BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Nine Number Three August 2023 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 Twenty Nine Number Three August 2023 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 29 Number 3 August 2023 Pages 1741–2597 ISI/BS

Volume 29 Number 3   August 2023   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 2023 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **29**(3), 2023, 1741–1763
https://doi.org/10.3150/22-BEJ1519

Power enhancement and phase transitions for global testing of the mixed membership stochastic block model

LOUIS V. CAMMARATA[a] and ZHENG TRACY KE[b]

Statistics Department, Harvard University, Cambridge, MA 02138, USA, [a]*louis_cammarata@g.harvard.edu,* [b]*zke@fas.harvard.edu*

The mixed-membership stochastic block model (MMSBM) is a common model for social networks. Given an n-node symmetric network generated from a K-community MMSBM, we would like to test $K = 1$ versus $K > 1$. We first study the degree-based χ^2 test and the orthodox Signed Quadrilateral (oSQ) test. These two statistics estimate an order-2 polynomial and an order-4 polynomial of a "signal" matrix, respectively. We derive the asymptotic null distribution and power for both tests. However, for each test, there exists a parameter regime where its power is unsatisfactory. It motivates us to propose a power enhancement (PE) test to combine the strengths of both tests. We show that the PE test has a tractable null distribution and improves the power of both tests. To assess the optimality of PE, we consider a randomized setting, where the n membership vectors are independently drawn from a distribution on the standard simplex. We show that the success of global testing is governed by a quantity $\beta_n(K, P, h)$, which depends on the community structure matrix P and the mean vector h of memberships. For each given (K, P, h), a test is called *optimal* if it distinguishes two hypotheses when $\beta_n(K, P, h) \to \infty$. A test is called *optimally adaptive* if it is optimal for all (K, P, h). We show that the PE test is optimally adaptive, while many existing tests are only optimal for some particular (K, P, h), hence, not optimally adaptive.

Keywords: Chi-square test; degree matching; mixed memberships; phase transition; signed cycles; stochastic block model

References

[1] Airoldi, E., Blei, D., Fienberg, S. and Xing, E. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.
[2] Arias-Castro, E. and Verzelen, N. (2014). Community detection in dense random networks. *Ann. Statist.* **42** 940–969. MR3210992 https://doi.org/10.1214/14-AOS1208
[3] Banerjee, D. (2018). Contiguity and non-reconstruction results for planted partition models: The dense case. *Electron. J. Probab.* **23** Paper No. 18. MR3771755 https://doi.org/10.1214/17-EJP128
[4] Banerjee, D. and Ma, Z. (2017). Optimal hypothesis testing for stochastic block models with growing degrees. Available at arXiv:1705.05305.
[5] Banerjee, D. and Ma, Z. (2022). Asymptotic normality and analysis of variance of log-likelihood ratios in spiked random matrix models. *Ann. Statist.* To appear.
[6] Banks, J., Moore, C., Neeman, J. and Netrapalli, P. (2016). Information-theoretic thresholds for community detection in sparse networks. In *Conference on Learning Theory* 383–416. PMLR.
[7] Bubeck, S., Ding, J., Eldan, R. and Rácz, M.Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Structures Algorithms* **49** 503–532. MR3545825 https://doi.org/10.1002/rsa.20633
[8] Cammarata, L. Ke, Z.T. (2023). Supplement to "Power enhancement and phase transitions for global testing of the mixed membership stochastic block model." https://doi.org/10.3150/22-BEJ1519SUPP
[9] Gao, C. and Lafferty, J. (2017). Testing for global network structure using small subgraph statistics. Available at arXiv:1710.00862.

[10] Gao, C. and Ma, Z. (2021). Minimax rates in network analysis: Graphon estimation, community detection and hypothesis testing. *Statist. Sci.* **36** 16–33. MR4194201 https://doi.org/10.1214/19-STS736

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

[12] Ji, P., Jin, J., Ke, Z.T. and Li, W. (2022). Co-citation and co-authorship networks of statisticians. *J. Bus. Econom. Statist.* **40** 469–485.

[13] Jin, J., Ke, Z.T. and Luo, S. (2017). Estimating network memberships by simplex vertex hunting. Available at arXiv:1708.07852.

[14] Jin, J., Ke, Z.T. and Luo, S. (2018). Network global testing by counting graphlets. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018* 2338–2346.

[15] Jin, J., Ke, Z.T. and Luo, S. (2021). Optimal adaptivity of signed-polygon statistics for network testing. *Ann. Statist.* **49** 3408–3433. MR4352535 https://doi.org/10.1214/21-aos2089

[16] Jin, J., Ke, Z.T., Luo, S. and Wang, M. (2022). Optimal estimation of the number of network communities. *J. Amer. Statist. Assoc.* 1–16.

[17] Jin, J., Ke, Z.T. and Wang, W. (2017). Phase transitions for high dimensional clustering and related problems. *Ann. Statist.* **45** 2151–2189. MR3718165 https://doi.org/10.1214/16-AOS1522

[18] Lei, J. (2016). A goodness-of-fit test for stochastic block models. *Ann. Statist.* **44** 401–424. MR3449773 https://doi.org/10.1214/15-AOS1370

[19] Lu, C. and Sen, S. (2020). Contextual stochastic block model: Sharp thresholds and contiguity. Available at arXiv:2011.09841.

[20] Mossel, E., Neeman, J. and Sly, A. (2015). Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields* **162** 431–461. MR3383334 https://doi.org/10.1007/s00440-014-0576-6

[21] Wang, Y.X.R. and Bickel, P.J. (2017). Likelihood-based model selection for stochastic block models. *Ann. Statist.* **45** 500–528. MR3650391 https://doi.org/10.1214/16-AOS1457

[22] Yuan, M., Liu, R., Feng, Y. and Shang, Z. (2022). Testing community structure for hypergraphs. *Ann. Statist.* **50** 147–169. MR4382012 https://doi.org/10.1214/21-aos2099

[23] Zhang, A.Y. and Zhou, H.H. (2016). Minimax rates of community detection in stochastic block models. *Ann. Statist.* **44** 2252–2280. MR3546450 https://doi.org/10.1214/15-AOS1428

[24] Zhang, P., Moore, C. and Newman, M.E.J. (2016). Community detection in networks with unequal groups. *Phys. Rev. E* **93** 012303. MR3678034 https://doi.org/10.1103/physreve.93.012303

Bernoulli **29**(3), 2023, 1764–1791
https://doi.org/10.3150/22-BEJ1520

Testing for linearity in boundary regression models with application to maximal life expectancies

JÜRGEN KAMPF[1,a] and ALEXANDER MEISTER[2,b]

[1] *Clinic for Cardiology and Angiology, University Hospital of Essen, Hufelandstr. 55, 45122 Essen, Germany,*
[a] *juergen.kampf@uk-essen.de*
[2] *Institute of Mathematics, University of Rostock, Ulmenstr. 69/3, 18057 Rostock, Germany,*
[b] *alexander.meister@uni-rostock.de*

We consider a regression model with errors that are a.s. negative. Thus the regression function is not the expected value of the observations but the right endpoint of their support. We develop two goodness-of-fit tests for the hypotheses that the regression function is an affine function, study the asymptotic distributions of the test statistics in order to approximately fix the sizes of the tests, derive their finite-sample properties based on simulations and apply them to life expectancy data.

Keywords: Boundary regression; goodness of fit test; life expectancy

References

[1] Cuesta-Albertos, J.A., Fraiman, R. and Ransford, T. (2007). A sharp form of the Cramér-Wold theorem. *J. Theoret. Probab.* **20** 201–209. MR2324526 https://doi.org/10.1007/s10959-007-0060-7
[2] Dong, C., Li, G. and Feng, X. (2019). Lack-of-fit tests for quantile regression models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 629–648. MR3961501
[3] Drees, H., Neumeyer, N. and Selk, L. (2019). Estimation and hypotheses testing in boundary regression models. *Bernoulli* **25** 424–463. MR3892325 https://doi.org/10.3150/17-bej992
[4] González-Manteiga, W. and Crujeiras, R.M. (2013). An updated review of goodness-of-fit tests for regression models. *TEST* **22** 361–411. MR3093195 https://doi.org/10.1007/s11749-013-0327-5
[5] Hall, P. and Van Keilegom, I. (2009). Nonparametric "regression" when errors are positioned at end-points. *Bernoulli* **15** 614–633. MR2555192 https://doi.org/10.3150/08-BEJ173
[6] Huet, S. and Kuhn, E. (2015). Goodness-of-fit test for Gaussian regression with block correlated errors. *Statistics* **49** 239–266. MR3325359 https://doi.org/10.1080/02331888.2014.913047
[7] Human Mortality Database University of California at Berkeley and the Max Planck Institute for Demographic Research. www.mortality.org.
[8] Jirak, M., Meister, A. and Reiss, M. (2014). Adaptive function estimation in nonparametric regression with one-sided errors. *Ann. Statist.* **42** 1970–2002. MR3262474 https://doi.org/10.1214/14-AOS1248
[9] Kampf, J., Meister, A. (2023). Supplement to "Testing for linearity in boundary regression models with application to maximal life expectancies." https://doi.org/10.3150/22-BEJ1520SUPP
[10] Mammen, E., Van Keilegom, I. and Yu, K. (2019). Expansion for moments of regression quantiles with applications to nonparametric testing. *Bernoulli* **25** 793–827. MR3920357 https://doi.org/10.3150/17-bej986
[11] Oeppen, J. and Vaupel, J. (2002). Broken limits to life expectancy. *Science* **296** 1029–1031.
[12] Philipp, W. (1969). The central limit problem for mixing sequences of random variables. *Z. Wahrsch. Verw. Gebiete* **12** 155–171. MR0246356 https://doi.org/10.1007/BF00531648
[13] Reiss, M. and Wahl, M. (2019). Functional estimation and hypothesis testing in nonparametric boundary models. *Bernoulli* **25** 2597–2619. MR4003559 https://doi.org/10.3150/18-BEJ1064

1350-7265

[14] 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

[15] Selk, L., Tillier, C. and Marigliano, O. (2022). Multivariate boundary regression models. *Scand. J. Stat.* **49** 400–426. MR4391059 https://doi.org/10.1111/sjos.12519

[16] Zheng, J.X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* **14** 123–138. MR1613710 https://doi.org/10.1017/S0266466698141051

Bernoulli **29**(3), 2023, 1792–1820
https://doi.org/10.3150/22-BEJ1521

Power variations in fractional Sobolev spaces for a class of parabolic stochastic PDEs

CARSTEN CHONG[1,a] and ROBERT C. DALANG[2,b]

[1]*Department of Statistics, Columbia University, New York, USA,* [a]*carsten.chong@columbia.edu*
[2]*Institut de mathématiques, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland,* [b]*robert.dalang@epfl.ch*

We consider a class of parabolic stochastic PDEs on bounded domains $D \subseteq \mathbb{R}^d$ that includes the stochastic heat equation but with a fractional power γ of the Laplacian. Viewing the solution as a process with values in a scale of fractional Sobolev spaces H_r, with $r < \gamma - d/2$, we study its power variations in H_r along regular partitions of the time-axis. As the mesh size tends to zero, we find a phase transition at $r = -d/2$: the solutions have a nontrivial quadratic variation when $r < -d/2$ and a nontrivial pth order variation for $p = 2\gamma/(\gamma - d/2 - r) > 2$ when $r > -d/2$. More generally, normalized power variations of any order satisfy a genuine law of large numbers in the first case and a degenerate limit theorem in the second case. When $r < -d/2$, the quadratic variation is given explicitly via an expression that involves the spectral zeta function, which reduces to the Riemann zeta function when $d = 1$ and D is an interval.

Keywords: Stochastic heat equation; stochastic partial differential equation; fractional Laplacian; power variations; Riemann zeta function; spectral zeta function

References

[1] Barndorff-Nielsen, O.E., Corcuera, J.M. and Podolskij, M. (2011). Multipower variation for Brownian semistationary processes. *Bernoulli* **17** 1159–1194. MR2854768 https://doi.org/10.3150/10-BEJ316

[2] Basse-O'Connor, A., Lachièze-Rey, R. and Podolskij, M. (2017). Power variation for a class of stationary increments Lévy driven moving averages. *Ann. Probab.* **45** 4477–4528. MR3737916 https://doi.org/10.1214/16-AOP1170

[3] Bibinger, M. and Trabs, M. (2019). On central limit theorems for power variations of the solution to the stochastic heat equation. In *Stochastic Models, Statistics and Their Applications* (A. Steland, E. Rafajłowicz and O. Okhrin, eds.) 69–84. Cham: Springer.

[4] Bibinger, M. and Trabs, M. (2020). Volatility estimation for stochastic PDEs using high-frequency observations. *Stochastic Process. Appl.* **130** 3005–3052. MR4080737 https://doi.org/10.1016/j.spa.2019.09.002

[5] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1324786

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

[7] Chong, C. (2019). High-frequency analysis of parabolic stochastic PDEs with multiplicative noise. Available at arXiv:1908.04145.

[8] Chong, C. (2020). High-frequency analysis of parabolic stochastic PDEs. *Ann. Statist.* **48** 1143–1167. MR4102691 https://doi.org/10.1214/19-AOS1841

[9] Chong, C., Dalang, R.C. (2023). Supplement to "Power variations in fractional Sobolev spaces for a class of parabolic stochastic PDEs." https://doi.org/10.3150/22-BEJ1521SUPP

[10] Cialenco, I. and Huang, Y. (2020). A note on parameter estimation for discretely sampled SPDEs. *Stoch. Dyn.* **20** 2050016. MR4101083 https://doi.org/10.1142/S0219493720500161

1350-7265

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

[12] Corcuera, J.M., Nualart, D. and Woerner, J.H.C. (2006). Power variation of some integral fractional processes. *Bernoulli* **12** 713–735. MR2248234 https://doi.org/10.3150/bj/1155735933

[13] Corcuera, J.M., Nualart, D. and Woerner, J.H.C. (2009). Convergence of certain functionals of integral fractional processes. *J. Theoret. Probab.* **22** 856–870. MR2558655 https://doi.org/10.1007/s10959-008-0158-6

[14] Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge: Cambridge Univ. Press. MR3236753 https://doi.org/10.1017/CBO9781107295513

[15] Dalang, R.C. and Sanz-Solé, M. (2005). Regularity of the sample paths of a class of second-order spde's. *J. Funct. Anal.* **227** 304–337. MR2168077 https://doi.org/10.1016/j.jfa.2004.11.015

[16] Dalang, R.C. and Sanz-Solé, M. (2020). An Introduction to Stochastic Partial Differential Equations. In preparation.

[17] Elizalde, E. (2012). *Ten Physical Applications of Spectral Zeta Functions*, 2nd ed. *Lecture Notes in Physics* **855**. Heidelberg: Springer. MR2934635 https://doi.org/10.1007/978-3-642-29405-1

[18] Foondun, M., Khoshnevisan, D. and Mahboubi, P. (2015). Analysis of the gradient of the solution to a stochastic heat equation via fractional Brownian motion. *Stoch. Partial Differ. Equ. Anal. Comput.* **3** 133–158. MR3350450 https://doi.org/10.1007/s40072-015-0045-y

[19] Giacomin, G., Olla, S. and Spohn, H. (2001). Equilibrium fluctuations for $\nabla\phi$ interface model. *Ann. Probab.* **29** 1138–1172. MR1872740 https://doi.org/10.1214/aop/1015345600

[20] Gilbarg, D. and Trudinger, N.S. (2001). *Elliptic Partial Differential Equations of Second Order*. *Classics in Mathematics*. Berlin: Springer. MR1814364

[21] Hawking, S.W. (1977). Zeta function regularization of path integrals in curved spacetime. *Comm. Math. Phys.* **55** 133–148. MR0524257

[22] Hildebrandt, F. and Trabs, M. (2021). Parameter estimation for SPDEs based on discrete observations in time and space. *Electron. J. Stat.* **15** 2716–2776. MR4280160 https://doi.org/10.1214/21-ejs1848

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

[24] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

[25] 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

[26] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. *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

[27] McCullagh, P. (2018). *Tensor Methods in Statistics*, 2nd ed. Mineola: Dover.

[28] Nourdin, I. (2008). Asymptotic behavior of weighted quadratic and cubic variations of fractional Brownian motion. *Ann. Probab.* **36** 2159–2175. MR2478679 https://doi.org/10.1214/07-AOP385

[29] Nourdin, I., Nualart, D. and Tudor, C.A. (2010). Central and non-central limit theorems for weighted power variations of fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 1055–1079. MR2744886 https://doi.org/10.1214/09-AIHP342

[30] Nualart, D. and Zeineddine, R. (2018). Symmetric weighted odd-power variations of fractional Brownian motion and applications. *Commun. Stoch. Anal.* **12** 37–58. MR3843909 https://doi.org/10.31390/cosa.12.1.04

[31] Nualart, E. (2018). Moment bounds for some fractional stochastic heat equations on the ball. *Electron. Commun. Probab.* **23** 41. MR3841402 https://doi.org/10.1214/18-ECP147

[32] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Milan: Springer. MR2791919 https://doi.org/10.1007/978-88-470-1679-8

[33] 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

[34] Pospíšil, J. and Tribe, R. (2007). Parameter estimates and exact variations for stochastic heat equations driven by space-time white noise. *Stoch. Anal. Appl.* **25** 593–611. MR2321899 https://doi.org/10.1080/07362990701282849

[35] Prévôt, C. and Röckner, M. (2007). *A Concise Course on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1905**. Berlin: Springer. MR2329435

[36] Sanz-Solé, M. and Sarrà, M. (2002). Hölder continuity for the stochastic heat equation with spatially correlated noise. In *Seminar on Stochastic Analysis, Random Fields and Applications, III* (R.C. Dalang, M. Dozzi and F. Russo, eds.). *Progress in Probability* **52** 259–268. Basel: Birkhäuser.

[37] Sheffield, S. (2007). Gaussian free fields for mathematicians. *Probab. Theory Related Fields* **139** 521–541. MR2322706 https://doi.org/10.1007/s00440-006-0050-1

[38] Suquet, C. (1999). Tightness in Schauder decomposable Banach spaces. In *Proceedings of the St. Petersburg Mathematical Society, Vol. V*. *Amer. Math. Soc. Transl. Ser. 2* **193** 201–224. Providence, RI: Amer. Math. Soc. MR1736910 https://doi.org/10.1090/trans2/193/08

[39] Swanson, J. (2007). Variations of the solution to a stochastic heat equation. *Ann. Probab.* **35** 2122–2159.

[40] van Neerven, J.M.A.M., Veraar, M.C. and Weis, L. (2008). Stochastic evolution equations in UMD Banach spaces. *J. Funct. Anal.* **255** 940–993. MR2433958 https://doi.org/10.1016/j.jfa.2008.03.015

[41] Vere-Jones, D. (1997). Alpha-permanents and their applications to multivariate gamma, negative binomial and ordinary binomial distributions. *New Zealand J. Math.* **26** 125–149. MR1450811

[42] Voros, A. (1987). Spectral functions, special functions and the Selberg zeta function. *Comm. Math. Phys.* **110** 439–465. MR0891947

[43] Voros, A. (1992). Spectral zeta functions. In *Zeta Functions in Geometry (Tokyo, 1990)*. *Adv. Stud. Pure Math.* **21** 327–358. Tokyo: Kinokuniya. MR1210795 https://doi.org/10.2969/aspm/02110327

[44] 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

Bernoulli **29**(3), 2023, 1821–1848
https://doi.org/10.3150/22-BEJ1522

Kolmogorov bounds for decomposable random variables and subgraph counting by the Stein–Tikhomirov method

PETER EICHELSBACHER[a] and BENEDIKT REDNOSS[b]

Faculty of Mathematics, Ruhr University Bochum, Germany, [a]*peter.eichelsbacher@rub.de,* [b]*benedikt.rednoss@rub.de*

We derive normal approximation bounds in the Kolmogorov distance for random variables possessing decompositions of Barbour, Karoński, and Ruciński (*J. Combin. Theory Ser. B* **47** (1989) 125–145). We highlight the example of standardized subgraph counts in the Erdős–Rényi random graph. We prove a bound by generalizing the argumentation of Röllin (*Probab. Engrg. Inform. Sci.* (2022) 747–773), who used the Stein–Tikhomirov method to prove a bound in the special case of standardized triangle counts. Our bounds match the best available Wasserstein bounds.

Keywords: Berry–Esseen bound; central limit theorem; characteristic function; Erdős–Rényi random graph; Kolmogorov distance; normal approximation; Stein's method; Stein–Tikhomirov method; subgraph count

References

[1] Arras, B. and Houdré, C. (2019). *On Stein's Method for Infinitely Divisible Laws with Finite First Moment. SpringerBriefs in Probability and Mathematical Statistics*. Cham: Springer. MR3931309 https://doi.org/10.1007/978-3-030-15017-4

[2] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2017). A new approach to the Stein-Tikhomirov method with applications to the second Wiener chaos and Dickman convergence. Version 2. Preprint. Available at arXiv:1605.06819v2.

[3] Barbour, A.D., Karoński, M. and Ruciński, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** 125–145. MR1047781 https://doi.org/10.1016/0095-8956(89)90014-2

[4] Chen, L.H.Y. (1978). Two central limit problems for dependent random variables. *Z. Wahrsch. Verw. Gebiete* **43** 223–243. MR0517439 https://doi.org/10.1007/BF00536204

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

[6] Ebner, B. (2021). On combining the zero bias transform and the empirical characteristic function to test normality. *ALEA Lat. Am. J. Probab. Math. Stat.* **18** 1029–1045. MR4282180 https://doi.org/10.30757/alea.v18-38

[7] Esary, J.D., Proschan, F. and Walkup, D.W. (1967). Association of random variables, with applications. *Ann. Math. Stat.* **38** 1466–1474. MR0217826 https://doi.org/10.1214/aoms/1177698701

[8] Féray, V., Méliot, P.-L. and Nikeghbali, A. (2019). Mod-ϕ convergence, II: Estimates on the speed of convergence. In *Séminaire de Probabilités L* (C. Donati-Martin, A. Lejay and A. Rouault, eds.) 405–477. Cham: Springer. MR4291787 https://doi.org/10.1007/978-3-030-28535-7_15

[9] Gilbert, E.N. (1959). Random graphs. *Ann. Math. Stat.* **30** 1141–1144. MR0108839 https://doi.org/10.1214/aoms/1177706098

[10] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. MR1484792 https://doi.org/10.1214/aoap/1043862419

[11] Janson, S., Łuczak, T. and Rucinski, A. (2000). *Random Graphs*. *Wiley-Interscience Series in Discrete Mathematics and Optimization*. New York: Wiley Interscience. MR1782847 https://doi.org/10.1002/9781118032718

[12] Krokowski, K., Reichenbachs, A. and Thäle, C. (2017). Discrete Malliavin–Stein method: Berry–Esseen bounds for random graphs and percolation. *Ann. Probab.* **45** 1071–1109. MR3630293 https://doi.org/10.1214/15-AOP1081

[13] Loève, M. (1977). *Probability Theory. I*, 4th ed. *Graduate Texts in Mathematics, Vol. 45*. New York: Springer. MR0651017 https://doi.org/10.1007/978-1-4684-9464-8

[14] McGinley, W.G. and Sibson, R. (1975). Dissociated random variables. *Math. Proc. Cambridge Philos. Soc.* **77** 185–188. MR0370714 https://doi.org/10.1017/S0305004100049513

[15] Newman, C.M. (1980). Normal fluctuations and the FKG inequalities. *Comm. Math. Phys.* **74** 119–128. MR0576267 https://doi.org/10.1007/BF01197754

[16] Privault, N. and Serafin, G. (2020). Normal approximation for sums of weighted U-statistics—application to Kolmogorov bounds in random subgraph counting. *Bernoulli* **26** 587–615. MR4036045 https://doi.org/10.3150/19-BEJ1141

[17] Raič, M. (2003). Normal approximation by Stein's method. In *Proceedings of the Seventh Young Statisticans Meeting* **21** 71–97.

[18] Röllin, A. (2022). Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős–Rényi random graph. *Probab. Engrg. Inform. Sci.* **36** 587–615. MR4459700 https://doi.org/10.1017/S0269964821000061

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

[20] Ruciński, A. (1988). When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields* **78** 1–10. MR0940863 https://doi.org/10.1007/BF00718031

[21] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Volume II: Probability Theory* (L. Le Cam, J. Neyman and E. Scott, eds.) 583–602. Berkeley, CA: Univ. California Press. MR0402873

[22] Tikhomirov, A.N. (1980). On the rate of convergence in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25** 790–809.

Bernoulli **29**(3), 2023, 1849–1876
https://doi.org/10.3150/22-BEJ1523

Adaptive lasso and Dantzig selector for spatial point processes intensity estimation

ACHMAD CHOIRUDDIN[1,a], JEAN-FRANÇOIS COEURJOLLY[2,b] and FRÉDÉRIQUE LETUÉ[2,c]

[1]*Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS), Indonesia,* [a]*choiruddin@its.ac.id*
[2]*Laboratory Jean Kuntzmann, University Grenoble Alpes, France,*
[b]*jean-francois.coeurjolly@univ-grenoble-alpes.fr,* [c]*frederique.letue@univ-grenoble-alpes.fr*

Lasso and Dantzig selector are standard procedures able to perform variable selection and estimation simultaneously. This paper is concerned with extending these procedures to spatial point process intensity estimation. We propose adaptive versions of these procedures, develop efficient computational methodologies and derive asymptotic results for a large class of spatial point processes under an original setting where the number of parameters, i.e. the number of spatial covariates considered, increases with the expected number of data points. Both procedures are compared theoretically, in a simulation study, and in a real data example.

Keywords: Estimating equations; high-dimensional statistics; linear programming; regularization methods; spatial point pattern

References

[1] Antoniadis, A., Fryzlewicz, P. and Letué, F. (2010). The Dantzig selector in Cox's proportional hazards model. *Scand. J. Stat.* **37** 531–552. MR2779635 https://doi.org/10.1111/j.1467-9469.2009.00685.x

[2] Baddeley, A., Rubak, E. and Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Boca Raton: CRC Press.

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

[4] Biscio, C.A.N. and Waagepetersen, R. (2019). A general central limit theorem and a subsampling variance estimator for α-mixing point processes. *Scand. J. Stat.* **46** 1168–1190. MR4033808 https://doi.org/10.1111/sjos.12389

[5] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge: Cambridge Univ. Press. MR2061575 https://doi.org/10.1017/CBO9780511804441

[6] Candes, E. and Romberg, J. (2005). ℓ^1-Magic: Recovery of Sparse Signals via Convex Programming. Preprint.

[7] Candes, E. and Tao, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n. *Ann. Statist.* **35** 2313–2351. MR2382644 https://doi.org/10.1214/009053606000001523

[8] Choiruddin, A., Aisah, Trisnisa, F. and Iriawan, N. (2021). Quantifying the effect of geological factors on distribution of earthquake occurrences by inhomogeneous Cox processes. *Pure Appl. Geophys.* **178** 1579–1592.

[9] Choiruddin, A., Coeurjolly, J.-F. and Letué, F. (2023). Supplement to "Adaptive lasso and Dantzig selector for spatial point processes intensity estimation." https://doi.org/10.3150/22-BEJ1523SUPP

[10] Choiruddin, A., Coeurjolly, J.-F. and Letué, F. (2017). Spatial point processes intensity estimation with a diverging number of covariates. ArXiv Preprint. Available at arXiv:1712.09562.

[11] Choiruddin, A., Coeurjolly, J.-F. and Letué, F. (2018). Convex and non-convex regularization methods for spatial point processes intensity estimation. *Electron. J. Stat.* **12** 1210–1255. MR3780731 https://doi.org/10.1214/18-EJS1408

[12] Choiruddin, A., Coeurjolly, J.-F. and Waagepetersen, R. (2021). Information criteria for inhomogeneous spatial point processes. *Aust. N. Z. J. Stat.* **63** 119–143. MR4296145 https://doi.org/10.1111/anzs.12327

1350-7265

[13] Choiruddin, A., Cuevas-Pacheco, F., Coeurjolly, J.-F. and Waagepetersen, R. (2020). Regularized estimation for highly multivariate log Gaussian Cox processes. *Stat. Comput.* **30** 649–662. MR4065224 https://doi.org/10.1007/s11222-019-09911-y

[14] Coeurjolly, J.-F. and Lavancier, F. (2019). Understanding spatial point patterns through intensity and conditional intensities. In *Stochastic Geometry*. *Lecture Notes in Math.* **2237** 45–85. Cham: Springer. MR3931583

[15] Daniel, J., Horrocks, J. and Umphrey, G.J. (2018). Penalized composite likelihoods for inhomogeneous Gibbs point process models. *Comput. Statist. Data Anal.* **124** 104–116. MR3787616 https://doi.org/10.1016/j.csda.2018.02.005

[16] Dicker, L. (2010). Regularized regression methods for variable selection and estimation Ph.D. thesis Harvard Univ.

[17] Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* **32** 928–961. MR2065194 https://doi.org/10.1214/009053604000000256

[18] Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33** 1–22.

[19] Guan, Y., Jalilian, A. and Waagepetersen, R. (2015). Quasi-likelihood for spatial point processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 677–697. MR3351450 https://doi.org/10.1111/rssb.12083

[20] Guan, Y. and Shen, Y. (2010). A weighted estimating equation approach for inhomogeneous spatial point processes. *Biometrika* **97** 867–880. MR2746157 https://doi.org/10.1093/biomet/asq043

[21] Hubbell, S.P., Condit, R. and Foster, R.B. (2005). Barro Colorado forest census plot data. Available at http://ctfs.si.edu/webatlas/datasets/bci/.

[22] James, G.M. and Radchenko, P. (2009). A generalized Dantzig selector with shrinkage tuning. *Biometrika* **96** 323–337. MR2507146 https://doi.org/10.1093/biomet/asp013

[23] James, G.M., Radchenko, P. and Lv, J. (2009). DASSO: Connections between the Dantzig selector and lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 127–142. MR2655526 https://doi.org/10.1111/j.1467-9868.2008.00668.x

[24] Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. Boca Raton, FL: CRC Press/CRC. MR2004226

[25] Rakshit, S., McSwiggan, G., Nair, G. and Baddeley, A. (2021). Variable selection using penalised likelihoods for point patterns on a linear network. *Aust. N. Z. J. Stat.* **63** 417–454. MR4374507 https://doi.org/10.1111/anzs.12341

[26] Waagepetersen, R. and Guan, Y. (2009). Two-step estimation for inhomogeneous spatial point processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 685–702. MR2749914 https://doi.org/10.1111/j.1467-9868.2008.00702.x

[27] Waagepetersen, R.P. (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics* **63** 252–258. MR2345595 https://doi.org/10.1111/j.1541-0420.2006.00667.x

[28] Yue, Y. and Loh, J.M. (2015). Variable selection for inhomogeneous spatial point process models. *Canad. J. Statist.* **43** 288–305. MR3353384 https://doi.org/10.1002/cjs.11244

[29] Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Amer. Statist. Assoc.* **101** 1418–1429. MR2279469 https://doi.org/10.1198/016214506000000735

Bernoulli **29**(3), 2023, 1877–1904
https://doi.org/10.3150/22-BEJ1524

Forward integration of bounded variation coefficients with respect to Hölder continuous processes

JOHANNA GARZÓN[1,a], JORGE A. LEÓN[2,b] and SOLEDAD TORRES[3,c]

[1]*Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia,*
[a]*mjgarzonm@unal.edu.co*
[2]*Departamento de Control Automático, Cinvestav-IPN, Ciudad de México, México,* [b]*jleon@ctrl.cinvestav.mx*
[3]*Facultad de Ingeniería, CIMFAV, Universidad de Valparaíso, Valparaíso, Chile,* [c]*soledad.torres@uv.cl*

In this article, we study the forward integral, in the Russo and Vallois sense, with respect to Hölder continuous stochastic processes Y with exponent bigger than $1/2$. Here, the integrands have the form $f(Y)$, where f is a bounded variation function. As a consequence of our results, we show that this integral agrees with the generalized Stieltjes integral given by Zähle and that, in the case that Y is fractional Brownian motion, this forward integral is equal to the divergence operator plus a trace term, which is related to the local time of Y. Moreover, the definition of the forward integral allows us to obtain a representation of the solutions to forward stochastic differential equations with a possibly discontinuous coefficient and, as a consequence of our analysis, to figure out some explicit solutions.

