BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Five Number Two May 2019 ISSN: 1350-7265

CONTENTS

(continued)

The papers published in Bernoulli are indexed or abstracted in *Current Index to Statistics*, *Mathematical Reviews*, *Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z)*, and *Zentralblatt für Mathematik* (also avalaible on the *MATH via STN* database and *Compact MATH CD-ROM*). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Five Number Two May 2019 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 25 Number 2 May 2019 Pages 793–1601 ISI/BS

Volume 25 Number 2   May 2019   ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope
BERNOULLI is the journal of the Bernoulli Society for Mathematical Statistics and Probability, issued four times per year. The journal provides a comprehensive account of important developments in the fields of statistics and probability, offering an international forum for both theoretical and applied work.

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: http://www.bernoulli-society.org/index.php/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

The papers published in Bernoulli are indexed or abstracted in *Current Index to Statistics, Mathematical Reviews, Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z), Thomson Scientific* and *Zentralblatt für Mathematik* (also available on the *MATH via STN* database and *Compact MATH CD-ROM*). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

In 2019 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **25**(2), 2019, 793–827
https://doi.org/10.3150/17-BEJ986

Expansion for moments of regression quantiles with applications to nonparametric testing

ENNO MAMMEN[1], INGRID VAN KEILEGOM[2] and KYUSANG YU[3]

[1]*Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany. E-mail: mammen@math.uni-heidelberg.de*
[2]*ORSTAT, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. E-mail: ingrid.vankeilegom@kuleuven.be*
[3]*Department of Applied Statistics, Konkuk University, Seoul 143-701, Korea. E-mail: kyusangu@konkuk.ac.kr*

We discuss nonparametric tests for parametric specifications of regression quantiles. The test is based on the comparison of parametric and nonparametric fits of these quantiles. The nonparametric fit is a Nadaraya–Watson quantile smoothing estimator.

An asymptotic treatment of the test statistic requires the development of new mathematical arguments. An approach that makes only use of plugging in a Bahadur expansion of the nonparametric estimator is not satisfactory. It requires too strong conditions on the dimension and the choice of the bandwidth.

Our alternative mathematical approach requires the calculation of moments of Nadaraya–Watson quantile regression estimators. This calculation is done by application of higher order Edgeworth expansions.

Keywords: Bahadur expansions; goodness-of-fit tests; kernel smoothing; nonparametric regression; nonparametric testing; quantiles

References

[1] Aït-Sahalia, Y., Fan, J. and Peng, H. (2009). Nonparametric transition-based tests for jump diffusions. *J. Amer. Statist. Assoc.* **104** 1102–1116. MR2750239
[2] Angrist, J., Chernozhukov, V. and Fernández-Val, I. (2006). Quantile regression under misspecification, with an application to the U.S. wage structure. *Econometrica* **74** 539–563. MR2207400
[3] Bhattacharya, R.N. and Ranga Rao, R. (1976). *Normal Approximation and Asymptotic Expansions*. *Wiley Series in Probability and Mathematical Statistics* Wiley, New York. MR0436272
[4] Bierens, H.J. and Ginther, D. (2001). Integrated conditional moment testing of quantile regression models. *Empir. Econ.* **26** 307–324.
[5] Chao, S.-K., Volgushev, S. and Cheng, G. (2017). Quantile processes for semi and nonparametric regression. *Electron. J. Stat.* **11** 3272–3331. MR3708539
[6] Chaudhuri, P. (1991). Nonparametric estimates of regression quantiles and their local Bahadur representation. *Ann. Statist.* **19** 760–777. MR1105843
[7] Conde-Amboage, M., Sánchez-Sellero, C. and González-Manteiga, W. (2015). A lack-of-fit test for quantile regression models with high-dimensional covariates. *Comput. Statist. Data Anal.* **88** 128–138. MR3332022
[8] De Backer, M., El Ghouch, A. and Van Keilegom, I. (2017). Semiparametric copula quantile regression for complete or censored data. *Electron. J. Stat.* **11** 1660–1698. MR3639560

[9] de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277. MR0885466
[10] Dette, H. and Spreckelsen, I. (2004). Some comments on specification tests in nonparametric absolutely regular processes. *J. Time Series Anal.* **25** 159–172. MR2045571
[11] El Ghouch, A. and Van Keilegom, I. (2009). Local linear quantile regression with dependent censored data. *Statist. Sinica* **19** 1621–1640. MR2589201
[12] Engel, E. (1857). Die vorherrschenden Gewerbszweige in den Gerichtsämtern mit Beziehung auf die Productions- und Consumtionsverhältnisse des Königreichs Sachsen. II. Das Gesetz der Dichtigkeit. *Zeitschrift des statistischen Bureaus des Königlich Sächsischen Ministerium des Inneren* 3 **8–9** 153–182.
[13] Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.* **29** 153–193. MR1833962
[14] Gao, J. and Hong, Y. (2008). Central limit theorems for generalized U-statistics with applications in nonparametric specification. *J. Nonparametr. Stat.* **20** 61–76. MR2396276
[15] González Manteiga, W. and Cao, R. (1993). Testing the hypothesis of a general linear model using nonparametric regression estimation. *TEST* **2** 161–188. MR1265489
[16] Guerre, E. and Lavergne, P. (2002). Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory* **18** 1139–1171. MR1926017
[17] Guerre, E. and Sabbah, C. (2012). Uniform bias study and Bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory* **28** 87–129. MR2899215
[18] Haag, B.R. (2008). Non-parametric regression tests using dimension reduction techniques. *Scand. J. Stat.* **35** 719–738. MR2468872
[19] Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21** 1926–1947. MR1245774
[20] He, X. and Ng, P. (1999). Quantile splines with several covariates. *J. Statist. Plann. Inference* **75** 343–352. The Seventh Eugene Lukacs Conference (Bowling Green, OH, 1997). MR1678981
[21] He, X., Ng, P. and Portnoy, S. (1998). Bivariate quantile smoothing splines. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 537–550. MR1625950
[22] He, X. and Zhu, L.-X. (2003). A lack-of-fit test for quantile regression. *J. Amer. Statist. Assoc.* **98** 1013–1022. MR2041489
[23] Hjellvik, V., Yao, Q. and Tjøstheim, D. (1998). Linearity testing using local polynomial approximation. *J. Statist. Plann. Inference* **68** 295–321. MR1629587
[24] Hoderlein, S. and Mammen, E. (2009). Identification and estimation of local average derivatives in non-separable models without monotonicity. *Econom. J.* **12** 1–25. MR2500194
[25] Hong, S.-Y. (2003). Bahadur representation and its applications for local polynomial estimates in nonparametric M-regression. *J. Nonparametr. Stat.* **15** 237–251. MR1981463
[26] Horowitz, J.L. and Spokoiny, V.G. (2002). An adaptive, rate-optimal test of linearity for median regression models. *J. Amer. Statist. Assoc.* **97** 822–835. MR1941412
[27] Ingster, Yu.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I. *Math. Methods Statist.* **2** 85–114. MR1257978
[28] Ingster, Yu.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. II. *Math. Methods Statist.* **2** 171–189. MR1257983
[29] Ingster, Yu.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. III. *Math. Methods Statist.* **2** 249–268. MR1259685
[30] Koenker, R. (2005). *Quantile Regression*. *Econometric Society Monographs* **38**. Cambridge Univ. Press, Cambridge. MR2268657
[31] Koenker, R. and Machado, J.A.F. (1999). Goodness of fit and related inference processes for quantile regression. *J. Amer. Statist. Assoc.* **94** 1296–1310. MR1731491

[32] Koenker, R. and Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica* **70** 1583–1612. MR1929979

[33] Kong, E., Linton, O. and Xia, Y. (2010). Uniform Bahadur representation for local polynomial estimates of M-regression and its application to the additive model. *Econometric Theory* **26** 1529–1564. MR2684794

[34] Kreiss, J.-P., Neumann, M.H. and Yao, Q. (2008). Bootstrap tests for simple structures in nonparametric times series regression. *Stat*. *Interface* **1** 367–380. MR2476752

[35] Lee, Y.K. and Lee, E.R. (2008). Kernel methods for estimating derivatives of conditional quantiles. *J*. *Korean Statist*. *Soc*. **37** 365–373. MR2467903

[36] Leucht, A. (2012). Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli* **18** 552–585. MR2922461

[37] Li, Q. and Racine, J.S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *J*. *Bus*. *Econom*. *Statist*. **26** 423–434. MR2459343

[38] Rothe, C. and Wied, D. (2013). Misspecification testing in a class of conditional distributional models. *J*. *Amer*. *Statist*. *Assoc*. **108** 314–324. MR3174622

[39] Su, L. and White, H.L. (2012). Conditional independence specification testing for dependent processes with local polynomial quantile regression. In *Essays in Honor of Jerry Hausman*. *Adv*. *Econom*. **29** 355–434. Emerald, Bingley. MR3494865

[40] Truong, Y.K. (1989). Asymptotic properties of kernel estimators based on local medians. *Ann*. *Statist*. **17** 606–617. MR0994253

[41] Volgushev, S., Birke, M., Dette, H. and Neumeyer, N. (2013). Significance testing in quantile regression. *Electron*. *J*. *Stat*. **7** 105–145. MR3020416

[42] Volgushev, S., Chao, S. and Cheng, G. (2017). Distributed inference for quantile regression processes. Preprint. Available at arXiv:1701.060088v3.

[43] Yu, K. and Jones, M.C. (1998). Local linear quantile regression. *J*. *Amer*. *Statist*. *Assoc*. **93** 228–237. MR1614628

[44] Zheng, J.X. (1996). A consistent test of functional form via nonparametric estimation techniques. *J*. *Econometrics* **75** 263–289. MR1413644

[45] Zheng, J.X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* **14** 123–138. MR1613710

Bernoulli **25**(2), 2019, 828–847
https://doi.org/10.3150/17-BEJ997

On squared Bessel particle systems

PIOTR GRACZYK[1] and JACEK MAŁECKI[2]

[1]*LAREMA, Université d'Angers, 2 Bd Lavoisier, 49045 Angers cedex 1, France.*
E-mail: piotr.graczyk@univ-angers.fr
[2]*Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, ul. Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland. E-mail: jacek.malecki@pwr.edu.pl*

We study the existence and uniqueness of solutions of SDEs describing squared Bessel particle systems in full generality. We define nonnegative and non-colliding squared Bessel particle systems and we study their properties. Particle systems dissatisfying non-colliding and unicity properties are pointed out. The structure of squared Bessel particle systems is described.

Keywords: non-colliding solution; particle system; squared Bessel process; stochastic differential equation; Wishart process

References

[1] Bru, M.-F. (1989). Diffusions of perturbed principal component analysis. *J. Multivariate Anal.* **29** 127–136. MR0991060
[2] Bru, M.-F. (1991). Wishart processes. *J. Theoret. Probab.* **4** 725–751. MR1132135
[3] Donati-Martin, C., Doumerc, Y., Matsumoto, H. and Yor, M. (2004). Some properties of the Wishart processes and a matrix extension of the Hartman–Watson laws. *Publ. Res. Inst. Math. Sci.* **40** 1385–1412. MR2105711
[4] Göing-Jaeschke, A. and Yor, M. (2003). A survey and some generalizations of Bessel processes. *Bernoulli* **9** 313–349.
[5] Graczyk, P. and Małecki, J. (2013). Multidimensional Yamada–Watanabe theorem and its applications to particle systems. *J. Math. Phys.* **54** 021503, 15 pp.
[6] Graczyk, P. and Małecki, J. (2014). Strong solutions of non-colliding particle systems. *Electron. J. Probab.* **19** 1–21. MR3296535
[7] Graczyk, P., Małecki, J. and Mayerhofer, E. (2018). A characterizations of Wishart processes and Wishart distributions. *Stoch. Proc. Appl.* **128** 1386–1404.
[8] Katori, M. (2016). *Bessel Processes, Schramm–Loewner Evolution, and the Dyson Model*. Tokyo: Springer.
[9] Katori, M. and Tanemura, H. (2004). Symmetry of matrix-valued stochastic processes and non-colliding diffusion particle systems. *J. Math. Phys.* **45** 3058–3085.
[10] König, W. and O'Connell, N. (2001). Eigenvalues of the Laguerre process as noncolliding squared Bessel process. *Electron. Commun. Probab.* **6** 107–114.
[11] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*. New York: Springer.
[12] Rogers, L.C.G. and Shi, Z. (1993). Interacting Brownian particles and the Wigner law. *Probab. Theory Related Fields* **95** 555–570.
[13] Ross, K.A. and Wright, C.R.B. (2003). *Discrete Mathematics*, 5th ed. New York: Prentice Hall.
[14] Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** 155–167.

Bernoulli **25**(2), 2019, 848–876
https://doi.org/10.3150/17-BEJ1005

Smooth, identifiable supermodels of discrete DAG models with latent variables

ROBIN J. EVANS[1] and THOMAS S. RICHARDSON[2]

[1]*Department of Statistics, University of Oxford, 24–29 St Giles', Oxford, OX1 3LB, UK. E-mail: evans@stats.ox.ac.uk*
[2]*Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195, USA. E-mail: thomasr@u.washington.edu*

We provide a parameterization of the discrete nested Markov model, which is a supermodel that approximates DAG models (Bayesian network models) with latent variables. Such models are widely used in causal inference and machine learning. We explicitly evaluate their dimension, show that they are curved exponential families of distributions, and fit them to data. The parameterization avoids the irregularities and unidentifiability of latent variable models. The parameters used are all fully identifiable and causally-interpretable quantities.

Keywords: Bayesian network; DAG; nested Markov model; parameterization

References

[1] Bishop, C.M. (2007). *Pattern Recognition and Machine Learning*. *Information Science and Statistics*. New York: Springer. MR2247587
[2] Darwiche, A. (2009). *Modeling and Reasoning with Bayesian Networks*. Cambridge: Cambridge Univ. Press. MR2572244
[3] Dawid, A.P. (2002). Influence diagrams for causal modelling and inference. *Int*. *Stat*. *Rev*. **70** 161–189.
[4] Drton, M. (2009). Discrete chain graph models. *Bernoulli* **15** 736–753. MR2555197
[5] Drton, M. (2009). Likelihood ratio tests and singularities. *Ann*. *Statist*. **37** 979–1012. MR2502658
[6] Drton, M. and Richardson, T.S. (2008). Binary models for marginal independence. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **70** 287–309. MR2424754
[7] Evans, R.J. (2018). Margins of discrete Bayesian networks. *Ann*. *Statist*. **46** 2623–2656.
[8] Evans, R.J. and Richardson, T.S. (2010). Maximum likelihood fitting of acyclic directed mixed graphs to binary data. In *Proceedings of the* 26*th Conference on Uncertainty in Artificial Intelligence* 177–184.
[9] Evans, R.J. and Richardson, T.S. (2013). Marginal log-linear parameters for graphical Markov models. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **75** 743–768. MR3091657
[10] Evans, R.J. and Richardson, T.S. (2014). Markovian acyclic directed mixed graphs for discrete data. *Ann*. *Statist*. **42** 1452–1482. MR3262457
[11] Hauser, R.M., Sewell, W.H. and Herd, P. Wisconsin Longitudinal Study (WLS), 1957–2012. Available at http://www.ssc.wisc.edu/wlsresearch/documentation/. Version 13.03, Univ. Wisconsin–Madison, WLS.
[12] Huang, J.C. and Frey, B.J. (2008). Cumulative distribution networks and the derivative-sum-product algorithm. In *Proceedings of the* 24*th Conference on Uncertainty in Artificial Intelligence* 290–297.

1350-7265

[13] Mond, D., Smith, J. and van Straten, D. (2003). Stochastic factorizations, sandwiched simplices and the topology of the space of explanations. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **459** 2821–2845. MR2015992
[14] Pearl, J. and Verma, T.S. (1992). A statistical semantics for causation. *Stat. Comput.* **2** 91–95.
[15] Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2548166
[16] Richardson, T. (2003). Markov properties for acyclic directed mixed graphs. *Scand. J. Stat.* **30** 145–157. MR1963898
[17] Richardson, T.S., Evans, R.J., Robins, J.M. and Shpitser, I. (2017). Nested Markov properties for acyclic directed mixed graphs. Preprint. Available at arXiv:1701.06686.
[18] Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period – Application to control of the healthy worker survivor effect. *Math. Model.* **7** 1393–1512. MR0877758
[19] Shpitser, I., Evans, R.J., Richardson, T.S. and Robins, J.M. (2013). Sparse nested Markov models with log-linear parameters. In *Proceedings of the* 29*th Conference on Uncertainty in Artificial Intelligence* 576–585.
[20] Shpitser, I., Evans, R.J., Richardson, T.S. and Robins, J.M. (2014). Introduction to nested Markov models. *Behaviormetrika* **41** 3–39.
[21] Shpitser, I. and Pearl, J. (2008). Dormant independence. Technical Report R-340, Cognitive Systems Laboratory, University of California, Los Angeles.
[22] Shpitser, I., Richardson, T.S., Robins, J.M. and Evans, R.J. (2011). Parameter and structure learning in mixed graph models of post-truncation independence. Draft.
[23] Silva, R., Blundell, C. and Teh, Y.W. (2011). Mixed cumulative distribution networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (*AISTATS*) **15** 670–678.
[24] Silva, R. and Ghahramani, Z. (2009). The hidden life of latent variables: Bayesian learning with mixed graph models. *J. Mach. Learn. Res.* **10** 1187–1238. MR2520804
[25] Richardson, T.S. Spirtes, P.L. and (2002). Ancestral graph Markov models. *Ann. Statist.* **30** 962-1030. MR1926166
[26] Tian, J. (2002). Studies in causal reasoning and learning. Ph.D. thesis, University of California, Los Angeles.
[27] Tian, J. and Pearl, J. (2002). A general identification condition for causal effects. In *Proceedings of the* 18*th National Conference on Artificial Intelligence*. AAAI.
[28] Tian, J. and Pearl, J. (2002). On the testable implications of causal models with hidden variables. In *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence* (*UAI*-02) 519–527. Morgan Kaufmann Publishers Inc.
[29] Verma, T.S. and Pearl, J. (1991). Equivalence and synthesis of causal models. In *Proceedings of the* 7*th Conference on Uncertainty in Artificial Intelligence* (*UAI*-91) 255–268.
[30] Wermuth, N. (2011). Probability distributions with summary graph structure. *Bernoulli* **17** 845–879. MR2817608

Bernoulli **25**(2), 2019, 877–901
https://doi.org/10.3150/17-BEJ1007

Bayesian consistency for a nonparametric stationary Markov model

MINWOO CHAE[1] and STEPHEN G. WALKER[2]

[1]*Department of Mathematics, Applied Mathematics and Statistics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA. E-mail: minwooo.chae@gmail.com*
[2]*Department of Mathematics, The University of Texas at Austin, 2515 Speedway, Austin, TX 78712, USA E-mail: s.g.walker@math.utexas.edu*

We consider posterior consistency for a Markov model with a novel class of nonparametric prior. In this model, the transition density is parameterized via a mixing distribution function. Therefore, the Wasserstein distance between mixing measures can be used to construct neighborhoods of a transition density. The Wasserstein distance is sufficiently strong, for example, if the mixing distributions are compactly supported, it dominates the sup-L_1 metric. We provide sufficient conditions for posterior consistency with respect to the Wasserstein metric provided that the true transition density is also parametrized via a mixing distribution. In general, when it is not be parameterized by a mixing distribution, we show the posterior distribution is consistent with respect to the average L_1 metric. Also, we provide a prior whose support is sufficiently large to contain most smooth transition densities.

Keywords: Kullback–Leibler support; mixtures; nonparametric Markov model; posterior consistency; Wasserstein metric

References

[1] Antoniano-Villalobos, I. and Walker, S.G. (2015). Bayesian consistency for Markov models. *Sankhya A* **77** 106–125. MR3317483
[2] Antoniano-Villalobos, I. and Walker, S.G. (2016). A nonparametric model for stationary time series. *J. Time Series Anal.* **37** 126–142. MR3439535
[3] Bissiri, P.G. and Ongaro, A. (2014). On the topological support of species sampling priors. *Electron. J. Stat.* **8** 861–882.
[4] Bruni, C. and Koch, G. (1985). Identifiability of continuous mixtures of unknown Gaussian distributions. *Ann. Probab.* **13** 1341–1357.
[5] Chen, X. and Fan, Y. (2006). Estimation of copula-based semiparametric time series models. *J. Econometrics* **130** 307–335. MR2211797
[6] Chen, X., Wu, W.B. and Yi, Y. (2009). Efficient estimation of copula-based semiparametric Markov models. *Ann. Statist.* **37** 4214–4253.
[7] Darsow, W.F., Nguyen, B. and Olsen, E.T. (1992). Copulas and Markov processes. *Illinois J. Math.* **36** 600–642. MR1215798
[8] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949
[9] Ghosal, S. and Tang, Y. (2006). Bayesian consistency for Markov processes. *Sankhyā* **68** 227–239.
[10] Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.

[11] Glynn, P.W. and Ormoneit, D. (2002). Hoeffding's inequality for uniformly ergodic Markov chains. *Statist*. *Probab*. *Lett*. **56** 143–146. MR1881167
[12] Hjort, N. L., Holmes, C., Müller, P. and Walker, S. G. (2010). *Bayesian Nonparametrics*. Cambridge University Press.
[13] Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Boca Raton: CRC Press.
[14] Kantorovich, L.V. and Rubinstein, G.Š. (1958). On a space of completely additive functions. *Vestnik Leningrad Univ*. *Math*. **13** 52–59. MR0102006
[15] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates: I. Density estimates. *Ann*. *Statist*. **12** 351–357.
[16] Loève, M. (1963). *Probability Theory*. 3rd ed. D. Van Nostrand: Princeton, N.J.-Toronto, Ont.-London. MR0203748
[17] Majumdar, S. (1992). On topological support of Dirichlet prior. *Statist*. *Probab*. *Lett*. **15** 385–388.
[18] Mena, R.H. and Walker, S.G. (2005). Stationary autoregressive models via a Bayesian nonparametric approach. *J*. *Time Series Anal*. **26** 789–805.
[19] Merkle, M. (2000). Topics in weak convergence of probability measures. *Zb*. *Rad*. (*Beogr*.) **9** 235–274.
[20] Meyn, S.P. and Tweedie, R.L. (2012). *Markov Chains and Stochastic Stability*. New York: Springer Science & Business Media.
[21] Nelsen, R.B. (2003). Properties and applications of copulas: A brief survey. In *Proceedings of the First Brazilian Conference on Statistical Modeling in Insurance and Finance* (J. Dhaene, N. Kolev and P. Morettin, eds.) 10–28. Sao Paulo: Univ. Press USP.
[22] Nguyen, X. (2013). Convergence of latent mixing measures in finite and infinite mixture models. *Ann*. *Statist*. **41** 370–400.
[23] Peel, D. and McLachlan, G.J. (2000). Robust mixture modelling using the t distribution. *Stat*. *Comput*. **10** 339–348.
[24] Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris* **8** 229–231.
[25] Tallis, G. and Chesson, P. (1982). Identifiability of mixtures. *J*. *Aust*. *Math*. *Soc*. **32** 339–348.
[26] Tang, Y. and Ghosal, S. (2007). Posterior consistency of Dirichlet mixtures for estimating a transition density. *J*. *Statist*. *Plann*. *Inference* **137** 1711–1726. MR2323858
[27] Teicher, H. (1960). On the mixture of distributions. *Ann*. *Math*. *Stat*. **31** 55–73.
[28] Teicher, H. (1963). Identifiability of finite mixtures. *Ann*. *Math*. *Stat*. **34** 1265–1269.
[29] Teicher, H. et al. (1961). Identifiability of mixtures. *Ann*. *Math*. *Stat*. **32** 244–248.
[30] Walker, S. (2003). On sufficient conditions for Bayesian consistency. *Biometrika* **90** 482–488. MR1986664
[31] Walker, S. (2004). New approaches to Bayesian consistency. *Ann*. *Statist*. **32** 2028–2043.
[32] Wu, J., Wang, X. and Walker, S.G. (2014). Bayesian nonparametric inference for a multivariate copula function. *Methodol*. *Comput*. *Appl*. *Probab*. **16** 747–763. MR3239818
[33] Wu, J., Wang, X. and Walker, S.G. (2015). Bayesian nonparametric estimation of a copula. *J*. *Stat*. *Comput*. *Simul*. **85** 103–116.
[34] Wu, Y. and Ghosal, S. (2010). The L_1-consistency of Dirichlet mixtures in multivariate Bayesian density estimation. *J*. *Multivariate Anal*. **101** 2411–2419. MR2719871

Bernoulli **25**(2), 2019, 902–931
https://doi.org/10.3150/17-BEJ1008

Low-frequency estimation of continuous-time moving average Lévy processes

DENIS BELOMESTNY[1,2], VLADIMIR PANOV[2] and JEANNETTE H.C. WOERNER[3]

[1]*University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany. E-mail: denis.belomestny@uni-due.de*
[2]*National Research University Higher School of Economics, Shabolovka, 26, 119049 Moscow, Russia. E-mail: vpanov@hse.ru*
[3]*Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany. E-mail: jwoerner@mathematik.uni-dortmund.de*

In this paper, we study the problem of statistical inference for a continuous-time moving average Lévy process of the form

$$Z_t = \int_{\mathbb{R}} \mathcal{K}(t-s)\, dL_s, \qquad t \in \mathbb{R},$$

with a deterministic kernel $\mathcal{K}$ and a Lévy process L. Especially the estimation of the Lévy measure ν of L from low-frequency observations of the process Z is considered. We construct a consistent estimator, derive its convergence rates and illustrate its performance by a numerical example. On the mathematical level, we establish some new results on exponential mixing for continuous-time moving average Lévy processes.

