN ZHOU[3,c]

[1]*School of Mathematical Sciences, Nankai University, Tianjin, China,* [a]*yjwang@nankai.edu.cn,* [b]*cxyan@mail.nankai.edu.cn*
[2]*School of Business, Nankai University, Tianjin, China*
[3]*Department of Mathematics and Statistics, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Canada,* [c]*xiaowen.zhou@concordia.ca*

For a class of non-linear stochastic heat equations driven by α-stable white noises for $\alpha \in (1,2)$ with Lipschitz coefficients, we prove the existence and pathwise uniqueness of L^p-valued càdlàg solution to such an equation for $p \in (\alpha, 2]$ by considering a sequence of approximating stochastic heat equations driven by truncated α-stable white noises obtained by removing the big jumps from the original α-stable white noise. If the α-stable white noise is spectrally one-sided, under additional monotonicity assumption on noise coefficients, we further prove a comparison theorem on the L^2-valued càdlàg solutions to such an equation. As a consequence, the non-negativity of the L^2-valued càdlàg solution is established for the above stochastic heat equation with non-negative initial function.

Keywords: α-stable white noises; comparison principle; non-negative solutions; stochastic heat equations; truncated α-stable white noises

References

Albeverio, S., Wu, J.-L. and Zhang, T.-S. (1998). Parabolic SPDEs driven by Poisson white noise. *Stochastic Process. Appl.* **74** 21–36. MR1624076 https://doi.org/10.1016/S0304-4149(97)00112-9

Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge: Cambridge Univ. Press. MR2512800 https://doi.org/10.1017/CBO9780511809781

Balan, R.M. (2014). SPDEs with α-stable Lévy noise: A random field approach. *Int. J. Stoch. Anal.* 793275. MR3166748 https://doi.org/10.1155/2014/793275

Bao, J. and Yuan, C. (2011). Comparison theorem for stochastic differential delay equations with jumps. *Acta Appl. Math.* **116** 119–132. MR2842984 https://doi.org/10.1007/s10440-011-9633-7

Bo, L. and Wang, Y. (2006). Stochastic Cahn-Hilliard partial differential equations with Lévy spacetime white noises. *Stoch. Dyn.* **6** 229–244. MR2239091 https://doi.org/10.1142/S0219493706001736

Chen, L. and Huang, J. (2019). Comparison principle for stochastic heat equation on $\mathbb{R}^d$. *Ann. Probab.* **47** 989–1035. MR3916940 https://doi.org/10.1214/18-AOP1277

Chen, L. and Kim, K. (2017). On comparison principle and strict positivity of solutions to the nonlinear stochastic fractional heat equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 358–388. MR3606745 https://doi.org/10.1214/15-AIHP719

Chong, C., Dalang, R.C. and Humeau, T. (2019). Path properties of the solution to the stochastic heat equation with Lévy noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **7** 123–168. MR3916265 https://doi.org/10.1007/s40072-018-0124-y

Dareiotis, K.A. and Gyöngy, I. (2014). A comparison principle for stochastic integro-differential equations. *Potential Anal.* **41** 1203–1222. MR3269720 https://doi.org/10.1007/s11118-014-9416-7

Denis, L., Matoussi, A. and Stoica, L. (2009). Maximum principle and comparison theorem for quasi-linear stochastic PDE's. *Electron. J. Probab.* **14** 500–530. MR2480551 https://doi.org/10.1214/EJP.v14-629

1350-7265

Donati-Martin, C. and Pardoux, É. (1993). White noise driven SPDEs with reflection. *Probab. Theory Related Fields* **95** 1–24. MR1207304 https://doi.org/10.1007/BF01197335

Evans, L.C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Providence, RI: Amer. Math. Soc. MR2597943 https://doi.org/10.1090/gsm/019

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. MR0270403

Gyöngy, I. (1982). On stochastic equations with respect to semimartingales. III. *Stochastics* **7** 231–254. MR0674448 https://doi.org/10.1080/17442508208833220

Kotelenez, P. (1992). Comparison methods for a class of function valued stochastic partial differential equations. *Probab. Theory Related Fields* **93** 1–19. MR1172936 https://doi.org/10.1007/BF01195385

Lin, S. (2013). Generalized Gronwall inequalities and their applications to fractional differential equations. *J. Inequal. Appl.* **2013** 549. MR3212979 https://doi.org/10.1186/1029-242X-2013-549

Mandrekar, V. and Rüdiger, B. (2015). *Stochastic Integration in Banach Spaces – Theory and Applications*. *Probability Theory and Stochastic Modelling* **73**. Cham: Springer. MR3243582 https://doi.org/10.1007/978-3-319-12853-5

Moreno Flores, G.R. (2014). On the (strict) positivity of solutions of the stochastic heat equation. *Ann. Probab.* **42** 1635–1643. MR3262487 https://doi.org/10.1214/14-AOP911

Mytnik, L. (2002). Stochastic partial differential equation driven by stable noise. *Probab. Theory Related Fields* **123** 157–201. MR1900321 https://doi.org/10.1007/s004400100180

Mytnik, L. and Perkins, E. (2003). Regularity and irregularity of $(1+\beta)$-stable super-Brownian motion. *Ann. Probab.* **31** 1413–1440. MR1989438 https://doi.org/10.1214/aop/1055425785

Niu, M. and Xie, B. (2019). Comparison theorem and correlation for stochastic heat equations driven by Lévy space-time white noises. *Discrete Contin. Dyn. Syst. Ser. B* **24** 2989–3009. MR3986189 https://doi.org/10.3934/dcdsb.2018296

Peng, S. and Zhu, X. (2006). Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations. *Stochastic Process. Appl.* **116** 370–380. MR2199554 https://doi.org/10.1016/j.spa.2005.08.004

Peszat, S. and Zabczyk, J. (2006). Stochastic heat and wave equations driven by an impulsive noise. In *Stochastic Partial Differential Equations and Applications—VII. Lect. Notes Pure Appl. Math.* **245** 229–242. Boca Raton, FL: CRC Press/CRC. MR2227232 https://doi.org/10.1201/9781420028720.ch19

Peszat, S. and Zabczyk, J. (2007). *Stochastic Partial Differential Equations with Lévy Noise*. *Encyclopedia of Mathematics and Its Applications* **113**. Cambridge: Cambridge Univ. Press. MR2356959 https://doi.org/10.1017/CBO9780511721373

Truman, A. and Wu, J.-L. (2003). Stochastic Burgers equation with Lévy space-time white noise. In *Probabilistic Methods in Fluids* 298–323. River Edge, NJ: World Sci. Publ. MR2083380 https://doi.org/10.1142/9789812703989_0020

Walsh, J.B. (1986). An introduction to stochastic partial differential equations. In *École D'été de Probabilités de Saint-Flour, XIV—1984*. *Lecture Notes in Math.* **1180** 265–439. Berlin: Springer. MR0876085 https://doi.org/10.1007/BFb0074920

Wang, Y., Yan, C. and Zhou, X. (2022). Existence of weak solutions to stochastic heat equations driven by truncated α-stable white noises with non-Lipschitz coefficients. arXiv:2208.00820v1.

Xiong, J. and Yang, X. (2019). Existence and pathwise uniqueness to an SPDE driven by α-stable colored noise. *Stochastic Process. Appl.* **129** 2681–2722. MR3980141 https://doi.org/10.1016/j.spa.2018.08.003

Xiong, J. and Yang, X. (2023). SPDEs with non-Lipschitz coefficients and nonhomogeous boundary conditions. Forthcoming on Bernoulli.

Yang, X. and Zhou, X. (2017). Pathwise uniqueness for an SPDE with Hölder continuous coefficient driven by α-stable noise. *Electron. J. Probab.* **22** 4. MR3613697 https://doi.org/10.1214/16-EJP23

Bernoulli **30**(2), 2024, 1401–1415
https://doi.org/10.3150/23-BEJ1636

Berry-Esseen bound and Cramér moderate deviation expansion for a supercritical branching random walk

THI THUY BUI[a], ION GRAMA[b] and QUANSHENG LIU[c]

Univ Bretagne Sud, CNRS UMR 6205, LMBA, F-56000 Vannes, France, [a]*thuybui041191@gmail.com,* [b]*ion.grama@univ-ubs.fr,* [c]*quansheng.liu@univ-ubs.fr*

We consider a supercritical branching random walk where each particle gives birth to a random number of particles of the next generation, which move on the real line, according to a fixed law. Let Z_n be the counting measure which counts the number of particles of n-th generation situated in a given region. Under suitable conditions, we establish a Berry-Esseen bound and a Cramér type moderate deviation expansion for Z_n with suitable norming.

Keywords: Branching random walk; central limit theorem; Berry-Esseen bound; large and moderate deviations; branching processes; random walks

References

[1] Aidekon, E. and Shi, Z. (2014). The Seneta-Heyde scaling for the branching random walk. *Ann. Probab.* **42** 959–993. MR3189063 https://doi.org/10.1214/12-AOP809

[2] Asmussen, S. and Kaplan, N. (1976). Branching random walks. I. *Stochastic Process. Appl.* **4** 1–13. MR0400429 https://doi.org/10.1016/0304-4149(76)90022-3

[3] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. *Die Grundlehren der Mathematischen Wissenschaften, Band 196*. New York: Springer. MR0373040

[4] Barral, J., Hu, Y. and Madaule, T. (2018). The minimum of a branching random walk outside the boundary case. *Bernoulli* **24** 801–841. MR3706777 https://doi.org/10.3150/15-BEJ784

[5] Barral, J. and Jin, X. (2014). On exact scaling log-infinitely divisible cascades. *Probab. Theory Related Fields* **160** 521–565. MR3278915 https://doi.org/10.1007/s00440-013-0534-8

[6] Biggins, J.D. (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14** 25–37. MR0433619 https://doi.org/10.2307/3213258

[7] Biggins, J.D. (1977). Chernoff's theorem in the branching random walk. *J. Appl. Probab.* **14** 630–636. MR0464415 https://doi.org/10.1017/s0021900200025900

[8] Biggins, J.D. (1979). Growth rates in the branching random walk. *Z. Wahrsch. Verw. Gebiete* **48** 17–34. MR0533003 https://doi.org/10.1007/BF00534879

[9] Biggins, J.D. (1990). The central limit theorem for the supercritical branching random walk, and related results. *Stochastic Process. Appl.* **34** 255–274. MR1047646 https://doi.org/10.1016/0304-4149(90)90018-N

[10] Biggins, J.D. (1991). Uniform convergence of martingales in the one-dimensional branching random walk. In *Selected Proceedings of the Sheffield Symposium on Applied Probability (Sheffield, 1989). Institute of Mathematical Statistics Lecture Notes—Monograph Series* **18** 159–173. Hayward, CA: IMS. MR1193068 https://doi.org/10.1214/lnms/1215459294

[11] Biggins, J.D. (1992). Uniform convergence of martingales in the branching random walk. *Ann. Probab.* **20** 137–151. MR1143415

[12] Biggins, J.D. and Kyprianou, A.E. (1997). Seneta-Heyde norming in the branching random walk. *Ann. Probab.* **25** 337–360. MR1428512 https://doi.org/10.1214/aop/1024404291

[13] Buraczewski, D., Damek, E., Guivarc'h, Y. and Mentemeier, S. (2014). On multidimensional Mandelbrot cascades. *J. Difference Equ. Appl.* **20** 1523–1567. MR3268907 https://doi.org/10.1080/10236198.2014.950259

[14] Buraczewski, D., Damek, E. and Mikosch, T. (2016). *Stochastic Models with Power-Law Tails: The Equation $X = AX + B$. Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3497380 https://doi.org/10.1007/978-3-319-29679-1

[15] Buraczewski, D. and Maślanka, M. (2019). Large deviation estimates for branching random walks. *ESAIM Probab. Stat.* **23** 823–840. MR4045541 https://doi.org/10.1051/ps/2019006

[16] Chen, X. (2001). Exact convergence rates for the distribution of particles in branching random walks. *Ann. Appl. Probab.* **11** 1242–1262. MR1878297 https://doi.org/10.1214/aoap/1015345402

[17] Chen, X. and He, H. (2019). On large deviation probabilities for empirical distribution of supercritical branching random walks with unbounded displacements. *Probab. Theory Related Fields* **175** 255–307. MR4009709 https://doi.org/10.1007/s00440-018-0891-4

[18] Gao, Z. and Liu, Q. (2016). Exact convergence rates in central limit theorems for a branching random walk with a random environment in time. *Stochastic Process. Appl.* **126** 2634–2664. MR3522296 https://doi.org/10.1016/j.spa.2016.02.013

[19] Gao, Z. and Liu, Q. (2018). Second and third orders asymptotic expansions for the distribution of particles in a branching random walk with a random environment in time. *Bernoulli* **24** 772–800. MR3706776 https://doi.org/10.3150/16-BEJ895

[20] Grübel, R. and Kabluchko, Z. (2017). Edgeworth expansions for profiles of lattice branching random walks. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 2103–2134. MR3729649 https://doi.org/10.1214/16-AIHP785

[21] Harris, T.E. (1963). *The Theory of Branching Processes*. *Die Grundlehren der Mathematischen Wissenschaften, Band 119*. Berlin: Springer. MR0163361

[22] Hu, Y. (2016). How big is the minimum of a branching random walk? *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 233–260. MR3449302 https://doi.org/10.1214/14-AIHP651

[23] Hu, Y. and Shi, Z. (2009). Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* **37** 742–789. MR2510023 https://doi.org/10.1214/08-AOP419

[24] Iksanov, A. (2016). *Renewal Theory for Perturbed Random Walks and Similar Processes*. *Probability and Its Applications*. Cham: Birkhäuser/Springer. MR3585464 https://doi.org/10.1007/978-3-319-49113-4

[25] Iksanov, A., Kolesko, K. and Meiners, M. (2019). Stable-like fluctuations of Biggins' martingales. *Stochastic Process. Appl.* **129** 4480–4499. MR4013869 https://doi.org/10.1016/j.spa.2018.11.022

[26] Kahane, J.-P. and Peyrière, J. (1976). Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.* **22** 131–145. MR0431355 https://doi.org/10.1016/0001-8708(76)90151-1