Keywords: Fractional Brownian motion; forward integral; stochastic differential equations; Malliavin calculus

References

[1] 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

[2] 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

[3] Azmoodeh, E., Mishura, Y. and Valkeila, E. (2009). On hedging European options in geometric fractional Brownian motion market model. *Statist. Decisions* **27** 129–143. MR2662719 https://doi.org/10.1524/stnd.2009.1021

[4] Azmoodeh, E. and Viitasaari, L. (2015). Rate of convergence for discretization of integrals with respect to fractional Brownian motion. *J. Theoret. Probab.* **28** 396–422. MR3320973 https://doi.org/10.1007/s10959-013-0495-y

[5] Besalú, M. and Nualart, D. (2011). Estimates for the solution to stochastic differential equations driven by a fractional Brownian motion with Hurst parameter $H \in (\frac{1}{3}, \frac{1}{2})$. *Stoch. Dyn.* **11** 243–263. MR2836524 https://doi.org/10.1142/S0219493711003267

[6] Chen, Z., Leskelä, L. and Viitasaari, L. (2019). Pathwise Stieltjes integrals of discontinuously evaluated stochastic processes. *Stochastic Process. Appl.* **129** 2723–2757. MR3980142 https://doi.org/10.1016/j.spa.2018.08.002

[7] Coutin, L., Nualart, D. and Tudor, C.A. (2001). Tanaka formula for the fractional Brownian motion. *Stochastic Process. Appl.* **94** 301–315. MR1840834 https://doi.org/10.1016/S0304-4149(01)00085-0

[8] Garzón, J., León, J.A. and Torres, S. (2017). Fractional stochastic differential equation with discontinuous diffusion. *Stoch. Anal. Appl.* **35** 1113–1123. MR3740753 https://doi.org/10.1080/07362994.2017.1358643

[9] Garzón, J., León, J.A. and Torres, S. (2021). Representation of solutions to sticky stochastic differential equations. *Stoch. Dyn.* To appear. https://doi.org/10.1142/S0219493723500053

1350-7265

[10] Hytönen, T., van Neerven, J., Veraar, M. and Weis, L. (2016). *Analysis in Banach Spaces. Vol. i. Martingales and Littlewood-Paley Theory. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]* **63**. Cham: Springer. MR3617205

[11] León, J.A., Nualart, D. and Tindel, S. (2017). Young differential equations with power type nonlinearities. *Stochastic Process. Appl.* **127** 3042–3067. MR3682123 https://doi.org/10.1016/j.spa.2017.01.007

[12] Nualart, D. (2003). Stochastic integration with respect to fractional Brownian motion and applications. In *Stochastic Models (Mexico City, 2002). Contemp. Math.* **336** 3–39. Providence, RI: Amer. Math. Soc. MR2037156 https://doi.org/10.1090/conm/336/06025

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

[14] Nualart, D. and Răşcanu, A. (2002). Differential equations driven by fractional Brownian motion. *Collect. Math.* **53** 55–81. MR1893308

[15] Ouknine, Y. (1988). Généralisation d'un lemme de S. Nakao et applications. *Stochastics* **23** 149–157. MR0928352 https://doi.org/10.1080/17442508808833487

[16] Ouknine, Y. and Rutkowski, M. (1990). Strong comparison of solutions of one-dimensional stochastic differential equations. *Stochastic Process. Appl.* **36** 217–230. MR1084976 https://doi.org/10.1016/0304-4149(90)90092-7

[17] 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

[18] 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

[19] Tikanmäki, H. (2012). Integral representations of some functionals of fractional Brownian motion. *Commun. Stoch. Anal.* **6** 193–212. MR2927700 https://doi.org/10.31390/cosa.6.2.02

[20] Zähle, M. (1998). Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Related Fields* **111** 333–374. MR1640795 https://doi.org/10.1007/s004400050171

[21] Zähle, M. (2002). Forward integrals and stochastic differential equations. In *Seminar on Stochastic Analysis, Random Fields and Applications, III (Ascona, 1999). Progress in Probability* **52** 293–302. Basel: Birkhäuser. MR1958824

Bernoulli **29**(3), 2023, 1905–1932
https://doi.org/10.3150/22-BEJ1525

Sample canonical correlation coefficients of high-dimensional random vectors with finite rank correlations

ZONGMING MA[a] and FAN YANG[b]

Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, PA 19104, USA,
[a]*zongming@wharton.upenn.edu,* [b]*fyang75@wharton.upenn.edu*

Consider two random vectors $\widetilde{\mathbf{x}} = A\mathbf{z} + \mathbf{C}_1^{1/2}\mathbf{x} \in \mathbb{R}^p$ and $\widetilde{\mathbf{y}} = B\mathbf{z} + \mathbf{C}_2^{1/2}\mathbf{y} \in \mathbb{R}^q$, where $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{y} \in \mathbb{R}^q$ and $\mathbf{z} \in \mathbb{R}^r$ are independent random vectors with i.i.d. entries of zero mean and unit variance, $\mathbf{C}_1$ and $\mathbf{C}_2$ are $p \times p$ and $q \times q$ deterministic population covariance matrices, and A and B are $p \times r$ and $q \times r$ deterministic factor loading matrices. With n independent observations of $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{y}}$, we study the sample canonical correlations between them. Under the sharp fourth moment condition on the entries of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, we prove the BBP transition for the sample canonical correlation coefficients (CCCs). More precisely, if a population CCC is below a threshold, then the corresponding sample CCC converges to the right edge of the bulk eigenvalue spectrum of the sample canonical correlation matrix and satisfies the famous Tracy-Widom law; if a population CCC is above the threshold, then the corresponding sample CCC converges to an outlier that is detached from the bulk eigenvalue spectrum. We prove our results in full generality, in the sense that they also hold for near-degenerate population CCCs and population CCCs that are close to the threshold.

Keywords: Canonical correlation analysis; BBP transition; Tracy-Widom law; edge eigenvalues

References

[1] Alt, J. (2017). Singularities of the density of states of random Gram matrices. *Electron. Commun. Probab.* **22** Paper No. 63, 13. MR3734102 https://doi.org/10.1214/17-ECP97

[2] Alt, J., Erdős, L. and Krüger, T. (2017). Local law for random Gram matrices. *Electron. J. Probab.* **22** Paper No. 25, 41. MR3622895 https://doi.org/10.1214/17-EJP42

[3] Bai, Z. and Yao, J. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474. MR2451053 https://doi.org/10.1214/07-AIHP118

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

[5] Baik, J., Ben Arous, G. and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. MR2165575 https://doi.org/10.1214/009117905000000233

[6] Baik, J. and Silverstein, J.W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680 https://doi.org/10.1016/j.jmva.2005.08.003

[7] Bao, Z., Hu, J., Pan, G. and Zhou, W. (2019). Canonical correlation coefficients of high-dimensional Gaussian vectors: Finite rank case. *Ann. Statist.* **47** 612–640. MR3909944 https://doi.org/10.1214/18-AOS1704

[8] Belinschi, S.T., Bercovici, H., Capitaine, M. and Février, M. (2017). Outliers in the spectrum of large deformed unitarily invariant models. *Ann. Probab.* **45** 3571–3625. MR3729610 https://doi.org/10.1214/16-AOP1144

[9] Benaych-Georges, F., Guionnet, A. and Maida, M. (2011). Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electron. J. Probab.* **16** 1621–1662. MR2835249 https://doi.org/10.1214/EJP.v16-929

1350-7265

[10] Benaych-Georges, F. and Nadakuditi, R.R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. MR2782201 https://doi.org/10.1016/j.aim.2011.02.007

[11] Bloemendal, A., Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2014). Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.* **19** no. 33, 53. MR3183577 https://doi.org/10.1214/ejp.v19-3054

[12] Bloemendal, A., Knowles, A., Yau, H.-T. and Yin, J. (2016). On the principal components of sample covariance matrices. *Probab. Theory Related Fields* **164** 459–552. MR3449395 https://doi.org/10.1007/s00440-015-0616-x

[13] Capitaine, M., Donati-Martin, C. and Féral, D. (2009). The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** 1–47. MR2489158 https://doi.org/10.1214/08-AOP394

[14] Capitaine, M., Donati-Martin, C. and Féral, D. (2012). Central limit theorems for eigenvalues of deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 107–133. MR2919200 https://doi.org/10.1214/10-AIHP410

[15] Ding, X. and Yang, F. (2020). Tracy-Widom distribution for the edge eigenvalues of Gram type random matrices. arXiv:2008.04166.

[16] Ding, X. and Yang, F. (2021). Spiked separable covariance matrices and principal components. *Ann. Statist.* **49** 1113–1138. MR4255121 https://doi.org/10.1214/20-aos1995

[17] Ding, X. and Yang, F. (2022). Edge statistics of large dimensional deformed rectangular matrices. *J. Multivariate Anal.* **192** Paper No. 105051. MR4438212 https://doi.org/10.1016/j.jmva.2022.105051

[18] Erdős, L., Knowles, A. and Yau, H.-T. (2013). Averaging fluctuations in resolvents of random band matrices. *Ann. Henri Poincaré* **14** 1837–1926. MR3119922 https://doi.org/10.1007/s00023-013-0235-y

[19] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. MR3098073 https://doi.org/10.1214/11-AOP734

[20] Féral, D. and Péché, S. (2007). The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.* **272** 185–228. MR2291807 https://doi.org/10.1007/s00220-007-0209-3

[21] Féral, D. and Péché, S. (2009). The largest eigenvalues of sample covariance matrices for a spiked population: Diagonal case. *J. Math. Phys.* **50** 073302, 33. MR2548630 https://doi.org/10.1063/1.3155785

[22] Forrester, P.J. (1993). The spectrum edge of random matrix ensembles. *Nuclear Phys. B* **402** 709–728. MR1236195 https://doi.org/10.1016/0550-3213(93)90126-A

[23] Fujikoshi, Y. (2017). High-dimensional asymptotic distributions of characteristic roots in multivariate linear models and canonical correlation analysis. *Hiroshima Math. J.* **47** 249–271. MR3719444

[24] Gao, C., Ma, Z., Ren, Z. and Zhou, H.H. (2015). Minimax estimation in sparse canonical correlation analysis. *Ann. Statist.* **43** 2168–2197. MR3396982 https://doi.org/10.1214/15-AOS1332

[25] Gao, C., Ma, Z. and Zhou, H.H. (2017). Sparse CCA: Adaptive estimation and computational barriers. *Ann. Statist.* **45** 2074–2101. MR3718162 https://doi.org/10.1214/16-AOS1519

[26] Han, X., Pan, G. and Yang, Q. (2018). A unified matrix model including both CCA and F matrices in multivariate analysis: The largest eigenvalue and its applications. *Bernoulli* **24** 3447–3468. MR3788178 https://doi.org/10.3150/17-BEJ965

[27] Han, X., Pan, G. and Zhang, B. (2016). The Tracy-Widom law for the largest eigenvalue of F type matrices. *Ann. Statist.* **44** 1564–1592. MR3519933 https://doi.org/10.1214/15-AOS1427

[28] Hotelling, H. (1936). Relations between two sets of variates. *Biometrika* **28** 321–377.

[29] Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961 https://doi.org/10.1214/aos/1009210544

[30] Johnstone, I.M. (2008). Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy-Widom limits and rates of convergence. *Ann. Statist.* **36** 2638–2716. MR2485010 https://doi.org/10.1214/08-AOS605

[31] Johnstone, I.M. and Onatski, A. (2020). Testing in high-dimensional spiked models. *Ann. Statist.* **48** 1231–1254. MR4124321 https://doi.org/10.1214/18-AOS1697

[32] Knowles, A. and Yin, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1750. MR3103909 https://doi.org/10.1002/cpa.21450

[33] Knowles, A. and Yin, J. (2014). The outliers of a deformed Wigner matrix. *Ann. Probab.* **42** 1980–2031. MR3262497 https://doi.org/10.1214/13-AOP855

[34] Knowles, A. and Yin, J. (2017). Anisotropic local laws for random matrices. *Probab. Theory Related Fields* **169** 257–352. MR3704770 https://doi.org/10.1007/s00440-016-0730-4
[35] Oda, R., Yanagihara, H. and Fujikoshi, Y. (2019). Asymptotic null and non-null distributions of test statistics for redundancy in high-dimensional canonical correlation analysis. *Random Matrices Theory Appl.* **8** 1950001, 26. MR3892805 https://doi.org/10.1142/S2010326319500011
[36] Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865
[37] Péché, S. (2006). The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* **134** 127–173. MR2221787 https://doi.org/10.1007/s00440-005-0466-z
[38] Tracy, C.A. and Widom, H. (1994). Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159** 151–174. MR1257246
[39] Tracy, C.A. and Widom, H. (1996). On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177** 727–754. MR1385083
[40] Wachter, K.W. (1980). The limiting empirical measure of multiple discriminant ratios. *Ann. Statist.* **8** 937–957. MR0585695
[41] Wang, Q. and Yao, J. (2017). Extreme eigenvalues of large-dimensional spiked Fisher matrices with application. *Ann. Statist.* **45** 415–460. MR3611497 https://doi.org/10.1214/16-AOS1463
[42] Xi, H., Yang, F. and Yin, J. (2017). Local circular law for the product of a deterministic matrix with a random matrix. *Electron. J. Probab.* **22** Paper No. 60, 77. MR3683369 https://doi.org/10.1214/17-EJP76
[43] Yang, F. (2019). Edge universality of separable covariance matrices. *Electron. J. Probab.* **24** Paper No. 123, 57. MR4029426 https://doi.org/10.1214/19-ejp381
[44] Ma, Z. and Yang, F. (2023). Supplement to "Sample canonical correlation coefficients of high-dimensional random vectors with finite rank correlations." https://doi.org/10.3150/22-BEJ1525SUPP
[45] Yang, F. (2022). Limiting distribution of the sample canonical correlation coefficients of high-dimensional random vectors. *Electron. J. Probab.* **27**. MR4455877 https://doi.org/10.1214/22-ejp814
[46] Yang, F. (2022). Sample canonical correlation coefficients of high-dimensional random vectors: Local law and Tracy-Widom limit. *Random Matrices Theory Appl.* **11** Paper No. 2250007, 70. MR4391830 https://doi.org/10.1142/S2010326322500071
[47] Yang, F., Liu, S., Dobriban, E. and Woodruff, D.P. (2021). How to reduce dimension with PCA and random projections? *IEEE Trans. Inf. Theory* **67** 8154–8189. MR4346082 https://doi.org/10.1109/tit.2021.3112821
[48] Yang, Y. and Pan, G. (2012). The convergence of the empirical distribution of canonical correlation coefficients. *Electron. J. Probab.* **17** no. 64, 13. MR2968671 https://doi.org/10.1214/EJP.v17-2239
[49] Yang, Y. and Pan, G. (2015). Independence test for high dimensional data based on regularized canonical correlation coefficients. *Ann. Statist.* **43** 467–500. MR3316187 https://doi.org/10.1214/14-AOS1284

Bernoulli **29**(3), 2023, 1933–1958
https://doi.org/10.3150/22-BEJ1526

Ergodicity of supercritical SDEs driven by α-stable processes and heavy-tailed sampling

XIAOLONG ZHANG[1,a] and XICHENG ZHANG[1,2,b]

[1] *School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China,*
[a] *zhangxl.math@whu.edu.cn*
[2] *School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China,*
[b] *XichengZhang@gmail.com*

Let $\alpha \in (0,2)$ and $d \in \mathbb{N}$. We consider the stochastic differential equation (SDE) driven by an α-stable process

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_{t-})\mathrm{d}L_t^\alpha, \quad X_0 = x \in \mathbb{R}^d,$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are locally γ-Hölder continuous with $\gamma \in (0 \vee (1-\alpha)^+, 1]$, and L_t^α is a d-dimensional symmetric rotationally invariant α-stable process. Under certain dissipative and non-degenerate assumptions on b and σ, we show the V-uniformly exponential ergodicity for the semigroup P_t associated with $\{X_t(x), t \geq 0\}$. Our proofs are mainly based on the heat kernel estimates recently established in (*J. Éc. Polytech. Math.* **9** (2022) 537–579) to demonstrate the strong Feller property and irreducibility of P_t. Interestingly, when α tends to zero, the diffusion coefficient σ can increase faster than the drift b. As an application, we put forward a new heavy-tailed sampling scheme.

Keywords: α-stable processes; ergodicity; heavy-tailed distribution; irreducibility; strong Feller property

References

[1] 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

[2] Arnold, B.C. (2015). *Pareto Distributions*, 2nd ed. *Monographs on Statistics and Applied Probability* **140**. Boca Raton, FL: CRC Press. MR3618736

[3] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9

[4] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR0898871 https://doi.org/10.1017/CBO9780511721434

[5] Bubeck, S., Eldan, R. and Lehec, J. (2018). Sampling from a log-concave distribution with projected Langevin Monte Carlo. *Discrete Comput. Geom.* **59** 757–783. MR3802303 https://doi.org/10.1007/s00454-018-9992-1

[6] Carmona, P. (2007). Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths. *Stochastic Process. Appl.* **117** 1076–1092. MR2340880 https://doi.org/10.1016/j.spa.2006.12.003

[7] Cerrai, S. (2001). *Second Order PDE's in Finite and Infinite Dimension. A Probabilistic Approach*. *Lecture Notes in Math.* **1762**. Berlin: Springer. MR1840644 https://doi.org/10.1007/b80743

[8] Chen, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. *Probability and Its Applications (New York)*. London: Springer. MR2105651

[9] Chen, Z.-Q., Zhang, X. and Zhao, G. (2021). Supercritical SDEs driven by multiplicative stable-like Lévy processes. *Trans. Amer. Math. Soc.* **374** 7621–7655. MR4328678 https://doi.org/10.1090/tran/8343

[10] Cheng, X., Chatterji, N.S., Abbasi-Yadkori, Y., Bartlett, P.L. and Jordan, M.I. (2020). Sharp convergence rates for Langevin dynamics in the nonconvex setting. Available at arXiv:1805.01648.

[11] 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

[12] Da Prato, G. and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions*. *Encyclopedia of Mathematics and Its Applications* **44**. Cambridge: Cambridge Univ. Press. MR1207136 https://doi.org/10.1017/CBO9780511666223

[13] Dalalyan, A.S. and Karagulyan, A. (2019). User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient. *Stochastic Process. Appl.* **129** 5278–5311. MR4025705 https://doi.org/10.1016/j.spa.2019.02.016

[14] Dalalyan, A.S. and Tsybakov, A.B. (2012). Sparse regression learning by aggregation and Langevin Monte-Carlo. *J. Comput. System Sci.* **78** 1423–1443. MR2926142 https://doi.org/10.1016/j.jcss.2011.12.023

[15] Down, D., Meyn, S.P. and Tweedie, R.L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. MR1379163

[16] Durrett, R. (2010). *Probability: Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics* **31**. Cambridge: Cambridge Univ. Press. MR2722836 https://doi.org/10.1017/CBO9780511779398

[17] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. *Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613 https://doi.org/10.1007/978-3-642-33483-2

[18] 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

[19] Foss, S., Korshunov, D. and Zachary, S. (2011). *An Introduction to Heavy-Tailed and Subexponential Distributions*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer.

[20] Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 123–214. MR2814492 https://doi.org/10.1111/j.1467-9868.2010.00765.x

[21] Goldys, B. and Maslowski, B. (2006). Exponential ergodicity for stochastic reaction-diffusion equations. In *Stochastic Partial Differential Equations and Applications—VII*. *Lect. Notes Pure Appl. Math.* **245** 115–131. Boca Raton, FL: CRC Press/CRC. MR2227225 https://doi.org/10.1201/9781420028720.ch12

[22] Gürbüzbalaban, M., Simsekli, U. and Zhu, L. (2021). The heavy-tail phenomenon in SGD. *International Conference on Machine Learning*. *PMLR* 3964–3975.

[23] Huang, L.-J., Majka, M.B. and Wang, J. (2021). Approximation of heavy-tailed distributions via stable-driven SDEs. *Bernoulli* **27** 2040–2068. MR4278802 https://doi.org/10.3150/20-bej1300

[24] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

[25] Kulik, A.M. (2009). Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.* **119** 602–632. MR2494006 https://doi.org/10.1016/j.spa.2008.02.006

[26] Kwaśnicki, M. (2017). Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.* **20** 7–51. MR3613319 https://doi.org/10.1515/fca-2017-0002

[27] Liang, M., Majka, M.B. and Wang, J. (2021). Exponential ergodicity for SDEs and McKean–Vlasov processes with Lévy noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1665–1701. MR4291453 https://doi.org/10.1214/20-aihp1123

[28] Liang, M. and Wang, J. (2020). Gradient estimates and ergodicity for SDEs driven by multiplicative Lévy noises via coupling. *Stochastic Process. Appl.* **130** 3053–3094. MR4080736 https://doi.org/10.1016/j.spa.2019.09.001

[29] Luo, D. and Wang, J. (2019). Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stochastic Process. Appl.* **129** 3129–3173. MR3985558 https://doi.org/10.1016/j.spa.2018.09.003

[30] Majka, M.B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stochastic Process. Appl.* **127** 4083–4125. MR3718107 https://doi.org/10.1016/j.spa.2017.03.020
[31] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010
[32] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266 https://doi.org/10.1016/S0304-4149(02)00150-3
[33] Menozzi, S. and Zhang, X. (2022). Heat kernel of supercritical nonlocal operators with unbounded drifts. *J. Éc. Polytech. Math.* **9** 537–579. MR4394312 https://doi.org/10.5802/jep.189
[34] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2509253 https://doi.org/10.1017/CBO9780511626630
[35] Nguyen, T., Simsekli, U. and Richard, D. (2019). Non-asymptotic analysis of fractional Langevin Monte Carlo for non-convex optimization. *International Conference on Machine Learning. PMLR.* **97** 4810–4819.
[36] 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
[37] Roberts, G.O. and Tweedie, R.L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* **2** 341–363. MR1440273 https://doi.org/10.2307/3318418
[38] Simsekli, U. (2017). Fractional Langevin Monte Carlo: Exploring Lévy driven stochastic differential equations for Markov chain Monte Carlo. In *International Conference on Machine Learning. PMLR* 3200–3209.
[39] Simsekli, U., Zhu, L., Teh, Y.W. and Gürbüzbalaban, M. (2020). Fractional underdamped Langevin dynamics: Retargeting sgd with momentum under heavy-tailed gradient noise. *International Conference on Machine Learning. PMLR.* 8970–8980.
[40] Stein, E.M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series, No. 30*. Princeton, NJ: Princeton Univ. Press. MR0290095
[41] Villani, C. (2008). *Optimal Transport, Old and New*. Tokyo: Springer.
[42] Wang, F. (2006). *Functional Inequalities Markov Semigroups and Spectral Theory*. Amsterdam: Elsevier.
[43] Wang, J. (2016). L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli* **22** 1598–1616. MR3474827 https://doi.org/10.3150/15-BEJ705
[44] Xie, L. and Zhang, X. (2020). Ergodicity of stochastic differential equations with jumps and singular coefficients. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 175–229. MR4058986 https://doi.org/10.1214/19-AIHP959
[45] Ye, N. and Zhu, Z. (2018). Stochastic fractional Hamiltonian Monte Carlo. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence* 3019–3025.

Bernoulli **29**(3), 2023, 1959–1982
https://doi.org/10.3150/22-BEJ1527

Optimal false discovery control of minimax estimators

QIFAN SONG[1,a] and GUANG CHENG[2,b]

[1]*Purdue University, West Lafayette, USA,* [a]*qfsong@purdue.edu*
[2]*University of California, Los Angeles, USA,* [b]*guangcheng@ucla.edu*

Two major research tasks lie at the heart of high dimensional data analysis: accurate parameter estimation and correct support recovery. The existing literature mostly aims for either the best parameter estimation or the best model selection result, however little has been done to understand the potential interaction between the estimation precision and the selection behavior. In this work, our minimax result shows that an estimator's performance of type I error control directly links with its L_2 estimation error rate, and reveals a trade-off phenomenon between the rate of convergence and the false discovery control: to achieve better accuracy, one risks yielding more false discoveries. In particular, we characterize the false discovery control behavior of rate optimal and rate suboptimal estimators under different sparsity regimes, and discover a rigid dichotomy between these two estimators under near-linear and linear sparsity settings. In addition, this work provides a rigorous explanation to the incompatibility phenomenon between selection consistency and rate minimaxity which has been frequently observed in the high dimensional literature.

Keywords: False discovery control; high dimension analysis; rate minimaxity; selection consistency

References

[1] Abramovich, F., Benjamini, Y., Donoho, D.L. and Johnstone, I.M. (2006). Adapting to unknown sparsity by controlling the false discovery rate. *Ann. Statist.* **34** 584–653. MR2281879 https://doi.org/10.1214/009053606000000074

[2] Arias-Castro, E. and Chen, S. (2017). Distribution-free multiple testing. *Electron. J. Stat.* **11** 1983–2001. MR3651021 https://doi.org/10.1214/17-EJS1277

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

[4] Bhattacharya, A., Pati, D., Pillai, N.S. and Dunson, D.B. (2015). Dirichlet-Laplace priors for optimal shrinkage. *J. Amer. Statist. Assoc.* **110** 1479–1490. MR3449048 https://doi.org/10.1080/01621459.2014.960967

[5] Birgé, L. and Massart, P. (2001). Gaussian model selection. *J. Eur. Math. Soc. (JEMS)* **3** 203–268. MR1848946 https://doi.org/10.1007/s100970100031

[6] Bogdan, M., Chakrabarti, A., Frommlet, F. and Ghosh, J.K. (2011). Asymptotic Bayes-optimality under sparsity of some multiple testing procedures. *Ann. Statist.* **39** 1551–1579. MR2850212 https://doi.org/10.1214/10-AOS869

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

[8] Butucea, C., Ndaoud, M., Stepanova, N.A. and Tsybakov, A.B. (2018). Variable selection with Hamming loss. *Ann. Statist.* **46** 1837–1875. MR3845003 https://doi.org/10.1214/17-AOS1572

[9] Castillo, I. and Roquain, É. (2020). On spike and slab empirical Bayes multiple testing. *Ann. Statist.* **48** 2548–2574. MR4152112 https://doi.org/10.1214/19-AOS1897

[10] Castillo, I. and van der Vaart, A. (2012). Needles and straw in a haystack: Posterior concentration for possibly sparse sequences. *Ann. Statist.* **40** 2069–2101. MR3059077 https://doi.org/10.1214/12-AOS1029

[11] Donoho, D.L. and Johnstone, I.M. (1994). Minimax risk over l_p-balls for l_q-error. *Probab. Theory Related Fields* **99** 277–303. MR1278886 https://doi.org/10.1007/BF01199026

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

[13] Foster, D.P. and Stine, R.A. (1999). Local asymptotic coding and the minimum description length. *IEEE Trans. Inf. Theory* **45** 1289–1293. MR1686271 https://doi.org/10.1109/18.761287

[14] Gao, C., van der Vaart, A.W. and Zhou, H.H. (2020). A general framework for Bayes structured linear models. *Ann. Statist.* **48** 2848–2878. MR4152123 https://doi.org/10.1214/19-AOS1909

[15] George, E.I. and Foster, D.P. (2000). Calibration and empirical Bayes variable selection. *Biometrika* **87** 731–747. MR1813972 https://doi.org/10.1093/biomet/87.4.731

[16] Johnstone, I.M. and Silverman, B.W. (2004). Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences. *Ann. Statist.* **32** 1594–1649. MR2089135 https://doi.org/10.1214/009053604000000030

[17] Koda, S., Onda, Y., Matsui, H., Takahagi, K., Uehara-Yamaguchi, Y., Shimizu, M., Inoue, K., Yoshida, T., Sakurai, T., Honda, H., Eguchi, S., Nishii, R. and Mochida, K. (2017). Diurnal transcriptome and gene network represented through sparse modeling in brachypodium distachyon. *Front. Plant Sci.* **8** 2055.

[18] Kos, M. and Bogdan, M. (2020). On the asymptotic properties of SLOPE. *Sankhya A* **82** 499–532. MR4140775 https://doi.org/10.1007/s13171-020-00212-5

[19] Ndaoud, M. (2019). Interplay of minimax estimation and minimax support recovery under sparsity. In *Algorithmic Learning Theory 2019*. *Proc. Mach. Learn. Res. (PMLR)* **98** 647–668. Proceedings of Machine Learning Research PMLR. MR3932863

[20] Raskutti, G., Wainwright, M.J. and Yu, B. (2011). Minimax rates of estimation for high-dimensional linear regression over ℓ_q-balls. *IEEE Trans. Inf. Theory* **57** 6976–6994. MR2882274 https://doi.org/10.1109/TIT.2011.2165799

[21] Ročková, V. (2018). Bayesian estimation of sparse signals with a continuous spike-and-slab prior. *Ann. Statist.* **46** 401–437. MR3766957 https://doi.org/10.1214/17-AOS1554

[22] Salomond, J.-B. (2017). Risk quantification for the thresholding rule for multiple testing using Gaussian scale mixtures. Preprint. Available at arXiv:1711.08705.

[23] Song, Q. (2020). Bayesian shrinkage towards sharp minimaxity. *Electron. J. Stat.* **14** 2714–2741. MR4125855 https://doi.org/10.1214/20-EJS1732

[24] Song, Q. and Cheng, G. (2023). Supplement to "Optimal false discovery control of minimax estimators." https://doi.org/10.3150/22-BEJ1527SUPP

[25] Song, Q. and Liang, F. (2015). High-dimensional variable selection with reciprocal L_1-regularization. *J. Amer. Statist. Assoc.* **110** 1607–1620. MR3449058 https://doi.org/10.1080/01621459.2014.984812

[26] Song, Q. and Liang, F. (2017). Nearly optimal Bayesian shrinkage for high dimensional regression. Available at arXiv:1712.08964.