Keywords: low-frequency estimation; Mellin transform; moving average

References

[1] Barndorff-Nielsen, O.E., Benth, F.E. and Veraart, A.E.D. (2015). Cross-commodity modelling by multivariate ambit fields. In *Commodities*, *Energy and Environmental Finance*. *Fields Inst*. *Commun*. **74** 109–148. Fields Inst. Res. Math. Sci., Toronto, ON. MR3380393
[2] Barndorff-Nielsen, O.E. and Schmiegel, J. (2009). Brownian semistationary processes and volatility/intermittency. In *Advanced Financial Modelling*. *Radon Ser. Comput*. *Appl*. *Math*. **8** 1–25. Walter de Gruyter, Berlin. MR2648456
[3] Basse, A. and Pedersen, J. (2009). Lévy driven moving averages and semimartingales. *Stochastic Process*. *Appl*. **119** 2970–2991. MR2554035
[4] Basse-O'Connor, A., Lachieze-Rey, R. and Podolskij, M. (2015). Limit theorems for stationary increments Lévy driven moving averages. *CREATES Research Papers* **2015**.
[5] Basse-O'Connor, A. and Rosiński, J. (2016). On infinitely divisible semimartingales. *Probab. Theory Related Fields* **164** 133–163.
[6] Belomestny, D. and Goldenschluger, A. (2017). Nonparametric density estimation from observations with multiplicative measurement errors. Available at arXiv:1709.00629.

1350-7265

[7] Belomestny, D. and Reiss, M. (2015). Estimation and calibration of Lévy models via Fourier methods. In *Lévy matters IV. Estimation for discretly observed Lévy processes*. 1–76. Springer.

[8] Belomestny, D. and Schoenmakers, J. (2016). Statistical inference for time-changed Lévy processes via Mellin transform approach. *Stochastic Process. Appl.* **126** 2092–2122. MR3483748

[9] Bender, C., Lindner, A. and Schicks, M. (2012). Finite variation of fractional Lévy processes. *J. Theoret. Probab.* **25** 594–612.

[10] Brockwell, P. and Lindner, A. (2012). Ornstein-Uhlenbeck related models driven by Lévy processes. In *Statistical Methods for Stochastic Differential Equations. Monogr. Statist. Appl. Probab.* **124** 383–427. Boca Raton, FL: CRC Press.

[11] Cohen, S. and Lindner, A. (2013). A central limit theorem for the sample autocorrelations of a Lévy driven continuous time moving average process. *J. Statist. Plann. Inference* **143** 1295–1306.

[12] Glaser, S. (2015). A law of large numbers for the power variation of fractional Lévy processes. *Stoch. Anal. Appl.* **33** 1–20. MR3285245

[13] Merlevéde, F., Peligrad, M. and Rio, E. (2009). Bernstein inequality and moderate deviation under strong mixing conditions. In *High Dimensional Probability*, *IMS Collections* 273–292. IMS.

[14] Oberhettinger, F. (1974). *Tables of Mellin Transforms*. Berlin: Springer.

[15] Podolskij, M. (2015). Ambit fields: Survey and new challenges. In *XI Symposium on Probability and Stochastic Processes* 241–279. Springer.

[16] Rajput, B.S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields* **82** 451–487. MR1001524

[17] Schnurr, A. and Woerner, J.H.C. (2011). Well-balanced Lévy driven Ornstein-Uhlenbeck processes. *Stat. Risk Model.* **28** 343–357. MR2877570

[18] Zhang, S., Lin, Z. and Zhang, X. (2015). A least squares estimator for Lévy-driven moving averages based on discrete time observations. *Comm. Statist. Theory Methods* **44** 1111–1129. MR3325371

Bernoulli **25**(2), 2019, 932–976
https://doi.org/10.3150/17-BEJ1009

Fréchet means and Procrustes analysis in Wasserstein space

YOAV ZEMEL* and VICTOR M. PANARETOS**

Institut de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.
E-mail: *yoav.zemel@epfl.ch; **victor.panaretos@epfl.ch

We consider two statistical problems at the intersection of functional and non-Euclidean data analysis: the determination of a Fréchet mean in the Wasserstein space of multivariate distributions; and the optimal registration of deformed random measures and point processes. We elucidate how the two problems are linked, each being in a sense dual to the other. We first study the finite sample version of the problem in the continuum. Exploiting the tangent bundle structure of Wasserstein space, we deduce the Fréchet mean via gradient descent. We show that this is equivalent to a Procrustes analysis for the registration maps, thus only requiring successive solutions to pairwise optimal coupling problems. We then study the population version of the problem, focussing on inference and stability: in practice, the data are i.i.d. realisations from a law on Wasserstein space, and indeed their observation is discrete, where one observes a proxy finite sample or point process. We construct regularised nonparametric estimators, and prove their consistency for the population mean, and uniform consistency for the population Procrustes registration maps.

Keywords: functional data analysis; manifold statistics; Monge–Kantorovich problem; multimarginal transportation; optimal transportation; phase variation; point process; random measure; registration; shape theory; warping

References

[1] Afsari, B., Tron, R. and Vidal, R. (2013). On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM J. Control Optim.* **51** 2230–2260. MR3057324
[2] Agueh, M. and Carlier, G. (2011). Barycenters in the Wasserstein space. *SIAM J. Math. Anal.* **43** 904–924. MR2801182
[3] Alberti, G. and Ambrosio, L. (1999). A geometrical approach to monotone functions in $\mathbf{R}^n$. *Math. Z.* **230** 259–316. MR1676726
[4] Allassonnière, S., Amit, Y. and Trouvé, A. (2007). Towards a coherent statistical framework for dense deformable template estimation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 3–29.
[5] Álvarez-Esteban, P.C., del Barrio, E., Cuesta-Albertos, J.A. and Matrán, C. (2011). Uniqueness and approximate computation of optimal incomplete transportation plans. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 358–375.
[6] Álvarez-Esteban, P.C., del Barrio, E., Cuesta-Albertos, J.A. and Matrán, C. (2016). A fixed-point approach to barycenters in Wasserstein space. *J. Math. Anal. Appl.* **441** 744–762.
[7] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. London: Springer.
[8] Amit, Y., Grenander, U. and Piccioni, M. (1991). Structural image restoration through deformable templates. *J. Amer. Statist. Assoc.* **86** 376–387.

[9] Anderes, E., Borgwardt, S. and Miller, J. (2016). Discrete Wasserstein barycenters: Optimal transport for discrete data. *Math. Methods Oper. Res.* 1–21.
[10] Benamou, J.-D. and Brenier, Y. (2000). A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84** 375–393. MR1738163
[11] Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L. and Peyré, G. (2015). Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.* **37** A1111–A1138.
[12] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* 1196–1217.
[13] Bigot, J., Gouet, R., Klein, T. and López, A. (2013). Geodesic PCA in the Wasserstein space. Preprint. Available at arXiv:1307.7721.
[14] Bigot, J. and Klein, T. (2012). Consistent estimation of a population barycenter in the wasserstein space. ArXiv e-prints.
[15] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. New York: Wiley. MR1700749
[16] Boissard, E., Le Gouic, T., Loubes, J.-M. et al. (2015). Distribution's template estimate with Wasserstein metrics. *Bernoulli* **21** 740–759.
[17] Bolstad, B.M., Irizarry, R.A., Åstrand, M. and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19** 185–193.
[18] Bonneel, N., Peyré, G. and Cuturi, M. (2016). Wasserstein barycentric coordinates: Histogram regression using optimal transport. *ACM Trans. Graph.* **35** 71–1.
[19] Bonneel, N., Rabin, J., Peyré, G. and Pfister, H. (2015). Sliced and Radon Wasserstein barycenters of measures. *J. Math. Imaging Vision* **51** 22–45. MR3300482
[20] Bookstein, F.L. (1997). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge Univ. Press. MR1469220
[21] Caffarelli, L.A. (1992). The regularity of mappings with a convex potential. *J. Amer. Math. Soc.* **5** 99–104.
[22] Carlier, G., Oberman, A. and Oudet, É. (2015). Numerical methods for matching for teams and Wasserstein barycenters. *ESAIM Math. Model. Numer. Anal.* **49** 1621–1642.
[23] Chartrand, R., Wohlberg, B., Vixie, K.R. and Bollt, E.M. (2009). A gradient descent solution to the Monge–Kantorovich problem. *Appl. Math. Sci.* (*Ruse*) **3** 1071–1080. MR2524965
[24] Chiu, S.N., Stoyan, D., Kendall, W.S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*. New York: Wiley.
[25] Cuesta-Albertos, J.A., Matrán, C. and Tuero-Diaz, A. (1997). Optimal transportation plans and convergence in distribution. *J. Multivariate Anal.* **60** 72–83.
[26] Cuturi, M. and Doucet, A. (2014). Fast computation of Wasserstein barycenters. *Proceedings of the International Conference on Machine Learning* 2014, *JMLR W&CP* **32** 685–693.
[27] Cuturi, M. and Peyré, G. (2016). A smoothed dual approach for variational Wasserstein problems. *SIAM J. Imaging Sci.* **9** 320–343. MR3466197
[28] Dowson, D. and Landau, B. (1982). The Fréchet distance between multivariate normal distributions. *J. Multivariate Anal.* **12** 450–455.
[29] Dryden, I.L. and Mardia, K.V. (1998). *Statistical Shape Analysis*. Chichester: Wiley. MR1646114
[30] Fiedler, M. (1971). Bounds for the determinant of the sum of Hermitian matrices. *Proc. Amer. Math. Soc.* 27–31.
[31] Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré* **10** 215–310. MR0027464
[32] Fréchet, M. (1957). Sur la distance de deux lois de probabilité. *C. R. Math. Acad. Sci. Paris* **244** 689–692.

[33] Freitag, G. and Munk, A. (2005). On Hadamard differentiability in k-sample semiparametric models – With applications to the assessment of structural relationships. *J. Multivariate Anal.* **94** 123–158.

[34] Gallón, S., Loubes, J.-M. and Maza, E. (2013). Statistical properties of the quantile normalization method for density curve alignment. *Math. Biosci.* **242** 129–142. MR3068678

[35] Gangbo, W. and Święch, A. (1998). Optimal maps for the multidimensional Monge–Kantorovich problem. *Comm. Pure Appl. Math.* **51** 23–45.

[36] Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 285–339.

[37] Gower, J.C. (1975). Generalized Procrustes analysis. *Psychometrika* **40** 33–51. MR0405725

[38] Groisser, D. (2005). On the convergence of some Procrustean averaging algorithms. *Stochastics* **77** 31–60.

[39] Haber, E., Rehman, T. and Tannenbaum, A. (2010). An efficient numerical method for the solution of the L_2 optimal mass transfer problem. *SIAM J. Sci. Comput.* **32** 197–211.

[40] Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Chichester: Wiley. MR3379106

[41] Huckemann, S., Hotz, T. and Munk, A. (2010). Intrinsic shape analysis: Geodesic PCA for Riemannian manifolds modulo isometric Lie group actions. *Statist. Sinica* **20** 1–58. MR2640651

[42] Huckemann, S. and Ziezold, H. (2006). Principal component analysis for Riemannian manifolds, with an application to triangular shape spaces. *Adv. in Appl. Probab.* 299–319.

[43] Kallenberg, O. (1986). *Random Measures*, 4th ed. Berlin: Akademie-Verlag. MR0854102

[44] Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.* **30** 509–541. MR0442975

[45] Kendall, W.S. (2010). A survey of Riemannian centres of mass for data. In *Proceedings 59th ISI World Statistics Congress*.

[46] Kendall, W.S. and Le, H. (2011). Limit theorems for empirical Fréchet means of independent and non-identically distributed manifold-valued random variables. *Braz. J. Probab. Stat.* **25** 323–352. MR2832889

[47] Krantz, S. (2014). *Convex Analysis. Textbooks in Mathematics*. Boca Raton: CRC Press.

[48] Le, H. (1998). On the consistency of procrustean mean shapes. *Adv. in Appl. Probab.* 53–63.

[49] Le, H. (2001). Locating Fréchet means with application to shape spaces. *Adv. in Appl. Probab.* 324–338.

[50] Le, H.L. (1995). Mean size-and-shapes and mean shapes: A geometric point of view. *Adv. in Appl. Probab.* **27** 44–55. MR1315576

[51] Le Gouic, T. and Loubes, J.-M. (2016). Existence and consistency of Wasserstein barycenters. *Probab. Theory Related Fields* 1–17.

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

[53] Molchanov, I. and Zuyev, S. (2002). Steepest descent algorithms in a space of measures. *Stat. Comput.* **12** 115–123.

[54] Munk, A. and Czado, C. (1998). Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 223–241.

[55] Munk, A., Paige, R., Pang, J., Patrangenaru, V. and Ruymgaart, F. (2008). The one-and multi-sample problem for functional data with application to projective shape analysis. *J. Multivariate Anal.* **99** 815–833.

[56] Oberman, A.M. and Ruan, Y. (2015). An efficient linear programming method for optimal transportation. Preprint. Available at arXiv:1509.03668.

[57] Olkin, I. and Pukelsheim, F. (1982). The distance between two random vectors with given dispersion matrices. *Linear Algebra Appl.* **48** 257–263.

[58] Panaretos, V.M. and Zemel, Y. (2016). Amplitude and phase variation of point processes. *Ann. Statist.* **44** 771–812.

[59] Pass, B. (2013). Optimal transportation with infinitely many marginals. *J. Funct. Anal.* **264** 947–963. MR3004954

[60] Patrangenaru, V. and Ellingson, L. (2016). *Nonparametric Statistics on Manifolds and Their Applications to Object Data Analysis*. Boca Raton, FL: CRC Press. MR3444169

[61] Pollard, D. (2012). *Convergence of Stochastic Processes*. New York: Springer Science & Business Media.

[62] Rippl, T., Munk, A. and Sturm, A. (2016). Limit laws of the empirical Wasserstein distance: Gaussian distributions. *J. Multivariate Anal.* **151** 90–109.

[63] Rockafellar, R.T. (1970). *Convex Analysis*. *Princeton Mathematical Series*, **28**. Princeton, NJ: Princeton Univ. Press. MR0274683

[64] Rolet, A., Cuturi, M. and Peyré, G. (2016). Fast dictionary learning with a smoothed Wasserstein loss. In *Proceedings of the* 19*th International Conference on Artificial Intelligence and Statistics* (A. Gretton and C.C. Robert, eds.). *Proceedings of Machine Learning Research* **51** 630–638. Cadiz, Spain.

[65] Schachermayer, W. and Teichmann, J. (2009). Characterization of optimal transport plans for the Monge–Kantorovich problem. *Proc. Amer. Math. Soc.* **137** 519–529. MR2448572

[66] Solomon, J., De Goes, F., Peyré, G., Cuturi, M., Butscher, A., Nguyen, A., Du, T. and Guibas, L. (2015). Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Trans. Graph.* **34** 66.

[67] Sommerfeld, M. and Munk, A. (2016). Inference for empirical Wasserstein distances on finite spaces. Preprint. Available at arXiv:1610.03287.

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

[69] Tameling, C., Sommerfeld, M. and Munk, A. (2017). Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications. Preprint. Available at arXiv:1707.00973.

[70] Villani, C. (2003). *Topics in Optimal Transportation* **58**. Providence: AMS.

[71] Wang, W., Slepčev, D., Basu, S., Ozolek, J.A. and Rohde, G.K. (2013). A linear optimal transportation framework for quantifying and visualizing variations in sets of images. *Int. J. Comput. Vis.* **101** 254–269. MR3021062

[72] Zemel, Y. and Panaretos, V.M. (2019). Supplement to "Fréchet means and Procrustes analysis in Wasserstein space." DOI:10.3150/17-BEJ1009SUPP.

[73] Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321. MR3546451

Bernoulli **25**(2), 2019, 977–1012
https://doi.org/10.3150/17-BEJ1010

Are there needles in a moving haystack? Adaptive sensing for detection of dynamically evolving signals

RUI M. CASTRO[1] and ERVIN TÁNCZOS[2]

[1] *Eindhoven University of Technology, P.O. Box 513, Eindhoven, 5600 MB, The Netherlands. E-mail: rmcastro@tue.nl*
[2] *University of Wisconsin – Madison, 330 North Orchard Street, Madison WI 53715, USA. E-mail: tanczos@wisc.edu*

In this paper, we investigate the problem of detecting dynamically evolving signals. We model the signal as an n dimensional vector that is either zero or has s non-zero components. At each time step $t \in \mathbb{N}$ the nonzero components change their location independently with probability p. The statistical problem is to decide whether the signal is a zero vector or in fact it has non-zero components. This decision is based on m noisy observations of individual signal components collected at times $t = 1, \ldots, m$. We consider two different sensing paradigms, namely adaptive and non-adaptive sensing. For non-adaptive sensing, the choice of components to measure has to be decided before the data collection process started, while for adaptive sensing one can adjust the sensing process based on observations collected earlier. We characterize the difficulty of this detection problem in both sensing paradigms in terms of the aforementioned parameters, with special interest to the speed of change of the active components. In addition, we provide an adaptive sensing algorithm for this problem and contrast its performance to that of non-adaptive detection algorithms.

Keywords: adaptive sensing; dynamically evolving signals; sequential experimental design; sparse signals

References

[1] Addario-Berry, L., Broutin, N., Devroye, L. and Lugosi, G. (2010). On combinatorial testing problems. *Ann. Statist.* **38** 3063–3092. MR2722464
[2] Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606. MR1935648
[3] Bayraktar, E. and Lai, L. (2015). Byzantine fault tolerant distributed quickest change detection. *SIAM J. Control Optim.* **53** 575–591.
[4] Caromi, R., Xin, Y. and Lai, L. (2013). Fast multiband spectrum scanning for cognitive radio systems. *IEEE Trans. Commun.* **61** 63–75.
[5] Castro, R.M. (2014). Adaptive sensing performance lower bounds for sparse signal detection and support estimation. *Bernoulli* **20** 2217–2246. MR3263103
[6] Castro, R.M. and Tánczos, E. (2015). Adaptive sensing for estimation of structured sparse signals. *IEEE Trans. Inform. Theory* **61** 2060–2080. MR3332997
[7] Castro, R.M. and Tánczos, E. (2017). Adaptive compressed sensing for support recovery of structured sparse sets. *IEEE Trans. Inform. Theory* **63** 1535–1554. MR3625979
[8] Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195

[9] Dragalin, V. (1996). A simple and effective scanning rule for a multi-channel system. *Metrika* **43** 165–182.
[10] Enikeeva, F., Munk, A. and Werner, F. (2018). Bump detection in heterogeneous Gaussian regression. *Bernoulli* **24** 1266–1306. MR3706794
[11] Flenner, A. and Hewer, G. (2011). A Helmholtz principle approach to parameter free change detection and coherent motion using exchangeable random variables. *SIAM J. Imaging Sci*. **4** 243–276.
[12] Gwadera, R., Atallah, M.J. and Szpankowski, W. (2005). Reliable detection of episodes in event sequences. *Knowl. Inf. Syst*. **7** 415–437.
[13] Hadjiliadis, O., Zhang, H. and Poor, H.V. (2008). One shot schemes for decentralized quickest change detection. In 11*th International Conference on Information Fusion* 1–8.
[14] Haupt, J., Castro, R.M. and Nowak, R. (2011). Distilled sensing: Adaptive sampling for sparse detection and estimation. *IEEE Trans. Inform. Theory* **57** 6222–6235.
[15] Huang, L., Kulldorff, M. and Gregorio, D. (2007). A spatial scan statistic for survival data. *Biometrics* **63** 109–118, 311–312. MR2345580
[16] Ingster, Y.I. (1997). Some problems of hypothesis testing leading to infinitely divisible distributions. *Math. Methods Statist*. **6** 47–69.
[17] Ingster, Y.I. and Suslina, I.A. (2000). Minimax nonparametric hypothesis testing for ellipsoids and Besov bodies. *ESAIM Probab. Stat*. **4** 53–135.
[18] Ingster, Y.I. and Suslina, I.A. (2002). On the detection of a signal with a known shape in a multichannel system. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **294** 88–112, 261. MR1976749
[19] Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables, with applications. *Ann. Statist*. **11** 286–295. MR0684886
[20] Klimko, E.M. and Yackel, J. (1975). Optimal search strategies for Wiener processes. *Stochastic Process. Appl*. **3** 19–33.
[21] Kulldorff, M., Heffernan, R., Hartman, J., Assunçao, R. and Mostashari, F. (2005). A space–time permutation scan statistic for disease outbreak detection. *PLoS Med*. **2** 216–224.
[22] Kulldorff, M., Huang, L. and Konty, K. (2009). A scan statistic for continuous data based on the normal probability model. *Int. J. Health Geogr.* **8** 58.
[23] Li, H. (2009). Restless watchdog: Selective quickest spectrum sensing in multichannel cognitive radio systems. *EURASIP J. Adv. Signal Process*. **2009** Article ID: 417457.
[24] Luo, W. and Tay, W.P. (2013). Finding an infection source under the SIS model. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*) 2930–2934.
[25] Malloy, M. and Nowak, R. (2011). On the limits of sequential testing in high dimensions. In *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers* (*ASILOMAR*), 2011 1245–1249.
[26] Malloy, M.L. and Nowak, R.D. (2014). Sequential testing for sparse recovery. *IEEE Trans. Inform. Theory* **60** 7862–7873.
[27] Neill, D.B. and Moore, A.W. (2004). A fast multi-resolution method for detection of significant spatial disease clusters. In *Advances in Neural Information Processing Systems* 16 651–658. MIT Press.
[28] Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A. and Ploner, A. (2005). False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21** 3017–3024.
[29] Phoha, V.V. (2007). *Internet Security Dictionary*. Springer Science & Business Media.
[30] Raghavan, V. and Veeravalli, V.V. (2010). Quickest change detection of a Markov process across a sensor array. *IEEE Trans. Inform. Theory* **56** 1961–1981.
[31] Shah, D. and Zaman, T. (2011). Rumors in a network: Who's the culprit? *IEEE Trans. Inform. Theory* **57** 5163–5181. MR2849111

[32] Thompson, D.R., Burke-Spolaor, S., Deller, A.T. et al. (2014). Real-time adaptive event detection in astronomical data streams. *IEEE Intell*. *Syst*. **29** 48–55.
[33] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Mathematics & Applications* **41**. Berlin: Springer. MR2013911
[34] Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann*. *Math*. *Stat*. **16** 117–186.
[35] Wang, H., Tang, M., Park, Y. and Priebe, C.E. (2014). Locality statistics for anomaly detection in time series of graphs. *IEEE Trans*. *Signal Process*. **62** 703–717. MR3160307
[36] Zhao, Q. and Ye, J. (2010). Quickest detection in multiple on–off processes. *IEEE Trans*. *Signal Process*. **58** 5994–6006.
[37] Zhu, K. and Ying, L. (2013). Information source detection in the SIR model: A sample path based approach. In *Information Theory and Applications Workshop* (*ITA*) 1–9.
[38] Zigangirov, K.Š. (1966). On a problem of optimal scanning. *Theory Probab*. *Appl*. **11** 294–298. MR0200090

Bernoulli **25**(2), 2019, 1013–1044
https://doi.org/10.3150/17-BEJ1011

Towards a general theory for nonlinear locally stationary processes

RAINER DAHLHAUS[1,*] STEFAN RICHTER[1,**] and WEI BIAO WU[2]

[1]*Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany. E-mail:* [*]*dahlhaus@statlab.uni-heidelberg.de;* [**]*stefan.richter@iwr.uni-heidelberg.de*
[2]*Department of Statistics, University of Chicago, 5734 S. University Avenue, Chicago, IL 60637, USA. E-mail: wbwu@galton.uchicago.edu*

In this paper, some general theory is presented for locally stationary processes based on the stationary approximation and the stationary derivative. Laws of large numbers, central limit theorems as well as deterministic and stochastic bias expansions are proved for processes obeying an expansion in terms of the stationary approximation and derivative. In addition it is shown that this applies to some general nonlinear non-stationary Markov-models. In addition the results are applied to derive the asymptotic properties of maximum likelihood estimates of parameter curves in such models.