[27] Kaplan, N. and Asmussen, S. (1976). Branching random walks. II. *Stochastic Process. Appl.* **4** 15–31. MR0400430 https://doi.org/10.1016/0304-4149(76)90023-5

[28] Klebaner, C.F. (1982). Branching random walk in varying environments. *Adv. in Appl. Probab.* **14** 359–367. MR0650128 https://doi.org/10.2307/1426526

[29] Liang, X. and Liu, Q. (2020). Regular variation of fixed points of the smoothing transform. *Stochastic Process. Appl.* **130** 4104–4140. MR4102261 https://doi.org/10.1016/j.spa.2019.11.011

[30] Liu, Q. (1997). Sur une équation fonctionnelle et ses applications: Une extension du théorème de Kesten-Stigum concernant des processus de branchement. *Adv. in Appl. Probab.* **29** 353–373. MR1450934 https://doi.org/10.2307/1428007

[31] Liu, Q. (2000). On generalized multiplicative cascades. *Stochastic Process. Appl.* **86** 263–286. MR1741808 https://doi.org/10.1016/S0304-4149(99)00097-6

[32] Lyons, R. (1997). A simple path to Biggins' martingale convergence for branching random walk. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*. *IMA Vol. Math. Appl.* **84** 217–221. New York: Springer. MR1601749 https://doi.org/10.1007/978-1-4612-1862-3_17

[33] Mentemeier, S. (2016). The fixed points of the multivariate smoothing transform. *Probab. Theory Related Fields* **164** 401–458. MR3449394 https://doi.org/10.1007/s00440-015-0615-y

[34] Petrov, V.V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete [Results in Mathematics and Related Areas], Band 82*. New York: Springer. Translated from the Russian by A.A. Brown. MR0388499

[35] Révész, P. (1994). *Random Walks of Infinitely Many Particles*. River Edge, NJ: World Scientific Co., Inc. MR1645302 https://doi.org/10.1142/2376

[36] Shi, Z. (2015). *Branching Random Walks*. *Lecture Notes in Math.* **2151**. Cham: Springer. Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR3444654 https://doi.org/10.1007/978-3-319-25372-5

[37] Stam, A.J. (1966). On a conjecture by Harris. *Z. Wahrsch. Verw. Gebiete* **5** 202–206. MR0202201 https://doi.org/10.1007/BF00533055

Bernoulli **30**(2), 2024, 1416–1440
https://doi.org/10.3150/23-BEJ1638

Inadmissibility of the corrected Akaike information criterion

TAKERU MATSUDA[1,2,a]

[1]*Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan,* [a]*matsuda@mist.i.u-tokyo.ac.jp*
[2]*Statistical Mathematics Unit, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan*

For the multivariate linear regression model with unknown covariance, the corrected Akaike information criterion is the minimum variance unbiased estimator of the expected Kullback–Leibler discrepancy. In this study, based on the loss estimation framework, we show its inadmissibility as an estimator of the Kullback–Leibler discrepancy itself, instead of the expected Kullback–Leibler discrepancy. We provide improved estimators of the Kullback–Leibler discrepancy that work well in reduced-rank situations and examine their performance numerically.

Keywords: Admissibility; Akaike information criterion; corrected Akaike information criterion; Kullback–Leibler discrepancy; loss estimation

References

Aitchison, J. (1975). Goodness of prediction fit. *Biometrika* **62** 547–554. MR0391353 https://doi.org/10.1093/biomet/62.3.547

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory (Tsahkadsor, 1971)* 267–281. Budapest: Akadémiai Kiadó. MR0483125

Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley. MR0091588

Bedrick, E.J. and Tsai, C.-L. (1994). Model selection for multivariate regression in small samples. *Biometrics* **50** 226–231.

Bellec, P.C. and Zhang, C.-H. (2021). Second-order Stein: SURE for SURE and other applications in high-dimensional inference. *Ann. Statist.* **49** 1864–1903. MR4319234 https://doi.org/10.1214/20-aos2005

Boisbunon, A., Canu, S., Fourdrinier, D., Strawderman, W. and Wells, M.T. (2014). Akaike's information criterion, C_p and estimators of loss for elliptically symmetric distributions. *Int. Stat. Rev.* **82** 422–439. MR3280984 https://doi.org/10.1111/insr.12052

Burnham, K.P. and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. New York: Springer. MR1919620

Cavanaugh, J.E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statist. Probab. Lett.* **33** 201–208. MR1458291 https://doi.org/10.1016/S0167-7152(96)00128-9

Davies, S.L., Neath, A.A. and Cavanaugh, J.E. (2006). Estimation optimality of corrected AIC and modified Cp in linear regression. *Int. Stat. Rev.* **74** 161–168.

Efron, B. and Morris, C. (1972). Empirical Bayes on vector observations: An extension of Stein's method. *Biometrika* **59** 335–347. MR0334386 https://doi.org/10.1093/biomet/59.2.335

Efron, B. and Morris, C. (1976). Multivariate empirical Bayes and estimation of covariance matrices. *Ann. Statist.* **4** 22–32. MR0394960

Fourdrinier, D. and Strawderman, W.E. (2003). On Bayes and unbiased estimators of loss. *Ann. Inst. Statist. Math.* **55** 803–816. MR2028618 https://doi.org/10.1007/BF02523394

Fourdrinier, D., Strawderman, W.E. and Wells, M.T. (2018). *Shrinkage Estimation*. *Springer Series in Statistics*. Cham: Springer. MR3887633 https://doi.org/10.1007/978-3-030-02185-6

Fourdrinier, D. and Wells, M.T. (2012). On improved loss estimation for shrinkage estimators. *Statist. Sci.* **27** 61–81. MR2953496 https://doi.org/10.1214/11-STS380

Fujikoshi, Y., Sakurai, T. and Yanagihara, H. (2014). Consistency of high-dimensional AIC-type and C_p-type criteria in multivariate linear regression. *J. Multivariate Anal.* **123** 184–200. MR3130429 https://doi.org/10.1016/j.jmva.2013.09.006

Fujikoshi, Y. and Satoh, K. (1997). Modified AIC and C_p in multivariate linear regression. *Biometrika* **84** 707–716. MR1603952 https://doi.org/10.1093/biomet/84.3.707

Gupta, A.K. and Nagar, D.K. (2000). *Matrix Variate Distributions*. *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics* **104**. Boca Raton, FL: CRC Press/CRC. MR1738933

Hansen, B.E. (2007). Least squares model averaging. *Econometrica* **75** 1175–1189. MR2333497 https://doi.org/10.1111/j.1468-0262.2007.00785.x

Hurvich, C.M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika* **76** 297–307. MR1016020 https://doi.org/10.1093/biomet/76.2.297

Johnstone, I. (1988). On inadmissibility of some unbiased estimates of loss. In *Statistical Decision Theory and Related Topics, IV, Vol. 1 (West Lafayette, Ind., 1986)* 361–379. New York: Springer. MR0927112

Kitagawa, G. (1997). Information criteria for the predictive evaluation of Bayesian models. *Comm. Statist. Theory Methods* **26** 2223–2246. MR1484250 https://doi.org/10.1080/03610929708832043

Kobayashi, K. and Komaki, F. (2008). Bayesian shrinkage prediction for the regression problem. *J. Multivariate Anal.* **99** 1888–1905. MR2466542 https://doi.org/10.1016/j.jmva.2008.01.014

Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*. *Springer Series in Statistics*. New York: Springer. MR2367855 https://doi.org/10.1007/978-0-387-71887-3

Lehmann, E.L. and Casella, G. (2006). *Theory of Point Estimation*. New York: Springer. MR1639875

Lu, K.L. and Berger, J.O. (1989). Estimation of normal means: Frequentist estimation of loss. *Ann. Statist.* **17** 890–906. MR0994274 https://doi.org/10.1214/aos/1176347149

Matsuda, T. and Strawderman, W.E. (2016). Pitman closeness properties of Bayes shrinkage procedures in estimation and prediction. *Statist. Probab. Lett.* **119** 21–29. MR3555265 https://doi.org/10.1016/j.spl.2016.07.005

Matsuda, T. and Strawderman, W.E. (2019). Improved loss estimation for a normal mean matrix. *J. Multivariate Anal.* **169** 300–311. MR3875601 https://doi.org/10.1016/j.jmva.2018.10.001

Matsuda, T. and Strawderman, W.E. (2022). Estimation under matrix quadratic loss and matrix superharmonicity. *Biometrika* **109** 503–519. MR4430971 https://doi.org/10.1093/biomet/asab025

Narayanan, R. and Wells, M.T. (2015). Improved loss estimation for the Lasso: A variable selection tool. *Sankhya B* **77** 45–74. MR3337562 https://doi.org/10.1007/s13571-015-0095-1

Reinsel, G.C. and Velu, R.P. (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

Reschenhofer, E. (1999). Improved estimation of the expected Kullback-Leibler discrepancy in case of misspecification. *Econometric Theory* **15** 377–387. MR1704229 https://doi.org/10.1017/S0266466699153052

Rosset, S. and Tibshirani, R.J. (2020). From Fixed-X to Random-X regression: Bias-variance decompositions, covariance penalties, and prediction error estimation. *J. Amer. Statist. Assoc.* **115** 138–151. MR4078450 https://doi.org/10.1080/01621459.2018.1424632

Sandved, E. (1968). Ancillary statistics and prediction of the loss in estimation problems. *Ann. Math. Stat.* **39** 1756–1758. MR0231481 https://doi.org/10.1214/aoms/1177698162

Stein, C. (1974). Estimation of the mean of a multivariate normal distribution. In *Proceedings of the Prague Symposium on Asymptotic Statistics (Charles Univ., Prague, 1973), Vol. II* 345–381. Prague: Charles Univ. MR0381062

Styan, G.P.H. (1989). Three useful expressions for expectations involving a Wishart matrix and its inverse. In *Statistical Data Analysis and Inference (Neuchâtel, 1989)* 283–296. Amsterdam: North-Holland. MR1089643 https://doi.org/10.1016/B978-0-444-88029-1.50032-0

Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Comm. Statist. Theory Methods* **7** 13–26.

Wan, A.T.K., Zhang, X. and Zou, G. (2010). Least squares model averaging by Mallows criterion. *J. Econometrics* **156** 277–283. MR2609932 https://doi.org/10.1016/j.jeconom.2009.10.030

Wan, A.T.K. and Zou, G. (2004). On unbiased and improved loss estimation for the mean of a multivariate normal distribution with unknown variance. *J. Statist. Plann. Inference* **119** 17–22. MR2018447 https://doi.org/10.1016/S0378-3758(02)00406-8

Yanagihara, H., Wakaki, H. and Fujikoshi, Y. (2015). A consistency property of the AIC for multivariate linear models when the dimension and the sample size are large. *Electron. J. Stat.* **9** 869–897. MR3338666 https://doi.org/10.1214/15-EJS1022

Bernoulli **30**(2), 2024, 1441–1457
https://doi.org/10.3150/23-BEJ1639

Characteristic kernels on Hilbert spaces, Banach spaces, and on sets of measures

JOHANNA ZIEGEL[a], DAVID GINSBOURGER[b] and LUTZ DÜMBGEN[c]

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland,
[a]*johanna.ziegel@stat.unibe.ch,* [b]*david.ginsbourger@stat.unibe.ch,* [c]*lutz.duembgen@stat.unibe.ch*

We present new classes of positive definite kernels on non-standard spaces that are integrally strictly positive definite or characteristic. In particular, we discuss radial kernels on separable Hilbert spaces, and introduce broad classes of kernels on Banach spaces and on metric spaces of strong negative type. The general results are used to give explicit classes of kernels on separable L^p spaces and on sets of measures.

Keywords: Characteristic kernel; integrally strictly positive definite kernel; kernel on Hilbert space; kernel on Banach space; kernel on measures

References

Bachoc, F., Suvorikova, A., Ginsbourger, D., Loubes, J.-M. and Spokoiny, V. (2020). Gaussian processes with multidimensional distribution inputs via optimal transport and Hilbertian embedding. *Electron. J. Stat.* **14** 2742–2772. MR4125856 https://doi.org/10.1214/20-EJS1725

Banerjee, B. (2023). Testing distributional equality for functional random variables. Preprint. Available at arXiv:2303.10973.

Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *J. Multivariate Anal.* **88** 190–206. MR2021870 https://doi.org/10.1016/S0047-259X(03)00079-4

Bauer, H. (2001). *Measure and Integration Theory*. *De Gruyter Studies in Mathematics* **26**. Berlin: de Gruyter. MR1897176 https://doi.org/10.1515/9783110866209

Baxendale, P. (1976). Gaussian measures on function spaces. *Amer. J. Math.* **98** 891–952. MR0467809 https://doi.org/10.2307/2374035

Benton, G.W., Maddox, W.J., Salkey, J.P., Albinati, J. and Wilson, A.G. (2019). Function-space distributions over kernels. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems* **1340**.

Berg, C., Christensen, J.P.R. and Ressel, P. (1984). *Harmonic Analysis on Semigroups*. *Graduate Texts in Mathematics* **100**. New York: Springer. MR0747302 https://doi.org/10.1007/978-1-4612-1128-0

Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Boston, MA: Kluwer Academic. MR2239907 https://doi.org/10.1007/978-1-4419-9096-9

Betken, A. and Dehling, H. (2021). Distance correlation for long-range dependent time series. Preprint. Available at arXiv:2107.03041.

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

Brehmer, J.R., Gneiting, T., Schlather, M. and Strokorb, K. (2021). Using scoring functions to evaluate point process forecasts. Preprint. Available at arXiv:2103.11884.

Buathong, P., Ginsbourger, D. and Krityakierne, F. (2020). Kernels over sets of finite sets using RKHS embeddings, with application to Bayesian (combinatorial) optimization. In *Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS)*. *Proceedings of Machine Learning Research* **108** 2731–2741.

Christensen, J.P.R. (1978). The small ball theorem for Hilbert spaces. *Math. Ann.* **237** 273–276. MR0508757 https://doi.org/10.1007/BF01420181

Christmann, A. and Steinwart, I. (2010). Universal kernels on non-standard input spaces. In *Advances in Neural Information Processing Systems* **23**.