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

[28] Su, W. and Candès, E. (2016). SLOPE is adaptive to unknown sparsity and asymptotically minimax. *Ann. Statist.* **44** 1038–1068. MR3485953 https://doi.org/10.1214/15-AOS1397

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

[30] Tibshirani, R. and Knight, K. (1999). The covariance inflation criterion for adaptive model selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 529–546. MR1707859 https://doi.org/10.1111/1467-9868.00191

[31] van der Pas, S.L., Kleijn, B.J.K. and van der Vaart, A.W. (2014). The horseshoe estimator: Posterior concentration around nearly black vectors. *Electron. J. Stat.* **8** 2585–2618. MR3285877 https://doi.org/10.1214/14-EJS962

[32] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170

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

[34] Weinstein, A., Barber, R. and Candes, E. (2017). A power and prediction analysis for knockoffs with lasso statistics. Available at arXiv:1712.06465.

[35] Wilkins, O., Hafemeister, C., Plessis, A., Holloway-Phillips, M.-M., Pham, G.M., Nicotra, A.B., Gregorio, G.B., Jagadish, S.K., Septiningsih, E.M., Bonneau, R. and Purugganan, M. (2016). EGRINs (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. *The Plant Cell* **28** 2365–2384.

[36] Wu, Z. and Zhou, H.H. (2013). Model selection and sharp asymptotic minimaxity. *Probab. Theory Related Fields* **156** 165–191. MR3055256 https://doi.org/10.1007/s00440-012-0424-5

[37] Ye, F. and Zhang, C.-H. (2010). Rate minimaxity of the Lasso and Dantzig selector for the ℓ_q loss in ℓ_r balls. *J. Mach. Learn. Res.* **11** 3519–3540. MR2756192

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

Bernoulli **29**(3), 2023, 1983–2007
https://doi.org/10.3150/22-BEJ1528

Rates of convergence for the number of zeros of random trigonometric polynomials

LAURE COUTIN[1,a] and LILIANA PERALTA[2,b]

[1] *Institut de mathématiques de Toulouse, Université Paul Sabatier, 118, route de Narbonne F-31062 Toulouse cedex 9, France,* [a] *laure.coutin@math.univ-toulouse.fr*
[2] *Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior, C.U., C.P. 04510 CDMX, Mexico,* [b] *lylyaanaa@ciencias.unam.mx*

In this paper, we quantify the rate of convergence between the distribution of number of zeros of random trigonometric polynomials (RTP) with i.i.d. centered random coefficients and the number of zeros of a stationary centered Gaussian process G, whose covariance function is given by the sinc function. First, we find the convergence of the RTP towards G in the Wasserstein−1 distance, which in turn is a consequence of Donsker Theorem. Then, we use this result to derive the rate of convergence between their respective number of zeros. Since the number of real zeros of the RTP is not a continuous function, we use the Kac-Rice formula to express it as the limit of an integral and, in this way, we approximate it by locally Lipschitz continuous functions.

Keywords: Random trigonometric polynomials; Wassertein distance; Donsker Theorem; Stein method

References

[1] Angst, J., Pham, V.-H. and Poly, G. (2018). Universality of the nodal length of bivariate random trigonometric polynomials. *Trans. Amer. Math. Soc.* **370** 8331–8357. MR3864378 https://doi.org/10.1090/tran/7255
[2] Azaïs, J.-M., Dalmao, F., León, J.R., Nourdin, I. and Poly, G. (2018). Local universality of the number of zeros of random trigonometric polynomials with continuous coefficients. Preprint.
[3] Azaïs, J.-M. and León, J.R. (2013). CLT for crossings of random trigonometric polynomials. *Electron. J. Probab.* **18** no. 68, 17. MR3084654 https://doi.org/10.1214/EJP.v18-2403
[4] Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: Wiley. MR2478201 https://doi.org/10.1002/9780470434642
[5] Barbour, A.D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659 https://doi.org/10.1007/BF01197887
[6] Bogomolny, E., Bohigas, O. and Leboeuf, P. (1996). Quantum chaotic dynamics and random polynomials. *J. Stat. Phys.* **85** 639–679. MR1418808 https://doi.org/10.1007/BF02199359
[7] Chernozhukov, V., Chetverikov, D. and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **45** 2309–2352. MR3693963 https://doi.org/10.1214/16-AOP1113
[8] Coutin, L. and Decreusefond, L. (2013). Stein's method for Brownian approximations. *Commun. Stoch. Anal.* **7** 349–372. MR3167403 https://doi.org/10.31390/cosa.7.3.01
[9] Coutin, L. and Decreusefond, L. (2020). Donsker's theorem in Wasserstein-1 distance. *Electron. Commun. Probab.* **25** Paper No. 27, 13. MR4089734 https://doi.org/10.1214/20-ecp308
[10] Coutin, L. and Peralta, L. (2023). Supplement to "Rates of convergence for the number of zeros of random trigonometric polynomials." https://doi.org/10.3150/22-BEJ1528SUPP
[11] Dunnage, J.E.A. (1966). The number of real zeros of a random trigonometric polynomial. *Proc. Lond. Math. Soc. (3)* **16** 53–84. MR0192532 https://doi.org/10.1112/plms/s3-16.1.53
[12] Erdös, P. and Offord, A.C. (1956). On the number of real roots of a random algebraic equation. *Proc. Lond. Math. Soc. (3)* **6** 139–160. MR0073870 https://doi.org/10.1112/plms/s3-6.1.139
[13] Farahmand, K. (1987). On the number of real zeros of a random trigonometric polynomial: Coefficients with nonzero infinite mean. *Stoch. Anal. Appl.* **5** 379–386. MR0912864 https://doi.org/10.1080/07362998708809125

[14] Farahmand, K. (1995). Level crossings of a random trigonometric polynomial with dependent coefficients. *J. Aust. Math. Soc. A* **58** 39–46. MR1313884

[15] Farahmand, K. (1997). On the variance of the number of real zeros of a random trigonometric polynomial. *J. Appl. Math. Stoch. Anal.* **10** 57–66. MR1437951 https://doi.org/10.1155/S1048953397000051

[16] Farahmand, K. and Sambandham, M. (1997). On the expected number of real zeros of random trigonometric polynomials. *Analysis (Oxford)* **17** 345–353. MR1600362 https://doi.org/10.1524/anly.1997.17.4.345

[17] Flasche, H. (2017). Expected number of real roots of random trigonometric polynomials. *Stochastic Process. Appl.* **127** 3928–3942. MR3718101 https://doi.org/10.1016/j.spa.2017.03.018

[18] Iksanov, A., Kabluchko, Z. and Marynych, A. (2016). Local universality for real roots of random trigonometric polynomials. *Electron. J. Probab.* **21** Paper No. 63, 19. MR3563891 https://doi.org/10.1214/16-EJP9

[19] Jacod, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. *Lecture Notes in Math.* **714**. Berlin: Springer. MR0542115

[20] Kac, M. (1949). On the average number of real roots of a random algebraic equation. II. *Proc. Lond. Math. Soc. (2)* **50** 390–408. MR0030713 https://doi.org/10.1112/plms/s2-50.5.390

[21] Littlewood, J.E. and Offord, A.C. (1938). On the Number of Real Roots of a Random Algebraic Equation. *J. Lond. Math. Soc.* **13** 288–295. MR1574980 https://doi.org/10.1112/jlms/s1-13.4.288

[22] Nguyen, O. and Vu, V. (2022). Roots of random functions: A framework for local universality. *Amer. J. Math.* **144** 1–74. MR4367414 https://doi.org/10.1353/ajm.2022.0000

[23] Villani, C. (2009). *Optimal Transport: Old and New*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9

Bernoulli **29**(3), 2023, 2008–2034
https://doi.org/10.3150/22-BEJ1530

Bounds for the chi-square approximation of Friedman's statistic by Stein's method

ROBERT E. GAUNT[1,a] and GESINE REINERT[2,b]

[1]*Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK,* [a]*robert.gaunt@manchester.ac.uk*
[2]*Department of Statistics, University of Oxford, 24-29 St. Giles', Oxford OX1 3LB, UK,* [b]*reinert@stats.ox.ac.uk*

Friedman's chi-square test is a non-parametric statistical test for r treatments across n trials to assess the null hypothesis that there is no treatment effect. We use Stein's method with an exchangeable pair coupling to derive a bound on the distance between the distribution of Friedman's statistic and its limiting chi-square distribution, measured using smooth test functions. Our bound is of the optimal order n^{-1}, and also has an optimal dependence on the parameter r, in that the bound tends to zero if and only if $r/n \to 0$. From this bound, we deduce a Kolmogorov distance bound that decays to zero under the weaker condition $r^{1/2}/n \to 0$.

Keywords: Stein's method; Friedman's statistic; chi-square approximation; rate of convergence; exchangeable pair

References

[1] Anastasiou, A. and Reinert, G. (2020). Bounds for the asymptotic distribution of the likelihood ratio. *Ann. Appl. Probab.* **30** 608–643. MR4108117 https://doi.org/10.1214/19-AAP1510
[2] Gaunt, R.E. and Reinert, G. (2023). Supplement to "Bounds for the chi-square approximation of Friedman's statistic by Stein's method." https://doi.org/10.3150/22-BEJ1530SUPP
[3] Asylbekov, Zh.A., Zubov, V.N. and Ul'yanov, V.V. (2011). On the approximation of some statistics of goodness-of-fit tests for the case of discrete three-dimensional data. *Sibirsk. Mat. Zh.* **52** 571–584. MR2883210 https://doi.org/10.1134/S0037446611040021
[4] Azmoodeh, E., Eichelsbacher, P. and Knichel, L. (2020). Optimal gamma approximation on Wiener space. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** 101–132. MR4057185 https://doi.org/10.30757/alea.v17-05
[5] Barbour, A.D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659 https://doi.org/10.1007/BF01197887
[6] Barbour, A.D. and Hall, P. (1985). On bounds to the rate of convergence in the central limit theorem. *Bull. Lond. Math. Soc.* **17** 151–156. MR0806241 https://doi.org/10.1112/blms/17.2.151
[7] Braverman, A., Dai, J.G. and Fang, X. (2020). High order steady-state diffusion approximations. arXiv:2012.02824.
[8] Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. MR2435859 https://doi.org/10.1214/07-AOP370
[9] Chatterjee, S., Fulman, J. and Röllin, A. (2011). Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** 197–223. MR2802856
[10] Chatterjee, S. and Meckes, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. MR2453473
[11] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications (New York)*. Heidelberg: Springer. MR2732624 https://doi.org/10.1007/978-3-642-15007-4
[12] Diaconis, P. and Zabell, S. (1991). Closed form summation for classical distributions: Variations on a theme of de Moivre. *Statist. Sci.* **6** 284–302. MR1144242
[13] Dicker, L.H. and Erdogdu, M.A. (2017). Flexible results for quadratic forms with applications to variance components estimation. *Ann. Statist.* **45** 386–414. MR3611496 https://doi.org/10.1214/16-AOS1456

1350-7265

[14] Döbler, C. and Peccati, G. (2018). The gamma Stein equation and noncentral de Jong theorems. *Bernoulli* **24** 3384–3421. MR3788176 https://doi.org/10.3150/17-BEJ963

[15] Fathi, M. (2021). Higher-order Stein kernels for Gaussian approximation. *Studia Math.* **256** 241–258. MR4172696 https://doi.org/10.4064/sm190415-28-10

[16] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–701.

[17] Gaunt, R.E. (2013). *Rates of Convergence of Variance-Gamma Approximations via Stein's Method*. DPhil thesis, University of Oxford.

[18] Gaunt, R.E. (2020). Stein's method for functions of multivariate normal random variables. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1484–1513. MR4076792 https://doi.org/10.1214/19-AIHP1011

[19] Gaunt, R.E. (2022). Bounds for the chi-square approximation of the power divergence family of statistics. *J. Appl. Probab.* To appear.

[20] Gaunt, R.E., Pickett, A.M. and Reinert, G. (2017). Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann. Appl. Probab.* **27** 720–756. MR3655852 https://doi.org/10.1214/16-AAP1213

[21] Gaunt, R.E. and Reinert, G. (2021). Bounds for the chi-square approximation of Friedman's statistic by Stein's method Technical report.

[22] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. MR1484792 https://doi.org/10.1214/aoap/1043862419

[23] Goldstein, L. and Reinert, G. (2005). Zero biasing in one and higher dimensions, and applications. In *Stein's Method and Applications*. *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 1–18. Singapore: Singapore Univ. Press. MR2201883 https://doi.org/10.1142/9789812567673_0001

[24] Goldstein, L. and Rinott, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17. MR1371949 https://doi.org/10.1017/s0021900200103675

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

[26] Götze, F. and Ulyanov, V.V. (2003). *Asymptotic Distribution of χ^2-Type Statistics*. *Research Group Spectral Analysis, Asymptotic Distributions and Stochastic Dynamics*. Bielefeld: Bielefeld Univ. Preprint 03–033.

[27] Jensen, D.R. (1977). On approximating the distributions of Friedman's χ_r^2 and related statistics. *Metrika* **24** 75–85. MR0443173 https://doi.org/10.1007/BF01893394

[28] Lefèvre, C. and Utev, S. (2003). Exact norms of a Stein-type operator and associated stochastic orderings. *Probab. Theory Related Fields* **127** 353–366. MR2018920 https://doi.org/10.1007/s00440-003-0288-9

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

[30] Loh, W.-L. (1992). Stein's method and multinomial approximation. *Ann. Appl. Probab.* **2** 536–554. MR1177898

[31] Luk, H.M. (1994). *Stein's Method for the Gamma Distribution and Related Statistical Applications*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–University of Southern California. MR2693204

[32] 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

[33] Mijoule, G., Reinert, G. and Swan, Y. (2021). Stein's density method for multivariate continuous distributions. arXiv:2101.05079.

[34] Nourdin, I. and Peccati, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. MR2520122 https://doi.org/10.1007/s00440-008-0162-x

[35] Peköz, E.A. and Röllin, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** 587–608. MR2789507 https://doi.org/10.1214/10-AOP559

[36] Pickett, A. (2004). Rates of Convergence of χ^2 Approximations via Stein's Method. DPhil thesis, Univ. Oxford.

[37] Reinert, G. and Röllin, A. (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** 2150–2173. MR2573554 https://doi.org/10.1214/09-AOP467

[38] Shevtsova, I.G. (2014). On absolute constants in inequalities of Berry-Esseen type. *Dokl. Akad. Nauk* **456** 650–654. MR3287912 https://doi.org/10.1134/s1064562414030338

[39] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory* 583–602. Berkeley, CA: Univ. California Press. MR0402873

[40] Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **7**. Hayward, CA: IMS. MR0882007

[41] Ulyanov, V.V. and Zubov, V.N. (2009). Refinement on the convergence of one family of goodness-of-fit statistics to chi-squared distribution. *Hiroshima Math. J.* **39** 133–161. MR2499200

[42] Yarnold, J.K. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. *Ann. Math. Stat.* **43** 1566–1580. MR0372967 https://doi.org/10.1214/aoms/1177692389

Bernoulli **29**(3), 2023, 2035–2061
https://doi.org/10.3150/22-BEJ1531

Parameter estimation for semilinear SPDEs from local measurements

RANDOLF ALTMEYER[1,a], IGOR CIALENCO[2,c] and GREGOR PASEMANN[1,b]

[1]*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany,* [a]*altmeyrx@math.hu-berlin.de,* [b]*gregor.pasemann@hu-berlin.de*
[2]*Department of Applied Mathematics, Illinois Institute of Technology, 10 W 32nd Str, Building REC, Room 220, Chicago, IL 60616, USA,* [c]*cialenco@iit.edu*

This work contributes to the limited literature on estimating the diffusivity or drift coefficient of nonlinear SPDEs driven by additive noise. Assuming that the solution is measured locally in space and over a finite time interval, we show that the augmented maximum likelihood estimator introduced in (*Ann. Appl. Probab.* **31** (2021) 1–38) for linear SPDEs remains rate-optimal when applied to a large class of semilinear SPDEs. The obtained abstract results are applied to several important classes of SPDEs, including stochastic reaction-diffusion equations. Moreover, we also study the stochastic Burgers equation, as an example with first order nonlinearity, which is a borderline case of the general results. The optimal statistical results are obtained through a precise control of the spatial regularity of the solution and by using higher order fractional L^p-Sobolev type spaces. We conclude with numerical examples that validate the theoretical results.

Keywords: Stochastic partial differential equations; semilinear SPDEs; augmented MLE; stochastic Burgers; stochastic reaction-diffusion; optimal regularity; inference; drift estimation; viscosity estimation; central limit theorem; local measurements

References

[1] Adams, R.A. (1975). *Sobolev Spaces*. San Diego: Academic Press.
[2] Alonso, S., Stange, M. and Beta, C. (2018). Modeling random crawling, membrane deformation and intracellular polarity of motile amoeboid cells. *PLoS ONE* **13** 1–22.
[3] Altmeyer, R., Bretschneider, T., Janák, J. and Reiß, M. (2022). Parameter estimation in an SPDE model for cell repolarization. *SIAM/ASA J. Uncertain. Quantificat.* **10** 179–199. MR4376300 https://doi.org/10.1137/20M1373347
[4] Altmeyer, R., Cialenco, I. and Pasemann, G. (2023). Supplement to “Parameter estimation for semilinear SPDEs from local measurements.” https://doi.org/10.3150/22-BEJ1531SUPP
[5] Altmeyer, R. and Reiss, M. (2021). Nonparametric estimation for linear SPDEs from local measurements. *Ann. Appl. Probab.* **31** 1–38. MR4254472 https://doi.org/10.1214/20-aap1581
[6] Bibinger, M. and Trabs, M. (2019). On central limit theorems for power variations of the solution to the stochastic heat equation. In *Stochastic Models, Statistics and Their Applications*. *Springer Proc. Math. Stat.* **294** 69–84. Cham: Springer. MR4043170
[7] Bibinger, M. and Trabs, M. (2020). Volatility estimation for stochastic PDEs using high-frequency observations. *Stochastic Process. Appl.* **130** 3005–3052. MR4080737 https://doi.org/10.1016/j.spa.2019.09.002
[8] Cahn, J. and Allen, S. (1977). A microscopic theory for domain wall motion and its experimental verification in fe-al alloy domain growth kinetics. *J. Phys. Colloq.* **38** 51–54.
[9] Chong, C. (2019). High-frequency analysis of parabolic stochastic PDEs with multiplicative noise: Part I. Preprint. arXiv:1908.04145.
[10] Chong, C. (2020). High-frequency analysis of parabolic stochastic PDEs. *Ann. Statist.* **48** 1143–1167. MR4102691 https://doi.org/10.1214/19-AOS1841
[11] Cialenco, I. (2018). Statistical inference for SPDEs: An overview. *Stat. Inference Stoch. Process.* **21** 309–329. MR3824970 https://doi.org/10.1007/s11203-018-9177-9

1350-7265

[12] Cialenco, I., Delgado-Vences, F. and Kim, H.-J. (2020). Drift estimation for discretely sampled SPDEs. *Stoch. Partial Differ. Equ. Anal. Comput.* **8** 895–920. MR4174073 https://doi.org/10.1007/s40072-019-00164-4
[13] Cialenco, I. and Glatt-Holtz, N. (2011). Parameter estimation for the stochastically perturbed Navier-Stokes equations. *Stochastic Process. Appl.* **121** 701–724. MR2770904 https://doi.org/10.1016/j.spa.2010.12.007
[14] Cialenco, I. and Huang, Y. (2020). A note on parameter estimation for discretely sampled SPDEs. *Stoch. Dyn.* **20** 2050016. MR4101083 https://doi.org/10.1142/S0219493720500161
[15] Cialenco, I. and Kim, H.-J. (2022). Parameter estimation for discretely sampled stochastic heat equation driven by space-only noise. *Stochastic Process. Appl.* **143** 1–30. MR4332773 https://doi.org/10.1016/j.spa.2021.09.012
[16] Cont, R. (2005). Modeling term structure dynamics: An infinite dimensional approach. *Int. J. Theor. Appl. Finance* **8** 357–380. MR2144706 https://doi.org/10.1142/S0219024905003049
[17] Cotter, C., Crisan, D., Holm, D.D., Pan, W. and Shevchenko, I. (2019). Numerically modeling stochastic Lie transport in fluid dynamics. *Multiscale Model. Simul.* **17** 192–232. MR3904409 https://doi.org/10.1137/18M1167929
[18] Da Prato, G., Debussche, A. and Temam, R. (1994). Stochastic Burgers' equation. *NoDEA Nonlinear Differential Equations Appl.* **1** 389–402. MR1300149 https://doi.org/10.1007/BF01194987
[19] Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge: Cambridge Univ. Press. MR3236753 https://doi.org/10.1017/CBO9781107295513
[20] Debussche, A., de Moor, S. and Hofmanová, M. (2015). A regularity result for quasilinear stochastic partial differential equations of parabolic type. *SIAM J. Math. Anal.* **47** 1590–1614. MR3340199 https://doi.org/10.1137/130950549
[21] Fitzhugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1** 445–466.
[22] Frankignoul, C. (1985). Sst anomalies, planetary waves and rc in the middle rectitudes. *Rev. Geophys.* **23** 357–390.
[23] Hairer, M. and Voss, J. (2011). Approximations to the stochastic Burgers equation. *J. Nonlinear Sci.* **21** 897–920. MR2860933 https://doi.org/10.1007/s00332-011-9104-3
[24] Janson, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge: Cambridge Univ. Press. MR1474726 https://doi.org/10.1017/CBO9780511526169
[25] Kaino, Y. and Uchida, M. (2021). Parametric estimation for a parabolic linear SPDE model based on discrete observations. *J. Statist. Plann. Inference* **211** 190–220. MR4153870 https://doi.org/10.1016/j.jspi.2020.05.004
[26] Liptser, R.S. and Shiryayev, A.N. (1989). *Theory of Martingales*. *Mathematics and Its Applications (Soviet Series)* **49**. Dordrecht: Kluwer Academic. MR1022664 https://doi.org/10.1007/978-94-009-2438-3
[27] Liu, W. and Röckner, M. (2015). *Stochastic Partial Differential Equations: An Introduction*. *Universitext*. Cham: Springer. MR3410409 https://doi.org/10.1007/978-3-319-22354-4
[28] Lockley, R., Ladds, G. and Bretschneider, T. (2015). Image based validation of dynamical models for cell reorientation. *Cytometry, Part A* **87** 471–480.
[29] Lord, G.J., Powell, C.E. and Shardlow, T. (2014). *An Introduction to Computational Stochastic PDEs*. *Cambridge Texts in Applied Mathematics*. New York: Cambridge Univ. Press. MR3308418 https://doi.org/10.1017/CBO9781139017329
[30] Mahdi Khalil, Z. and Tudor, C. (2019). Estimation of the drift parameter for the fractional stochastic heat equation via power variation. *Mod. Stoch. Theory Appl.* **6** 397–417. MR4047392
[31] Nagumo, A., Arimoto, S. and Yoshizawa, S.S. (1962). An active pulse transmission line simulating nerve axon. *Proc. IRE* **50** 2061–2070.
[32] Nüsken, N., Reich, S. and Rozdeba, P.J. (2019). State and parameter estimation from observed signal increments. *Entropy* **21** 505. MR3975241 https://doi.org/10.3390/e21050505
[33] Pasemann, G., Flemming, S., Alonso, S., Beta, C. and Stannat, W. (2021). Diffusivity estimation for activator-inhibitor models: Theory and application to intracellular dynamics of the actin cytoskeleton. *J. Nonlinear Sci.* **31** 59. MR4255684 https://doi.org/10.1007/s00332-021-09714-4
[34] Pasemann, G. and Stannat, W. (2020). Drift estimation for stochastic reaction-diffusion systems. *Electron. J. Stat.* **14** 547–579. MR4056266 https://doi.org/10.1214/19-EJS1665

[35] Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. *Applied Mathematical Sciences* **44**. New York: Springer. MR0710486 https://doi.org/10.1007/978-1-4612-5561-1

[36] Reich, S. and Rozdeba, P. (2020). Posterior contraction rates for non-parametric state and drift estimation. *Found. Data Sci.* **2** 333–349.

[37] Schlögl, F. (1972). Chemical reaction models for non-equilibrium phase transitions. *Z. Phys.* **253** 147–161.

[38] Triebel, H. (1983). *Theory of Function Spaces*. *Monographs in Mathematics* **78**. Basel: Birkhäuser. MR0781540 https://doi.org/10.1007/978-3-0346-0416-1

[39] Tsybakov, A.B. (2008). *Introduction to Nonparametric Estimation*. Berlin: Springer.

[40] Walsh, J.B. (1981). A stochastic model of neural response. *Adv. in Appl. Probab.* **13** 231–281. MR0612203 https://doi.org/10.2307/1426683

[41] Yagi, A. (2010). *Abstract Parabolic Evolution Equations and Their Applications*. *Springer Monographs in Mathematics*. Berlin: Springer. MR2573296 https://doi.org/10.1007/978-3-642-04631-5

[42] Yan, D. (2019). Bayesian Inference for Gaussian Models: Inverse Problems and Evolution Equations Ph.D. thesis, Universiteit Leiden.

Bernoulli **29**(3), 2023, 2062–2083
https://doi.org/10.3150/22-BEJ1532

Nonparametric estimation of locally stationary Hawkes processes

ENNO MAMMEN[a] and MARILENA MÜLLER[b]

Institute of Applied Mathematics, Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany, [a]*mammen@math.uni-heidelberg.de,* [b]*m.mueller@math.uni-heidelberg.de*

In this paper we consider multivariate Hawkes processes with baseline conditional intensities and reproduction functions that depend on time. This defines a class of locally stationary processes. We discuss estimation of the time-dependent baseline intensities and reproduction functions based on a localized criterion. Theory on stationary Hawkes processes is extended to develop asymptotic theory for the estimator in the locally stationary model.

Keywords: Hawkes processes; nonparametric estimation

References

Bacry, E., Dayri, K. and Muzy, J.F. (2012). Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *Eur. Phys. J. B* **85** 157.

Bacry, E., Mastromatteo, I. and Muzy, J.-F. (2015). Hawkes processes in finance. *Mark. Microstruct. Liq.* **01** 1550005.

Bacry, E. and Muzy, J.-F. (2016). First- and second-order statistics characterization of Hawkes processes and non-parametric estimation. *IEEE Trans. Inf. Theory* **62** 2184–2202. MR3480107 https://doi.org/10.1109/TIT.2016.2533397

Bai, F., Chen, F. and Chen, K. (2015). Semiparametric estimation of a self-exciting regression model with an application in recurrent event data analysis. *Statist. Sinica* **25** 1503–1526. MR3409078

Chen, F. and Hall, P. (2013). Inference for a nonstationary self-exciting point process with an application in ultra-high frequency financial data modeling. *J. Appl. Probab.* **50** 1006–1024. MR3161370 https://doi.org/10.1239/jap/1389370096

Chen, F. and Hall, P. (2016). Nonparametric estimation for self-exciting point processes—a parsimonious approach. *J. Comput. Graph. Statist.* **25** 209–224. MR3474044 https://doi.org/10.1080/10618600.2014.1001491

Clinet, S. and Potiron, Y. (2018). Statistical inference for the doubly stochastic self-exciting process. *Bernoulli* **24** 3469–3493. MR3788179 https://doi.org/10.3150/17-BEJ966

Dahlhaus, R. (1996). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62** 139–168. MR1388767 https://doi.org/10.1016/0304-4149(95)00090-9

Daley, D.J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. *Springer Series in Statistics*. New York: Springer. MR0950166

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

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

Fox, E.W., Short, M.B., Schoenberg, F.P., Coronges, K.D. and Bertozzi, A.L. (2016). Modeling e-mail networks and inferring leadership using self-exciting point processes. *J. Amer. Statist. Assoc.* **111** 564–584. MR3538687 https://doi.org/10.1080/01621459.2015.1135802

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. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Probab.* **11** 493–503. MR0378093 https://doi.org/10.2307/3212693

Kirchner, M. (2016). Hawkes and INAR(∞) processes. *Stochastic Process. Appl.* **126** 2494–2525. MR3505235 https://doi.org/10.1016/j.spa.2016.02.008

Kirchner, M. (2017). An estimation procedure for the Hawkes process. *Quant. Finance* **17** 571–595. MR3620953 https://doi.org/10.1080/14697688.2016.1211312

Mammen, E. and Müller, M. (2023). Supplement to "Nonparametric estimation of locally stationary Hawkes processes." https://doi.org/10.3150/22-BEJ1532SUPP

Mohler, G.O., Short, M.B., Brantingham, P.J., Schoenberg, F.P. and Tita, G.E. (2011). Self-exciting point process modeling of crime. *J. Amer. Statist. Assoc.* **106** 100–108. MR2816705 https://doi.org/10.1198/jasa.2011.ap09546

Omi, T., Hirata, Y. and Aihara, K. (2017). Hawkes process model with a time-dependent background rate and its application to high-frequency financial data. *Phys. Rev. E* **96** 012303.

Rambaldi, M., Bacry, E. and Lillo, F. (2017). The role of volume in order book dynamics: A multivariate Hawkes process analysis. *Quant. Finance* **17** 999–1020. MR3660377 https://doi.org/10.1080/14697688.2016.1260759

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

Rosenthal, H.P. (1970). On the subspaces of L_p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303. MR0271721 https://doi.org/10.1007/BF02771562

Roueff, F. and von Sachs, R. (2019). Time-frequency analysis of locally stationary Hawkes processes. *Bernoulli* **25** 1355–1385. MR3920375 https://doi.org/10.3150/18-bej1023

Roueff, F., von Sachs, R. and Sansonnet, L. (2016). Locally stationary Hawkes processes. *Stochastic Process. Appl.* **126** 1710–1743. MR3483734 https://doi.org/10.1016/j.spa.2015.12.003

Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. *J. Roy. Statist. Soc. Ser. B* **32** 1–62. MR0272087

Bernoulli **29**(3), 2023, 2084–2100
https://doi.org/10.3150/22-BEJ1533

Local exchangeability

TREVOR CAMPBELL[1,a], SAIFUDDIN SYED[1,b], CHIAO-YU YANG[2,c], MICHAEL I. JORDAN[2,d] and TAMARA BRODERICK[3,e]

[1]*Department of Statistics, University of British Columbia, Vancouver, Canada,* [a]*trevor@stat.ubc.ca,* [b]*saif.syed@stat.ubc.ca*
[2]*Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, USA,* [c]*chiaoyu@berkeley.edu,* [d]*jordan@eecs.berkeley.edu*
[3]*Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, USA,* [e]*tbroderick@mit.edu*

Exchangeability—in which the distribution of an infinite sequence is invariant to reorderings of its elements—implies the existence of a simple conditional independence structure that may be leveraged in the design of statistical models and inference procedures. In this work, we study a relaxation of exchangeability in which this invariance need not hold precisely. We introduce the notion of *local exchangeability*—where swapping data associated with nearby covariates causes a bounded change in the distribution. We prove that locally exchangeable processes correspond to independent observations from an underlying measure-valued stochastic process. Using this main probabilistic result, we show that the *local empirical measure* of a finite collection of observations provides an approximation of the underlying measure-valued process and Bayesian posterior predictive distributions. The paper concludes with applications of the main theoretical results to a model from Bayesian nonparametrics and covariate-dependent permutation tests.