Keywords: derivative processes; non-stationary processes

References

[1] Chen, X., Xu, M. and Wu, W.B. (2013). Covariance and precision matrix estimation for high-dimensional time series. *Ann. Statist.* **41** 2994–3021. MR3161455
[2] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916
[3] Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762–1794. MR1835040
[4] Dahlhaus, R. (2012). Locally stationary processes. In *Time Series Analysis*: *Methods and Applications* (T.S. Rao, S.S. Rao and C.R. Rao, eds.). *Handbook of Statistics* **30** 351–413. Amsterdam: Elsevier.
[5] Dahlhaus, R. and Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *J. Time Series Anal.* **19** 629–655. MR1665941
[6] Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15** 1–39. MR2546797
[7] Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Supplement to "Towards a general theory for nonlinear locally stationary processes." DOI:10.3150/17-BEJ1011SUPP.
[8] Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying ARCH processes. *Ann. Statist.* **34** 1075–1114. MR2278352
[9] Eichler, M., Motta, G. and von Sachs, R. (2011). Fitting dynamic factor models to non-stationary time series. *J. Econometrics* **163** 51–70. MR2803666
[10] Jones, D.A. (1978). Nonlinear autoregressive processes. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **360** 71–95. MR0501672
[11] Koo, B. and Linton, O. (2012). Estimation of semiparametric locally stationary diffusion models. *J. Econometrics* **170** 210–233. MR2955950

1350-7265

[12] Kreiss, J.-P. and Paparoditis, E. (2015). Bootstrapping locally stationary processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 267–290. MR3299408
[13] Liu, W., Xiao, H. and Wu, W.B. (2013). Probability and moment inequalities under dependence. *Statist. Sinica* **23** 1257–1272. MR3114713
[14] Martin, W. and Flandrin, P. (1985). Wigner-Ville spectral analysis of nonstationary processes. *IEEE Trans. Acoust. Speech Signal Process.* **33** 1461–1470.
[15] Motta, G., Hafner, C.M. and von Sachs, R. (2011). Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory* **27** 1279–1319. MR2868840
[16] Palma, W. and Olea, R. (2010). An efficient estimator for locally stationary Gaussian long-memory processes. *Ann. Statist.* **38** 2958–2997. MR2722461
[17] Preuß, P., Vetter, M. and Dette, H. (2013). A test for stationarity based on empirical processes. *Bernoulli* **19** 2715–2749. MR3160569
[18] Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions. *J. Theoret. Probab.* **22** 146–163. MR2472010
[19] Roueff, F. and von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Process. Appl.* **121** 813–844. MR2770908
[20] Sedro, J. (2017). A regularity result for fixed points, with applications to linear response. Preprint. Available at arXiv:1705.04078.
[21] Sergides, M. and Paparoditis, E. (2008). Bootstrapping the local periodogram of locally stationary processes. *J. Time Series Anal.* **29** 264–299. MR2392774
[22] Sergides, M. and Paparoditis, E. (2009). Frequency domain tests of semi-parametric hypotheses for locally stationary processes. *Scand. J. Stat.* **36** 800–821. MR2573309
[23] Shao, X. and Wu, W.B. (2007). Asymptotic spectral theory for nonlinear time series. *Ann. Statist.* **35** 1773–1801. MR2351105
[24] Subba Rao, S. (2006). On some nonstationary, nonlinear random processes and their stationary approximations. *Adv. in Appl. Probab.* **38** 1155–1172. MR2285698
[25] Truquet, L. (2016). Local stationarity and time-inhomogeneous Markov chains. Preprint. Available at arXiv:1610.01290.
[26] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247
[27] Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614
[28] Wu, W. and Zhou, Z. (2017). Nonparametric inference for time-varying coefficient quantile regression. *J. Bus. Econom. Statist.* **35** 98–109. MR3591532
[29] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. MR2172215
[30] Wu, W.B. (2007). M-Estimation of linear models with dependent errors. *Ann. Statist.* **35** 495–521. MR2336857
[31] Wu, W.B. (2008). Empirical processes of stationary sequences. *Statist. Sinica* **18** 313–333. MR2384990
[32] Wu, W.B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4** 207–226. MR2812816
[33] Wu, W.B. and Zhou, Z. (2011). Gaussian approximations for non-stationary multiple time series. *Statist. Sinica* **21** 1397–1413. MR2827528
[34] Zhou, Z. (2014). Inference of weighted V-statistics for nonstationary time series and its applications. *Ann. Statist.* **42** 87–114. MR3161462
[35] Zhou, Z. (2014). Nonparametric specification for non-stationary time series regression. *Bernoulli* **20** 78–108. MR3160574
[36] Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. MR2541444

Bernoulli **25**(2), 2019, 1045–1075
https://doi.org/10.3150/17-BEJ1012

Properties of switching jump diffusions: Maximum principles and Harnack inequalities

XIAOSHAN CHEN[1], ZHEN-QING CHEN[2], KY TRAN[3] and GEORGE YIN[4]

[1]*School of Mathematical Sciences, South China Normal University, Guangdong, China.*
E-mail: xschen@m.scnu.edu.cn
[2]*Departments of Mathematics, University of Washington, Seattle, WA 98195, USA.*
E-mail: zqchen@uw.edu
[3]*Department of Mathematics, College of Education, Hue University, Hue city, Vietnam.*
E-mail: quankysp@gmail.com
[4]*Department of Mathematics, Wayne State University, Detroit, MI 48202, USA.*
E-mail: gyin@math.wayne.edu

This work examines a class of switching jump diffusion processes. The main effort is devoted to proving the maximum principle and obtaining the Harnack inequalities. Compared with the diffusions and switching diffusions, the associated operators for switching jump diffusions are non-local, resulting in more difficulty in treating such systems. Our study is carried out by taking into consideration of the interplay of stochastic processes and the associated systems of integro-differential equations.

Keywords: Harnack inequality; jump diffusion; maximum principle; regime switching

References

[1] Arapostathis, A., Ghosh, M.K. and Marcus, S.I. (1999). Harnack's inequality for cooperative weakly coupled elliptic systems. *Comm. Partial Differential Equations* **24** 1555–1571. MR1708101
[2] Athreya, S. and Ramachandran, K. (2017). Harnack inequality for non-local Schrödinger operators. *Potential Anal*. To appear. DOI:10.1007/s11118-017-9646-6.
[3] Bass, R.F. and Kassmann, M. (2005). Harnack inequalities for non-local operators of variable order. *Trans. Amer. Math. Soc.* **357** 837–850. MR2095633
[4] Bass, R.F., Kassmann, M. and Kumagai, T. (2010). Symmetric jump processes: Localization, heat kernels and convergence. *Ann. Inst. Henri Poincaré B, Probab. Stat.* **46** 59–71. MR2641770
[5] Bass, R.F. and Levin, D.A. (2002). Harnack inequalities for jump processes. *Potential Anal.* **17** 375–388. MR1918242
[6] Caffarelli, L. and Silvestre, L. (2009). Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math.* **62** 597–638. MR2494809
[7] Chen, X., Chen, Z.-Q., Tran, K. and Yin, G. (2017). Recurrence and ergodicity for a class of regime-switching jump diffusions. *Appl. Math. Optim.* To appear. DOI:10.1007/s00245-017-9470-9.
[8] Chen, Z.-Q., Hu, E., Xie, L. and Zhang, X. (2017). Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations* **263** 6576–6634. MR3693184
[9] Chen, Z.-Q. and Kumagai, T. (2003). Heat kernel estimates for stable-like processes on d-sets. *Stochastic Process. Appl.* **108** 27–62.

[10] Chen, Z.-Q. and Kumagai, T. (2008). Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Related Fields* **140** 277–317. MR2357678
[11] Chen, Z.-Q. and Kumagai, T. (2010). A priori Hölder estimate, parabolic Harnack principle and heat kernel estimates for diffusions with jumps. *Rev. Mat. Iberoam.* **26** 551–589. MR2677007
[12] Chen, Z.-Q., Wang, H. and Xiong, J. (2012). Interacting superprocesses with discontinuous spatial motion. *Forum Math.* **24** 1183–1223.
[13] Chen, Z.-Q. and Zhao, Z. (1996). Potential theory for elliptic systems. *Ann. Probab.* **24** 293–319.
[14] Chen, Z.-Q. and Zhao, Z. (1997). Harnack principle for weakly coupled elliptic systems. *J. Differential Equations* **139** 261–282. MR1472349
[15] Evans, L.C. (2010). *Partial Differential Equations*, 2nd ed. Providence, RI: Amer. Math. Soc.
[16] Foondun, M. (2009). Harmonic functions for a class of integro-differential operators. *Potential Anal.* **31** 21–44. MR2507444
[17] Ikeda, N., Nagasawa, M. and Watanabe, S. (1966). A construction of Markov process by piecing out. *Proc. Jpn. Acad.* **42** 370–375.
[18] Jasso-Fuentes, H. and Yin, G. (2013). *Advanced Criteria for Controlled Markov-Modulated Diffusions in an Infinite Horizon: Overtaking, Bias, and Blackwell Optimality*. Beijing: Science Press.
[19] Komatsu, T. (1973). Markov processes associated with certain integro-differential. *Osaka J. Math.* **10** 271–303.
[20] Krylov, N.V. (1987). *Nonlinear Elliptic and Parabolic Equations of the Second Order*. *Mathematics and Its Applications (Soviet Series)* **7**. Dordrecht: D. Reidel Publishing Co. Translated from the Russian by P. L. Buzytsky [P. L. Buzytskiĭ]. MR0901759
[21] Kushner, H.J. (1990). *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*. *Systems & Control: Foundations & Applications* **3**. Boston, MA: Birkhäuser, Inc. MR1102242
[22] Liu, R. (2016). Optimal stopping of switching diffusions with state dependent switching rates. *Stochastics* **88** 586–605.
[23] Mao, X. and Yuan, C. (2006). *Stochastic Differential Equations with Markovian Switching*. London: Imperial College Press. MR2256095
[24] Meyer, P. (1975). Renaissance, recollements, mélanges, relentissement de processus de Markov. *Ann. Inst. Fourier (Grenoble)* **25** 465–497.
[25] Mikulevičius, R. and Pragarauskas, H. (1988). On Hölder continuity of solutions of certain integro-differential equations. *Ann. Acad. Sci. Fenn., Ser. A* 1 *Math.* **13** 231–238. MR0986325
[26] Negoro, A. and Tsuchiya, M. (1989). Stochastic processes and semigroups associate with degenerate Lévy generating operators. *Stoch. Stoch. Rep.* **26** 29–61.
[27] Protter, M.H. and Weinberger, H.F. (1967). *Maximum Principles in Differential Equations*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
[28] Sharpe, M. (1986). *General Theory of Markov Processes*. New York: Academic.
[29] Song, R. and Vondraček, Z. (2005). Harnack inequality for some discontinuous Markov processes with a diffusion part. *Glas. Mat. Ser. III* **40** 177–187.
[30] Wang, J.-M. (2014). Martingale problems for switched processes. *Math. Nachr.* **287** 1186–1201. MR3231533
[31] Xi, F. (2009). Asymptotic properties of jump-diffusion processes with state-dependent switching. *Stoch. Process. Appl.* **119** 2198–2221. MR2531089
[32] Yin, G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*. New York: Springer.

Bernoulli **25**(2), 2019, 1076–1104
https://doi.org/10.3150/17-BEJ1013

Error bounds in local limit theorems using Stein's method

A.D. BARBOUR[1], ADRIAN RÖLLIN[2] and NATHAN ROSS[3]

[1]*Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 Zürich, Switzerland. E-mail: a.d.barbour@math.uzh.ch; url: http://user.math.uzh.ch/barbour/*
[2]*Department of Statistics and Applied Probability, 6 Science Drive 2, Singapore, 117546, Singapore. E-mail: adrian.roellin@nus.edu.sg; url: https://blog.nus.edu.sg/roellin/*
[3]*School of Mathematics and Statistics, Peter Hall Building, University of Melbourne, VIC, 3010, Australia. E-mail: nathan.ross@unimelb.edu.au; url: http://www.nathanrossprob.com*

We provide a general result for bounding the difference between point probabilities of integer supported distributions and the translated Poisson distribution, a convenient alternative to the discretized normal. We illustrate our theorem in the context of the Hoeffding combinatorial central limit theorem with integer valued summands, of the number of isolated vertices in an Erdős–Rényi random graph, and of the Curie–Weiss model of magnetism, where we provide optimal or near optimal rates of convergence in the local limit metric. In the Hoeffding example, even the discrete normal approximation bounds seem to be new. The general result follows from Stein's method, and requires a new bound on the Stein solution for the Poisson distribution, which is of general interest.

Keywords: approximation error; Curie–Weiss model; Erdős–Rényi random graph; Hoeffding combinatorial statistic; local limit theorem

References

[1] Arratia, R. and Baxendale, P. (2015). Bounded size bias coupling: A Gamma function bound, and universal Dickman-function behavior. *Probab. Theory Related Fields* **162** 411–429. MR3383333
[2] Barbour, A.D. (1980). Equilibrium distributions for Markov population processes. *Adv. in Appl. Probab.* **12** 591–614. MR0578839
[3] Barbour, A.D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. New York: The Clarendon Press, Oxford Univ. Press. MR1163825
[4] 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.
[5] Barbour, A.D. and Xia, A. (1999). Poisson perturbations. *ESAIM Probab. Stat.* **3** 131–150. MR1716120
[6] Bartroff, J., Goldstein, L. and Işlak, Ü. (2015). Bounded size biased couplings, log concave distributions and concentration of measure for occupancy models. *Bernoulli*. To appear. Available at arXiv:1402.6769v2.
[7] Bolthausen, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** 379–386. MR0751577
[8] Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072

1350-7265

[9] Chatterjee, S. and Dey, P.S. (2010). Applications of Stein's method for concentration inequalities. *Ann. Probab.* **38** 2443–2485. MR2683635

[10] Chatterjee, S. and Shao, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964

[11] Chen, L.H.Y. and Fang, X. (2015). On the error bound in a combinatorial central limit theorem. *Bernoulli* **21** 335–359. MR3322321

[12] Chen, L.H.Y., Fang, X. and Shao, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293. MR3059199

[13] Chen, L.H.Y. and Röllin, A. (2010). Stein couplings for normal approximation. Preprint. Available at arXiv:1003.6039v2.

[14] Dembo, A. and Montanari, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563

[15] Eichelsbacher, P. and Löwe, M. (2010). Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15** 962–988. MR2659754

[16] Ellis, R.S. (2006). *Entropy, Large Deviations, and Statistical Mechanics. Classics in Mathematics.* Berlin: Springer. Reprint of the 1985 original. MR2189669

[17] Ellis, R.S. and Newman, C.M. (1978). Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **44** 117–139. MR0503333

[18] Ellis, R.S., Newman, C.M. and Rosen, J.S. (1980). Limit theorems for sums of dependent random variables occurring in statistical mechanics. II. Conditioning, multiple phases, and metastability. *Z. Wahrsch. Verw. Gebiete* **51** 153–169. MR0566313

[19] Esseen, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125. MR0014626

[20] Fang, X. (2014). Discretized normal approximation by Stein's method. *Bernoulli* **20** 1404–1431. MR3217448

[21] Goldstein, L. (2005). Berry–Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **42** 661–683. MR2157512

[22] Goldstein, L. (2013). A Berry–Esseen bound with applications to vertex degree counts in the Erdős–Rényi random graph. *Ann. Appl. Probab.* **23** 617–636. MR3059270

[23] Goldstein, L. and Işlak, Ü. (2014). Concentration inequalities via zero bias couplings. *Statist. Probab. Lett.* **86** 17–23. MR3162712

[24] Goldstein, L. and Xia, A. (2006). Zero biasing and a discrete central limit theorem. *Ann. Probab.* **34** 1782–1806. MR2271482

[25] Hoeffding, W. (1951). A combinatorial central limit theorem. *Ann. Math. Stat.* **22** 558–566. MR0044058

[26] Kordecki, W. (1990). Normal approximation and isolated vertices in random graphs. In *Random Graphs '87* (*Poznań*, 1987) 131–139. Chichester: Wiley. MR1094128

[27] 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

[28] McDonald, D.R. (1979). On local limit theorem for integer valued random variables. *Teor. Veroyatn. Primen.* **24** 607–614. MR0541375

[29] Petrov, V.V. (1975). *Sums of Independent Random Variables. Ergebnisse der Mathematik und ihrer Grenzgebiete* **82**. New York: Springer. Translated from the Russian by A.A. Brown. MR0388499

[30] Röllin, A. (2005). Approximation of sums of conditionally independent variables by the translated Poisson distribution. *Bernoulli* **11** 1115–1128. MR2189083

[31] Röllin, A. (2007). Translated Poisson approximation using exchangeable pair couplings. *Ann. Appl. Probab.* **17** 1596–1614. MR2358635

[32] Röllin, A. (2008). Symmetric and centered binomial approximation of sums of locally dependent random variables. *Electron*. *J*. *Probab*. **13** 756–776. MR2399295

[33] Röllin, A. (2017). Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős–Rényi random graph. Preprint. Available at arXiv:1704.00410v1.

[34] Röllin, A. and Ross, N. (2015). Local limit theorems via Landau–Kolmogorov inequalities. *Bernoulli* **21** 851–880. MR3338649

[35] Ross, N. (2011). Fundamentals of Stein's method. *Probab*. *Surv*. **8** 210–293. MR2861132

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

[37] Wald, A. and Wolfowitz, J. (1944). Statistical tests based on permutations of the observations. *Ann*. *Math*. *Stat*. **15** 358–372. MR0011424

Bernoulli **25**(2), 2019, 1105–1140
https://doi.org/10.3150/17-BEJ1014

Stability for gains from large investors' strategies in M_1/J_1 topologies

DIRK BECHERER*, TODOR BILAREV** and PETER FRENTRUP†

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany.
E-mail: *becherer@math.hu-berlin.de; **bilarev@math.hu-berlin.de; †frentrup@math.hu-berlin.de

We prove continuity of a controlled SDE solution in Skorokhod's M_1 and J_1 topologies and also uniformly, in probability, as a nonlinear functional of the control strategy. The functional comes from a finance problem to model price impact of a large investor in an illiquid market. We show that M_1-continuity is the key to ensure that proceeds and wealth processes from (self-financing) càdlàg trading strategies are determined as the continuous extensions for those from continuous strategies. We demonstrate by examples how continuity properties are useful to solve different stochastic control problems on optimal liquidation and to identify asymptotically realizable proceeds.

Keywords: continuity of proceeds; illiquid markets; no-arbitrage; optimal liquidation; Skorokhod space; Skorokhod topologies; stability; stochastic differential equation; transient price impact

References

[1] Alfonsi, A., Fruth, A. and Schied, A. (2010). Optimal execution strategies in limit order books with general shape functions. *Quant. Finance* **10** 143–157. MR2642960

[2] Bank, P. and Baum, D. (2004). Hedging and portfolio optimization in financial markets with a large trader. *Math. Finance* **14** 1–18. MR2030833

[3] Becherer, D., Bilarev, T. and Frentrup, P. (2015). Multiplicative limit order markets with transient impact and zero spread. Available at arXiv:1501.01892v1.

[4] Becherer, D., Bilarev, T. and Frentrup, P. (2018). Optimal asset liquidation with multiplicative transient price impact. *Appl. Math. Optim.* To appear. DOI:10.1007/s00245-017-9418-0.

[5] Becherer, D., Bilarev, T. and Frentrup, P. (2018). Optimal liquidation under stochastic liquidity. *Finance Stoch.* **22** 39–68.

[6] Billingsley, P. (1999). *Convergence of Probability Measures. Wiley Series in Probability and Statistics*. New York: Wiley. MR1700749

[7] Blümmel, T. and Rheinländer, T. (2017). Financial markets with a large trader. *Ann. Appl. Probab.* **27** 3735–3786. MR3737937

[8] Borodin, A.N. and Salminen, P. (2002). *Handbook of Brownian Motion—Facts and Formulae. Probability and Its Applications*. Basel: Birkhäuser. MR1912205

[9] Bouchard, B., Loeper, G. and Zou, Y. (2016). Almost-sure hedging with permanent price impact. *Finance Stoch.* **20** 741–771. MR3519167

[10] Carr, P., Geman, H., Madan, D.B. and Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *J. Bus.* **75** 305–332.