Dawid, A.P. (2007). The geometry of proper scoring rules. *Ann. Inst. Statist. Math.* **59** 77–93. MR2396033 https://doi.org/10.1007/s10463-006-0099-8

Diestel, J. and Uhl, J.J. Jr. (1977). *Vector Measures*. *Mathematical Surveys, No. 15*. Providence, RI: Amer. Math. Soc. MR0453964

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

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

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

Hamid, S., Schulze, S., Osborne, M.A. and Roberts, S.J. (2021). Marginalising over stationary kernels with Bayesian quadrature. Preprint. Available at arXiv:2106.07452.

Hayati, S., Fukumizu, K. and Parvardeh, A. (2020). Kernel mean embeddings of probability measures and its applications to functional data analysis. Preprint. Available at arXiv:2011.02315.

Heinrich, C., Schneider, M., Guttorp, P. and Thorarinsdottir, T. (2019). Validation of point process forecasts. Preprint. Available at https://www.nr.no/directdownload/1564572954/PointProcessValidation-Heinrich.pdf.

Hofmann, T., Schölkopf, B. and Smola, A.J. (2008). Kernel methods in machine learning. *Ann. Statist.* **36** 1171–1220. MR2418654 https://doi.org/10.1214/009053607000000677

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

Kom Samo, Y.L. and Roberts, S. (2015). Generalized spectral kernels. Preprint. Available at arXiv:1506.02236.

Linde, W. (1986). Uniqueness theorems for measures in L_r and $C_0(\Omega)$. *Math. Ann.* **274** 617–626. MR0848507 https://doi.org/10.1007/BF01458597

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

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

Matsui, M., Mikosch, T. and Samorodnitsky, G. (2017). Distance covariance for stochastic processes. *Probab. Math. Statist.* **37** 355–372. MR3745391 https://doi.org/10.19195/0208-4147.37.2.9

Mérigot, Q., Delalande, A. and Chazal, F. (2020). Quantitative stability of optimal transport maps and linearization of the 2-Wasserstein space. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. *Proceedings of Machine Learning Research* **108** 3186–3196.

Muandet, K., Fukumizu, K., Dinuzzo, F. and Schölkopf, B. (2012). Learning from distributions via support vector machines. In *Advances in Neural Information Processing Systems* 10–18.

Pan, W., Tian, Y., Wang, X. and Zhang, H. (2018). Ball divergence: Nonparametric two sample test. *Ann. Statist.* **46** 1109–1137. MR3797998 https://doi.org/10.1214/17-AOS1579

Panaretos, V.M. and Zemel, Y. (2020). *An Invitation to Statistics in Wasserstein Space*. *SpringerBriefs in Probability and Mathematical Statistics*. Cham: Springer. MR4350694 https://doi.org/10.1007/978-3-030-38438-8

Purves, R. (1966). Bimeasurable functions. *Fund. Math.* **58** 149–157. MR0199339 https://doi.org/10.4064/fm-58-2-149-157

Rudin, W. (1970). *Real and Complex Analysis*. London: McGraw-Hill.

Schoenberg, I.J. (1938). Metric spaces and completely monotone functions. *Ann. of Math. (2)* **39** 811–841. MR1503439 https://doi.org/10.2307/1968466

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

Simon-Gabriel, C.-J. and Schölkopf, B. (2018). Kernel distribution embeddings: Universal kernels, characteristic kernels and kernel metrics on distributions. *J. Mach. Learn. Res.* **19** 44. MR3874152

Simon-Gabriel, C.J., Barp, A., Schölkopf, B. and Mackey, L. (2020). Metrizing weak convergence with maximum mean discrepancies. Preprint. Available at arXiv:2006.09268.

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, M.L. (1999). *Interpolation of Spatial Data, Some Theory for Kriging*. *Springer Series in Statistics*. New York: Springer. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

Steinwart, I. (2001). On the influence of the kernel on the consistency of support vector machines. *J. Mach. Learn. Res.* **2** 67–93.

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

Steinwart, I. and Scovel, C. (2012). Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr. Approx.* **35** 363–417. MR2914365 https://doi.org/10.1007/s00365-012-9153-3

Steinwart, I. and Ziegel, J.F. (2021). Strictly proper kernel scores and characteristic kernels on compact spaces. *Appl. Comput. Harmon. Anal.* **51** 510–542. MR4196451 https://doi.org/10.1016/j.acha.2019.11.005

Stewart, J. (1976). Positive definite functions and generalizations, an historical survey. *Rocky Mountain J. Math.* **6** 409–434. MR0430674 https://doi.org/10.1216/RMJ-1976-6-3-409

Sutherland, D. (2016). Scalable, Flexible and Active Learning on Distributions Ph.D. thesis School of Computer Science, Carnegie Mellon Univ.

Szabó, Z., Sriperumbudur, B.K., Póczos, B. and Gretton, A. (2016). Learning theory of distribution regression. *J. Mach. Learn. Res.* **17** 152. MR3555043

Székely, G.J. and Rizzo, M. (2004). Testing for equal distribution in high dimension. *InterStat* **5**.

Székely, G.J. and Rizzo, M.L. (2005). A new test for multivariate normality. *J. Multivariate Anal.* **93** 58–80. MR2119764 https://doi.org/10.1016/j.jmva.2003.12.002

Székely, G.J. and Rizzo, M.L. (2009). Brownian distance covariance. *Ann. Appl. Stat.* **3** 1236–1265. MR2752127 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

Wendland, H. (2005). *Scattered Data Approximation*. *Cambridge Monographs on Applied and Computational Mathematics* **17**. Cambridge: Cambridge Univ. Press. MR2131724

Werner, D. (2002). *Funktionalanalysis*, 3rd ed. Berlin: Springer.

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

Bernoulli **30**(2), 2024, 1458–1481
https://doi.org/10.3150/23-BEJ1640

Empirical likelihood ratio tests for non-nested model selection based on predictive losses

JIANCHENG JIANG[1,a], XUEJUN JIANG[2,b] and HAOFENG WANG[2,c]

[1]*Department of Mathematics and Statistics & School of Data Science, University of North Carolina at Charlotte, NC, 28223, USA,* [a]*jjiang1@uncc.edu*
[2]*Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, 518055, China,* [b]*jiangxj@sustech.edu.cn,* [c]*11749262@mail.sustech.edu.cn*

We propose an empirical likelihood ratio (ELR) test for comparing any two supervised learning models, which may be nested, non-nested, overlapping, misspecified, or correctly specified. The test compares the prediction losses of models based on the cross-validation. We determine the asymptotic null and alternative distributions of the ELR test for comparing two nonparametric learning models under a general framework of convex loss functions. However, the prediction losses from the cross-validation involve repeatedly fitting the models with one observation left out, which leads to a heavy computational burden. We introduce an easy-to-implement ELR test which requires fitting the models only once and shares the same asymptotics as the original one. The proposed tests are applied to compare additive models with varying-coefficient models. Furthermore, a scalable distributed ELR test is proposed for testing the importance of a group of variables in possibly misspecified additive models with massive data. Simulations show that the proposed tests work well and have favorable finite-sample performance compared to some existing approaches. The methodology is validated on an empirical application.

Keywords: Cross-validation; nonparametric smoothing; scalable distributed test

References

Battey, H., Fan, J., Liu, H., Lu, J. and Zhu, Z. (2018). Distributed estimation and inference with statistical guarantees. *Ann. Statist.* **46** 1352–1382.

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

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3** 1–122.

Cai, Z., Fan, J. and Yao, Q. (2000). Functional-coefficient regression models for nonlinear time series. *J. Amer. Statist. Assoc.* **95** 941–956. MR1804449 https://doi.org/10.2307/2669476

Chang, J., Tang, C.Y. and Wu, Y. (2016). Local independence feature screening for nonparametric and semiparametric models by marginal empirical likelihood. *Ann. Statist.* **44** 515–539. MR3476608 https://doi.org/10.1214/15-AOS1374

Chang, J., Chen, S.X., Tang, C.Y. and Wu, T.T. (2021). High-dimensional empirical likelihood inference. *Biometrika* **108** 127–147. MR4226194 https://doi.org/10.1093/biomet/asaa051

Chen, X., Hong, H. and Shum, M. (2007). Nonparametric likihood ratio model selection tests between parametric likelihood and moment condition models. *J. Econometrics* **141** 109–140. MR2411739 https://doi.org/10.1016/j.jeconom.2007.01.010

Chen, X., Liu, W. and Zhang, Y. (2019). Quantile regression under memory constraint. *Ann. Statist.* **47** 3244–3273. MR4025741 https://doi.org/10.1214/18-AOS1777

Chen, S.X. and Peng, L. (2021). Distributed statistical inference for massive data. *Ann. Statist.* **49** 2851–2869. MR4338895 https://doi.org/10.1214/21-aos2062

1350-7265

Chen, S.X. and Van Keilegom, I. (2009). A review on empirical likelihood methods for regression. *TEST* **18** 415–447. MR2566404 https://doi.org/10.1007/s11749-009-0159-5

Chen, X. and Xie, M. (2014). A split-and-conquer approach for analysis of extraordinarily large data. *Statist. Sinica* **24** 1655–1684. MR3308656

Cox, D.R. (1961). Tests of separate families of hypotheses. *Berkeley Symp. Math. Stat. Prob.* **4** 105–123.

Cox, D.R. (1962). Further results on tests of separate families of hypotheses. *J. Roy. Statist. Soc. Ser. B* **24** 406–424. MR0156409

Dave et al. (2004). Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. *N. Engl. J. Med.* **351** 2159–2169.

Fan, J. and Huang, T. (2005). Profile likelihood inferences on semiparametric varying-coefficient partially linear models. *Bernoulli* **11** 1031–1057. MR2189080 https://doi.org/10.3150/bj/1137421639

Fan, J. and Jiang, J. (2005). Nonparametric inferences for additive models. *J. Amer. Statist. Assoc.* **100** 890–907. MR2201017 https://doi.org/10.1198/016214504000001439

Fan, J. and Jiang, J. (2007). Nonparametric inference with generalized likelihood ratio tests (with discussion). *TEST* **16** 409–444.

Fan, J. and Zhang, W. (1999). Statistical estimation in varying coefficient models. *Ann. Statist.* **27** 1491–1518. MR1742497 https://doi.org/10.1214/aos/1017939139

Fan, J. and Zhang, J. (2004). Sieve empirical likelihood ratio tests for nonparametric functions. *Ann. Statist.* **32** 1858–1907. MR2102496 https://doi.org/10.1214/009053604000000210

Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.* **29** 153–193. MR1833962 https://doi.org/10.1214/aos/996986505

Freund, Y. and Schapire, R.E. (1997). A decision-theoretic generalization of on- line learning and an application to boosting. *J. Comput. System Sci.* **55** 119–139.

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Statist.* **29** 1189–1232. MR1873328 https://doi.org/10.1214/aos/1013203451

Hall, P. (1990). Pseudo-likelihood theory for empirical likelihood. *Ann. Statist.* **18** 121–140. MR1041388 https://doi.org/10.1214/aos/1176347495

Hall, P. and La Scala, B. (1990). Methodology and algorithms of empirical likelihood. *Int. Stat. Rev.* **58** 109–127.

Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. *Monographs on Statistics and Applied Probability* **43**. London: CRC Press. MR1082147

Hastie, T.J. and Tibshirani, R.J. (1993). Varying-coefficient models (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **55** 757–796.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second ed. New York: Springer.

He, X. and Shi, P. (1996). Bivariate tensor-product B-splines in a partly linear model. *J. Multivariate Anal.* **58** 162–181. MR1405586 https://doi.org/10.1006/jmva.1996.0045

Huang, J.Z. and Shen, H. (2004). Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scand. J. Stat.* **31** 515–534. MR2101537 https://doi.org/10.1111/j.1467-9469.2004.00404.x

Jiang, J., Jiang, X. and Wang, H. (2024). Supplement to "Empirical likelihood ratio tests for non-nested model selection based on predictive losses." https://doi.org/10.3150/23-BEJ1640SUPP

Jiang, J. and Li, J. (2008). Two-stage local M-estimation of additive models. *Sci. China Ser. A* **51** 1315–1338. MR2417497 https://doi.org/10.1007/s11425-007-0173-6

Jiang, J., Zhou, H., Jiang, X. and Peng, J. (2007). Generalized likelihood ratio tests for the structure of semiparametric additive models. *Canad. J. Statist.* **35** 381–398. MR2396026 https://doi.org/10.1002/cjs.5550350304

Leeb, H. and Pötscher, B.M. (2005). Model selection and inference: Facts and fiction. *Econometric Theory* **21** 21–59. MR2153856 https://doi.org/10.1017/S0266466605050036

Lewis, S.M. and Raftery, A.E. (1997). Estimating Bayes factors via posterior simulation with the Laplace-Metropolis estimator. *J. Amer. Statist. Assoc.* **92** 648–655. MR1467855 https://doi.org/10.2307/2965712

Li, R. and Liang, H. (2008). Variable selection in semiparametric regression modeling. *Ann. Statist.* **36** 261–286. MR2387971 https://doi.org/10.1214/009053607000000604

Liao, Z. and Shi, X. (2020). A nondegenerate Vuong test and post selection confidence intervals for semi/nonparametric models. *Quant. Econ.* **11** 983–1017. MR4131178 https://doi.org/10.3982/qe1312

McElroy, T.S. (2016). Nonnested model comparisons for time series. *Biometrika* **103** 905–914. MR3620447 https://doi.org/10.1093/biomet/asw048

Opsomer, J.D. and Ruppert, D. (1998). A fully automated bandwidth selection method for fitting additive models. *J. Amer. Statist. Assoc.* **93** 605–619. MR1631333 https://doi.org/10.2307/2670112

Owen, A.B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75** 237–249. MR0946049 https://doi.org/10.1093/biomet/75.2.237

Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120. MR1041387 https://doi.org/10.1214/aos/1176347494

Owen, A.B. (2001). *Empirical Likelihood*. New York: CRC Press/CRC.

Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22** 300–325. MR1272085 https://doi.org/10.1214/aos/1176325370

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *Econom. J.* **5** 1–39. MR1909299 https://doi.org/10.1111/1368-423X.t01-1-00071

Schennach, S.M. and Wilhelm, D. (2017). A simple parametric model selection test. *J. Amer. Statist. Assoc.* **112** 1663–1674. MR3750889 https://doi.org/10.1080/01621459.2016.1224716

Schumaker, L.L. (1981). *Spline Functions: Basic Theory*. New York: Wiley.