Keywords: Exchangeability; local; representation; de Finetti; Bayesian nonparametrics

References

Aldous, D.J. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** 581–598. MR0637937 https://doi.org/10.1016/0047-259X(81)90099-3

Aldous, D.J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII—1983*. *Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. MR0883646 https://doi.org/10.1007/BFb0099421

Austin, T. and Panchenko, D. (2014). A hierarchical version of the de Finetti and Aldous-Hoover representations. *Probab. Theory Related Fields* **159** 809–823. MR3230009 https://doi.org/10.1007/s00440-013-0521-0

Baiocchi, M., Small, D.S., Lorch, S. and Rosenbaum, P.R. (2010). Building a stronger instrument in an observational study of perinatal care for premature infants. *J. Amer. Statist. Assoc.* **105** 1285–1296. MR2796550 https://doi.org/10.1198/jasa.2010.ap09490

Berti, P., Pratelli, L. and Rigo, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32** 2029–2052. MR2073184 https://doi.org/10.1214/009117904000000676

Blei, D. and Lafferty, J. (2006). Dynamic topic models. In *International Conference on Machine Learning*.

Borgs, C., Chayes, J.T., Cohn, H. and Holden, N. (2017). Sparse exchangeable graphs and their limits via graphon processes. *J. Mach. Learn. Res.* **18** Paper No. 210. MR3827098

Broderick, T., Pitman, J. and Jordan, M.I. (2013). Feature allocations, probability functions, and paintboxes. *Bayesian Anal.* **8** 801–836. MR3150470 https://doi.org/10.1214/13-BA823

Cai, D., Campbell, T. and Broderick, T. (2016). Edge-exchangeable graphs and sparsity. In *Advances in Neural Information Processing Systems*.

Camerlenghi, F., Lijoi, A., Orbanz, P. and Prünster, I. (2019). Distribution theory for hierarchical processes. *Ann. Statist.* **47** 67–92. MR3909927 https://doi.org/10.1214/17-AOS1678

Campbell, T., Cai, D. and Broderick, T. (2018). Exchangeable trait allocations. *Electron. J. Stat.* **12** 2290–2322. MR3832093 https://doi.org/10.1214/18-EJS1455

Campbell, T., Syed, S., Yang, C.-Y., Jordan, M.I., Broderick, T. (2023). Supplement to "Local Exchangeability." https://doi.org/10.3150/22-BEJ1533SUPP

Caron, F. and Fox, E.B. (2017). Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1295–1366. MR3731666 https://doi.org/10.1111/rssb.12233

Chen, C., Rao, V., Buntine, W. and Teh, Y. (2013). Dependent normalized random measures. In *International Conference on Machine Learning*.

Cochran, W.G. (1965). The planning of observational studies of human populations. *J. R. Stat. Soc., A* **128** 234–266.

Crane, H. and Dempsey, W. (2016). Edge exchangeable models for network data. Available at arXiv:1603.04571v3.

Crane, H. and Dempsey, W. (2019). Relational exchangeability. *J. Appl. Probab.* **56** 192–208. MR3981153 https://doi.org/10.1017/jpr.2019.13

de Finetti, B. (1931). Funzione caratteristica di un fenomeno aleatorio. *Atti della R. Academia Nazionale dei Lincei, Serie 6.* **4** 251–299. In Italian.

de Finetti, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036

de Finetti, B. (1938). Sur la condition d'equivalence partielle. *Actualites Scientifiques et Industrielles* **739**. In French; translated as "On the condition of partial exchangeability," P. Benacerraf and R. Jeffrey (eds) in *Studies in Inductive Logic and Probability II*, 193–205, Berkeley, University of California Press, 1980.

Devroye, L., Mehrabian, A. and Reddad, T. (2020). The total variation distance between high-dimensional Gaussians. Available at arXiv:1810.08693.

Diaconis, P. (1977). Finite forms of de Finetti's theorem on exchangeability: Foundations of probability and statistics, II. *Synthese* **36** 271–281. MR0517222 https://doi.org/10.1007/BF00486116

Diaconis, P. (1988). Recent progress on de Finetti's notions of exchangeability. In *Bayesian Statistics, 3 (Valencia, 1987)*. *Oxford Sci. Publ.* 111–125. New York: Oxford Univ. Press. MR1008047

Diaconis, P. and Freedman, D. (1978). de Finetti's generalizations of exchangeability. Technical Report No. 109, Univ. California, Berkeley.

Diaconis, P. and Freedman, D. (1980b). Finite exchangeable sequences. *Ann. Probab.* **8** 745–764. MR0577313

Diaconis, P. and Freedman, D. (1980c). de Finetti's theorem for Markov chains. *Ann. Probab.* **8** 115–130. MR0556418

Ernst, M.D. (2004). Permutation methods: A basis for exact inference. *Statist. Sci.* **19** 676–685. MR2185589 https://doi.org/10.1214/088342304000000396

Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

Fisher, R. (1966). *The Design of Experiments*, 8th ed. Edinburgh: Oliver & Boyd.

Gibbs, A. and Su, F. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.

Greevy, R., Lu, B., Silber, J. and Rosenbaum, P. (2004). Optimal multivariate matching before randomization. *Biostatistics* **5** 263–275.

Hansen, B.B. (2004). Full matching in an observational study of coaching for the SAT. *J. Amer. Statist. Assoc.* **99** 609–618. MR2086387 https://doi.org/10.1198/016214504000000647

Hansen, B.B. and Klopfer, S.O. (2006). Optimal full matching and related designs via network flows. *J. Comput. Graph. Statist.* **15** 609–627. MR2280151 https://doi.org/10.1198/106186006X137047

Hewitt, E. and Savage, L.J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501. MR0076206 https://doi.org/10.2307/1992999

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

Hoover, D. (1979). Relations on probability spaces and arrays of random variables. Technical Report, Institute for Advanced Study, Princeton Univ.

Janson, S. (2018). On edge exchangeable random graphs. *J. Stat. Phys.* **173** 448–484. MR3876897 https://doi.org/10.1007/s10955-017-1832-9

Jordan, M. (2010). Bayesian nonparametric learning: Expressive priors for intelligent systems. In *Heuristics, Probability and Causality: A Tribute to Judea Pearl* (R. Dechter, H. Geffner and J. Halpern, eds.). College Publications.

Jung, P., Lee, J., Staton, S. and Yang, H. (2021). A generalization of hierarchical exchangeability on trees to directed acyclic graphs. *Ann. Henri Lebesgue* **4** 325–368. MR4213163 https://doi.org/10.5802/ahl.74

Kallenberg, O. (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16** 508–534. MR0929061
Kallenberg, O. (1990). Exchangeable random measures in the plane. *J. Theoret. Probab.* **3** 81–136. MR1031426 https://doi.org/10.1007/BF01063330
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
Kallenberg, O. (2005). *Probabilistic Symmetries and Invariance Principles*. *Probability and Its Applications (New York)*. New York: Springer. MR2161313
Kingman, J.F.C. (1978). The representation of partition structures. *J. Lond. Math. Soc. (2)* **18** 374–380. MR0509954 https://doi.org/10.1112/jlms/s2-18.2.374
Lauritzen, S. (1974). On the interrelationships among sufficiency, total sufficiency, and some related concepts. Technical Report, Institute of Mathematical Statistics, Univ. Copenhagen.
Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR2135927
Lin, D. and Fisher, J. (2010). Construction of dependent Dirichlet processes based on Poisson processes. In *Advances in Neural Information Processing Systems*.
Lu, B. and Rosenbaum, P.R. (2004). Optimal pair matching with two control groups. *J. Comput. Graph. Statist.* **13** 422–434. MR2063993 https://doi.org/10.1198/1061860043470
Lu, B., Greevy, R., Xu, X. and Beck, C. (2011). Optimal nonbipartite matching and its statistical applications. *Amer. Statist.* **65** 21–30. MR2899649 https://doi.org/10.1198/tast.2011.08294
MacEachern, S. (1999). Dependent nonparametric processes. Technical Report, The Ohio State Univ.
MacEachern, S. (2000). Dependent Dirichlet processes. Technical Report, The Ohio State Univ.
Neal, R.M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* **9** 249–265. MR1823804 https://doi.org/10.2307/1390653
Orbanz, P. and Roy, D. (2015). Bayesian models of graphs, arrays, and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 437–461.
Pitman, E. (1937a). Significance tests which may be applied to samples from any populations. *Suppl. J. R. Stat. Soc.* **4** 119–130.
Pitman, E. (1937b). Significance tests which may be applied to samples from any populations II: The correlation coefficient test. *Suppl. J. R. Stat. Soc.* **4** 225–232.
Pitman, E. (1937c). Significance tests which may be applied to samples from any populations III: The analysis of variance test. *Biometrika* **29** 322–335.
Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249 https://doi.org/10.1007/BF01213386
Potthoff, J. (2009). Sample properties of random fields. II. Continuity. *Commun. Stoch. Anal.* **3** 331–348. MR2604006 https://doi.org/10.31390/cosa.3.3.02
Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR2514435
Ren, L., Wang, Y., Dunson, D. and Carin, L. (2011). The kernel beta process. In *Advances in Neural Information Processing Systems*.
Rosenbaum, P.R. (1984). Conditional permutation tests and the propensity score in observational studies. *J. Amer. Statist. Assoc.* **79** 565–574. MR0763575
Rosenbaum, P. (1989). Optimal matching for observational studies. *J. Amer. Statist. Assoc.* **84** 1024–1032.
Rosenbaum, P.R. (2002). Covariance adjustment in randomized experiments and observational studies. *Statist. Sci.* **17** 286–327. MR1962487 https://doi.org/10.1214/ss/1042727942
Rubin, D. (1973b). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics* **29** 185–203.
Rubin, D. (1973a). Matching to remove bias in observational studies. *Biometrics* **29** 159–183.
Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica* **4** 639–650. MR1309433
Varadarajan, V.S. (1958). On the convergence of sample probability distributions. *Sankhyā* **19** 23–26. MR0094839
Veitch, V. and Roy, D. (2015). The class of random graphs arising from exchangeable random measures. Available at arXiv:1512.03099.
Wang, C., Blei, D. and Heckerman, D. (2008). Continuous time dynamic topic models. In *Uncertainty in Artificial Intelligence*.

Bernoulli **29**(3), 2023, 2101–2118
https://doi.org/10.3150/22-BEJ1534

Consistent and unbiased variable selection under indepedent features using Random Forest permutation importance

BURIM RAMOSAJ[a] and MARKUS PAULY[b]

Faculty of Statistics, Institute of Mathematical Statistics and Applications in Industry, Technical University of Dortmund, Joseph-von-Fraunhofer-Straße 2-4, 44227 Dortmund, Germany, [a]*burim.ramosaj@tu-dortmund.de,* [b]*markus.pauly@tu-dortmund.de*

Variable selection in sparse regression models is an important task as applications ranging from biomedical research to econometrics have shown. Especially for higher dimensional regression problems, for which the regression function as the link between response and covariates cannot be directly detected, the selection of informative variables is challenging. Under these circumstances, the Random Forest method is an helpful tool to predict new outcomes while delivering measures for variable selection. One common approach is the usage of the permutation importance. Due to its intuitive idea and flexible usage, it is important to explore circumstances, for which the permutation importance based on Random Forest correctly indicates informative covariates. Regarding the latter, we deliver theoretical guarantees for the validity of the permutation importance measure under specific assumptions such as the mutual independence of the features and prove its (asymptotic) unbiasedness, while under slightly stricter assumptions, consistency of the permutation importance measure is established. An extensive simulation study supports our findings.

Keywords: Random Forest; permutation importance; unbiasedness; consistency; Out-of-Bag samples; statistical learning

References

Altmann, A., Toloşi, L., Sander, O. and Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics* **26** 1340–1347.

Archer, K.J. and Kimes, R.V. (2008). Empirical characterization of random forest variable importance measures. *Comput. Statist. Data Anal.* **52** 2249–2260. MR2418502 https://doi.org/10.1016/j.csda.2007.08.015

Breiman, L. (2001). Random Forests. *Mach. Learn.* **45** 5–32.

Chi, C.-M., Vossler, P., Fan, Y. and Lv, J. (2020). Asymptotic Properties of High-Dimensional Random Forests. Preprint. Available at arXiv:2004.13953.

Coleman, T., Peng, W. and Mentch, L. (2019). Scalable and Efficient Hypothesis Testing with Random Forests. Preprint. Available at arXiv:1904.07830.

Genuer, R., Poggi, J.-M. and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recogn. Lett.* **31** 2225–2236.

Gregorutti, B., Michel, B. and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Stat. Comput.* **27** 659–678. MR3613591 https://doi.org/10.1007/s11222-016-9646-1

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3** 1157–1182.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2722294 https://doi.org/10.1007/978-0-387-84858-7

Louppe, G., Wehenkel, L., Sutera, A. and Geurts, P. (2013). Understanding variable importances in forests of randomized trees. In *Advances in Neural Information Processing Systems* 431–439.

Mentch, L. and Hooker, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *J. Mach. Learn. Res.* **17** Paper No. 26, 41 pp. MR3491120

Nicodemus, K.K. (2011). Letter to the editor: On the stability and ranking of predictors from random forest variable importance measures. *Brief. Bioinform.* **12** 369–373.

Nicodemus, K.K. and Malley, J.D. (2009). Predictor correlation impacts machine learning algorithms: Implications for genomic studies. *Bioinformatics* **25** 1884–1890.

Nicodemus, K.K., Malley, J.D., Strobl, C. and Ziegler, A. (2010). The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinform.* **11** 110. https://doi.org/10.1186/1471-2105-11-110

Ramosaj, B. (2021). Interpretable Machines: Constructing Valid Prediction Intervals with Random Forests. Preprint. Available at arXiv:2103.05766.

Ramosaj, B. and Pauly, M. (2019). Consistent estimation of residual variance with random forest out-of-bag errors. *Statist. Probab. Lett.* **151** 49–57. MR3937367 https://doi.org/10.1016/j.spl.2019.03.017

Ramosaj, B. and Pauly, M. (2023). Supplement to "Consistent and unbiased variable selection under indepedent features using random forest permutation importance." https://doi.org/10.3150/22-BEJ1534SUPP

Scornet, E., Biau, G. and Vert, J.-P. (2015). Consistency of random forests. *Ann. Statist.* **43** 1716–1741. MR3357876 https://doi.org/10.1214/15-AOS1321

Strobl, C., Boulesteix, A.-L., Zeileis, A. and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **8** 25.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. and Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinform.* **9** 307. https://doi.org/10.1186/1471-2105-9-307

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Statist. Assoc.* **113** 1228–1242. MR3862353 https://doi.org/10.1080/01621459.2017.1319839

Wager, S., Hastie, T. and Efron, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *J. Mach. Learn. Res.* **15** 1625–1651. MR3225243

Zhu, R., Zeng, D. and Kosorok, M.R. (2015). Reinforcement learning trees. *J. Amer. Statist. Assoc.* **110** 1770–1784. MR3449072 https://doi.org/10.1080/01621459.2015.1036994

Bernoulli **29**(3), 2023, 2119–2141
https://doi.org/10.3150/22-BEJ1535

On adaptive confidence sets for the Wasserstein distances

NEIL DEO[1,a] and THIBAULT RANDRIANARISOA[2,b]

[1]*Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK,*
[a]*and30@cam.ac.uk*
[2]*Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, Paris, France,*
[b]*thibault.randrianarisoa@sorbonne-universite.fr*

In the density estimation model, we investigate the problem of constructing *adaptive honest confidence sets* with diameter measured in Wasserstein distance W_p, $p \geq 1$, and for densities with unknown regularity measured on a Besov scale. As sampling domains, we focus on the d-dimensional torus $\mathbb{T}^d$, in which case $1 \leq p \leq 2$, and $\mathbb{R}^d$, for which $p = 1$. We identify necessary and sufficient conditions for the existence of adaptive confidence sets with diameters of the order of the regularity-dependent W_p-minimax estimation rate. Interestingly, it appears that the possibility of such adaptation of the diameter depends on the dimension of the underlying space. In low dimensions, $d \leq 4$, adaptation to any regularity is possible. In higher dimensions, adaptation is possible if and only if the underlying regularities belong to some bounded interval, whose width can be chosen to be at least $d/(d-4)$. This contrasts with the L_p-theory where, independently of the dimension, adaptation occurs only if regularities lie in a small fixed-width window. When possible, we explicitly construct confidence regions via the method of risk estimation. These are the first results in a statistical approach to adaptive uncertainty quantification with Wasserstein distances. Our analysis and methods extend to weak losses such as Sobolev norms with negative smoothness indices.

Keywords: Wasserstein distance; uncertainty quantification; nonparametric confidence sets

References

[1] Armstrong, T.B. (2021). Adaptation bounds for confidence bands under self-similarity. *Bernoulli* **27** 1348–1370. MR4255237 https://doi.org/10.3150/20-bej1277
[2] Bull, A.D. (2012). Honest adaptive confidence bands and self-similar functions. *Electron. J. Stat.* **6** 1490–1516. MR2988456 https://doi.org/10.1214/12-EJS720
[3] Bull, A.D. and Nickl, R. (2013). Adaptive confidence sets in L^2. *Probab. Theory Related Fields* **156** 889–919. MR3078289 https://doi.org/10.1007/s00440-012-0446-z
[4] Cai, T.T. and Low, M.G. (2006). Adaptive confidence balls. *Ann. Statist.* **34** 202–228. MR2275240 https://doi.org/10.1214/009053606000000146
[5] Carpentier, A. (2013). Honest and adaptive confidence sets in L_p. *Electron. J. Stat.* **7** 2875–2923. MR3148371 https://doi.org/10.1214/13-EJS867
[6] 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
[7] Castillo, I. and Nickl, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. MR3262473 https://doi.org/10.1214/14-AOS1246
[8] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Anti-concentration and honest, adaptive confidence bands. *Ann. Statist.* **42** 1787–1818. MR3262468 https://doi.org/10.1214/14-AOS1235
[9] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999 https://doi.org/10.1214/aop/1022677394

1350-7265

[10] Deo, N. and Randrianarisoa, T. (2023). Supplement to "On Adaptive Confidence Sets for the Wasserstein Distances." https://doi.org/10.3150/22-BEJ1535SUPP
[11] Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24** 508–539. MR1394974 https://doi.org/10.1214/aos/1032894451
[12] Dudley, R.M. (1968). The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Stat.* **40** 40–50. MR0236977 https://doi.org/10.1214/aoms/1177697802
[13] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7
[14] Giné, E. and Nickl, R. (2010). Confidence bands in density estimation. *Ann. Statist.* **38** 1122–1170. MR2604707 https://doi.org/10.1214/09-AOS738
[15] 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
[16] Hoffmann, M. and Nickl, R. (2011). On adaptive inference and confidence bands. *Ann. Statist.* **39** 2383–2409. MR2906872 https://doi.org/10.1214/11-AOS903
[17] Juditsky, A. and Lambert-Lacroix, S. (2003). Nonparametric confidence set estimation. *Math. Methods Statist.* **12** 410–428. MR2054156
[18] Kantorovič, L.V. and Rubinšteĭn, G.Š. (1958). On a space of completely additive functions. *Vestn. Leningr. Univ.* **13** 52–59. MR0102006
[19] Kantorovitch, L. (1942). On the translocation of masses. *C. R. (Dokl.) Acad. Sci. URSS* **37** 199–201. MR0009619
[20] Lepskiĭ, O.V. (1990). A problem of adaptive estimation in Gaussian white noise. *Teor. Veroyatn. Primen.* **35** 459–470. MR1091202 https://doi.org/10.1137/1135065
[21] Low, M.G. (1997). On nonparametric confidence intervals. *Ann. Statist.* **25** 2547–2554. MR1604412 https://doi.org/10.1214/aos/1030741084
[22] McDonald, S. and Campbell, D. (2021). A review of uncertainty quantification for density estimation. *Stat. Surv.* **15** 1–71. MR4255286 https://doi.org/10.1214/21-ss130
[23] Monge, G. (1781). *Mémoire sur la Théorie des Déblais et des Remblais*. De l'Imprimerie Royale.
[24] Nickl, R. and Szabó, B. (2016). A sharp adaptive confidence ball for self-similar functions. *Stochastic Process. Appl.* **126** 3913–3934. MR3565485 https://doi.org/10.1016/j.spa.2016.04.017
[25] Panaretos, V.M. and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** 405–431. MR3939527 https://doi.org/10.1146/annurev-statistics-030718-104938
[26] Peyré, G. and Cuturi, M. (2019). Computational optimal transport. *Found. Trends Mach. Learn.* **11** 355–607.
[27] Picard, D. and Tribouley, K. (2000). Adaptive confidence interval for pointwise curve estimation. *Ann. Statist.* **28** 298–335. MR1762913 https://doi.org/10.1214/aos/1016120374
[28] 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
[29] Robins, J. and van der Vaart, A. (2006). Adaptive nonparametric confidence sets. *Ann. Statist.* **34** 229–253. MR2275241 https://doi.org/10.1214/009053605000000877
[30] Rousseau, J. and Szabo, B. (2020). Asymptotic frequentist coverage properties of Bayesian credible sets for sieve priors. *Ann. Statist.* **48** 2155–2179. MR4134790 https://doi.org/10.1214/19-AOS1881
[31] Szabó, B., van der Vaart, A.W. and van Zanten, J.H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428. MR3357861 https://doi.org/10.1214/14-AOS1270
[32] Villani, C. (2009). *Optimal Transport: Old and New*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9
[33] Weed, J. and Bach, F. (2019). Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** 2620–2648. MR4003560 https://doi.org/10.3150/18-BEJ1065
[34] Weed, J. and Berthet, Q. (2019). Estimation of smooth densities in Wasserstein distance. In *Proceedings of the Thirty-Second Conference on Learning Theory* (A. Beygelzimer and D. Hsu, eds.). *Proceedings of Machine Learning Research* **99** 3118–3119. PMLR.

Bernoulli **29**(3), 2023, 2142–2166
https://doi.org/10.3150/22-BEJ1536

A unified construction for series representations and finite approximations of completely random measures

JUHO LEE[1,a], XENIA MISCOURIDOU[2,b] and FRANÇOIS CARON[3,c]

[1]*Graduate school of AI, KAIST, Daejeon, Republic of Korea,* [a]*juholee@kaist.ac.kr*
[2]*Department of Mathematics, Imperial College London, England,* [b]*xeniamisc@gmail.com*
[3]*Department of Statistics, University of Oxford, England,* [c]*caron@stats.ox.ac.uk*

Infinite-activity completely random measures (CRMs) have become important building blocks of complex Bayesian nonparametric models. They have been successfully used in various applications such as clustering, density estimation, latent feature models, survival analysis or network science. Popular infinite-activity CRMs include the (generalised) gamma process and the (stable) beta process. However, except in some specific cases, exact simulation or scalable inference with these models is challenging and finite-dimensional approximations are often considered. In this work, we propose a general and unified framework to derive both series representations and finite-dimensional approximations of CRMs. Our framework can be seen as a generalisation of constructions based on size-biased sampling of Poisson point process (*Probab. Theory Related Fields* **92** (1992) 21–39). It includes as special cases several known series representations and finite approximations as well as novel ones. In particular, we show that one can get novel series representations for the generalised gamma process and the stable beta process. We show how these constructions can be used to derive novel algorithms for posterior inference, including a generalisation of the slice sampler for normalised CRMs mixture models introduced by (*J. Comput. Graph. Statist.* **20** (2011) 241–259). We also provide some analysis of the truncation error.

Keywords: Bayesian nonparametrics; Poisson random measures; completely random measures; generalised gamma process; slice sampling

References

[1] 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
[2] Arbel, J. and Prünster, I. (2017). A moment-matching Ferguson & Klass algorithm. *Stat. Comput.* **27** 3–17. MR3598905 https://doi.org/10.1007/s11222-016-9676-8
[3] Argiento, R., Bianchini, I. and Guglielmi, A. (2016). A blocked Gibbs sampler for NGG-mixture models via a priori truncation. *Stat. Comput.* **26** 641–661. MR3489862 https://doi.org/10.1007/s11222-015-9549-6
[4] Argiento, R., Bianchini, I. and Guglielmi, A. (2016). Posterior sampling from ε-approximation of normalized completely random measure mixtures. *Electron. J. Stat.* **10** 3516–3547. MR3572858 https://doi.org/10.1214/16-EJS1168
[5] Ayed, F. and Caron, F. (2021). Nonnegative Bayesian nonparametric factor models with completely random measures. *Stat. Comput.* **31** Paper No. 63. MR4303601 https://doi.org/10.1007/s11222-021-10037-3
[6] Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 167–241. MR1841412 https://doi.org/10.1111/1467-9868.00282
[7] Barrios, E., Lijoi, A., Nieto-Barajas, L.E. and Prünster, I. (2013). Modeling with normalized random measure mixture models. *Statist. Sci.* **28** 313–334. MR3135535 https://doi.org/10.1214/13-STS416

[8] Bertoin, J., Fujita, T., Roynette, B. and Yor, M. (2006). On a particular class of self-decomposable random variables: The durations of Bessel excursions straddling independent exponential times. *Probab. Math. Statist.* **26** 315–366. MR2325310
[9] Blei, D.M. and Jordan, M.I. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Anal.* **1** 121–143. MR2227367 https://doi.org/10.1214/06-BA104
[10] Bondesson, L. (1982). On simulation from infinitely divisible distributions. *Adv. in Appl. Probab.* **14** 855–869. MR0677560 https://doi.org/10.2307/1427027
[11] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.* **31** 929–953. MR1747450 https://doi.org/10.1239/aap/1029955251
[12] Cai, D., Campbell, T. and Broderick, T. (2016). Edge-exchangeable graphs and sparsity. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., **29** 4249–4257. https://proceedings.neurips.cc/paper/2016/file/1a0a283bfe7c549dee6c638a05200e32-Paper.pdf.
[13] Campbell, T., Cai, D. and Broderick, T. (2018). Exchangeable trait allocations. *Electron. J. Stat.* **12** 2290–2322. MR3832093 https://doi.org/10.1214/18-EJS1455
[14] Campbell, T., Huggins, J.H., How, J.P. and Broderick, T. (2019). Truncated random measures. *Bernoulli* **25** 1256–1288. MR3920372 https://doi.org/10.3150/18-bej1020
[15] Canale, A., Corradin, R. and Nipoti, B. (2022). Importance conditional sampling for Pitman–Yor mixtures. *Stat. Comput.* **32** Paper No. 40. MR4423541 https://doi.org/10.1007/s11222-022-10096-0
[16] Caron, F. (2012). Bayesian nonparametric models for bipartite graphs. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc. **2** 2051–2059. https://proceedings.neurips.cc/paper/2012/file/0768281a05da9f27df178b5c39a51263-Paper.pdf.
[17] Caron, F. and Fox, E.B. (2017). Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1295–1366. MR3731666 https://doi.org/10.1111/rssb.12233
[18] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: CRC Press/CRC. MR2042661
[19] Devroye, L. and James, L. (2014). On simulation and properties of the stable law. *Stat. Methods Appl.* **23** 307–343. MR3233961 https://doi.org/10.1007/s10260-014-0260-0
[20] Doshi, F., Miller, K., Van Gael, J. and Teh, Y.W. (2009). Variational inference for the Indian buffet process. In *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*. PMLR **5** 137–144. http://proceedings.mlr.press/v5/doshi09a/doshi09a.pdf.
[21] Favaro, S., Lomeli, M., Nipoti, B. and Teh, Y.W. (2014). On the stick-breaking representation of σ-stable Poisson-Kingman models. *Electron. J. Stat.* **8** 1063–1085. MR3263112 https://doi.org/10.1214/14-EJS921
[22] Favaro, S. and Teh, Y.W. (2013). MCMC for normalized random measure mixture models. *Statist. Sci.* **28** 335–359. MR3135536 https://doi.org/10.1214/13-STS422
[23] Favaro, S. and Walker, S.G. (2013). Slice sampling σ-stable Poisson-Kingman mixture models. *J. Comput. Graph. Statist.* **22** 830–847. MR3173745 https://doi.org/10.1080/10618600.2012.681211
[24] Ferguson, T.S. and Klass, M.J. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Stat.* **43** 1634–1643. MR0373022 https://doi.org/10.1214/aoms/1177692395
[25] Foti, N. and Williamson, S. (2012). Slice sampling normalized kernel-weighted completely random measure mixture models. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc. **25** 2240–2248. https://proceedings.neurips.cc/paper/2012/file/8f1d43620bc6bb580df6e80b0dc05c48-Paper.pdf.
[26] Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–511.
[27] Gopalan, P., Ruiz, F.J., Ranganath, R. and Blei, D.M. (2014). Bayesian nonparametric Poisson factorization for recommendation systems. In *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics*. PMLR **33** 275–283. http://proceedings.mlr.press/v33/gopalan14.pdf.
[28] Griffin, J.E. (2017). Sequential Monte Carlo methods for mixtures with normalized random measures with independent increments priors. *Stat. Comput.* **27** 131–145. MR3598913 https://doi.org/10.1007/s11222-015-9612-3
[29] Griffin, J.E. and Walker, S.G. (2011). Posterior simulation of normalized random measure mixtures. *J. Comput. Graph. Statist.* **20** 241–259. MR2816547 https://doi.org/10.1198/jcgs.2010.08176

[30] Griffiths, T.L. and Ghahramani, Z. (2005). Infinite latent feature models and the Indian buffet process. In *Advances in Neural Information Processing Systems*, MIT Press, **18**, https://proceedings.neurips.cc/paper/2005/file/2ef35a8b78b572a47f56846acbeef5d3-Paper.pdf.

[31] 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

[32] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73** 387–396. MR0855898 https://doi.org/10.1093/biomet/73.2.387

[33] Ishwaran, H. and James, L.F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96** 161–173. MR1952729 https://doi.org/10.1198/016214501750332758

[34] Ishwaran, H. and Zarepour, M. (2002). Exact and approximate sum representations for the Dirichlet process. *Canad. J. Statist.* **30** 269–283. MR1926065 https://doi.org/10.2307/3315951

[35] James, L.F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33** 1771–1799. MR2166562 https://doi.org/10.1214/009053605000000336

[36] James, L.F. (2017). Bayesian Poisson calculus for latent feature modeling via generalized Indian buffet process priors. *Ann. Statist.* **45** 2016–2045. MR3718160 https://doi.org/10.1214/16-AOS1517

[37] James, L.F., Lijoi, A. and Prünster, I. (2009). Posterior analysis for normalized random measures with independent increments. *Scand. J. Stat.* **36** 76–97. MR2508332 https://doi.org/10.1111/j.1467-9469.2008.00609.x

[38] Kalli, M., Griffin, J.E. and Walker, S.G. (2011). Slice sampling mixture models. *Stat. Comput.* **21** 93–105. MR2746606 https://doi.org/10.1007/s11222-009-9150-y

[39] Khintchine, A. (1937). Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze. *Rec. Math. Moscou, N. Ser* **2** 79–117.