[11] Çetin, U., Jarrow, R.A. and Protter, P.E. (2004). Liquidity risk and arbitrage pricing theory. *Finance Stoch.* **8** 311–341. MR2213255

1350-7265

[12] Çetin, U., Soner, H.M. and Touzi, N. (2010). Option hedging for small investors under liquidity costs. *Finance Stoch*. **14** 317–341. MR2670416
[13] Chan, L.K.C. and Lakonishok, J. (1995). The behavior of stock prices around institutional trades. *J. Finance* **50** 1147–1174.
[14] Cont, R. and De Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM J. Financial Math*. **4** 1–25. MR3032934
[15] Dellacherie, C. and Meyer, P.-A. (1982). *Probabilities and Potential*. *B*. Amsterdam: North-Holland. MR745449
[16] Esche, F. and Schweizer, M. (2005). Minimal entropy preserves the Lévy property: How and why. *Stochastic Process*. *Appl*. **115** 299–327. MR2111196
[17] Friz, P. and Chevyrev, I. (2018). Canonical RDEs and general semimartingales as rough paths. *Ann. Probab.* To appear.
[18] Guo, X. and Zervos, M. (2015). Optimal execution with multiplicative price impact. *SIAM J. Financial Math*. **6** 281–306. MR3335493
[19] Henderson, V. and Hobson, D. (2011). Optimal liquidation of derivative portfolios. *Math*. *Finance* **21** 365–382. MR2830426
[20] Hindy, A., Huang, C.-F. and Kreps, D. (1992). On intertemporal preferences in continuous time: The case of certainty. *J. Math*. *Econom*. **21** 401–440. MR1183611
[21] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*. Berlin: Springer. MR1943877
[22] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications*. New York: Springer. MR1876169
[23] Kallsen, J. and Shiryaev, A.N. (2002). The cumulant process and Esscher's change of measure. *Finance Stoch*. **6** 397–428. MR1932378
[24] Kardaras, C. (2013). On the closure in the Emery topology of semimartingale wealth-process sets. *Ann*. *Appl*. *Probab*. **23** 1355–1376. MR3098435
[25] Klein, O., Maug, E. and Schneider, C. (2017). Trading strategies of corporate insiders. *J. Financ*. *Mark*. **34** 48–68.
[26] Kurtz, T.G., Pardoux, É. and Protter, P. (1995). Stratonovich stochastic differential equations driven by general semimartingales. *Ann*. *Inst*. *Henri Poincaré Probab. Stat*. **31** 351–377. MR1324812
[27] Kurtz, T.G. and Protter, P.E. (1996). Weak convergence of stochastic integrals and differential equations. In *Probabilistic Models for Nonlinear Partial Differential Equations* (D. Talay and L. Tubaro, eds.). *Lecture Notes in Math*. **1627** 1–41. Berlin: Springer. MR1431298
[28] Løkka, A. (2012). Optimal execution in a multiplicative limit order book. Preprint, London School of Economics.
[29] Løkka, A. (2014). Optimal liquidation in a limit order book for a risk-averse investor. *Math*. *Finance* **24** 696–727. MR3274929
[30] Lorenz, C. and Schied, A. (2013). Drift dependence of optimal trade execution strategies under transient price impact. *Finance Stoch*. **17** 743–770. MR3105932
[31] Marcus, S.I. (1981). Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics* **4** 223–245. MR605630
[32] Obizhaeva, A. and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *J. Financ*. *Mark*. **16** 1–32.
[33] Pang, G., Talreja, R. and Whitt, W. (2007). Martingale proofs of many-server heavy-traffic limits for Markovian queues. *Probab. Surv*. **4** 193–267. MR2368951
[34] Pang, G. and Whitt, W. (2010). Continuity of a queueing integral representation in the M_1 topology. *Ann*. *Appl*. *Probab*. **20** 214–237. MR2582647

[35] Predoiu, S., Shaikhet, G. and Shreve, S. (2011). Optimal execution in a general one-sided limit-order book. *SIAM J*. *Financial Math*. **2** 183–212. MR2775411
[36] Protter, P.E. (2005). *Stochastic Integration and Differential Equations*. *Stochastic Modelling and Applied Probability* **21**. Berlin: Springer. Second edition. Version 2.1, corrected third printing. MR2273672
[37] Roch, A.F. and Soner, H.M. (2013). Resilient price impact of trading and the cost of illiquidity. *Int*. *J*. *Theor. Appl*. *Finance* **16** 1350037 (27 pages). MR3117871
[38] Roch, A.F. (2011). Liquidity risk, price impacts and the replication problem. *Finance Stoch*. **15** 399–419. MR2833094
[39] Skorokhod, A.V. (1956). Limit theorems for stochastic processes. *Theory Probab*. *Appl*. **1** 261–290.
[40] Whitt, W. (2002). *Stochastic-Process Limits*: *An Introduction to Stochastic-Process Limits and Their Application to Queues*. *Springer Series in Operations Research*. New York: Springer. MR1876437
[41] Wong, E. and Zakai, M. (1965). On the convergence of ordinary integrals to stochastic integrals. *Ann*. *Math*. *Stat*. **36** 1560–1564. MR0195142

Bernoulli **25**(2), 2019, 1141–1159
https://doi.org/10.3150/17-BEJ1016

Convergence rates for a class of estimators based on Stein's method

CHRIS J. OATES[1,4], JON COCKAYNE[2], FRANÇOIS-XAVIER BRIOL[2,3] and MARK GIROLAMI[3,4]

[1]*School of Mathematics and Statistics, Newcastle University, UK. E-mail: chris.oates@ncl.ac.uk*
[2]*Department of Statistics, University of Warwick, UK*
[3]*Department of Mathematics, Imperial College London, UK*
[4]*Alan Turing Institute, UK*

Gradient information on the sampling distribution can be used to reduce the variance of Monte Carlo estimators via Stein's method. An important application is that of estimating an expectation of a test function along the sample path of a Markov chain, where gradient information enables convergence rate improvement at the cost of a linear system which must be solved. The contribution of this paper is to establish theoretical bounds on convergence rates for a class of estimators based on Stein's method. Our analysis accounts for (i) the degree of smoothness of the sampling distribution and test function, (ii) the dimension of the state space, and (iii) the case of non-independent samples arising from a Markov chain. These results provide insight into the rapid convergence of gradient-based estimators observed for low-dimensional problems, as well as clarifying a curse-of-dimension that appears inherent to such methods.

Keywords: asymptotics; control functionals; reproducing kernel; scattered data; variance reduction

References

[1] Andradóttir, S., Heyman, D.P. and Ott, T.J. (1993). Variance reduction through smoothing and control variates for Markov chain simulations. *ACM Trans. Model. Comput. Simul.* **3** 167–189.
[2] Assaraf, R. and Caffarel, M. (1999). Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.* **83** 4682–4685.
[3] Assaraf, R. and Caffarel, M. (2003). Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces. *J. Chem. Phys.* **119** 10536.
[4] Azaïs, R., Delyon, B. and Portier, F. (2016). Integral estimation based on Markovian design. Available at arXiv:1609.01165.
[5] Bahvalov, N.S. (1959). Approximate computation of multiple integrals. *Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him.* **1959** 3–18. MR0115275
[6] Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Boston: Kluwer Academic.
[7] Briol, F.-X., Oates, C.J., Cockayne, J., Chen, W.Y. and Girolami, M. (2017). On the sampling problem for kernel quadrature. In *Proceedings of the* 34*th International Conference on Machine Learning*.
[8] Briol, F.-X., Oates, C.J., Girolami, M., Osborne, M.A. and Sejdinovic, D. (2017). Probabilistic integration: A role in statistical computation? Available at arXiv:1512.00933.
[9] Carpenter, B., Hoffman, M.D., Brubaker, M., Lee, D., Li, P. and Betancourt, M. (2015). The Stan math library: Reverse-mode automatic differentiation in C++. Available at arXiv:1509.07164.

1350-7265

[10] Chwialkowski, K., Strathmann, H. and Gretton, A. (2016). A kernel test of goodness of fit. In *Proceedings of the* 33*rd International Conference on Machine Learning*.
[11] Cockayne, J., Oates, C.J., Sullivan, T. and Girolami, M. (2016). Probabilistic numerical methods for PDE-constrained Bayesian inverse problems. In *Proceedings of the* 36*th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*.
[12] Cockayne, J., Oates, C.J., Sullivan, T. and Girolami, M. (2017). Probabilistic meshless methods for Bayesian inverse problems. Available at arXiv:1605.07811.
[13] Cockayne, J., Oates, C.J., Sullivan, T. and Girolami, M. (2017). Bayesian probabilistic numerical methods. Available at arXiv:1702.03673.
[14] Dellaportas, P. and Kontoyiannis, I. (2012). Control variates for estimation based on reversible Markov chain Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 133–161.
[15] Delyon, B. and Portier, F. (2016). Integral approximation by kernel smoothing. *Bernoulli* **22** 2177–2208. MR3498027
[16] Dick, J., Gantner, R.N., Gia, Q.T.L. and Schwab, C. (2016). Higher order quasi-Monte Carlo integration for Bayesian estimation. Available at arXiv:1602.07363.
[17] Gorham, J., Duncan, A.B., Vollmer, S.J. and Mackey, L. (2016). Measuring sample quality with diffusions. Available at arXiv:1611.06972.
[18] Gorham, J. and Mackey, L. (2015). Measuring sample quality with Stein's method. In *Proceedings of the* 28*th Annual Conference on Neural Information Processing Systems*.
[19] Gorham, J. and Mackey, L. (2017). Measuring sample quality with kernels. In *Proceedings of the* 34*th International Conference on Machine Learning*.
[20] Hammer, H. and Tjelmeland, H. (2008). Control variates for the Metropolis–Hastings algorithm. *Scand. J. Stat.* **35** 400–414.
[21] Kanagawa, M., Sriperumbudur, B.K. and Fukumizu, K. (2016). Convergence guarantees for kernel-based quadrature rules in misspecified settings. In *Proceedings of the* 29*th Annual Conference on Neural Information Processing Systems*.
[22] Li, W., Chen, R. and Tan, Z. (2016). Efficient sequential Monte Carlo with multiple proposals and control variates. *J. Amer. Statist. Assoc.* **111** 298–313. MR3494661
[23] Liu, Q. and Lee, J.D. (2017). Black-box importance sampling. In *Proceedings of the* 21*st International Conference on Artificial Intelligence and Statistics*.
[24] Liu, Q., Lee, J.D. and Jordan, M.I. (2016). A kernelized Stein discrepancy for goodness-of-fit tests and model evaluation. In *Proceedings of the* 33*rd International Conference on Machine Learning*.
[25] Maclaurin, D., Duvenaud, D., Johnson, M. and Adams, R.P. (2015). Autograd: Reverse-mode differentiation of native Python. Available at http://github.com/HIPS/autograd+.
[26] Meyn, S.P. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press.
[27] Micchelli, C.A., Xu, Y. and Zhang, H. (2006). Universal kernels. *J. Mach. Learn. Res.* **7** 2651–2667.
[28] Migliorati, G., Nobile, F. and Tempone, R. (2015). Convergence estimates in probability and in expectation for discrete least squares with noisy evaluations at random points. *J. Multivariate Anal.* **142** 167–182.
[29] Mijatović, A. and Vogrinc, J. (2015). On the Poisson equation for Metropolis–Hastings chains. Available at arXiv:1511.07464.
[30] Mijatović, A. and Vogrinc, J. (2017). Asymptotic variance for random walk Metropolis chains in high dimensions: Logarithmic growth via the Poisson equation. Available at arXiv:1707.08510.
[31] Mira, A., Solgi, R. and Imparato, D. (2013). Zero variance Markov chain Monte Carlo for Bayesian estimators. *Stat. Comput.* **23** 653–662. MR3094805
[32] Narcowich, F.J., Ward, J.D. and Wendland, H. (2005). Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting. *Math. Comp.* **74** 743–763. MR2114646

[33] Niederreiter, H. (2010). *Quasi-Monte Carlo Methods*. New York: Wiley.
[34] Oates, C.J., Cockayne, J., Briol, F.-X. and Girolami, M. (2019). Supplement to “Convergence rates for a class of estimators based on Stein’s method.” DOI:10.3150/17-BEJ1016SUPP.
[35] Oates, C.J., Girolami, M. and Chopin, N. (2017). Control functionals for Monte Carlo integration. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 695–718. MR3641403
[36] Oates, C.J., Papamarkou, T. and Girolami, M. (2016). The controlled thermodynamic integral for Bayesian model evidence evaluation. *J. Amer. Statist. Assoc.* **111** 634–645.
[37] Robert, C. and Casella, G. (2013). *Monte Carlo Statistical Methods*. New York: Springer.
[38] Roberts, G.O. and Rosenthal, J.S. (1998). On convergence rates of Gibbs samplers for uniform distributions. *Ann. Appl. Probab.* **8** 1291–1302.
[39] Rubinstein, R.Y. and Marcus, R. (1985). Efficiency of multivariate control variates in Monte Carlo simulation. *Oper. Res.* **33** 661–677.
[40] Schölkopf, B., Herbrich, R. and Smola, A.J. (2001). A generalized representer theorem. *Lecture Notes in Comput. Sci.* **2111** 416–426.
[41] 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
[42] Steinwart, I. (2001). On the influence of the kernel on the consistency of support vector machines. *J. Mach. Learn. Res.* **2** 67–93.
[43] Stuart, A.M. and Teckentrup, A.L. (2018). Posterior consistency for Gaussian process approximations of Bayesian posterior distributions. *Math. Comp.* To appear.
[44] Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.* **4** 389–396. MR1366510
[45] Wendland, H. (2004). *Scattered Data Approximation*. Cambridge: Cambridge Univ. Press.

Bernoulli **25**(2), 2019, 1160–1188
https://doi.org/10.3150/17-BEJ1017

Mallows and generalized Mallows model for matchings

EKHINE IRUROZKI[1,2,*] BORJA CALVO[2] and JOSE A. LOZANO[1,2,**]

[1]*Basque Center for Applied Mathematics (BCAM), Mazarredo 14, 48009 Bilbao, Spain. E-mail:* [*]*eirurozki@bcamath.org;* [**]*jlozano@bcamath.org*
[2]*Intelligent Systems Group, University of the Basque Country, Manuel Lardizabal, 1, 20018 Donostia/San Sebastián, Spain. E-mail: borja.calvo@ehu.eus*

The Mallows and Generalized Mallows Models are two of the most popular probability models for distributions on permutations. In this paper, we consider both models under the Hamming distance. This models can be seen as models for matchings instead of models for rankings. These models cannot be factorized, which contrasts with the popular MM and GMM under Kendall's-τ and Cayley distances. In order to overcome the computational issues that the models involve, we introduce a novel method for computing the partition function. By adapting this method we can compute the expectation, joint and conditional probabilities. All these methods are the basis for three sampling algorithms, which we propose and analyze. Moreover, we also propose a learning algorithm. All the algorithms are analyzed both theoretically and empirically, using synthetic and real data from the context of e-learning and Massive Open Online Courses (MOOC).

Keywords: Generalized Mallows Model; hamming; learning; Mallows Model; matching; sampling

References

[1] Ali, A. and Meila, M. (2012). Experiments with Kemeny ranking: What works when? *Math*. *Social Sci*. **64** 28–40.
[2] Baker, F.B. and Harwell, M.R. (1996). Computing elementary symmetric functions and their derivatives: A didactic. *Appl*. *Psychol*. *Meas*. **20** 169–192.
[3] Belongie, S., Malik, J. and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Trans*. *Pattern Anal*. *Mach*. *Intell*. **24** 509–522.
[4] Busse, L.M. and Buhmann, J.M. (2007). Cluster analysis of heterogeneous rank data. In *International Conference on Machine Learning* 113–120.
[5] Ceberio, J., Irurozki, E., Mendiburu, A. and Lozano, J.A. (2014). Extending distance-based ranking models in estimation of distribution algorithms. In *IEEE Congress on Evolutionary Computation* 2459–2466.
[6] Ceberio, J., Mendiburu, A. and Lozano, J.A. (2015). Kernels of Mallows models for solving permutation-based problems. In *Genetic and Evolutionary Computation Conference* (*GECCO*-2015) 505–512. Madrid.
[7] Chen, L. and Pu, P. (2004). Survey of preference elicitation methods. Technical report.
[8] Cheng, W. and Hüllermeier, E. (2009). A new instance-based label ranking approach using the Mallows model. In *Advances in Neural Networks* (*ISNN*). *Lecture Notes in Computer Science* **5551** 707–716. Springer.
[9] Cheng, W. and Hullermeier, E. (2009). A simple instance-based approach to multilabel classification using the Mallows model. In *Workshop Proceedings of Learning from Multi-Label Data* 28–38. Bled, Slovenia.

1350-7265

[10] Coppersmith, D., Fleischer, L.K. and Rurda, A. (2010). Ordering by weighted number of wins gives a good ranking for weighted tournaments. *ACM Trans*. *Algorithms* **6** Art. 55, 13. MR2682624
[11] Critchlow, D.E., Fligner, M.A. and Verducci, J.S. (1991). Probability models on rankings. *J*. *Math*. *Psych*. **35** 294–318.
[12] Csiszár, V. (2008). Conditional independence relations and log-linear models for random matchings. *Acta Math*. *Hungar.* **122** 131–152.
[13] Csiszár, V. (2009). On L-decomposability of random orderings. *J*. *Math*. *Psych*. **53** 294–297.
[14] D'Elia, A. and Piccolo, D. (2005). A mixture model for preferences data analysis. *Comput*. *Statist*. *Data Anal*. **49** 917–934. MR2141426
[15] Deza, M. and Huang, T. (1998). Metrics on permutations, a survey. *J*. *Comb*. *Inf*. *Syst*. *Sci*. **23** 173–185. MR1737796
[16] Diaconis, P. (1988). *Group Representations in Probability and Statistics*. IMS.
[17] Farah, M. and Vanderpooten, D. (2007). An outranking approach for rank aggregation in information retrieval. In *Conference on Research and Development in Information Retrieval* (*ACM SIGIR*), *SIGIR* '07 591–598. New York, NY, USA: ACM.
[18] Flajolet, P., Zimmermann, P. and Van Cutsem, B. (1994). A calculus for the random generation of labelled combinatorial structures. *Theoret*. *Comput*. *Sci*. **132** 1–35.
[19] Fligner, M.A. and Verducci, J.S. (1986). Distance based ranking models. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **48** 359–369.
[20] Fligner, M.A. and Verducci, J.S. (1990). Posterior probabilities for a consensus ordering. *Psychometrika* **55** 53–63. MR1060264
[21] Gnedin, A. and Olshanski, G. (2012). The two-sided infinite extension of the Mallows model for random permutations. *Adv*. *in Appl*. *Math*. **48** 615–639.
[22] Gupta, J. and Damien, P. (2002). Conjugacy class prior distributions on metric-based ranking models. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **64** 433–445.
[23] Huang, J. and Guestrin, C. (2012). Uncovering the riffled independence structure of rankings. *Electron*. *J*. *Stat*. **6** 199–230.
[24] Huang, J., Guestrin, C., Jiang, X. and Guibas, L. (2009). Exploiting probabilistic independence for permutations. In *Artificial Intelligence and Statistics* (*AISTATS*).
[25] Irurozki, E. (2014). Sampling and Learning Distance-Based Probability Models for Permutation Spaces. Ph.D. thesis, University of the Basque Country.
[26] Irurozki, E., Calvo, B. and Lozano, J.A. (2016). PerMallows: An R package for Mallows and generalized Mallows models. *J*. *Stat*. *Softw*. **71** 1–30.
[27] Irurozki, E., Calvo, B. and Lozano, J.A. (2019). Supplement to "Mallows and generalized Mallows model for matchings." DOI:10.3150/17-BEJ1017SUPP.
[28] Jonker, R. and Volgenant, A. (1987). A shortest augmenting path algorithm for dense and sparse linear assignment problems. *The Netherlands Computing* **38** 325–340.
[29] Kammerdiner, A., Krokhmal, P.A. and Pardalos, P.M. (2009). On the Hamming distance in combinatorial optimization problems on hypergraph matchings. *Optim*. *Lett*. **4** 609–617.
[30] Kondor, R., Howard, A. and Jebara, T. (2007). Multi-object tracking with representations of the symmetric group. In *International Conference on Artificial Intelligence and Statistics* 211–218.
[31] Kuhn, H.W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly* **2** 83–97.
[32] Kuncheva, L.I. (2010). Full-class set classification using the Hungarian algorithm. *Int*. *J*. *Mach*. *Learn*. *Cybern*. **1** 53–61.
[33] Lee, P.H. and Yu, P.L.H. (2012). Mixtures of weighted distance-based models for ranking data with applications in political studies. *Comput*. *Statist*. *Data Anal*. **56** 2486–2500.

[34] Lu, T. and Boutilier, C. (2011). Learning Mallows models with pairwise preferences. In *International Conference on Machine Learning* (*ICML*) 145–152.

[35] Luce, R.D. (1959). *Individual Choice Behavior*: *A Theoretical Analysis*. New York: Wiley. MR0108411

[36] Mallows, C.L. (1957). Non-null ranking models. *Biometrika* **44** 114–130.

[37] Mandhani, B. and Meila, M. (2009). Tractable search for learning exponential models of rankings. *J. Mach. Learn. Res.* **5** 392–399.

[38] Mao, Y. and Lebanon, G. (2008). Non-parametric modeling of partially ranked data. *J. Mach. Learn. Res.* **9** 2401–2429.

[39] Meila, M. and Bao, L. (2008). Estimation and clustering with infinite rankings. In *Uncertainty in Artificial Intelligence* (*UAI*) 393–402. Corvallis, Oregon: AUAI Press.

[40] Meila, M. and Chen, H. (2010). Dirichlet process mixtures of generalized Mallows models. In *Uncertainty in Artificial Intelligence* (*UAI*) 285–294.

[41] Meila, M. and Chen, H. (2016). Bayesian non-parametric clustering of ranking data. *IEEE Trans. Pattern Anal. Mach. Intell.* **38** 2156–2169.

[42] Murphy, T.B. and Martin, D. (2003). Mixtures of distance-based models for ranking data. *Comput. Statist. Data Anal.* **41** 645–655. MR1973732

[43] Plackett, R.L. (1975). The analysis of permutations. *J. R. Stat. Soc.*, *A* **24** 193–202.

[44] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). *Numerical Recipes in C*: *The Art of Scientific Computing*, 2nd ed. New York, NY: Cambridge Univ. Press.

[45] Schiavinotto, T. and Stützle, T. (2007). A review of metrics on permutations for search landscape analysis. *Comput. Oper. Res.* **34** 3143–3153.

[46] Sloane, N.J.A. (2009). Subfactorial or rencontres numbers, or derangements. Available at https://oeis.org/A000166.

[47] Stanley, R.P. (1986). *Enumerative Combinatorics*. Belmont, CA, USA: Wadsworth Publishing Company.

[48] Thurstone, L.L. (1927). A law of comparative judgment. *Psychological Review* **24** 273–286.

[49] UNWTO (2016). UNWTO Tourism Highlights, 2016 Edition.

[50] Vitelli, V., Sørensen, Ø., Crispino, M., Frigessi, A. and Arjas, E. (2014). Probabilistic preference learning with the Mallows rank model. Preprint. Available at arXiv:1405.7945.