Shi, X. (2015). A nondegenerate Vuong test. *Quant. Econ.* **6** 85–121. MR3337817 https://doi.org/10.3982/QE382

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 583–639.

Stone, C.J. (1986). Comment on “Generalized additive models” by Hastie and Tibshirani. *Statist. Sci.* **1** 312–314.

Vuong, Q.H. (1989). Likelihood ratio tests for model selection and nonnested hypotheses. *Econometrica* **57** 307–333. MR0996939 https://doi.org/10.2307/1912557

White, H. (1982a). Regularity conditions for Cox’s test of nonnested hypothesis. *J. Econometrics* **19** 301–318. MR0672058 https://doi.org/10.1016/0304-4076(82)90007-0

White, H. (1982b). Maximum likelihood estimation of misspecified models. *Econometrica* **50** 1–25. MR0640163 https://doi.org/10.2307/1912526

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 49–67. MR2212574 https://doi.org/10.1111/j.1467-9868.2005.00532.x

Bernoulli **30**(2), 2024, 1482–1501
https://doi.org/10.3150/23-BEJ1641

Sectional Voronoi tessellations: Characterization and high-dimensional limits

ANNA GUSAKOVA[1,a], ZAKHAR KABLUCHKO[1,b] and CHRISTOPH THÄLE[2,c]

[1] *Institute of Mathematical Stochastic, Münster University, Münster, Germany,* [a]*gusakova@uni-muenster.de,* [b]*zakhar.kabluchko@uni-muenster.de*
[2] *Faculty of Mathematics, Ruhr University Bochum, Germany,* [c]*christoph.thaele@rub.de*

The intersections of beta-Voronoi, beta-prime-Voronoi and Gaussian-Voronoi tessellations in $\mathbb{R}^d$ with an ℓ-dimensional affine subspaces, $1 \le \ell \le d-1$, are shown to be random tessellations of the same type but with different model parameters. In particular, the intersection of a classical Poisson-Voronoi tessellation with an affine subspace is shown to have the same distribution as a certain beta-Voronoi tessellation. The geometric properties of the typical cell and, more generally, typical k-faces, of the sectional Poisson-Voronoi tessellation are studied in detail. It is proved that in high dimensions, that is as $d \to \infty$, the intersection of the d-dimensional Poison-Voronoi tessellation with an affine subspace of fixed dimension ℓ converges to the ℓ-dimensional Gaussian-Voronoi tessellation.

Keywords: Beta-Voronoi tessellation; Gaussian-Voronoi tessellation; high-dimensional limit; Laguerre tessellation; Poisson point process; Poisson-Voronoi tessellation; sectional tessellation; stochastic geometry; typical cell

References

[1] Calka, P., Schreiber, T. and Yukich, J.E. (2013). Brownian limits, local limits and variance asymptotics for convex hulls in the ball. *Ann. Probab.* **41** 50–108. MR3059193 https://doi.org/10.1214/11-AOP707

[2] Calka, P. and Yukich, J.E. (2014). Variance asymptotics for random polytopes in smooth convex bodies. *Probab. Theory Related Fields* **158** 435–463. MR3152787 https://doi.org/10.1007/s00440-013-0484-1

[3] Calka, P. and Yukich, J.E. (2015). Variance asymptotics and scaling limits for Gaussian polytopes. *Probab. Theory Related Fields* **163** 259–301. MR3405618 https://doi.org/10.1007/s00440-014-0592-6

[4] Calka, P. and Yukich, J.E. (2017). Variance asymptotics and scaling limits for random polytopes. *Adv. Math.* **304** 1–55. MR3558204 https://doi.org/10.1016/j.aim.2016.08.006

[5] Calka, P. and Yukich, J.E. (2021). Convex hulls of perturbed random point sets. *Ann. Appl. Probab.* **31** 1598–1632. MR4312840 https://doi.org/10.1214/20-aap1627

[6] Chiu, S.N., van de Weygaert, R. and Stoyan, D. (1996). The sectional Poisson Voronoi tessellation is not a Voronoi tessellation. *Adv. in Appl. Probab.* **28** 356–376. MR1387880 https://doi.org/10.2307/1428061

[7] Gusakova, A., Kabluchko, Z. and Thäle, C. (2022). The β-Delaunay tessellation: Description of the model and geometry of typical cells. *Adv. in Appl. Probab.* **54** 1252–1290. MR4505686 https://doi.org/10.1017/apr.2022.6

[8] Gusakova, A., Kabluchko, Z. and Thäle, C. (2022). The β-Delaunay tessellation II: The Gaussian limit tessellation. *Electron. J. Probab.* **27** 62. MR4419469 https://doi.org/10.1214/22-ejp782

[9] Gusakova, A., Kabluchko, Z. and Thäle, C. (2022). The β-Delaunay tessellation III: Kendall's problem and limit theorems in high dimensions. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 23–50. MR4359784 https://doi.org/10.30757/alea.v19-02

[10] Gusakova, A., Kabluchko, Z. and Thäle, C. (2023). The β-Delaunay tessellation IV: Mixing properties and central limit theorems. *Stoch. Dyn.* **23** 2350021. MR4605623 https://doi.org/10.1142/S0219493723500211

[11] Kabluchko, Z. (2021). Angles of random simplices and face numbers of random polytopes. *Adv. Math.* **380** Paper No. 107612, 68. MR4205707 https://doi.org/10.1016/j.aim.2021.107612

[12] Kabluchko, Z. (2021). Recursive scheme for angles of random simplices, and applications to random polytopes. *Discrete Comput. Geom.* **66** 902–937. MR4310599 https://doi.org/10.1007/s00454-020-00259-z
[13] Kabluchko, Z., Temesvari, D. and Thäle, C. (2019). Expected intrinsic volumes and facet numbers of random beta-polytopes. *Math. Nachr.* **292** 79–105. MR3909222 https://doi.org/10.1002/mana.201700255
[14] Kabluchko, Z. and Thäle, C. (2021). The typical cell of a Voronoi tessellation on the sphere. *Discrete Comput. Geom.* **66** 1330–1350. MR4333294 https://doi.org/10.1007/s00454-021-00315-2
[15] Kabluchko, Z., Thäle, C. and Zaporozhets, D. (2020). Beta polytopes and Poisson polyhedra: f-vectors and angles. *Adv. Math.* **374** 107333, 63. MR4131401 https://doi.org/10.1016/j.aim.2020.107333
[16] 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
[17] Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge: Cambridge Univ. Press. MR3791470
[18] Lautensack, C. (2007). Random Laguerre Tessellations Ph.D. thesis.
[19] Lautensack, C. and Zuyev, S. (2008). Random Laguerre tessellations. *Adv. in Appl. Probab.* **40** 630–650. MR2454026 https://doi.org/10.1239/aap/1222868179
[20] Miles, R.E. Sectional Voronoi tessellations. *Rev. Union Math. Argent.* **29** 301–327.
[21] Nagel, W. (2010). Stereology. In *New Perspectives in Stochastic Geometry* 451–475. Oxford: Oxford Univ. Press. MR2654687
[22] Okabe, A., Boots, B., Sugihara, K. and Chiu, S.N. (2000). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed. *Wiley Series in Probability and Statistics*. Chichester: Wiley. With a foreword by D. G. Kendall. MR1770006 https://doi.org/10.1002/9780470317013
[23] Reiss, R.-D. (1993). *A Course on Point Processes*. *Springer Series in Statistics*. New York: Springer. MR1199815 https://doi.org/10.1007/978-1-4613-9308-5
[24] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. Reprint of the 1987 original. MR2364939
[25] Schlottmann, M. (1993). Periodic and quasi-periodic Laguerre tilings. *Internat. J. Modern Phys. B* **7** 1351–1363. MR1215338 https://doi.org/10.1142/S0217979293002365
[26] Schneider, R. and Weil, W. (2000). *Stochastische Geometrie*. *Teubner Skripten zur Mathematischen Stochastik. [Teubner Texts on Mathematical Stochastics]*. Stuttgart: B. G. Teubner. MR1794753 https://doi.org/10.1007/978-3-322-80106-7
[27] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. *Probability and Its Applications (New York)*. Berlin: Springer. MR2455326 https://doi.org/10.1007/978-3-540-78859-1
[28] Schreiber, T. and Yukich, J.E. (2008). Variance asymptotics and central limit theorems for generalized growth processes with applications to convex hulls and maximal points. *Ann. Probab.* **36** 363–396. MR2370608 https://doi.org/10.1214/009117907000000259
[29] Stoyan, D., Kendall, W.S. and Mecke, J. (1987). *Stochastic Geometry and Its Applications*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. Chichester: Wiley. With a foreword by D. G. Kendall. MR0895588

Bernoulli **30**(2), 2024, 1502–1524
https://doi.org/10.3150/23-BEJ1642

On Z-mean reflected BSDEs

JOFFREY DERCHU[1,a] and THIBAUT MASTROLIA[2,b]

[1]*École Polytechnique, CMAP, Palaiseau, France,* [a]*derchu@polytechnique.edu*
[2]*UC Berkeley, IEOR, Berkeley, USA,* [b]*mastrolia@berkeley.edu*

In this paper we investigate the existence of (minimal) supersolutions to BSDEs with mean-reflection on the Z component. We first prove that classical methods to obtain conditions for the existence of supersolutions to BSDEs cannot be applied for this type of constraints. We show that, contrary to BSDEs with mean-reflections on the Y component, we cannot expect a supersolution with a deterministic increasing process K. Nonetheless, we give conditions for the existence of a supersolution for a stochastic component K and under various constraints. Finally, we turn to the existence of minimal supersolution by formalizing some previous arguments on the time-inconsistency of such problems. We formalize some previous arguments on the time-inconsistency of such problems, proving that a minimal supersolution is necessarily a solution in our framework. We apply the results to a replication problem with consumption-investment strategy under law constraints on the investment strategy. We show that the only strategy that might be optimal is the one with no investment.

Keywords: Constrained BSDEs; Malliavin calculus

References

Agram, N., Hu, Y. and Øksendal, B. (2022). Mean-field backward stochastic differential equations and applications. *Systems Control Lett.* **162** Paper No. 105196, 7 pp. MR4400089 https://doi.org/10.1016/j.sysconle.2022.105196

Ansel, J.-P. and Stricker, C. (1994). Couverture des actifs contingents et prix maximum. *Ann. Inst. Henri Poincaré Probab. Stat.* **30** 303–315. MR1277002

Antonelli, F. (1993). Backward-forward stochastic differential equations. *Ann. Appl. Probab.* **3** 777–793. MR1233625

Bouchard, B., Cheridito, P. and Hu, Y. (2018). BSDE formulation of combined regular and singular stochastic control problems. Preprint. Available at arXiv:1801.03336.

Bouchard, B., Elie, R. and Moreau, L. (2018). Regularity of BSDEs with a convex constraint on the gains-process. *Bernoulli* **24** 1613–1635. MR3757510 https://doi.org/10.3150/16-BEJ806

Briand, P., Elie, R. and Hu, Y. (2018). BSDEs with mean reflection. *Ann. Appl. Probab.* **28** 482–510. MR3770882 https://doi.org/10.1214/17-AAP1310

Briand, P., Cardaliaguet, P., Chaudru de Raynal, P. and Hu, Y. (2020). Forward and backward stochastic differential equations with normal constraints in law. *Stochastic Process. Appl.* **130** 7021–7097. MR4167200 https://doi.org/10.1016/j.spa.2020.07.007

Buckdahn, R. and Hu, Y. (1998). Hedging contingent claims for a large investor in an incomplete market. *Adv. in Appl. Probab.* **30** 239–255. MR1618845 https://doi.org/10.1239/aap/1035228002

Buckdahn, R., Li, J. and Peng, S. (2009). Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Process. Appl.* **119** 3133–3154. MR2568268 https://doi.org/10.1016/j.spa.2009.05.002

Chen, Y., Hamadène, S. and Mu, T. (2023). Mean-field doubly reflected backward stochastic differential equations. *Numer. Algebra Control Optim.* **13** 431–460. MR4588426 https://doi.org/10.3934/naco.2022012

Cvitanić, J. and Karatzas, I. (1993). Hedging contingent claims with constrained portfolios. *Ann. Appl. Probab.* **3** 652–681. MR1233619

Cvitanić, J., Karatzas, I. and Soner, H.M. (1998). Backward stochastic differential equations with constraints on the gains-process. *Ann. Probab.* **26** 1522–1551. MR1675035 https://doi.org/10.1214/aop/1022855872

1350-7265

Djehiche, B., Elie, R. and Hamadène, S. (2023). Mean-field reflected backward stochastic differential equations. *Ann. Appl. Probab.* **33** 2493–2518. MR4612648 https://doi.org/10.1214/20-aap1657

El Karoui, N., Peng, S. and Quenez, M.C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71. MR1434407 https://doi.org/10.1111/1467-9965.00022

El Karoui, N. and Quenez, M.-C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.* **33** 29–66. MR1311659 https://doi.org/10.1137/S0363012992232579

El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S. and Quenez, M.C. (1997). Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *Ann. Probab.* **25** 702–737. MR1434123 https://doi.org/10.1214/aop/1024404416

Essaky, E.H. and Hassani, M. (2011). General existence results for reflected BSDE and BSDE. *Bull. Sci. Math.* **135** 442–466. MR2817457 https://doi.org/10.1016/j.bulsci.2011.04.003

Föllmer, H. and Schweizer, M. (1991). Hedging of contingent claims under incomplete information. In *Applied Stochastic Analysis (London, 1989)*. *Stochastics Monogr.* **5** 389–414. New York: Gordon and Breach. MR1108430

Hamadène, S. and Lepeltier, J.P. (1995). Backward equations, stochastic control and zero-sum stochastic differential games. *Stoch. Stoch. Rep.* **54** 221–231. MR1382117 https://doi.org/10.1080/17442509508834006

Hu, Y., Moreau, R. and Wang, F. (2022). General Mean Reflected BSDEs. Preprint. Available at arXiv:2211.01187.

Imkeller, P., Dos Reis, G. and Zhang, J. (2010). Results on numerics for FBSDE with drivers of quadratic growth. In *Contemporary Quantitative Finance* 159–182. Berlin: Springer. MR2732845 https://doi.org/10.1007/978-3-642-03479-4_9

Jacka, S.D. (1992). A martingale representation result and an application to incomplete financial markets. *Math. Finance* **2** 239–250.