[40] Kingman, J.F.C. (1967). Completely random measures. *Pacific J. Math.* **21** 59–78. MR0210185

[41] Kingman, J.F.C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. Oxford University Press, New York: The Clarendon Press. MR1207584

[42] Kingman, J.F.C., Taylor, S.J., Hawkes, A.G., Walker, A.M., Cox, D.R., Smith, A.F.M., Hill, B.M., Burville, P.J. and Leonard, T. (1975). Random discrete distribution. *J. Roy. Statist. Soc. Ser. B* **37** 1–22. MR0368264

[43] Lee, J., James, L.F. and Choi, S. (2016). Finite-dimensional BFRY priors and variational Bayesian inference for power-law models. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc. **29** 3170–3178. https://proceedings.neurips.cc/paper/2016/file/0d4f4805c36dc6853edfa4c7e1638b48-Paper.pdf.

[44] Lee, J., Miscouridou, X. and Caron, F. (2023). Supplement to "A unified construction for series representations and finite approximations of completely random measures." https://doi.org/10.3150/22-BEJ1536SUPP

[45] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Hierarchical mixture modeling with normalized inverse-Gaussian priors. *J. Amer. Statist. Assoc.* **100** 1278–1291. MR2236441 https://doi.org/10.1198/016214505000000132

[46] Lijoi, A., Mena, R.H. and Prünster, I. (2007). Controlling the reinforcement in Bayesian non-parametric mixture models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 715–740. MR2370077 https://doi.org/10.1111/j.1467-9868.2007.00609.x

[47] Lijoi, A. and Prünster, I. (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (N.L. Hjort, C. Holmes, P. Müller and S.G. Walker, eds.). *Camb. Ser. Stat. Probab. Math.* **28** 80–136. Cambridge: Cambridge Univ. Press. MR2730661

[48] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. MR0733519 https://doi.org/10.1214/aos/1176346412

[49] Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. Boca Raton, FL: CRC Press/CRC. MR2004226

[50] Neal, R.M. (2003). Slice sampling. *Ann. Statist.* **31** 705–767. MR1994729 https://doi.org/10.1214/aos/1056562461

[51] Nguyen, T.D., Huggins, J., Masoero, L., Mackey, L. and Broderick, T. (2020). Independent finite approximations for Bayesian nonparametric inference. ArXiv Preprint. Available at arXiv:2009.10780.

[52] Nieto-Barajas, L.E., Prünster, I. and Walker, S.G. (2004). Normalized random measures driven by increasing additive processes. *Ann. Statist.* **32** 2343–2360. MR2153987 https://doi.org/10.1214/009053604000000625

[53] Paisley, J., Blei, D. and Jordan, M.L. (2012). Stick-breaking beta processes and the Poisson process. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*. PMLR **22** 850–858. http://proceedings.mlr.press/v22/paisley12/paisley12.pdf.

[54] Paisley, J. and Carin, L. (2009). Nonparametric factor analysis with beta process priors. In *Proceedings of the 26th International Conference on Machine Learning*. ACM 777–784. https://doi.org/10.1145/1553374.1553474.

[55] Paisley, J., Carin, L. and Blei, D. (2011). Variational inference for stick-breaking beta process priors, In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*. Omnipress 889–896. https://icml.cc/2011/papers/475_icmlpaper.pdf.

[56] Perman, M., Pitman, J. and Yor, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields* **92** 21–39. MR1156448 https://doi.org/10.1007/BF01205234

[57] Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. https://doi.org/10.1214/aos/1051027881

[58] Rosiński, J. (1990). On series representations of infinitely divisible random vectors. *Ann. Probab.* **18** 405–430. MR1043955

[59] Rosiński, J. (2001). Contribution to the discussion of the paper "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics" by O. Barndorff-Nielsen and N. Shephard. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 230–231.

[60] Rosiński, J. (2001). Series representations of Lévy processes from the perspective of point processes. In *Lévy Processes* (O.E. Barndoff-Nielson, T. Mikosch and S.I. Resnick, eds.) 401–415. Boston, MA: Birkhäuser. MR1833707

[61] Rosiński, J. (2007). Tempering stable processes. *Stochastic Process. Appl.* **117** 677–707. MR2327834 https://doi.org/10.1016/j.spa.2006.10.003

[62] Roychowdhury, A. and Kulis, B. (2015). Gamma processes, stick-breaking, and variational inference. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*. PMLR **38** 800–808. http://proceedings.mlr.press/v38/roychowdhury15.pdf.

[63] Teh, Y.W. and Gorur, D. (2009). Indian buffet processes with power-law behavior. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc. **22** 1838–1846. https://proceedings.neurips.cc/paper/2009/file/f1b6f2857fb6d44dd73c7041e0aa0f19-Paper.pdf.

[64] Teh, Y.W., Görür, D. and Ghahramani, Z. (2007). Stick-breaking construction for the Indian buffet process. In *Proceedings of the 11th International Conference on Artificial Intelligence and Statistics*. PMLR **2** 556–563. http://proceedings.mlr.press/v2/teh07a/teh07a.pdf.

[65] Titsias, M.K. (2007). The infinite gamma-Poisson feature model. In *Advances in Neural Information Processing Systems 20*. Curran Associates, Inc. **20** 1513–1520. https://proceedings.neurips.cc/paper/2007/file/6da37dd3139aa4d9aa55b8d237ec5d4a-Paper.pdf.

[66] Walker, S.G. (2007). Sampling the Dirichlet mixture model with slices. *Comm. Statist. Simulation Comput.* **36** 45–54. MR2370888 https://doi.org/10.1080/03610910601096262

[67] Wolpert, R.L. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85** 251–267. MR1649114 https://doi.org/10.1093/biomet/85.2.251

[68] Wolpert, R.L. and Ickstadt, K. (1998). Simulation of Lévy random fields. In *Practical Nonparametric and Semiparametric Bayesian Statistics*. *Lect. Notes Stat.* **133** 227–242. New York: Springer. MR1630084 https://doi.org/10.1007/978-1-4612-1732-9_12

[69] Zhou, M. (2015). Infinite edge partition models for overlapping community detection and link prediction. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*. PMLR **38** 1135–1143. http://proceedings.mlr.press/v38/zhou15a.pdf.

[70] Zhou, M., Hannah, L., Dunson, D. and Carin, L. (2012). Beta-negative binomial process and Poisson factor analysis. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*. PMLR **22** 1462–1471. http://proceedings.mlr.press/v22/zhou12c/zhou12c.pdf.

Bernoulli **29**(3), 2023, 2167–2191
https://doi.org/10.3150/22-BEJ1537

Multivariate self-exciting jump processes with applications to financial data

HEIDAR EYJOLFSSON[1,a] and DAG TJØSTHEIM[2,b]

[1]*Department of Engineering, Reykjavik University, Menntavegi 1, 102 Reykjavik, Iceland,* [a]*heidare@ru.is*
[2]*Department of Mathematics, University of Bergen, Postbox 7803, 5020 Bergen, Norway,*
[b]*Dag.Tjostheim@uib.no*

The paper discusses multivariate self- and cross-exciting processes. We define a class of multivariate point processes via their corresponding stochastic intensity processes that are driven by stochastic jumps. Essentially, there is a jump in an intensity process whenever the corresponding point process records an event. An attribute of our modelling class is that not only a jump is recorded at each instance, but also its magnitude. This allows large jumps to influence the intensity to a larger degree than smaller jumps. We give conditions which guarantee that the process is stable, in the sense that it does not explode, and provide a detailed discussion on when particular subclasses, which include both linear and non-linear models, are stable. Finally, we fit our model to financial time series data from the S&P 500 and Nikkei 225 indices respectively. We conclude that a non-linear variant from our modelling class fits the data best. This supports the observation that in times of crises (high intensity) jumps tend to arrive in clusters, whereas there are typically longer times between jumps when the markets are calmer. We moreover observe more variability in jump sizes when the intensity is high, than when it is low.

Keywords: Multivariate self-exciting processes; Markov processes; jump processes; financial time series

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2015). Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ.* **117** 585–606.

[2] Behr, M., Benner, P. and Heiland, J. (2019). Solution formulas for differential Sylvester and Lyapunov equations. *Calcolo* **56** Paper No. 51, 33 pp. MR4034914 https://doi.org/10.1007/s10092-019-0348-x

[3] Bormetti, G., Calcagnile, L.M., Treccani, M., Corsi, F., Marmi, S. and Lillo, F. (2015). Modelling systemic price cojumps with Hawkes factor models. *Quant. Finance* **15** 1137–1156. MR3354128 https://doi.org/10.1080/14697688.2014.996586

[4] Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. *Springer Series in Statistics*. New York–Berlin: Springer. MR0636252

[5] Carbonell, F., Jímenez, J.C. and Pedroso, L.M. (2008). Computing multiple integrals involving matrix exponentials. *J. Comput. Appl. Math.* **213** 300–305. MR2382698 https://doi.org/10.1016/j.cam.2007.01.007

[6] Clinet, S. and Yoshida, N. (2017). Statistical inference for ergodic point processes and application to limit order book. *Stochastic Process. Appl.* **127** 1800–1839. MR3646432 https://doi.org/10.1016/j.spa.2016.09.014

[7] Dahl, K.R. and Eyjolfsson, H. (2022). Self-exciting jump processes and their asymptotic behaviour. *Stochastics* 1–20. https://doi.org/10.1080/17442508.2022.2028789

[8] Daley, D.J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. *Springer Series in Statistics*. New York: Springer. MR0950166

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

[10] Davis, R.A., Holan, S.H., Lund, R. and Ravishanker, N., eds. (2016). *Handbook of Discrete-Valued Time Series*. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. Boca Raton, FL: CRC Press. MR3642975

[11] Debaly, Z.M. and Truquet, L. (2021). Multivariate time series models for mixed data.

1350-7265

[12] Debaly, Z.M. and Truquet, L. (2021). A note on the stability of multivariate non-linear time series with an application to time series of counts. *Statist. Probab. Lett.* **179** Paper No. 109196, 7 pp. MR4293765 https://doi.org/10.1016/j.spl.2021.109196
[13] 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
[14] Engle, R.F. and Russell, J.R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* **66** 1127–1162. MR1639411 https://doi.org/10.2307/2999632
[15] Eyjolfsson, H. and Tjøstheim, D. (2018). Self-exciting jump processes with applications to energy markets. *Ann. Inst. Statist. Math.* **70** 373–393. MR3768938 https://doi.org/10.1007/s10463-016-0591-8
[16] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression. *J. Amer. Statist. Assoc.* **104** 1430–1439. MR2596998 https://doi.org/10.1198/jasa.2009.tm08270
[17] Fokianos, K., Støve, B., Tjøstheim, D. and Doukhan, P. (2020). Multivariate count autoregression. *Bernoulli* **26** 471–499. MR4036041 https://doi.org/10.3150/19-BEJ1132
[18] Hawkes, A.G. (1971). Point spectra of some mutually exciting point processes. *J. Roy. Statist. Soc. Ser. B* **33** 438–443. MR0358976
[19] Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. MR0278410 https://doi.org/10.1093/biomet/58.1.83
[20] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. MR1943877 https://doi.org/10.1007/978-3-662-05265-5
[21] Meyn, S.P. and Tweedie, R.L. (1992). Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.* **24** 542–574. MR1174380 https://doi.org/10.2307/1427479
[22] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** 518–548. MR1234295 https://doi.org/10.2307/1427522
[23] Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Statist. Math.* **30** 243–261. MR0514494 https://doi.org/10.1007/BF02480216
[24] Protter, P.E. (2005). *Stochastic Integration and Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **21**. Berlin: Springer. Version 2.1, Corrected third printing. MR2273672 https://doi.org/10.1007/978-3-662-10061-5
[25] Swishchuk, A., Zagst, R. and Zeller, G. (2021). Hawkes processes in insurance: Risk model, application to empirical data and optimal investment. *Insurance Math. Econom.* **101** 107–124. MR4330906 https://doi.org/10.1016/j.insmatheco.2020.12.005
[26] Zhuang, J., Ogata, Y. and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *J. Amer. Statist. Assoc.* **97** 369–380. MR1941459 https://doi.org/10.1198/016214502760046925
[27] Zhuang, J., Ogata, Y. and Vere-Jones, D. (2004). Analyzing earthquake clustering features by using stochastic reconstruction. *J. Geophys. Res.* **109** B05301. https://doi.org/10.1029/2003JB002879

Bernoulli **29**(3), 2023, 2192–2218
https://doi.org/10.3150/22-BEJ1538

Bootstrap inference for a class of non-regular estimators

MIHAI GIURCANU[1,a] and BRETT PRESNELL[2,b]

[1]*Department of Public Health Sciences, University of Chicago, Chicago, Illinois, USA,* [a]*giurcanu@uchicago.edu*
[2]*Department of Statistics, University of Florida, Gainesville, Florida, USA,* [b]*presnell@ufl.edu*

We introduce a class of non-regular estimators in arbitrary normed spaces which include estimators for a non-negative mean, for the squared mean, as well as for the Hodges and Stein estimators. In these cases, the nonparametric bootstrap is consistent on all but a small subset of the underlying parameter space. Bootstrap remedies, such as the *m*-out-of-*n* bootstrap and the oracle bootstrap, have been proposed to mainly solve the inconsistency of the nonparametric bootstrap under a fixed parameter setting. We study the local asymptotic behavior of the estimators and of their bootstrap distributions by allowing the underlying parameter to approach a fixed value at various rates. Our theoretical results determine the precise limiting behavior of the estimators and of their bootstrap distributions in these problems. Simulation results examining the finite sample local performance of the bootstrap estimators are provided.

Keywords: Nonparametric bootstrap; *m*-out-of-*n* bootstrap; oracle bootstrap; Hodges estimator; Stein estimator

References

[1] Andrews, D.W.K. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* **68** 399–405. MR1748009 https://doi.org/10.1111/1468-0262.00114
[2] Andrews, D.W.K. and Guggenberger, P. (2010). Asymptotic size and a problem with subsampling and with the *m* out of *n* bootstrap. *Econometric Theory* **26** 426–468. MR2600570 https://doi.org/10.1017/S0266466609100051
[3] Athreya, K.B. (1987). Bootstrap of the mean in the infinite variance case. *Ann. Statist.* **15** 724–731. MR0888436 https://doi.org/10.1214/aos/1176350371
[4] Athreya, K.B. and Fukuchi, J. (1997). Confidence intervals for endpoints of a c.d.f. via bootstrap. *J. Statist. Plann. Inference* **58** 299–320. MR1450018 https://doi.org/10.1016/S0378-3758(96)00087-0
[5] Athreya, K.B., Lahiri, S.N. and Wu, W. (1998). Inference for heavy tailed distributions. *J. Statist. Plann. Inference* **66** 61–75. MR1616996 https://doi.org/10.1016/S0378-3758(97)00077-3
[6] Babu, G.J. (1984). Bootstrapping statistics with linear combinations of chi-squares as weak limit. *Sankhyā Ser. A* **46** 85–93. MR0768919
[7] Beran, R. (1982). Estimated sampling distributions: The bootstrap and competitors. *Ann. Statist.* **10** 212–225. MR0642733
[8] Beran, R. (1997). Diagnosing bootstrap success. *Ann. Inst. Statist. Math.* **49** 1–24. MR1450689 https://doi.org/10.1023/A:1003114420352
[9] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217. MR0630103
[10] Bickel, P.J., Götze, F. and van Zwet, W.R. (1997). Resampling fewer than *n* observations: Gaines, losses, and remedies for losses. *Statist. Sinica* **7** 1–31.
[11] Bickel, P.J. and Sakov, A. (2008). On the choice of *m* in the *m* out of *n* bootstrap and confidence bounds for extrema. *Statist. Sinica* **18** 967–985. MR2440400
[12] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396
[13] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1324786

1350-7265

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

[15] Cheung, K.Y., Lee, S.M.S. and Young, G.A. (2006). Stein confidence sets based on non-iterated and iterated parametric bootstraps. *Statist. Sinica* **16** 45–75. MR2256079

[16] Datta, S. (1995). On a modified bootstrap for certain asymptotically nonnormal statistics. *Statist. Probab. Lett.* **24** 91–98. MR1354361 https://doi.org/10.1016/0167-7152(94)00152-X

[17] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*. *Cambridge Series in Statistical and Probabilistic Mathematics* **1**. Cambridge: Cambridge Univ. Press. MR1478673 https://doi.org/10.1017/CBO9780511802843

[18] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26. MR0515681

[19] Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Canad. J. Statist.* **9** 139–172. MR0640014 https://doi.org/10.2307/3314608

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

[21] Fristedt, B. and Gray, L. (1997). *A Modern Approach to Probability Theory*. *Probability and Its Applications*. Boston, MA: Birkhäuser, Inc. MR1422917 https://doi.org/10.1007/978-1-4899-2837-5

[22] Ghosh, J.K. and Ramamoorthi, R.V. (2003). *Bayesian Nonparametrics*. *Springer Series in Statistics*. New York: Springer. MR1992245

[23] Giurcanu, M. and Presnell, B. (2018). Bootstrap inference for misspecified moment condition models. *Ann. Inst. Statist. Math.* **70** 605–630. MR3785709 https://doi.org/10.1007/s10463-017-0604-2

[24] Giurcanu, M. and Presnell, B. (2019). Bootstrapping LASSO-type estimators in regression models. *J. Statist. Plann. Inference* **199** 114–125. MR3857816 https://doi.org/10.1016/j.jspi.2018.05.007

[25] Giurcanu, M. and Presnell, B. (2023). Supplement A to "Bootstrap inference for a class of non-regular estimators." https://doi.org/10.3150/22-BEJ1538SUPPA, https://doi.org/10.3150/22-BEJ1538SUPPB

[26] Giurcanu, M. and Presnell, B. (2023). Supplement B to "Bootstrap inference for a class of non-regular estimators." https://doi.org/10.3150/22-BEJ1538SUPPB

[27] Giurcanu, M.C. (2012). Bootstrapping in non-regular smooth function models. *J. Multivariate Anal.* **111** 78–93. MR2944407 https://doi.org/10.1016/j.jmva.2012.04.016

[28] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. *Springer Series in Statistics*. New York: Springer. MR1145237 https://doi.org/10.1007/978-1-4612-4384-7

[29] Hall, P. and Presnell, B. (1999). Intentionally biased bootstrap methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 143–158. MR1664116 https://doi.org/10.1111/1467-9868.00168

[30] Kosorok, M.R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. *Springer Series in Statistics*. New York: Springer. MR2724368 https://doi.org/10.1007/978-0-387-74978-5

[31] Lindley, D.V. (1962). Discussion of "Confidence sets for the mean of a multivariate distribution" by C. Stein. *J. Roy. Statist. Soc. Ser. B* **24** 285–287.

[32] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18** 90–120. MR1041387 https://doi.org/10.1214/aos/1176347494

[33] Owen, A. (2001). *Empirical Likelihood*. Boca Raton, FL: CRC Press / CRC.

[34] 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

[35] Politis, D.N. and Romano, J.P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann. Statist.* **22** 2031–2050. MR1329181 https://doi.org/10.1214/aos/1176325770

[36] Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*. *Springer Series in Statistics*. New York: Springer. MR1707286 https://doi.org/10.1007/978-1-4612-1554-7

[37] Pötscher, B.M. and Leeb, H. (2009). On the distribution of penalized maximum likelihood estimators: The LASSO, SCAD, and thresholding. *J. Multivariate Anal.* **100** 2065–2082. MR2543087 https://doi.org/10.1016/j.jmva.2009.06.010

[38] Putter, H. and van Zwet, W.R. (1996). Resampling: Consistency of substitution estimators. *Ann. Statist.* **24** 2297–2318. MR1425955 https://doi.org/10.1214/aos/1032181156

[39] Read, T.R.C. and Cressie, N.A.C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. *Springer Series in Statistics*. New York: Springer. MR0955054 https://doi.org/10.1007/978-1-4612-4578-0

[40] Samworth, R. (2003). A note on methods of restoring consistency to the bootstrap. *Biometrika* **90** 985–990. MR2024773 https://doi.org/10.1093/biomet/90.4.985

[41] Shao, J. and Tu, D.S. (1995). *The Jackknife and Bootstrap*. *Springer Series in Statistics*. New York: Springer. MR1351010 https://doi.org/10.1007/978-1-4612-0795-5

[42] Sobel, M.E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociol. Method.* **13** 290–312.

[43] R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

[44] 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

[45] Varadarajan, V.S. (1963). Groups of automorphisms of Borel spaces. *Trans. Amer. Math. Soc.* **109** 191–220. MR0159923 https://doi.org/10.2307/1993903

[46] Wu, C.F.J. (1990). On the asymptotic properties of the jackknife histogram. *Ann. Statist.* **18** 1438–1452. MR1062718 https://doi.org/10.1214/aos/1176347759

[47] Zhou, M. and Yang, Y. (2015). Emplik: Empirical Likelihood Ratio for Censored/Truncated Data. R package version 0.9-9-6.

[48] Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Amer. Statist. Assoc.* **101** 1418–1429. MR2279469 https://doi.org/10.1198/016214506000000735

Bernoulli **29**(3), 2023, 2219–2246
https://doi.org/10.3150/22-BEJ1539

Moments and tails of hitting times of Bessel processes and convolutions of elementary mixtures of exponential distributions

WITOLD M. BEDNORZ[1,a] and RAFAŁ M. ŁOCHOWSKI[2,b]

[1] *University of Warsaw, Warsaw, Poland,* [a]*wbednorz@mimuw.edu.pl*
[2] *Warsaw School of Economics, Warsaw, Poland,* [b]*rlocho@sgh.waw.pl*

We present explicit estimates of right and left tails and exact (up to universal, multiplicative constants) estimates of tails and moments of hitting times of Bessel processes. The latter estimates are obtained from more general estimates of moments and tails established for convolutions of elementary mixtures of exponential distributions.

Keywords: Bessel processes; hitting times; elementary mixtures of exponential distributions; moments; tails

References

[1] Breen, S. (1995). Uniform upper and lower bounds on the zeros of Bessel functions of the first kind. *J. Math. Anal. Appl.* **196** 1–17. MR1359929 https://doi.org/10.1006/jmaa.1995.1395
[2] Byczkowski, T., Graczyk, P. and Stós, A. (2007). Poisson kernels of half-spaces in real hyperbolic spaces. *Rev. Mat. Iberoam.* **23** 85–126. MR2351127 https://doi.org/10.4171/RMI/487
[3] Byczkowski, T., Małecki, J. and Ryznar, M. (2013). Hitting times of Bessel processes. *Potential Anal.* **38** 753–786. MR3034599 https://doi.org/10.1007/s11118-012-9296-7
[4] Byczkowski, T. and Ryznar, M. (2006). Hitting distributions of geometric Brownian motion. *Studia Math.* **173** 19–38. MR2204460 https://doi.org/10.4064/sm173-1-2
[5] Ciesielski, Z. and Taylor, S.J. (1962). First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path. *Trans. Amer. Math. Soc.* **103** 434–450. MR0143257 https://doi.org/10.2307/1993838
[6] DeBlassie, R.D. (1987). Stopping times of Bessel processes. *Ann. Probab.* **15** 1044–1051. MR0893912
[7] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics (New York)* **38**. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
[8] Gluskin, E.D. and Kwapień, S. (1995). Tail and moment estimates for sums of independent random variables with logarithmically concave tails. *Studia Math.* **114** 303–309. MR1338834 https://doi.org/10.4064/sm-114-3-303-309
[9] Grzywny, T., Ryznar, M. and Trojan, B. (2019). Asymptotic behaviour and estimates of slowly varying convolution semigroups. *Int. Math. Res. Not. IMRN* **23** 7193–7258. MR4039012 https://doi.org/10.1093/imrn/rnx324
[10] Hamana, Y. and Matsumoto, H. (2012). The probability densities of the first hitting times of Bessel processes. *J. Math-for-Ind.* **4B** 91–95. MR3072321
[11] Hamana, Y. and Matsumoto, H. (2013). The probability distributions of the first hitting times of Bessel processes. *Trans. Amer. Math. Soc.* **365** 5237–5257. MR3074372 https://doi.org/10.1090/S0002-9947-2013-05799-6
[12] Jakubowski, J. and Wiśniewolski, M. (2018). Invariance formulas for stopping times of squared Bessel process. *Stoch. Anal. Appl.* **36** 671–699. MR3854537 https://doi.org/10.1080/07362994.2018.1443014
[13] Jedidi, W. and Simon, T. (2015). Diffusion hitting times and the bell-shape. *Statist. Probab. Lett.* **102** 38–41. MR3343781 https://doi.org/10.1016/j.spl.2015.03.008

[14] Kent, J.T. (1980). Eigenvalue expansions for diffusion hitting times. *Z. Wahrsch. Verw. Gebiete* **52** 309–319. MR0576891 https://doi.org/10.1007/BF00538895

[15] Latała, R. (1997). Estimation of moments of sums of independent real random variables. *Ann. Probab.* **25** 1502–1513. MR1457628 https://doi.org/10.1214/aop/1024404522

[16] Lévy, P. (1953). La mesure de Hausdorff de la courbe du mouvement brownien. *G. Ist. Ital. Attuari* **16** 1–37. MR0064344

[17] Mörters, P. and Peres, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge: Cambridge Univ. Press. MR2604525 https://doi.org/10.1017/CBO9780511750489

[18] Piessens, R. (1984). A series expansion for the first positive zero of the Bessel functions. *Math. Comp.* **42** 195–197. MR0725995 https://doi.org/10.2307/2007570

[19] Revuz, D. and Yor, M. (2005). *Continuous Martingales and Brownian Motion* 3rd ed. *Grundlehren der Mathematischen Wissenschaften* **293**. Springer-Verlag, Berlin. MR1083357

[20] Rösler, U. (1980). Unimodality of passage times for one-dimensional strong Markov processes. *Ann. Probab.* **8** 853–859. MR0577322

[21] Serafin, G. (2017). Exit times densities of the Bessel process. *Proc. Amer. Math. Soc.* **145** 3165–3178. MR3637962 https://doi.org/10.1090/proc/13419

[22] Watson, G.N. (1944). *A Treatise on the Theory of Bessel Functions*. Cambridge: Cambridge Univ. Press. MR0010746

[23] Yamazato, M. (1990). Hitting time distributions of single points for 1-dimensional generalized diffusion processes. *Nagoya Math. J.* **119** 143–172. MR1071905 https://doi.org/10.1017/S0027763000003172

Bernoulli **29**(3), 2023, 2247–2271
https://doi.org/10.3150/22-BEJ1540

Linear functional estimation under multiplicative measurement error

SERGIO BRENNER MIGUEL[1,a], FABIENNE COMTE[2,c] and JAN JOHANNES[1,b]

[1]*Ruprecht-Karls-Universität Heidelberg, Institut für Angewandte Mathematik und Interdisciplinary Center for Scientific Computing (IWR), Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany,* [a]*brennermiguel@math.uni-heidelberg.de,* [b]*johannes@math.uni-heidelberg.de*
[2]*Université Paris Cité, CNRS, MAP5, F-75006 Paris, France,* [c]*fabienne.comte@parisdescartes.fr*

We study the non-parametric estimation of the value of a linear functional evaluated at an unknown density function f with support on $\mathbb{R}_+$ based on an i.i.d. sample with multiplicative measurement errors. The proposed plug-in estimation procedure combines the estimation of the Mellin transform of the density f and a regularisation of the inverse of the Mellin transform by a spectral cut-off. The attainable accuracy of the estimator is essentially determined by the decay of the upcoming Mellin transforms and the smoothness of the linear functional which we illustrate by different scenarios. As usual the choice of the cut-off parameter is crucial and we propose its data-driven selection inspired by the work of (Goldenshluger and Lepski *Ann. Statist.* **39** (2011) 1608–1632). By proving matching lower bounds we show that the plug-in estimator with optimally chosen cut-off parameter attains minimax-optimal rates of convergence over *Mellin-Sobolev spaces*. Furthermore the rate of convergence of the data-driven estimator features at most a deterioration by a logarithmic factor which is widely considered as an acceptable price for adaptation. In particular, our theory covers point-wise estimation of the density f, its derivative and Laplace transform, its associated survival and cumulative distribution function as well as the point-wise estimation of the mean residual life.