[51] Wang, Y., Makedon, F., Ford, J. and Huang, H. (2004). A bipartite graph matching framework for finding correspondences between structural elements in two proteins. In *Engineering in Medicine and Biology Society*, 2004. *IEMBS'*04. 26*th Annual International Conference of the IEEE* **2** 2972–2975. IEEE.

Bernoulli **25**(2), 2019, 1189–1224
https://doi.org/10.3150/17-BEJ1018

Stable limit theorems for empirical processes under conditional neighborhood dependence

JI HYUNG LEE[1] and KYUNGCHUL SONG[2]

[1]*Department of Economics, University of Illinois, 1407 W. Gregory Dr., 214 David Kinley Hall, Urbana, IL 61801, USA. E-mail: jihyung@illinois.edu*
[2]*Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver, BC, Canada, V6T 1L4. E-mail: kysong@mail.ubc.ca*

This paper introduces a new concept of stochastic dependence among many random variables which we call conditional neighborhood dependence (CND). Suppose that there are a set of random variables and a set of sigma algebras where both sets are indexed by the same set endowed with a neighborhood system. When the set of random variables satisfies CND, any two non-adjacent sets of random variables are conditionally independent given sigma algebras having indices in one of the two sets' neighborhood. Random variables with CND include those with conditional dependency graphs and a class of Markov random fields with a global Markov property. The CND property is useful for modeling cross-sectional dependence governed by a complex, large network. This paper provides two main results. The first result is a stable central limit theorem for a sum of random variables with CND. The second result is a Donsker-type result of stable convergence of empirical processes indexed by a class of functions satisfying a certain bracketing entropy condition when the random variables satisfy CND.

Keywords: conditional neighborhood dependence; dependency graphs; empirical processes; Markov random fields; maximal inequalities; stable central limit theorem

References

[1] Aldous, D.J. and Eagleson, G.K. (1978). On mixing and stability of limit theorems. *Ann. Probab.* **6** 325–331.
[2] Andrews, D.W.K. (1994). Empirical process methods in econometrics. In *Handbook of Econometrics* 2247–2294. Elsevier.
[3] Baldi, P. and Rinott, Y. (1989). On normal approximations of distributions in terms of dependency graphs. *Ann. Probab.* **17** 1646–1650.
[4] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286** 509–512.
[5] Berti, P., Pratelli, L. and Rigo, P. (2012). Limit theorems for empirical processes based on dependent data. *Electron. J. Probab.* **9** 1–18.
[6] Cai, T.T., Liu, W. and Zhou, H.H. (2016). Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *Ann. Statist.* **44** 455–488. MR3476606
[7] Canen, N., Schwartz, J. and Song, K. (2017). Estimating local interactions among many agents who observe their neighbors. Working paper.
[8] Chen, L. and Wu, W.B. (2016). Stability and asymptotics for autoregressive processes. *Electron. J. Stat.* **10** 3723–3751.

[9] Chen, L.H.Y. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028.
[10] Dawid, P.A. (1979). Conditional independence in statistical theory. *J. Roy. Statist. Soc. Ser. B* **41** 1–31.
[11] Dudley, R.M. (1985). An extended Wichura theorem, definitions of donsker class, and weighted empirical distributions. In *Probability in Banach Spaces V* (A. Beck, R. Dudley, M. Hahn, J. Kuelbs and M. Marcus, eds.) 141–178. Springer.
[12] Hahn, J., Kuersteiner, G. and Mazzocco, M. (2016). Central limit theory for combined cross-section and time series. Preprint. Available at arXiv:1610.01697 [stat.ME].
[13] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*. New York: Academic Press, Inc. MR0624435
[14] Häusler, E. and Luschgy, H. (2010). *Stable Convergence and Stable Limit Theorems*. New York, USA: Springer Science+Business Media.
[15] Ibragimov, L.A. and Linnik, Y.V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff.
[16] Janson, S. (1988). Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.* **16** 305–312. MR0920273
[17] Janson, S. (2004). Large deviations for sums of partly dependent random variables. *Random Structures Algorithms* **24** 234–248.
[18] Jenish, N. and Prucha, I.R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *J. Econometrics* **150** 86–98.
[19] Jirak, M. (2016). Berry–Esseen theorems under weak dependence. *Ann. Probab.* **44** 2024–2063. MR3502600
[20] Kallenberg, O. (1997). *Foundations of Modern Probability*. New York: Springer.
[21] Koller, D. and Friedman, N. (2009). *Probabilistic Graphical Models*: *Principles and Techniques*. Cambridge, Massachusetts: The MIT Press.
[22] Kuersteiner, G.M. and Prucha, I.R. (2013). Limit theory for panel data models with cross sectional dependence and sequential exogeneity. *J. Econometrics* **174** 107–126.
[23] Lauritzen, S.L. (1996). *Graphical Models*. New York: Springer.
[24] Lauritzen, S.L., Dawid, A.P., Larsen, B.N. and Leimer, H.-G. (1990). Independence properties of directed Markov fields. *Networks* **20** 491–505.
[25] Leung, M.P. (2016). Treatment and spillover effects under network interference. Working paper.
[26] Meinshausen, N. and Bühlmann, P. (2008). High-dimensional graphs and variable selection with the LASSO. *Ann. Statist.* **34** 1436–1462.
[27] Penrose, M. (2003). *Random Geometric Graphs*. Oxford, UK: Oxford Univ. Press.
[28] Pollard, D. (1990). *Empirical Processes*: *Theorey and Applications*. *NSF-CBMS Regional Conference Series in Probability and Statistics* **2**. Hayward, USA: IMS.
[29] Rinott, Y. and Rotar, V. (1996). A multivariate CLT for local dependence with $n^{1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56** 333–350.
[30] Song, K. (2015). Measuring the graph concordance of locally dependent observations. Preprint. Available at arXiv:1504.03712v2 [stat.ME].
[31] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. 583–602. MR0402873
[32] Tihomirov, A.N. (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25** 790–809.
[33] van der Vaart, A.W. (1996). New donsker classes. *Ann. Statist.* **24** 2128–2140.
[34] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. New York, USA: Springer.

[35] Wellner, J.A. (2005). Empirical processes: Theory and applications. Special topics course notes, Delft Technical University.

[36] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154.

[37] Yuan, D. and Lei, L. (2016). Some results following from conditional characteristic functions. *Comm. Statist. Theory Methods* **45** 3706–3720.

Bernoulli **25**(2), 2019, 1225–1255
https://doi.org/10.3150/18-BEJ1019

Oracle inequalities for high-dimensional prediction

JOHANNES LEDERER[1,*], LU YU[1,**] and IRINA GAYNANOVA[2]

[1]*Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195, USA.*
E-mail: [*]*ledererj@uw.edu;* [**]*luyu92@uw.edu*
[2]*Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843, USA.*
E-mail: irinag@stat.tamu.edu

The abundance of high-dimensional data in the modern sciences has generated tremendous interest in penalized estimators such as the lasso, scaled lasso, square-root lasso, elastic net, and many others. In this paper, we establish a general oracle inequality for prediction in high-dimensional linear regression with such methods. Since the proof relies only on convexity and continuity arguments, the result holds irrespective of the design matrix and applies to a wide range of penalized estimators. Overall, the bound demonstrates that generic estimators can provide consistent prediction with any design matrix. From a practical point of view, the bound can help to identify the potential of specific estimators, and they can help to get a sense of the prediction accuracy in a given application.

Keywords: high-dimensional regression; oracle inequalities; prediction

References

[1] Bellec, P., Dalalyan, A., Grappin, E. and Paris, Q. (2016). On the prediction loss of the lasso in the partially labeled setting. Available at arXiv:1606.06179.

[2] Bellec, P., Lecué, G. and Tsybakov, A. (2016). Slope meets Lasso: Improved oracle bounds and optimality. Available at arXiv:1605.08651.

[3] Belloni, A., Chernozhukov, V. and Wang, L. (2011). Square-root lasso: Pivotal recovery of sparse signals via conic programming. *Biometrika* **98** 791–806.

[4] Bickel, P., Ritov, Y. and Tsybakov, A. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732.

[5] Bien, J., Taylor, J. and Tibshirani, R. (2013). A lasso for hierarchical interactions. *Ann. Statist.* **41** 1111–1141.

[6] 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

[7] Bondell, H. and Reich, B. (2008). Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR. *Biometrics* **64** 115–123.

[8] Bühlmann, P. (2013). Statistical significance in high-dimensional linear models. *Bernoulli* **19** 1212–1242.

[9] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Berlin: Springer.

[10] Bunea, F., Lederer, J. and She, Y. (2014). The Group Square-Root Lasso: Theoretical Properties and Fast Algorithms. *IEEE Trans. Inform. Theory* **60** 1313–1325.

[11] Bunea, F., Tsybakov, A. and Wegkamp, M. (2007). Sparsity oracle inequalities for the Lasso. *Electron. J. Stat.* **1** 169–194.

[12] Chatterjee, S. (2013). Assumptionless consistency of the lasso. Available at arXiv:1303.5817.

[13] Chatterjee, S. (2014). A new perspective on least squares under convex constraint. *Ann. Statist.* **42** 2340–2381. MR3269982

[14] Chételat, D., Lederer, J. and Salmon, J. (2017). Optimal two-step prediction in regression. *Electron. J. Stat.* **11** 2519–2546.

[15] Chichignoud, M., Lederer, J. and Wainwright, M. (2016). A practical scheme and fast algorithm to tune the lasso with optimality guarantees. *J. Mach. Learn. Res.* **17** 1–20.

[16] Dalalyan, A., Hebiri, M. and Lederer, J. (2017). On the prediction performance of the Lasso. *Bernoulli* **23** 552–581. MR3556784

[17] Giraud, C., Huet, S. and Verzelen, N. (2012). High-dimensional regression with unknown variance. *Statist. Sci.* **27** 500–518. MR3025131

[18] Greenshtein, E. and Ritov, Y. (2004). Persistence in high-dimensional linear predictor selection and the virtue of overparametrization. *Bernoulli* **10** 971–988. MR2108039

[19] Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. Boca Raton, FL: Chapman & Hall.

[20] Hebiri, M. and Lederer, J. (2013). How Correlations Influence Lasso Prediction. *IEEE Trans. Inform. Theory* **59** 1846–1854.

[21] Hebiri, M. and van de Geer, S. (2011). The smooth-lasso and other $\ell_1 + \ell_2$-penalized methods. *Electron. J. Stat.* **5** 1184–1226.

[22] Jacob, L., Obozinski, G. and Vert, J.-P. (2009). Group lasso with overlap and graph lasso. In *ICML* 2009 433–440.

[23] Kim, S.-J., Koh, K., Boyd, S. and Gorinevsky, D. (2009). l_1 trend filtering. *SIAM Rev.* **51** 339–360. MR2505584

[24] Koltchinskii, V. (2009). Sparse recovery in convex hulls via entropy penalization. *Ann. Statist.* **37** 1332–1359.

[25] Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*. Berlin: Springer.

[26] Koltchinskii, V., Lounici, K. and Tsybakov, A. (2011). Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Ann. Statist.* **39** 2302–2329.

[27] Lederer, J. and Müller, C. (2015). Don't fall for tuning parameters: Tuning-free variable selection in high dimensions with the TREX. In *AAAI*-15 2729–2735.

[28] Lederer, J. and van de Geer, S. (2014). New concentration inequalities for suprema of empirical processes. *Bernoulli* **20** 2020–2038. MR3263097

[29] Massart, P. and Meynet, C. (2011). The Lasso as an ℓ_1-ball model selection procedure. *Electron. J. Stat.* **5** 669–687. MR2820635

[30] Rigollet, P. and Tsybakov, A. (2011). Exponential Screening and optimal rates of sparse estimation. *Ann. Statist.* **39** 731–771.

[31] Rudin, L., Osher, S. and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D* **60** 259–268.

[32] Simon, N., Friedman, J., Hastie, T. and Tibshirani, R. (2013). A sparse-group lasso. *J. Comput. Graph. Statist.* **22** 231–245.

[33] Su, W. and Candès, E. (2016). SLOPE is adaptive to unknown sparsity and asymptotically minimax. *Ann. Statist.* **44** 1038–1068. MR3485953

[34] Sun, T. and Zhang, C.-H. (2012). Scaled sparse linear regression. *Biometrika* **99** 879–898. MR2999166

[35] Sun, T. and Zhang, C.-H. (2013). Sparse matrix inversion with scaled lasso. *J. Mach. Learn. Res.* **14** 3385–3418.
[36] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **58** 267–288.
[37] Tibshirani, R. (2014). Adaptive piecewise polynomial estimation via trend filtering. *Ann. Statist.* **42** 285–323.
[38] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J. and Knight, K. (2005). Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 91–108.
[39] Tsybakov, A. (2009). *Introduction to Nonparametric Estimation*. New York: Springer.
[40] van de Geer, S. and Muro, A. (2014). On higher order isotropy conditions and lower bounds for sparse quadratic forms. *Electron. J. Stat.* **8** 3031–3061. MR3301300
[41] van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes*. New York: Springer.
[42] van de Geer, S. (2007). The deterministic Lasso. In 2007 *Proc. Amer. Math. Soc.* [*CD-ROM*]. Available at www.stat.math.ethz.ch/~geer/lasso.pdf.
[43] van de Geer, S. and Bühlmann, P. (2009). On the conditions used to prove oracle results for the Lasso. *Electron. J. Stat.* **3** 1360–1392.
[44] van de Geer, S. and Lederer, J. (2013). The Lasso, correlated design, and improved oracle inequalities. *Inst. Math. Stat. Collect.* **9** 303–316.
[45] Wainwright, M. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). *IEEE Trans. Inform. Theory* **55** 2183–2202.
[46] Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 49–67.
[47] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 301–320. MR2137327

Bernoulli **25**(2), 2019, 1256–1288
https://doi.org/10.3150/18-BEJ1020

Truncated random measures

TREVOR CAMPBELL[1,*], JONATHAN H. HUGGINS[1,**], JONATHAN P. HOW[2] and TAMARA BRODERICK[1,†]

[1]*Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail:* [*]*tdjc@mit.edu;* [**]*jhuggins@mit.edu;* [†]*tbroderick@csail.mit.edu*
[2]*Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: jhow@mit.edu*

Completely random measures (CRMs) and their normalizations are a rich source of Bayesian nonparametric priors. Examples include the beta, gamma, and Dirichlet processes. In this paper, we detail two major classes of sequential CRM representations—*series representations* and *superposition representations*—within which we organize both novel and existing sequential representations that can be used for simulation and posterior inference. These two classes and their constituent representations subsume existing ones that have previously been developed in an ad hoc manner for specific processes. Since a complete infinite-dimensional CRM cannot be used explicitly for computation, sequential representations are often truncated for tractability. We provide truncation error analyses for each type of sequential representation, as well as their normalized versions, thereby generalizing and improving upon existing truncation error bounds in the literature. We analyze the computational complexity of the sequential representations, which in conjunction with our error bounds allows us to directly compare representations and discuss their relative efficiency. We include numerous applications of our theoretical results to commonly-used (normalized) CRMs, demonstrating that our results enable a straightforward representation and analysis of CRMs that has not previously been available in a Bayesian nonparametric context.

Keywords: Bayesian nonparametrics; completely random measure; normalized completely random measure; Poisson point process; truncation

References

[1] Abramowitz, M. and Stegun, I.A., eds. (1964). *Handbook of Mathematical Functions*, *with Formulas*, *Graphs*, *and Mathematical Tables*. New York: Dover Publications, Inc. MR0208797
[2] Airoldi, E.M., Blei, D., Erosheva, E.A. and Fienberg, S.E. (2014). *Handbook of Mixed Membership Models and Their Applications*. Boca Raton, FL: CRC Press.
[3] Arbel, J. and Prünster, I. (2017). A moment-matching Ferguson & Klass algorithm. *Stat*. *Comput*. **27** 3–17. MR3598905
[4] 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
[5] Banjevic, D., Ishwaran, H. and Zarepour, M. (2002). A recursive method for functionals of Poisson processes. *Bernoulli* **8** 295–311. MR1913109
[6] Blei, D.M. and Jordan, M.I. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Anal*. **1** 121–143. MR2227367
[7] Bondesson, L. (1982). On simulation from infinitely divisible distributions. *Adv. in Appl*. *Probab.* **14** 855–869. MR0677560
[8] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl*. *Probab.* **31** 929–953. MR1747450

1350-7265

[9] Broderick, T., Jordan, M.I. and Pitman, J. (2012). Beta processes, stick-breaking and power laws. *Bayesian Anal*. **7** 439–475. MR2934958
[10] Broderick, T., Mackey, L., Paisley, J. and Jordan, M.I. (2015). Combinatorial clustering and the beta negative binomial process. *IEEE Trans*. *Pattern Anal*. *Mach*. *Intell*. **37** 290–306.
[11] Broderick, T., Wilson, A.C. and Jordan, M.I. (2018). Posteriors, conjugacy, and exponential families for completely random measures. *Bernoulli* **24** 3181–3221. MR3788171
[12] Campbell, T., Huggins, J.H., How, J.P. and Broderick, T. (2019). Supplement to "Truncated random measures". DOI:10.3150/18-BEJ1020SUPP.
[13] Doshi-Velez, F., Miller, K.T., Van Gael, J. and Teh, Y.W. (2009). Variational inference for the Indian buffet process. In *International Conference on Artificial Intelligence and Statistics*.
[14] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann*. *Statist*. **1** 209–230. MR0350949
[15] Ferguson, T.S. and Klass, M.J. (1972). A representation of independent increment processes without Gaussian components. *Ann*. *Math*. *Stat*. **43** 1634–1643. MR0373022
[16] Gelfand, A.E. and Kottas, A. (2002). A computational approach for full nonparametric Bayesian inference under Dirichlet process mixture models. *J*. *Comput*. *Graph*. *Statist*. **11** 289–305. MR1938136
[17] Gumbel, E.J. (1954). *Statistical Theory of Extreme Values and Some Practical Applications*. *A Series of Lectures*. *National Bureau of Standards Applied Mathematics Series* **33**. Washington, DC: U.S. Government Printing Office. MR0061342
[18] Hjort, N.L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann*. *Statist*. **18** 1259–1294. MR1062708
[19] Ishwaran, H. and James, L.F. (2001). Gibbs sampling methods for stick-breaking priors. *J*. *Amer*. *Statist*. *Assoc*. **96** 161–173. MR1952729
[20] Ishwaran, H. and James, L.F. (2002). Approximate Dirichlet process computing in finite normal mixtures: Smoothing and prior information. *J*. *Comput*. *Graph*. *Statist*. **11** 508–532. MR1938445
[21] Ishwaran, H. and Zarepour, M. (2002). Exact and approximate sum representations for the Dirichlet process. *Canad*. *J*. *Statist*. **30** 269–283. MR1926065
[22] James, L.F. (2002). Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics. Preprint. Available at arXiv:0205093.
[23] James, L.F. (2013). Stick-breaking PG(α, ζ)-generalized gamma processes. Preprint. Available at arXiv:1308.6570.
[24] James, L.F. (2014). Poisson latent feature calculus for generalized Indian buffet processes. Preprint. Available at arXiv:1411.2936v3.
[25] 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
[26] Kim, Y. (1999). Nonparametric Bayesian estimators for counting processes. *Ann*. *Statist*. **27** 562–588. MR1714717
[27] Kingman, J.F.C. (1967). Completely random measures. *Pacific J*. *Math*. **21** 59–78. MR0210185
[28] Kingman, J.F.C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. New York: The Clarendon Press, Oxford Univ. Press. MR1207584
[29] 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. With a discussion by S.J. Taylor, A.G. Hawkes, A.M. Walker, D.R. Cox, A.F.M. Smith, B.M. Hill, P.J. Burville, T. Leonard and a reply by the author. MR0368264
[30] Lijoi, A., Mena, R.H. and Prünster, I. (2005). Bayesian nonparametric analysis for a generalized Dirichlet process prior. *Stat*. *Inference Stoch*. *Process*. **8** 283–309. MR2177315
[31] Lijoi, A., Mena, R.H. and Prünster, I. (2007). Controlling the reinforcement in Bayesian nonparametric mixture models. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **69** 715–740. MR2370077

[32] Lijoi, A. and Prünster, I. (2003). On a normalized random measure with independent increments relevant to Bayesian nonparametric inference. In *Proceedings of the* 13*th European Young Statisticians Meeting* 123–124.
[33] Lijoi, A. and Prünster, I. (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (N.L. Hjort, C. Holmes, P. Müller and S. Walker, eds.). *Camb. Ser. Stat. Probab. Math.* **28** 80–136. Cambridge: Cambridge Univ. Press. MR2730661
[34] Maddison, C., Tarlow, D. and Minka, T.P. (2014). A* sampling. In *Advances in Neural Information Processing Systems*.
[35] Muliere, P. and Tardella, L. (1998). Approximating distributions of random functionals of Ferguson–Dirichlet priors. *Canad. J. Statist.* **26** 283–297. MR1648431
[36] Orbanz, P. (2010). Conjugate projective limits. Preprint. Available at arXiv:1012.0363.
[37] Paisley, J.W., Blei, D.M. and Jordan, M.I. (2012). Stick-breaking beta processes and the Poisson process. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*. *Proceedings of Machine Learning Research* **22** 850–858.
[38] Paisley, J.W., Zaas, A.K., Woods, C.W., Ginsburg, G.S. and Carin, L. (2010). A stick-breaking construction of the beta process. In *International Conference on Machine Learning*.
[39] Perman, M. (1993). Order statistics for jumps of normalised subordinators. *Stochastic Process. Appl.* **46** 267–281. MR1226412
[40] 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
[41] Pitman, J. (2003). Poisson–Kingman partitions. In *Statistics and Science: A Festschrift for Terry Speed. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **40** 1–34. Beachwood, OH: IMS. MR2004330
[42] Regazzini, E., Lijoi, A. and Prünster, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. Dedicated to the memory of Herbert E. Robbins. MR1983542
[43] Rosiński, J. (1990). On series representations of infinitely divisible random vectors. *Ann. Probab.* **18** 405–430. MR1043955
[44] Rosiński, J. (2001). Series representations of Lévy processes from the perspective of point processes. In *Lévy Processes* (O. Barndorff-Nielson, S. Resnick and T. Mikosch, eds.) 401–415. Boston, MA: Birkhäuser. MR1833707
[45] Roy, D. (2014). The continuum-of-urns scheme, generalized beta and Indian buffer processes, and hierarchies thereof. Preprint. Available at arXiv:1501.00208.
[46] Roychowdhury, A. and Kulis, B. (2015). Gamma processes, stick-breaking, and variational inference. In *International Conference on Artificial Intelligence and Statistics*.
[47] Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica* **4** 639–650. MR1309433
[48] Teh, Y.W. and Görür, D. (2009). Indian buffet processes with power-law behavior. In *Advances in Neural Information Processing Systems*.
[49] Thibaux, R. and Jordan, M.I. (2007). Hierarchical beta processes and the Indian buffet process. In *International Conference on Artificial Intelligence and Statistics*.
[50] Titsias, M. (2008). The infinite gamma-Poisson feature model. In *Advances in Neural Information Processing Systems*.
[51] Zhou, M., Hannah, L., Dunson, D. and Carin, L. (2012). Beta-negative binomial process and Poisson factor analysis. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*. *Proceedings of Machine Learning Research* **22** 1462–1471.