Kobylanski, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28** 558–602. MR1782267 https://doi.org/10.1214/aop/1019160253

Mastrolia, T., Possamaï, D. and Réveillac, A. (2017). On the Malliavin differentiability of BSDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 464–492. MR3606749 https://doi.org/10.1214/15-AIHP723

Nakatsu, T. (2013). Absolute continuity of the laws of a multi-dimensional stochastic differential equation with coefficients dependent on the maximum. *Statist. Probab. Lett.* **83** 2499–2506. MR3144031 https://doi.org/10.1016/j.spl.2013.07.011

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Berlin: Springer. MR2200233

Pardoux, É. and Peng, S.G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61. MR1037747 https://doi.org/10.1016/0167-6911(90)90082-6

Pardoux, É. and Peng, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications (Charlotte, NC, 1991)*. *Lect. Notes Control Inf. Sci.* **176** 200–217. Berlin: Springer. MR1176785 https://doi.org/10.1007/BFb0007334

Peng, S. (1999). Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob–Meyer's type. *Probab. Theory Related Fields* **113** 473–499. MR1717527 https://doi.org/10.1007/s004400050214

Peng, S. and Xu, M. (2010). Reflected BSDE with a constraint and its applications in an incomplete market. *Bernoulli* **16** 614–640. MR2730642 https://doi.org/10.3150/09-BEJ227

Peng, S. and Xu, M. (2013). Constrained BSDEs, viscosity solutions of variational inequalities and their applications. *Math. Control Relat. Fields* **3** 233–244. MR3031142 https://doi.org/10.3934/mcrf.2013.3.233

Bernoulli **30**(2), 2024, 1525–1554
https://doi.org/10.3150/23-BEJ1643

Bayesian multiscale analysis of the Cox model

BO Y.-C. NING[1,a] and ISMAËL CASTILLO[2,b]

[1]*Department of Statistics, University of California, Davis, California 95616, USA,* [a]*bycning@ucdavis.edu*
[2]*Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, Paris 75252, France ,*
[b]*ismael.castillo@upmc.fr*

Piecewise constant priors are routinely used in the Bayesian Cox proportional hazards model for survival analysis. Despite its popularity, large sample properties of this Bayesian method are not yet well understood. This work provides a unified theory for posterior distributions in this setting, not requiring the priors to be conjugate. We first derive contraction rate results for wide classes of histogram priors on the unknown hazard function and prove asymptotic normality of linear functionals of the posterior hazard in the form of Bernstein–von Mises theorems. Second, using recently developed multiscale techniques, we derive functional limiting results for the cumulative hazard and survival function. Frequentist coverage properties of Bayesian credible sets are investigated: we prove that certain easily computable credible bands for the survival function are optimal frequentist confidence bands. We conduct simulation studies that confirm these predictions, with an excellent behavior particularly in finite samples. Our results suggest that the Bayesian approach can provide an easy solution to obtain both the coefficients estimate and the credible bands for survival function in practice.

Keywords: Bayesian Cox model; frequentist analysis of Bayesian procedures; piecewise constant prior; parametric and nonparametric Bernstein–von Mises theorems; survival analysis; supremum-norm contraction rate

References

Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. *Springer Series in Statistics*. New York: Springer. MR1198884 https://doi.org/10.1007/978-1-4612-4348-9

Breslow, N.E. (1972). Contribution to the discussion of the paper by D. R. Cox. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **34** 216–217.

Brilleman, S.L., Wolfe, R., Moreno-Betancur, M. and Crowther, M.J. (2020). Simulating survival data using the simsurv R package. *J. Stat. Softw.* **97** 1–27. https://doi.org/10.18637/jss.v097.i03

Burridge, J. (1981). Empirical Bayes analysis of survival time data. *J. Roy. Statist. Soc. Ser. B* **43** 65–75. MR0610379

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

Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. MR3127856 https://doi.org/10.1214/13-AOS1133

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

Castillo, I. and Rousseau, J. (2015a). A Bernstein–von Mises theorem for smooth functionals in semiparametric models. *Ann. Statist.* **43** 2353–2383. MR3405597 https://doi.org/10.1214/15-AOS1336

Castillo, I. and Rousseau, J. (2015b). Supplement to “A Bernstein–von Mises theorem for smooth functionals in semiparametric models”. *Ann. Statist.* **43** 1–21.

Castillo, I. and van der Pas, S. (2021). Multiscale Bayesian survival analysis. *Ann. Statist.* **49** 3559–3582. MR4352541 https://doi.org/10.1214/21-aos2097

Cox, D.R. (1972). Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220. MR0341758

Damien, P., Laud, P.W. and Smith, A.F.M. (1996). Implementation of Bayesian non-parametric inference based on beta processes. *Scand. J. Stat.* **23** 27–36. MR1380479

De Blasi, P. and Hjort, N.L. (2009). The Bernstein–von Mises theorem in semiparametric competing risks models. *J. Statist. Plann. Inference* **139** 2316–2328. MR2507993 https://doi.org/10.1016/j.jspi.2008.10.018

De Blasi, P., Peccati, G. and Prünster, I. (2009). Asymptotics for posterior hazards. *Ann. Statist.* **37** 1906–1945. MR2533475 https://doi.org/10.1214/08-AOS631

Equeter, L., Ducobu, F., Rivière-Lorphèvre, E., Serra, R. and Dehombreux, P. (2020). An analytic approach to the Cox proportional hazards model for estimating the lifespan of cutting tools. *J. Manuf. Mater. Process.* **27** 4.

Fisher, L.D. and Lin, D.Y. (1999). Time-dependent covariates in the Cox proportional-hazards regression model. *Annu. Rev. Public Health* **20** 145–157.

Florens, J.P., Mouchart, M. and Rolin, J.M. (1999). Semi- and nonparametric Bayesian analysis of duration models with Dirichlet priors: A survey. *Int. Stat. Rev.* **67** 187–210.

Gerds, T.A. and Kattan, M.W. (2021). *Medical Risk Prediction Models: With Ties to Machine Learning*, 1st ed. Boca Raton: CRC Press/CRC.

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

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

Guilloux, A., Lemler, S. and Taupin, M.-L. (2016). Adaptive kernel estimation of the baseline function in the Cox model with high-dimensional covariates. *J. Multivariate Anal.* **148** 141–159. MR3493026 https://doi.org/10.1016/j.jmva.2016.03.002

Hjort, N.L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259–1294. MR1062708 https://doi.org/10.1214/aos/1176347749

Ibrahim, J.G., Chen, M.-H. and Sinha, D. (2001). *Bayesian Survival Analysis*. *Springer Series in Statistics*. New York: Springer. MR1876598 https://doi.org/10.1007/978-1-4757-3447-8

Isobe, T., Feigelson, E.D. and Nelson, P.I. (1986). Statistical methods for astronomical data with upper limits. II. Correlation and regression. *Astrophys. J.* **306** 490–507.

Kalbfleisch, J.D. (1978). Non-parametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc. Ser. B* **40** 214–221. MR0517442

Kim, Y. (2006). The Bernstein–von Mises theorem for the proportional hazard model. *Ann. Statist.* **34** 1678–1700. MR2283713 https://doi.org/10.1214/009053606000000533

Kim, Y. and Lee, J. (2001). On posterior consistency of survival models. *Ann. Statist.* **29** 666–686. MR1865336 https://doi.org/10.1214/aos/1009210685

Kim, Y. and Lee, J. (2004). A Bernstein–von Mises theorem in the nonparametric right-censoring model. *Ann. Statist.* **32** 1492–1512. MR2089131 https://doi.org/10.1214/009053604000000526

Li, J. and Ma, S. (2013). *Survival Analysis in Medicine and Genetics*. *Chapman & Hall/CRC Biostatistics Series*. Boca Raton, FL: CRC Press. MR3616722

Lin, D.Y. (2007). On the Breslow estimator. *Lifetime Data Anal.* **13** 471–480. MR2416534 https://doi.org/10.1007/s10985-007-9048-y

Lin, D.Y., Fleming, T.R. and Wei, L.J. (1994). Confidence bands for survival curves under the proportional hazards model. *Biometrika* **81** 73–81. MR1279657 https://doi.org/10.2307/2337051

Ning, B.Y.-C. and Castillo, I. (2024). Supplement to “Bayesian multiscale analysis of the Cox model.” https://doi.org/10.3150/23-BEJ1643SUPP

Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* **11** 453–466. MR0696058 https://doi.org/10.1214/aos/1176346152

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

Rivoirard, V. and Rousseau, J. (2012). Bernstein–von Mises theorem for linear functionals of the density. *Ann. Statist.* **40** 1489–1523. MR3015033 https://doi.org/10.1214/12-AOS1004

Scheike, T.H. and Zhang, M.-J. (2002). An additive-multiplicative Cox–Aalen regression model. *Scand. J. Stat.* **29** 75–88. MR1894382 https://doi.org/10.1111/1467-9469.00065

Scheike, T.H. and Zhang, M.-J. (2008). Flexible competing risks regression modeling and goodness-of-fit. *Lifetime Data Anal.* **14** 464–483. MR2464770 https://doi.org/10.1007/s10985-008-9094-0

Schemper, M. (2002). Cox analysis of survival data with non-proportional hazard functions. *Statistician* **41** 455–465.

Sparapani, R.A., Logan, B.R., McCulloch, R.E. and Laud, P.W. (2016). Nonparametric survival analysis using Bayesian Additive Regression Trees (BART). *Stat. Med.* **35** 2741–2753. MR3513715 https://doi.org/10.1002/sim.6893

Subbotin, M.T. (1923). On the law of frequency of error. *Mat. Sb.* **31** 296–301.

van der Pas, S. and Castillo, I. (2021). BayesSurvival: Bayesian Survival Analysis for Right Censored Data. R package version 0.2.0.

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

Xu, Y., Thall, P.F., Hua, W. and Andersson, B.S. (2019). Bayesian non-parametric survival regression for optimizing precision dosing of intravenous busulfan in allogeneic stem cell transplantation. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **68** 809–828. MR3937475

Bernoulli **30**(2), 2024, 1555–1585
https://doi.org/10.3150/23-BEJ1645

Spine for interacting populations and sampling

VINCENT BANSAYE[a]

CMAP, Ecole polytechnique, Palaiseau, France, [a]*vincent.bansaye@polytechnique.edu*

We consider some Markov jump processes which model structured populations with interactions via density dependence. We propose a Markov construction involving a distinguished individual (spine) which allows us to describe the random tree and random sample at a given time via a change of probability. This spine construction involves the extension of the type space of individuals to include the state of the population. The jump rates off the spine individual can also be modified. We exploit this approach to study issues concerning population dynamics. For single type populations, we derive the phase diagram of a growth fragmentation model with competition as well as the growth of the size of transient birth and death processes which permit multiple births. We also describe the ancestral lineages of a uniform sample in multitype populations.

Keywords: Interactions; jump Markov process; martingales; populations; positive semigroup; random tree; spine

References

[1] Addario-Berry, L. and Penington, S. (2017). The front location in branching Brownian motion with decay of mass. *Ann. Probab.* **45** 3752–3794. MR3729614 https://doi.org/10.1214/16-AOP1148

[2] Anderson, D.F. and Kurtz, T.G. (2011). *Continuous Time Markov Chain Models for Chemical Reaction Networks* 3–42. New York: Springer.

[3] Athreya, K.B. (2000). Change of measures for Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli* **6** 323–338. MR1748724 https://doi.org/10.2307/3318579

[4] Bansaye, V., Cloez, B., Gabriel, P. and Marguet, A. (2022). A non-conservative Harris ergodic theorem. *J. Lond. Math. Soc. (2)* **106** 2459–2510. MR4498558 https://doi.org/10.1112/jlms.12639

[5] Bansaye, V., Delmas, J.-F., Marsalle, L. and Tran, V.C. (2011). Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *Ann. Appl. Probab.* **21** 2263–2314. MR2895416 https://doi.org/10.1214/10-AAP757

[6] Bansaye, V. and Méléard, S. (2015). *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems* **1**. Cham: Springer; Columbus, OH: MBI Mathematical Biosciences Institute, Ohio State Univ. MR3380810 https://doi.org/10.1007/978-3-319-21711-6

[7] Berestycki, J., Fittipaldi, M.C. and Fontbona, J. (2018). Ray-Knight representation of flows of branching processes with competition by pruning of Lévy trees. *Probab. Theory Related Fields* **172** 725–788. MR3877546 https://doi.org/10.1007/s00440-017-0819-4

[8] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge: Cambridge Univ. Press. MR2253162 https://doi.org/10.1017/CBO9780511617768

[9] Bertoin, J. (2017). Markovian growth-fragmentation processes. *Bernoulli* **23** 1082–1101. MR3606760 https://doi.org/10.3150/15-BEJ770

[10] Bertoin, J. and Watson, A.R. (2018). A probabilistic approach to spectral analysis of growth-fragmentation equations. *J. Funct. Anal.* **274** 2163–2204. MR3767431 https://doi.org/10.1016/j.jfa.2018.01.014

[11] Calvez, V., Henry, B., Méléard, S. and Tran, V.C. (2022). Dynamics of lineages in adaptation to a gradual environmental change. *Ann. Henri Lebesgue* **5** 729–777. MR4482341 https://doi.org/10.5802/ahl.135

[12] Chauvin, B. and Rouault, A. (1988). KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Theory Related Fields* **80** 299–314. MR0968823 https://doi.org/10.1007/BF00356108

1350-7265

[13] Cloez, B. (2017). Limit theorems for some branching measure-valued processes. *Adv. in Appl. Probab.* **49** 549–580. MR3668388 https://doi.org/10.1017/apr.2017.12

[14] Del Moral, P. (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. *Probability and Its Applications (New York)*. New York: Springer. MR2044973 https://doi.org/10.1007/978-1-4684-9393-1

[15] Eckhoff, M., Kyprianou, A.E. and Winkel, M. (2015). Spines, skeletons and the strong law of large numbers for superdiffusions. *Ann. Probab.* **43** 2545–2610. MR3395469 https://doi.org/10.1214/14-AOP944