Keywords: Linear functional model; multiplicative measurement error; Mellin transform; Mellin-Sobolev space; minimax theory; inverse problem; adaptation

References

[1] Andersen, K.E. and Hansen, M.B. (2001). Multiplicative censoring: Density estimation by a series expansion approach. *J. Statist. Plann. Inference* **98** 137–155. MR1860231 https://doi.org/10.1016/S0378-3758(00)00237-8

[2] Asgharian, M. and Wolfson, D.B. (2005). Asymptotic behavior of the unconditional NPMLE of the length-biased survivor function from right censored prevalent cohort data. *Ann. Statist.* **33** 2109–2131. MR2211081 https://doi.org/10.1214/009053605000000372

[3] Belomestny, D., Comte, F. and Genon-Catalot, V. (2016). Nonparametric Laguerre estimation in the multiplicative censoring model. *Electron. J. Stat.* **10** 3114–3152. MR3571964 https://doi.org/10.1214/16-EJS1203

[4] Belomestny, D. and Goldenshluger, A. (2020). Nonparametric density estimation from observations with multiplicative measurement errors. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 36–67. MR4058981 https://doi.org/10.1214/18-AIHP954

[5] Birgé, L. and Massart, P. (1998). Minimum contrast estimators on sieves: Exponential bounds and rates of convergence. *Bernoulli* **4** 329–375. MR1653272 https://doi.org/10.2307/3318720

[6] Brenner Miguel, S. (2022). Anisotropic spectral cut-off estimation under multiplicative measurement errors. *J. Multivariate Anal.* **190** Paper No. 104990, 18. MR4396579 https://doi.org/10.1016/j.jmva.2022.104990

[7] Brenner Miguel, S., Comte, F. and Johannes, J. (2021). Spectral cut-off regularisation for density estimation under multiplicative measurement errors. *Electron. J. Stat.* **15** 3551–3573. MR4298979 https://doi.org/10.1214/21-ejs1870

1350-7265

[8] Brenner Miguel, S. and Phandoidaen, N. (2022). Multiplicative deconvolution in survival analysis under dependency. *Statistics* **56** 297–328. MR4415439 https://doi.org/10.1080/02331888.2022.2058503
[9] Brenner Miguel, S. (2021). Multiplicative deconvolution estimator based on a ridge approach. Preprint. Available at arXiv:2108.01523.
[10] Brown, L.D. and Low, M.G. (1996). A constrained risk inequality with applications to nonparametric functional estimation. *Ann. Statist.* **24** 2524–2535. MR1425965 https://doi.org/10.1214/aos/1032181166
[11] Brunel, E., Comte, F. and Genon-Catalot, V. (2016). Nonparametric density and survival function estimation in the multiplicative censoring model. *TEST* **25** 570–590. MR3531843 https://doi.org/10.1007/s11749-016-0479-1
[12] Butucea, C. and Comte, F. (2009). Adaptive estimation of linear functionals in the convolution model and applications. *Bernoulli* **15** 69–98. MR2546799 https://doi.org/10.3150/08-BEJ146
[13] Butucea, C. and Tsybakov, A.B. (2007). Sharp optimality in density deconvolution with dominating bias. I. *Teor. Veroyatn. Primen.* **52** 111–128. MR2354572 https://doi.org/10.1137/S0040585X97982840
[14] Comte, F. (2017). *Nonparametric Estimation*. *Master and Research*. Paris: Spartacus-Idh.
[15] Comte, F. and Dion, C. (2016). Nonparametric estimation in a multiplicative censoring model with symmetric noise. *J. Nonparametr. Stat.* **28** 768–801. MR3555457 https://doi.org/10.1080/10485252.2016.1225737
[16] Donoho, D.L. (1994). Statistical estimation and optimal recovery. *Ann. Statist.* **22** 238–270. MR1272082 https://doi.org/10.1214/aos/1176325367
[17] Donoho, D.L. and Low, M.G. (1992). Renormalization exponents and optimal pointwise rates of convergence. *Ann. Statist.* **20** 944–970. MR1165601 https://doi.org/10.1214/aos/1176348665
[18] Engl, H.W., Hanke, M. and Neubauer, A. (1996). *Regularization of Inverse Problems*. *Mathematics and Its Applications* **375**. Dordrecht: Kluwer Academic. MR1408680
[19] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19** 1257–1272. MR1126324 https://doi.org/10.1214/aos/1176348248
[20] Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39** 1608–1632. MR2850214 https://doi.org/10.1214/11-AOS883
[21] Goldenshluger, A. and Pereverzev, S.V. (2000). Adaptive estimation of linear functionals in Hilbert scales from indirect white noise observations. *Probab. Theory Related Fields* **118** 169–186. MR1790080 https://doi.org/10.1007/s440-000-8013-3
[22] Guess, F. and Proschan, F. (1988). 12 mean residual life: Theory and applications. *Handb. Statist.* **7** 215–224.
[23] Hall, P. and Meister, A. (2007). A ridge-parameter approach to deconvolution. *Ann. Statist.* **35** 1535–1558. MR2351096 https://doi.org/10.1214/009053607000000028
[24] Ibragimov, I.A. and Khas'minskiĭ, R.Z. (1984). Nonparametric estimation of the value of a linear functional in Gaussian white noise. *Teor. Veroyatn. Primen.* **29** 19–32. MR0739497
[25] Kappus, J. (2014). Adaptive nonparametric estimation for Lévy processes observed at low frequency. *Stochastic Process. Appl.* **124** 730–758. MR3131312 https://doi.org/10.1016/j.spa.2013.08.010
[26] Laurent, B., Ludeña, C. and Prieur, C. (2008). Adaptive estimation of linear functionals by model selection. *Electron. J. Stat.* **2** 993–1020. MR2448602 https://doi.org/10.1214/07-EJS127
[27] Li, K.C. (1982). Minimaxity of the method of regularization on stochastic processes. *Ann. Statist.* **10** 937–942. MR0663444
[28] Mabon, G. (2016). Adaptive deconvolution of linear functionals on the nonnegative real line. *J. Statist. Plann. Inference* **178** 1–23. MR3531524 https://doi.org/10.1016/j.jspi.2016.04.006
[29] Meister, A. (2009). *Deconvolution Problems in Nonparametric Statistics*. *Lecture Notes in Statistics* **193**. Berlin: Springer. MR2768576 https://doi.org/10.1007/978-3-540-87557-4
[30] Paris, R.B. and Kaminski, D. (2001). *Asymptotics and Mellin-Barnes Integrals*. *Encyclopedia of Mathematics and Its Applications* **85**. Cambridge: Cambridge Univ. Press. MR1854469 https://doi.org/10.1017/CBO9780511546662
[31] Pensky, M. (2017). Minimax theory of estimation of linear functionals of the deconvolution density with or without sparsity. *Ann. Statist.* **45** 1516–1541. MR3670187 https://doi.org/10.1214/16-AOS1498
[32] Petrov, V.V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*. *Oxford Studies in Probability* **4**. New York: The Clarendon Press. MR1353441

[33] Speckman, P. (1979). Minimax Estimation of Linear Functionals in a Hilbert space. Unpublished manuscript.

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

[35] van Es, B., Klaassen, C.A.J. and Oudshoorn, K. (2000). Survival analysis under cross-sectional sampling: Length bias and multiplicative censoring. *J. Statist. Plann. Inference* **91** 295–312. MR1814785 https://doi.org/10.1016/S0378-3758(00)00183-X

[36] Vardi, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: Nonparametric estimation. *Biometrika* **76** 751–761. MR1041420 https://doi.org/10.1093/biomet/76.4.751

[37] Vardi, Y. and Zhang, C.-H. (1992). Large sample study of empirical distributions in a random-multiplicative censoring model. *Ann. Statist.* **20** 1022–1039. MR1165604 https://doi.org/10.1214/aos/1176348668

Bernoulli **29**(3), 2023, 2272–2294
https://doi.org/10.3150/22-BEJ1541

Compound Poisson disorder problem with uniformly distributed disorder time

CAGIN URU[1,a], SAVAS DAYANIK[2,b] and SEMIH O. SEZER[3,c]

[1]*Duke University, Durham, NC, USA,* [a]*cagin.uru@duke.edu*
[2]*Bilkent University, Ankara, Turkey,* [b]*sdayanik@bilkent.edu.tr*
[3]*Sabancı University, Istanbul, Turkey,* [c]*sezer@sabanciuniv.edu*

Suppose that the arrival rate and the jump distribution of a compound Poisson process change suddenly at an unknown and unobservable time. We want to detect the change as quickly as possible to take counteractions, e.g., to assure top quality of products in a production system, or to stop credit card fraud in a banking system. If we have no prior information about future disorder time, then we typically assume that the disorder is equally likely to happen any time – or has uniform distribution – over a long but finite time horizon. We solve this so-called compound Poisson disorder problem for the practically important case of unknown, unobserved, but uniformly distributed disorder time. The solution hinges on the complete separation of information flow from the hard time horizon constraint, by describing the former with an autonomous time-homogeneous one-dimensional Markov process in terms of which the detection problem translates into a finite horizon optimal stopping problem. For any given finite horizon, the solution is two-dimensional. For cases where the horizon is large and one is unwilling to set a fixed value for it, we give a one-dimensional approximation. Also, we discuss an extension where the disorder may not happen on the given interval with a positive probability. In this extended model, if no detection decision is made by the end of the horizon, then a second level hypothesis testing problem is solved to determine the local parameters of the observed process.

Keywords: Compound Poisson process; optimal stopping; Poisson disorder problem; quickest detection

References

Arslan, A.M., Frenk, J.B.G. and Sezer, S.O. (2015). On the single-leg airline revenue management problem in continuous time. *Math. Methods Oper. Res.* **81** 27–52. MR3312429 https://doi.org/10.1007/s00186-014-0485-6

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

Bayraktar, E., Dayanik, S. and Karatzas, I. (2005). The standard Poisson disorder problem revisited. *Stochastic Process. Appl.* **115** 1437–1450. MR2158013 https://doi.org/10.1016/j.spa.2005.04.011

Bayraktar, E., Dayanik, S. and Karatzas, I. (2006). Adaptive Poisson disorder problem. *Ann. Appl. Probab.* **16** 1190–1261. MR2260062 https://doi.org/10.1214/105051606000000312

Bayraktar, E. and Sezer, S. (2009). Online change detection for a Poisson process with a phase-type change-time prior distribution. *Sequential Anal.* **28** 218–250. MR2518831 https://doi.org/10.1080/07474940902816833

Brown, M. (2008). Bayesian detection of changes of a Poisson process monitored at discrete time points where the arrival rates are unknown. *Sequential Anal.* **27** 68–77. MR2400238 https://doi.org/10.1080/07474940701801994

Brown, M. (2016). Detecting changes in a Poisson process monitored at random time intervals. *Sequential Anal.* **35** 358–369. MR3547972 https://doi.org/10.1080/07474946.2016.1206381

Buonaguidi, B., Mira, A., Bucheli, H. and Vitanis, V. (2022). Bayesian quickest detection of credit card fraud. *Bayesian Anal.* **17** 261–290. MR4377143 https://doi.org/10.1214/20-ba1254

Davis, M. (1976). A note on the Poisson disorder problem. *Banach Center Publ.* **1** 65–72.

Davis, M.H.A. (1993). *Markov Models and Optimization*. *Monographs on Statistics and Applied Probability* **49**. London: CRC Press. MR1283589 https://doi.org/10.1201/9780203748039

1350-7265

Dayanik, S. and Parlar, M. (2013). Dynamic bidding strategies in search-based advertising. *Ann. Oper. Res.* **211** 103–136. MR3151429 https://doi.org/10.1007/s10479-013-1427-z

Dayanik, S., Poor, H.V. and Sezer, S.O. (2008). Multisource Bayesian sequential change detection. *Ann. Appl. Probab.* **18** 552–590. MR2399705 https://doi.org/10.1214/07-AAP463

Dayanik, S. and Sezer, S.O. (2006a). Compound Poisson disorder problem. *Math. Oper. Res.* **31** 649–672. MR2281222 https://doi.org/10.1287/moor.1060.0223

Dayanik, S. and Sezer, S.O. (2006b). Sequential testing of simple hypotheses about compound Poisson processes. *Stochastic Process. Appl.* **116** 1892–1919. MR2307064 https://doi.org/10.1016/j.spa.2006.05.004

Ekström, E. and Wang, Y. (2022). Multi-dimensional sequential testing and detection. *Stochastics* **94** 789–806. MR4445993 https://doi.org/10.1080/17442508.2021.1993852

El Karoui, N., Loisel, S. and Salhi, Y. (2017). Minimax optimality in robust detection of a disorder time in doubly-stochastic Poisson processes. *Ann. Appl. Probab.* **27** 2515–2538. MR3693532 https://doi.org/10.1214/16-AAP1266

Figueroa-López, J.E. and Ólafsson, S. (2019). Change-point detection for Lévy processes. *Ann. Appl. Probab.* **29** 717–738. MR3910015 https://doi.org/10.1214/17-AAP1368

Gal'chuk, L.I. and Rozovskii, B.L. (1971). The "disorder" problem for a Poisson process. *Theory Probab. Appl.* **16** 712–716.

Gapeev, P.V. (2005). The disorder problem for compound Poisson processes with exponential jumps. *Ann. Appl. Probab.* **15** 487–499. MR2115049 https://doi.org/10.1214/105051604000000981

Gugerli, U.S. (1986). Optimal stopping of a piecewise-deterministic Markov process. *Stochastics* **19** 221–236. MR0872462 https://doi.org/10.1080/17442508608833426

Herberts, T. and Jensen, U. (2004). Optimal detection of a change point in a Poisson process for different observation schemes. *Scand. J. Stat.* **31** 347–366. MR2087830 https://doi.org/10.1111/j.1467-9469.2004.02-102.x

Kilic, K., Saygi, M.G. and Sezer, S.O. (2017). A mathematical model for personalized advertisement in virtual reality environments. *Math. Methods Oper. Res.* **85** 241–264. MR3671557 https://doi.org/10.1007/s00186-016-0567-8

Ludkovski, M. and Sezer, S.O. (2012). Finite horizon decision timing with partially observable Poisson processes. *Stoch. Models* **28** 207–247. MR2926586 https://doi.org/10.1080/15326349.2012.672143

Peskir, G. and Shiryaev, A.N. (2002). Solving the Poisson disorder problem. In *Advances in Finance and Stochastics* 295–312. Berlin: Springer. MR1929384

Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary Problems*. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2256030

Poor, H.V. and Hadjiliadis, O. (2009). *Quickest Detection*. Cambridge: Cambridge Univ. Press. MR2482527

Shiryaev, A.N. (1996). *Probability*, 2nd ed. *Graduate Texts in Mathematics* **95**. New York: Springer. MR1368405 https://doi.org/10.1007/978-1-4757-2539-1

Shiryaev, A.N. (2008). *Optimal Stopping Rules*. *Stochastic Modelling and Applied Probability* **8**. Berlin: Springer. MR2374974

Shiryaev, A.N. (2019). *Stochastic Disorder Problems*. *Probability Theory and Stochastic Modelling* **93**. Cham: Springer. MR3890022 https://doi.org/10.1007/978-3-030-01526-8

Shiryaev, A.N., Zhitlukhin, M.V. and Ziemba, W.T. (2014). When to sell Apple and the NASDAQ? Trading bubbles with a stochastic disorder model. *J. Portf. Manag.* **40** 54–63.

Shiryaev, A.N., Zhitlukhin, M.V. and Ziemba, W.T. (2015). Land and stock bubbles, crashes and exit strategies in Japan circa 1990 and in 2013. *Quant. Finance* **15** 1449–1469. MR3378085 https://doi.org/10.1080/14697688.2014.989897

Sokko, A.A. (2015). Disorder problem for a Brownian motion on a segment in the case of uniformly distributed moment of disorder. *Moscow Univ. Math. Bull.* **70** 103–110. MR3401232 https://doi.org/10.3103/S0027132215030018

Tartakovsky, A., Nikiforov, I. and Basseville, M. (2015). *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. *Monographs on Statistics and Applied Probability* **136**. Boca Raton, FL: CRC Press. MR3241619

Ürü, Ç. (2019). Compound Poisson disorder problem with uniformly distributed disorder time. Master's thesis, Bilkent Univ.

Zhitlukhin, M.V. and Shiryaev, A.N. (2013). Bayesian disorder problems on filtered probability spaces. *Theory Probab. Appl.* **57** 497–511. MR3196783 https://doi.org/10.1137/S0040585X97986072

Zhitlukhin, M.V. and Shiryaev, A.N. (2014). Optimal stopping problems for a Brownian motion with disorder on a segment. *Theory Probab. Appl.* **58** 164–171. MR3267290 https://doi.org/10.1137/S0040585X97986448

Bernoulli **29**(3), 2023, 2295–2317
https://doi.org/10.3150/22-BEJ1542

Harmonic analysis meets stationarity: A general framework for series expansions of special Gaussian processes

MOHAMED NDAOUD[a]

Department of Decisions Sciences, ESSEC Business School 95000 Cergy, France, [a]ndaoud@essec.edu

In this paper, we present a new approach to derive series expansions for some Gaussian processes based on harmonic analysis of their covariance function. In particular, we propose a new simple rate-optimal series expansion for fractional Brownian motion. The convergence of the latter series holds in mean square and uniformly almost surely, with a rate-optimal decay of the remainder of the series. We also develop a general framework of convergent series expansions for certain classes of Gaussian processes with stationarity. Finally, an application to optimal functional quantization is described.

Keywords: Fractional Brownian motion; fractional Ornstein-Uhlenbeck; Karhunen-Loève decomposition; Fourier series; functional quantization

References

Ayache, A. and Taqqu, M.S. (2003). Rate optimality of wavelet series approximations of fractional Brownian motion. *J. Fourier Anal. Appl.* **9** 451–471. MR2027888 https://doi.org/10.1007/s00041-003-0022-0

Bronski, J.C. (2003a). Asymptotics of Karhunen-Loeve eigenvalues and tight constants for probability distributions of passive scalar transport. *Comm. Math. Phys.* **238** 563–582. MR1993386 https://doi.org/10.1007/s00220-003-0835-3

Bronski, J.C. (2003b). Small ball constants and tight eigenvalue asymptotics for fractional Brownian motions. *J. Theoret. Probab.* **16** 87–100. MR1956822 https://doi.org/10.1023/A:1022226420564

Chigansky, P. and Kleptsyna, M. (2018). Exact asymptotics in eigenproblems for fractional Brownian covariance operators. *Stochastic Process. Appl.* **128** 2007–2059. MR3797652 https://doi.org/10.1016/j.spa.2017.08.019

Corlay, S. et al. (2010). Functional quantization-based stratified sampling methods. ArXiv Preprint. arXiv:1008.4441.

Deheuvels, P. (2007). A Karhunen-Loève expansion for a mean-centered Brownian bridge. *Statist. Probab. Lett.* **77** 1190–1200. MR2392790 https://doi.org/10.1016/j.spl.2007.03.011

Dietrich, C.R. and Newsam, G.N. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* **18** 1088–1107. MR1453559 https://doi.org/10.1137/S1064827592240555

Dzhaparidze, K. and van Zanten, H. (2004). A series expansion of fractional Brownian motion. *Probab. Theory Related Fields* **130** 39–55. MR2092872 https://doi.org/10.1007/s00440-003-0310-2

Iglói, E. (2005). A rate-optimal trigonometric series expansion of the fractional Brownian motion. *Electron. J. Probab.* **10** 1381–1397. MR2183006 https://doi.org/10.1214/EJP.v10-287

Itô, K. and Nisio, M. (1968). On the convergence of sums of independent Banach space valued random variables. *Osaka Math. J.* **5** 35–48. MR0235593

Junglen, S. and Luschgy, H. (2010). A constructive sharp approach to functional quantization of stochastic processes. *J. Appl. Math.* 378519. MR2739750 https://doi.org/10.1155/2010/378519

Kühn, T. and Linde, W. (2002). Optimal series representation of fractional Brownian sheets. *Bernoulli* **8** 669–696. MR1935652

1350-7265

León, J.A. (2020). Stratonovich type integration with respect to fractional Brownian motion with Hurst parameter less than $1/2$. *Bernoulli* **26** 2436–2462. MR4091115 https://doi.org/10.3150/20-BEJ1202

Li, Y. and Dai, H. (2011). Approximations of fractional Brownian motion. *Bernoulli* **17** 1195–1216. MR2854769 https://doi.org/10.3150/10-BEJ319

Luschgy, H. and Pagès, G. (2002). Functional quantization of Gaussian processes. *J. Funct. Anal.* **196** 486–531. MR1943099 https://doi.org/10.1016/S0022-1236(02)00010-1

Luschgy, H. and Pagès, G. (2004). Sharp asymptotics of the functional quantization problem for Gaussian processes. *Ann. Probab.* **32** 1574–1599. MR2060310 https://doi.org/10.1214/009117904000000324

Luschgy, H. and Pagès, G. (2007). High-resolution product quantization for Gaussian processes under sup-norm distortion. *Bernoulli* **13** 653–671. MR2348745 https://doi.org/10.3150/07-BEJ6025

Luschgy, H. and Pagès, G. (2009). Expansions for Gaussian processes and Parseval frames. *Electron. J. Probab.* **14** 1198–1221. MR2511282 https://doi.org/10.1214/EJP.v14-649

Nazarov, A.I. and Nikitin, Y.Y. (2005). Logarithmic L2-small ball asymptotics for some fractional Gaussian processes. *Theory Probab. Appl.* **49** 645–658.

Pagès, G. and Printems, J. (2005). Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods Appl.* **11** 407–446. MR2186817 https://doi.org/10.1163/156939605777438578

Wang, H.T. (1985). Convex functions and Fourier coefficients. *Proc. Amer. Math. Soc.* **94** 641–646. MR0792276 https://doi.org/10.2307/2044879

Bernoulli **29**(3), 2023, 2318–2341
https://doi.org/10.3150/22-BEJ1543

Density estimation under local differential privacy and Hellinger loss

MATHIEU SART[a]

Université Jean Monnet Saint-Étienne, CNRS, Institut Camille Jordan UMR 5208, F-42023, Saint-Étienne, France, [a]*mathieu.sart@univ-st-etienne.fr*

In this paper, we carry out a piecewise constant estimator of the density for privatised data. We establish a non-asymptotic oracle inequality for the Hellinger loss and deduce that our estimator is adaptive and rate optimal over a wide range of Besov classes (up to possible logarithmic factors). We also get better estimation rates when the density is not only in a Besov class but also bounded away from 0. These rates are optimal within possible log factors. This result is in contrast to what happens with the $\mathbb{L}^2$ loss where the privatised minimax rates over Besov classes can be improved in some cases by assuming the target bounded from above.

Keywords: Besov spaces; density estimation; local differential privacy; minimax estimation

References

[1] Akakpo, N. and Lacour, C. (2011). Inhomogeneous and anisotropic conditional density estimation from dependent data. *Electron. J. Stat.* **5** 1618–1653. MR2870146 https://doi.org/10.1214/11-EJS653

[2] Baraud, Y. (2011). Estimator selection with respect to Hellinger-type risks. *Probab. Theory Related Fields* **151** 353–401. MR2834722 https://doi.org/10.1007/s00440-010-0302-y

[3] Baraud, Y. and Birgé, L. (2009). Estimating the intensity of a random measure by histogram type estimators. *Probab. Theory Related Fields* **143** 239–284. MR2449129 https://doi.org/10.1007/s00440-007-0126-6

[4] Barron, A., Birgé, L. and Massart, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113** 301–413. MR1679028 https://doi.org/10.1007/s004400050210

[5] Berrett, T.B., Györfi, L. and Walk, H. (2021). Strongly universally consistent nonparametric regression and classification with privatised data. *Electron. J. Stat.* **15** 2430–2453. MR4255329 https://doi.org/10.1214/21-ejs1845

[6] Birgé, L. (2006). Model selection via testing: An alternative to (penalized) maximum likelihood estimators. *Ann. Inst. Henri Poincaré Probab. Stat.* **42** 273–325. MR2219712 https://doi.org/10.1016/j.anihpb.2005.04.004

[7] Birgé, L. (2014). Model selection for density estimation with $\mathbb{L}_2$-loss. *Probab. Theory Related Fields* **158** 533–574. MR3176358 https://doi.org/10.1007/s00440-013-0488-x

[8] Blanchard, G., Schäfer, C. and Rozenholc, Y. (2004). Oracle bounds and exact algorithm for dyadic classification trees. In *Learning Theory*. *Lecture Notes in Computer Science* **3120** 378–392. Berlin: Springer. MR2177922 https://doi.org/10.1007/978-3-540-27819-1_26

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

[10] Bousquet, O., Boucheron, S. and Lugosi, G. (2003). Introduction to statistical learning theory. In *Summer School on Machine Learning* 169–207. Springer.

[11] Butucea, C., Dubois, A., Kroll, M. and Saumard, A. (2020). Local differential privacy: Elbow effect in optimal density estimation and adaptation over Besov ellipsoids. *Bernoulli* **26** 1727–1764. MR4091090 https://doi.org/10.3150/19-BEJ1165

[12] Butucea, C. and Issartel, Y. (2021). Locally differentially private estimation of functionals of discrete distributions. *Adv. Neural Inf. Process. Syst.* **34**.

1350-7265

[13] Butucea, C., Rohde, A. and Steinberger, L. (2020). Interactive versus non-interactive locally differentially private estimation: Two elbows for the quadratic functional. ArXiv Preprint.
[14] Castellan, G. (1999). Modified Akaike's criterion for histogram density estimation. *Technical Report*.
[15] DeVore, R.A. and Lorentz, G.G. (1993). *Constructive Approximation*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **303**. Berlin: Springer. MR1261635
[16] DeVore, R.A. and Yu, X.M. (1990). Degree of adaptive approximation. *Math. Comp.* **55** 625–635. MR1035930 https://doi.org/10.2307/2008437
[17] Devroye, L. and Lugosi, G. (2001). *Combinatorial Methods in Density Estimation*. *Springer Series in Statistics*. New York: Springer. MR1843146 https://doi.org/10.1007/978-1-4613-0125-7
[18] Dubois, A., Berrett, T. and Butucea, C. (2021). Goodness-of-fit testing for Hölder continuous densities under local differential privacy. ArXiv preprint.
[19] Duchi, J.C., Jordan, M.I. and Wainwright, M.J. (2018). Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* **113** 182–201. MR3803452 https://doi.org/10.1080/01621459.2017.1389735
[20] Györfi, L. and Kroll, M. (2021). Multivariate density estimation from privatised data: Universal consistency and minimax rates. *ArXiv Preprint*.
[21] Hall, R., Rinaldo, A. and Wasserman, L. (2013). Differential privacy for functions and functional data. *J. Mach. Learn. Res.* **14** 703–727. MR3033345
[22] Kroll, M. (2020). Adaptive spectral density estimation by model selection under local differential privacy. ArXiv preprint.
[23] Kroll, M. (2021). On density estimation at a fixed point under local differential privacy. *Electron. J. Stat.* **15** 1783–1813. MR4255316 https://doi.org/10.1214/21-ejs1830
[24] Lam-Weil, J., Laurent, B. and Loubes, J.-M. (2022). Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint. *Bernoulli* **28** 579–600. MR4337717 https://doi.org/10.3150/21-bej1358
[25] Lepskiĭ, O.V. (1991). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36** 682–697.
[26] Massart, P. (2007). *Concentration Inequalities and Model Selection*. *Lecture Notes in Math.* **1896**. Berlin: Springer. MR2319879
[27] Rohde, A. and Steinberger, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.* **48** 2646–2670. MR4152116 https://doi.org/10.1214/19-AOS1901
[28] Sart, M. (2014). Estimation of the transition density of a Markov chain. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 1028–1068. MR3224298 https://doi.org/10.1214/13-AIHP551
[29] Sart, M. (2021). Minimax bounds for Besov classes in density estimation. *Electron. J. Stat.* **15** 3184–3216. MR4280167 https://doi.org/10.1214/21-ejs1856
[30] Sart, M. (2021). Estimating a density, a hazard rate, and a transition intensity via the ρ-estimation method. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 195–249. MR4255173 https://doi.org/10.1214/20-aihp1076
[31] Sart, M. (2023). Supplement to "Density estimation under local differential privacy and Hellinger loss." https://doi.org/10.3150/22-BEJ1543SUPP
[32] Wasserman, L. and Zhou, S. (2010). A statistical framework for differential privacy. *J. Amer. Statist. Assoc.* **105** 375–389. MR2656057 https://doi.org/10.1198/jasa.2009.tm08651

Bernoulli **29**(3), 2023, 2342–2366
https://doi.org/10.3150/22-BEJ1544

Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation

TEPPEI OGIHARA[1,a] and YUMA UEHARA[2,b]

[1]*Graduate School of Information Science and Technology, University of Tokyo, Tokyo, Japan,*
[a]*ogihara@mist.i.u-tokyo.ac.jp*
[2]*Department of Mathematics, Faculty of Engineering Science, Kansai University, Osaka, Japan,*
[b]*y-uehara@kansai-u.ac.jp*

We show local asymptotic normality (LAN) for a statistical model of discretely observed ergodic jump-diffusion processes, where the drift coefficient, diffusion coefficient, and jump structure are parametrized. Under the LAN property, we can discuss the asymptotic efficiency of regular estimators, and the quasi-maximum-likelihood and Bayes-type estimators proposed in Shimizu and Yoshida (*Stat. Inference Stoch. Process.* **9** (2006) 227–277) and Ogihara and Yoshida (*Stat. Inference Stoch. Process.* **14** (2011) 189–229) are shown to be asymptotically efficient in this model. Moreover, we can construct asymptotically uniformly most powerful tests for the parameters. Unlike with a model for diffusion processes, Aronson-type estimates of the transition density functions do not hold, which makes it difficult to prove LAN. Therefore, instead of Aronson-type estimates, we employ the idea of Theorem 1 in Jeganathan (*Sankhyā Ser. A* **44** (1982) 173–212) and use the L^2 regularity condition. Moreover, we show that local asymptotic mixed normality of a statistical model is implied from that for a model generated by approximated transition density functions under suitable conditions. Together with density approximation by means of thresholding techniques, the LAN property for the jump-diffusion processes is proved.

Keywords: Asymptotically efficient estimator; asymptotically uniformly most powerful test; jump-diffusion processes; local asymptotic mixed normality; L^2 regularity condition; Malliavin calculus; thresholding techniques

References

[1] Aït-Sahalia, Y. and Jacob, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** 355–392. MR2332279 https://doi.org/10.1214/009053606000001190
[2] 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
[3] Basawa, I.V. and Scott, D.J. (1983). *Asymptotic Optimal Inference for Nonergodic Models*. *Lecture Notes in Statistics* **17**. New York-Berlin: Springer. MR0688650
[4] Bibinger, M. and Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *J. Econometrics* **184** 361–378. MR3291008 https://doi.org/10.1016/j.jeconom.2014.10.004
[5] Choi, S., Hall, W.J. and Schick, A. (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.* **24** 841–861. MR1394992 https://doi.org/10.1214/aos/1032894469
[6] Clément, E., Delattre, S. and Gloter, A. (2014). Asymptotic lower bounds in estimating jumps. *Bernoulli* **20** 1059–1096. MR3217438 https://doi.org/10.3150/13-BEJ515
[7] Clément, E. and Gloter, A. (2015). Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable Lévy processes. *Stochastic Process. Appl.* **125** 2316–2352. MR3322866 https://doi.org/10.1016/j.spa.2015.01.002
[8] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: CRC Press/CRC. MR2042661

1350-7265

[9] Eguchi, S. and Masuda, H. (2018). Schwarz type model comparison for LAQ models. *Bernoulli* **24** 2278–2327. MR3757530 https://doi.org/10.3150/17-BEJ928
[10] Fukasawa, M. and Ogihara, T. Malliavin calculus techniques for local asymptotic mixed normality and their application to degenerate diffusions. Available at arXiv:2005.14599.
[11] Gloter, A., Loukianova, D. and Mai, H. (2018). Jump filtering and efficient drift estimation for Lévy-driven SDEs. *Ann. Statist.* **46** 1445–1480. MR3819106 https://doi.org/10.1214/17-AOS1591
[12] 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
[13] 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
[14] 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
[15] 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, Calif.: Univ. California Press. MR0400513
[16] Ibragimov, I.A. and Has'minskiĭ, R.Z. (1981). *Statistical Estimation*. *Applications of Mathematics* **16**. New York-Berlin: Springer. Asymptotic theory, Translated from the Russian by Samuel Kotz. MR0620321
[17] Inatsugu, H. and Yoshida, N. (2021). Global jump filters and quasi-likelihood analysis for volatility. *Ann. Inst. Statist. Math.* **73** 555–598. MR4247071 https://doi.org/10.1007/s10463-020-00768-x
[18] 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
[19] Kawai, R. and Masuda, H. (2013). Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling. *ESAIM Probab. Stat.* **17** 13–32. MR3002994 https://doi.org/10.1051/ps/2011101
[20] 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
[21] Mai, H. (2014). Efficient maximum likelihood estimation for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli* **20** 919–957. MR3178522 https://doi.org/10.3150/13-BEJ510
[22] Mancini, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Process. Appl.* **121** 845–855. MR2770909 https://doi.org/10.1016/j.spa.2010.12.001
[23] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010
[24] Masuda, H. (2008). On stability of diffusions with compound-Poisson jumps. *Bull. Inform. Cybernet.* **40** 61–74. MR2484446
[25] 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
[26] Ogihara, T. and Uehara, Y. (2023). Supplement to "Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation." https://doi.org/10.3150/22-BEJ1544SUPP
[27] Ogihara, T. and Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.* **14** 189–229. MR2832818 https://doi.org/10.1007/s11203-011-9057-z
[28] Prakasa Rao, B.L.S. (1999). *Semimartingales and Their Statistical Inference*. *Monographs on Statistics and Applied Probability* **83**. Boca Raton, FL: CRC Press/CRC. MR1689166
[29] Shimizu, Y. (2008). A practical inference for discretely observed jump-diffusions from finite samples. *J. Japan Statist. Soc.* **38** 391–413. MR2510946 https://doi.org/10.14490/jjss.38.391
[30] Shimizu, Y. and Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.* **9** 227–277. MR2252242 https://doi.org/10.1007/s11203-005-8114-x
[31] Sweeting, T.J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8** 1375–1381. MR0594652

Bernoulli **29**(3), 2023, 2367–2391
https://doi.org/10.3150/22-BEJ1545

A frequency domain bootstrap for general multivariate stationary processes

MARCO MEYER[1,a] and EFSTATHIOS PAPARODITIS[2,b]

[1]*Department of Mathematics, Technische Universität Braunschweig, Braunschweig, Germany,* [a]*marco.meyer@tu-bs.de*
[2]*Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus,* [b]*stathisp@ucy.ac.cy*

Developing valid frequency domain bootstrap procedures for integrated periodogram statistics for multivariate time series is a challenging problem. This is mainly due to the fact that the distribution of such statistics depends on the fourth-order moment structure of the underlying multivariate process in nearly every scenario. Exceptions are some very special cases like nonparametric estimators of the spectral density matrix or Gaussian time series. In contrast to the univariate case, even additional structural assumptions – such as linearity of the multivariate process or a standardization of the statistic of interest – do not solve the problem. This paper proposes a new frequency domain bootstrap procedure for multivariate time series, the multivariate frequency domain hybrid bootstrap (MFHB), for integrated periodogram statistics as well as for functions thereof. Asymptotic validity of the MFHB procedure is established for these statistics and for a class of stationary multivariate processes satisfying rather weak dependence conditions ranging clearly beyond linear processes. The finite sample performance of the MFHB is investigated by means of simulations.