Bernoulli **25**(2), 2019, 1289–1325
https://doi.org/10.3150/18-BEJ1021

Minimax optimal estimation in partially linear additive models under high dimension

ZHUQING YU*, MICHAEL LEVINE** and GUANG CHENG†

Department of Statistics, Purdue University, 250 N. University Street, West Lafayette, IN 47907, USA.
E-mail: *zhuqing.yu.stat@gmail.com; **mlevins@purdue.edu; †chengg@stat.purdue.edu

In this paper, we derive minimax rates for estimating both parametric and nonparametric components in partially linear additive models with high dimensional sparse vectors and smooth functional components. The minimax lower bound for Euclidean components is the typical sparse estimation rate that is independent of nonparametric smoothness indices. However, the minimax lower bound for each component function exhibits an interplay between the dimensionality and sparsity of the parametric component and the smoothness of the relevant nonparametric component. Indeed, the minimax risk for smooth nonparametric estimation can be slowed down to the sparse estimation rate whenever the smoothness of the nonparametric component or dimensionality of the parametric component is sufficiently large. In the above setting, we demonstrate that penalized least square estimators can nearly achieve minimax lower bounds.

Keywords: high dimension; minimax optimal; partial linear additive model; semiparametric

References

[1] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. *Johns Hopkins Series in the Mathematical Sciences*. Baltimore, MD: Johns Hopkins Univ. Press. MR1245941

[2] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data*: *Methods*, *Theory and Applications*. *Springer Series in Statistics*. Heidelberg: Springer. MR2807761

[3] Cheng, G., Zhang, H.H. and Shang, Z. (2015). Sparse and efficient estimation for partial spline models with increasing dimension. *Ann*. *Inst*. *Statist*. *Math*. **67** 93–127. MR3297860

[4] Gilbert, E.N. (1952). A comparison of signalling alphabets. *Bell Syst*. *Tech*. *J*. **31** 504–522.

[5] Härdle, W., Liang, H. and Gao, J. (2000). *Partially Linear Models*. *Contributions to Statistics*. Heidelberg: Physica-Verlag. MR1787637

[6] Horowitz, J., Klemelä, J. and Mammen, E. (2006). Optimal estimation in additive regression models. *Bernoulli* **12** 271–298. MR2218556

[7] Koltchinskii, V. and Yuan, M. (2010). Sparsity in multiple kernel learning. *Ann*. *Statist*. **38** 3660–3695. MR2766864

[8] Ma, C. and Huang, J. (2016). Asymptotic properties of lasso in high-dimensional partially linear models. *Sci*. *China Math*. **59** 769–788. MR3474502

[9] Massart, P. (2007). *Concentration Inequalities and Model Selection*. *Lecture Notes in Math*. **1896**. Berlin: Springer. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard. MR2319879

[10] Müller, P. and van de Geer, S. (2015). The partial linear model in high dimensions. *Scand*. *J*. *Stat*. **42** 580–608. MR3345123

1350-7265

[11] Nickl, R. and van de Geer, S. (2013). Confidence sets in sparse regression. *Ann. Statist.* **41** 2852–2876. MR3161450
[12] Nussbaum, M. (1985). Spline smoothing in regression models and asymptotic efficiency in L_2. *Ann. Statist.* **13** 984–997. MR0803753
[13] Pinsker, M.S. Optimal filtration of square-integrable signals in Gaussian noise. MR0624591
[14] Raskutti, G., Wainwright, M.J. and Yu, B. (2011). Minimax rates of estimation for high-dimensional linear regression over ℓ_q-balls. *IEEE Trans. Inform. Theory* **57** 6976–6994. MR2882274
[15] Raskutti, G., Wainwright, M.J. and Yu, B. (2012). Minimax-optimal rates for sparse additive models over kernel classes via convex programming. *J. Mach. Learn. Res.* **13** 389–427. MR2913704
[16] Stone, C.J. (1985). Additive regression and other nonparametric models. *Ann. Statist.* **13** 689–705. MR0790566
[17] Suzuki, T. and Sugiyama, M. (2013). Fast learning rate of multiple kernel learning: Trade-off between sparsity and smoothness. *Ann. Statist.* **41** 1381–1405. MR3113815
[18] Tsybakov, A.B. (2009). *Introduction to nonparametric estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359
[19] van de Geer, S. (2014). On the uniform convergence of empirical norms and inner products, with application to causal inference. *Electron. J. Stat.* **8** 543–574. MR3211024
[20] van de Geer, S. and Muro, A. (2015). Penalized least squares estimation in the additive model with different smoothness for the components. *J. Statist. Plann. Inference* **162** 43–61. MR3323103
[21] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170
[22] Verzelen, N. (2012). Minimax risks for sparse regressions: Ultra-high dimensional phenomenons. *Electron. J. Stat.* **6** 38–90. MR2879672
[23] Xie, H. and Huang, J. (2009). SCAD-penalized regression in high-dimensional partially linear models. *Ann. Statist.* **37** 673–696. MR2502647
[24] 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
[25] Yu, K., Mammen, E. and Park, B.U. (2011). Semi-parametric regression: Efficiency gains from modeling the nonparametric part. *Bernoulli* **17** 736–748. MR2787613
[26] Yuan, M. and Zhou, D.-X. (2016). Minimax optimal rates of estimation in high dimensional additive models. *Ann. Statist.* **44** 2564–2593. MR3576554
[27] Zhang, H.H., Cheng, G. and Liu, Y. (2011). Linear or nonlinear? Automatic structure discovery for partially linear models. *J. Amer. Statist. Assoc.* **106** 1099–1112. MR2894767
[28] Zhu, Y. (2017). Nonasymptotic analysis of semiparametric regression models with high-dimensional parametric coefficients. *Ann. Statist.* **45** 2274–2298. MR3718169

Bernoulli **25**(2), 2019, 1326–1354
https://doi.org/10.3150/18-BEJ1022

Strong Gaussian approximation of the mixture Rasch model

FRIEDRICH LIESE*, ALEXANDER MEISTER** and JOHANNA KAPPUS†

Institut für Mathematik, Universität Rostock, D-18051 Rostock, Germany.
E-mail: *friedrich.liese@uni-rostock.de; **alexander.meister@uni-rostock.de;*
†*johanna_kappus@t-online.de*

We consider the famous Rasch model, which is applied to psychometric surveys when n persons under test answer m questions. The score is given by a realization of a random binary $n \times m$-matrix. Its (j,k)th component indicates whether or not the answer of the jth person to the kth question is correct. In the mixture, Rasch model one assumes that the persons are chosen randomly from a population. We prove that the mixture Rasch model is asymptotically equivalent to a Gaussian observation scheme in Le Cam's sense as n tends to infinity and m is allowed to increase slowly in n. For that purpose, we show a general result on strong Gaussian approximation of the sum of independent high-dimensional binary random vectors. As a first application, we construct an asymptotic confidence region for the difficulty parameters of the questions.

Keywords: asymptotic equivalence of statistical experiments; high-dimensional central limit theorem; item response model; Le Cam distance; psychometrics

References

[1] Alagumalai, S., Curtis, D.D. and Hungi, N. (2005). *Applied Rasch Measurement: A Book of Exemplars*. Berlin: Springer.
[2] Andersen, E.B. (1977). Sufficient statistics and latent trait models. *Psychometrika* **42** 69–81. MR0483255
[3] Andersen, E.B. (1980). Comparing latent distributions. *Psychometrika* **45** 121–134. MR0570773
[4] Andrich, D. (2010). Sufficiency and conditional estimation of person parameters in the polytomous Rasch model. *Psychometrika* **75** 292–308. MR2719929
[5] Bezruczko, N. (2005). *Rasch Measurement in Health Sciences*. Maple Grove, MN: JAM Press.
[6] Biehler, M., Holling, H. and Doebler, P. (2015). Saddlepoint approximations of the distribution of the person parameter in the two parameter logistic model. *Psychometrika* **80** 665–688. MR3392024
[7] Brown, L.D., Carter, A.V., Low, M.G. and Zhang, C.-H. (2004). Equivalence theory for density estimation, Poisson processes and Gaussian white noise with drift. *Ann. Statist.* **32** 2074–2097. MR2102503
[8] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958
[9] Cai, T.T. and Zhou, H.H. (2009). Asymptotic equivalence and adaptive estimation for robust nonparametric regression. *Ann. Statist.* **37** 3204–3235. MR2549558
[10] Carter, A.V. (2002). Deficiency distance between multinomial and multivariate normal experiments. *Ann. Statist.* **30** 708–730. MR1922539
[11] Carter, A.V. (2006). A continuous Gaussian approximation to a nonparametric regression in two dimensions. *Bernoulli* **12** 143–156. MR2202326

1350-7265

[12] DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*. *Springer Texts in Statistics*. New York: Springer. MR2664452

[13] de Leeuw, J. and Verhelst, N. (1986). Maximum likelihood estimation in generalized Rasch models. *J. Educ. Behav. Stat.* **11** 183–196.

[14] Doebler, A., Doebler, P. and Holling, H. (2012). Optimal and most exact confidence intervals for person parameters in item response theory models. *Psychometrika* **77** 98–115. MR3042821

[15] Fischer, G.H. and Molenaar, I.W., eds. (1995). *Rasch Models*. New York: Springer. MR1367341

[16] Genon-Catalot, V. and Larédo, C. (2014). Asymptotic equivalence of nonparametric diffusion and Euler scheme experiments. *Ann. Statist.* **42** 1145–1165. MR3224284

[17] Haberman, S.J. (1977). Maximum likelihood estimates in exponential response models. *Ann. Statist.* **5** 815–841. MR0501540

[18] Hoderlein, S., Mammen, E. and Yu, K. (2011). Non-parametric models in binary choice fixed effects panel data. *Econom. J.* **14** 351–367. MR2853231

[19] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. *Springer Series in Statistics*. New York: Springer. MR0856411

[20] Le Cam, L. and Yang, G.L. (2000). *Asymptotics in Statistics: Some Basic Concepts*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1784901

[21] Liese, F. and Miescke, K.-J. (2008). *Statistical Decision Theory*. *Springer Series in Statistics*. New York: Springer. MR2421720

[22] Lindsay, B., Clogg, C.C. and Grego, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *J. Amer. Statist. Assoc.* **86** 96–107. MR1137102

[23] Mariucci, E. (2016). Asymptotic equivalence for pure jump Lévy processes with unknown Lévy density and Gaussian white noise. *Stochastic Process. Appl.* **126** 503–541. MR3434992

[24] Meister, A. (2011). Asymptotic equivalence of functional linear regression and a white noise inverse problem. *Ann. Statist.* **39** 1471–1495. MR2850209

[25] Meister, A. and Reiß, M. (2013). Asymptotic equivalence for nonparametric regression with non-regular errors. *Probab. Theory Related Fields* **155** 201–229. MR3010397

[26] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399–2430. MR1425959

[27] Pfanzagl, J. (1993). On the consistency of conditional maximum likelihood estimators. *Ann. Inst. Statist. Math.* **45** 703–719. MR1252949

[28] Pfanzagl, J. (1994). On the identifiability of structural parameters in mixtures: Applications to psychological tests. *J. Statist. Plann. Inference* **38** 309–326. MR1261806

[29] Rasch, G. (1960/1980). *Probabilistic Models for Some Intelligence and Attainment Tests* (*Expanded Edition*). Chicago: Univ. Chicago Press.

[30] Reiß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Statist.* **39** 772–802. MR2816338

[31] Rice, K.M. (2004). Equivalence between conditional and mixture approaches to the Rasch model and matched case-control studies, with applications. *J. Amer. Statist. Assoc.* **99** 510–522. MR2062836

[32] Rohde, A. (2004). On the asymptotic equivalence and rate of convergence of nonparametric regression and Gaussian white noise. *Statist. Decisions* **22** 235–243. MR2125610

[33] Schmidt-Hieber, J. (2014). Asymptotic equivalence for regression under fractional noise. *Ann. Statist.* **42** 2557–2585. MR3277671

[34] Shiryaev, A.N. and Spokoiny, V.G. (2000). *Statistical Experiments and Decisions: Asymptotic theory*. *Advanced Series on Statistical Science & Applied Probability* **8**. River Edge, NJ: World Scientific. MR1791434

[35] Strasser, H. (1985). *Mathematical Theory of Statistics*: *Statistical Experiments and Asymptotic Decision Theory*. *De Gruyter Studies in Mathematics* **7**. Berlin: de Gruyter. MR0812467
[36] Strasser, H. (2012a). The covariance structure of cml-estimates in the Rasch model. *Stat*. *Risk Model*. **29** 315–326. MR2997038
[37] Strasser, H. (2012b). Asymptotic expansions for conditional moments of Bernoulli trials. *Stat*. *Risk Model*. **29** 327–343. MR2997039
[38] von Davier, M. and Carstensen, C.H. (2007). *Multivariate and Mixture Distribution Rasch Models – Extensions and Applications*. *Statistics for Social and Behavioral Sciences*. New York: Springer.

Bernoulli **25**(2), 2019, 1355–1385
https://doi.org/10.3150/18-BEJ1023

Time-frequency analysis of locally stationary Hawkes processes

FRANÇOIS ROUEFF[1] and RAINER VON SACHS[2]

[1]*LTCI, Télécom ParisTech, Université Paris-Saclay, 46, rue Barrault, 75013, Paris, France. E-mail: roueff@telecom-paristech.fr*

[2]*Institut de statistique, biostatistique et sciences actuarielles (ISBA) IMMAQ, Université catholique de Louvain, Voie du Roman Pays 20/L1.04.01, 1348 Louvain-la-Neuve. E-mail: rvs@uclouvain.be*

Locally stationary Hawkes processes have been introduced in order to generalise classical Hawkes processes away from stationarity by allowing for a time-varying second-order structure. This class of self-exciting point processes has recently attracted a lot of interest in applications in the life sciences (seismology, genomics, neuro-science, ...), but also in the modeling of high-frequency financial data. In this contribution, we provide a fully developed nonparametric estimation theory of both local mean density and local Bartlett spectra of a locally stationary Hawkes process. In particular, we apply our kernel estimation of the spectrum localised both in time and frequency to two data sets of transaction times revealing pertinent features in the data that had not been made visible by classical non-localised approaches based on models with constant fertility functions over time.

Keywords: high frequency financial data; locally stationary time series; non-parametric kernel estimation; self-exciting point processes; time frequency analysis

References

[1] Aalen, O.O. (1975). *Statistical inference for a family of counting processes*. Ann Arbor, MI: ProQuest LLC. Ph.D. thesis, Univ. California, Berkeley. MR2625917

[2] Andersen, P.K. and Borgan, Ø. (1985). Counting process models for life history data: A review. *Scand. J. Stat.* **12** 97–158. With discussion and a reply by the authors. MR0808151

[3] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* **13** 65–77. MR3005350

[4] Birr, S., Volgushev, S., Kley, T., Dette, H. and Hallin, M. (2017). Quantile spectral analysis for locally stationary time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1619–1643. MR3731679

[5] 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

[6] Chen, F. and Hall, P. (2016). Nonparametric estimation for self-exciting point processes – A parsimonious approach. *J. Comput. Graph. Statist.* **25** 209–224. MR3474044

[7] Dahlhaus, R. (1996). On the Kullback–Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.* **62** 139–168. MR1388767

[8] Dahlhaus, R. (1996). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. In *Athens Conference on Applied Probability and Time Series Analysis*, *Vol. II* (1995). *Lecture Notes in Statistics* **115** 145–159. New York: Springer. MR1466743

[9] Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28** 1762–1794. MR1835040

[10] Dahlhaus, R. (2009). Local inference for locally stationary time series based on the empirical spectral measure. *J. Econometrics* **151** 101–112. MR2559818

[11] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*, *Vol. I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1950431

[12] Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence*: *With Examples and Applications*. *Lecture Notes in Statistics* **190**. New York: Springer. MR2338725

[13] Giraud, C., Roueff, F. and Sanchez-Perez, A. (2015). Aggregation of predictors for nonstationary sub-linear processes and online adaptive forecasting of time varying autoregressive processes. *Ann. Statist.* **43** 2412–2450. MR3405599

[14] Hansen, N.R., Reynaud-Bouret, P. and Rivoirard, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21** 83–143. MR3322314

[15] Mammen, E. (2017). Nonparametric estimation of locally stationary Hawkes processe. arXiv e-prints.

[16] Neumann, M.H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.* **25** 38–76. MR1429917

[17] Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* **11** 453–466. MR0696058

[18] 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

[19] Richter, S. and Dahlhaus, R. (2017). Cross validation for locally stationary processes. arXiv e-prints.

[20] Roueff, F. and von Sachs, R. (2019). Supplement to "Time-frequency analysis of locally stationary Hawkes processes". DOI:10.3150/18-BEJ1023SUPP.

[21] Roueff, F., von Sachs, R. and Sansonnet, L. (2016). Locally stationary Hawkes processes. *Stochastic Process. Appl.* **126** 1710–1743. MR3483734

[22] van Delft, A. and Eichler, M. (2015). Data-adaptive estimation of time-varying spectral densities. arXiv e-prints.

[23] Zheng, B., Roueff, F. and Abergel, F. (2014). Modelling bid and ask prices using constrained Hawkes processes: Ergodicity and scaling limit. *SIAM J. Financial Math.* **5** 99–136. MR3164121

[24] Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. MR2541444

Bernoulli **25**(2), 2019, 1386–1411
https://doi.org/10.3150/18-BEJ1024

Quenched central limit theorem rates of convergence for one-dimensional random walks in random environments

SUNG WON AHN[1] and JONATHON PETERSON[2]

[1]*Department of Mathematics and Actuarial Science, Roosevelt University, 430 S. Michigan Ave., Chicago, IL 60605, USA. E-mail: sahn02@roosevelt.edu*
[2]*Department of Mathematics, Purdue University, 150 N University Street, West Lafayette, IN 47907, USA. E-mail: peterson@purdue.edu; url: http://www.math.purdue.edu/~peterson*

Unlike classical simple random walks, one-dimensional random walks in random environments (RWRE) are known to have a wide array of potential limiting distributions. Under certain assumptions, however, it is known that CLT-like limiting distributions hold for the walk under both the quenched and averaged measures. We give upper bounds on the rates of convergence for the quenched central limit theorems for both the hitting time and position of the RWRE with polynomial rates of convergence that depend on the distribution on environments.

Keywords: quenched central limit theorem; rates of convergence

References

[1] Ahn, S.W. and Peterson, J. (2016). Oscillations of quenched slowdown asymptotics for ballistic one-dimensional random walk in a random environment. *Electron. J. Probab.* **21** Paper No. 16, 27. MR3485358

[2] Alili, S. (1999). Asymptotic behaviour for random walks in random environments. *J. Appl. Probab.* **36** 334–349. MR1724844

[3] Armstrong, S.N. and Smart, C.K. (2014). Quantitative stochastic homogenization of elliptic equations in nondivergence form. *Arch. Ration. Mech. Anal.* **214** 867–911. MR3269637

[4] Berry, A.C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** 122–136. MR0003498

[5] Caffarelli, L.A. and Souganidis, P.E. (2010). Rates of convergence for the homogenization of fully nonlinear uniformly elliptic pde in random media. *Invent. Math.* **180** 301–360. MR2609244

[6] Dembo, A., Peres, Y. and Zeitouni, O. (1996). Tail estimates for one-dimensional random walk in random environment. *Comm. Math. Phys.* **181** 667–683. MR1414305

[7] Dolgopyat, D. and Goldsheid, I. (2012). Quenched limit theorems for nearest neighbour random walks in 1D random environment. *Comm. Math. Phys.* **315** 241–277. MR2966946

[8] Enriquez, N., Sabot, C., Tournier, L. and Zindy, O. (2013). Quenched limits for the fluctuations of transient random walks in random environment on $\mathbb{Z}^1$. *Ann. Appl. Probab.* **23** 1148–1187. MR3076681

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

[10] Gantert, N. and Peterson, J. (2011). Maximal displacement for bridges of random walks in a random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 663–678. MR2841070

[11] Gantert, N. and Shi, Z. (2002). Many visits to a single site by a transient random walk in random environment. *Stochastic Process. Appl.* **99** 159–176. MR1901151
[12] Gantert, N. and Zeitouni, O. (1998). Quenched sub-exponential tail estimates for one-dimensional random walk in random environment. *Comm. Math. Phys.* **194** 177–190. MR1628294
[13] Gloria, A., Neukamm, S. and Otto, F. (2015). Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199** 455–515. MR3302119
[14] Gloria, A. and Otto, F. (2011). An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.* **39** 779–856. MR2789576
[15] Goldsheid, I.Ya. (2007). Simple transient random walks in one-dimensional random environment: The central limit theorem. *Probab. Theory Related Fields* **139** 41–64. MR2322691
[16] Katz, M.L. (1963). Note on the Berry-Esseen theorem. *Ann. Math. Stat.* **34** 1107–1108. MR0151996
[17] Kesten, H., Kozlov, M.V. and Spitzer, F. (1975). A limit law for random walk in a random environment. *Compos. Math.* **30** 145–168. MR0380998
[18] Mayer-Wolf, E., Roitershtein, A. and Zeitouni, O. (2004). Limit theorems for one-dimensional transient random walks in Markov environments. *Ann. Inst. Henri Poincaré Probab. Stat.* **40** 635–659. MR2086017
[19] Mourrat, J.-C. (2012). A quantitative central limit theorem for the random walk among random conductances. *Electron. J. Probab.* **17** no. 97, 17. MR2994845
[20] Papanicolaou, G.C. and Varadhan, S.R.S. (1982). Diffusions with random coefficients. In *Statistics and Probability: Essays in Honor of C.R. Rao* 547–552. Amsterdam: North-Holland. MR0659505
[21] Peterson, J. (2009). Quenched limits for transient, ballistic, sub-Gaussian one-dimensional random walk in random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 685–709. MR2548499
[22] Peterson, J. and Samorodnitsky, G. (2012). Weak weak quenched limits for the path-valued processes of hitting times and positions of a transient, one-dimensional random walk in a random environment. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** 531–569. MR3069376
[23] Peterson, J. and Samorodnitsky, G. (2013). Weak quenched limiting distributions for transient one-dimensional random walk in a random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 722–752. MR3112432
[24] Peterson, J. and Zeitouni, O. (2009). Quenched limits for transient, zero speed one-dimensional random walk in random environment. *Ann. Probab.* **37** 143–188. MR2489162
[25] Peterson, J. (2008). Limiting distributions and large deviations for random walks in random environments. PhD thesis, University of Minnesota. Available at http://arxiv.org/abs/0810.0257.
[26] Petrov, V.V. (1975). *Sums of Independent Random Variables*. New York–Heidelberg: Springer. Translated from the Russian by A.A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82. MR0388499
[27] Sinaĭ, Ya.G. (1983). The limit behavior of a one-dimensional random walk in a random environment. *Theory Probab. Appl.* **27** 256–268.
[28] Solomon, F. (1975). Random walks in a random environment. *Ann. Probab.* **3** 1–31. MR0362503
[29] Yurinskiĭ, V.V. (1982). Averaging of second-order nondivergent equations with random coefficients. *Sibirsk. Mat. Zh.* **23** 176–188, 217. MR0652234
[30] Yurinskiĭ, V.V. (1988). On the error of averaging of multidimensional diffusions. *Teor. Veroyatn. Primen.* **33** 14–24. MR0939985
[31] Zeitouni, O. (2004). Random walks in random environment. In *Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1837** 189–312. Berlin: Springer. MR2071631

Bernoulli **25**(2), 2019, 1412–1450
https://doi.org/10.3150/18-BEJ1025

From random partitions to fractional Brownian sheets

OLIVIER DURIEU[1] and YIZAO WANG[2]

[1] *Laboratoire de Mathématiques et Physique Théorique, UMR-CNRS 7350, Fédération Denis Poisson, FR-CNRS 2964, Université de Tours, Parc de Grandmont, 37200 Tours, France.*
E-mail: olivier.durieu@lmpt.univ-tours.fr
[2] *Department of Mathematical Sciences, University of Cincinnati, 2815 Commons Way, Cincinnati, OH 45221-0025, USA. E-mail: yizao.wang@uc.edu*

We propose discrete random-field models that are based on random partitions of $\mathbb{N}^2$. The covariance structure of each random field is determined by the underlying random partition. Functional central limit theorems are established for the proposed models, and fractional Brownian sheets, with full range of Hurst indices, arise in the limit. Our models could be viewed as discrete analogues of fractional Brownian sheets, in the same spirit that the simple random walk is the discrete analogue of the Brownian motion.