[16] Engländer, J. (2015). *Spatial Branching in Random Environments and with Interaction*. *Advanced Series on Statistical Science & Applied Probability* **20**. Hackensack, NJ: World Scientific Co. Pte. Ltd. MR3362353 https://doi.org/10.1142/8991

[17] Engländer, J., Harris, S.C. and Kyprianou, A.E. (2010). Strong law of large numbers for branching diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 279–298. MR2641779 https://doi.org/10.1214/09-AIHP203

[18] Etheridge, A.M. and Kurtz, T.G. (2019). Genealogical constructions of population models. *Ann. Probab.* **47** 1827–1910. MR3980910 https://doi.org/10.1214/18-AOP1266

[19] 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

[20] Georgii, H.-O. and Baake, E. (2003). Supercritical multitype branching processes: The ancestral types of typical individuals. *Adv. in Appl. Probab.* **35** 1090–1110. MR2014271 https://doi.org/10.1239/aap/1067436336

[21] Harris, S.C., Hesse, M. and Kyprianou, A.E. (2016). Branching Brownian motion in a strip: Survival near criticality. *Ann. Probab.* **44** 235–275. MR3456337 https://doi.org/10.1214/14-AOP972

[22] Harris, S.C., Johnston, S.G.G. and Roberts, M.I. (2020). The coalescent structure of continuous-time Galton-Watson trees. *Ann. Appl. Probab.* **30** 1368–1414. MR4133376 https://doi.org/10.1214/19-AAP1532

[23] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. Amsterdam: North-Holland. MR1011252

[24] Jagers, P. and Nerman, O. (1996). The asymptotic composition of supercritical multi-type branching populations. In *Séminaire de Probabilités, XXX*. *Lecture Notes in Math.* **1626** 40–54. Berlin: Springer. MR1459475 https://doi.org/10.1007/BFb0094640

[25] Karlin, S. and McGregor, J.L. (1957). The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.* **85** 489–546. MR0091566 https://doi.org/10.2307/1992942

[26] Karlin, S. and Taylor, H.M. (1975). *A First Course in Stochastic Processes*, 2nd ed. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0356197

[27] Keller, G., Kersting, G. and Rösler, U. (1987). On the asymptotic behaviour of discrete time stochastic growth processes. *Ann. Probab.* **15** 305–343. MR0877606

[28] Klebaner, F.C. (1984). Geometric rate of growth in population-size-dependent branching processes. *J. Appl. Probab.* **21** 40–49. MR0732669 https://doi.org/10.2307/3213662

[29] Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y. (1997). A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*. *IMA Vol. Math. Appl.* **84** 181–185. New York: Springer. MR1601737 https://doi.org/10.1007/978-1-4612-1862-3_14

[30] Kurtz, T.G. (1981). *Approximation of Population Processes*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **36**. Philadelphia, PA: SIAM. MR0610982

[31] Kurtz, T.G. and Rodrigues, E.R. (2011). Poisson representations of branching Markov and measure-valued branching processes. *Ann. Probab.* **39** 939–984. MR2789580 https://doi.org/10.1214/10-AOP574

[32] Küster, P. (1985). Asymptotic growth of controlled Galton-Watson processes. *Ann. Probab.* **13** 1157–1178. MR0806215

[33] Lambert, A. (2010). The contour of splitting trees is a Lévy process. *Ann. Probab.* **38** 348–395. MR2599603 https://doi.org/10.1214/09-AOP485

[34] Le, V., Pardoux, E. and Wakolbinger, A. (2013). "Trees under attack": A Ray-Knight representation of Feller's branching diffusion with logistic growth. *Probab. Theory Related Fields* **155** 583–619. MR3034788 https://doi.org/10.1007/s00440-011-0408-x

[35] Lyons, R. (1997). A simple path to Biggins' martingale convergence for branching random walk. In *Classical and Modern Branching Processes (Minneapolis, MN, 1994)*. *IMA Vol. Math. Appl.* **84** 217–221. New York: Springer. MR1601749 https://doi.org/10.1007/978-1-4612-1862-3_17
[36] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164
[37] Marguet, A. (2019). Uniform sampling in a structured branching population. *Bernoulli* **25** 2649–2695. MR4003561 https://doi.org/10.3150/18-BEJ1066
[38] Mischler, S. and Scher, J. (2016). Spectral analysis of semigroups and growth-fragmentation equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **33** 849–898. MR3489637 https://doi.org/10.1016/j.anihpc.2015.01.007
[39] Norris, J.R. (1997). *Markov Chains*. *Cambridge Series in Statistical and Probabilistic Mathematics* **2**. Cambridge: Cambridge Univ. Press. MR1600720

Bernoulli **30**(2), 2024, 1586–1610
https://doi.org/10.3150/23-BEJ1646

Strong and weak convergence for the averaging principle of DDSDE with singular drift

MENGYU CHENG[1,2,a], ZIMO HAO[2,3,b] and MICHAEL RÖCKNER[2,4,c]

[1]*School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China,* [a]*mengyu.cheng@hotmail.com*
[2]*Fakultät für Mathematik, Universität Bielefeld, 33615 Bielefeld, Germany,* [b]*zhao@math.uni-bielefeld.de,* [c]*roeckner@math.uni-bielefeld.de*
[3]*School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China*
[4]*Academy of Mathematics and Systems Science, CAS, Beijing 100190, China*

In this paper, we study the averaging principle for distribution dependent stochastic differential equations with drift in localized L^p spaces. Using Zvonkin's transformation and estimates for solutions to Kolmogorov equations, we prove that the solutions of the original system strongly and weakly converge to the solution of the averaged system as the time scale ε goes to zero. Moreover, we obtain rates of the strong and weak convergence that depend on p.

Keywords: Averaging principle; distribution dependent SDE; heat kernel

References

[1] Bakhtin, V. and Kifer, Y. (2004). Diffusion approximation for slow motion in fully coupled averaging. *Probab. Theory Related Fields* **129** 157–181. MR2063374 https://doi.org/10.1007/s00440-003-0326-7

[2] Barbu, V. and Röckner, M. (2018). Probabilistic representation for solutions to nonlinear Fokker-Planck equations. *SIAM J. Math. Anal.* **50** 4246–4260. MR3835244 https://doi.org/10.1137/17M1162780

[3] Barbu, V. and Röckner, M. (2020). From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48** 1902–1920. MR4124528 https://doi.org/10.1214/19-AOP1410

[4] Barbu, V. and Röckner, M. (2022). Nonlinear Fokker-Planck equations with fractional Laplacian and McKean-Vlasov SDEs with Lévy-Noise. Available at arXiv:2210.05612.

[5] Barbu, V. and Röckner, M. (2023). Uniqueness for nonlinear Fokker-Planck equations and for McKean-Vlasov SDEs: The degenerate case. *J. Funct. Anal.* **285** Paper No. 109980. MR4583740 https://doi.org/10.1016/j.jfa.2023.109980

[6] Bogoliubov, N.N. and Mitropolsky, Y.A. (1961). *Asymptotic Methods in the Theory of Non-linear Oscillations*. New York: Gordon and Breach Science Publishers.

[7] Carmona, R. and Delarue, F. (2013). Probabilistic analysis of mean-field games. *SIAM J. Control Optim.* **51** 2705–2734. MR3072222 https://doi.org/10.1137/120883499

[8] Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. II. Mean Field Games with Common Noise and Master Equations*. *Probability Theory and Stochastic Modelling* **84**. Cham: Springer. MR3753660

[9] Carrillo, J.A., Gvalani, R.S., Pavliotis, G.A. and Schlichting, A. (2020). Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. *Arch. Ration. Mech. Anal.* **235** 635–690. MR4062483 https://doi.org/10.1007/s00205-019-01430-4

[10] Cerrai, S. (2009). A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *Ann. Appl. Probab.* **19** 899–948. MR2537194 https://doi.org/10.1214/08-AAP560

[11] Cerrai, S. and Freidlin, M. (2009). Averaging principle for a class of stochastic reaction-diffusion equations. *Probab. Theory Related Fields* **144** 137–177. MR2480788 https://doi.org/10.1007/s00440-008-0144-z

[12] Cerrai, S. and Lunardi, A. (2017). Averaging principle for nonautonomous slow-fast systems of stochastic reaction-diffusion equations: The almost periodic case. *SIAM J. Math. Anal.* **49** 2843–2884. MR3679916 https://doi.org/10.1137/16M1063307

1350-7265

[13] Cheng, M., Hao, Z. and Röckner, M. (2022). Strong and weak convergence for averaging principle of DDSDE with singular drift. Available at arXiv:2207.12108.

[14] Cheng, M., Hao, Z. and Röckner, M. (2024). Supplement to "Strong and weak convergence for the averaging principle of DDSDE with singular drift." https://doi.org/10.3150/23-BEJ1646SUPP

[15] Cheng, M. and Liu, Z. (2023). The second Bogolyubov theorem and global averaging principle for SPDEs with monotone coefficients. *SIAM J. Math. Anal.* **55** 1100–1144. MR4579723 https://doi.org/10.1137/21M1443698

[16] Crippa, G. and De Lellis, C. (2008). Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.* **616** 15–46. MR2369485 https://doi.org/10.1515/CRELLE.2008.016

[17] Dareiotis, K. and Gerencsér, M. (2020). On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift. *Electron. J. Probab.* **25** 82. MR4125787 https://doi.org/10.1214/20-ejp479

[18] Duan, J. and Wang, W. (2014). *Effective Dynamics of Stochastic Partial Differential Equations*. *Elsevier Insights*. Amsterdam: Elsevier. MR3289240

[19] Freidlin, M.I. and Wentzell, A.D. (2012). *Random Perturbations of Dynamical Systems*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Heidelberg: Springer. Translated from the 1979 Russian original by Joseph Szücs. MR2953753 https://doi.org/10.1007/978-3-642-25847-3

[20] Hairer, M. and Li, X.-M. (2020). Averaging dynamics driven by fractional Brownian motion. *Ann. Probab.* **48** 1826–1860. MR4124526 https://doi.org/10.1214/19-AOP1408

[21] Hammersley, W.R.P., Šiška, D. and Szpruch, Ł. (2021). McKean-Vlasov SDEs under measure dependent Lyapunov conditions. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1032–1057. MR4260494 https://doi.org/10.1214/20-aihp1106

[22] Han, Y. (2022). Solving McKean-Vlasov SDEs via relative entropy. Available at arXiv:2204.05709.

[23] Hao, Z., Röckner, M. and Zhang, X. (2021). Euler scheme for density dependent stochastic differential equations. *J. Differ. Equ.* **274** 996–1014. MR4189000 https://doi.org/10.1016/j.jde.2020.11.018

[24] Hao, Z., Röckner, M. and Zhang, X. (2022). Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions. Available at arXiv:2204.07952.

[25] Has'minskiĭ, R.Z. (1968). On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)* **4** 260–279. MR0260052

[26] Hong, W., Li, S. and Liu, W. (2022). Strong convergence rates in averaging principle for slow-fast McKean-Vlasov SPDEs. *J. Differ. Equ.* **316** 94–135. MR4374850 https://doi.org/10.1016/j.jde.2022.01.039

[27] Kifer, Y. (2004). Some recent advances in averaging. In *Modern Dynamical Systems and Applications* 385–403. Cambridge: Cambridge Univ. Press. MR2093312

[28] Krylov, N.V. and Bogolyubov, N.N. (1943). *Introduction to Non-linear Mechanics*. Princeton, NJ: Princeton Univ. Press.

[29] Krylov, N.V. and Röckner, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** 154–196. MR2117951 https://doi.org/10.1007/s00440-004-0361-z

[30] Lacker, D. (2018). On a strong form of propagation of chaos for McKean-Vlasov equations. *Electron. Commun. Probab.* **23** 45. MR3841406 https://doi.org/10.1214/18-ECP150

[31] Lacker, D. (2023). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.* **4** 377–432. MR4595391 https://doi.org/10.2140/pmp.2023.4.377

[32] Lê, K. and Ling, C. (2021). Taming singular stochastic differential equations: A numerical method. Available at arXiv:2110.01343.

[33] Maslowski, B., Seidler, J. and Vrkoč, I. (1991). An averaging principle for stochastic evolution equations. II. *Math. Bohem.* **116** 191–224. MR1112004

[34] McKean, H.P. Jr. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907

[35] Mishura, Y. and Veretennikov, A. (2020). Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *Theory Probab. Math. Statist.* **103** 59–101. MR4421344 https://doi.org/10.1090/tpms/1135

[36] Pei, B., Inahama, Y. and Xu, Y. (2021). Averaging principle for fast-slow system driven by mixed fractional Brownian rough path. *J. Differ. Equ.* **301** 202–235. MR4305934 https://doi.org/10.1016/j.jde.2021.08.006

[37] Röckner, M., Sun, X. and Xie, L. (2019). Strong and weak convergence in the averaging principle for SDEs with Hölder coefficients. Available at arXiv:1907.09256.

[38] Röckner, M., Sun, X. and Xie, Y. (2021). Strong convergence order for slow-fast McKean-Vlasov stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 547–576. MR4255184 https://doi.org/10.1214/20-aihp1087

[39] Röckner, M., Xie, L. and Yang, L. (2023). Asymptotic behavior of multiscale stochastic partial differential equations with Hölder coefficients.. *J. Funct. Anal.* **285** Paper No. 110103. https://doi.org/10.1016/j.jfa.2023.110103

[40] Röckner, M. and Zhang, X. (2021). Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli* **27** 1131–1158. MR4255229 https://doi.org/10.3150/20-bej1268

[41] Sanders, J.A. and Verhulst, F. (1985). *Averaging Methods in Nonlinear Dynamical Systems*. *Applied Mathematical Sciences* **59**. New York: Springer. MR0810620 https://doi.org/10.1007/978-1-4757-4575-7

[42] Scheutzow, M. (2013). A stochastic Gronwall lemma. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **16** 1350019. MR3078830 https://doi.org/10.1142/S0219025713500197

[43] Shen, G., Xiang, J. and Wu, J.-L. (2022). Averaging principle for distribution dependent stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *J. Differ. Equ.* **321** 381–414. MR4396023 https://doi.org/10.1016/j.jde.2022.03.015

[44] Sznitman, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX—1989*. *Lecture Notes in Math.* **1464** 165–251. Berlin: Springer. MR1108185 https://doi.org/10.1007/BFb0085169

[45] Veretennikov, A.Y. (1999). On large deviations in the averaging principle for SDEs with a "full dependence". *Ann. Probab.* **27** 284–296. MR1681106 https://doi.org/10.1214/aop/1022677263

[46] Wang, F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.* **128** 595–621. MR3739509 https://doi.org/10.1016/j.spa.2017.05.006

[47] Wang, W. and Roberts, A.J. (2012). Average and deviation for slow-fast stochastic partial differential equations. *J. Differ. Equ.* **253** 1265–1286. MR2927381 https://doi.org/10.1016/j.jde.2012.05.011

[48] 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

[49] Zhang, X. (2010). Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *J. Funct. Anal.* **258** 1361–1425. MR2565842 https://doi.org/10.1016/j.jfa.2009.11.006

[50] Zhang, X. (2023). Weak solutions of McKean-Vlasov SDEs with supercritical drifts. *Commun. Math. Stat.* https://doi.org/10.1007/s40304-021-00277-0

[51] Zhang, X. and Zhao, G. (2018). Singular Brownian diffusion processes. *Commun. Math. Stat.* **6** 533–581. MR3877717 https://doi.org/10.1007/s40304-018-0164-7

[52] Zhao, G. (2020). On distribution depend SDEs with singular drifts. Available at arXiv:2003.04829v3.