Keywords: Bootstrap; periodogram; spectral means; stationary processes

References

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory*, 2nd ed. *Holden-Day Series in Time Series Analysis*. Oakland, CA: Holden-Day, Inc. MR0595684

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2839251

Dahlhaus, R. (1985). On a spectral density estimate obtained by averaging periodograms. *J. Appl. Probab.* **22** 598–610. MR0799283 https://doi.org/10.1017/s0021900200029351

Dahlhaus, R. and Janas, D. (1996). A frequency domain bootstrap for ratio statistics in time series analysis. *Ann. Statist.* **24** 1934–1963. MR1421155 https://doi.org/10.1214/aos/1069362304

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26. MR0515681

Fragkeskou, M. and Paparoditis, E. (2018). Extending the range of validity of the autoregressive (sieve) bootstrap. *J. Time Series Anal.* **39** 356–379. MR3796524 https://doi.org/10.1111/jtsa.12275

Francq, C. and Zakoïan, J.-M. (2016). Estimating multivariate volatility models equation by equation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 613–635. MR3506795 https://doi.org/10.1111/rssb.12126

Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *Ann. Statist.* **20** 121–145. MR1150337 https://doi.org/10.1214/aos/1176348515

Hurvich, C.M. and Zeger, S.L. (1987). Frequency domain bootstrap methods for time series. Preprint, Dept. of Stat. and Operations Research, New York Univ.

Jentsch, C. and Kreiss, J.-P. (2010). The multiple hybrid bootstrap – Resampling multivariate linear processes. *J. Multivariate Anal.* **101** 2320–2345. MR2719865 https://doi.org/10.1016/j.jmva.2010.06.005

Kreiss, J.-P. and Paparoditis, E. (2012). The hybrid wild bootstrap for time series. *J. Amer. Statist. Assoc.* **107** 1073–1084. MR3010895 https://doi.org/10.1080/01621459.2012.695664

Kreiss, J.-P., Paparoditis, E. and Politis, D.N. (2011). On the range of validity of the autoregressive sieve bootstrap. *Ann. Statist.* **39** 2103–2130. MR2893863 https://doi.org/10.1214/11-AOS900

1350-7265

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

Meyer, M. and Paparoditis, E. (2023). Supplement to "A frequency domain bootstrap for general multivariate stationary processes." https://doi.org/10.3150/22-BEJ1545SUPP

Meyer, M., Paparoditis, E. and Kreiss, J.-P. (2020). Extending the validity of frequency domain bootstrap methods to general stationary processes. *Ann. Statist.* **48** 2404–2427. MR4134800 https://doi.org/10.1214/19-AOS1892

Priestley, M.B. (1981). *Spectral Analysis and Time Series. Vol. 1*. *Probability and Mathematical Statistics*. London: Academic Press. MR0628735

Robinson, P.M. (1991). Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica* **59** 1329–1363. MR1133037 https://doi.org/10.2307/2938370

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Boston, MA: Birkhäuser, Inc. MR0885090 https://doi.org/10.1007/978-1-4612-5156-9

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. *Springer Series in Statistics*. New York: Springer. MR1785484 https://doi.org/10.1007/978-1-4612-1162-4

Tewes, J., Politis, D.N. and Nordman, D.J. (2019). Convolved subsampling estimation with applications to block bootstrap. *Ann. Statist.* **47** 468–496. MR3909939 https://doi.org/10.1214/18-AOS1695

Tsay, R.S. (2014). *Multivariate Time Series Analysis*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR3236787

Withers, C.S. and Nadarajah, S. (2012). Moments and cumulants for the complex Wishart. *J. Multivariate Anal.* **112** 242–247. MR2957300 https://doi.org/10.1016/j.jmva.2012.05.002

Wu, W.B. and Zaffaroni, P. (2018). Asymptotic theory for spectral density estimates of general multivariate time series. *Econometric Theory* **34** 1–22. MR3739234 https://doi.org/10.1017/S0266466617000068

Bernoulli **29**(3), 2023, 2392–2416
https://doi.org/10.3150/22-BEJ1546

Conditional quantiles: An operator-theoretical approach

LUCIANO DE CASTRO[1,a], BRUNO N. COSTA[3,4,c], ANTONIO F. GALVAO[2,b] and JORGE P. ZUBELLI[3,4,d]

[1]*Department of Economics, University of Iowa, Iowa City, USA,* [a]*decastro.luciano@gmail.com*
[2]*Department of Economics, Michigan State University, East Lansing, USA,* [b]*agalvao@msu.edu*
[3]*IMPA, Rio de Janeiro, Brazil,* [c]*brunonunes@impa.br*
[4]*Department of Mathematics, Khalifa University, Abu Dhabi, UAE,* [d]*jorge.zubelli@ku.ac.ae*

This paper derives several novel properties of conditional quantiles viewed as nonlinear operators. The results are organized in parallel to the usual properties of the expectation operator. We first define a τ-conditional quantile random set, relative to any sigma-algebra, as a set of solutions of an optimization problem. Then, well-known properties of unconditional quantiles, as translation invariance, comonotonicity, and equivariance to monotone transformations, are generalized to the conditional case. Moreover, a simple proof for Jensen's inequality for conditional quantiles is provided. We also investigate continuity of conditional quantiles as operators with respect to different topologies and obtain a novel Fatou's lemma for quantiles. Conditions for continuity in L^p and weak continuity are also derived. Then, the differentiability properties of quantiles are addressed. We demonstrate the validity of Leibniz's rule for conditional quantiles for the cases of monotone, as well as separable functions. Finally, although the law of iterated quantiles does not hold in general, we characterize the maximum set of random variables for which this law holds, and investigate its consequences for the infinite composition of conditional quantiles.

Keywords: Conditional quantiles; continuity for quantiles; Fatou's lemma for quantiles; Leibniz's rule

References

[1] Apiwatcharoenkul, W., Lake, L.W. and Jensen, J.L. (2016). Uncertainty in proved reserves estimates by decline curve analysis. In *SPE/IAEE Hydrocarbon Economics and Evaluation Symposium* Society of Petroleum Engineers. https://doi.org/10.2118/179992-MS

[2] Chambers, C.P. (2009). An axiomatization of quantiles on the domain of distribution functions. *Math. Finance* **19** 335–342. MR2509291 https://doi.org/10.1111/j.1467-9965.2009.00369.x

[3] Cheung, K.C. (2007). Characterizations of conditional comonotonicity. *J. Appl. Probab.* **44** 607–617. MR2355579 https://doi.org/10.1239/jap/1189717532

[4] de Castro, L., Costa, B.N., Galvao, A.F. and Zubelli, J.P. (2023). Supplement to "Conditional Quantiles: An Operator-Theoretical Approach." https://doi.org/10.3150/22-BEJ1546SUPP

[5] de Castro, L. and Galvao, A.F. (2019). Dynamic quantile models of rational behavior. *Econometrica* **87** 1893–1939. MR4051682 https://doi.org/10.3982/ecta15146

[6] de Castro, L. and Galvao, A.F. (2022). Static and dynamic quantile preferences. *Econom. Theory* **73** 747–779. MR4414440 https://doi.org/10.1007/s00199-021-01355-8

[7] de Castro, L., Galvao, A.F., Montes-Rojas, G. and Olmo, J. (2022). Portfolio selection in quantile decision models. *Ann. Finance* **18** 133–181. MR4434243 https://doi.org/10.1007/s10436-021-00405-4

[8] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory *Insurance Math. Econom.* **31** 3–33. MR1956509 https://doi.org/10.1016/S0167-6687(02)00134-8

[9] Duffie, D. and Pan, J. (1997). An overview of value at risk. *J. of Derivatives* **4** 7–49.

[10] Embrechts, P., Höing, A. and Juri, A. (2003). Using copulae to bound the value-at-risk for functions of dependent risks. *Finance Stoch.* **7** 145–167. MR1968943 https://doi.org/10.1007/s007800200085
[11] Fanchi, J.R. and Christiansen, R.L. (2017). *Introduction to Petroleum Engineering*. New York: Wiley Online Library.
[12] Ferguson, T.S. (1996). *A Course in Large Sample Theory*. *Texts in Statistical Science Series*. London: CRC Press. MR1699953 https://doi.org/10.1007/978-1-4899-4549-5
[13] Föllmer, H. and Schied, A. (2002). *Stochastic Finance. An Introduction in Discrete Time*. *De Gruyter Studies in Mathematics* **27**. Berlin: de Gruyter. MR1925197 https://doi.org/10.1515/9783110198065
[14] Jorion, P. (2007). *Value at Risk. The New Benchmark for Managing Financial Risk*, 3rd ed. New York: McGraw-Hill.
[15] Jouini, E. and Napp, C. (2004). Conditional comonotonicity. *Decis. Econ. Finance* **27** 153–166. MR2104639 https://doi.org/10.1007/s10203-004-0049-y
[16] Koenker, R. (2005). *Quantile Regression*. *Econometric Society Monographs* **38**. Cambridge: Cambridge Univ. Press. MR2268657 https://doi.org/10.1017/CBO9780511754098
[17] Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50. MR0474644 https://doi.org/10.2307/1913643
[18] Koenker, R. and Bassett, G.W. Jr. (2010). March madness, quantile regression bracketology, and the Hayek hypothesis. *J. Bus. Econom. Statist.* **28** 26–35. MR2650598 https://doi.org/10.1198/jbes.2009.07093
[19] Koenker, R., Chernozhukov, V., He, X. and Peng, L. (2017). *Handbook of Quantile Regression*. New York, New York: CRC Press/CRC.
[20] Le Gall, J.-F. (2013). *Mouvement Brownien, Martingales et Calcul Stochastique*. *Mathématiques & Applications (Berlin) [Mathematics & Applications]* **71**. Heidelberg: Springer. MR3184878 https://doi.org/10.1007/978-3-642-31898-6
[21] Le Gall, J.-F. (2006). Intégration, Probabilités et Processus Aléatories. Mimeo, Département Mathématiques et Applications Ecole normale supérieure de Paris.
[22] Manski, C.F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision* **25** 79–104. MR0951879 https://doi.org/10.1007/BF00129169
[23] Mendelson, H. (1987). Quantile-preserving spread. *J. Econom. Theory* **42** 334–351. MR0898524 https://doi.org/10.1016/0022-0531(87)90091-3
[24] Merkle, M. (2005). Jensen's inequality for medians. *Statist. Probab. Lett.* **71** 277–281. MR2126412 https://doi.org/10.1016/j.spl.2004.11.010
[25] Molchanov, I. (2017). *Theory of Random Sets*, 2nd ed. London: Springer.
[26] Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary Problems*. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2256030
[27] Pollard, D. (2006). *A User's Guide to Measure Theoretic Probability*. Cambridge, MA: Cambridge Univ. Press.
[28] Rostek, M. (2010). Quantile maximization in decision theory. *Rev. Econ. Stud.* **77** 339–371. MR2599290 https://doi.org/10.1111/j.1467-937X.2009.00564.x
[29] Schmeidler, D. (1986). Integral representation without additivity. *Proc. Amer. Math. Soc.* **97** 255–261. MR0835875 https://doi.org/10.2307/2046508
[30] Shreve, S.E. (2004). *Stochastic Calculus for Finance. I: The Binomial Asset Pricing Model*. *Springer Finance*. New York: Springer. MR2049045
[31] Sweeting, T.J. (1989). On conditional weak convergence. *J. Theoret. Probab.* **2** 461–474. MR1011199 https://doi.org/10.1007/BF01051878
[32] Tasche, D. (2000). Conditional expectation as quantile derivative. Working Paper.
[33] Valadier, M. (1984). La multi-application médianes conditionnelles. *Probab. Theory Related Fields* **67** 279–282. MR0762082 https://doi.org/10.1007/BF00535005
[34] Williams, D. (1991). *Probability with Martingales*. *Cambridge Mathematical Textbooks*. Cambridge: Cambridge Univ. Press. MR1155402 https://doi.org/10.1017/CBO9780511813658
[35] Zhao, M., Yang, X., He, Q., Zhou, Z. and Ge, X. (2021). Quantile Jensen's inequalities. *J. Inequal. Appl.* **86** 1–8.

Bernoulli **29**(3), 2023, 2417–2441
https://doi.org/10.3150/22-BEJ1547

A probabilistic view of latent space graphs and phase transitions

SUQI LIU[a] and MIKLÓS Z. RÁCZ[b]

Department of Operations Research and Financial Engineering, Princeton University, Sherrerd Hall, Charlton Street, Princeton, NJ 08544, USA, [a]*suqil@princeton.edu,* [b]*mracz@princeton.edu*

We study random graphs with latent geometric structure, where the probability of each edge depends on the underlying random positions corresponding to the two endpoints. We consider the setting where this conditional probability is a general monotone increasing function of the inner product of two vectors; such a function can naturally be viewed as the cumulative distribution function of some independent random variable. A one-parameter family of random graphs, characterized by the variance of this random variable, that smoothly interpolates between a random dot product graph and an Erdős–Rényi random graph, is investigated. Focusing on the dense regime, we prove phase transitions of detecting geometry in these graphs, in terms of the dimension of the underlying geometric space and the variance parameter: When the dimension is high or the variance is large, the graph is similar to an Erdős–Rényi graph with the same edge density; in other parameter regimes, there is a computationally efficient signed triangle statistic that can distinguish them. The proofs make use of information-theoretic inequalities and concentration of measure phenomena.

Keywords: Random graph; random dot product graph; high-dimensional geometric structure; signed triangle

References

[1] Athreya, A., Fishkind, D.E., Tang, M., Priebe, C.E., Park, Y., Vogelstein, J.T., Levin, K., Lyzinski, V., Qin, Y. and Sussman, D.L. (2017). Statistical inference on random dot product graphs: A survey. *J. Mach. Learn. Res.* **18** Paper No. 226, 92. MR3827114

[2] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. With a foreword by Michel Ledoux. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

[3] Brennan, M., Bresler, G. and Huang, B. (2021). De Finetti-Style Results for Wishart Matrices: Combinatorial Structure and Phase Transitions. arXiv preprint arXiv:2103.14011.

[4] Brennan, M., Bresler, G. and Nagaraj, D. (2020). Phase transitions for detecting latent geometry in random graphs. *Probab. Theory Related Fields* **178** 1215–1289. MR4168398 https://doi.org/10.1007/s00440-020-00998-3

[5] Bubeck, S., Ding, J., Eldan, R. and Rácz, M.Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Structures Algorithms* **49** 503–532. MR3545825 https://doi.org/10.1002/rsa.20633

[6] Bubeck, S. and Ganguly, S. (2018). Entropic CLT and phase transition in high-dimensional Wishart matrices. *Int. Math. Res. Not. IMRN* **2** 588–606. MR3801440 https://doi.org/10.1093/imrn/rnw243

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

[8] Devroye, L., György, A., Lugosi, G. and Udina, F. (2011). High-dimensional random geometric graphs and their clique number. *Electron. J. Probab.* **16** 2481–2508. MR2861682 https://doi.org/10.1214/EJP.v16-967

[9] Duchemin, Q. and De Castro, Y. (2021). Random Geometric Graph: Some recent developments and perspectives. Preprint.

[10] Eldan, R. and Mikulincer, D. (2020). Information and dimensionality of anisotropic random geometric graphs. In *Geometric Aspects of Functional Analysis. Vol. I*. *Lecture Notes in Math.* **2256** 273–324. Cham: Springer. MR4175752 https://doi.org/10.1007/978-3-030-36020-7_13

[11] Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.
[12] Gilbert, E.N. (1961). Random plane networks. *J. Soc. Indust. Appl. Math.* **9** 533–543. MR0132566
[13] Hoff, P.D., Raftery, A.E. and Handcock, M.S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. MR1951262 https://doi.org/10.1198/016214502388618906
[14] Janson, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge: Cambridge Univ. Press. MR1474726 https://doi.org/10.1017/CBO9780511526169
[15] Jiang, T. and Li, D. (2015). Approximation of rectangular beta-Laguerre ensembles and large deviations. *J. Theoret. Probab.* **28** 804–847. MR3413957 https://doi.org/10.1007/s10959-013-0519-7
[16] Liu, S. and Rácz, M.Z. (2021). Phase transition in noisy high-dimensional random geometric graphs. arXiv preprint arXiv:2103.15249.
[17] Liu, S. and Rácz, M.Z. (2023). Supplement to "A probabilistic view of latent space graphs and phase transitions." https://doi.org/10.3150/22-BEJ1547SUPP
[18] Ma, Z., Ma, Z. and Yuan, H. (2020). Universal latent space model fitting for large networks with edge covariates. *J. Mach. Learn. Res.* **21** Paper No. 4, 67. MR4071187 https://doi.org/10.1109/tnnls.2020.3010690
[19] Marchal, O. and Arbel, J. (2017). On the sub-Gaussianity of the beta and Dirichlet distributions. *Electron. Commun. Probab.* **22** Paper No. 54, 14. MR3718704 https://doi.org/10.1214/17-ECP92
[20] Penrose, M. (2003). *Random Geometric Graphs*. *Oxford Studies in Probability* **5**. Oxford: Oxford Univ. Press. MR1986198 https://doi.org/10.1093/acprof:oso/9780198506263.001.0001
[21] Penrose, M.D. (2016). Connectivity of soft random geometric graphs. *Ann. Appl. Probab.* **26** 986–1028. MR3476631 https://doi.org/10.1214/15-AAP1110
[22] Polyanskiy, Y. and Wu, Y. (2012-2017). Lecture notes on Information Theory.
[23] Rácz, M.Z. and Bubeck, S. (2017). Basic models and questions in statistical network analysis. *Stat. Surv.* **11** 1–47. MR3696007 https://doi.org/10.1214/17-SS117
[24] Rácz, M.Z. and Richey, J. (2019). A smooth transition from Wishart to GOE. *J. Theoret. Probab.* **32** 898–906. MR3959632 https://doi.org/10.1007/s10959-018-0808-2
[25] Stein, C.M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151. MR0630098
[26] Tang, M., Sussman, D.L. and Priebe, C.E. (2013). Universally consistent vertex classification for latent positions graphs. *Ann. Statist.* **41** 1406–1430. MR3113816 https://doi.org/10.1214/13-AOS1112
[27] van Handel, R. (2016). Probability in High Dimension. APC 550 Lecture Notes (Princeton University).
[28] Wainwright, M.J. (2019). *High-Dimensional Statistics: A Non-asymptotic Viewpoint*. *Cambridge Series in Statistical and Probabilistic Mathematics* **48**. Cambridge: Cambridge Univ. Press. MR3967104 https://doi.org/10.1017/9781108627771
[29] Wendel, J.G. (1948). Note on the gamma function. *Amer. Math. Monthly* **55** 563–564. MR0029448 https://doi.org/10.2307/2304460
[30] Young, S.J. and Scheinerman, E.R. (2007). Random dot product graph models for social networks. In *Algorithms and Models for the Web-Graph*. *Lecture Notes in Computer Science* **4863** 138–149. Berlin: Springer. MR2504912 https://doi.org/10.1007/978-3-540-77004-6_11
[31] Zhang, A.R. and Zhou, Y. (2020). On the non-asymptotic and sharp lower tail bounds of random variables. *Stat* **9** e314, 11. MR4193419 https://doi.org/10.1002/sta4.314

Bernoulli **29**(3), 2023, 2442–2465
https://doi.org/10.3150/22-BEJ1548

Exponential ergodicity for damping Hamiltonian dynamics with state-dependent and non-local collisions

JIANHAI BAO[1,a] and JIAN WANG[2,b]

[1] *Center for Applied Mathematics, Tianjin University, 300072 Tianjin, P.R. China,* [a]*jianhaibao@tju.edu.cn*
[2] *School of Mathematics and Statistics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, 350007 Fuzhou, P.R. China,* [b]*jianwang@fjnu.edu.cn*

In this paper, we investigate the exponential ergodicity in a Wasserstein-type distance for a damping Hamiltonian dynamics with state-dependent and non-local collisions, which indeed is a special case of piecewise deterministic Markov processes that is very popular in numerous modelling situations including stochastic algorithms. The approach adopted in this work is based on a combination of the refined basic coupling and the refined reflection coupling for non-local operators. In a certain sense, the main result developed in the present paper is a continuation of the counterpart in (*Stochastic Process. Appl.* (2022) **146** 114–142) on exponential ergodicity of stochastic Hamiltonian systems with Lévy noises and a complement of (*Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022a) 916–944) upon exponential ergodicity for Andersen dynamics with constant jump rate functions.

Keywords: Coupling; damping Hamiltonian dynamics; exponential ergodicity; non-local collision; Wasserstein-type distance

References

Andrieu, C., Durmus, A., Nüsken, N. and Roussel, J. (2021). Hypocoercivity of piecewise deterministic Markov process-Monte Carlo. *Ann. Appl. Probab.* **31** 2478–2517. MR4332703 https://doi.org/10.1214/20-aap1653

Bao, J. and Wang, J. (2022). Coupling approach for exponential ergodicity of stochastic Hamiltonian systems with Lévy noises. *Stochastic Process. Appl.* **146** 114–142. MR4370566 https://doi.org/10.1016/j.spa.2021.12.014

Benaïm, M., Le Borgne, S., Malrieu, F. and Zitt, P.-A. (2015). Qualitative properties of certain piecewise deterministic Markov processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1040–1075. MR3365972 https://doi.org/10.1214/14-AIHP619

Bierkens, J., Grazzi, S., van der Meulen, F. and Schauer, M. (2021). A piecewise deterministic Monte Carlo method for diffusion bridges. *Stat. Comput.* **31** Paper No. 37, 21. MR4249043 https://doi.org/10.1007/s11222-021-10008-8

Bou-Rabee, N. and Eberle, A. (2022a). Couplings for Andersen dynamics. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 916–944. MR4421613 https://doi.org/10.1214/21-aihp1197

Bou-Rabee, N. and Eberle, A. (2022b). Couplings for Andersen dynamics. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 916–944. MR4421613 https://doi.org/10.1214/21-aihp1197

Bou-Rabee, N. and Sanz-Serna, J.M. (2017). Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **27** 2159–2194. MR3693523 https://doi.org/10.1214/16-AAP1255

Cattiaux, P., León, J.R. and Prieur, C. (2014). Estimation for stochastic damping Hamiltonian systems under partial observation—I. Invariant density. *Stochastic Process. Appl.* **124** 1236–1260. MR3148012 https://doi.org/10.1016/j.spa.2013.10.008

Costa, O.L.V. and Dufour, F. (2008). Stability and ergodicity of piecewise deterministic Markov processes. *SIAM J. Control Optim.* **47** 1053–1077. MR2385873 https://doi.org/10.1137/060670109

Czapla, D., Horbacz, K. and Wojewódka-Ściążko, H. (2020). Ergodic properties of some piecewise-deterministic Markov process with application to gene expression modelling. *Stochastic Process. Appl.* **130** 2851–2885. MR4080732 https://doi.org/10.1016/j.spa.2019.08.006

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

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

Durmus, A., Guillin, A. and Monmarché, P. (2021). Piecewise deterministic Markov processes and their invariant measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1442–1475. MR4291455 https://doi.org/10.1214/20-aihp1125

Eberle, A., Guillin, A. and Zimmer, R. (2019). Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47** 1982–2010. MR3980913 https://doi.org/10.1214/18-AOP1299

Fearnhead, P., Bierkens, J., Pollock, M. and Roberts, G.O. (2018). Piecewise deterministic Markov processes for continuous-time Monte Carlo. *Statist. Sci.* **33** 386–412. MR3843382 https://doi.org/10.1214/18-STS648

Hairer, M., Mattingly, J.C. and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** 223–259. MR2773030 https://doi.org/10.1007/s00440-009-0250-6

Jacobsen, M. (2006). *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. *Probability and Its Applications*. Boston, MA: Birkhäuser, Inc. MR2189574

Lemaire, V., Thieullen, M. and Thomas, N. (2020). Thinning and multilevel Monte Carlo methods for piecewise deterministic (Markov) processes with an application to a stochastic Morris-Lecar model. *Adv. in Appl. Probab.* **52** 138–172. MR4092810 https://doi.org/10.1017/apr.2019.55

Liang, M., Schilling, R.L. and Wang, J. (2020). A unified approach to coupling SDEs driven by Lévy noise and some applications. *Bernoulli* **26** 664–693. MR4036048 https://doi.org/10.3150/19-BEJ1148

Lu, Y. and Mattingly, J.C. (2020). Geometric ergodicity of Langevin dynamics with Coulomb interactions. *Nonlinearity* **33** 675–699. MR4048775 https://doi.org/10.1088/1361-6544/ab514a

Luo, D. and Wang, J. (2019). Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stochastic Process. Appl.* **129** 3129–3173. MR3985558 https://doi.org/10.1016/j.spa.2018.09.003

Malrieu, F. (2015). Some simple but challenging Markov processes. *Ann. Fac. Sci. Toulouse Math. (6)* **24** 857–883. MR3434260 https://doi.org/10.5802/afst.1468

Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266 https://doi.org/10.1016/S0304-4149(02)00150-3

Talay, D. (2002). Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** 163–198. MR1924934

Wu, L. (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** 205–238. MR1807683 https://doi.org/10.1016/S0304-4149(00)00061-2

Zhang, X. (2014). Fundamental solution of kinetic Fokker-Planck operator with anisotropic nonlocal dissipativity. *SIAM J. Math. Anal.* **46** 2254–2280. MR3225504 https://doi.org/10.1137/130926493

Bernoulli **29**(3), 2023, 2466–2491
https://doi.org/10.3150/22-BEJ1549

Cramér-type moderate deviation for quadratic forms with a fast rate

XIAO FANG[1,a], SONG-HAO LIU[2,b] and QI-MAN SHAO[3,c]

[1]*Department of Statistics, The Chinese University of Hong Kong, Hong Kong SAR,* [a]*xfang@sta.cuhk.edu.hk*
[2]*Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, 518055, Guangdong, China,* [b]*liush@sustech.edu.cn*
[3]*Department of Statistics and Data Science, SICM, National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology, Shenzhen, 518055, Guangdong, China,* [c]*shaoqm@sustech.edu.cn*

Let $X_1,\dots,X_n$ be independent and identically distributed random vectors in $\mathbb{R}^d$. Suppose $\mathbb{E}X_1 = 0$, $\mathrm{Cov}(X_1) = I_d$, where I_d is the $d \times d$ identity matrix. Suppose further that there exist positive constants t_0 and c_0 such that $\mathbb{E}e^{t_0|X_1|} \le c_0 < \infty$, where $|\cdot|$ denotes the Euclidean norm. Let $W = \sum_{i=1}^n X_i/\sqrt{n}$ and let Z be a d-dimensional standard normal random vector. Let Q be a $d \times d$ symmetric positive definite matrix whose largest eigenvalue is 1. We prove that for $0 \le x \le \varepsilon n^{1/6}$,

$$\left|\frac{\mathbb{P}(|Q^{1/2}W| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} - 1\right| \le C\left(\frac{1+x^5}{\det(Q^{1/2})n} + \frac{x^6}{n}\right) \quad \text{for } d \ge 5$$

and

$$\left|\frac{\mathbb{P}(|Q^{1/2}W| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} - 1\right| \le C\left(\frac{1+x^3}{\det(Q^{1/2})n^{\frac{d}{d+1}}} + \frac{x^6}{n}\right) \quad \text{for } 1 \le d \le 4,$$

where ε and C are positive constants depending only on d, t_0, and c_0. This is a first extension of Cramér-type moderate deviation to the multivariate setting with a faster convergence rate than $1/\sqrt{n}$. The range of $x = o(n^{1/6})$ for the relative error to vanish and the dimension requirement $d \ge 5$ for the $1/n$ rate are both optimal. We prove our result using a new change of measure, a two-term Edgeworth expansion for the changed measure, and cancellation by symmetry for terms of the order $1/\sqrt{n}$.

Keywords: Asymptotic expansion; central limit theorem; change of measure; moderate deviations; quadratic forms

References

Aleškevičienè, A. and Statulevičius, V. (1997). Probabilities of large deviations in the approximation by χ^2-law. *Lith. Math. J.* **37** 301–309.
Bentkus, V. (2004). A Lyapunov type bound in $\mathbf{R}^d$. *Theory Probab. Appl.* **49** 311–323.
Bentkus, V. and Götze, F. (1997). Uniform rates of convergence in the CLT for quadratic forms in multidimensional spaces. *Probab. Theory Related Fields* **109** 367–416. MR1481126 https://doi.org/10.1007/s004400050136
Bentkus, V. and Götze, F. (1999). Lattice point problems and distribution of values of quadratic forms. *Ann. of Math. (2)* **150** 977–1027. MR1740988 https://doi.org/10.2307/121060
Berry, A.C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** 122–136. MR0003498 https://doi.org/10.2307/1990053
Bhattacharya, R.N. and Rao, R.R. (2010). *Normal Approximation and Asymptotic Expansions*. *Classics in Applied Mathematics* **64**. Philadelphia, PA: SIAM. MR3396213 https://doi.org/10.1137/1.9780898719895.ch1
Cramér, H. (1938). Sur un nouveau théoreme-limite de la théorie des probabilités. *Actual. Sci. Ind.* **736** 5–23.