Keywords: fractional Brownian motion; fractional Brownian sheet; invariance principle; long-range dependence; random field; random partition; regular variation

References

[1] Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Stat.* **42** 1656–1670. MR0383482
[2] Biermé, H. and Durieu, O. (2014). Invariance principles for self-similar set-indexed random fields. *Trans. Amer. Math. Soc.* **366** 5963–5989. MR3256190
[3] Biermé, H., Durieu, O. and Wang, Y. (2017). Invariance principles for operator-scaling Gaussian random fields. *Ann. Appl. Probab.* **27** 1190–1234. MR3655864
[4] Biermé, H., Meerschaert, M.M. and Scheffler, H.-P. (2007). Operator scaling stable random fields. *Stochastic Process. Appl.* **117** 312–332. MR2290879
[5] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. New York: Wiley. MR1700749
[6] 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
[7] Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields. *Ann. Probab.* **10** 1047–1050. MR0672305
[8] Dedecker, J. (2001). Exponential inequalities and functional central limit theorems for a random fields. *ESAIM Probab. Stat.* **5** 77–104. MR1875665
[9] Durieu, O. and Wang, Y. (2016). From infinite urn schemes to decompositions of self-similar Gaussian processes. *Electron. J. Probab.* **21** Paper No. 43, 23. MR3530320
[10] Enriquez, N. (2004). A simple construction of the fractional Brownian motion. *Stochastic Process. Appl.* **109** 203–223. MR2031768
[11] Gnedin, A., Hansen, B. and Pitman, J. (2007). Notes on the occupancy problem with infinitely many boxes: General asymptotics and power laws. *Probab. Surv.* **4** 146–171. MR2318403

1350-7265

[12] Hammond, A. and Sheffield, S. (2013). Power law Pólya's urn and fractional Brownian motion. *Probab. Theory Related Fields* **157** 691–719. MR3129801

[13] Hu, Y., Øksendal, B. and Zhang, T. (2000). Stochastic partial differential equations driven by multiparameter fractional white noise. In *Stochastic Processes*, *Physics and Geometry*: *New Interplays*, *II* (*Leipzig*, 1999). *CMS Conf. Proc*. **29** 327–337. Providence, RI: Amer. Math. Soc. MR1803426

[14] Kallenberg, O. (1997). *Foundations of Modern Probability*. New York: Springer. MR1464694

[15] Kamont, A. (1996). On the fractional anisotropic Wiener field. *Probab. Math. Statist*. **16** 85–98. MR1407935

[16] Karlin, S. (1967). Central limit theorems for certain infinite urn schemes. *J. Math. Mech*. **17** 373–401. MR0216548

[17] Klüppelberg, C. and Kühn, C. (2004). Fractional Brownian motion as a weak limit of Poisson shot noise processes – With applications to finance. *Stochastic Process. Appl*. **113** 333–351. MR2087964

[18] Kolmogoroff, A.N. (1940). Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *Dokl. Akad. Nauk SSSR* **26** 115–118. MR0003441

[19] Lavancier, F. (2007). Invariance principles for non-isotropic long memory random fields. *Stat. Inference Stoch. Process*. **10** 255–282. MR2321311

[20] Lei, P. and Nualart, D. (2009). A decomposition of the bifractional Brownian motion and some applications. *Statist. Probab. Lett*. **79** 619–624. MR2499385

[21] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR0242239

[22] McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab*. **2** 620–628. MR0358933

[23] Mikosch, T. and Samorodnitsky, G. (2007). Scaling limits for cumulative input processes. *Math. Oper. Res*. **32** 890–918. MR2363203

[24] Øksendal, B. and Zhang, T. (2001). Multiparameter fractional Brownian motion and quasi-linear stochastic partial differential equations. *Stoch. Stoch. Rep*. **71** 141–163. MR1922562

[25] Peligrad, M. and Sethuraman, S. (2008). On fractional Brownian motion limits in one dimensional nearest-neighbor symmetric simple exclusion. *ALEA Lat. Am. J. Probab. Math. Stat*. **4** 245–255. MR2448774

[26] Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics* **45**. Cambridge: Cambridge Univ. Press. MR3729426

[27] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math*. **1875**. Berlin: Springer. MR2245368

[28] Puplinskaitė, D. and Surgailis, D. (2015). Scaling transition for long-range dependent Gaussian random fields. *Stochastic Process. Appl*. **125** 2256–2271. MR3322863

[29] Puplinskaitė, D. and Surgailis, D. (2016). Aggregation of autoregressive random fields and anisotropic long-range dependence. *Bernoulli* **22** 2401–2441. MR3498033

[30] Samorodnitsky, G. (2016). *Stochastic Processes and Long Range Dependence*. Cham: Springer. MR3561100

[31] Shen, Y. and Wang, Y. (2017). Operator-scaling Gaussian random fields via aggregation. Preprint. Available at https://arxiv.org/abs/1712.07082.

[32] Wang, Y. (2014). An invariance principle for fractional Brownian sheets. *J. Theoret. Probab*. **27** 1124–1139. MR3278934

[33] Xiao, Y. (2009). Sample path properties of anisotropic Gaussian random fields. In *A Minicourse on Stochastic Partial Differential Equations*. *Lecture Notes in Math*. **1962** 145–212. Berlin: Springer. MR2508776

Bernoulli **25**(2), 2019, 1451–1471
https://doi.org/10.3150/18-BEJ1026

A Bernstein-type inequality for functions of bounded interaction

ANDREAS MAURER

Adalbertstrasse 55, D 80799 München, Germany. E-mail: am@andreas-maurer.eu

We give a distribution-dependent concentration inequality for functions of independent variables. The result extends Bernstein's inequality from sums to more general functions, whose variation in any argument does not depend too much on the other arguments. Applications sharpen existing bounds for U-statistics and the generalization error of regularized least squares.

Keywords: Bernstein inequality; concentration; u-statistics

References

[1] Adamczak, R. (2006). Moment inequalities for U-statistics. *Ann. Probab.* **34** 2288–2314. MR2294982
[2] Arcones, M.A. (1995). A Bernstein-type inequality for U-statistics and U-processes. *Statist. Probab. Lett.* **22** 239–247. MR1323145
[3] Bernstein, S. (1924). On a modification of Chebyshev's inequality and of the error formula of Laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math* **1** 38-49.
[4] Boucheron, S., Lugosi, G. and Massart, P. (2003). Concentration inequalities using the entropy method. *Ann. Probab.* **31** 1583–1614. MR1989444
[5] 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
[6] Bousquet, O. and Elisseeff, A. (2002). Stability and generalization. *J. Mach. Learn. Res.* **2** 499–526. MR1929416
[7] de la Peña, V.H. (1992). Decoupling and Khintchine's inequalities for U-statistics. *Ann. Probab.* **20** 1877–1892. MR1188046
[8] Efron, B. and Stein, C. (1981). The jackknife estimate of variance. *Ann. Statist.* **9** 586–596. MR0615434
[9] Giné, E., Latała, R. and Zinn, J. (2000). Exponential and moment inequalities for U-statistics. In *High Dimensional Probability, II (Seattle, WA, 1999). Progress in Probability* **47** 13–38. Boston, MA: Birkhäuser. MR1857312
[10] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325. MR0026294
[11] Houdré, C. (1997). The iterated jackknife estimate of variance. *Statist. Probab. Lett.* **35** 197–201. MR1483274
[12] Houdré, C. and Reynaud-Bouret, P. (2003). Exponential inequalities, with constants, for U-statistics of order two. In *Stochastic Inequalities and Applications. Progress in Probability* **56** 55–69. Basel: Birkhäuser. MR2073426
[13] Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347

[14] Maurer, A. (2006). Concentration inequalities for functions of independent variables. *Random Structures Algorithms* **29** 121–138. MR2245497
[15] Maurer, A. (2012). Thermodynamics and concentration. *Bernoulli* **18** 434–454. MR2922456
[16] Maurer, A. (2017). A second-order look at stability and generalization. In *Conference on Learning Theory* 1461–1475.
[17] McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*. *Algorithms Combin*. **16** 195–248. Berlin: Springer. MR1678578
[18] Steele, J.M. (1986). An Efron–Stein inequality for nonsymmetric statistics. *Ann*. *Statist*. **14** 753–758. MR0840528
[19] v. Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann*. *Math*. *Stat*. **18** 309–348. MR0022330

Bernoulli **25**(2), 2019, 1472–1503
https://doi.org/10.3150/18-BEJ1027

An extreme-value approach for testing the equality of large U-statistic based correlation matrices

CHENG ZHOU[1,*], FANG HAN[2,**], XIN-SHENG ZHANG[1,†] and HAN LIU[3,‡]

[1]*Department of Statistics, Management School, Fudan University, Shanghai, China.*
E-mail: [*]chengzhmike@gmail.com; [†]xszhang@fudan.edu.cn
[2]*Department of Statistics, University of Washington, Seattle, WA 98195, USA. E-mail: [**]fanghan@uw.edu*
[3]*Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA. E-mail: [‡]hanliu@northwestern.edu*

There has been an increasing interest in testing the equality of large Pearson's correlation matrices. However, in many applications it is more important to test the equality of large rank-based correlation matrices since they are more robust to outliers and nonlinearity. Unlike the Pearson's case, testing the equality of large rank-based statistics has not been well explored and requires us to develop new methods and theory. In this paper, we provide a framework for testing the equality of two large U-statistic based correlation matrices, which include the rank-based correlation matrices as special cases. Our approach exploits extreme value statistics and the Jackknife estimator for uncertainty assessment and is valid under a fully nonparametric model. Theoretically, we develop a theory for testing the equality of U-statistic based correlation matrices. We then apply this theory to study the problem of testing large Kendall's tau correlation matrices and demonstrate its optimality. For proving this optimality, a novel construction of least favorable distributions is developed for the correlation matrix comparison.

Keywords: extreme value type I distribution; hypothesis testing; Jackknife variance estimator; Kendall's tau; U-statistics

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley-Interscience [John Wiley & Sons]. MR1990662
[2] Aslam, S. and Rocke, D.M. (2005). A robust testing procedure for the equality of covariance matrices. *Comput*. *Statist*. *Data Anal*. **49** 863–874. MR2141423
[3] Bai, Z., Jiang, D., Yao, J.-F. and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann*. *Statist*. **37** 3822–3840. MR2572444
[4] Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statist*. *Sinica* **18** 425–442. MR2411613
[5] Bai, Z.D. and Yin, Y.Q. (1993). Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann*. *Probab*. **21** 1275–1294. MR1235416
[6] Bao, Z., Lin, L.-C., Pan, G. and Zhou, W. (2015). Spectral statistics of large dimensional Spearman's rank correlation matrix and its application. *Ann*. *Statist*. **43** 2588–2623. MR3405605
[7] Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606. MR1935648

1350-7265

[8] Bickel, P.J. and Levina, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227. MR2387969
[9] Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **106** 672–684. MR2847949
[10] Cai, T., Liu, W. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.* **108** 265–277. MR3174618
[11] Cai, T., Liu, W. and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 349–372. MR3164870
[12] El Maache, H. and Lepage, Y. (2003). Spearman's rho and Kendall's tau for multivariate data sets. In *Mathematical Statistics and Applications: Festschrift for Constance van Eeden. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **42** 113–130. IMS, Beachwood, OH. MR2138289
[13] Embrechts, P., Lindskog, F. and McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. *Handbook of Heavy Tailed Distributions in Finance* **8** 329–384.
[14] Fang, H.-B., Fang, K.-T. and Kotz, S. (2002). The meta-elliptical distributions with given marginals. *J. Multivariate Anal.* **82** 1–16. MR1918612
[15] Giedd, J.N., Blumenthal, J., Molloy, E. and Castellanos, F.X. (2001). Brain imaging of attention deficit/hyperactivity disorder. *Ann. N.Y. Acad. Sci.* **931** 33–49.
[16] Han, F., Chen, S. and Liu, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306
[17] Han, F., Xu, S. and Zhou, W.-X. (2018). On Gaussian comparison inequality and its application to spectral analysis of large random matrices. *Bernoulli* **24** 1787–1833. MR3757515
[18] Han, F., Zhao, T. and Liu, H. (2013). CODA: High dimensional copula discriminant analysis. *J. Mach. Learn. Res.* **14** 629–671. MR3033343
[19] Ho, J.W.K., Stefani, M., dos Remedios, C.G. and Charleston, M.A. (2008). Differential variability analysis of gene expression and its application to human diseases. *Bioinformatics* **24** i390–i398.
[20] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325. MR0026294
[21] Hu, R., Qiu, X. and Glazko, G. (2010). A new gene selection procedure based on the covariance distance. *Bioinformatics* **26** 348–354.
[22] Hu, R., Qiu, X., Glazko, G., Klebanov, L. and Yakovlev, A. (2009). Detecting intergene correlation changes in microarray analysis: A new approach to gene selection. *BMC Bioinform.* **10** 20.
[23] Jiang, D., Jiang, T. and Yang, F. (2012). Likelihood ratio tests for covariance matrices of high-dimensional normal distributions. *J. Statist. Plann. Inference* **142** 2241–2256. MR2911842
[24] Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93.
[25] Klüppelberg, C. and Kuhn, G. (2009). Copula structure analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 737–753. MR2749917
[26] Kruskal, W.H. (1958). Ordinal measures of association. *J. Amer. Statist. Assoc.* **53** 814–861. MR0100941
[27] Li, J. and Chen, S.X. (2012). Two sample tests for high-dimensional covariance matrices. *Ann. Statist.* **40** 908–940. MR2985938
[28] Liu, H., Han, F., Yuan, M., Lafferty, J. and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. *Ann. Statist.* **40** 2293–2326. MR3059084
[29] Lopez-Paz, D., Hennig, P. and Schölkopf, B. (2013). The randomized dependence coefficient. In *Advances in Neural Information Processing Systems* 1–9.
[30] Lou, H., Henriksen, L. and Bruhn, P. (1990). Focal cerebral dysfunction in developmental learning disabilities. *Lancet* **335** 8–11.
[31] Mai, Q. and Zou, H. (2015). Sparse semiparametric discriminant analysis. *J. Multivariate Anal.* **135** 175–188. MR3306434

[32] Markowitz, H.M. (1991). Foundations of portfolio theory. *J. Finance* **46** 469–477.
[33] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. New York: John Wiley & Sons, Inc. Wiley Series in Probability and Mathematical Statistics. MR0652932
[34] Nagao, H. (1973). On some test criteria for covariance matrix. *Ann. Statist.* **1** 700–709. MR0339405
[35] O'Brien, P.C. (1992). Robust procedures for testing equality of covariance matrices. *Biometrics* **48** 819–827.
[36] Ravikumar, P., Wainwright, M.J., Raskutti, G. and Yu, B. (2011). High-dimensional covariance estimation by minimizing ℓ_1-penalized log-determinant divergence. *Electron. J. Stat.* **5** 935–980. MR2836766
[37] Roy, S.N. (1957). *Some Aspects of Multivariate Analysis*. Calcutta: Wiley, New York; Indian Statistical Institute. MR0092296
[38] Schott, J.R. (2007). A test for the equality of covariance matrices when the dimension is large relative to the sample sizes. *Comput. Statist. Data Anal.* **51** 6535–6542. MR2408613
[39] Shafritz, K.M., Marchione, K.E., Gore, J.C., Shaywitz, S.E. and Shaywitz, B.A. (2004). The effects of methylphenidate on neural systems of attention in attention deficit hyperactivity disorder. *Amer. J. Psychiatry* **161** 1990–1997.
[40] Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychology* **15** 72–101.
[41] Srivastava, M.S. and Du, M. (2008). A test for the mean vector with fewer observations than the dimension. *J. Multivariate Anal.* **99** 386–402. MR2396970
[42] Srivastava, M.S. and Yanagihara, H. (2010). Testing the equality of several covariance matrices with fewer observations than the dimension. *J. Multivariate Anal.* **101** 1319–1329. MR2609494
[43] Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. MR2382665
[44] Yufeng, Z., Yong, H., Chaozhe, Z., Qingjiu, C., Manqiu, S., Meng, L., Lixia, T., Tianzi, J. and Yufeng, W. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain and Development* **29** 83–91.
[45] Zhao, T., Roeder, K. and Liu, H. (2014). Positive semidefinite rank-based correlation matrix estimation with application to semiparametric graph estimation. *J. Comput. Graph. Statist.* **23** 895–922. MR3270703
[46] Zhou, C., Han, F., Zhang, X.-S. and Liu, H. (2019). Supplement to "An extreme-value approach for testing the equality of large U-statistic based correlation matrices." DOI:10.3150/18-BEJ1027SUPP.
[47] Zhou, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363. MR2327033
[48] Zou, Q., Zhu, C., Yang, Y., Zuo, X., Long, X., Cao, Q., Wang, Y. and Zang, Y. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J. Neurosci. Methods* **172** 137–141.

Bernoulli **25**(2), 2019, 1504–1535
https://doi.org/10.3150/18-BEJ1028

Numerically stable online estimation of variance in particle filters

JIMMY OLSSON[1] and RANDAL DOUC[2]

[1]*Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden.*
E-mail: jimmyol@kth.se
[2]*Département CITI, TELECOM SudParis, 9 rue Charles Fourier, 91000 EVRY, France.*
E-mail: randal.douc@it-sudparis.eu

This paper discusses variance estimation in sequential Monte Carlo methods, alternatively termed particle filters. The variance estimator that we propose is a natural modification of that suggested by H.P. Chan and T.L. Lai [*Ann. Statist.* **41** (2013) 2877–2904], which allows the variance to be estimated in a single run of the particle filter by tracing the genealogical history of the particles. However, due particle lineage degeneracy, the estimator of the mentioned work becomes numerically unstable as the number of sequential particle updates increases. Thus, by tracing only a part of the particles' genealogy rather than the full one, our estimator gains long-term numerical stability at the cost of a bias. The scope of the genealogical tracing is regulated by a lag, and under mild, easily checked model assumptions, we prove that the bias tends to zero geometrically fast as the lag increases. As confirmed by our numerical results, this allows the bias to be tightly controlled also for moderate particle sample sizes.

Keywords: asymptotic variance; Feynman–Kac models; hidden Markov models; particle filters; sequential Monte Carlo methods; state-space models; variance estimation

References

[1] Cappé, O., Moulines, E. and Rydén, T. (2005). *Inference in Hidden Markov Models*. *Springer Series in Statistics*. New York: Springer. MR2159833
[2] Chan, H.P. and Lai, T.L. (2013). A general theory of particle filters in hidden Markov models and some applications. *Ann. Statist.* **41** 2877–2904. MR3161451
[3] Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.* **32** 2385–2411. MR2153989
[4] Crisan, D. and Heine, K. (2008). Stability of the discrete time filter in terms of the tails of noise distributions. *J. Lond. Math. Soc.* (2) **78** 441–458. MR2439634
[5] Del Moral, P. (2004). *Feynman–Kac Formulae*: *Genealogical and Interacting Particle Systems with Applications*. *Probability and Its Applications* (*New York*). New York: Springer. MR2044973
[6] Del Moral, P. (2013). *Mean Field Simulation for Monte Carlo Integration*. *Monographs on Statistics and Applied Probability* **126**. Boca Raton, FL: CRC Press. MR3060209
[7] Del Moral, P. and Guionnet, A. (1999). Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9** 275–297. MR1687359
[8] Del Moral, P. and Guionnet, A. (2001). On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. Henri Poincaré Probab. Stat.* **37** 155–194. MR1819122
[9] Douc, R., Fort, G., Moulines, E. and Priouret, P. (2009). Forgetting the initial distribution for hidden Markov models. *Stochastic Process. Appl.* **119** 1235–1256. MR2508572

[10] Douc, R. and Moulines, E. (2008). Limit theorems for weighted samples with applications to sequential Monte Carlo methods. *Ann. Statist.* **36** 2344–2376. MR2458190

[11] Douc, R. and Moulines, E. (2012). Asymptotic properties of the maximum likelihood estimation in misspecified hidden Markov models. *Ann. Statist.* **40** 2697–2732. MR3097617

[12] Douc, R., Moulines, E. and Olsson, J. (2014). Long-term stability of sequential Monte Carlo methods under verifiable conditions. *Ann. Appl. Probab.* **24** 1767–1802. MR3226163

[13] Doucet, A., De Freitas, N. and Gordon, N., eds. (2001). *Sequential Monte Carlo Methods in Practice*. New York: Springer.

[14] Gordon, N., Salmond, D. and Smith, A.F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc., F, Radar Signal Process.* **140** 107–113.

[15] Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance* **42** 281–300.

[16] Jacob, P.E., Murray, L.M. and Rubenthaler, S. (2015). Path storage in the particle filter. *Stat. Comput.* **25** 487–496. MR3306720

[17] Kitagawa, G. and Sato, S. (2001). Monte Carlo smoothing and self-organising state-space model. In *Sequential Monte Carlo Methods in Practice* (A. Doucet, N. De Freitas and N. Gordon, eds.). *Stat. Eng. Inf. Sci.* 177–195. New York: Springer. MR1847792

[18] Künsch, H.R. (2005). Recursive Monte Carlo filters: Algorithms and theoretical analysis. *Ann. Statist.* **33** 1983–2021. MR2211077

[19] Lee, A. and Whiteley, N. (2016). Variance estimation in particle filters. Preprint. Available at arXiv:1509.00394.

[20] Lindsten, F., Schön, T.B. and Olsson, J. (2011). An explicit variance reduction expression for the Rao–Blackwellised particle filter. In *Proceedings of the* 18*th IFAC World Congress* 11979–11984.

[21] Olsson, J., Cappé, O., Douc, R. and Moulines, E. (2008). Sequential Monte Carlo smoothing with application to parameter estimation in nonlinear state space models. *Bernoulli* **14** 155–179. MR2401658

[22] Olsson, J. and Ströjby, J. (2011). Particle-based likelihood inference in partially observed diffusion processes using generalised Poisson estimators. *Electron. J. Stat.* **5** 1090–1122. MR2836770

[23] Ristic, B., Arulampalam, M. and Gordon, A. (2004). *Beyond Kalman Filters: Particle Filters for Target Tracking*. Norwood: Artech House.