[53] Zvonkin, A.K. (1974). A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb. (N.S.)* **93** 129–149. (Russian).

Bernoulli **30**(2), 2024, 1611–1639
https://doi.org/10.3150/23-BEJ1647

Gaussian Whittle–Matérn fields on metric graphs

DAVID BOLIN[1,a], ALEXANDRE B. SIMAS[1,b] and JONAS WALLIN[2,c]

[1]*Statistics Program, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia,* [a]*david.bolin@kaust.edu.sa,* [b]*alexandre.simas@kaust.edu.sa*

[2]*Department of Statistics, Lund University, Lund, Sweden,* [c]*jonas.wallin@stat.lu.se*

We define a new class of Gaussian processes on compact metric graphs such as street or river networks. The proposed models, the Whittle–Matérn fields, are defined via a fractional stochastic differential equation on the compact metric graph and are a natural extension of Gaussian fields with Matérn covariance functions on Euclidean domains to the non-Euclidean metric graph setting. Existence of the processes, as well as some of their main properties, such as sample path regularity are derived. The model class in particular contains differentiable processes. To the best of our knowledge, this is the first construction of a differentiable Gaussian process on general compact metric graphs. Further, we prove an intrinsic property of these processes: that they do not change upon addition or removal of vertices with degree two. Finally, we obtain Karhunen–Loève expansions of the processes, provide numerical experiments, and compare them to Gaussian processes with isotropic covariance functions.

Keywords: Gaussian processes; networks; quantum graphs; stochastic partial differential equations

References

[1] Anderes, E., Møller, J. and Rasmussen, J.G. (2020). Isotropic covariance functions on graphs and their edges. *Ann. Statist.* **48** 2478–2503. MR4134803 https://doi.org/10.1214/19-AOS1896
[2] Arioli, M. and Benzi, M. (2018). A finite element method for quantum graphs. *IMA J. Numer. Anal.* **38** 1119–1163. MR3829156 https://doi.org/10.1093/imanum/drx029
[3] Baddeley, A., Nair, G., Rakshit, S. and McSwiggan, G. (2017). "Stationary" point processes are uncommon on linear networks. *Stat* **6** 68–78. MR3613182 https://doi.org/10.1002/sta4.135
[4] Bakka, H., Vanhatalo, J., Illian, J.B., Simpson, D. and Rue, H. (2019). Non-stationary Gaussian models with physical barriers. *Spat. Stat.* **29** 268–288. MR3903698 https://doi.org/10.1016/j.spasta.2019.01.002
[5] Berkolaiko, G. and Kuchment, P. (2013). *Introduction to Quantum Graphs*. *Mathematical Surveys and Monographs* **186**. Providence, RI: Amer. Math. Soc. MR3013208 https://doi.org/10.1090/surv/186
[6] Bolin, D. (2014). Spatial Matérn fields driven by non-Gaussian noise. *Scand. J. Stat.* **41** 557–579. MR3249417 https://doi.org/10.1111/sjos.12046
[7] Bolin, D. and Kirchner, K. (2020). The rational SPDE approach for Gaussian random fields with general smoothness. *J. Comput. Graph. Statist.* **29** 274–285. MR4116041 https://doi.org/10.1080/10618600.2019.1665537
[8] Bolin, D., Kirchner, K. and Kovács, M. (2020). Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA J. Numer. Anal.* **40** 1051–1073. MR4092278 https://doi.org/10.1093/imanum/dry091
[9] Bolin, D. and Wallin, J. (2020). Multivariate type G Matérn stochastic partial differential equation random fields. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 215–239. MR4060983
[10] Borovitskiy, V., Azangulov, I., Terenin, A., Mostowsky, P., Deisenroth, M. and Durrande, N. (2021). Matérn Gaussian processes on graphs. In *International Conference on Artificial Intelligence and Statistics* 2593–2601. PMLR.
[11] Chandler-Wilde, S.N., Hewett, D.P. and Moiola, A. (2015). Interpolation of Hilbert and Sobolev spaces: Quantitative estimates and counterexamples. *Mathematika* **61** 414–443. MR3343061 https://doi.org/10.1112/S0025579314000278

1350-7265

[12] Cox, S.G. and Kirchner, K. (2020). Regularity and convergence analysis in Sobolev and Hölder spaces for generalized Whittle-Matérn fields. *Numer. Math.* **146** 819–873. MR4182088 https://doi.org/10.1007/s00211-020-01151-x
[13] Cronie, O., Moradi, M. and Mateu, J. (2020). Inhomogeneous higher-order summary statistics for point processes on linear networks. *Stat. Comput.* **30** 1221–1239. MR4137248 https://doi.org/10.1007/s11222-020-09942-w
[14] Daon, Y. and Stadler, G. (2018). Mitigating the influence of the boundary on PDE-based covariance operators. *Inverse Probl. Imaging* **12** 1083–1102. MR3836580 https://doi.org/10.3934/ipi.2018045
[15] Demengel, F. and Demengel, G. (2007). *Functional Spaces for the Theory of Elliptic Partial Differential Equations*. *Universitext*. London: Springer. MR2895178 https://doi.org/10.1007/978-1-4471-2807-6
[16] Dunson, D.B., Wu, H.-T. and Wu, N. (2022). Graph based Gaussian processes on restricted domains. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 414–439. MR4412992 https://doi.org/10.1111/rssb.12486
[17] Evans, L.C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Providence, RI: Amer. Math. Soc. MR2597943 https://doi.org/10.1090/gsm/019
[18] Hildeman, A., Bolin, D. and Rychlik, I. (2021). Deformed SPDE models with an application to spatial modeling of significant wave height. *Spat. Stat.* **42** Paper No. 100449, 27 pp. MR4233262 https://doi.org/10.1016/j.spasta.2020.100449
[19] Kostenko, A., Mugnolo, D. and Nicolussi, N. (2022). Self-adjoint and Markovian extensions of infinite quantum graphs. *J. Lond. Math. Soc. (2)* **105** 1262–1313. MR4400947 https://doi.org/10.1112/jlms.12539
[20] Krätschmer, V. and Urusov, M. (2023). A Kolmogorov-Chentsov type theorem on general metric spaces with applications to limit theorems for Banach-Valued processes. *J. Theoret. Probab.* **36** 1454–1486. MR4621071 https://doi.org/10.1007/s10959-022-01207-8
[21] Lindgren, F., Bolin, D. and Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. *Spat. Stat.* **50** Paper No. 100599, 29 pp. MR4439328 https://doi.org/10.1016/j.spasta.2022.100599
[22] Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. MR2853727 https://doi.org/10.1111/j.1467-9868.2011.00777.x
[23] Lunardi, A. (2018). *Interpolation Theory*. *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]* **16**. Pisa: Edizioni della Normale. MR3753604 https://doi.org/10.1007/978-88-7642-638-4
[24] Matérn, B. (1960). *Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations*. *Meddelanden Från Statens Skogsforskningsinstitut* **49**. Stockholm: Statens Skogsforskningsinstitut. MR0169346
[25] McLean, W. (2000). *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge: Cambridge Univ. Press. MR1742312
[26] Odžak, A. and Šćeta, L. (2019). On the Weyl law for quantum graphs. *Bull. Malays. Math. Sci. Soc.* **42** 119–131. MR3894619 https://doi.org/10.1007/s40840-017-0469-9
[27] Okabe, A. and Sugihara, K. (2012). *Spatial Analysis Along Networks: Statistical and Computational Methods*. New York: Wiley.
[28] Sanz-Alonso, D. and Yang, R. (2022). The SPDE approach to Matérn fields: Graph representations. *Statist. Sci.* **37** 519–540. MR4497230 https://doi.org/10.1214/21-sts838
[29] Serio, A. (2021). On extremal eigenvalues of the graph Laplacian. *J. Phys. A* **54** Paper No. 015202, 14 pp. MR4190125 https://doi.org/10.1088/1751-8121/abc59c
[30] Steinwart, I. and Scovel, C. (2012). Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr. Approx.* **35** 363–417. MR2914365 https://doi.org/10.1007/s00365-012-9153-3
[31] Ver Hoef, J.M., Peterson, E. and Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environ. Ecol. Stat.* **13** 449–464. MR2297373 https://doi.org/10.1007/s10651-006-0022-8
[32] Whittle, P. (1963). Stochastic processes in several dimensions. *Bull. Inst. Int. Stat.* **40** 974–994. MR0173287
[33] Zeidler, E. (1995). *Applied Functional Analysis: Applications to Mathematical Physics*. *Applied Mathematical Sciences* **108**. New York: Springer. MR1347691

Bernoulli **30**(2), 2024, 1640–1668
https://doi.org/10.3150/23-BEJ1648

Exact detection thresholds and minimax optimality of Chatterjee's correlation coefficient

ARNAB AUDDY[1,a], NABARUN DEB[2,b] and SAGNIK NANDY[3,c]

[1]*Department of Biostatistics, University of Pennsylvania, Philadelphia, USA,*
[a]*arnab.auddy@pennmedicine.upenn.edu*
[2]*Econometrics and Statistics, Booth School of Business, University of Chicago, Chicago, USA,*
[b]*nabarun.deb@chicagobooth.edu*
[3]*Department of Statistics, University of Pennsylvania, Philadelphia, USA,* [c]*sagnik@wharton.upenn.edu*

Recently, Chatterjee (2021) introduced a new rank-based correlation coefficient which can be used to measure the strength of dependence between two random variables. This coefficient has already attracted much attention as it converges to the Dette-Siburg-Stoimenov measure (see Dette et al. (2013)), which equals 0 if and only if the variables are independent and 1 if and only if one variable is a function of the other. Further, Chatterjee's coefficient is computable in (near) linear time, which makes it appropriate for large-scale applications. In this paper, we expand the theoretical understanding of Chatterjee's coefficient in two directions: (a) First we consider the problem of testing for independence using Chatterjee's correlation. We obtain its asymptotic distribution under any changing sequence of alternatives converging to the null hypothesis (of independence). We further obtain a general result that gives exact detection thresholds and limiting power for Chatterjee's test of independence under natural nonparametric alternatives converging to the null. As applications of this general result, we prove a $n^{-1/4}$ detection boundary for this test and compute explicitly the limiting local power on the detection boundary for popularly studied alternatives in the literature. (b) We then construct a test for non-trivial levels of dependence using Chatterjee's coefficient. In contrast to testing for independence, we prove that, in this case, Chatterjee's coefficient indeed yields a minimax optimal procedure with a $n^{-1/2}$ detection boundary. Our proof techniques rely on Stein's method of exchangeable pairs, a non-asymptotic projection result, and information theoretic lower bounds.

Keywords: Independence testing; Kantorovic-Wasserstein distance; Le Cam's two-point method; local power; Stein's method for locally dependent structures

References

Ansari, J. and Fuchs, S. (2022). A simple extension of Azadkia & Chatterjee's rank correlation to a vector of endogenous variables. Preprint. Available at arXiv:2212.01621.

Arias-Castro, E., Huang, R. and Verzelen, N. (2020). Detection of sparse positive dependence. *Electron. J. Stat.* **14** 702–730. MR4057605 https://doi.org/10.1214/19-EJS1675

Auddy, A., Deb, N. and Nandy, S. (2024). Supplement to "Exact detection thresholds and minimax optimality of Chatterjee's correlation coefficient." https://doi.org/10.3150/23-BEJ1648SUPP

Azadkia, M. and Chatterjee, S. (2021). A simple measure of conditional dependence. *Ann. Statist.* **49** 3070–3102. MR4352523 https://doi.org/10.1214/21-aos2073

Azadkia, M., Taeb, A. and Bühlmann, P. (2021). A Fast Non-parametric Approach for Causal Structure Learning in Polytrees. Preprint. Available at arXiv:2111.14969.

Bergsma, W. and Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20** 1006–1028. MR3178526 https://doi.org/10.3150/13-BEJ514

Berrett, T.B., 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

1350-7265

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

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

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Stat.* **21** 593–600. MR0039190 https://doi.org/10.1214/aoms/1177729754

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

Cao, S. and Bickel, P.J. (2020). Correlations with tailored extremal properties. Preprint. Available at arXiv:2008.10177.

Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. MR2435859 https://doi.org/10.1214/07-AOP370

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

Chatterjee, S. (2022). A survey of some recent developments in measures of association. Preprint. Available at arXiv:2211.04702.

Chatterjee, S., Salimi, A. and Lee, J.Y. (2020). Insights into amyotrophic lateral sclerosis linked Pro525Arg mutation in the fused in sarcoma protein through in silico analysis and molecular dynamics simulation. *J. Biomol. Struct. Dyn.* 1–14.

Chatterjee, S. and Vidyasagar, M. (2022). Estimating large causal polytree skeletons from small samples. Preprint. Available at arXiv:2209.07028.