Esseen, C.-G. (1942). On the Liapounoff limit of error in the theory of probability. *Ark. Mat. Astron. Fys.* **28A** 19. MR0011909

Esseen, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* **77** 1–125. MR0014626 https://doi.org/10.1007/BF02392223

Götze, F. and Zaitsev, A.Y. (2014). Explicit rates of approximation in the CLT for quadratic forms. *Ann. Probab.* **42** 354–397. MR3161488 https://doi.org/10.1214/13-AOP839

Landau, E. (1924). Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Z.* **21** 126–132. MR1544690 https://doi.org/10.1007/BF01187457

Fang, X., Liu, S.-H. and Shao, Q.-M. (2023). Supplement to "Cramér-type Moderate Deviation for Quadratic Forms with a Fast Rate." https://doi.org/10.3150/22-BEJ1549SUPP

von Bahr, B. (1967). Multi-dimensional integral limit theorems for large deviations. *Ark. Mat.* **7** 89–99. MR0224129 https://doi.org/10.1007/BF02591679

Walfisz, A. (1924). Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Z.* **19** 300–307. MR1544657 https://doi.org/10.1007/BF01181080

Walfisz, A. (1957). *Gitterpunkte in Mehrdimensionalen Kugeln*. *Monografie Matematyczne, Vol. 33*. Warsaw: Państwowe Wydawnictwo Naukowe. MR0097357

Bernoulli **29**(3), 2023, 2492–2518
https://doi.org/10.3150/22-BEJ1550

Empirical approximation to invariant measures for McKean–Vlasov processes: Mean-field interaction vs self-interaction

KAI DU[1,2,a], YIFAN JIANG[3,b] and JINFENG LI[4,c]

[1]*Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China,* [a]*kdu@fudan.edu.cn*
[2]*Shanghai Artificial Intelligence Laboratory, 701 Yunjin Road, Shanghai 200232, China*
[3]*Mathematical Institute, University of Oxford, Oxford, United Kingdom,* [b]*yifan.jiang@maths.ox.ac.uk*
[4]*School of Mathematical Sciences, Fudan University, Shanghai, China,* [c]*lijinfeng@fudan.edu.cn*

This paper proves that, under a monotonicity condition, the invariant probability measure of a McKean–Vlasov process can be approximated by weighted empirical measures of some processes including itself. These processes are described by distribution dependent or empirical measure dependent stochastic differential equations constructed from the equation for the McKean–Vlasov process. Convergence of empirical measures is characterized by upper bound estimates for their Wasserstein distances to the invariant measure. Numerical simulations of the mean-field Ornstein–Uhlenbeck process are implemented to demonstrate the theoretical results.

Keywords: Empirical measures; invariant measures; McKean-Vlasov processes; Wasserstein distances

References

[1] Ahmed, N.U. and Ding, X. (1993). On invariant measures of nonlinear Markov processes. *J. Appl. Math. Stoch. Anal.* **6** 385–406. MR1255258 https://doi.org/10.1155/S1048953393000310
[2] AlRachid, H., Bossy, M., Ricci, C. and Szpruch, L. (2019). New particle representations for ergodic McKean-Vlasov SDEs. In *CEMRACS 2017—Numerical Methods for Stochastic Models: Control, Uncertainty Quantification, Mean-Field. ESAIM Proc. Surveys* **65** 68–83. Les Ulis: EDP Sci. MR3968537 https://doi.org/10.1051/proc/201965068
[3] Ambrosio, L., Stra, F. and Trevisan, D. (2019). A PDE approach to a 2-dimensional matching problem. *Probab. Theory Related Fields* **173** 433–477. MR3916111 https://doi.org/10.1007/s00440-018-0837-x
[4] Bao, J., Scheutzow, M. and Yuan, C. (2022). Existence of invariant probability measures for functional McKean-Vlasov SDEs. *Electron. J. Probab.* **27** Paper No. 43, 14. MR4404942 https://doi.org/10.1214/22-ejp773
[5] Benachour, S., Roynette, B. and Vallois, P. (1998). Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl.* **75** 203–224. MR1632197 https://doi.org/10.1016/S0304-4149(98)00019-2
[6] Benaïm, M., Ledoux, M. and Raimond, O. (2002). Self-interacting diffusions. *Probab. Theory Related Fields* **122** 1–41. MR1883716 https://doi.org/10.1007/s004400100161
[7] Birkhoff, G.D. (1931). Proof of the ergodic theorem. *Proc. Natl. Acad. Sci. USA* **17** 656–660.
[8] Bogachev, V.I., Röckner, M. and Shaposhnikov, S.V. (2019). On convergence to stationary distributions for solutions of nonlinear Fokker-Planck-Kolmogorov equations. *J. Math. Sci. (N. Y.)* **242** 69–84. MR3999207 https://doi.org/10.1007/s10958-019-04467-8
[9] Boissard, E. and Le Gouic, T. (2014). On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 539–563. MR3189084 https://doi.org/10.1214/12-AIHP517
[10] Buckdahn, R., Li, J., Peng, S. and Rainer, C. (2017). Mean-field stochastic differential equations and associated PDEs. *Ann. Probab.* **45** 824–878. MR3630288 https://doi.org/10.1214/15-AOP1076

1350-7265

[11] Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I–II*. Springer.

[12] Carrillo, J.A., McCann, R.J. and Villani, C. (2003). Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoam.* **19** 971–1018. MR2053570 https://doi.org/10.4171/RMI/376

[13] Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge: Cambridge Univ. Press. MR1417491 https://doi.org/10.1017/CBO9780511662829

[14] Dawson, D.A. (1983). Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Stat. Phys.* **31** 29–85. MR0711469 https://doi.org/10.1007/BF01010922

[15] Driver, R.D. (2012). *Ordinary and Delay Differential Equations* **20**. Springer Science & Business Media. MR0477368

[16] Durrett, R.T. and Rogers, L.C.G. (1992). Asymptotic behavior of Brownian polymers. *Probab. Theory Related Fields* **92** 337–349. MR1165516 https://doi.org/10.1007/BF01300560

[17] Eberle, A., Guillin, A. and Zimmer, R. (2019). Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.* **371** 7135–7173. MR3939573 https://doi.org/10.1090/tran/7576

[18] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7

[19] Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–998. With discussion. MR0757767

[20] Gu, H., Guo, X., Wei, X. and Xu, R. (2021). Mean-field multi-agent reinforcement learning: A decentralized network approach. arXiv preprint arXiv:2108.02731.

[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] Hernández-Lerma, O. and Lasserre, J.B. (2012). *Markov Chains and Invariant Probabilities* **211**. Birkhäuser. MR1974383 https://doi.org/10.1007/978-3-0348-8024-4

[23] Herrmann, S. and Tugaut, J. (2010). Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Process. Appl.* **120** 1215–1246. MR2639745 https://doi.org/10.1016/j.spa.2010.03.009

[24] Horowitz, J. and Karandikar, R.L. (1994). Mean rates of convergence of empirical measures in the Wasserstein metric. *J. Comput. Appl. Math.* **55** 261–273. MR1329874 https://doi.org/10.1016/0377-0427(94)90033-7

[25] Huang, X., Röckner, M. and Wang, F.-Y. (2019). Nonlinear Fokker-Planck equations for probability measures on path space and path-distribution dependent SDEs. *Discrete Contin. Dyn. Syst.* **39** 3017–3035. MR3959419 https://doi.org/10.3934/dcds.2019125

[26] Kac, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. III* 171–197. Berkeley-Los Angeles, Calif.: Univ. California Press. MR0084985

[27] Kleptsyn, V. and Kurtzmann, A. (2012). Ergodicity of self-attracting motion. *Electron. J. Probab.* **17** no. 50, 37. MR2946157 https://doi.org/10.1214/EJP.v17-2121

[28] Lamberton, D. and Pagès, G. (2002). Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8** 367–405. MR1913112 https://doi.org/10.1142/S0219493703000838

[29] Lasry, J.-M. and Lions, P.-L. (2007). Mean field games. *Jpn. J. Math.* **2** 229–260. MR2295621 https://doi.org/10.1007/s11537-007-0657-8

[30] Liu, W., Wu, L. and Zhang, C. (2021). Long-time behaviors of mean-field interacting particle systems related to McKean-Vlasov equations. *Comm. Math. Phys.* **387** 179–214. MR4312362 https://doi.org/10.1007/s00220-021-04198-5

[31] 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

[32] Mehri, S., Scheutzow, M., Stannat, W. and Zangeneh, B.Z. (2020). Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.* **30** 175–207. MR4068309 https://doi.org/10.1214/19-AAP1499

[33] 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
[34] Pemantle, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79. MR2282181 https://doi.org/10.1214/07-PS094
[35] Prévôt, C. and Röckner, M. (2007). *A Concise Course on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1905**. Berlin: Springer. MR2329435
[36] Ren, P. and Wang, F.-Y. (2021). Exponential convergence in entropy and Wasserstein for McKean-Vlasov SDEs. *Nonlinear Anal.* **206** Paper No. 112259, 20. MR4206077 https://doi.org/10.1016/j.na.2021.112259
[37] Riekert, A. (2022). Convergence rates for empirical measures of Markov chains in dual and Wasserstein distances. *Statist. Probab. Lett.* **189** Paper No. 109605, 8. MR4452774 https://doi.org/10.1016/j.spl.2022.109605
[38] Sharrock, L., Kantas, N., Parpas, P. and Pavliotis, G.A. (2021). Parameter estimation for the McKean–Vlasov stochastic differential equation. arXiv preprint arXiv:2106.13751.
[39] Spitzer, F. (1991). Interaction of Markov processes. In *Random Walks, Brownian Motion, and Interacting Particle Systems* 66–110. Springer.
[40] 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
[41] Touboul, J. (2014). Propagation of chaos in neural fields. *Ann. Appl. Probab.* **24** 1298–1328. MR3199987 https://doi.org/10.1214/13-AAP950
[42] Veretennikov, A.Y. (2006). On ergodic measures for McKean-Vlasov stochastic equations. In *Monte Carlo and Quasi-Monte Carlo Methods 2004* 471–486. Berlin: Springer. MR2208726 https://doi.org/10.1007/3-540-31186-6_29
[43] Villani, C. (2009). *Optimal Transport: Old and New*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9
[44] 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
[45] Wang, F.-Y. (2021). Exponential ergodicity for non-dissipative McKean–Vlasov SDEs. arXiv preprint arXiv:2101.12562.
[46] Wang, F.-Y. (2022). Wasserstein convergence rate for empirical measures on noncompact manifolds. *Stochastic Process. Appl.* **144** 271–287. MR4347493 https://doi.org/10.1016/j.spa.2021.11.006
[47] Wang, F.-Y. and Zhu, J.-X. (2019). Limit theorems in Warsserstein distance for empirical measures of diffusion processes on Riemannian manifolds. arXiv preprint arXiv:1906.03422.
[48] Yang, Y., Luo, R., Li, M., Zhou, M., Zhang, W. and Wang, J. (2018). Mean field multi-agent reinforcement learning. In *Proceedings of the 35th International Conference on Machine Learning* (J. Dy and A. Krause, eds.). *Proceedings of Machine Learning Research* **80** 5571–5580. PMLR.
[49] Zhang, S.-Q. (2021). Existence and non-uniqueness of stationary distributions for distribution dependent SDEs. arXiv preprint arXiv:2105.04226.

Bernoulli **29**(3), 2023, 2519–2543
https://doi.org/10.3150/22-BEJ1551

Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics

WEI-LIEM LOH[a] and SAIFEI SUN[b]

Department of Statistics and Data Science, National University of Singapore, Singapore, [a]*stalohwl@nus.edu.sg,* [b]*saifeisun@u.nus.edu*

This article considers parameter estimation for a class of Gaussian random fields on $[0,1)^d$ that are observed with measurement error and irregularly spaced design sites. This class comprises Gaussian random fields with suitably smooth mean functions and isotropic powered exponential, Matérn or generalized Wendland covariance functions. Under fixed-domain asymptotics, consistent estimators are proposed for three microergodic parameters, namely the nugget, the smoothness parameter and a parameter related to the coefficient of the principal irregular term of the isotropic covariance function. Upper bounds for the convergence rate of these estimators are established. Simulations are conducted to study the finite sample accuracy of the proposed estimators.

Keywords: Convergence rate; fixed-domain asymptotics; Gaussian random field; generalized Wendland; multivariate discrete differentiation; nugget; powered exponential; quadratic variation; Matérn; space-filling design

References

Andrews, G.E., Askey, R. and Roy, R. (1999). *Special Functions*. *Encyclopedia of Mathematics and Its Applications* **71**. Cambridge: Cambridge Univ. Press. MR1688958 https://doi.org/10.1017/CBO9781107325937

Bevilacqua, M., Faouzi, T., Furrer, R. and Porcu, E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics. *Ann. Statist.* **47** 828–856. MR3909952 https://doi.org/10.1214/17-AOS1652

Chang, C.-H., Huang, H.-C. and Ing, C.-K. (2017). Mixed domain asymptotics for a stochastic process model with time trend and measurement error. *Bernoulli* **23** 159–190. MR3556770 https://doi.org/10.3150/15-BEJ740

Chen, H.-S., Simpson, D.G. and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statist. Sinica* **10** 141–156. MR1742105

Cressie, N.A.C. (1991). *Statistics for Spatial Data*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1127423

Du, J., Zhang, H. and Mandrekar, V.S. (2009). Fixed-domain asymptotic properties of tapered maximum likelihood estimators. *Ann. Statist.* **37** 3330–3361. MR2549562 https://doi.org/10.1214/08-AOS676

Golub, G.H. and Van Loan, C.F. (1983). *Matrix Computations*. *Johns Hopkins Series in the Mathematical Sciences* **3**. Baltimore, MD: Johns Hopkins Univ. Press. MR0733103

Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864 https://doi.org/10.1214/aoms/1177693335

Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183 https://doi.org/10.1017/CBO9780511810817

Kaufman, C.G. and Shaby, B.A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika* **100** 473–484. MR3068447 https://doi.org/10.1093/biomet/ass079

Keshavarz, H., Nguyen, X. and Scott, C. (2019). Local inversion-free estimation of spatial Gaussian processes. *Electron. J. Stat.* **13** 4224–4279. MR4021265 https://doi.org/10.1214/19-EJS1592

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

Loh, W.-L. (1996). On Latin hypercube sampling. *Ann. Statist.* **24** 2058–2080. MR1421161 https://doi.org/10.1214/aos/1069362310

Loh, W.-L. (2015). Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. *Ann. Statist.* **43** 2766–2794. MR3405611 https://doi.org/10.1214/15-AOS1365

Loh, W.-L. and Lam, T.-K. (2000). Estimating structured correlation matrices in smooth Gaussian random field models. *Ann. Statist.* **28** 880–904. MR1792792 https://doi.org/10.1214/aos/1015952003

Loh, W.-L. and Sun, S. (2023). Supplement to "Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics." https://doi.org/10.3150/22-BEJ1551SUPP

Loh, W.-L., Sun, S. and Wen, J. (2021). On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. *Ann. Statist.* **49** 3127–3152. MR4352525 https://doi.org/10.1214/21-aos2077

McKay, M.D., Beckman, R.J. and Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245. MR0533252 https://doi.org/10.2307/1268522

Owen, A.B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statist. Sinica* **2** 439–452. MR1187952

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

Sadhanala, V., Wang, Y.-X., Hu, A.J. and Tibshirani, R.J. (2021). Multivariate trend filtering for lattice data. Available at arXiv:2112.14758v1.

Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer. MR2160708 https://doi.org/10.1007/978-1-4757-3799-8

Stein, M. (1987). Large sample properties of simulations using Latin hypercube sampling. *Technometrics* **29** 143–151. MR0887702 https://doi.org/10.2307/1269769

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

Stein, M.L. (2012). Simulation of Gaussian random fields with one derivative. *J. Comput. Graph. Statist.* **21** 155–173. MR2913361 https://doi.org/10.1198/jcgs.2010.10069

Sun, S. (2020). *Parameter Estimation for Isotropic Gaussian Random Fields with the Generalized Wendland Covariance Function or the Matérn Covariance Function with Nugget.* PhD thesis, Department of Statistics and Applied Probability, National University of Singapore.

Tang, W., Zhang, L. and Banerjee, S. (2021). On identifiability and consistency of the nugget in Gaussian spatial process models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 1044–1070. MR4349127 https://doi.org/10.1111/rssb.12472

Wang, D. and Loh, W.-L. (2011). On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electron. J. Stat.* **5** 238–269. MR2792553 https://doi.org/10.1214/11-EJS607

Wen, J., Sun, S. and Loh, W.-L. (2021). Smoothness estimation of nonstationary Gaussian random fields from irregularly spaced data observed along a curve. *Electron. J. Stat.* **15** 6071–6150. MR4355704 https://doi.org/10.1214/21-ejs1941

Wood, A.T.A. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0,1]^d$. *J. Comput. Graph. Statist.* **3** 409–432. MR1323050 https://doi.org/10.2307/1390903

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. MR2054303 https://doi.org/10.1198/016214504000000241

Bernoulli **29**(3), 2023, 2544–2573
https://doi.org/10.3150/22-BEJ1552

Consistency of p-norm based tests in high dimensions: Characterization, monotonicity, domination

ANDERS BREDAHL KOCK[1,a] and DAVID PREINERSTORFER[2,b]

[1] *University of Oxford, Aarhus University and CREATES, 10 Manor Rd, Oxford OX1 3UQ, United Kingdom,*
[a] *anders.kock@economics.ox.ac.uk*
[2] *University of St. Gallen, SEPS-SEW, Varnbüelstrasse 14, 9000 St. Gallen, Switzerland,*
[b] *david.preinerstorfer@unisg.ch*

Many commonly used test statistics are based on a norm measuring the evidence against the null hypothesis. To understand how the choice of that norm affects power properties of tests in high dimensions, we study the *consistency sets* of p-norm based tests in the prototypical framework of sequence models with unrestricted parameter spaces, the null hypothesis being that all observations have zero mean. The consistency set of a test is here defined as the set of all arrays of alternatives the test is consistent against as the dimension of the parameter space diverges. We characterize the consistency sets of p-norm based tests and find, in particular, that the consistency against an array of alternatives cannot be determined solely in terms of the p-norm of the alternative. Our characterization also reveals an unexpected monotonicity result: namely that the consistency set is strictly increasing in $p \in (0, \infty)$, such that tests based on higher p strictly dominate those based on lower p in terms of consistency. This monotonicity property allows us to construct novel tests that dominate, with respect to their consistency behavior, *all* p-norm based tests without sacrificing asymptotic size.

Keywords: Consistency; high-dimensional testing problems; norm-based tests; power enhancement principle; sequence models

References

Anderson, T.W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176. MR0069229 https://doi.org/10.2307/2032333

Birnbaum, A. (1955). Characterizations of complete classes of tests of some multiparametric hypotheses, with applications to likelihood ratio tests. *Ann. Math. Stat.* **26** 21–36. MR0067438 https://doi.org/10.1214/aoms/1177728590

Bogachev, L. (2006). Limit laws for norms of IID samples with Weibull tails. *J. Theoret. Probab.* **19** 849–873. MR2279606 https://doi.org/10.1007/s10959-006-0036-z

Fan, J., Liao, Y. and Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica* **83** 1497–1541. MR3384226 https://doi.org/10.3982/ECTA12749

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications. Vol. I*, 3rd ed. New York: Wiley. MR0228020

Feng, L., Jiang, T., Liu, B. and Xiong, W. (2022). Max-sum tests for cross-sectional independence of high-dimensional panel data. *Ann. Statist.* **50** 1124–1143. MR4404930 https://doi.org/10.1214/21-aos2142

Giessing, A. and Fan, J. (2020). Bootstrapping ℓ_p-statistics in high dimensions. Preprint. Available at arXiv:2006.13099.

He, Y., Xu, G., Wu, C. and Pan, W. (2021). Asymptotically independent U-statistics in high-dimensional testing. *Ann. Statist.* **49** 154–181. MR4206673 https://doi.org/10.1214/20-AOS1951

Ingster, Y.I. and Suslina, I.A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. New York: Springer. MR1991446 https://doi.org/10.1007/978-0-387-21580-8

1350-7265

Jammalamadaka, S.R., Meintanis, S. and Verdebout, T. (2020). On Sobolev tests of uniformity on the circle with an extension to the sphere. *Bernoulli* **26** 2226–2252. MR4091107 https://doi.org/10.3150/19-BEJ1191

Janßen, A. (2010). Limit laws for power sums and norms of i.i.d. samples. *Probab. Theory Related Fields* **146** 515–533. MR2574737 https://doi.org/10.1007/s00440-008-0198-y

Jiang, F., Wang, R. and Shao, X. (2023). Robust inference for change points in high dimension. *J. Multivariate Anal.* **193** Paper No. 105114. MR4499388 https://doi.org/10.1016/j.jmva.2022.105114

Kock, A.B. and Preinerstorfer, D. (2019). Power in high-dimensional testing problems. *Econometrica* **87** 1055–1069. MR3957338 https://doi.org/10.3982/ECTA15844

Kock, A.B. and Preinerstorfer, D. (2023). Supplement to "Consistency of p-norm based tests in high dimensions: characterization, monotonicity, domination." https://doi.org/10.3150/22-BEJ1552SUPP

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. *Springer Series in Statistics*. New York–Berlin: Springer. MR0691492

Pinelis, I. (2010). Asymptotic efficiency of p-mean tests for means in high dimensions. Preprint. Available at arXiv:1006.0505.

Pinelis, I. (2014). Schur^2-concavity properties of Gaussian measures, with applications to hypotheses testing. *J. Multivariate Anal.* **124** 384–397. MR3147333 https://doi.org/10.1016/j.jmva.2013.11.011

Schlather, M. (2001). Limit distributions of norms of vectors of positive i.i.d. random variables. *Ann. Probab.* **29** 862–881. MR1849180 https://doi.org/10.1214/aop/1008956695

Stein, C. (1956). The admissibility of Hotelling's T^2-test. *Ann. Math. Stat.* **27** 616–623. MR0080413 https://doi.org/10.1214/aoms/1177728171

Wang, H.J., McKeague, I.W. and Qian, M. (2018). Testing for marginal linear effects in quantile regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 433–452. MR3763699 https://doi.org/10.1111/rssb.12258

Xu, G., Lin, L., Wei, P. and Pan, W. (2016). An adaptive two-sample test for high-dimensional means. *Biometrika* **103** 609–624. MR3551787 https://doi.org/10.1093/biomet/asw029

Yang, Q. and Pan, G. (2017). Weighted statistic in detecting faint and sparse alternatives for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **112** 188–200. MR3646565 https://doi.org/10.1080/01621459.2015.1122602

Yu, X., Li, D. and Xue, L. (2020). Fisher's combined probability test for high-dimensional covariance matrices. Preprint. Available at arXiv:2006.00426.

Yu, X., Li, D., Xue, L. and Li, R. (2021). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. Preprint. Available at arXiv:2109.15287.

Bernoulli **29**(3), 2023, 2574–2597
https://doi.org/10.3150/22-BEJ1553

Deep stable neural networks: Large-width asymptotics and convergence rates

STEFANO FAVARO[1,a], SANDRA FORTINI[2,b] and STEFANO PELUCHETTI[3,c]

[1]*Department of Economics and Statistics, University of Torino and Collegio Carlo Alberto, Torino, 10134, Italy,* [a]*stefano.favaro@unito.it*
[2]*Department of Decision Sciences, Bocconi University, Milano, 20136 Italy,* [b]*sandra.fortini@unibocconi.it*
[3]*Cogent Labs 36F 3-2-1 Roppongi, Minato-ku, Tokyo, Japan,* [c]*speluchetti@cogent.co.jp*

In modern deep learning, there is a recent and growing literature on the interplay between large-width asymptotic properties of deep Gaussian neural networks (NNs), i.e. deep NNs with Gaussian-distributed weights, and Gaussian stochastic processes (SPs). Motivated by empirical analyses that show the potential of replacing Gaussian distributions with Stable distributions for the NN's weights, in this paper we present a rigorous analysis of the large-width asymptotic behaviour of (fully connected) feed-forward deep Stable NNs, i.e. deep NNs with Stable-distributed weights. We show that as the width goes to infinity jointly over the NN's layers, i.e. the "joint growth" setting, a rescaled deep Stable NN converges weakly to a Stable SP whose distribution is characterized recursively through the NN's layers. Because of the non-triangular structure of the NN, this is a non-standard asymptotic problem, to which we propose an inductive approach of independent interest. Then, we establish sup-norm convergence rates of the rescaled deep Stable NN to the Stable SP, under the "joint growth" and a "sequential growth" of the width over the NN's layers. Such a result provides the difference between the "joint growth" and the "sequential growth" settings, showing that the former leads to a slower rate than the latter, depending on the depth of the layer and the number of inputs of the NN. Our work extends some recent results on infinitely wide limits for deep Gaussian NNs to the more general deep Stable NNs, providing the first result on convergence rates in the "joint growth" setting.

Keywords: Bayesian inference; deep neural network; depth limit; exchangeable sequence; Gaussian stochastic process; neural tangent kernel; infinitely wide limit; Stable stochastic process; spectral measure; sup-norm convergence rate

References

[1] Aitken, K. and Gur-Ari, G. (2020). On the asymptotics of wide networks with polynomial activations. Preprint. Available at arXiv:2006.06687.
[2] Andreassen, A. and Dyer, E. (2020). Asymptotics of wide convolutional neural networks. Preprint. Available at arXiv:2008.08675.
[3] Antognini, J.M. (2019). Finite size corrections for neural network gaussian processes. Preprint. Available at arXiv:1908.10030.
[4] Arora, S., Du, S.S., Hu, W., Li, Z., Salakhutdinov, R.R. and Wang, R. (2019). On exact computation with an infinitely wide neural net. In *Advances in Neural Information Processing Systems*.
[5] Basteri, A. and Trevisan, D. (2022). Quantitative Gaussian approximation of randomly initialized deep neural networks. Preprint. Available at arXiv:2203.07379.
[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1700749 https://doi.org/10.1002/9780470316962
[7] Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. *Ann. Math. Stat.* **33** 882–886. MR0149577 https://doi.org/10.1214/aoms/1177704456
[8] Blum, J.R., Chernoff, H., Rosenblatt, M. and Teicher, H. (1958). Central limit theorems for interchangeable processes. *Canad. J. Math.* **10** 222–229. MR0096298 https://doi.org/10.4153/CJM-1958-026-0

1350-7265

[9] Bordino, A., Favaro, S. and Fortini (2022). Infinite-wide limits for Stable deep neural networks: Sub-linear, linear and super-linear activation functions. Preprint available upon request.

[10] Byczkowski, T., Nolan, J.P. and Rajput, B. (1993). Approximation of multidimensional stable densities. *J. Multivariate Anal.* **46** 13–31. MR1231241 https://doi.org/10.1006/jmva.1993.1044

[11] Cohen, A., Cont, R., Rossier, A. and Xu, R. (2021). Scaling properties of deep residual networks. In *International Conference on Machine Learning*.

[12] Der, R. and Lee, D. (2006). Beyond Gaussian processes: On the distributions of infinite networks. In *Advances in Neural Information Processing Systems*.

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

[14] Favaro, S., Fortini, S. and Peluchetti, S. (2020). Stable behaviour of infinitely wide deep neural networks. In *International Conference on Artificial Intelligence and Statistics*.

[15] Favaro, S., Fortini, S. and Peluchetti, S. (2022). Neural tangent kernel analysis of shallow α-Stable ReLU neural networks. Preprint. Available at arXiv:2206.08065.

[16] Favaro, S., Fortini, S. and Peluchetti, S. (2023). Supplement to "Deep Stable neural networks: large-width asymptotics and convergence rates." https://doi.org/10.3150/22-BEJ1553SUPP

[17] Fortuin, V., Garriga-Alonso, A., Wenzel, F., Ratsch, G., Turner, R.E., van der Wilk, M. and Aitchison, L. (2020). Bayesian neural network priors revisited. In *Advances in Neural Information Processing Systems*.

[18] Garriga-Alonso, A., Rasmussen, C.E. and Aitchison, L. (2018). Deep convolutional networks as shallow Gaussian processes. In *International Conference on Learning Representation*.

[19] Gnedenko, B.V. and Kolmogorov, A.N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Cambridge, MA: Addison-Wesley. MR0062975

[20] Hayou, S., Doucet, A. and Rousseau, J. (2019). On the impact of the activation function on deep neural networks training. In *International Conference on Machine Learning*.

[21] Hazan, T. and Jaakkola, T. (2015). Steps toward deep kernel methods from infinite neural networks. Preprint. Available at arXiv:1508.05133.

[22] Hodgkinson, L. and Mahoney, M. (2021). Multiplicative noise and heavy tails in stochastic optimization. In *International Conference on Machine Learning*.

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

[24] Klukowski, A. (2021). Rate of convergence of polynomial networks to Gaussian processes. Preprint. Available at arXiv:2111.03175.

[25] LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature* **521** 436–444.

[26] Lee, J., Schoenholz, S., Pennington, J., Adlam, B., Xiao, L., Novak, R. and Sohl-Dickstein, J. (2020). Finite versus infinite neural networks: An empirical study. In *Advances in Neural Information Processing Systems*.

[27] Lee, J., Sohldickstein, J., Pennington, J., Novak, R., Schoenholz, S. and Bahri, Y. (2018). Deep neural networks as Gaussian processes. In *International Conference on Learning Representation*.

[28] Lee, J., Xiao, L., Schoenholz, S., Bahri, Y., Sohl-Dickstein, J. and Pennington, J. (2019). Wide neural networks of any depth evolve as linear models under gradient descent. In *Advances in Neural Information Processing Systems*.

[29] Li, C., Dunlop, M. and Stadler, G. (2022). Bayesian neural network priors for edge-preserving inversion. *Inverse Probl. Imaging* **16** 1229–1254. MR4475851 https://doi.org/10.3934/ipi.2022022

[30] Li, M.B., Nica, M. and Roy, D.M. (2021). The future is log-Gaussian: ResNets and their infinite-depth-and-width limit at initialization. Preprint. Available at arXiv:2106.04013.

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

[32] Neal, R.M. (1996). *Bayesian Learning for Neural Networks*. Berlin: Springer.

[33] Nolan, J.P. (2010). Metrics for multivariate stable distributions. In *Stability in Probability*. *Banach Center Publ.* **90** 83–102. Warsaw: Polish Acad. Sci. Inst. Math. MR2798853 https://doi.org/10.4064/bc90-0-6

[34] Nolan, J.P. (2010). An overview of multivariate Stable distributions. Department of Mathematics and Statistics at American University. Preprint.

[35] Novak, R., Xiao, L., Bahri, Y., Lee, J., Yang, G., Hron, J., Abolafia, D., Pennington, J. and Sohldickstein, J. (2018). Bayesian deep convolutional networks with many channels are Gaussian processes. In *International Conference on Learning Representation*.

[36] Poole, B., Lahiri, S., Raghu, M., Sohl-Dickstein, J. and Ganguli, S. (2016). Exponential expressivity in deep neural networks through transient chaos. In *Advances in Neural Information Processing Systems*.

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

[38] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes:Stochastic Models with Infinite Variance*. *Stochastic Modeling*. New York: CRC Press. MR1280932

[39] Schoenholz, S., Gilmer, J., Ganguli, S. and Sohl-Dickstein, J. (2017). Deep information propagation. In *International Conference on Learning Representation*.

[40] von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Stat.* **36** 299–303. MR0170407 https://doi.org/10.1214/aoms/1177700291

[41] Williams, C.K. (1997). Computing with infinite networks. In *Advances in Neural Information Processing Systems*.

[42] Yang, G. (2019). Scaling limits of wide neural networks with weight sharing: Gaussian process behavior, gradient independence, and neural tangent kernel derivation. Preprint. Available at arXiv:1902.04760.

[43] Yang, G. (2019). Tensor programs I: Wide feedforward or recurrent neural networks of any architecture are Gaussian processes. Preprint. Available at arXiv:1910.12478.