Bernoulli **25**(2), 2019, 1536–1567
https://doi.org/10.3150/18-BEJ1029

New tests of uniformity on the compact classical groups as diagnostics for weak-* mixing of Markov chains

AMIR SEPEHRI

Department of Statistics, Sequoia Hall, Stanford, CA 94305, USA. E-mail: asepehri@stanford.edu

This paper introduces two new families of non-parametric tests of goodness-of-fit on the compact classical groups. One of them is a family of tests for the eigenvalue distribution induced by the uniform distribution, which is consistent against all fixed alternatives. The other is a family of tests for the uniform distribution on the entire group, which is again consistent against all fixed alternatives. The construction of these tests heavily employs facts and techniques from the representation theory of compact groups. In particular, new Cauchy identities are derived and proved for the characters of compact classical groups, in order to accommodate the computation of the test statistic. We find the asymptotic distribution under the null and general alternatives. The tests are proved to be asymptotically admissible. Local power is derived and the global properties of the power function against local alternatives are explored.

The new tests are validated on two random walks for which the mixing-time is studied in the literature. The new tests, and several others, are applied to the Markov chain sampler proposed by Jones, Osipov and Rokhlin [*Proc. Natl. Acad. Sci.* **108** (2011) 15679–15686], providing strong evidence supporting the claim that the sampler mixes quickly.

Keywords: Cauchy identity; goodness-of-fit; mixing-diagnostics for Markov Chains; non-parametric hypothesis testing; random rotation generators; representation theory of compact groups; spectral analysis

References

[1] Abramowitz, M. and Stegun, I.A. (1966). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. **55**. New York: Dover Publications. MR0208797
[2] Ajne, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika* **55** 343–354. MR0235662
[3] Andrews, G.E. (1998). *The Theory of Partitions*. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. MR1634067
[4] Arias-Castro, E., Pelletier, B. and Saligrama, V. (2018). Remember the curse of dimensionality: The case of goodness-of-fit testing in arbitrary dimension. *J. Nonparametr. Stat.* **30** 448–471. MR3794401
[5] Baringhaus, L. (1991). Testing for spherical symmetry of a multivariate distribution. *Ann. Statist.* **19** 899–917. MR1105851
[6] Beran, R. (1975). Tail probabilities of noncentral quadratic forms. *Ann. Statist.* **3** 969–974. MR0381122
[7] Beran, R.J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Probab.* **5** 177–195. MR0228098
[8] 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

1350-7265

[9] Bump, D. (2004). *Lie Groups*. *Graduate Texts in Mathematics* **225**. New York: Springer. MR2062813
[10] Coram, M. and Diaconis, P. (2003). New tests of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function. *J. Phys. A* **36** 2883–2906. MR1986397
[11] Diaconis, P. (2003). Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture. *Bull. Amer. Math. Soc.* (*N.S.*) **40** 155–178. MR1962294
[12] Diaconis, P. and Mallows, C. (1986). On the trace of random orthogonal matrices. Unpublished Manuscript. Results Summarized in Diaconis (1990).
[13] Diaconis, P. and Shahshahani, M. (1986). Products of random matrices as they arise in the study of random walks on groups. In *Random Matrices and Their Applications* (*Brunswick*, *Maine*, 1984). *Contemp. Math.* **50** 183–195. Providence, RI: Amer. Math. Soc. MR0841092
[14] Diaconis, P. and Shahshahani, M. (1987). The subgroup algorithm for generating uniform random variables. *Probab. Engrg. Inform. Sci.* **1** 15–32.
[15] Downs, T.D. (1972). Orientation statistics. *Biometrika* **59** 665–676. MR0345334
[16] Giné, E.M. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. *Ann. Statist.* **3** 1243–1266. MR0388663
[17] Giné, E.M. (1975). The addition formula for the eigenfunctions of the Laplacian. *Adv. Math.* **18** 102–107. MR0377997
[18] Goodman, R. and Wallach, N.R. (2009). *Symmetry*, *Representations*, *and Invariants*. *Graduate Texts in Mathematics* **255**. Dordrecht: Springer. MR2522486
[19] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109. MR3363437
[20] Hermans, M. and Rasson, J.-P. (1985). A new Sobolev test for uniformity on the circle. *Biometrika* **72** 698–702. MR0817587
[21] Janssen, A. (1995). Principal component decomposition of non-parametric tests. *Probab. Theory Related Fields* **101** 193–209. MR1318192
[22] Jones, P.W., Osipov, A. and Rokhlin, V. (2011). Randomized approximate nearest neighbors algorithm. *Proc. Natl. Acad. Sci.* **108** 15679–15686.
[23] Jupp, P.E. and Spurr, B.D. (1983). Sobolev tests for symmetry of directional data. *Ann. Statist.* **11** 1225–1231. MR0720267
[24] Jupp, P.E. and Spurr, B.D. (1985). Sobolev tests for independence of directions. *Ann. Statist.* **13** 1140–1155. MR0803763
[25] Kac, M. (1959). *Probability and Related Topics in Physical Sciences* **1**. American Mathematical Soc.
[26] Kerkyacharian, G., Nickl, R. and Picard, D. (2012). Concentration inequalities and confidence bands for needlet density estimators on compact homogeneous manifolds. *Probab. Theory Related Fields* **153** 363–404. MR2925578
[27] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR2135927
[28] Liberty, E., Woolfe, F., Martinsson, P.-G., Rokhlin, V. and Tygert, M. (2007). Randomized algorithms for the low-rank approximation of matrices. *Proc. Natl. Acad. Sci. USA* **104** 20167–20172. MR2366406
[29] Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR1828667
[30] Mehta, M.L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics* (*Amsterdam*) **142**. Amsterdam: Elsevier/Academic Press. MR2129906
[31] Oliveira, R.I. (2009). On the convergence to equilibrium of Kac's random walk on matrices. *Ann. Appl. Probab.* **19** 1200–1231. MR2537204
[32] Pak, I. and Sidenko, S. (2007). Convergence of Kac's random walk. Preprint. Available at http://www-math.mit.edu/~pak/research.html.

[33] Pillai, N.S. and Smith, A. (2016). On the mixing time of Kac's walk and other high-dimensional Gibbs samplers with constraints. Preprint. Available at ArXiv:1605.08122.
[34] Porod, U. (1996). The cut-off phenomenon for random reflections. *Ann. Probab.* **24** 74–96. MR1387627
[35] Prentice, M.J. (1978). On invariant tests of uniformity for directions and orientations. *Ann. Statist.* **6** 169–176. MR0458721
[36] Rayleigh, L. (1880). XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Philos. Mag.* **10** 73–78.
[37] Rokhlin, V. and Tygert, M. (2008). A fast randomized algorithm for overdetermined linear least-squares regression. *Proc. Natl. Acad. Sci. USA* **105** 13212–13217. MR2443725
[38] Römisch, W. (2005). Delta method, infinite dimensional. *Encyclopedia of Statistical Sciences*.
[39] Rosenthal, J.S. (1994). Random rotations: Characters and random walks on $\mathrm{SO}(N)$. *Ann. Probab.* **22** 398–423. MR1258882
[40] Sengupta, A. and Pal, C. (2001). On optimal tests for isotropy against the symmetric wrapped stable-circular uniform mixture family. *J. Appl. Stat.* **28** 129–143. MR1819907
[41] Sepehri, A. (2019). Supplement to "New tests of uniformity on the compact classical groups as diagnostics for weak-* mixing of Markov chains." DOI:10.3150/18-BEJ1029SUPP.
[42] Strasser, H. (1985). *Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory. De Gruyter Studies in Mathematics* **7**. Berlin: de Gruyter. MR0812467
[43] Thomson, B.S., Bruckner, J.B. and Bruckner, A.M. (2008). *Elementary Real Analysis*. Available at ClassicalRealAnalysis.com.
[44] van der Vaart, A.W. and Wellner, J.A. (1996). Weak convergence. In *Weak Convergence and Empirical Processes* 16–28. New York: Springer. MR1385671
[45] Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48** 109–114. MR0131930
[46] Watson, G.S. (1962). Goodness-of-fit tests on a circle. II. *Biometrika* **49** 57–63. MR0138179
[47] Watson, G.S. (1967). Another test for the uniformity of a circular distribution. *Biometrika* **54** 675–677. MR0221649
[48] Wellner, J.A. (1979). Permutation tests for directional data. *Ann. Statist.* **7** 929–943. MR0536498
[49] Weyl, H. (1939). *The Classical Groups. Their Invariants and Representations*. Princeton, NJ: Princeton Univ. Press. MR0000255

Bernoulli **25**(2), 2019, 1568–1601
https://doi.org/10.3150/18-BEJ1030

Macroscopic analysis of determinantal random balls

JEAN-CHRISTOPHE BRETON*, ADRIEN CLARENNE** and RENAN GOBARD†

Univ Rennes, CNRS, IRMAR – UMR 6625, F-35000 Rennes, France.
*E-mail: *jean-christophe.breton@univ-rennes1.fr; **adrien.clarenne@univ-rennes1.fr; †renan.gobard@univ-rennes1.fr*

We consider a collection of Euclidean random balls in $\mathbb{R}^d$ generated by a determinantal point process inducing inhibitory interaction into the balls. We study this model at a macroscopic level obtained by a zooming-out and three different regimes – Gaussian, Poissonian and stable – are exhibited as in the Poissonian model without interaction. This shows that the macroscopic behaviour erases the interactions induced by the determinantal point process.

Keywords: determinantal point processes; generalized random fields; limit theorem; point processes; stable fields

References

[1] Arfken, G.B. and Weber, H.J. (2001). *Mathematical Methods for Physicists*, 5th ed. Burlington, MA: Harcourt/Academic Press. MR1810939
[2] Biermé, H. and Estrade, A. (2006). Poisson random balls: Self-similarity and x-ray images. *Adv. in Appl. Probab.* **38** 853–872. MR2285684
[3] Biermé, H., Estrade, A. and Kaj, I. (2010). Self-similar random fields and rescaled random balls models. *J. Theoret. Probab.* **23** 1110–1141. MR2735739
[4] Breton, J.-C. and Dombry, C. (2009). Rescaled weighted random ball models and stable self-similar random fields. *Stochastic Process. Appl.* **119** 3633–3652. MR2568289
[5] Chiu, S.N., Stoyan, D., Kendall, W.S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*, 3rd ed. Chichester: Wiley. MR3236788
[6] Daley, D.J. and Vere-Jones, D. (2002). *Introduction to Point Processes. Volumes* 1 *and* 2, 2nd ed.
[7] Deng, N., Zhou, W. and Haenggi, M. (2014). The Ginibre point process as a model for wireless networks with repulsion. Available at arXiv:1401.3677.
[8] Dragomir, S.S. (2017). Some trace inequalities for operators in Hilbert spaces. *Kragujevac J. Math.* **41** 33–55. MR3668251
[9] Gobard, R. (2015). Random balls model with dependence. *J. Math. Anal. Appl.* **423** 1284–1310. MR3278199
[10] Gobard, R. (2015). Fluctuations dans les modèles de boules aléatoires. Ph.D., Université de Rennes 1. Available at https://hal.inria.fr/IRMAR/tel-01167520v1.
[11] Heinrich, L. and Schmidt, V. (1985). Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. in Appl. Probab.* **17** 709–730. MR0809427
[12] Hough, J.B., Krishnapur, M., Peres, Y. and Virág, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes. University Lecture Series* **51**. Providence, RI: Amer. Math. Soc. MR2552864

[13] Kaj, I., Leskelä, L., Norros, I. and Schmidt, V. (2007). Scaling limits for random fields with long-range dependence. *Ann. Probab.* **35** 528–550. MR2308587

[14] Kaj, I. and Taqqu, M.S. (2008). Convergence to fractional Brownian motion and to the Telecom process: The integral representation approach. In *In and Out of Equilibrium. 2. Progress in Probability* **60** 383–427. Birkhäuser, Basel. MR2477392

[15] Klüppelberg, C. and Mikosch, T. (1995). Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* **1** 125–147. MR1354458

[16] Lane, J.A. (1984). The central limit theorem for the Poisson shot-noise process. *J. Appl. Probab.* **21** 287–301. MR0741131

[17] Li, Y., Baccelli, F., Dhillon, H.S. and Andrews, J.G. (2015). Statistical modeling and probabilistic analysis of cellular networks model with determinantal point processes. *IEE Transactions on Communications* **63** 3405–3422.

[18] Meester, R. and Roy, R. (1996). *Continuum Percolation. Cambridge Tracts in Mathematics* **119**. Cambridge: Cambridge Univ. Press. MR1409145

[19] Mikosch, T., Resnick, S., Rootzén, H. and Stegeman, A. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.* **12** 23–68. MR1890056

[20] Miyoshi, N. and Shirai, T. (2014). A cellular network model with Ginibre configured base stations. *Adv. in Appl. Probab.* **46** 832–845. MR3254344

[21] Miyoshi, N. and Shirai, T. (2017). Tail asymptotics of signal-to-interference ratio distribution in spatial cellular network models. *Probab. Math. Statist.* **37** 431–453. MR3745394

[22] Reed, M. and Simon, B. (1972). *Methods of Modern Mathematical Physics. I. Functional Analysis*. New York: Academic Press. MR0493419

[23] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman & Hall. MR1280932

[24] Shirai, T. and Takahashi, Y. (2003). Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *J. Funct. Anal.* **205** 414–463. MR2018415

[25] Yang, X. and Petropulu, A.P. (2003). Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications. *IEEE Trans. Signal Process.* **51** 64–76. MR1956093

Bernoulli Forthcoming Papers

THAI, M.-N.
Birth and death process in mean field type interaction

ZHOU, S.
Sparse Hanson–Wright inequalities for subgaussian quadratic forms

BEIGLBÖCK, M., LIM,, T. and OBŁÓJ, J.
Dual attainment for the martingale transport problem

YAROSLAVTSEV, I.S.
Martingale decompositions and weak differential subordination in UMD Banach spaces

DOMBRY, C. and FERREIRA, A.
Maximum likelihood estimators based on the block maxima method

POINAS, A., DELYON, B. and LAVANCIER, F.
Mixing properties and central limit theorem for associated point processes

KÜHN, F.
Perpetual integrals via random time changes

BRUNEL, V.-E.
Uniform behaviors of random polytopes under the Hausdorff metric

SAUMARD, A. and WELLNER, J.A.
On the isoperimetric constant, covariance inequalities and L_p-Poincaré inequalities in dimension one

KELLNER, J. and CELISSE, A.
A one-sample test for normality with kernel methods

BAI, Z., LI, H. and PAN, G.
Central limit theorem for linear spectral statistics of large dimensional separable sample covariance matrices

FUKASAWA, M. and TAKABATAKE, T.
Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations

BELOMESTNY, D., TRABS, M. and TSYBAKOV, A.B.
Sparse covariance matrix estimation in high-dimensional deconvolution

CHAKRABORTY, A. and PANARETOS, V.M.
Hybrid regularisation and the (in)admissibility of ridge regression in infinite dimensional Hilbert spaces

MARIN, J.-M., PUDLO, P. and SEDKI, M.
Consistency of adaptive importance sampling and recycling schemes

LI, C., LIN, L. and DUNSON, D.B.
On posterior consistency of tail index for Bayesian kernel mixture models

Continues

Bernoulli Forthcoming Papers—*Continued*

GRAHOVAC, D., LEONENKO, N.N., SIKORSKII, A. and TAQQU, M.S.
The unusual properties of aggregated superpositions of Ornstein–Uhlenbeck type processes

HENNING, F., KRAAIJ, R.C. and KÜLSKE, C.
Gibbs–non-Gibbs transitions in the fuzzy Potts model with a Kac-type interaction: Closing the Ising gap

LUGOSI, G. and MENDELSON, S.
Regularization, sparse recovery, and median-of-means tournaments

TRUQUET, L.
Root-n consistent estimation of the marginal density in semiparametric autoregressive time series models

DEYA, A. and SCHOTT, R.
Integration with respect to the non-commutative fractional Brownian motion

HE, Y., LIN, C.D. and SUN, F.
Construction of marginally coupled designs by subspace theory

KOSKELA, J., SPANÒ, D. and JENKINS, P.A.
Consistency of Bayesian nonparametric inference for discretely observed jump diffusions

FANG, B. and GUNTUBOYINA, A.
On the risk of convex-constrained least squares estimators under misspecification

GRANELLI, A. and VERAART, A.E.D.
A central limit theorem for the realised covariation of a bivariate Brownian semistationary process

WANG, L., XIANG, K. and ZOU, L.
The first order correction to harmonic measure for random walks of rotationally invariant step distribution

HE, H. and WINKEL, M.
Gromov–Hausdorff–Prokhorov convergence of vertex cut-trees of n-leaf Galton–Watson trees

YOO, W.W. and GHOSAL, S.
Bayesian mode and maximum estimation and accelerated rates of contraction

JENTSCH, C. and WEISS, C.H.
Bootstrapping INAR models

LEI, T.
Scaling limit of random forests with prescribed degree sequences

ICHIBA, T. and SARANTSEV, A.
Stationary distributions and convergence for Walsh diffusions

Continues

Bernoulli Forthcoming Papers—*Continued*

LIMIC, V.
The eternal multiplicative coalescent encoding via excursions of Lévy-type processes

BUHL, S., DAVIS, R.A., KLÜPPELBERG, C. and STEINKOHL, C.
Semiparametric estimation for isotropic max-stable space-time processes

GÖTZE, F., NAUMOV, A., SPOKOINY, V. and ULYANOV, V.
Large ball probabilities, Gaussian comparison and anti-concentration

FANG, X., PENG, S., SHAO, Q.-M. and SONG, Y.
Limit theorems with rate of convergence under sublinear expectations

REISS, M. and WAHL, M.
Functional estimation and hypothesis testing in nonparametric boundary models

WEED, J. and BACH, F.
Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance

MARGUET, A.
Uniform sampling in a structured branching population

ABRAHAM, K.
Nonparametric Bayesian posterior contraction rates for scalar diffusions with high-frequency data

HUESMANN, M. and TREVISAN, D.
A Benamou–Brenier formulation of martingale optimal transport

BHATTACHARJEE, C. and GOLDSTEIN, L.
Dickman approximation in simulation, summations and perpetuities

FAN, X., GRAMA, I., LIU, Q. and SHAO, Q.-M.
Self-normalized Cramér type moderate deviations for martingales

RAIČ, M.
A multivariate Berry–Esséen theorem with explicit constants

BECT, J., BACHOC, F. and GINSBOURGER, D.
A supermartingale approach to Gaussian process based sequential design of experiments

ANH, V., LEONENKO, N., OLENKO, A. and VASKOVYCH, V.
On rate of convergence in non-central limit theorems

LIU, Z., BLANCHET, J.H., DIEKER, A.B. and MIKOSCH, T.
On logarithmically optimal exact simulation of max-stable and related random fields on a compact set

DOLERA, E. and REGAZZINI, E.
Uniform rates of the Glivenko–Cantelli convergence and their use in approximating Bayesian inferences

Continues

Bernoulli Forthcoming Papers—*Continued*

BELLEC, P.C.
Localized Gaussian width of M-convex hulls with applications to Lasso and convex aggregation

DITZHAUS, M.
Signal detection via Phi-divergences for general mixtures

BALAN, R.M. and SONG, J.
Second order Lyapunov exponents for parabolic and hyperbolic Anderson models

NGUYEN, V.H.
Φ-entropy inequalities and asymmetric covariance estimates for convex measures

LIVINGSTONE, S., BETANCOURT, M., BYRNE, S. and GIROLAMI, M.
On the geometric ergodicity of Hamiltonian Monte Carlo

ALONSO-GUTIÉRREZ, D., PROCHNO, J. and THÄLE, C.
Gaussian fluctuations for high-dimensional random projections of ℓ_p^n-balls

CHZHEN, E., HEBIRI, M. and SALMON, J.
On Lasso refitting strategies

MERLEVÈDE, F., PELIGRAD, M. and UTEV, S.
Functional CLT for martingale-like nonstationary dependent structures

VARVENNE, M.
Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory

BALABDAOUI, F., DUROT, C. and JANKOWSKI, H.
Least squares estimation in the monotone single index model

HONDA, T., ING, C.-K. and WU, W.-Y.
Adaptively weighted group Lasso for semiparametric quantile regression models

ALETTI, G., CRIMALDI, I. and GHIGLIETTI, A.
Networks of reinforced stochastic processes: Asymptotics for the empirical means

BUTLER, R.W. and WOOD, A.T.A.
Limiting saddlepoint relative errors in large deviation regions under purely Tauberian conditions

JIANG, W. and ZHANG, C.-H.
Rate of divergence of the nonparametric likelihood ratio test for Gaussian mixtures

GREENSHTEIN, E., MANSURA, A. and RITOV, Y.
Nonparametric empirical Bayes improvement of shrinkage estimators with applications to time series

MALLEIN, B. and RAMASSAMY, S.
Two-sided infinite-bin models and analyticity for Barak–Erdős graphs

MATZINGER, H. and HAUSER, R.
Microscopic Path Structure of Optimally Aligned Random Sequences

Continues

Bernoulli Forthcoming Papers—*Continued*

CHENG, D., CAMMAROTA, V., FANTAYE, Y., MARRINUCCI, D. and SCHWARTZMAN, A.
Multiple testing of local maxima for detection of peaks on the (celestial) sphere

DURMUS, A. and MOULINES, E.
High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm

BERHARD, H. and DAS, B.
Heavy-tailed random walks, buffered queues and hidden large deviations

DAVIS, R.A., NIELSEN, M.S. and ROHDE, V.
Stochastic differential equations with a fractionally filtered delay: a semimartingale model for long-range dependent processes

HO, N., NGUYEN, X., RITOV, Y.
Robust estimation of mixing measures in finite mixture models

BLANCHARD, G. and ZADOROZHNYI, O.
Concentration of weakly dependent Banach-valued sums and applications to kernel learning methods

CÉNAC, P., DE LOYNES, B., OFFRET, Y. and ROUSSELLE, A.
Recurrence of Multidimensional Persistent Random Walks. Fourier and Series Criteria

PILAVAKAS, D., PAPARODITIS, E. and SAPATINAS, T.
Moving Block and Tapered Block Bootstrap for Functional Time Series With an Application to the-K-Sample Mean Problem

HAMZA, K.
Convergence of the age structure of general schemes of population processes

CLÉMENÇON, S. and BERTAIL, P.
Bernstein-type exponential inequalities in survey sampling: conditional Poisson sampling schemes

BARTHELME, S., AMBLAED, P.-O., and TREMBLAY, N.
Asymptotic Equivalence of Fixed-size and Varying-size Determinantal Point Processes

HEINY, J. and MIKOSCH, T.
The eigenstructure of the sample covariance matrices of high-dimensional stochastic volatility models with heavy tails

PITMAN, J. and YAKUBOVICH, Y.
Gaps and interleaving of point processes in sampling from a residual allocation model

NOURDIN, I., PECCATI, G. and SEURET, S.
Sojourn Time Dimensions of Fractional Brownian Motion

LIN, S.
Harmonic measure for biased random walk in a supercritical Galton–Watson tree

Continues

Bernoulli Forthcoming Papers—*Continued*

FRANCESCHI, S. and RASCHEL, K.
Explicit expression for the stationary distribution of reflected Brownian motion in a wedge

GOUÉRÉ, J.-B. and THÉRET, M.
Equivalence of some subcritical properties in continuum percolation

RAVNER, L., BOXMA, O. and MANDJES, M.
Estimating the input of a Lévy-driven queue by Poisson sampling of the workload process

COLLING, B. and VAN KEILEGOM, I.
Estimation of fully nonparametric transformation models