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

Deb, N., Ghosal, P. and Sen, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. Preprint. 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

Dette, H., Siburg, K.F. and Stoimenov, P.A. (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat.* **40** 21–41. MR3024030 https://doi.org/10.1111/j.1467-9469.2011.00767.x

Dhar, S.S., Dassios, A. and Bergsma, W. (2016). A study of the power and robustness of a new test for independence against contiguous alternatives. *Electron. J. Stat.* **10** 330–351. MR3466185 https://doi.org/10.1214/16-EJS1107

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

Even-Zohar, C. (2020). independence: Fast Rank Tests. Preprint. Available at arXiv:2010.09712.

Even-Zohar, C. and Leng, C. (2021). Counting small permutation patterns. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)* 2288–2302. Philadelphia, PA: SIAM. MR4262569 https://doi.org/10.1137/1.9781611976465.136

Farlie, D.J.G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47** 307–323. MR0119312 https://doi.org/10.1093/biomet/47.3-4.307

Farlie, D.J.G. (1961). The asymptotic efficiency of Daniels's generalized correlation coefficients. *J. Roy. Statist. Soc. Ser. B* **23** 128–142. MR0124108

Fruciano, C., Colangelo, P., Castiglia, R. and Franchini, P. (2020). Does divergence from normal patterns of integration increase as chromosomal fusions increase in number? A test on a house mouse hybrid zone. *Curr. Zool.* **66** 527–538. https://doi.org/10.1093/cz/zoaa035

Fuchs, S. (2021). Quantifying directed dependence via dimension reduction. Preprint. Available at arXiv:2112.10147.

Gamboa, F., Klein, T. and Lagnoux, A. (2018). Sensitivity analysis based on Cramér–von Mises distance. *SIAM/ASA J. Uncertain. Quantificat.* **6** 522–548. MR3784121 https://doi.org/10.1137/15M1025621

Gieser, P.W. (1993). *A New Nonparametric Test for Independence Between Two Sets of Variates*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–University of Florida. MR2691505

Gini, C. (1914). *L'ammontare e la Composizione della Ricchezza delle Nazioni* **62**. Fratelli Bocca.

Griessenberger, F., Junker, R.R. and Trutschnig, W. (2022). On a multivariate copula-based dependence measure and its estimation. *Electron. J. Stat.* **16** 2206–2251. MR4401220 https://doi.org/10.1214/22-ejs2005

Hájek, J., Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests*, 2nd ed. *Probability and Mathematical Statistics*. San Diego, CA: Academic Press. MR1680991

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

Han, F. and Huang, Z. (2022). Azadkia-Chatterjee's correlation coefficient adapts to manifold data. Preprint. Available at arXiv:2209.11156.

Heller, Y. and Heller, R. (2016). Computing the Bergsma Dassios sign-covariance. Preprint. Available at arXiv:1605.08732.

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

Holma, A. (2022). Correlation coefficient based feature screening: With applications to microarray data.

Huang, Z., Deb, N. and Sen, B. (2022). Kernel partial correlation coefficient—A measure of conditional dependence. *J. Mach. Learn. Res.* **23** Paper No. [216], 58 pp. MR4577169 https://doi.org/10.1086/287487

Ingster, Y.I. (1987). Minimax testing of nonparametric hypotheses on a distribution density in the L_p metrics. *Theory Probab. Appl.* **31** 333–337.

Ingster, Y.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I, II, III. *Math. Methods Statist.* **2** 85–114. MR1257978

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

Kim, I., Balakrishnan, S. and Wasserman, L. (2022). Minimax optimality of permutation tests. *Ann. Statist.* **50** 225–251. MR4382015 https://doi.org/10.1214/21-aos2103

Konijn, H.S. (1956). On the power of certain tests for independence in bivariate populations. *Ann. Math. Stat.* **27** 300–323. MR0079384 https://doi.org/10.1214/aoms/1177728260

Kössler, W. and Rödel, E. (2007). The asymptotic efficacies and relative efficiencies of various linear rank tests for independence. *Metrika* **65** 3–28. MR2288045 https://doi.org/10.1007/s00184-006-0055-x

Ledwina, T. (1986). On the limiting Pitman efficiency of some rank tests of independence. *J. Multivariate Anal.* **20** 265–271. MR0866074 https://doi.org/10.1016/0047-259X(86)90082-5

Lin, Z. and Han, F. (2022). Limit theorems of Chatterjee's rank correlation. Preprint. Available at arXiv:2204.08031.

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

Lin, Z. and Han, F. (2023b). On the failure of the bootstrap for Chatterjee's rank correlation. Preprint. Available at arXiv:2303.14088.

Morgenstern, D. (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsbl. Math. Statist.* **8** 234–235. MR0081575

Nandy, P., Weihs, L. and Drton, M. (2016). Large-sample theory for the Bergsma-Dassios sign covariance. *Electron. J. Stat.* **10** 2287–2311. MR3541972 https://doi.org/10.1214/16-EJS1166

Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests*. Cambridge: Cambridge Univ. Press. MR1335235 https://doi.org/10.1017/CBO9780511530081

Nikitin, Y.Y. and Stepanova, N. (2003). Pitman efficiency of independence tests based on weighted rank statistics. *J. Math. Sci.* **118** 5596–5606.

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **58** 240–242.

Rosenblatt, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.* **3** 1–14. MR0428579

Shi, H., Drton, M. and Han, F. (2020). On the power of Chatterjee rank correlation. Preprint. Available at arXiv:2008.11619.

Shi, H., Drton, M. and Han, F. (2021). On Azadkia-Chatterjee's conditional dependence coefficient. Preprint. Available at arXiv:2108.06827.

Spearman, C. (1906). Footrule for measuring correlation. *Br. J. Psychol.* **2** 89.

Spearman, C. (1961). The proof and measurement of association between two things.

Stepanova, N. and Wang, S. (2008). Asymptotic efficiency of the Blest-type tests for independence. *Aust. N. Z. J. Stat.* **50** 217–233. MR2455429 https://doi.org/10.1111/j.1467-842X.2008.00513.x

Strothmann, C., Dette, H. and Siburg, K.F. (2022). Rearranged dependence measures. Preprint. Available at arXiv:2201.03329.

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

Wang, X., Jiang, B. and Liu, J.S. (2017). Generalized R-squared for detecting dependence. *Biometrika* **104** 129–139. MR3626486 https://doi.org/10.1093/biomet/asw071

Weihs, L., Drton, M. and Leung, D. (2016). Efficient computation of the Bergsma-Dassios sign covariance. *Comput. Statist.* **31** 315–328. MR3481807 https://doi.org/10.1007/s00180-015-0639-x

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

Yanagimoto, T. (1970). On measures of association and a related problem. *Ann. Inst. Statist. Math.* **22** 57–63.

Zhang, Q. (2023). On the asymptotic null distribution of the symmetrized Chatterjee's correlation coefficient. *Statist. Probab. Lett.* **194** Paper No. 109759, 7 pp. MR4525660 https://doi.org/10.1016/j.spl.2022.109759

Bernoulli **30**(2), 2024, 1669–1693
https://doi.org/10.3150/23-BEJ1649

Estimating a regression function in exponential families by model selection

JUNTONG CHEN[a]

Department of mathematics, University of Luxembourg, Esch-sur-Alzette, Luxembourg, [a]*juntong.chen@uni.lu*

Let $(W_1,Y_1),\dots,(W_n,Y_n)$ be n pairs of independent random variables. We assume that, for each $i \in \{1,\dots,n\}$, the conditional distribution of Y_i given W_i belongs to a one-parameter exponential family with parameter $\boldsymbol{\gamma}^{\star}(W_i) \in \mathbb{R}$. The statistical goal is to estimate these conditional distributions. We consider a model selection procedure which works based on a general assumption that each of the model is VC-subgraph. We establish a non-asymptotic risk bound for the resulting estimator with respect to a Hellinger-type distance. By leveraging this result, we extend several findings previously explored in Gaussian regression to the regression in exponential families. Specifically, we address the curse of dimensionality by imposing structural assumptions, such as general additive and multiple index structures, on $\boldsymbol{\gamma}^{\star}$. We also study model selection for ReLU neural networks, and provide a concrete example of how ReLU neural networks can achieve a significantly faster convergence rate than traditional models. When $\boldsymbol{\gamma}^{\star}$ is close to a composition of several Hölder functions, we show that under a suitable parametrization of the exponential family, our estimator achieves the same rate of convergence as in the Gaussian case. Combining with a lower bound, the rate is minimax optimal up to a logarithmic term. Finally, we apply the model selection procedure to address adaptation and variable selection problems in exponential families.

Keywords: Generalized additive structure; model selection; multiple index structure; ReLU neural networks; regression in exponential family; variable selection

References

Akakpo, N. (2012). Adaptation to anisotropy and inhomogeneity via dyadic piecewise polynomial selection. *Math. Methods Statist.* **21** 1–28. MR2901269 https://doi.org/10.3103/S1066530712010012

Antoniadis, A. and Sapatinas, T. (2001). Wavelet shrinkage for natural exponential families with quadratic variance functions. *Biometrika* **88** 805–820. MR1859411 https://doi.org/10.1093/biomet/88.3.805

Baraud, Y. and Birgé, L. (2014). Estimating composite functions by model selection. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 285–314. MR3161532 https://doi.org/10.1214/12-AIHP516

Baraud, Y. and Birgé, L. (2018). Rho-estimators revisited: General theory and applications. *Ann. Statist.* **46** 3767–3804. MR3852668 https://doi.org/10.1214/17-AOS1675

Baraud, Y., Birgé, L. and Sart, M. (2017). A new method for estimation and model selection: ρ-estimation. *Invent. Math.* **207** 425–517. MR3595933 https://doi.org/10.1007/s00222-016-0673-5

Baraud, Y. and Chen, J. (2020). Robust estimation of a regression function in exponential families. arXiv preprint. Available at arXiv:2011.01657.

Bartlett, P.L., Maiorov, V. and Meir, R. (1998). Almost linear VC-dimension bounds for piecewise polynomial networks. *Neural Comput.* **10** 2159–2173.

Bartlett, P.L., Harvey, N., Liaw, C. and Mehrabian, A. (2019). Nearly-tight VC-dimension and pseudodimension bounds for piecewise linear neural networks. *J. Mach. Learn. Res.* **20** Paper No. 63, 17. MR3960917

Brown, L.D., Cai, T.T. and Zhou, H.H. (2010). Nonparametric regression in exponential families. *Ann. Statist.* **38** 2005–2046. MR2676882 https://doi.org/10.1214/09-AOS762

Chen, J. (2024). Supplement to “Estimating a regression function in exponential families by model selection.” https://doi.org/10.3150/23-BEJ1649SUPP

Dahmen, W., DeVore, R. and Scherer, K. (1980). Multidimensional spline approximation. *SIAM J. Numer. Anal.* **17** 380–402. MR0581486 https://doi.org/10.1137/0717033

Daubechies, I., DeVore, R., Foucart, S., Hanin, B. and Petrova, G. (2022). Nonlinear approximation and (deep) ReLU networks. *Constr. Approx.* **55** 127–172. MR4376561 https://doi.org/10.1007/s00365-021-09548-z

Fryźlewicz, P. and Nason, G.P. (2001). Poisson intensity estimation using wavelets and the Fisz transformation. Technical Report, 01/10, Department of Mathematics, Univ. Bristol, United Kingdom.

Fryzlewicz, P. and Nason, G.P. (2004). A Haar-Fisz algorithm for Poisson intensity estimation. *J. Comput. Graph. Statist.* **13** 621–638. MR2087718 https://doi.org/10.1198/106186004X2697

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR3617773

Hochmuth, R. (2002). Wavelet characterizations for anisotropic Besov spaces. *Appl. Comput. Harmon. Anal.* **12** 179–208. MR1884234 https://doi.org/10.1006/acha.2001.0377

Horowitz, J.L. and Mammen, E. (2007). Rate-optimal estimation for a general class of nonparametric regression models with unknown link functions. *Ann. Statist.* **35** 2589–2619. MR2382659 https://doi.org/10.1214/009053607000000415

Ibragimov, I.A. and Has'minskiĭ, R.Z. (1984). More on the estimation of distribution densities. *J. Sov. Math.* **25** 1155–1165.

Jia, J., Xie, F. and Xu, L. (2019). Sparse Poisson regression with penalized weighted score function. *Electron. J. Stat.* **13** 2898–2920. MR3998931 https://doi.org/10.1214/19-EJS1580

Kolaczyk, E.D. and Nowak, R.D. (2005). Multiscale generalised linear models for nonparametric function estimation. *Biometrika* **92** 119–133. MR2158614 https://doi.org/10.1093/biomet/92.1.119

Kroll, M. (2019). Non-parametric Poisson regression from independent and weakly dependent observations by model selection. *J. Statist. Plann. Inference* **199** 249–270. MR3857826 https://doi.org/10.1016/j.jspi.2018.07.003

Li, Y. and Cevher, V. (2015). Consistency of ℓ_1-regularized maximum-likelihood for compressive Poisson regression. In *2015 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)* 3606–3610.

Nunes, M.A. and Nason, G.P. (2009). A multiscale variance stabilization for binomial sequence proportion estimation. *Statist. Sinica* **19** 1491–1510. MR2589194

Nussbaum, M. (1987). Nonparametric estimation of a regression function that is smooth in a domain in $\mathbb{R}_k$. *Theory Probab. its Appl.* **31** 108–115.

Schmidt-Hieber, J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Statist.* **48** 1875–1897. MR4134774 https://doi.org/10.1214/19-AOS1875

Schumaker, L.L. (1981). *Spline Functions: Basic Theory*. *Pure and Applied Mathematics*. New York: Wiley. MR0606200

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

Suzuki, T. (2019). Adaptivity of deep ReLU network for learning in Besov and mixed smooth Besov spaces: Optimal rate and curse of dimensionality. In *7th International Conference on Learning Representations, ICLR*.

Suzuki, T. and Nitanda, A. (2021). Deep learning is adaptive to intrinsic dimensionality of model smoothness in anisotropic Besov space. In *35th Advances in Neural Information Processing Systems, NeurIPS*.

Triebel, H. (2006). *Theory of Function Spaces. III*. *Monographs in Mathematics* **100**. Basel: Birkhäuser. MR2250142

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

Yamaguti, M. and Hata, M. (1983). Weierstrass's function and chaos. *Hokkaido Math. J.* **12** 333–342. MR0719972 https://doi.org/10.14492/hokmj/1470081010